

WebSphere Application Server



Load Balancer - Guide d'administration

Version 6.1

WebSphere Application Server



Load Balancer - Guide d'administration

Version 6.1

Important

Avant d'utiliser le présent document et le produit associé, prenez connaissance des informations générales figurant à l'Annexe E, «Remarques», à la page 497.

Remarque

Certaines captures d'écran et certains graphiques de ce manuel ne sont pas disponibles en français à la date d'impression.

Première édition - mai 2006

Réf. US : GC31-6921-00

LE PRESENT DOCUMENT EST LIVRE EN L'ETAT SANS AUCUNE GARANTIE EXPLICITE OU IMPLICITE. IBM DECLINE NOTAMMENT TOUTE RESPONSABILITE RELATIVE A CES INFORMATIONS EN CAS DE CONTREFACON AINSI QU'EN CAS DE DEFAUT D'APTITUDE A L'EXECUTION D'UN TRAVAIL DONNE.

Ce document est mis à jour périodiquement. Chaque nouvelle édition inclut les mises à jour. Les informations qui y sont fournies sont susceptibles d'être modifiées avant que les produits décrits ne deviennent eux-mêmes disponibles. En outre, il peut contenir des informations ou des références concernant certains produits, logiciels ou services non annoncés dans ce pays. Cela ne signifie cependant pas qu'ils y seront annoncés.

Pour plus de détails, pour toute demande d'ordre technique, ou pour obtenir des exemplaires de documents IBM, référez-vous aux documents d'annonce disponibles dans votre pays, ou adressez-vous à votre partenaire commercial.

Vous pouvez également consulter les serveurs Internet suivants :

- <http://www.fr.ibm.com> (serveur IBM en France)
- <http://www.can.ibm.com> (serveur IBM au Canada)
- <http://www.ibm.com> (serveur IBM aux Etats-Unis)

*Compagnie IBM France
Direction Qualité
Tour Descartes
92066 Paris-La Défense Cedex 50*

© Copyright IBM France 2006. Tous droits réservés.

© Copyright International Business Machines Corporation 2006. All rights reserved.

Table des matières

Tableaux.	xiii
------------------	-------------

Figures	xv
----------------	-----------

A propos de ce guide.	xvii
------------------------------	-------------

A qui s'adresse ce guide	xvii
Informations connexes	xvii
Accessibilité.	xvii
Comment envoyer vos commentaires	xvii

Documents associés et sites Web.	xix
---	------------

Partie 1. Introduction à Load Balancer	1
---	----------

Chapitre 1. Présentation générale de Load Balancer	3
---	----------

Principe de Load Balancer	3
Composant(s) de Load Balancer utilisable(s)	3
Avantages de Load Balancer	4
Haute disponibilité fournie par Load Balancer	6
Dispatcher	6
CBR	6
Cisco CSS Controller ou Nortel Alteon Controller	6
Nouvelles fonctions	6

Chapitre 2. Présentation générale des composants de Load Balancer	9
--	----------

Composants de Load Balancer	9
Présentation générale du composant Dispatcher	9
Gestion de serveurs locaux avec Dispatcher	10
Gestion des serveurs à l'aide de Dispatcher et de Metric Server	11
Gestion de serveurs locaux et éloignés avec Dispatcher	12
Présentation générale du composant CBR (Content Based Routing)	12
Gestion des serveurs locaux avec CBR	13
Présentation générale du composant Site Selector	14
Gestion des serveurs locaux et éloignés avec Site Selector et Metric Server	15
Présentation générale du composant Cisco CSS Controller	15
Présentation générale du composant Nortel Alteon Controller	17

Chapitre 3. Gestion du réseau : Fonctions Load Balancer requises.	19
--	-----------

Gestionnaire, conseillers et système Metric Server (des composants Dispatcher, CBR et Site Selector)	19
Fonctions du composant Dispatcher	19
Administration à distance	19
Co-implantation	19

Haute disponibilité	20
Affinité client à serveur	20
Equilibrage de charge basé sur des règles	20
Routage par contenu à l'aide de la méthode d'acheminement CBR de Dispatcher	21
Equilibrage de charge d'un réseau étendu	22
Mappage de port	22
Configuration de Dispatcher sur un réseau privé	22
Cluster générique et port générique	22
Détection d'attaque de "refus de service"	22
Consignation binaire	23
Alertes	23

Fonctions du composant CBR (Content Based Routing)	23
--	----

Comparaison entre le composant CBR et la méthode d'acheminement CBR de Dispatcher	23
Administration à distance	24
Co-implantation	24
CBR et plusieurs instances de Caching Proxy	24
Routage par contenu pour les connexions SSL	24
Partitionnement du serveur	24
Equilibrage de charge basé sur des règles	24
Affinité client à serveur	25
Haute disponibilité à l'aide de Dispatcher et CBR	25
Consignation binaire	25
Alertes	25

Fonctions du composant Site Selector	25
Administration à distance	26
Co-implantation	26
Haute disponibilité	26
Affinité client à serveur	26
Equilibrage de charge basé sur des règles	26
Equilibrage de charge d'un réseau étendu	26
Alertes	27

Fonctions du composant Cisco CSS Controller	27
Administration à distance	27
Co-implantation	27
Haute disponibilité	27
Consignation binaire	28
Alertes	28

Fonctions du composant Nortel Alteon Controller	28
Administration à distance	28
Co-implantation	28
Haute disponibilité	28
Consignation binaire	28
Alertes	29

Chapitre 4. Installation de Load Balancer	31
--	-----------

Configuration requise et installation pour AIX	31
Configuration requise pour les systèmes AIX	31
Installation pour les systèmes AIX	32
Avant de commencer l'installation	32
Etapes de la procédure d'installation	33
Configuration requise et installation pour HP-UX	35

Configuration requise pour les systèmes HP-UX	35
Installation pour les systèmes HP-UX.	35
Avant de commencer l'installation.	35
Etapes de la procédure d'installation.	35
Configuration requise et installation pour Linux	36
Configuration requise pour les systèmes Linux	36
Installation pour les systèmes Linux	37
Avant de commencer l'installation.	37
Etapes de la procédure d'installation.	37
Configuration requise et installation pour Solaris	38
Configuration requise pour Solaris.	38
Installation pour Solaris	38
Avant de commencer l'installation.	39
Etapes de la procédure d'installation.	39
Configuration requise et installation pour Windows	40
Configuration requise pour les systèmes	
Windows	40
Installation pour les systèmes Windows	40
Avant de commencer l'installation.	40
Etapes de la procédure d'installation.	41

Partie 2. Composant Dispatcher . . . 43

Chapitre 5. Configuration de démarrage rapide . . . 45

Matériel requis	45
Préparation	46
Configuration du composant Dispatcher.	47
Configuration à partir de la ligne de commande	47
Test de vérification de la configuration	47
Configuration à l'aide de l'interface graphique	48
Assistant de configuration	48
Types de configurations de cluster, de port et de serveur	48

Chapitre 6. Planification de Dispatcher 51

Remarques relatives à la planification.	51
Méthodes d'acheminement	52
Réacheminement MAC de Dispatcher (méthode d'acheminement mac)	53
Réacheminement NAT/NAPT de Dispatcher (méthode d'acheminement nat)	53
Fonction CBR de Dispatcher (méthode d'acheminement cbr)	55
Etapes de configuration des méthodes d'acheminement nat ou cbr de Dispatcher	57
Partitionnement du serveur : serveurs logiques configurés pour un serveur physique (adresse IP)	58
Partitionnement de serveur à l'aide des conseillers HTTP ou HTTPS	58
Exemple de configuration d'un serveur physique en plusieurs serveurs logiques	59
Haute disponibilité	60
Haute disponibilité simple	60
Haute disponibilité réciproque	61

Chapitre 7. Configuration de Dispatcher 63

Présentation générale des tâches de configuration	63
Méthodes de configuration	63

Ligne de commande	63
Scripts	64
Interface graphique.	65
Configuration à l'aide de l'assistant de configuration	66
Configuration de la machine Dispatcher	66
Etape 1. Démarrage de la fonction serveur	68
Etape 2. Démarrage de la fonction exécuteur	68
Etape 3. Définition de l'adresse de non-réacheminement (si différente du nom d'hôte)	69
Etape 4. Définition et configuration des options du cluster	69
Etape 5. Affectation d'un alias à la carte d'interface réseau	69
Etape 6. Définition des ports et de leurs options	70
Etape 7. Définition des serveurs avec équilibrage de charge	71
Etape 8. Démarrage de la fonction gestionnaire (facultatif)	71
Etape 9. Démarrage de la fonction conseiller (facultatif)	71
Etape 10. Définition du niveau d'importance des informations requis pour le cluster	72
Configuration des serveurs pour l'équilibrage de la charge	72
Etape 1. Affectation d'un alias pour l'unité de bouclage	72
Etape 2. Vérification de l'existence d'une route supplémentaire	76
Etape 3. Suppression d'une route supplémentaire	77
Etape 4. Vérification de la configuration du serveur	77
Solutions alternatives pour l'affectation d'alias à l'unité de bouclage sous Linux lors de l'utilisation de la méthode d'acheminement MAC de Load Balancer	78

Chapitre 8. Déploiement de Dispatcher sur Load Balancer pour IPv4 et IPv6 . . 81

Plateformes prises en charge pour Load Balancer pour IPv4 et IPv6	82
Plateformes prises en charge pour l'équilibrage de charge dans l'espace utilisateur.	82
Remarques sur la plateforme Linux	82
Restrictions de serveur dorsal	82
Installation de Load Balancer pour IPv4 et IPv6	83
Limitations et remarques spéciales pour Load Balancer pour IPv4 et IPv6	84
Configuration de l'adresse lien-local IPv6	84
Paires cluster/serveur homogènes	84
Fonctions Dispatcher non prises en charge	84
Configuration de conseillers	85
Configuration de la haute disponibilité	85
Co-implantation de serveurs.	86
Fonction d'affinité pour les systèmes qui s'exécutent dans l'espace utilisateur (Linux)	86
Configuration de Metric Server	87
Activation du traitement des paquets IPv6 dans Load Balancer pour IPv4 et IPv6	88

Création d'un alias pour le périphérique d'interface dans Load Balancer pour IPv4 et IPv6	88
Etapes de configuration de cluster requises pour Linux sur zSeries	91
Commandes Dispatcher (dscontrol) pour Load Balancer pour IPv4 et IPv6	92
Différences entre les syntaxes des commandes	92
Commandes dscontrol prises en charge	92
Commandes dscontrol non prises en charge	95

Partie 3. Composant CBR (Content Based Routing) 97

Chapitre 9. Configuration de démarrage rapide 99

Matériel requis	99
Préparation	100
Configuration du composant CBR	100
Configuration à partir de la ligne de commande	100
Test de vérification de la configuration	102
Configuration à l'aide de l'interface graphique	102
Configuration à l'aide de l'assistant de configuration	102
Types de configurations de cluster, de port et de serveur	102

Chapitre 10. Planification de CBR (Content Based Routing) 105

Remarques relatives à la planification	105
Équilibrage de la charge des requêtes pour différents types de contenus	106
Division du contenu de votre site pour améliorer le temps de réponse	106
Copie de sauvegarde du contenu du serveur Web	107
Utilisation de plusieurs processus Caching Proxy pour optimiser l'utilisation de la CPU	107
Équilibrage de charge basé sur des règles avec CBR	107
Équilibrage de charge sur les connexions sécurisées (SSL)	107
Équilibrage de charge client-proxy dans SSL et proxy-serveur dans HTTP	108

Chapitre 11. Configuration de CBR (Content Based Routing) 109

Présentation générale des tâches de configuration	109
Méthodes de configuration	109
Ligne de commande	110
Scripts	111
Interface graphique	112
Assistant de configuration	113
Configuration du poste CBR	113
Etape 1. Configuration de Caching Proxy pour utiliser CBR	114
Etape 2. Démarrage de la fonction serveur	115
Etape 3. Démarrage de la fonction exécuteur	115
Etape 4. Définition et configuration des options du cluster	115

Etape 5. Affectation d'un alias à la carte d'interface réseau (facultatif)	116
Etape 6. Définition des ports et de leurs options	117
Etape 7. Définition des serveurs avec équilibrage de charge	117
Etape 8. Ajout de règles à la configuration	117
Etape 9. Ajout de serveurs à vos règles	117
Etape 10. Démarrage de la fonction gestionnaire (facultatif)	117
Etape 11. Démarrage de la fonction conseiller (facultatif)	118
Etape 12. Définition du niveau d'importance des informations requis pour le cluster	118
Etape 13. Démarrage de Caching Proxy	118
Exemple de configuration CBR	118

Partie 4. Composant Site Selector 121

Chapitre 12. Configuration de démarrage rapide 123

Matériel requis	123
Préparation	123
Configuration du composant Site Selector	124
Configuration à partir de la ligne de commande	124
Test de vérification de la configuration	125
Configuration à l'aide de l'interface graphique	125
Configuration à l'aide de l'assistant de configuration	125

Chapitre 13. Planification de Site Selector 127

Remarques relatives à la planification	127
Considérations relatives à la durée de vie (TTL)	129
Utilisation de la fonction de proximité réseau (Network Proximity)	129

Chapitre 14. Configuration de Site Selector 131

Présentation générale des tâches de configuration	131
Méthodes de configuration	131
Ligne de commande	131
Scripts	132
Interface graphique	132
Assistant de configuration	133
Installation de la machine Site Selector	134
Etape 1. Démarrage de la fonction serveur	134
Etape 2. Démarrage du serveur de noms	134
Etape 3. Définition d'un nom de site et définition des options du nom de site	134
Etape 4. Définition de serveurs avec équilibrage de charge	135
Etape 5. Démarrage de la fonction gestionnaire (facultatif)	135
Etape 6. Démarrage de la fonction conseiller (facultatif)	135
Etape 7. Définition des mesures du système (facultatif)	135
Etape 8. Définition du niveau d'importance des informations pour le nom de site	135

Configuration des serveurs pour l'équilibrage de la charge	136
--	-----

Partie 5. Composant Cisco CSS Controller 137

Chapitre 15. Configuration de démarrage rapide 139

Matériel requis	139
Préparation	139
Configuration du composant Cisco CSS Controller	140
Configuration à partir de la ligne de commande	140
Test de vérification de la configuration	141
Configuration à l'aide de l'interface graphique	141

Chapitre 16. Planification de Cisco CSS Controller 143

Configuration requise	143
Remarques relatives à la planification	143
Positionnement du consultant dans le réseau	144
Haute disponibilité	146
Calcul des pondérations	146
Identification des incidents	147

Chapitre 17. Configuration de Cisco CSS Controller 149

Présentation générale des tâches de configuration	149
Méthodes de configuration	149
Ligne de commande	149
XML	150
Interface graphique	151
Installation de la machine Contrôleur pour commutateurs Cisco CSS	152
Etape 1. Démarrage de la fonction serveur	152
Etape 2. Démarrage de l'interface de ligne de commande	152
Etape 3. Configuration du consultant	152
Etape 3. Configuration d'un contenu de propriétaire	153
Etape 4. Vérification de la définition des services	153
Etape 5. Configuration des mesures	153
Etape 6. Lancement du consultant	153
Etape 7. Lancement du système Metric Server (facultatif)	153
Etape 8. Configuration de la haute disponibilité (facultatif)	153
Test de vérification de la configuration	154

Partie 6. Composant Nortel Alteon Controller 155

Chapitre 18. Configuration de démarrage rapide 157

Matériel requis	157
Préparation	158
Configuration du composant Nortel Alteon Controller	158

Configuration à partir de la ligne de commande	158
Test de vérification de la configuration	159
Configuration à l'aide de l'interface graphique	159

Chapitre 19. Planification de Nortel Alteon Controller 161

Configuration requise	161
Remarques relatives à la planification	161
Positionnement du consultant dans le réseau	162
Attributs de serveur sur le commutateur (définis par le contrôleur)	164
Configuration de serveurs de secours	165
Configuration de groupes	166
Haute disponibilité	166
Optimisation	168
Identification des incidents	169

Chapitre 20. Configuration de Nortel Alteon Controller 171

Présentation générale des tâches de configuration	171
Méthodes de configuration	171
Ligne de commande	171
XML	172
Interface graphique	173
Installation de Nortel Alteon Controller	174
Etape 1. Démarrage de la fonction serveur	174
Etape 2. Démarrage de l'interface de ligne de commande	175
Etape 3. Définition d'un consultant de Nortel Alteon Web Switch	175
Etape 4. Ajout d'un service au consultant de commutateur	175
Etape 5. Configuration des mesures	175
Etape 6. Lancement du consultant	175
Etape 7. Configuration de la haute disponibilité (facultatif)	175
Etape 8. Lancement du système Metric Server (facultatif)	175
Etape 9. Régénération de la configuration de Nortel Alteon Controller	176
Test de vérification de la configuration	176

Partie 7. Fonctions et fonctions avancées de Load Balancer 177

Chapitre 21. Gestionnaire, conseillers et système Metric Server (des composants Dispatcher, CBR et Site Selector) 179

Optimisation de la fonction d'équilibrage de charge fournie par Load Balancer	180
Proportion de l'importance accordée aux données d'état	180
Pondérations	181
Intervalles gestionnaire	183
Seuil de sensibilité	183
Indice de lissage	184

Utilisation de scripts pour la génération d'une alerte ou d'une erreur du serveur d'enregistrement	184
Conseillers	185
Fonctionnement des conseillers	186
Démarrage et arrêt d'un conseiller	186
Intervalles conseiller	187
Délai de rapport du conseiller	187
Délai de connexion du conseiller et délai de réception pour les serveurs	188
Tentative du conseiller	188
Liste des conseillers	188
Configuration du conseiller HTTP ou HTTPS à l'aide de l'option de demande ou de réponse (URL)	190
Utilisation d'un conseiller Self dans une configuration WAN à deux niveaux	191
Création de conseillers personnalisés	192
Conseiller WAS	193
Convention d'attribution de nom	194
Compilation	194
Exécution	194
Sous-programmes requis	195
Ordre de recherche	195
Affectation du nom et du chemin	195
Conseiller type	196
Metric Server	196
Restrictions relatives à WLM	196
Conditions préalables	196
Conditions d'utilisation de Metric Server	196
Conseiller Workload Manager	198
Restrictions relatives à Metric Server	199

Chapitre 22. Fonctions avancées de Dispatcher, CBR et Site Selector . . . 201

Utilisation de serveurs implantés au même endroit	202
Pour le composant Dispatcher	203
Composant CBR	204
Pour le composant Site Selector	204
Haute disponibilité	204
Configuration de la haute disponibilité	205
Détections des incidents en utilisant signal de présence et cible à atteindre	208
Stratégie de reprise	208
Utilisation de scripts	209
Configuration de la co-implantation et de la haute disponibilité (systèmes Windows)	211
Configuration de l'équilibrage de charge basé sur des règles	212
Evaluation des règles	213
Utilisation de règles basées sur l'adresse IP des clients	213
Utilisation de règles basées sur le port du client	214
Utilisation de règles basées sur l'heure	214
Utilisation de règles basées sur le type de services (TOS)	214
Utilisation de règles basées sur le nombre de connexions par seconde	215
Utilisation de règles basées sur le nombre total de connexions actives	215

Utilisation de règles basées sur la largeur de bande réservée et sur la largeur de bande partagée	216
Règle Mesure de tous les serveurs	217
Règle Moyenne des mesures	218
Utilisation de règles toujours vraies	218
Utilisation de règles basées sur le contenu des demandes	219
Substitution d'affinité de port	219
Ajout de règles à la configuration	219
Option d'évaluation de serveur	220
Fonctionnement de la fonction d'affinité pour Load Balancer	221
Comportement lorsque l'affinité est désactivée	221
Comportement lorsque l'affinité est activée	222
Affinité de ports croisés	222
Masque d'adresse de l'affinité (masque de maintien de routage)	222
Mise au repos de la gestion des connexions serveur	223
Option d'affinité de la règle basé sur le contenu de la demande du client	224
Affinité de cookie actif	225
Affinité de cookie passif	226
Affinité d'URI	227
Configuration du support de réseau étendu pour Dispatcher	228
Syntaxe des commandes	229
Utilisation de conseillers éloignés avec le support de réseau étendu de Dispatcher	230
Exemple de configuration	232
Support GRE (Generic Routing Encapsulation)	234
Utilisation de liens explicites	235
Utilisation d'une configuration réseau privée	235
Utilisation d'un cluster générique pour combiner les configurations serveurs	236
Utilisation du cluster générique pour équilibrer la charge des pare-feux	237
Utilisation de cluster générique avec Caching Proxy pour le proxy transparent	238
Utilisation du port générique pour acheminer le trafic destiné à un port non configuré	238
Port générique pour le traitement du trafic FTP	238
Détection d'attaque de refus de service	239
Utilisation de la consignation binaire pour analyser les statistiques des serveurs	240
Utilisation d'un client co-implanté	241

Chapitre 23. Fonctions avancées de Cisco CSS Controller et Nortel Alteon Controller 243

Co-implantation	243
Haute disponibilité	243
Configuration	244
Détection des incidents	245
Stratégie de récupération	245
Exemples	246
Optimisation de la fonction d'équilibrage de charge fournie par Load Balancer	246

Importance accordée aux informations de mesure	246
Pondérations	247
Délai d'inactivité dans le calcul des pondérations	247
Seuil de sensibilité.	248
Conseillers	248
Fonctionnement des conseillers	248
Délai d'inactivité du conseiller.	249
Délai de connexion du conseiller et délai de réception pour les serveurs.	249
Tentative du conseiller	249
Création de conseillers personnalisés	250
Convention d'attribution de nom.	251
Compilation.	251
Exécution.	252
Sous-programmes requis	252
Ordre de recherche	252
Affectation du nom et du chemin.	253
Conseiller type	253
Système Metric Server	253
Conditions préalables	253
Conditions d'utilisation de Metric Server	253
Conseiller Workload manager	255
Utilisation de la consignment binaire pour analyser les statistiques des serveurs	255
Utilisation de scripts pour la génération d'une alerte ou d'une erreur du serveur d'enregistrement.	257

Partie 8. Administration et identification des incidents de Load Balancer 259

Chapitre 24. Exploitation et gestion de Load Balancer 261

Administration à distance de Load Balancer	261
RMI (Remote Method Invocation)	262
administration basée sur le Web	263
Utilisation des journaux Load Balancer.	265
Pour Dispatcher, CBR et Site Selector	265
Pour Cisco CSS Controller et Nortel Alteon Controller	267
Utilisation du composant Dispatcher	268
Démarrage et arrêt de Dispatcher	268
Utilisation de la valeur du délai d'attente	268
Contrôle du nettoyage des enregistrements de connexions à l'aide des paramètres finitimeout et staletimeout	269
Interface graphique — option de menu Contrôler.	269
Utilisation du protocole SNMP (Simple Network Management Protocol, protocole simplifié de gestion de réseau) avec le composant Dispatcher	269
Rejet de l'ensemble du trafic vers Load Balancer avec la fonction ipchains ou iptables (systèmes Linux).	276
Utilisation du composant CBR (Content Based Routing)	276
Démarrage et arrêt de CBR.	277

Contrôle de CBR	277
Utilisation des journaux de CBR	277
Utilisation du composant Site Selector	277
Démarrage et arrêt de Site Selector	277
Contrôle d'Site Selector	277
Utilisation des journaux Site Selector	278
Utilisation du composant Cisco CSS Controller	278
Démarrage et arrêt de Cisco CSS Controller	278
Contrôle de Cisco CSS Controller.	278
Utilisation des journaux Cisco CSS Controller	278
Utilisation du composant Nortel Alteon Controller	278
Démarrage et arrêt de Nortel Alteon Controller	278
Contrôle de Nortel Alteon Controller	278
Utilisation des journaux Nortel Alteon Controller	279
Utilisation du composant Metric Server	279
Démarrage et arrêt de Metric Server.	279
Utilisation des journaux Metric Server	279

Chapitre 25. Résolution des incidents 281

Collecte des informations de résolution des incidents	281
Informations générales (obligatoires)	281
Incidents liés à la haute disponibilité	282
Incidents liés aux conseillers	283
Incident liés au routage par contenu (CBR)	284
Impossibilité d'accéder au cluster.	284
Echec de toutes les tentatives de résolution des incidents	285
Mises à niveau	285
Code Java	285
Liens utiles	285
Tableaux de résolution des incidents.	285
Vérification des numéros de port Dispatcher	299
Vérification des numéros de port CBR	299
Vérification des numéros de port Site Selector	300
Vérification des numéros de port Cisco CSS Controller	301
Vérification des numéros de port Nortel Alteon Controller	301
Résolution des incidents courants—Dispatcher	302
Incident : Dispatcher ne fonctionne pas.	302
Incident : Le répartiteur et le serveur ne répondent pas	302
Incident : Les requêtes Dispatcher ne sont pas équilibrées	302
Incident : La fonction haute disponibilité de Dispatcher est inopérante	303
Incident : Impossible d'ajouter un signal de présence (plateforme Windows)	303
Incident : Routes supplémentaires (Windows 2000)	303
Incident : Les conseillers ne fonctionnent pas correctement	304
Incident : Dispatcher, Microsoft IIS et SSL ne fonctionnent pas (plateformeWindows).	304
Incident : Connexion du répartiteur à une machine éloignée	304
Incident : La commande dscontrol ou lbadmin n'a pas abouti	304

Incident : Affichage du message d'erreur "Fichier introuvable..." lorsque vous tentez de visualiser l'aide en ligne (plateforme Windows)	305
Incident : L'interface graphique ne démarre pas correctement	305
Incident : Erreur lors de l'exécution de Dispatcher lorsque Caching Proxy est installé	305
Incident : L'interface graphique ne s'affiche pas correctement	306
Incident : Sous Windows, les fenêtre d'aide disparaissent parfois sous d'autres fenêtres ouvertes	306
Incident : Load Balancer ne peut pas traiter et transmettre un cadre	306
Incident : Un écran bleu s'affiche lors du démarrage de l'exécuteur Load Balancer	306
Incident : La fonction Path MTU Discovery permet d'éviter le trafic retour avec Load Balancer	306
Incident : La fonction haute disponibilité de Load Balancer en mode réseau étendu est inopérante	307
Incident : Arrêt (ou comportement imprévu) de l'interface graphique lors de la tentative de chargement d'un fichier de configuration volumineux	308
Incident : lbadm se déconnecte du serveur après mise à jour de la configuration	309
Incident : Adresses IP non résolues correctement sur la connexion éloignée	309
Incident : L'interface coréenne de Load Balancer affiche sous AIX et Linux des polices non souhaitées ou qui se chevauchent	309
Incident : Sous Windows, adresse d'alias renvoyée au lieu de l'adresse locale lors de l'émission de commandes telles que hostname	310
Incident : sous Windows, comportement inattendu de l'interface graphique lors de l'utilisation de cartes vidéo Matrox AGP	310
Incident : Comportement inattendu lors de l'exécution de "rmmod ibmlb" (systèmes Linux)	310
Incident : Temps de réponse important lors de l'exécution de commandes sur la machine Dispatcher	310
Incident : Le conseiller SSL ou HTTPS n'enregistre pas les charges des serveurs (avec l'acheminement MAC)	311
Incident : Déconnexion de l'hôte lors du redimensionnement de la fenêtre du navigateur Netscape en cours d'administration Web	311
Incident : Regroupement de connexions activé et serveur Web établissant une liaison à 0.0.0.0	311
Incident : Sous Windows, des caractères nationaux Latin-1 endommagés apparaissent dans la fenêtre de la ligne de commande	312
Incident : Sous HP-UX, la mémoire est insuffisante pour Java et une erreur d'unité d'exécution est générée	312
Incident : Sous Windows, les conseillers et les cibles à contacter marquent tous les serveurs comme étant arrêtés	313

Incident : Sous Windows, la résolution de l'adresse IP en nom d'hôte n'est pas possible lorsque plusieurs adresses sont configurées sur un adaptateur	313
Incident : Sous Windows, les conseillers ne fonctionnent pas dans une configuration en haute disponibilité après une panne réseau	314
Incident : Sous Linux, n'utilisez pas la commande "IP address add" lors de l'affectation d'alias à plusieurs clusters de l'unité de bouclage	315
Incident : Message d'erreur "Adresse de routeur non spécifiée ou non valide pour la méthode port"	315
Incident : Sur les systèmes Solaris, les processus Load Balancer s'arrêtent lorsque vous quittez la fenêtre de terminal à partir de laquelle ils ont été lancés	316
Incident : Délai lors du chargement d'une configuration Load Balancer	316
Incident : Sur les systèmes Windows, un message d'erreur lié à un conflit d'adresses IP apparaît à l'écran	316
Incident : les machines principale et de secours sont toutes deux activées en mode haute disponibilité	317
Incident : Les demandes client échouent lors de la tentative de renvoi de réponses de grande page	317
Incident : Sous Windows, l'erreur "Le serveur ne répond pas" survient lors de l'exécution d'une commande dscontrol ou lbadm	318
Incident : Les machines Dispatcher à haute disponibilité risquent de ne pas être synchronisées sur les systèmes Linux pour S/390 avec des pilotes qeth	318
Incident : Conseils sur la configuration de la haute disponibilité	318
Incident : Sous Linux, limitations de la configuration Dispatcher lors de l'utilisation de serveurs zSeries ou S/390 dotés de cartes OSA (Open System Adapter)	320
Incident : Sur certaines versions Linux, une fuite de mémoire se produit lors de l'exécution de Dispatcher configuré avec le gestionnaire et les conseillers	322
Incident : Sur SUSE Linux Enterprise Server 9, Dispatcher achemine les paquets, mais ceux-ci n'arrivent pas jusqu'au serveur dorsal	322
Incident : Sur le système Windows, un message de conflit d'adresses IP apparaît pendant la reprise de la haute disponibilité	323
Incident : Les iptables de Linux peuvent interférer avec le routage de paquets	324
Incident : Impossible d'ajouter un serveur IPv6 à la configuration Load Balancer sur les systèmes Solaris	324
Un message d'avertissement Java s'affiche lors de l'installation de correctifs de service	324
Mise à niveau de l'ensemble de fichiers Java fourni avec l'installation Load Balancer	325

Résolution des incidents courants—CBR	325	Incident : sous Windows, comportement inattendu de l'interface graphique lors de l'utilisation de cartes vidéo Matrox AGP	332
Incident : CBR ne fonctionne pas	325	Incident : Réception d'une erreur de connexion lors de l'ajout d'un consultant	332
Incident : La commande cbrcontrol ou lbadmin n'a pas abouti	325	Incident : Pondérations non actualisées sur le commutateur	332
Incident : Les requêtes ne sont pas équilibrées	326	Incident : La commande de régénération n'a pas actualisé la configuration du consultant	332
Incident : Sur les systèmes Solaris, la commande cbrcontrol executor start n'aboutit pas	326	Incident : Déconnexion de l'hôte lors du redimensionnement de la fenêtre du navigateur Netscape en cours d'administration Web	333
Incident : erreur de syntaxe ou de configuration	326	Incident : Sous Windows, des caractères nationaux Latin-1 endommagés apparaissent dans la fenêtre de la ligne de commande	333
Incident : sous Windows, comportement inattendu de l'interface graphique lors de l'utilisation de cartes vidéo Matrox AGP	326	Incident : Sous HP-UX, la mémoire est insuffisante pour Java et une erreur d'unité d'exécution est générée	333
Incident : Déconnexion de l'hôte lors du redimensionnement de la fenêtre du navigateur Netscape en cours d'administration Web	327	Résolution des incidents courants—Nortel Alteon Controller	333
Incident : Sous Windows, des caractères nationaux Latin-1 endommagés apparaissent dans la fenêtre de la ligne de commande	327	Incident : nalservice ne démarre pas	333
Incident : Sous HP-UX, la mémoire est insuffisante pour Java et une erreur d'unité d'exécution est générée	327	Incident : la commande nalcontrol ou lbadmin n'a pas abouti	333
Incident : Sous Windows, les conseillers et les cibles à contacter marquent tous les serveurs comme étant arrêtés	327	Incident : Impossible de créer un registre sur le port 14099	334
Incident : Sur les systèmes Windows, la résolution de l'adresse IP en nom d'hôte n'est pas possible lorsque plusieurs adresses sont configurées sur un adaptateur	328	Incident : sous Windows, comportement inattendu de l'interface graphique lors de l'utilisation de cartes vidéo Matrox AGP	334
Résolution des incidents courants—Site Selector	328	Incident : Déconnexion de l'hôte lors du redimensionnement de la fenêtre du navigateur Netscape en cours d'administration Web	335
Incident : Site Selector ne s'exécute pas	328	Incident : Réception d'une erreur de connexion lors de l'ajout d'un consultant	335
Incident : Site Selector ne permet pas le trafic à permutation circulaire à partir des clients Solaris	328	Incident : Pondérations non actualisées sur le commutateur	335
Incident : la commande sscontrol ou lbadmin n'a pas abouti	328	Incident : La commande de régénération n'a pas actualisé la configuration du consultant	335
Incident : Echec du démarrage de sserver sous Windows	329	Incident : Sous Windows, des caractères nationaux Latin-1 endommagés apparaissent dans la fenêtre de la ligne de commande	335
Incident : Site Selector ayant des chemins en double pour lequel l'équilibrage de charge ne s'effectue pas correctement	329	Incident : Sous HP-UX, la mémoire est insuffisante pour Java et une erreur d'unité d'exécution est générée	336
Incident : sous Windows, comportement inattendu de l'interface graphique lors de l'utilisation de cartes vidéo Matrox AGP	329	Résolution des incidents courants—Metric Server	336
Incident : Déconnexion de l'hôte lors du redimensionnement de la fenêtre du navigateur Netscape en cours d'administration Web	330	Incident : IOException Metric Server sous Windows lors de l'exécution de fichiers de mesure utilisateur de format .bat or .cmd	336
Incident : Sous Windows, des caractères nationaux Latin-1 endommagés apparaissent dans la fenêtre de la ligne de commande	330	Incident : Metric Server n'indique pas la charge à la machine Load Balancer	336
Incident : Sous HP-UX, la mémoire est insuffisante pour Java et une erreur d'unité d'exécution est générée	330	Incident : Le journal de la machine Metric Server indique qu'une signature est nécessaire pour accéder à l'agent	337
Incident : Sous Windows, les conseillers et les cibles à contacter marquent tous les serveurs comme étant arrêtés	330	Incident : Sur les systèmes AIX, lorsque Metric Server s'exécute dans des conditions difficiles, il est possible que le résultat de la commande ps -vg soit altéré	337
Résolution des incidents courants—Cisco CSS Controller	331	Incident : Configuration de Metric Server dans une configuration de second niveau avec équilibrage de la charge entre des machines Dispatcher haute disponibilité par Site Selector	337
Incident : ccoserver ne démarre pas	331		
Incident : La commande ccocontrol ou lbadmin n'a pas abouti	331		
Incident : Impossible de créer un registre sur le port 13099	331		

Incident : Les scripts exécutés sur des machines Solaris dotées de plusieurs CPU génèrent des messages de console non souhaités	339
Incident : Avec Load Balancer pour IPv6, extraction impossible de valeurs de Metric Server sur des systèmes Linux.	339
Incident : Après le démarrage de Metric Server, la valeur de mesure renvoie -1	340

Partie 9. Guide des commandes 341

Chapitre 26. Lecture d'un schéma de syntaxe 343

Symboles et ponctuation.	343
Paramètres	343
Exemples de syntaxe	343

Chapitre 27. Guide des commandes Dispatcher et CBR 345

Différences de configuration entre CBR et Dispatcher	346
dscontrol advisor — Contrôle du conseiller	347
dscontrol binlog — Contrôle du fichier journal binaire.	352
dscontrol cluster — Configuration des clusters	353
dscontrol executor — Contrôle de l'exécuteur	357
dscontrol file — Gestion des fichiers de configuration	362
dscontrol help — Affichage ou impression de l'aide relative à cette commande	364
dscontrol highavailability — Contrôle de la haute disponibilité.	365
dscontrol host — Configuration d'une machine éloignée	369
dscontrol logstatus — Affichage des paramètres du journal du serveur.	370
dscontrol manager — Contrôle du gestionnaire	371
dscontrol metric — Configuration des mesures du système	377
dscontrol port — Configuration des ports	378
dscontrol rule — Configuration des règles.	384
dscontrol server — Configuration des serveurs	390
dscontrol set — Configuration du journal du serveur	396
dscontrol status — Indique par affichage si le gestionnaire et les conseillers sont en cours d'exécution	397
dscontrol subagent — Configuration du sous-agent SNMP.	398

Chapitre 28. Guide des commandes Site Selector. 401

sscontrol advisor — Contrôle du conseiller	402
sscontrol file — Gestion des fichiers de configuration	407
sscontrol help — Affichage ou impression de l'aide relative à cette commande	409
sscontrol logstatus — Affichage des paramètres du journal du serveur.	410

sscontrol manager — Contrôle du gestionnaire	411
sscontrol metric — Configuration des mesures du système	416
sscontrol nameserver — Contrôle de NameServer	417
sscontrol rule — Configuration des règles	418
sscontrol server — Configuration des serveurs	421
sscontrol set — Configuration du journal du serveur	423
sscontrol sitename — Configuration d'un nom de site	424
sscontrol status — Affiche si le gestionnaire et les conseillers sont en cours d'exécution	427

Chapitre 29. Guide des commandes Cisco CSS Controller 429

cococontrol consultant — Configuration et contrôle d'un consultant.	430
cococontrol controller — Gestion du contrôleur	433
cococontrol file — Gestion des fichiers de configuration	435
cococontrol help — Affichage ou impression de l'aide relative à cette commande	437
cococontrol highavailability — Contrôle de la haute disponibilité.	438
cococontrol metriccollector — Configuration du programme de collecte de mesures	441
cococontrol ownercontent — Contrôle du nom de propriétaire et de la règle de contenu	443
cococontrol service — Configuration d'un service	446

Chapitre 30. Guide des commandes Nortel Alteon Controller 449

nalcontrol consultant — Configuration et contrôle d'un consultant.	450
nalcontrol controller — Gestion du contrôleur	453
nalcontrol file — Gestion des fichiers de configuration	455
nalcontrol help — Affichage ou impression de l'aide relative à cette commande	457
nalcontrol highavailability — Contrôle de la haute disponibilité.	458
nalcontrol metriccollector — Configuration du programme de collecte de mesures	461
nalcontrol server — Configuration d'un serveur	463
nalcontrol service — Configuration d'un service	465

Annexe A. Interface graphique utilisateur : Instructions générales . . 469

Annexe B. Syntaxe des règles de contenu (modèle). 477

Syntaxe de règle de contenu (modèle) :	477
Mots clés réservés	477

Annexe C. Exemples de fichiers de configuration 481

Exemples de fichiers de configuration Load Balancer	481
---	-----

Dispatcher Fichier de configuration — systèmes AIX, Linux et Solaris	481
Dispatcher Fichier de configuration — systèmes Windows.	484
Conseiller type	487

Annexe D. Exemple de configuration de haute disponibilité à deux niveaux utilisant Dispatcher, CBR et Caching Proxy	493
--	------------

Configuration de la machine serveur	493
---	-----

Annexe E. Remarques	497
Marques	499

Glossaire	501
----------------------------	------------

Index	511
------------------------	------------

Tableaux

1. Images installp d'AIX	32	11. Tâches de configuration pour le composant Nortel Alteon Controller	171
2. Commandes d'installation AIX	34	12. Tâches de configuration avancées pour Load Balancer	179
3. Détails de l'installation des packages de Load Balancer sous HP-UX	35	13. Tâches de configuration avancées pour Load Balancer	201
4. Tâches de configuration pour la fonction Dispatcher	63	14. Tableau de résolution des incidents de Dispatcher	286
5. Commandes pour l'affectation d'un alias à l'unité de bouclage (lo0) pour Dispatcher	73	15. Tableau de résolution des incidents de CBR	292
6. Commandes de suppression d'une route supplémentaire pour Dispatcher	77	16. Tableau de résolution des incidents de Site Selector	293
7. Tâches de configuration pour le composant CBR	109	17. Tableau de résolution des incidents de Contrôleur pour commutateurs Cisco CSS	295
8. Commandes pour l'affectation d'un alias à la carte d'interface réseau	116	18. Tableau de résolution des incidents de Nortel Alteon Controller	296
9. Tâches de configuration pour le composant Site Selector	131	19. Tableau de dépannage du système Metric Server	297
10. Tâches de configuration du composant Cisco CSS Controller	149		

Figures

1. Exemple de représentation physique d'un site utilisant Dispatcher pour gérer les serveurs locaux	10	25. Configuration Cisco CSS Controller simple	139
2. Exemple de site utilisant Dispatcher et Metric Server pour gérer les serveurs	11	26. Exemple de consultant connecté derrière les commutateurs	145
3. Exemple de site utilisant Dispatcher pour gérer des serveurs locaux et éloignés	12	27. Exemple de consultant (avec partenaire haute disponibilité optionnel), configuré derrière le commutateur avec une interface graphique devant le commutateur	146
4. Exemple de site utilisant CBR pour gérer les serveurs locaux	13	28. Configuration Nortel Alteon Controller simple	157
5. Exemple de site utilisant Site Selector et Metric Server pour gérer les serveurs locaux et éloignés	15	29. Exemple de consultant connecté derrière le commutateur	163
6. Exemple de site utilisant Cisco CSS Controller et Metric Server pour gérer les services locaux .	17	30. Exemple de consultant connecté via un intranet situé devant le commutateur	164
7. Exemple de site utilisant Nortel Alteon Controller pour gérer les serveurs locaux . . .	18	31. Exemple de consultant derrière le commutateur et d'interface graphique devant le commutateur	164
8. Configuration Dispatcher locale simple	45	32. Exemple de consultant configuré avec des serveurs de secours	166
9. Exemple de composant Dispatcher configuré avec un cluster et 2 ports	48	33. Exemple de Nortel Alteon Controller et de Nortel Alteon Web Switch haute disponibilité .	168
10. Exemple de composant Dispatcher configuré avec deux clusters, chacun étant associé à un port	49	34. Exemple de configuration WAN à deux niveaux utilisant le conseiller self	192
11. Exemple de composant Dispatcher configuré avec 2 clusters, chacun étant associé à 2 ports .	50	35. Exemple de configuration consistant en un seul segment de réseau local	229
12. Exemple d'utilisation des méthodes d'acheminement nat ou cbr de Dispatcher . . .	57	36. Exemple de configuration utilisant des serveurs locaux et éloignés	229
13. Exemple de Dispatcher utilisant la haute disponibilité	60	37. Exemple de configuration en réseau étendu avec des composants Load Balancer éloignés .	232
14. Exemple de Dispatcher utilisant la haute disponibilité réciproque	61	38. Exemple de configuration en réseau étendu avec une plateforme serveur prenant en charge GRE	234
15. Exemple d'adresses IP nécessaires pour la machine Dispatcher	68	39. Exemple de réseau privé utilisant Dispatcher	236
16. Configuration CBR locale simple	99	40. Commandes SNMP pour les systèmes Linux et UNIX	270
17. Exemple de composant CBR configuré avec un cluster et 2 ports	102	41. Interface graphique affichant l'arborescence du composant Dispatcher	470
18. Exemple de composant CBR configuré avec deux clusters, chacun étant associé à un port .	103	42. Interface graphique affichant l'arborescence du composant CBR	471
19. Exemple de composant CBR configuré avec 2 clusters, chacun étant associé à 2 ports . . .	104	43. Interface graphique affichant l'arborescence du composant Site Selector	472
20. Fichier de configuration CBR pour les systèmes AIX, Linux et Solaris	114	44. Interface graphique affichant l'arborescence du composant Cisco CSS Controller	473
21. Fichier de configuration CBR pour les systèmes HP-UX	115	45. Interface graphique affichant l'arborescence du composant Nortel Alteon Controller . .	474
22. Fichier de configuration CBR pour les systèmes Windows	115	46. Exemple de configuration de haute disponibilité à deux niveaux utilisant Dispatcher, CBR et Caching Proxy	493
23. Configuration Site Selector simple	123		
24. Exemple d'environnement DNS	127		

A propos de ce guide

Le présent document explique comment installer, configurer, utiliser et dépanner IBM WebSphere Application Server Load Balancer pour AIX, HP-UX, Linux, Solaris et Windows. Ce produit était connu sous le nom Edge Server Network Dispatcher, SecureWay Network Dispatcher, eNetwork Dispatcher et Interactive Network Dispatcher.

A qui s'adresse ce guide

Le *Guide d'administration de Load Balancer* s'adresse à des administrateurs réseau et système chevronnés, qui connaissent parfaitement leurs systèmes d'exploitation et les services Internet fournis. Aucune connaissance préalable de Load Balancer n'est requise.

Ce manuel n'assure pas le support des versions antérieures de Load Balancer.

Informations connexes

Le site Web du centre de documentation Edge Components propose un lien vers la version courante du présent manuel aux formats HTML et PDF.

Pour obtenir les mises à jour les plus récentes de Load Balancer, visitez la page Web d'assistance et le lien vers le site Technote.

Pour accéder à ces sites ainsi qu'aux pages Web associées, cliquez sur les liens répertoriés à la section «Documents associés et sites Web», à la page xix.

Accessibilité

Des fonctions d'accessibilité permettent aux personnes à mobilité réduite ou malvoyantes d'utiliser sans problème les logiciels. Les principales fonctions d'accessibilité de Load Balancer sont les suivantes :

- Vous pouvez utiliser un lecteur d'écran ou un synthétiseur vocal pour entendre prononcer ce qui s'affiche à l'écran. Vous pouvez également utiliser un logiciel de reconnaissance vocale, tel que IBM ViaVoice, pour entrer des données et naviguer dans l'interface graphique.
- Vous pouvez activer les fonctions avec le clavier au lieu de la souris.
- Vous pouvez configurer et administrer les fonctions de Load Balancer à l'aide d'éditeurs de texte standard ou d'interfaces de ligne de commande, au lieu des interfaces graphiques fournies. Pour plus d'informations sur l'accessibilité d'une fonction particulière, reportez-vous à la documentation relative à cette fonction.

Comment envoyer vos commentaires

Vos commentaires sont importants dans la mesure où ils nous aident à offrir des informations précises et de qualité. Pour tout commentaire sur ce manuel ou sur toute autre documentation Edge :

- Envoyez un courrier électronique à fsdoc@us.ibm.com. N'oubliez pas d'inclure le nom du manuel, son numéro de référence, la version et éventuellement l'emplacement du texte sur lequel porte votre commentaire (par exemple, un numéro de page ou de tableau).

Documents associés et sites Web

- *Concepts, planification et installation pour Edge Components* GC11-2539-00
- *Programming Guide for Edge Components* GC31-6919-00
- *Caching Proxy - Guide d'administration* GC11-2540-00
- Site Web d'accueil IBM : www.ibm.com/
- Produit IBM WebSphere Application Server : www.ibm.com/software/webservers/appserv/
- Site Web de la bibliothèque IBM WebSphere Application Server : www.ibm.com/software/webservers/appserv/was/library/
- Site Web d'assistance IBM WebSphere Application Server : www.ibm.com/software/webservers/appserv/was/support/
- Centre de documentation IBM WebSphere Application Server : www.ibm.com/software/webservers/appserv/infocenter.html
- Centre de documentation IBM WebSphere Application Server Edge Components : www.ibm.com/software/webservers/appserv/ecinfocenter.html

Partie 1. Introduction à Load Balancer

Cette section présente Load Balancer et ses composants, décrit en détail les fonctions de configuration disponibles, répertorie les matériels et logiciels requis et fournit des instructions d'installation. Elle se compose des chapitres suivants :

- Chapitre 1, «Présentation générale de Load Balancer», à la page 3
- Chapitre 2, «Présentation générale des composants de Load Balancer», à la page 9
- Chapitre 3, «Gestion du réseau : Fonctions Load Balancer requises», à la page 19
- Chapitre 4, «Installation de Load Balancer», à la page 31

Chapitre 1. Présentation générale de Load Balancer

Ce chapitre contient une présentation générale de Load Balancer et comporte les sections suivantes :

- «Principe de Load Balancer»
- «Composant(s) de Load Balancer utilisable(s)»
- «Avantages de Load Balancer», à la page 4
- «Haute disponibilité fournie par Load Balancer», à la page 6
- «Nouvelles fonctions», à la page 6

Le Chapitre 3, «Gestion du réseau : Fonctions Load Balancer requises», à la page 19 contient une liste complète des fonctions avancées de configuration fournies par chacun des composants de Load Balancer, qui vous permettra de déterminer celles qui sont les mieux adaptées à la gestion de votre réseau.

Principe de Load Balancer

Load Balancer est une solution logicielle de distribution des demandes client entrantes à plusieurs serveurs. Il permet d'optimiser les performances en orientant les demandes de session TCP/IP vers différents serveurs au sein d'un groupe, assurant ainsi une répartition équilibrée des demandes entre tous les serveurs. Cette procédure d'équilibrage de charge est parfaitement transparente, tant pour l'utilisateur que pour les applications. Load Balancer s'avère particulièrement utile pour les applications telles que les serveurs de messagerie électronique, les serveurs Internet (WWW), les demandes de bases de données parallèles distribuées et autres applications TCP/IP.

Appliqué à des serveurs Web, Load Balancer peut contribuer à accroître les capacités d'un site en apportant une solution puissante, souple et modulable aux incidents liés à la surcharge des réseaux. Si les visiteurs ne peuvent pas accéder à votre site pendant les périodes d'affluence, Load Balancer peut déterminer automatiquement le serveur le mieux placé pour traiter les demandes en instance. De cette manière, la rentabilité de votre site augmente en même temps que la satisfaction de vos clients.

Composant(s) de Load Balancer utilisable(s)

IMPORTANT : Si vous utilisez Load Balancer pour IPv4 et IPv6, seul le composant Dispatcher est disponible. Pour plus d'informations, voir Chapitre 8, «Déploiement de Dispatcher sur Load Balancer pour IPv4 et IPv6», à la page 81.

Load Balancer comprend les cinq composants suivants qui, employés conjointement ou séparément, vous permettront d'obtenir les meilleurs résultats en matière d'équilibrage de charge :

- Le composant **Dispatcher** peut être utilisé seul pour équilibrer la charge des serveurs, dans le cadre d'un réseau local ou étendu, sur la base de mesures et d'évaluations définies de façon dynamique par Dispatcher. Ce composant assure l'équilibrage des charges au niveau des services spécifiques tels que HTTP, FTP, SSL, NNTP, IMAP, POP3, SMTP, SIP et Telnet. Il n'utilise pas de serveur de noms de domaine pour associer noms de domaine et adresses IP.

Pour le protocole HTTP, vous pouvez utiliser le composant fonction CBR de Dispatcher pour équilibrer la charge à partir du contenu de la demande du client. Le choix du serveur est fonction du résultat de la comparaison de l'URL à une règle donnée. Le routage en fonction du contenu de Dispatcher (méthode d'acheminement cbr) ne requiert *pas* Caching Proxy.

- Pour les protocoles HTTP et HTTPS (SSL), vous pouvez utiliser le composant **Content Based Routing** (CBR) pour équilibrer la charge à partir du contenu de la demande du client. Un client envoie une demande à Caching Proxy, qui la transmet au serveur approprié. Le choix du serveur est fonction du résultat de la comparaison de l'URL à une règle donnée.

Remarque : Le composant Content Based Routing (CBR) est disponible sur toutes les plateformes prises en charge à l'exception de celles exécutant une JVM 64 bits. Vous avez également la possibilité d'employer la méthode d'acheminement cbr du composant Dispatcher de Load Balancer pour permettre le routage CBR sans faire appel à Caching Proxy. Pour plus de détails, voir «Fonction CBR de Dispatcher (méthode d'acheminement cbr)», à la page 55.

- Le composant **Site Selector** peut être utilisé pour équilibrer la charge des serveurs, dans le cadre d'un réseau local ou étendu, par permutation circulaire des serveurs de noms de domaines ou par une méthode plus évoluée définie par l'utilisateur. Un serveur de noms de domaine est associé à Site Selector. Il met les noms DNS en correspondance avec les adresses IP.
- Vous pouvez utiliser le composant **Cisco CSS Controller** ou **Nortel Alteon Controller** qui génère des mesures de pondération de serveur respectivement envoyées au composant Cisco CSS Switch ou Nortel Alteon Web Switch pour une sélection optimale des serveurs, l'optimisation des charges et la tolérance aux pannes.

Pour plus d'informations sur les composants Dispatcher, CBR, Site Selector, Cisco CSS Controller et Nortel Alteon Controller, voir «Composants de Load Balancer», à la page 9.

Avantages de Load Balancer

Le nombre d'utilisateurs et de réseaux qui se connectent au réseau mondial Internet connaît une croissance exponentielle. Cette croissance entraîne des problèmes d'évolutivité pouvant limiter l'accès des utilisateurs aux sites les plus fréquentés.

Actuellement, les administrateurs de réseau utilisent diverses méthodes pour optimiser l'accessibilité. Avec certaines de ces méthodes, vous pouvez, par exemple, sélectionner un autre serveur de manière aléatoire lorsque le premier choisi répond trop lentement ou ne répond pas. Cette approche est peu pratique, peu conviviale et inefficace. Autre méthode utilisée, l'approche circulaire standard, dans laquelle le serveur de noms de domaine sélectionne tour à tour des serveurs pour traiter les demandes. Cette approche est sans doute meilleure que la première, mais reste inefficace dans la mesure où l'acheminement du trafic s'effectue sans tenir compte de la charge des serveurs. En outre, même si un serveur est défaillant, les demandes continuent de lui être adressées.

Né de ce besoin d'une solution plus puissante, Load Balancer apporte nombre d'améliorations par rapport aux solutions antérieures comparables :

Modularité

Pour répondre à l'augmentation du nombre de demandes client, IND permet d'ajouter des serveurs de manière dynamique, ouvrant ainsi l'accès à des dizaines de millions de demandes chaque jour sur des dizaines, voire des centaines, de serveurs.

Utilisation optimale des équipements

L'équilibrage de charge permet à chaque groupe de serveurs d'utiliser ses ressources matérielles de manière optimale en réduisant les surcharges qui se produisent fréquemment avec une méthode de permutation de serveurs circulaire classique.

Facilité d'intégration

Load Balancer utilise les protocoles TCP/IP ou UDP/IP standard. Il peut être ajouté à n'importe quel réseau sans qu'aucune modification matérielle soit nécessaire. C'est un produit simple à installer et à configurer.

Faible charge induite

Avec la méthode d'acheminement de niveau MAC simple qu'il utilise, Dispatcher se contente de surveiller les transmissions entrantes du client vers le serveur. Il n'effectue aucun contrôle des transmissions en sortie, du serveur vers le client. Si l'on compare à d'autres méthodes, cet aspect réduit sensiblement son impact sur les performances des applications et permet même d'accroître celles du réseau.

Haute disponibilité

Les composants Dispatcher, Cisco CSS Controller et Nortel Alteon Controller disposent d'une fonction intégrée assurant une haute disponibilité ; à tout moment, une machine de secours peut assurer l'équilibrage de charge en cas de défaillance du serveur principal. Si l'un des serveurs ne répond plus, le traitement des demandes se poursuit sur un autre serveur. L'arrêt d'un serveur ne constitue plus une défaillance majeure et le site conserve ainsi sa haute disponibilité.

Pour plus d'informations, voir «Haute disponibilité fournie par Load Balancer», à la page 6

Acheminement sur la base du contenu (avec le composant CBR ou Dispatcher)

Associé à Caching Proxy, le composant CBR peut relayer les demandes HTTP et HTTPS (SSL) vers des serveurs spécifiques en fonction du contenu demandé. Par exemple, si une demande contient la chaîne `"/cgi-bin/"` dans la partie répertoire de l'URL et que le serveur est un serveur local, CBR peut acheminer la demande vers un ensemble de serveurs spécialisés dans les demandes cgi et choisir parmi ceux-ci le serveur optimal.

Remarque : Le composant Content Based Routing (CBR) est disponible sur toutes les plateformes prises en charge à l'exception de celles exécutant une JVM 64 bits. Vous avez également la possibilité d'employer la méthode d'acheminement cbr du composant Dispatcher de Load Balancer pour permettre le routage CBR sans faire appel à Caching Proxy. Pour plus de détails, voir «Fonction CBR de Dispatcher (méthode d'acheminement cbr)», à la page 55.

Le composant Dispatcher assure aussi l'acheminement sur la base du contenu, mais ne nécessite pas que Caching Proxy soit installé. Etant donné que la fonction CBR du composant Dispatcher est exécutée dans le noyau à

la réception des paquets, l'acheminement est plus rapide que celui réalisé par le composant CBR. Le composant Dispatcher exécute la fonction fonction CBR (content-based routing) pour HTTP (avec la règle de type de contenu) et HTTPS (avec l'affinité des ID de session).

Remarque : Seul le composant CBR peut utiliser la règle de contenu pour HTTPS (SSL) pour un équilibrage de charge effectué sur la base du contenu de la demande HTTP, ce qui nécessite un déchiffrement et nouveau chiffrement des messages.

Haute disponibilité fournie par Load Balancer

Dispatcher

Dispatcher offre une fonctionnalité de haute disponibilité intégrée, de sorte que Dispatcher ne constitue plus un point unique de défaillance dans votre réseau. Cette dernière implique l'utilisation d'une deuxième machine Dispatcher, chargée de contrôler la machine principale (également appelée machine principale), et qui reste en attente, prête à assurer l'équilibrage de charge en cas d'incident sur la machine principale. Le composant Dispatcher offre également la fonction de haute disponibilité réciproque qui permet à deux machines de travailler simultanément en mode principal et secondaire l'une avec l'autre. Voir «Configuration de la haute disponibilité», à la page 205.

CBR

Vous pouvez également atteindre un niveau de haute disponibilité avec le composant CBR lorsque vous utilisez une configuration à deux niveaux avec une machine Dispatcher répartissant la charge sur plusieurs serveurs dotés de CBR.

Cisco CSS Controller ou Nortel Alteon Controller

Les contrôleurs étant dotés d'une fonctionnalité de haute disponibilité, chaque contrôleur ne constitue plus un point unique de défaillance. Le contrôleur d'un poste peut être configuré en tant que contrôleur principal et celui d'un autre poste en tant que contrôleur de secours. Le contrôleur de secours surveille le contrôleur principal et se tient prêt à fournir aux commutateurs les mesures de pondération de serveur adéquates en cas de défaillance du contrôleur principal. Pour plus d'informations, voir «Haute disponibilité», à la page 243.

Nouvelles fonctions

Load Balancer pour IBM WebSphere Application Server Version 6.1 contient un certain nombre de nouvelles fonctions. Les nouveautés les plus importantes sont présentées ci-après.

- **Prise en charge de l'exécution de processus d'équilibrage de charge dans l'espace utilisateur sur les systèmes Linux**

Une prise en charge a été ajoutée pour que les installations Load Balancer pour IPv4 et IPv6 puissent exécuter des processus d'équilibrage de charge dans l'espace utilisateur, plutôt que dans l'espace noyau. Les systèmes Linux perdent alors toute dépendance vis à vis du module du noyau.

Pour les informations les plus récentes sur les systèmes acceptant l'équilibrage de charge dans l'espace utilisateur (hors noyau), voir le site Web suivant : <http://www.ibm.com/support/docview.wss?rs=180&uid=swg27006921>.

Pour plus d'informations, voir «Plateformes prises en charge pour Load Balancer pour IPv4 et IPv6», à la page 82.

- **Prise en charge de HP 11iv2 sur PA-RISC (prise en charge de HP 11iv1 supprimée)**

Pour plus d'informations sur les conditions matérielles et logicielles prises en charge, accédez au site Web suivant : <http://www.ibm.com/support/docview.wss?rs=180&uid=swg27006921>

- **Prise en charge des systèmes Linux sur zSeries 64 bits**

La prise en charge des systèmes Linux sur zSeries 64 bits est fournie uniquement pour les installations Load Balancer pour IPv4 et IPv6.

Pour des informations sur Load Balancer pour IPv4 et IPv6 et des remarques sur l'exécution des systèmes Linux sur zSeries 64 bits, voir Chapitre 8, «Déploiement de Dispatcher sur Load Balancer pour IPv4 et IPv6», à la page 81.

Pour plus d'informations sur les conditions matérielles et logicielles prises en charge, accédez au site Web suivant : <http://www.ibm.com/support/docview.wss?rs=180&uid=swg27006921>

- **Prise en charge du conseiller SIP**

La prise en charge d'un conseiller SIP a été ajoutée. Le conseiller SIP qui est pris en charge s'exécute sur protocole TCP uniquement.

Pour plus d'informations, voir la page 188.

- **Pour les systèmes Linux, prise en charge des configurations de client co-implanté**

Cette fonction s'applique à tous les composants de Load Balancer.

Seuls les systèmes Linux acceptent des configurations dans lesquelles le client réside sur la même machine que Load Balancer.

Pour plus d'informations, voir «Utilisation d'un client co-implanté», à la page 241.

- **Prise en charge du navigateur Firefox**

Pour plus d'informations sur les versions de Firefox prises en charge et sur tous les navigateurs acceptés, accédez au site Web suivant : <http://www.ibm.com/support/docview.wss?rs=180&uid=swg27006921>.

Chapitre 2. Présentation générale des composants de Load Balancer

Ce chapitre contient une présentation générale des composants de Load Balancer et comporte les sections suivantes :

- «Composants de Load Balancer»
- «Présentation générale du composant Dispatcher»
- «Présentation générale du composant CBR (Content Based Routing)», à la page 12
- «Présentation générale du composant Site Selector», à la page 14
- «Présentation générale du composant Cisco CSS Controller», à la page 15
- «Présentation générale du composant Nortel Alteon Controller», à la page 17

Le Chapitre 3, «Gestion du réseau : Fonctions Load Balancer requises», à la page 19 contient une liste complète des fonctions avancées de configuration fournies par chacun des composants de Load Balancer, qui vous permettra de déterminer celles qui sont les mieux adaptées à la gestion de votre réseau.

Composants de Load Balancer

Les cinq composants de Load Balancer sont Dispatcher, Content Based Routing (CBR), Site Selector, Cisco CSS Controller et Nortel Alteon Controller. Load Balancer permet d'utiliser ces composants séparément ou ensemble, selon la configuration de votre site. La section qui suit présente ces composants.

IMPORTANT : Si vous utilisez Load Balancer pour IPv4 et IPv6, seul le composant Dispatcher est disponible. Pour plus d'informations, voir Chapitre 8, «Déploiement de Dispatcher sur Load Balancer pour IPv4 et IPv6», à la page 81.

Présentation générale du composant Dispatcher

Le composant Dispatcher assure l'équilibrage de la charge du trafic entre les serveurs via une combinaison unique de logiciels d'équilibrage de charge et de gestion. Dispatcher peut aussi détecter l'échec d'un serveur et canaliser le trafic sur les serveurs qui l'entourent. Dispatcher prend en charge les protocoles HTTP, FTP, SSL, SMTP, NNTP, IMAP, POP3, Telnet, SIP et toute application de type TCP ou UDP sans état.

Toutes les demandes de client adressées à la machine Dispatcher sont acheminées vers le serveur le mieux adapté, compte tenu de mesures définies de façon dynamique. Vous pouvez utiliser les valeurs par défaut associées à ces mesures ou les modifier au cours du processus de configuration.

Dispatcher offre trois méthodes d'acheminement (indiquées sur le port) :

- Méthode d'acheminement MAC (**mac**). Cette méthode d'acheminement permet à Dispatcher d'équilibrer la charge de la demande entrante adressée au serveur. Le serveur renvoie la réponse directement au client sans l'intervention de Dispatcher.
- Méthode d'acheminement NAT/NAPT (**nat**). Si vous utilisez NAT/NAPT (Network Address Translation/Network Address Port Translation), il n'est pas

nécessaire que les serveurs principaux se trouvent sur un réseau local. Si vous préférez disposer de serveurs éloignés, utilisez la méthode d'acheminement nat plutôt que la technique GRE/WAN (Generic Routing Encapsulation/Wide Area Network). Cette méthode d'acheminement permet à Dispatcher d'équilibrer la charge de la demande entrante adressée au serveur. Le serveur renvoie la réponse à Dispatcher. La machine Dispatcher renvoie ensuite la réponse au client.

- Méthode d'acheminement Content-Based Routing (**CBR**). Sans Caching Proxy, le composant Dispatcher permet d'exécuter la fonction CBR (content-based routing) pour HTTP (avec la règle de type de contenu) et HTTPS (avec l'affinité des ID de session SSL). Pour le trafic HTTP et HTTPS, le composant Dispatcher peut fournir une fonction CBR (content-based routing) plus rapide que le composant CBR. Cette méthode d'acheminement permet à Dispatcher d'équilibrer la charge de la demande entrante adressée au serveur. Le serveur renvoie la réponse à Dispatcher. La machine Dispatcher renvoie ensuite la réponse au client.

Le composant Dispatcher constitue la clé de voûte d'une gestion efficace et durable d'un réseau de serveurs étendu et modulable. Avec Dispatcher, vous pouvez lier différents serveurs en un seul serveur virtuel. De cette manière, le site est associé à une adresse IP unique. Dispatcher fonctionne indépendamment de tout serveur de noms de domaine ; toutes les demandes sont dirigées sur l'adresse IP de la machine Dispatcher.

Dispatcher présente des avantages spécifiques indéniables en matière d'équilibrage de charge sur des serveurs en cluster. Ces atouts permettent de mettre en oeuvre une gestion de site aussi efficace que stable.

Gestion de serveurs locaux avec Dispatcher

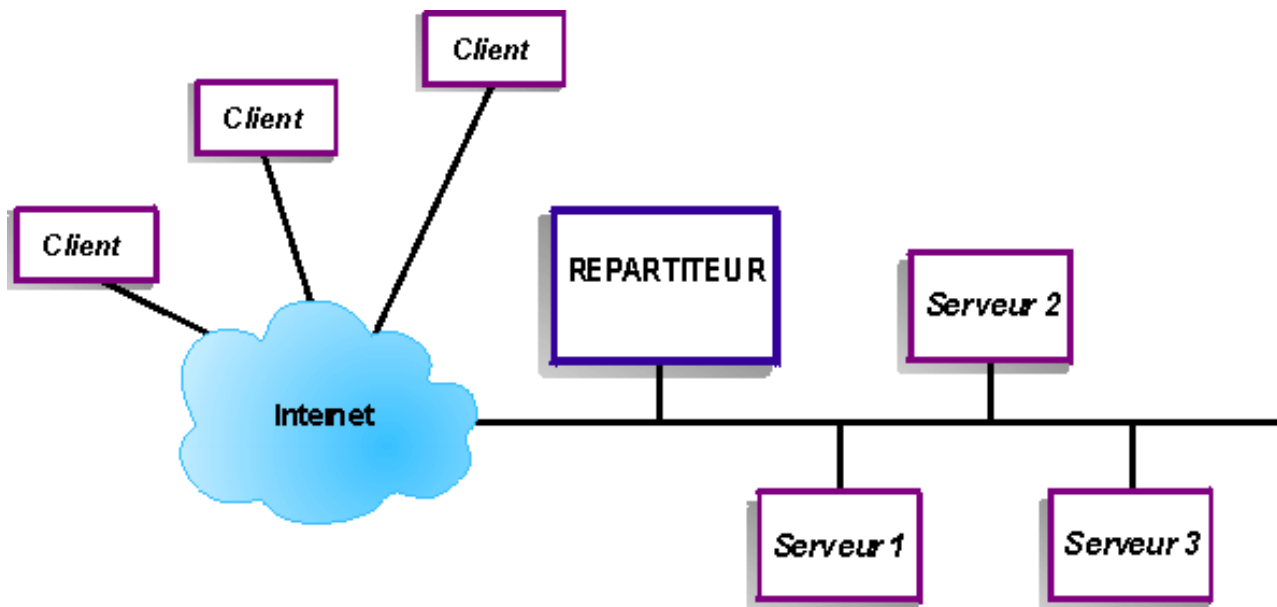


Figure 1. Exemple de représentation physique d'un site utilisant Dispatcher pour gérer les serveurs locaux

La figure 1 est la représentation physique d'un site utilisant une configuration de réseau Ethernet. La machine Dispatcher peut être installée sans apporter de changement physique à la physionomie du réseau. Après acheminement d'une demande de client au serveur optimal par Dispatcher, la réponse correspondante

est transmise directement du serveur au client sans intervention de Dispatcher lorsque la méthode d'acheminement MAC est utilisée.

Gestion des serveurs à l'aide de Dispatcher et de Metric Server

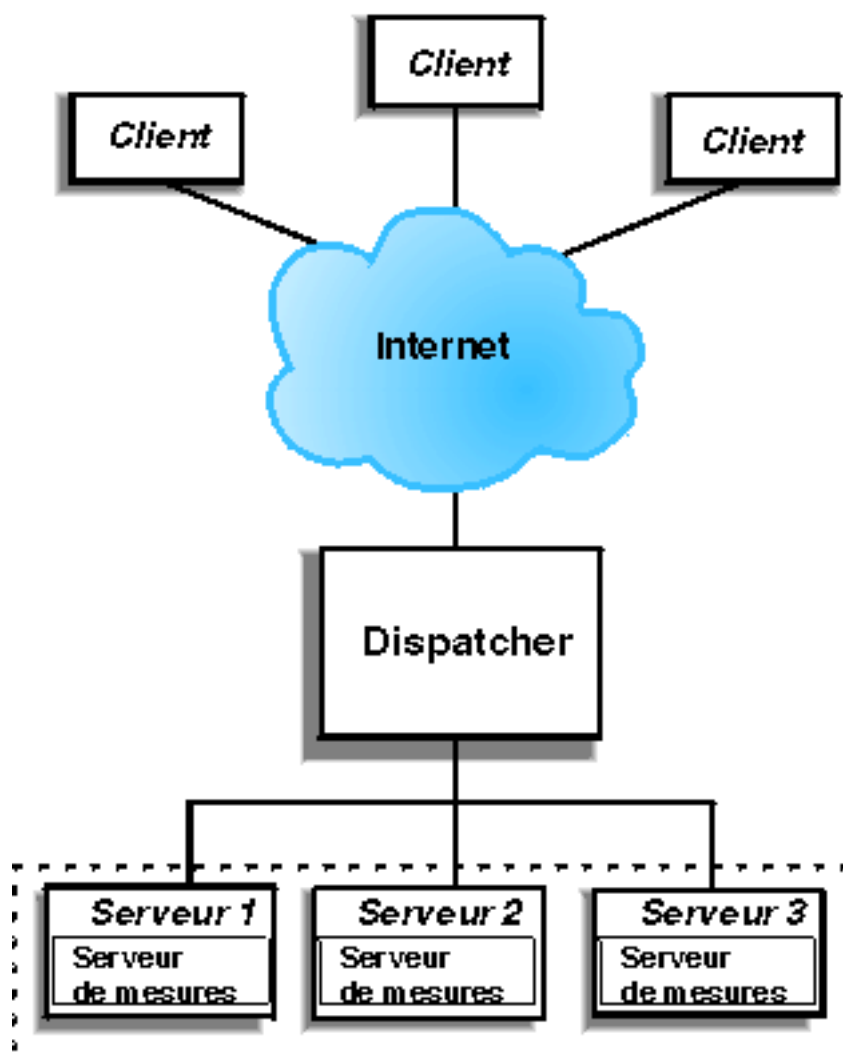


Figure 2. Exemple de site utilisant Dispatcher et Metric Server pour gérer les serveurs

La figure 2 représente un site dans lequel tous les serveurs se trouvent sur un réseau local. Le composant Dispatcher permet d'acheminer les demandes et le composant Metric Server permet de fournir les informations de charge du système au poste Dispatcher.

Dans cet exemple, le démon Metric Server est installé sur chaque serveur principal. Vous pouvez utiliser Metric Server avec le composant Dispatcher ou tout autre composant Load Balancer.

Gestion de serveurs locaux et éloignés avec Dispatcher

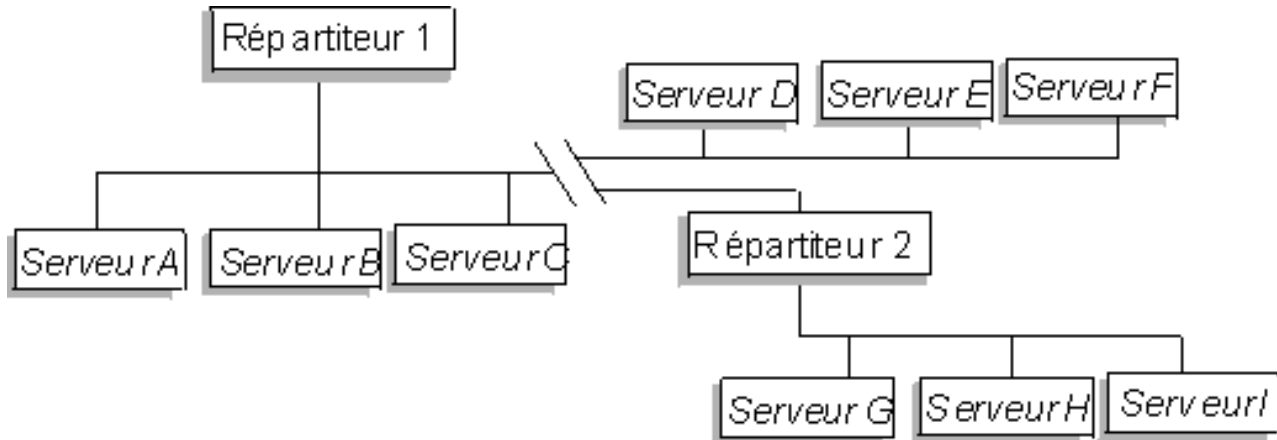


Figure 3. Exemple de site utilisant Dispatcher pour gérer des serveurs locaux et éloignés

La prise en charge de réseau étendu de Dispatcher permet d'utiliser à la fois des serveurs locaux et éloignés (c'est-à-dire des serveurs résidant sur différents sous-réseaux). La figure 3 présente une configuration dans laquelle un Dispatcher "local" (Dispatcher 1) sert de point d'entrée pour l'ensemble des demandes. Il distribue ces demandes entre ses propres serveurs locaux (Serveur A, Serveur B, Serveur C) et au serveur Dispatcher éloigné (Dispatcher 2), qui équilibre la charge sur ses serveurs locaux (Serveur G, Serveur H, Serveur I).

Lors de l'utilisation de la méthode d'acheminement NAT de Dispatcher ou du support GRE, le support de réseau étendu avec Dispatcher peut aussi être assuré sans utiliser de serveur Dispatcher sur le site éloigné (où se trouvent les serveurs D, E et F). Pour plus d'informations, voir «Réacheminement NAT/NAPT de Dispatcher (méthode d'acheminement nat)», à la page 53 et «Support GRE (Generic Routing Encapsulation)», à la page 234.

Présentation générale du composant CBR (Content Based Routing)

CBR coopère avec Caching Proxy pour relayer les demandes des clients aux serveurs HTTP ou HTTPS (SSL) indiqués. Il permet de manipuler les détails de la mémoire cache pour accélérer le rappel des documents Web avec une petite largeur de bande. CBR et Caching Proxy examinent les requêtes HTTP à l'aide des types de règle indiqués.

Remarque : Le composant Content Based Routing (CBR) est disponible sur toutes les plateformes prises en charge à l'exception de celles exécutant une JVM 64 bits. Vous avez également la possibilité d'employer la méthode d'acheminement cbr du composant Dispatcher de Load Balancer pour permettre le routage CBR sans faire appel à Caching Proxy. Pour plus de détails, voir «Fonction CBR de Dispatcher (méthode d'acheminement cbr)», à la page 55.

CBR permet de spécifier un ensemble de serveurs qui prend en charge une demande en fonction de son contenu. CBR vous permet d'indiquer plusieurs serveurs pour chaque type de requête. Par conséquent, les requêtes peuvent être équilibrées pour obtenir une réponse optimale du client. CBR peut aussi détecter les incidents qui se produisent sur un serveur et arrêter d'acheminer des demandes

vers ce dernier. L'algorithme d'équilibrage de charge utilisé par le composant CBR est identique à l'algorithme éprouvé utilisé par le composant Dispatcher.

Lorsqu'une demande est reçue par Caching Proxy, elle est comparée aux règles qui ont été définies dans le composant CBR. En cas de correspondance, l'un des serveurs associés à cette règle est désigné pour prendre en charge la demande. Caching Proxy continue alors son traitement normal pour acheminer la demande vers le serveur sélectionné.

CBR offre les mêmes fonctions que Dispatcher à l'exception des fonctions de haute disponibilité, de sous-agent SNMP, de réseau étendu et de quelques commandes de configuration.

Caching Proxy doit être en fonction pour permettre à CBR d'équilibrer la charge des demandes des clients.

Gestion des serveurs locaux avec CBR

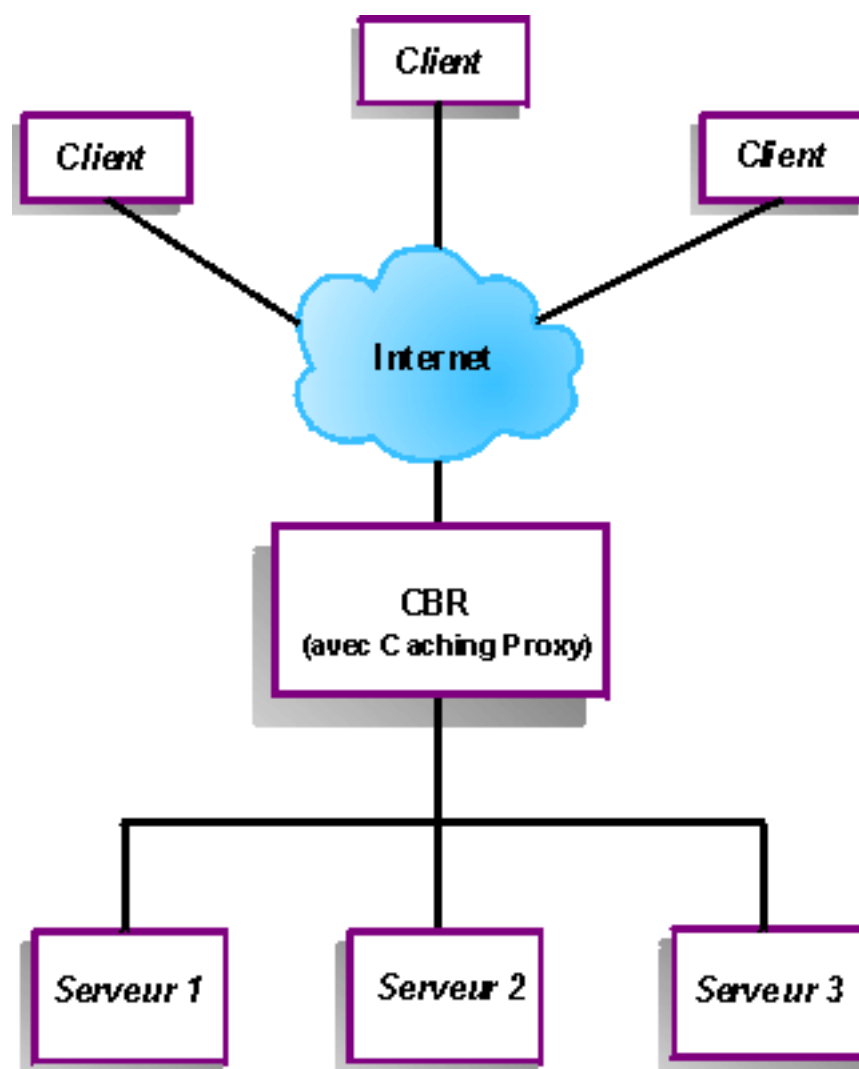


Figure 4. Exemple de site utilisant CBR pour gérer les serveurs locaux

La figure 4, à la page 13 montre la représentation logique d'un site utilisant CBR pour acheminer des demandes issues des serveurs locaux. Le composant CBR utilise Caching Proxy pour acheminer les demandes des clients (HTTP ou HTTPS) aux serveurs en fonction du contenu de l'URL.

Présentation générale du composant Site Selector

Site Selector fonctionne avec d'autres serveurs de noms pour équilibrer la charge sur un groupe de serveurs à l'aide des mesures et des pondérations recueillies. Vous pouvez créer une configuration de site pour assurer l'équilibrage de charge sur un groupe de serveurs sur la base du nom de domaine utilisé pour la demande d'un client.

Un client envoie une demande de résolution de nom de domaine à un serveur de noms appartenant au réseau. Le serveur de noms achemine la demande au poste Site Selector. Site Selector résout le nom de domaine en adresse IP de l'un des serveurs qui a été configuré sous le nom du site. Site Selector renvoie l'adresse IP du serveur sélectionné au serveur de noms. Le serveur de noms renvoie l'adresse IP au client.

Metric Server est un composant de Load Balancer qui surveille le système et doit être installé sur chaque serveur avec équilibrage de charge de votre configuration. Metric Server permet à Site Selector de surveiller le niveau d'activité d'un serveur, de détecter le moment où un serveur est le moins chargé et de détecter un serveur défaillant. Par charge, on entend le travail effectivement fourni par le serveur. En personnalisant les fichiers script de mesure du système, vous pouvez choisir le type de mesure utilisé pour évaluer la charge. Site Selector peut être configuré en fonction de chaque environnement, en tenant compte de facteurs tels que la fréquence des accès, le nombre total d'utilisateurs et les différents types d'accès (requêtes courtes, longues, à forte ou faible consommation de ressources de l'unité centrale).

Gestion des serveurs locaux et éloignés avec Site Selector et Metric Server

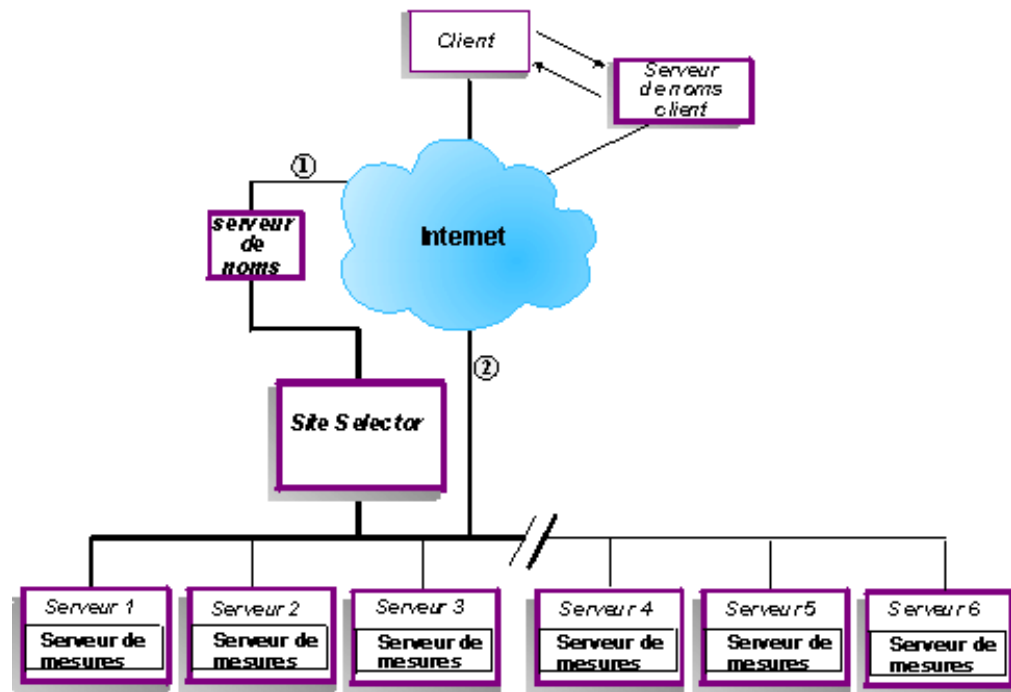


Figure 5. Exemple de site utilisant Site Selector et Metric Server pour gérer les serveurs locaux et éloignés

La figure 5 illustre un site utilisant le composant Site Selector pour répondre aux demandes. Serveur 1, Serveur 2 et Serveur 3 sont des serveurs locaux. Serveur 4, Serveur 5 et Serveur 6 sont des serveurs éloignés.

Un client envoie une demande de résolution de nom de domaine à un serveur de noms client. Le serveur de noms client achemine la demande au poste Site Selector (chemin d'accès 1) via DNS. Site Selector résout ensuite le nom de domaine en adresse IP de l'un des serveurs. Site Selector renvoie l'adresse IP du serveur sélectionné au serveur de noms client. Le serveur de noms renvoie l'adresse IP au client.

Une fois que le client a reçu l'adresse IP du serveur, il achemine les demandes d'application directement au serveur sélectionné (chemin d'accès 2).

Remarque : Dans cet exemple, le serveur Metric Server fournit les informations de charge du système au poste Site Selector. L'agent Metric Server est installé sur chaque serveur principal. Utilisez Metric Server conjointement à Site Selector, sinon Site Selector peut seulement utiliser une méthode de sélection par permutation circulaire pour l'équilibrage de charge.

Présentation générale du composant Cisco CSS Controller

Cisco CSS Controller constitue une solution complémentaire avec les commutateurs Cisco CSS série 11000. La solution combinée réunit de robustes fonctions d'acheminement de paquets et de routage de contenu à des algorithmes de reconnaissance sophistiqués pour déterminer la disponibilité et les informations de

charge du *service* (base de données ou application serveur principal). La fonction Cisco CSS Controller fait appel à l'algorithme de calcul de pondération, aux conseillers standard et personnalisés et à Metric Server pour déterminer les mesures, la santé et la charge du service. Cisco CSS Controller utilise ces informations pour générer les mesures de pondération du service, qu'il envoie au serveur Cisco CSS Switch pour la sélection du service optimal, l'optimisation de la charge et la tolérance aux pannes.

Cisco CSS Controller suit de nombreux critères, dont :

- les connexions actives et le débit de connexion (nombre de nouvelles connexions au cours d'un cycle de calcul de pondération),
- la disponibilité des applications et des bases de données, qui est facilitée par l'utilisation de conseillers standard et personnalisés et les agents résidant sur le service, adaptés à l'application spécifique,
- l'utilisation de la CPU,
- l'utilisation de la mémoire,
- les mesures du système personnalisables par l'utilisateur.

Lorsqu'un serveur Cisco CSS Switch, sans Cisco CSS Controller, détermine la santé d'un service fournisseur de contenu, il utilise les temps de réponse aux demandes de contenu ou d'autres mesures de réseau. Avec Cisco CSS Controller, le serveur Cisco CSS Switch se décharge de ces activités sur le serveur Cisco CSS Controller. Cisco CSS Controller influence la pondération du service ou sa faculté à servir le contenu, et active ou suspend un service, selon le cas, lorsque le service devient disponible ou indisponible.

Cisco CSS Controller:

- Utilise une interface SNMP publiée pour obtenir des informations de connexion à partir du serveur Cisco CSS Switch
- Utilise les données d'entrée du conseiller pour analyser la disponibilité et le temps de réponse du service
- Utilise les informations Metric Server pour analyser la charge du système
- Génère des pondérations pour chaque service de la configuration

Les pondérations définies s'appliquent à tous les services connectés sur un même port. Pour chaque port, les demandes sont réparties entre les services selon la pondération relative de chacun. Par exemple, si un service a une pondération (paramètre Weight) de 10 et un autre de 5, le premier recevra deux fois plus de demandes que le second. Ces pondérations sont fournies au serveur Cisco CSS Switch avec SNMP. Lorsque la valeur de pondération d'un service est augmentée, le serveur Cisco CSS Switch dirige davantage de demandes vers ce service.

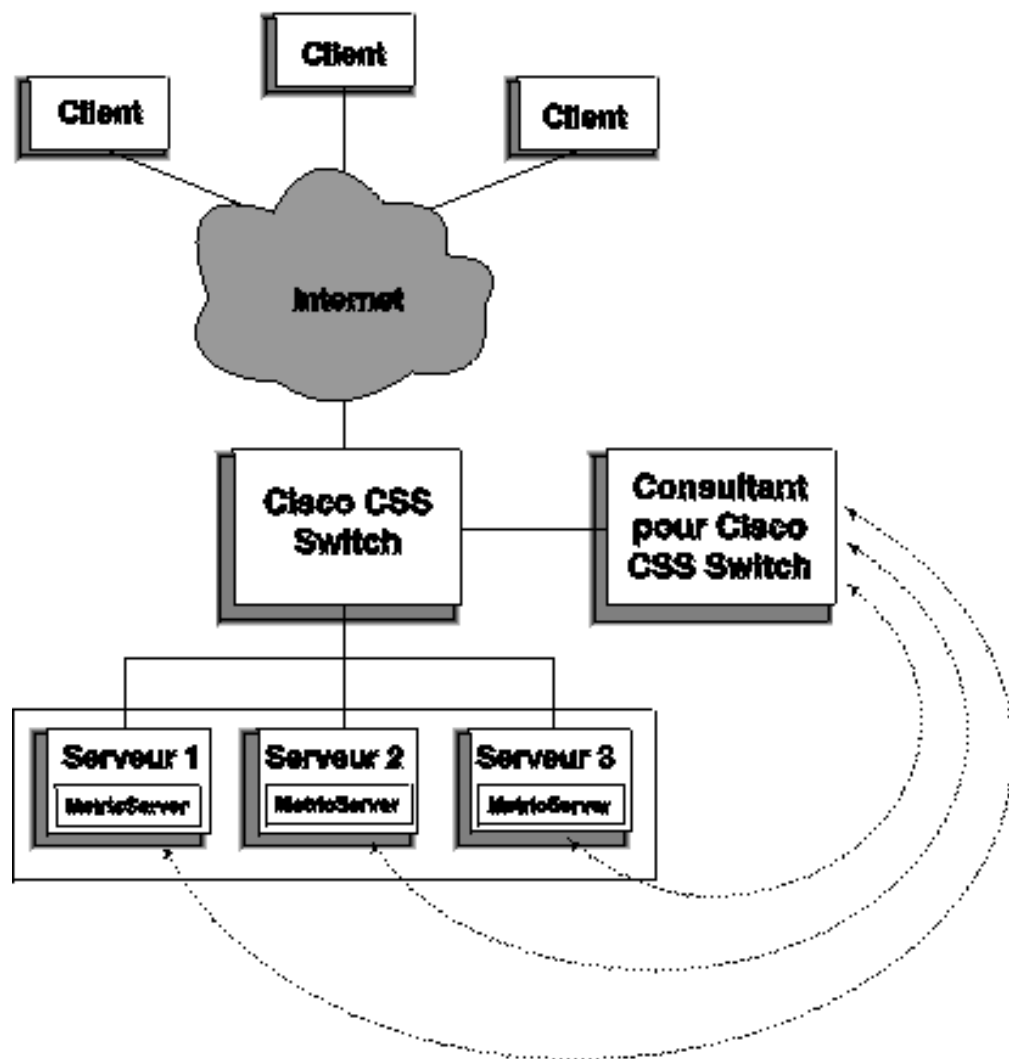


Figure 6. Exemple de site utilisant Cisco CSS Controller et Metric Server pour gérer les services locaux

Cisco CSS Controller, conjointement au serveur Cisco CSS Switch, constitue une solution idéale qui combine la commutation de contenu à haute vitesse avec la reconnaissance d'applications sophistiquées, la tolérance aux pannes et l'optimisation de la charge des services. Cisco CSS Controller appartient à une solution globale complémentaire située entre le serveur Cisco CSS Switch et la fonction Load Balancer d'IBM WebSphere Application Server .

Présentation générale du composant Nortel Alteon Controller

Nortel Alteon Controller, conjointement à la famille Nortel Alteon de commutateurs Web, constitue une solution complémentaire qui combine capacité et vitesse de transmission des paquets des commutateurs avec la reconnaissance des algorithmes sophistiqués de Load Balancer pour déterminer la pondération de charge des serveurs.

Nortel Alteon Controller permet de développer des conseillers personnalisés capables d'évaluations avec reconnaissance des applications plus intelligentes de la disponibilité et de la charge des applications utilisées pour déployer des services.

Le composant Metric Server fournit des informations relatives à la charge du système, telles que l'utilisation de la mémoire et de la CPU, ainsi qu'une structure vous permettant de développer des dispositifs personnalisés de mesure de la charge système.

Nortel Alteon Controller collecte de nombreux types de mesures pour déterminer les pondérations des serveurs dont la charge est équilibrée par des commutateurs Nortel Alteon Web Switch, notamment :

- les connexions actives et nouvelles,
- la disponibilité des applications et des bases de données, qui est facilitée par l'utilisation de conseillers standard et personnalisés et les agents résidant sur le serveur, adaptés à l'application spécifique,
- l'utilisation de la CPU,
- l'utilisation de la mémoire,
- les mesures du serveur personnalisables par l'utilisateur,
- accessibilité.

Nortel Alteon Controller utilise SNMP pour communiquer avec le commutateur. Les informations de configuration, d'état et de connexion sont extraites du commutateur. Lorsque le contrôleur a calculé les pondérations du serveur, celles-ci sont définies sur le commutateur. Le commutateur se sert des pondérations définies par le contrôleur pour sélectionner le serveur le mieux à même de traiter les demandes de service des clients.

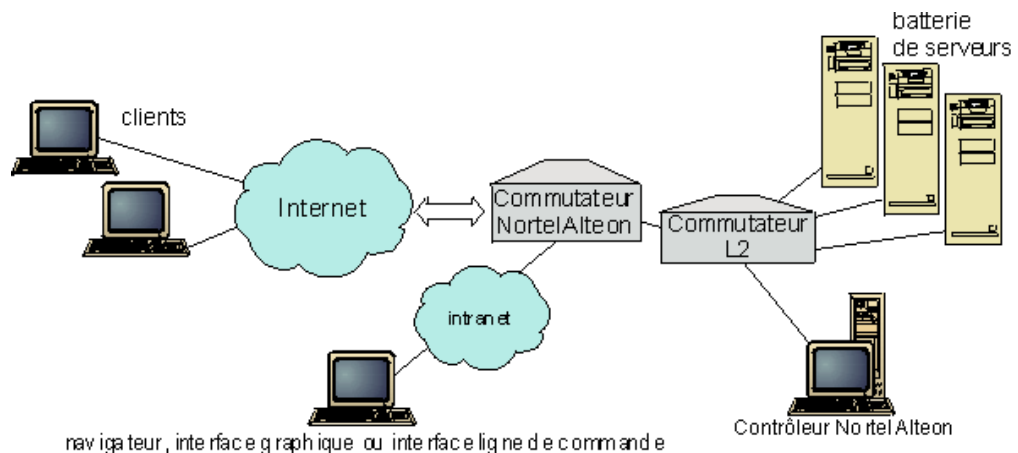


Figure 7. Exemple de site utilisant Nortel Alteon Controller pour gérer les serveurs locaux

La gestion du contrôleur peut s'effectuer à l'aide d'un navigateur, d'une interface graphique éloignée ou d'une interface de ligne de commande.

Nortel Alteon Controller, associé à la famille Nortel Alteon des commutateurs Web, constitue une solution idéale qui combine la commutation de paquets avec la reconnaissance d'applications sophistiquées, la tolérance aux pannes et l'optimisation de la charge des serveurs. Nortel Alteon Controller fait partie d'une solution complémentaire entre la famille Nortel Alteon des commutateurs et WebSphere d'IBM.

Chapitre 3. Gestion du réseau : Fonctions Load Balancer requises

Ce chapitre présente toutes les fonctions de configuration des composants de Load Balancer de sorte que vous pouvez déterminer celles qui sont les mieux adaptées à la gestion de votre réseau :

- «Gestionnaire, conseillers et système Metric Server (des composants Dispatcher, CBR et Site Selector)»
- «Fonctions du composant Dispatcher»
- «Fonctions du composant CBR (Content Based Routing)», à la page 23
- «Fonctions du composant Site Selector», à la page 25
- «Fonctions du composant Cisco CSS Controller», à la page 27
- «Fonctions du composant Nortel Alteon Controller», à la page 28

IMPORTANT : Si vous utilisez Load Balancer pour IPv4 et IPv6, seul le composant Dispatcher est disponible. Pour plus d'informations, voir Chapitre 8, «Déploiement de Dispatcher sur Load Balancer pour IPv4 et IPv6», à la page 81.

Gestionnaire, conseillers et système Metric Server (des composants Dispatcher, CBR et Site Selector)

Pour optimiser l'équilibrage de la charge sur les serveurs et garantir le choix du serveur approprié, voir :

- «Optimisation de la fonction d'équilibrage de charge fournie par Load Balancer», à la page 180
- «Conseillers», à la page 185
- «Metric Server», à la page 196

Fonctions du composant Dispatcher

Dispatcher assure l'équilibrage de charge sur vos serveurs pour les protocoles HTTP, FTP, SSL, SMTP, NNTP, IMAP, POP3, Telnet, SIP, et toute application de type TCP ou UDP sans état.

Administration à distance

- Pour configurer Load Balancer à partir d'un autre poste que celui où réside ce composant, voir «Administration à distance de Load Balancer», à la page 261.
(Si vous utilisez l'installation Load Balancer pour IPv4 et IPv6, cette fonction n'est pas disponible.)

Co-implantation

- Pour exécuter Dispatcher sur le même poste qu'un serveur Web dont vous équilibrez la charge, voir «Utilisation de serveurs implantés au même endroit», à la page 202.

Haute disponibilité

- Pour supprimer de votre réseau les restrictions liées au principe de point de défaillance unique à l'aide de Dispatcher, voir «Haute disponibilité simple», à la page 60 et «Haute disponibilité réciproque», à la page 61.
(Si vous utilisez l'installation Load Balancer pour IPv4 et IPv6, la fonction de haute disponibilité simple est disponible, mais pas la haute disponibilité réciproque.)

Affinité client à serveur

Lors de l'équilibrage de la charge du trafic SSL (HTTPS) :

- Pour vous assurer que le client utilise le même serveur SSL pour plusieurs connexions, voir «Fonctionnement de la fonction d'affinité pour Load Balancer», à la page 221.
- Pour vous assurer que le client utilise le même serveur pour le trafic HTTP et SSL, voir «Affinité de ports croisés», à la page 222.
(Si vous utilisez l'installation Load Balancer pour IPv4 et IPv6, la fonction d'affinité de ports croisés n'est pas disponible.)
- Pour vous assurer que le client utilise le même serveur pour plusieurs connexions, voir «Fonctionnement de la fonction d'affinité pour Load Balancer», à la page 221.
- Pour vous assurer qu'un groupe de clients utilise le même serveur pour plusieurs connexions, voir «Masque d'adresse de l'affinité (masque de maintien de routage)», à la page 222.
(Si vous utilisez l'installation Load Balancer pour IPv4 et IPv6, la fonction stickymask n'est pas disponible.)
- Pour supprimer un serveur de votre configuration (pour des besoins de maintenance, par exemple) sans interrompre le trafic client, voir «Mise au repos de la gestion des connexions serveur», à la page 223.

Equilibrage de charge basé sur des règles

Pour diriger les clients vers différents groupes de serveurs pour la même adresse Web, vous pouvez ajouter des "règles" à la configuration de Dispatcher. Pour plus d'informations, voir «Configuration de l'équilibrage de charge basé sur des règles», à la page 212.

- Pour diriger les clients vers différents groupes de serveurs en fonction de l'adresse IP source du client, voir «Utilisation de règles basées sur l'adresse IP des clients», à la page 213.
- Pour diriger les clients vers différents groupes de serveurs en fonction du port du client, voir «Utilisation de règles basées sur le port du client», à la page 214.
- Pour diriger les clients vers différents groupes de serveurs en fonction de l'heure, voir «Utilisation de règles basées sur l'heure», à la page 214.
- Pour diriger les clients des serveurs en fonction des bits TOS (Type Of Service, type de service) des paquets réseau, voir «Utilisation de règles basées sur le type de services (TOS)», à la page 214.
- Pour diriger les clients vers différents groupes de serveurs en fonction du trafic sur le site :
 - à l'aide du nombre de connexions par seconde, voir «Utilisation de règles basées sur le nombre de connexions par seconde», à la page 215,
 - à l'aide du nombre total de connexions actives, voir «Utilisation de règles basées sur le nombre total de connexions actives», à la page 215,

- par réservation et partage de la largeur de bande entre différentes adresses Web, voir «Utilisation de règles basées sur la largeur de bande réservée et sur la largeur de bande partagée», à la page 216,
- en vous assurant que le trafic est correctement évalué pour chaque ensemble de serveurs, voir «Option d'évaluation de serveur», à la page 220.
- Pour diriger le trafic excédentaire vers un ensemble de serveurs par défaut (par exemple, des serveurs qui répondront "site occupé"), voir «Utilisation de règles toujours vraies», à la page 218.
- Pour remplacer l'affinité client afin de garantir qu'un client ne reste pas "maintenu" à un serveur de débordement, voir «Substitution d'affinité de port», à la page 219.

Si vous utilisez l'installation Load Balancer pour IPv4 et IPv6, l'équilibrage de charge basé sur des règles n'est pas disponible.

Routage par contenu à l'aide de la méthode d'acheminement CBR de Dispatcher

Pour vous assurer que les clients SSL reviennent au même serveur, sur la base de l'ID SSL de la demande client

- Reportez-vous à la page 55.

Pour acheminer les clients HTTP vers différents groupes de serveurs à l'aide de règles basées sur la correspondance avec le contenu de l'URL de la demande client, voir «Fonction CBR de Dispatcher (méthode d'acheminement cbr)», à la page 55 et «Utilisation de règles basées sur le contenu des demandes», à la page 219.

- Pour différencier deux URL et leurs applications de service, voir «Partitionnement du serveur : serveurs logiques configurés pour un serveur physique (adresse IP)», à la page 58.
- Pour vous assurer que les clients reviennent au même serveur lors de la demande d'un contenu similaire dans plusieurs connexions à l'aide de cookies créés par vos serveurs Web, voir «Affinité de cookie passif», à la page 226.
- Pour équilibrer le trafic Web sur des serveurs relais avec mémoire cache permettant de placer un contenu unique en cache sur chaque serveur (augmentant ainsi la taille de la mémoire cache du site en éliminant les éléments superflus placés en mémoire cache sur plusieurs machines), voir «Affinité d'URI», à la page 227.

(Si vous utilisez l'installation Load Balancer pour IPv4 et IPv6, la méthode de transfert cbr de Dispatcher n'est pas disponible.)

Comparaison entre la méthode d'acheminement CBR de Dispatcher et le composant CBR

La méthode d'acheminement CBR de Dispatcher présente l'avantage de répondre plus rapidement aux requêtes client que le composant CBR. De plus, elle ne requiert *pas* l'installation et l'emploi du module Caching Proxy.

Si votre réseau inclut du trafic SSL (client via serveur) totalement sécurisé, l'utilisation du composant CBR (conjointement au module Caching Proxy) présente l'avantage de traiter le chiffrement et le déchiffrement requis pour effectuer un routage par contenu. Pour des connexions totalement sécurisées, la méthode d'acheminement CBR de Dispatcher ne peut être configurée qu'avec affinité d'ID

SSL car elle ne peut pas traiter le chiffrement et le déchiffrement pour effectuer un réel routage par contenu sur l'URL de la requête client.

Équilibrage de charge d'un réseau étendu

L'équilibrage de charge d'un réseau étendu peut être obtenu à l'aide de plusieurs méthodes distinctes.

- Pour équilibrer la charge sur des serveurs éloignés à l'aide de la fonction de réseau étendu de Dispatcher, voir «Configuration du support de réseau étendu pour Dispatcher», à la page 228 et «Support GRE (Generic Routing Encapsulation)», à la page 234.

Remarque : Si GRE n'est pas pris en charge sur le site éloigné, vous devez installer un module Dispatcher supplémentaire sur ce site.

- Pour équilibrer la charge sur des serveurs éloignés à l'aide de la méthode d'acheminement NAT de Dispatcher, voir «Réacheminement NAT/NAPT de Dispatcher (méthode d'acheminement nat)», à la page 53.

Remarque : Avec cette méthode, *aucun* module Dispatcher supplémentaire n'est requis sur le site éloigné.

(Si vous utilisez l'installation Load Balancer pour IPv4 et IPv6, la fonction d'équilibrage de charge d'un réseau étendu n'est pas disponible.)

Mappage de port

- Pour équilibrer la charge d'une adresse Web sur plusieurs démons de serveur d'une même machine, écoutant chacun un port différent, voir «Réacheminement NAT/NAPT de Dispatcher (méthode d'acheminement nat)», à la page 53.

(Si vous utilisez l'installation Load Balancer pour IPv4 et IPv6, cette fonction n'est pas disponible.)

Configuration de Dispatcher sur un réseau privé

- Pour placer le trafic de Dispatcher sur un autre réseau que celui du trafic client (afin d'améliorer les performances en réduisant les conflits sur le réseau externe), voir «Utilisation d'une configuration réseau privée», à la page 235.

Cluster générique et port générique

- Pour combiner plusieurs adresses Web en une seule configuration, voir «Utilisation d'un cluster générique pour combiner les configurations serveurs», à la page 236.
- Pour équilibrer la charge de pare-feu, voir «Utilisation du cluster générique pour équilibrer la charge des pare-feux», à la page 237.
- Pour acheminer le trafic de tous les ports de destination, voir «Utilisation du port générique pour acheminer le trafic destiné à un port non configuré», à la page 238.

Détection d'attaque de "refus de service"

- Pour détecter les éventuelles attaques de "refus de service", voir «Détection d'attaque de refus de service», à la page 239.

Consignation binaire

- Pour analyser le trafic du serveur, voir «Utilisation de la consignation binaire pour analyser les statistiques des serveurs», à la page 240.

Alertes

- Pour générer des alertes lorsque des serveurs sont marqués comme actifs ou inactifs, voir «Utilisation de scripts pour la génération d'une alerte ou d'une erreur du serveur d'enregistrement», à la page 184.

Fonctions du composant CBR (Content Based Routing)

CBR intègre l'équilibrage de charge au module Caching Proxy de WebSphere Application Server pour relayer les demandes des clients aux serveurs HTTP ou HTTPS (SSL) indiqués. Pour pouvoir utiliser CBR, vous devez installer et configurer le module Caching Proxy sur le même poste. Pour plus d'informations sur la configuration de Caching Proxy en vue d'utiliser CBR, voir «Etape 1. Configuration de Caching Proxy pour utiliser CBR», à la page 114.

Remarque : Le composant Content Based Routing (CBR) est disponible sur toutes les plateformes prises en charge à l'exception de celles exécutant une JVM 64 bits. Vous avez également la possibilité d'employer la méthode d'acheminement cbr du composant Dispatcher de Load Balancer pour permettre le routage CBR sans faire appel à Caching Proxy. Pour plus de détails, voir «Fonction CBR de Dispatcher (méthode d'acheminement cbr)», à la page 55.

Le composant CBR (ou la méthode d'acheminement CBR du composant Dispatcher), procure à vos clients les avantages suivants :

- Equilibrage de la charge des requêtes client pour différents types de contenus sur des groupes de serveurs. (Voir la section «Equilibrage de la charge des requêtes pour différents types de contenus», à la page 106.)
- Amélioration du temps de réponse par optimisation de la répartition du contenu de votre site entre vos serveurs Web. (Voir la section «Division du contenu de votre site pour améliorer le temps de réponse», à la page 106.)
- Préservation du trafic client en cas de défaillance du serveur par affectation de plusieurs serveurs à chaque partie de votre site. (Voir la section «Copie de sauvegarde du contenu du serveur Web», à la page 107.)

Comparaison entre le composant CBR et la méthode d'acheminement CBR de Dispatcher

Si votre réseau nécessite le trafic SSL (client via serveur) totalement sécurisé, l'utilisation du composant CBR (conjointement au module Caching Proxy) présente l'avantage de traiter le chiffrement/déchiffrement SSL requis pour effectuer un routage par contenu.

Pour des connexions SSL totalement sécurisées, la méthode d'acheminement CBR de Dispatcher ne peut être configurée qu'avec affinité d'ID SSL car elle ne peut pas traiter le chiffrement/déchiffrement pour effectuer un réel routage par contenu sur l'URL de la requête client.

Pour le trafic HTTP, l'utilisation de la méthode d'acheminement CBR de Dispatcher présente l'avantage de répondre plus rapidement aux requêtes client que le composant CBR. De plus, elle ne requiert *pas* l'installation et l'emploi du module Caching Proxy.

Administration à distance

- Pour configurer Load Balancer à partir d'un autre poste que celui où réside ce composant, voir «Administration à distance de Load Balancer», à la page 261.

Co-implantation

- CBR peut s'exécuter sur le même poste qu'un serveur dont vous équilibrez la charge. Pour plus d'informations, voir «Utilisation de serveurs implantés au même endroit», à la page 202.

CBR et plusieurs instances de Caching Proxy

- Pour optimiser l'utilisation de la CPU en exécutant plusieurs processus Caching Proxy, voir «Utilisation de plusieurs processus Caching Proxy pour optimiser l'utilisation de la CPU», à la page 107.

Routage par contenu pour les connexions SSL

Pour autoriser le routage par contenu du trafic SSL :

- Par le biais d'une connexion sécurisée à chaque extrémité (client à proxy et proxy à serveur), voir «Équilibrage de charge sur les connexions sécurisées (SSL)», à la page 107.
- Par le biais de connexions sécurisées côté client à proxy uniquement, voir «Équilibrage de charge client-proxy dans SSL et proxy-serveur dans HTTP», à la page 108.

Partitionnement du serveur

- Pour différencier deux URL et leurs applications de service, voir «Partitionnement du serveur : serveurs logiques configurés pour un serveur physique (adresse IP)», à la page 58.

Équilibrage de charge basé sur des règles

Pour diriger les clients vers différents groupes de serveurs pour la même adresse Web, vous pouvez ajouter des "règles" à la configuration de CBR. Pour plus d'informations, voir «Configuration de l'équilibrage de charge basé sur des règles», à la page 212.

- Pour diriger les clients vers différents groupes de serveurs en fonction du contenu de l'URL demandée, voir «Utilisation de règles basées sur le contenu des demandes», à la page 219.
- Pour diriger les clients vers différents groupes de serveurs en fonction de l'adresse IP source du client, voir «Utilisation de règles basées sur l'adresse IP des clients», à la page 213.
- Pour diriger les clients vers différents groupes de serveurs en fonction de l'heure, voir «Utilisation de règles basées sur l'heure», à la page 214.
- Pour diriger les clients vers différents groupes de serveurs en fonction du trafic sur le site :
 - à l'aide du nombre de connexions par seconde, voir «Utilisation de règles basées sur le nombre de connexions par seconde», à la page 215,

- à l'aide du nombre total de connexions actives, voir «Utilisation de règles basées sur le nombre total de connexions actives», à la page 215,
- Pour diriger le trafic excédentaire vers un ensemble de serveurs par défaut (par exemple, des serveurs qui répondront "site occupé"), voir «Utilisation de règles toujours vraies», à la page 218.
- Pour remplacer l'affinité client afin de garantir qu'un client ne reste pas "maintenu" à un serveur de débordement, voir «Substitution d'affinité de port», à la page 219.

Affinité client à serveur

- Pour vous assurer qu'un client revient au même serveur pour plusieurs connexions, voir «Fonctionnement de la fonction d'affinité pour Load Balancer», à la page 221.
- Pour supprimer un serveur de votre configuration (pour des besoins de maintenance, par exemple) sans interrompre le trafic client, voir «Mise au repos de la gestion des connexions serveur», à la page 223.
- Pour vous assurer que les clients reviennent au même serveur lors de la demande d'un contenu similaire dans plusieurs connexions sans se baser sur des cookies créés par vos serveurs Web, voir «Affinité de cookie actif», à la page 225.
- Pour vous assurer que les clients reviennent au même serveur lors de la demande d'un contenu similaire dans plusieurs connexions à l'aide de cookies créés par vos serveurs Web, voir «Affinité de cookie passif», à la page 226.
- Pour équilibrer le trafic Web sur des serveurs relais avec mémoire cache permettant de placer un contenu unique en cache sur chaque serveur (augmentant ainsi la taille de la mémoire cache du site en éliminant les éléments superflus placés en mémoire cache sur plusieurs machines), voir «Affinité d'URI», à la page 227.

Haute disponibilité à l'aide de Dispatcher et CBR

- Pour supprimer de votre réseau les restrictions liées au principe de point de défaillance unique à l'aide de Dispatcher utilisé dans une configuration de second niveau avec CBR, voir «Haute disponibilité fournie par Load Balancer», à la page 6.

Consignation binaire

- Pour analyser le trafic du serveur, voir «Utilisation de la consignation binaire pour analyser les statistiques des serveurs», à la page 240.

Alertes

- Pour générer des alertes lorsque des serveurs sont marqués comme actifs ou inactifs, voir «Utilisation de scripts pour la génération d'une alerte ou d'une erreur du serveur d'enregistrement», à la page 184.

Fonctions du composant Site Selector

Site Selector équilibre la charge d'une requête de service d'annuaire dans un groupe de serveurs

Administration à distance

- Pour configurer Load Balancer à partir d'un autre poste que celui où réside ce composant, voir «Administration à distance de Load Balancer», à la page 261.

Co-implantation

- Site Selector peut s'exécuter sur le même poste qu'un serveur dont vous équilibrez la charge sans configuration supplémentaire.

Haute disponibilité

- La haute disponibilité est inhérente aux méthodologies DNS (Domain Name System) qui utilisent plusieurs modules Site Selector redondants pour une parfaite configuration du serveur de noms parent et un positionnement approprié des méthodes de reprise DNS normales. La retransmission des demandes et le renouvellement des transferts de zone sont des exemples de méthodes de reprise DNS normales.
- Pour supprimer de votre réseau les restrictions liées au principe de point de défaillance unique à l'aide de Dispatcher utilisé dans une configuration de second niveau avec Site Selector, voir «Haute disponibilité fournie par Load Balancer», à la page 6.

Affinité client à serveur

- Pour vous assurer que le client utilise le même serveur pour plusieurs demandes de serveur de noms, voir «Fonctionnement de la fonction d'affinité pour Load Balancer», à la page 221.
- Pour garantir l'affinité client à serveur à l'aide de la méthode DNS standard de définition de la durée de vie (TTL, Time To Live), voir «Considérations relatives à la durée de vie (TTL)», à la page 129.

Equilibrage de charge basé sur des règles

Pour acheminer les requêtes des clients vers différents groupes de serveurs pour la résolution des noms de domaine, vous pouvez ajouter des "règles" à la configuration de Site Selector. Pour plus d'informations, voir «Configuration de l'équilibrage de charge basé sur des règles», à la page 212.

- Pour diriger les clients vers différents groupes de serveurs en fonction de l'adresse IP source du client, voir «Utilisation de règles basées sur l'adresse IP des clients», à la page 213.
- Pour diriger les clients vers différents groupes de serveurs en fonction de l'heure, voir «Utilisation de règles basées sur l'heure», à la page 214.
- Pour diriger les clients vers différents ensembles de serveurs en fonction des valeurs de charge de l'ensemble de serveurs, voir :
 - «Règle Mesure de tous les serveurs», à la page 217
 - «Règle Moyenne des mesures», à la page 218
- Pour diriger le trafic excédentaire vers un ensemble de serveurs par défaut (par exemple, des serveurs qui répondront "site occupé"), voir «Utilisation de règles toujours vraies», à la page 218.

Equilibrage de charge d'un réseau étendu

Site Selector peut s'exécuter dans un réseau local (LAN) comme dans un réseau étendu (WAN).

Dans un réseau étendu :

- Pour équilibrer la charge des demandes de serveur de noms des clients à l'aide d'une technique de permutation circulaire pondérée, aucune configuration supplémentaire n'est nécessaire.
- Pour évaluer la proximité au réseau du serveur de noms du client par rapport aux serveurs qui fournissent l'application requise (serveurs de destination), reportez-vous à la section «Utilisation de la fonction de proximité réseau (Network Proximity)», à la page 129.

Alertes

- Pour générer des alertes lorsque des serveurs sont marqués comme actifs ou inactifs, voir «Utilisation de scripts pour la génération d'une alerte ou d'une erreur du serveur d'enregistrement», à la page 184.

Fonctions du composant Cisco CSS Controller

Cisco CSS Controller améliore la fonction d'équilibrage des charges des serveurs Cisco Switch en prêtant plus d'attention aux applications et au système. Le contrôleur utilise des mesures plus appropriées aux applications et au système pour calculer dynamiquement les pondérations des serveurs. Les pondérations sont transmises au commutateur à l'aide de SNMP. Le commutateur utilise les pondérations lors du traitement des demandes des clients ce qui optimise la charge des serveurs et augmente la tolérance aux pannes.

Pour optimiser l'équilibrage de la charge sur les serveurs et garantir le choix du serveur approprié, voir :

- «Optimisation de la fonction d'équilibrage de charge fournie par Load Balancer», à la page 246
- «Conseillers», à la page 248 et «Création de conseillers personnalisés», à la page 250
- «Système Metric Server», à la page 253

Administration à distance

- Pour configurer Load Balancer à partir d'un autre poste que celui où réside ce composant, voir «Administration à distance de Load Balancer», à la page 261.

Co-implantation

- Cisco CSS Controller peut s'exécuter sur le même poste qu'un serveur dont vous équilibrez la charge sans configuration supplémentaire.

Haute disponibilité

- Pour supprimer de votre réseau les restrictions liées au principe de point de défaillance unique, le composant Cisco CSS Switch comme le composant Cisco CSS Controller intègrent des fonctions de haute disponibilité. Pour le commutateur, les fonctions de haute disponibilité sont accessibles via le protocole de redondance CSS. Pour le composant Cisco CSS Controller, un protocole propriétaire permet la configuration en secours automatique des deux contrôleurs.

Pour plus d'informations sur la fonction de haute disponibilité, voir «Haute disponibilité», à la page 146.

Consignation binaire

- Pour analyser le trafic du serveur, voir «Utilisation de la consignation binaire pour analyser les statistiques des serveurs», à la page 255.

Alertes

- Pour générer des alertes lorsque des serveurs sont marqués comme actifs ou inactifs, voir «Utilisation de scripts pour la génération d'une alerte ou d'une erreur du serveur d'enregistrement», à la page 257.

Fonctions du composant Nortel Alteon Controller

Nortel Alteon Controller améliore la fonction d'équilibrage des charges des serveurs Nortel Alteon Switch en prêtant plus d'attention aux applications et au système. Le contrôleur utilise des mesures plus appropriées aux applications et au système pour calculer dynamiquement les pondérations des serveurs. Les pondérations sont transmises au commutateur à l'aide de SNMP. Le commutateur utilise les pondérations lors du traitement des demandes des clients ce qui optimise la charge des serveurs et augmente la tolérance aux pannes.

Pour optimiser l'équilibrage de la charge sur les serveurs et garantir le choix du serveur approprié, voir :

- «Optimisation de la fonction d'équilibrage de charge fournie par Load Balancer», à la page 246
- «Conseillers», à la page 248 et «Création de conseillers personnalisés», à la page 250
- «Système Metric Server», à la page 253

Administration à distance

- Pour configurer Load Balancer à partir d'un autre poste que celui où réside ce composant, voir «Administration à distance de Load Balancer», à la page 261.

Co-implantation

- Nortel Alteon Controller peut s'exécuter sur le même poste qu'un serveur dont vous équilibrez la charge sans configuration supplémentaire.

Haute disponibilité

- Pour supprimer de votre réseau les restrictions liées au principe de point de défaillance unique, le commutateur Web comme le contrôleur Nortel Alteon intègrent des fonctions de haute disponibilité. Pour le commutateur, la haute disponibilité est accessible via le protocole de redondance pour les connexions aux serveurs et les services. Le contrôleur Nortel Alteon offre la haute disponibilité via un protocole propriétaire qui permet la configuration en secours automatique de deux contrôleurs.

Pour plus d'informations sur la fonction de haute disponibilité, voir «Haute disponibilité», à la page 166.

Consignation binaire

- Pour analyser le trafic du serveur, voir «Utilisation de la consignation binaire pour analyser les statistiques des serveurs», à la page 255.

Alertes

- Pour générer des alertes lorsque des serveurs sont marqués comme actifs ou inactifs, voir «Utilisation de scripts pour la génération d’une alerte ou d’une erreur du serveur d’enregistrement», à la page 257.

Chapitre 4. Installation de Load Balancer

Le présent chapitre décrit l'installation de Load Balancer à l'aide des outils de création de packages du système et la configuration requise pour tous les systèmes d'exploitation pris en charge.

- «Configuration requise et installation pour AIX»
- «Configuration requise et installation pour HP-UX», à la page 35
- «Configuration requise et installation pour Linux», à la page 36
- «Configuration requise et installation pour Solaris», à la page 38
- «Configuration requise et installation pour Windows», à la page 40

Pour les instructions d'installation à l'aide du programme d'installation du produit, reportez-vous au document intitulé *Concepts, planification et installation pour Edge Components*.

Le SDK Java 2 est automatiquement installé avec Load Balancer sur toutes les plateformes.

Si vous migrez à partir d'une version antérieure de Load Balancer ou réinstallez un système d'exploitation, avant d'effectuer l'installation, vous pouvez sauvegarder des fichiers de configuration ou des fichiers script antérieurs de Load Balancer.

- Après l'installation, placez les fichiers de configuration dans le répertoire `...ibm/edge/lb/servers/configurations/composant` (où *composant* correspond à dispatcher, cbr, ss, cco ou nal).
- Après l'installation, placez les fichiers script (tels que goIdle et goStandby) dans le répertoire `.../ibm/edge/lb/servers/bin` afin de pouvoir les exécuter.

Suivant le type d'installation, les packages du composant Load Balancer qui sont répertoriés dans cette section ne sont pas tous fournis.

- Pour les installations Edge Component pouvant fournir Load Balancer et Caching Proxy, tous les packages du composant d'installation de Load Balancer sont disponibles.
- Pour les installations Edge Component pouvant fournir Load Balancer mais pas Caching Proxy, les packages du composant CBR ne sont pas inclus avec Load Balancer.
- Pour les installations Edge Component pour IPv6 (Load Balancer pour IPv4 et IPv6), le package du composant Dispatcher est inclus avec Load Balancer. Les packages du composant CBR, Site Selector et Controller ne sont pas inclus. Pour l'ordre d'installation recommandé des packages Load Balancer pour IPv4 et IPv6, voir «Installation de Load Balancer pour IPv4 et IPv6», à la page 83.

Configuration requise et installation pour AIX

Configuration requise pour les systèmes AIX

Pour plus d'informations sur les conditions matérielles et logicielles requises, y compris les navigateurs pris en charge, accédez à la page Web suivante : <http://www.ibm.com/support/docview.wss?rs=180&uid=swg27006921>

Installation pour les systèmes AIX

Le tableau 1 répertorie les images installp de Load Balancer et l'ordre d'installation recommandé à l'aide de l'outil d'installation de packages système.

Tableau 1. Images installp d'AIX

Produit de base	ibmlb.base.rte
Administration (avec messages)	<ul style="list-style-type: none">ibmlb.admin.rteibmlb.msg.<i>langue</i>.admin
Pilote de périphérique	ibmlb.lb.driver
Licence	ibmlb.lb.license
Composants Load Balancer (avec messages)	<ul style="list-style-type: none">ibmlb.<i>composant</i>.rteibmlb.msg.<i>langue</i>.lb
Documentation (avec messages)	<ul style="list-style-type: none">ibmlb.doc.rteibmlb.msg.en_US.doc
Système Metric Server	ibmlb.ms.rte

Où *composant* correspond à disp (Dispatcher), cbr (CBR), ss (Site Selector), cco (Cisco CSS Controller) ou nal (Nortel Alteon Controller). Vous pouvez éventuellement sélectionner le ou les composant(s) à installer.

Où *langue* peut prendre les valeurs suivantes :

- en_US
- de_CH
- de_DE
- es_ES
- fr_CA
- fr_CH
- fr_FR
- it_CH
- it_IT
- ja_JP
- Ja_JP
- ko_KR
- pt_BR
- zh_CN
- ZH_CN
- zh_TW
- Zh_TW

Le package de la documentation contient uniquement de l'anglais. Les traductions de l'ensemble des documentations Load Balancer se trouvent sur le site Web : www.ibm.com/software/webserver/appserv/ecinfocenter.html.

Avant de commencer l'installation

Si une version antérieure est déjà installée sur votre système, désinstallez-la avant d'installer la version actuelle. Assurez-vous, tout d'abord, que tous les exécuteurs et serveurs sont arrêtés. Puis, pour désinstaller l'intégralité du produit, entrez

installp -u ibmlb (ou l'ancien nom, par exemple, **intnd**). Pour désinstaller certains jeux de fichiers, spécifiez-les au lieu d'entrer le nom du module.

Lorsque vous installez le produit, le choix vous est offert d'installer tout ou partie des composants suivants :

- Produit de base
- Administration (avec messages)
- Pilote de périphérique (requis)
- Licence (requis)
- Composant Dispatcher (avec messages)
- Composant CBR (avec messages)
- Composant Site Selector (avec messages)
- Composant Cisco CSS Controller (avec messages)
- Composant Nortel Alteon Controller (avec messages)
- Documentation (avec messages)
- Système Metric Server

Etapas de la procédure d'installation

Effectuez les opérations ci-dessous pour installer Load Balancer pour les systèmes AIX :

1. Connectez-vous en tant que superutilisateur.
2. Insérez le support fourni dans son lecteur ou, si vous l'installez à partir du Web, copiez les images des disquettes d'installation dans un répertoire.
3. Installez l'image d'installation. Utilisez SMIT pour installer Load Balancer pour AIX. Ainsi, tous les messages seront installés automatiquement.

A l'aide de **SMIT** :

Sélectionnez

Installation et maintenance de logiciels

Sélectionnez

Installation et mise à jour de logiciels

Sélectionnez

Installation et mise à jour de tous les logiciels disponibles

Entrez l'unité ou le répertoire contenant les images installp

Entrez sur la ligne *PROGICIEL à installer, les données correspondant aux options choisies (ou sélectionnez Liste)

Cliquez sur

OK

Après l'exécution de la commande, appuyez sur **Fin**, puis sélectionnez l'option permettant de **quitter Smit** à partir du menu de sortie ou appuyez sur **F12**. Si vous utilisez SMITTY, appuyez sur **F10** pour quitter le programme.

A partir de la ligne de commande :

Si vous effectuez l'installation à partir d'un CD-ROM, entrez les commandes suivantes pour le monter :

```
mkdir /cdrom  
mount -v cdrfs -p -r /dev/cd0 /cdrom
```

Pour savoir quelle(s) commande(s) utiliser pour installer les packages Load Balancer choisis pour les systèmes AIX, reportez-vous au tableau suivant :

Tableau 2. Commandes d'installation AIX

Base	installp -acXgd <i>périphérique</i> ibmlb.base.rte
Administration (avec messages)	installp -acXgd <i>périphérique</i> ibmlb.admin.rte ibmlb.msg. <i>langue</i> .admin
Pilote de périphérique	installp -acXgd <i>périphérique</i> ibmlb.lb.driver
Licence	installp -acXgd <i>périphérique</i> ibmlb.lb.license
Composants Load Balancer (avec msgs). Inclut : Dispatcher, CBR, Site Selector, Cisco CSS Controller et Nortel Alteon Controller	installp -acXgd <i>périphérique</i> ibmlb.composant.rte ibmlb.msg. <i>langue</i> .lb
Documentation (avec messages)	installp -acXgd <i>périphérique</i> ibmlb.doc.rte ibmlb.msg.en_US.lb
Système Metric Server	installp -acXgd <i>périphérique</i> ibmlb.ms.rte

La variable *unité* prend les valeurs suivantes :

- *"/cdrom"* si l'on installe à partir d'un lecteur de CD-ROM.
- *"/dir"* (le répertoire contenant les images d'installation) si l'on installe à partir d'un système de fichiers.

Assurez-vous que la colonne de résultat du résumé d'opération contient la mention "SUCCESS" (réussite) pour chaque composant de Load Balancer installé (état APPLY). Ne poursuivez pas avant d'avoir réussi à installer tous les composants choisis.

Remarque : Pour générer la liste des jeux de fichiers d'une image installp avec l'ensemble des catalogues de messages disponibles, entrez
installp -ld *unité*

La variable *unité* prend les valeurs suivantes :

- *"/cdrom"* si l'on installe à partir d'un lecteur de CD-ROM.
- *"/dir"* (le répertoire contenant les images d'installation) si l'on installe à partir d'un système de fichiers.

Pour démonter le CD-ROM, tapez la commande suivante :

```
umount /cdrom
```

4. Vérifiez que le produit est installé correctement. Entrez la commande suivante :

```
ls1pp -h | grep ibmlb
```

Si vous avez installé l'intégralité du produit, cette commande génère le résultat suivant :

```
ibmlb.base.rte
ibmlb.admin.rte
ibmlb.lb.driver
ibmlb.lb.license
ibmlb.<composant>.rte
ibmlb.doc.rte
ibmlb.ms.rte
ibmlb.msg.langue.admin
ibmlb.msg.en_US.doc
ibmlb.msg.langue.lb
```

Les chemins d'installation de Load Balancer sont les suivants :

- Administration - */opt/ibm/edge/lb/admin*
- Composants Load Balancer - */opt/ibm/edge/lb/servers*
- Metric Server - */opt/ibm/edge/lb/ms*

- Documentation (*Guide d'administration*) - `/opt/ibm/edge/lb/documentation`

Pour l'administration à distance de Load Balancer, à l'aide de l'invocation RMI (Remote Method Invocation), vous devez installer les modules Administration, Base, composant et Licence sur le client. Pour plus d'informations sur l'invocation RMI, voir «RMI (Remote Method Invocation)», à la page 262.

Configuration requise et installation pour HP-UX

Configuration requise pour les systèmes HP-UX

Pour plus d'informations sur les conditions matérielles et logicielles requises, y compris les navigateurs pris en charge, accédez à la page Web suivante : <http://www.ibm.com/support/docview.wss?rs=180&uid=swg27006921>

Installation pour les systèmes HP-UX

La présente section explique comment installer Load Balancer sur les systèmes HP-UX à partir du CD-ROM du produit.

Avant de commencer l'installation

Avant de commencer la procédure d'installation, soyez certain de disposer de droits d'accès root pour installer le logiciel.

Si une version antérieure est déjà installée sur votre système, supprimez-la avant d'installer la nouvelle. Vérifiez d'abord que l'exécuteur et le serveur sont arrêtés. Puis, pour désinstaller Load Balancer, voir «Instructions de désinstallation des packages», à la page 36.

Etapes de la procédure d'installation

Le tableau 3 répertorie les noms des packages d'installation de Load Balancer et l'ordre suivant lequel ils doivent être installés à l'aide de l'outil d'installation des packages du système.

Tableau 3. Détails de l'installation des packages de Load Balancer sous HP-UX

Description du package	Nom du package HP-UX
Produit de base	ibmlb.base
Administration et messages	ibmlb.admin ibmlb.nlv-lang
Licence de Load Balancer	ibmlb.lic
Composants de Load Balancer	ibmlb.composant
Documentation	ibmlb.doc
Metric Server	ibmlb.ms
Remarques : <ol style="list-style-type: none"> 1. La variable <i>lang</i> est remplacée par l'un des codes de langue suivants : de_DE, en_US, es_ES, fr_FR, it_IT, ja_JP, ko_KR, zh_CN, zh_TW. 2. La variable <i>composant</i> peut prendre les valeurs suivantes : disp (dispatcher), cbr (CBR), ss (Site Selector), cco (Cisco CSS Controller) ou nal (Nortel Alteon Controller). 3. Le package de la documentation (ibmlb.doc) contient uniquement de l'anglais. Les traductions de l'ensemble des documentations Load Balancer se trouvent sur le site Web : www.ibm.com/software/webservers/appserv/ecinfocenter.html. 	

Remarque : Les systèmes HP-UX ne prennent pas en charge le paramètre régional Portugais (Brésil) (pt_BR). Les paramètres régionaux pris en charge sont les suivants :

- de_DE.iso88591
- en_US.iso88591
- es_ES.iso88591
- fr_FR.iso88591
- it_IT.iso88591
- ja_JP.SJIS
- ko_KR.eucKR
- zh_CN.hp15CN
- zh_TW.big5

Instructions d'installation des packages

La procédure ci-après décrit les étapes requises pour effectuer cette tâche.

1. Connectez-vous en tant qu'utilisateur root.

```
su - root
Mot de passe : motdepasse
```

2. Exécutez la commande permettant d'installer les packages :

```
swinstall -s /source nom_package
```

source représentant le chemin d'accès absolu au répertoire du package et *nom_package*, le nom du package.

Pour installer le package de base de Load Balancer (ibmlb.base) à partir du répertoire racine du CD, entrez la commande suivante :

```
swinstall -s /source ibmlb.base
```

Pour installer tous les packages de Load Balancer, à partir du répertoire racine du CD, entrez la commande suivante :

```
swinstall -s /source ibmlb
```

3. Vérifiez l'installation des packages de Load Balancer

Exécutez la commande **swlist** pour répertorier tous les packages installés. Par exemple,

```
swlist -l fileset ibmlb
```

Instructions de désinstallation des packages

Utilisez la commande **swremove** pour désinstaller les packages. Supprimez les packages dans l'ordre inverse de leur installation. Vous pouvez par exemple exécuter les commandes suivantes :

- Pour désinstaller tous les packages de Load Balancer :

```
swremove ibmlb
```

Pour ne désinstaller qu'un package, par exemple le composant Dispatcher :

```
swremove ibmlb.disp
```

Configuration requise et installation pour Linux

Configuration requise pour les systèmes Linux

Pour plus d'informations sur les conditions matérielles et logicielles requises, y compris les navigateurs pris en charge, accédez à la page Web suivante :

<http://www.ibm.com/support/docview.wss?rs=180&uid=swg27006921>

Installation pour les systèmes Linux

La présente section indique comment installer Load Balancer sous Linux à partir du CD-ROM du produit.

Avant de commencer l'installation

Avant de commencer la procédure d'installation, veillez à disposer des droits d'accès de superutilisateur pour installer le logiciel.

Si une version antérieure est déjà installée sur votre système, supprimez-la avant d'installer la version actuelle. Assurez-vous, tout d'abord, que tous les exécuteurs et serveurs sont arrêtés. Puis, pour procéder à la désinstallation de l'ensemble du produit, entrez **rpm -e pkgname**. Procédez à la désinstallation dans l'ordre inverse des procédures du package d'installation, en vous assurant que les packages d'administration sont les derniers à être désinstallés.

Etapas de la procédure d'installation

Pour installer Load Balancer :

1. Préparez l'installation.

- Connectez-vous en tant que superutilisateur.
- Insérez le support du produit, ou téléchargez-le à partir de notre site Web, et installez l'image d'installation à l'aide de RPM (Gestionnaire de paquets Red Hat).

L'image d'installation est un fichier au format **eLBLX-version:tar.z**.

- Décompactez le fichier compacté dans un répertoire temporaire en entrant la commande suivante : **tar -xf eLBLX-version:tar.z**. Vous obtenez comme résultat un ensemble de fichiers d'extension .rpm.

Voici la liste des packages RPM qui peuvent être installés.

- **ibmlb-base-version-édition.hardw.rpm** (Base)
- **ibmlb-admin-version-édition.hardw.rpm** (Administration)
- **ibmlb-lic-version-édition.hardw.rpm** (Licence)
- **ibmlb-composant-version-édition.hardw.rpm** (composant LB)
- **ibmlb-doc-version-édition.hardw.rpm** (Documentation)
- **ibmlb-ms-version-édition.hardw.rpm** (Metric Server)

Où :

- *version-édition* correspond à l'édition actuelle, par exemple : 6.1-0
- *hardw* peut prendre l'une des valeurs suivantes : i386, ppc64, ppc, s390, s390x, x86_64
- *composant* accepte l'une des valeurs suivantes : disp (composant Dispatcher), cbr (composant CBR), ss (composant Site Selector), cco (contrôleur Cisco CSS), nal (contrôleur Nortel Alteon)

Le package de la documentation contient uniquement de l'anglais. Les traductions de l'ensemble des documentations Load Balancer se trouvent sur le site Web : www.ibm.com/software/webservers/appserv/ecinfocenter.html.

- L'ordre d'installation des packages est important. Voici la liste des packages nécessaires et leur ordre d'installation :
 - Produit de base (base)
 - Administration (admin)
 - Licence (lic)
 - Composants Load Balancer (disp, cbr, ss, cco, nal)

- Metric Server (ms)
- Documentation (doc)

La commande d'installation des packages doit être émise à partir du répertoire où sont situés les fichiers RPM. Entrez la commande suivante pour installer chaque package : **rpm -i package.rpm**.

Systèmes Red Hat Linux : Suite à un problème Red Hat Linux connu, vous devez également supprimer les fichiers RPM *_db**, sous peine de générer une erreur.

- Les chemins d'installation de Load Balancer sont les suivants :
 - Administration - **/opt/ibm/edge/lb/admin**
 - Composants Load Balancer - **/opt/ibm/edge/lb/servers**
 - Metric Server - **/opt/ibm/edge/lb/ms**
 - Documentation - **/opt/ibm/edge/lb/documentation**
- Procédez à la désinstallation des packages dans l'ordre inverse de l'installation, en vous assurant que le package d'administration est le dernier à être désinstallé.

2. Vérifiez que le produit est installé correctement. Entrez la commande suivante :
rpm -qa | grep ibmlb

L'installation du produit complet génère une liste semblable à la suivante :

- *ibmlb-base-version-édition*
- *ibmlb-admin-version-édition*
- *ibmlb-lic-version-édition*
- *ibmlb-dsp-version-édition*
- *ibmlb-cbr-version-édition*
- *ibmlb-ss-version-édition*
- *ibmlb-cco-version-édition*
- *ibmlb-nal-version-édition*
- *ibmlb-doc-version-édition*
- *ibmlb-ms-version-édition*

Pour l'administration à distance de Load Balancer, à l'aide de l'invocation RMI (Remote Method Invocation), vous devez installer les packages d'administration, de produit de base, de composant et de licence sur le client. Pour plus d'informations sur l'invocation RMI, voir «RMI (Remote Method Invocation)», à la page 262.

Configuration requise et installation pour Solaris

Configuration requise pour Solaris

Pour plus d'informations sur les conditions matérielles et logicielles requises, y compris les navigateurs pris en charge, accédez à la page Web suivante :
<http://www.ibm.com/support/docview.wss?rs=180&uid=swg27006921>

Installation pour Solaris

La présente section explique comment installer Load Balancer sur des systèmes Solaris à partir du CD-ROM du produit.

Avant de commencer l'installation

Avant de commencer la procédure d'installation, soyez certain de disposer de droits d'accès root pour installer le logiciel.

Si une version antérieure est déjà installée sur votre système, supprimez-la avant d'installer la nouvelle. Vérifiez d'abord que les exécuteurs et les serveurs sont arrêtés. Puis, pour désinstaller Load Balancer, entrez **pkgrm nom_package**.

Etapas de la procédure d'installation

Pour installer Load Balancer :

1. Préparez l'installation.

- Connectez-vous en tant que root.
- Insérez le CD-ROM contenant le logiciel Load Balancer dans le lecteur approprié.

A l'invite, entrez **pkgadd -d chemin d'accès** où *chemin d'accès* est le nom d'unité du lecteur de CD-ROM ou le répertoire du disque dur contenant le package ; par exemple, **pkgadd -d /cdrom/cdrom0/**.

Voici la liste des packages affichés et leur ordre d'installation recommandé.

- ibmlbbase (Produit de base)
- ibmlbadm (Administration)
- ibmlblic (Licence)
- ibmlbdisp (Composant Dispatcher)
- ibmlbcbr (Composant CBR)
- ibmlbss (Composant Site Selector)
- ibmlbcc (Composant Cisco CSS Controller)
- ibmlbna (Composant Nortel Alteon Controller)
- ibmlbdoc (Documentation)
- ibmlbms (Système Metric Server)

Le package de la documentation (ibmlbdoc) contient uniquement de l'anglais. Les traductions de l'ensemble des documentations Load Balancer se trouvent sur le site Web : www.ibm.com/software/webservers/appserv/ecinfocenter.html.

Pour installer tous les packages, entrez "all" et appuyez sur la touche Entrée. Pour installer uniquement certains composants, entrez le ou les noms correspondants aux packages à installer séparés par un espace ou par une virgule et appuyez sur la touche Entrée. Vous serez peut-être invité à changer vos droits d'accès à certains répertoires ou fichiers. Il suffit d'appuyer sur le bouton Entrée ou de répondre "yes". Vous devez installer les packages prérequis (car l'installation s'effectue suivant l'ordre alphabétique et non en fonction des éléments prérequis). Si vous indiquez "all" et que vous répondez "yes" à toutes les questions, l'installation se déroule sans incident.

Pour n'installer que le composant Dispatcher, la documentation et le système Metric Server, installez : ibmlbbase, ibmlbadm, ibmlblic, ibmlbdisp, ibmlbdoc et ibmlbms.

Pour l'administration à distance de Load Balancer, à l'aide de l'invocation RMI (Remote Method Invocation), vous devez installer les packages Administration, Base, composant et Licence sur le client. Pour plus d'informations sur l'invocation RMI, voir «RMI (Remote Method Invocation)», à la page 262.

Les chemins d'installation de Load Balancer sont les suivants :

- Les composants Load Balancer se trouvent dans le répertoire d'installation **/opt/ibm/edge/lb/servers**.
 - Le composant d'administration se trouve dans le répertoire **/opt/ibm/edge/lb/admin**
 - Metric Server se trouve dans le répertoire **/opt/ibm/edge/lb/ms**
 - La documentation se trouve dans le répertoire **/opt/ibm/edge/lb/documentation**
2. Vérifiez que le produit est installé correctement. Entrez la commande suivante : **pkginfo | grep ibm**.

Configuration requise et installation pour Windows

Configuration requise pour les systèmes Windows

Pour plus d'informations sur les conditions matérielles et logicielles requises, y compris les navigateurs pris en charge, accédez à la page Web suivante : <http://www.ibm.com/support/docview.wss?rs=180&uid=swg27006921>

Installation pour les systèmes Windows

La présente section indique comment installer Load Balancer sur des systèmes Windows à partir du CD-ROM du produit.

Modules d'installation

Un choix de packages à installer vous est proposé :

- Administration
- Licence
- Dispatcher
- CBR (Content Based Routing)
- Site Selector
- Cisco CSS Controller
- Nortel Alteon Controller
- Documentation
- Système Metric Server

Pour l'administration à distance de Load Balancer, à l'aide de l'invocation RMI (Remote Method Invocation), vous devez installer les modules d'administration, de composant et de licence sur le client. Pour plus d'informations sur l'invocation RMI, voir «RMI (Remote Method Invocation)», à la page 262.

Avant de commencer l'installation

Restrictions : La version Windows de Load Balancer ne peut pas être installée sur la même machine qu'IBM Firewall.

Avant de commencer la procédure d'installation, assurez-vous que vous vous êtes connecté en qualité d'administrateur ou d'utilisateur doté de privilèges administratifs.

Si une version antérieure est déjà installée sur votre système, supprimez-la avant d'installer la version actuelle. Pour effectuer une désinstallation en utilisant la fonction **Ajouter/Supprimer un programme**, procédez comme suit :

1. Cliquez sur **Démarrer** > **Paramètres** (pour Windows 2000) > **Panneau de configuration**
2. Cliquez deux fois sur **Ajout/Suppression de programmes**
3. Sélectionnez *IBM WebSphere Edge Components* (ou l'ancien nom, par exemple, *IBM Edge Server*)
4. Cliquez sur le bouton **Modifier/Supprimer**

Etapes de la procédure d'installation

Pour installer Load Balancer :

1. Insérez le CD-ROM de Load Balancer dans le lecteur de CD-ROM. La fenêtre d'installation doit s'afficher automatiquement.
2. Effectuez l'opération qui suit uniquement si le CD n'est pas lancé automatiquement sur votre machine. Cliquez avec le bouton gauche de la souris pour exécuter les tâches suivantes :
 - Cliquez sur **Démarrer**.
 - Sélectionnez **Exécuter**.
 - Indiquez l'unité de CD-ROM, suivie de setup.exe. Par exemple :
`E:\setup`
3. Sélectionnez la **langue** à utiliser pour l'installation.
4. Cliquez sur **OK**.
5. Suivez les instructions du programme d'installation.
6. Si vous désirez changer d'unité ou de répertoire de destination, cliquez sur **Parcourir**.
7. Vous pouvez choisir entre "Produit Load Balancer complet" ou "Les composants de votre choix".
8. Une fois l'installation terminée, un message vous demande de réamorcer le système avant d'utiliser Load Balancer. Cette opération est nécessaire pour vous assurer que tous les fichiers sont installés et que la variable d'environnement IBMLBPATH a bien été ajoutée au registre.

Les chemins d'installation de Load Balancer sont les suivants :

- Administration – **C:\Program Files\IBM\edge\lb\admin**
- Composants Load Balancer – **C:\Program Files\IBM\edge\lb\servers**
- Système Metric Server – **C:\Program Files\IBM\edge\lb\ms**
- Documentation (Guide d'administration) – **C:\Program Files\IBM\edge\lb\documentation**

Remarque : La documentation qui se trouve dans le répertoire d'installation contient uniquement de l'anglais. Les traductions de l'ensemble des documentations Load Balancer se trouvent sur le site Web :
www.ibm.com/software/webservers/appserv/ecinfocenter.html.

Partie 2. Composant Dispatcher

Cette section contient des informations pour la configuration d'un démarrage rapide ainsi que des remarques relatives à la planification, et présente les diverses méthodes de configuration du composant Dispatcher de Load Balancer. Elle se compose des chapitres suivants :

- Chapitre 5, «Configuration de démarrage rapide», à la page 45
- Chapitre 6, «Planification de Dispatcher», à la page 51
- Chapitre 7, «Configuration de Dispatcher», à la page 63
- Chapitre 8, «Déploiement de Dispatcher sur Load Balancer pour IPv4 et IPv6», à la page 81

Chapitre 5. Configuration de démarrage rapide

Cet exemple de démarrage rapide indique comment configurer trois postes de travail connectés en local avec la méthode d'acheminement mac de Dispatcher pour équilibrer la charge du trafic Web entre deux serveurs Web. Cette configuration est également valable pour l'équilibrage de tout autre trafic d'applications TCP ou UDP sans état.

IMPORTANT : Si vous utilisez Load Balancer pour IPv4 et IPv6, voir Chapitre 8, «Déploiement de Dispatcher sur Load Balancer pour IPv4 et IPv6», à la page 81.

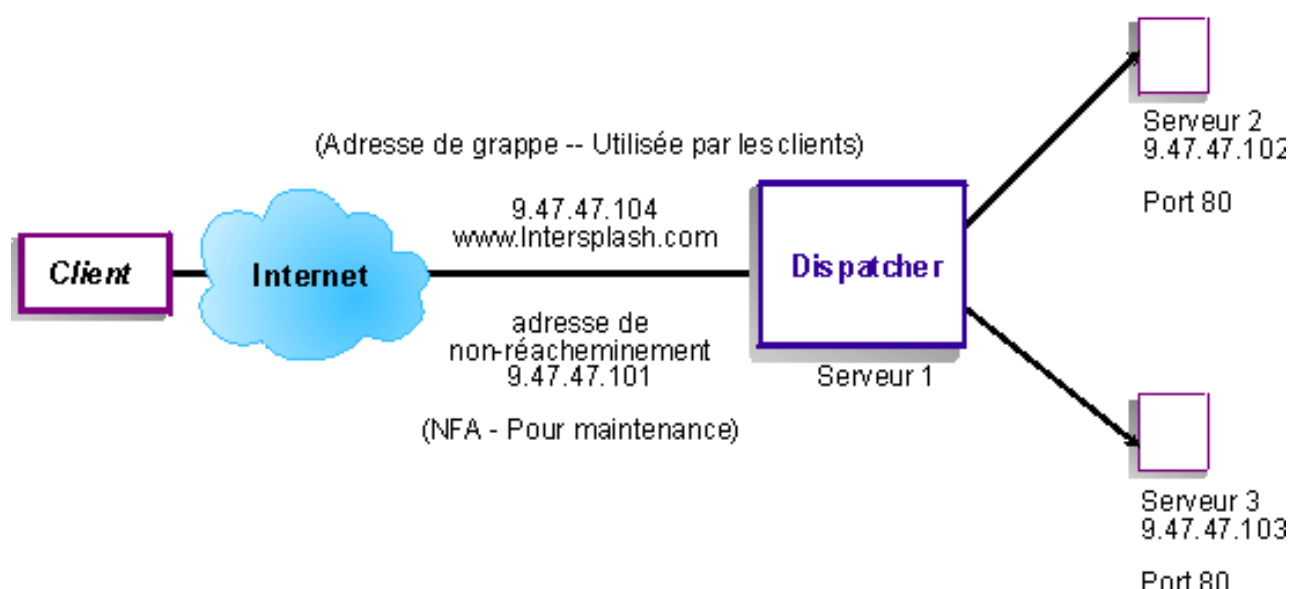


Figure 8. Configuration Dispatcher locale simple

La méthode d'acheminement mac est la méthode d'acheminement par défaut avec laquelle Dispatcher équilibre la charge des demandes entrantes sur le serveur, ce dernier renvoyant la réponse directement au client. Pour plus d'informations sur la méthode d'acheminement MAC de Dispatcher, voir «Réacheminement MAC de Dispatcher (méthode d'acheminement mac)», à la page 53.

Remarque : Vous pouvez effectuer la configuration avec seulement deux postes de travail, Dispatcher étant installé sur l'un des serveurs Web. Il s'agit d'une configuration co-implantée. Les procédures permettant de paramétrer des configurations plus complexes sont présentées à la section «Configuration de la machine Dispatcher», à la page 66.

Matériel requis

Pour l'exemple à démarrage rapide, vous devez disposer de trois postes de travail et de quatre adresses IP. Un poste de travail est la machine Dispatcher ; les deux autres postes sont les serveurs Web. Chaque serveur Web requiert une adresse IP. Le poste de travail Dispatcher requiert deux adresses : l'adresse de non-acheminement (adresse NFA) et l'adresse de la grappe (adresse dont la charge est équilibrée), que vous fournissez aux clients pour accéder à votre site Web.

Remarque : L'adresse NFA est celle que renvoie la commande **hostname**. Cette adresse est utilisée pour des besoins d'administration, comme la configuration à distance.

Préparation

1. Pour cet exemple de configuration avec connexion locale, configurez les postes de travail sur le même segment de réseau local. Vérifiez que le trafic réseau entre les trois machines n'a pas à traverser de routeurs ou de ponts. (Pour les configurations avec serveurs éloignés, voir «Configuration du support de réseau étendu pour Dispatcher», à la page 228.)
2. Configurez les cartes réseau de ces trois postes de travail. Dans cet exemple, nous supposons que vous disposez de la configuration réseau suivante :

Poste de travail	Nom	Adresse IP
1	server1.Intersplashx.com	9.47.47.101
2	server2.Intersplashx.com	9.47.47.102
3	server3.Intersplashx.com	9.47.47.103
Masque réseau = 255.255.255.0		

Chaque poste de travail ne contient qu'une carte d'interface réseau Ethernet standard.

3. Vérifiez que server1.Intersplashx.com peut contacter server2.Intersplashx.com et server3.Intersplashx.com (avec la commande ping).
4. Vérifiez que server2.Intersplashx.com et server3.Intersplashx.com peuvent contacter server1.Intersplashx.com (avec la commande ping).
5. Vérifiez que le contenu est identique sur les deux serveurs Web (Serveur 2 et Serveur 3). Pour cela, répliquez les données des deux postes de travail à l'aide d'un système de fichiers partagé tel que NFS, AFS ou DFS, ou à l'aide de tout autre moyen approprié pour votre site.
6. Vérifiez que les serveurs Web de server2.Intersplashx.com et server3.Intersplashx.com sont opérationnels. Utilisez un navigateur Web pour accéder directement aux pages à partir de **http://server2.Intersplashx.com** et **http://server3.Intersplashx.com**.
7. Cherchez une autre adresse IP valide pour ce segment de réseau local. Il s'agit de l'adresse que vous fournirez aux clients qui désirent accéder à votre site. Dans cet exemple, nous utiliserons :

Name= www.Intersplashx.com
IP=9.47.47.104

8. Configurez les deux serveurs Web pour qu'ils acceptent le trafic de www.Intersplashx.com.
Ajoutez un alias pour www.Intersplashx.com à l'interface de **bouclage** de server2.Intersplashx.com et server3.Intersplashx.com.
 - Pour les systèmes AIX :
ifconfig lo0 alias www.Intersplashx.com netmask 255.255.255.0
 - Pour les systèmes Solaris 9 :
ifconfig lo0:1 plumb www.Intersplashx.com netmask 255.255.255.0 up
 - Pour les autres systèmes d'exploitation, voir tableau 5, à la page 73.
9. Supprimez toute autre route créée à la suite de l'attribution d'un alias à l'interface de bouclage. Voir «Etape 2. Vérification de l'existence d'une route supplémentaire», à la page 76.

Vous venez de terminer les étapes de configuration requises pour les deux serveurs Web.

Configuration du composant Dispatcher

A l'aide de Dispatcher, vous pouvez créer une configuration à l'aide de la ligne de commande, de l'assistant de configuration ou de l'interface graphique.

Remarque : Les valeurs des paramètres doivent être saisies à l'aide de caractères anglais. Les seules exceptions sont les valeurs des paramètres des noms d'hôte et des noms de fichiers.

Configuration à partir de la ligne de commande

Si vous utilisez la ligne de commande, suivez la procédure ci-dessous.

1. Démarrez dsserver sur Dispatcher :
 - Pour les systèmes AIX, HP-UX, Linux ou Solaris, exécutez la commande suivante en tant que superutilisateur : **dsserver**
 - Pour les systèmes Windows, dsserver est exécuté en tant que service qui démarre automatiquement.
2. Démarrez la fonction exécuteur (executor) de Dispatcher :
dscontrol executor start
3. Ajoutez l'adresse du cluster à la configuration de Dispatcher :
dscontrol cluster add www.Intersplashx.com
4. Ajoutez le port du protocole HTTP à la configuration de Dispatcher :
dscontrol port add www.Intersplashx.com:80
5. Ajoutez chaque serveur Web à la configuration Dispatcher :
dscontrol server add www.Intersplashx.com:80:server2.Intersplashx.com
dscontrol server add www.Intersplashx.com:80:server3.Intersplashx.com
6. Configurez la machine pour accepter le trafic des autres adresses IP :
dscontrol executor configure www.Intersplashx.com
7. Démarrez la fonction gestionnaire (manager) de Dispatcher :
dscontrol manager start
Dispatcher procède maintenant à l'équilibrage de charge en fonction des performances des serveurs.
8. Démarrez la fonction conseiller (advisor) de Dispatcher :
dscontrol advisor start http 80
Dispatcher vérifie désormais que les demandes des clients ne sont pas envoyées vers un serveur Web arrêté.

La configuration de base comportant des serveurs liés en local est maintenant terminée.

Test de vérification de la configuration

Vérifiez que la configuration fonctionne :

1. A l'aide d'un navigateur Web, accédez à **http://www.Intersplashx.com**. Si une page s'affiche, la configuration fonctionne.
2. Rechargez la page dans le navigateur Web.

3. Vérifiez les résultats de la commande suivante : **dscontrol server report www.Intersplashx.com:80:**. La colonne du nombre total de connexions des deux serveurs doit contenir la valeur "2."

Configuration à l'aide de l'interface graphique

Pour plus d'informations sur l'utilisation de l'interface graphique, voir «Interface graphique», à la page 65 et Annexe A, «Interface graphique utilisateur : Instructions générales», à la page 469.

Assistant de configuration

Pour plus d'informations sur l'utilisation de l'assistant de configuration, voir «Configuration à l'aide de l'assistant de configuration», à la page 66.

Types de configurations de cluster, de port et de serveur

Il y a plusieurs manières de configurer Load Balancer pour assurer le support de votre site. Si votre site ne comprend qu'un seul nom de système hôte auquel tous vos clients se connectent, vous pouvez ne définir qu'un seul cluster de serveurs. Pour chaque serveur, configurez un port par l'intermédiaire duquel Load Balancer communique. Voir figure 9.

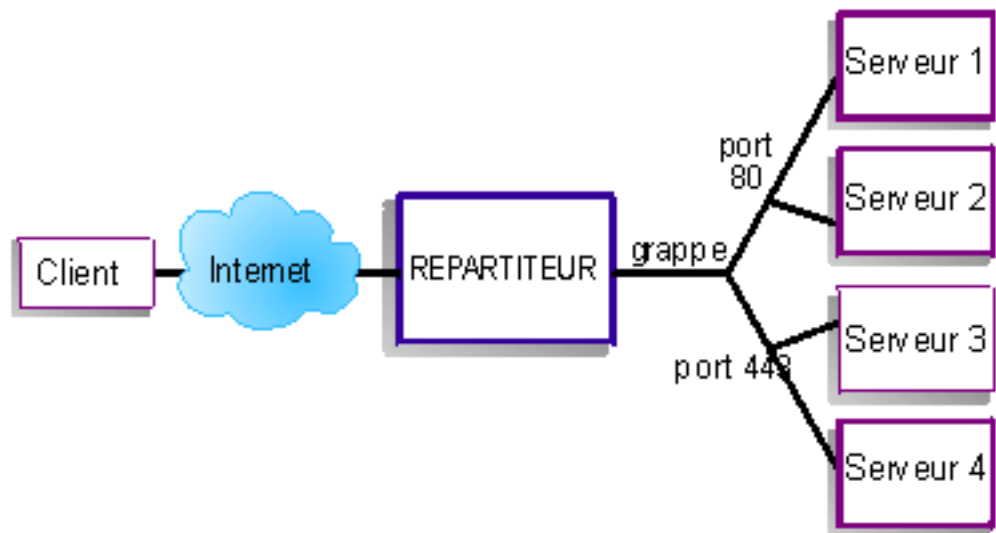


Figure 9. Exemple de composant Dispatcher configuré avec un cluster et 2 ports

Dans cet exemple de composant Dispatcher, un cluster est défini sur www.productworks.com. Il dispose de deux ports : le port 80 pour HTTP et le port 443 pour SSL. Un client adressant une requête à l'adresse <http://www.productworks.com> (port 80) accède à un autre serveur qu'un client s'adressant à <https://www.productworks.com> (port 443).

Si le site est très étendu et qu'il comporte un grand nombre de serveurs, chacun étant dédié à un protocole en particulier, une autre méthode de configuration de Load Balancer sera peut-être préférable. Dans ce dernier cas, il est souhaitable de définir un cluster pour chaque protocole, avec un seul port mais plusieurs serveurs, comme illustré à la figure 10, à la page 49.

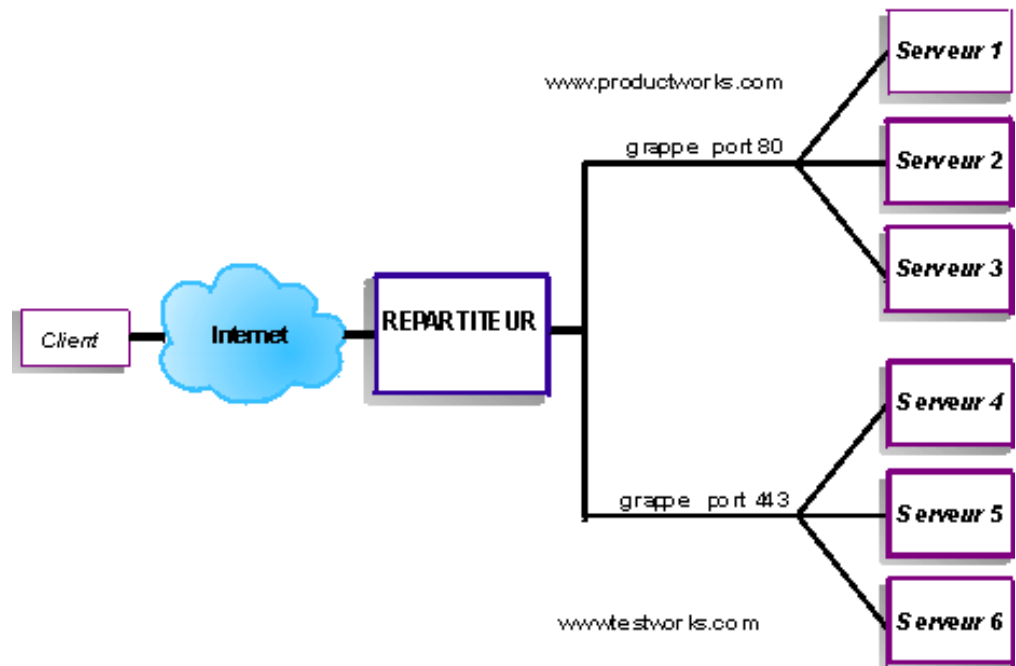


Figure 10. Exemple de composant Dispatcher configuré avec deux clusters, chacun étant associé à un port

Dans cet exemple de composant Dispatcher, deux clusters sont définis : `www.productworks.com` pour le port 80 (HTTP) et `www.testworks.com` pour le port 443 (SSL).

Une troisième configuration de Load Balancer pourra s'avérer nécessaire si votre site abrite plusieurs sociétés ou services, chacun accédant à votre site par une adresse URL distincte. Dans ce cas, vous pouvez définir un cluster pour chaque société ou service ainsi qu'un nombre de ports variable pour réceptionner les connexions de cette URL, comme illustré par la figure 11, à la page 50.

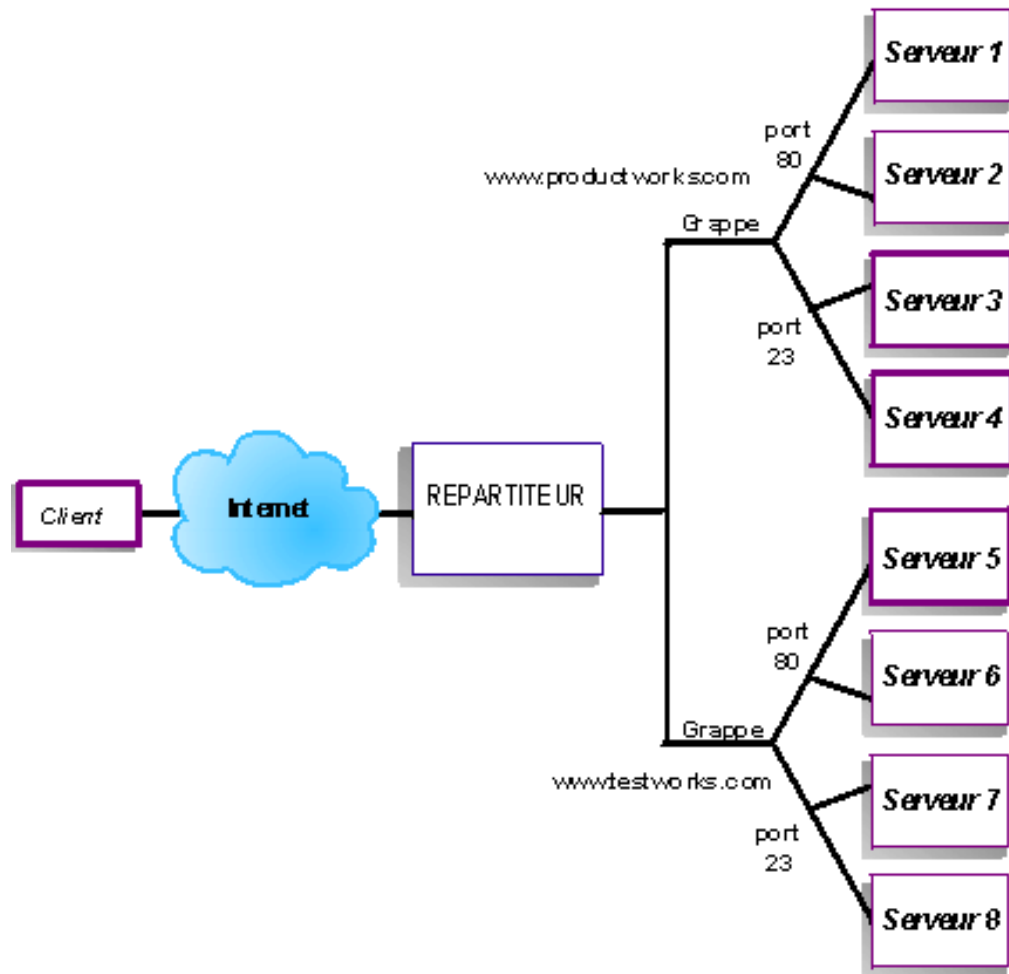


Figure 11. Exemple de composant Dispatcher configuré avec 2 clusters, chacun étant associé à 2 ports

Dans cet exemple de composant Dispatcher, deux clusters sont définis avec le port 80 pour HTTP et le port 23 pour Telnet pour chacun des sites www.productworks.com et www.testworks.com.

Chapitre 6. Planification de Dispatcher

Le présent chapitre décrit les aspects que l'administrateur de réseau doit prendre en compte avant d'installer et de configurer le composant Dispatcher.

- Voir Chapitre 3, «Gestion du réseau : Fonctions Load Balancer requises», à la page 19 pour une présentation des fonctions permettant de gérer votre réseau.
- Voir Chapitre 7, «Configuration de Dispatcher», à la page 63 pour obtenir des informations sur la configuration des paramètres d'équilibrage de charge de Dispatcher.
- Voir Chapitre 8, «Déploiement de Dispatcher sur Load Balancer pour IPv4 et IPv6», à la page 81 si vous utilisez Load Balancer pour IPv4 et IPv6.
- Voir Chapitre 22, «Fonctions avancées de Dispatcher, CBR et Site Selector», à la page 201 pour obtenir des informations sur la configuration de Load Balancer pour les fonctions avancées.
- Voir Chapitre 24, «Exploitation et gestion de Load Balancer», à la page 261, pour obtenir des informations sur l'administration authentifiée à distance, les fichiers journaux Load Balancer et sur l'utilisation des composants Load Balancer.

Il contient les sections suivantes :

- «Remarques relatives à la planification»
- «Réacheminement MAC de Dispatcher (méthode d'acheminement mac)», à la page 53
- «Réacheminement NAT/NAPT de Dispatcher (méthode d'acheminement nat)», à la page 53
- «Fonction CBR de Dispatcher (méthode d'acheminement cbr)», à la page 55
- «Partitionnement du serveur : serveurs logiques configurés pour un serveur physique (adresse IP)», à la page 58
- «Haute disponibilité», à la page 60

Remarque : Dans les versions antérieures où le produit se nommait Network Dispatcher, la commande de contrôle de Dispatcher était `ndcontrol`. Elle s'intitule désormais **`dscontrol`**.

Remarques relatives à la planification

IMPORTANT : Si vous utilisez Load Balancer pour IPv4 et IPv6, voir Chapitre 8, «Déploiement de Dispatcher sur Load Balancer pour IPv4 et IPv6», à la page 81.

Dispatcher se compose des fonctions suivantes :

- **`dsserver`** traite les demandes à partir de la ligne de commande adressées à l'exécuteur, au gestionnaire et aux conseillers.
- L'**exécuteur** assure l'équilibrage de la charge des connexions TCP et UDP sur la base des ports. Il peut transmettre des connexions à des serveurs en fonction du type de demande reçu (par exemple, HTTP, FTP, SSL, etc.) L'exécuteur s'exécute toujours lorsque le composant Dispatcher est utilisé pour l'équilibrage de charge.
- Le **gestionnaire** définit les mesures utilisées par l'exécuteur en fonction de plusieurs facteurs :
 - les décomptes internes de l'exécuteur,
 - le retour d'informations sur les serveurs fourni par les conseillers,

- le retour d’informations émanant d’un programme de contrôle système, tel que Metric Server ou WLM.

L’utilisation du gestionnaire n’est que facultative. Toutefois, s’il n’est pas utilisé, l’équilibrage de charge se fait sur la base d’une planification circulaire pondérée, elle-même basée sur les mesures de charge des serveurs et les conseillers ne sont pas disponibles.

- Les **conseillers** interrogent les serveurs puis analysent les résultats par protocole avant de demander au gestionnaire de régler les capacités comme il convient. Actuellement, il existe des conseillers pour les protocoles suivants : HTTP, FTP, SSL, SMTP, NNTP, IMAP, POP3, SIP et Telnet.

Dispatcher fournit également des conseillers qui n’échangent pas d’informations relatives aux protocoles, tels que le conseiller DB2 qui indique l’état des serveurs DB2 et le conseiller ping qui indique si le serveur répond à une commande ping. Pour connaître la liste complète des conseillers, voir «Liste des conseillers», à la page 188.

Vous avez également la possibilité de développer vos propres conseillers (voir «Création de conseillers personnalisés», à la page 192).

L’utilisation des conseillers est facultative mais recommandée.

- Pour configurer et gérer l’exécuteur, les conseillers et le gestionnaire, utilisez la ligne de commande (**dscontrol**) ou l’interface utilisateur graphique (**lbadmin**).
- Un **fichier de configuration exemple** est fourni et peut être utilisé pour la configuration et l’administration de la machine Dispatcher. Voir Annexe C, «Exemples de fichiers de configuration», à la page 481. Une fois le produit installé, ce fichier se trouve dans le sous-répertoire **...ibm/edge/lb/servers/samples** où réside Load Balancer.
- Le **sous-agent SNMP** permet à une application de gestion de type SNMP de contrôler l’état de Dispatcher.

Les trois fonctions clés de Dispatcher (l’exécuteur, le gestionnaire et les conseillers) agissent en collaboration pour équilibrer et répartir entre les serveurs les requêtes réceptionnées. Outre la gestion des requêtes d’équilibrage de charge, l’exécuteur contrôle le nombre de nouvelles connexions, de connexions actives et de connexions terminées. Il assure également le retrait des connexions terminées ou réinitialisées et transmet ces informations au gestionnaire.

Le gestionnaire recueille les informations transmises par l’exécuteur, les conseillers et par tout programme de contrôle tel que Metric Server. Sur la base de ces informations, le gestionnaire ajuste les capacités des machines serveurs, pour chaque port, et transmet ces données à l’exécuteur qui en tient compte pour l’équilibrage de charge des nouvelles connexions.

Les conseillers contrôlent chaque serveur relié au port dont ils ont la charge afin de déterminer leur temps de réponse et leur disponibilité, puis retournent ces informations au gestionnaire. Les conseillers détectent également si un serveur est opérationnel ou non. Sans la contribution du gestionnaire et des conseillers, l’exécuteur assure une planification circulaire basée sur les capacités courantes des serveurs.

Méthodes d’acheminement

Avec Dispatcher, vous pouvez choisir l’une des trois méthodes d’acheminement spécifiées au niveau du port : acheminement MAC, acheminement NAT/NAPT ou acheminement CBR (routage par contenu).

Réacheminement MAC de Dispatcher (méthode d'acheminement mac)

La méthode d'acheminement MAC de Dispatcher (qui est la méthode d'acheminement par défaut) permet d'équilibrer la charge de la demande entrante sur le serveur sélectionné et de faire en sorte que le serveur renvoie une réponse *directement* au client sans impliquer le composant Dispatcher. Ainsi, Dispatcher se contente de surveiller les flux entrants du client vers le serveur. Il n'effectue aucun contrôle des transmissions en sortie, du serveur vers le client. Cet aspect réduit sensiblement son impact sur les performances des applications et permet même d'accroître celles du réseau.

La méthode d'acheminement peut être sélectionnée lors de l'ajout d'un port à l'aide de la commande **dscontrol port add cluster:port method valeur**. La valeur de la méthode d'acheminement par défaut est **mac**. Vous ne pouvez spécifier le paramètre **method** que lorsque le port est ajouté. Une fois le port ajouté, vous ne pouvez pas modifier les paramètres de la méthode d'acheminement. Pour plus d'informations, voir «dscontrol port — Configuration des ports», à la page 378.

Restriction sous Linux : les systèmes Linux emploient un modèle fondé sur l'hôte pour afficher les adresses matérielles sous la forme d'adresses IP à l'aide du protocole ARP. Ce modèle n'est pas compatible avec les conditions requises par le serveur dorsal ou le serveur co-implanté à haute disponibilité pour la prise en charge de la méthode d'acheminement MAC de Load Balancer. Pour modifier le comportement du système Linux afin de le rendre compatible avec l'acheminement MAC de Load Balancer, voir les solutions présentées dans «Solutions alternatives pour l'affectation d'alias à l'unité de bouclage sous Linux lors de l'utilisation de la méthode d'acheminement MAC de Load Balancer», à la page 78.

Restriction sous Linux lors de l'utilisation de serveurs zSeries ou S/390 : Il existe des limitations lorsque vous utilisez des serveurs zSeries ou S/390 dotés de cartes OSA (Open System Adapter). Pour des solutions possibles, voir «Incident : Sous Linux, limitations de la configuration Dispatcher lors de l'utilisation de serveurs zSeries ou S/390 dotés de cartes OSA (Open System Adapter)», à la page 320.

Réacheminement NAT/NAPT de Dispatcher (méthode d'acheminement nat)

Si vous utilisez NAT ou NAPT, il n'est pas nécessaire que les serveurs d'équilibrage de charge se trouvent sur un réseau local. Si vous préférez disposer de serveurs à distance, utilisez la méthode d'acheminement NAT plutôt que la technique d'encapsulation GRE/WAN. Vous pouvez également utiliser la fonction NAPT pour accéder à plusieurs démons de serveur situés sur chaque machine serveur faisant l'objet d'un équilibrage de charge, où chaque démon écoute sur un port unique.

Vous pouvez configurer un serveur à plusieurs démons de deux façons différentes.

- Avec NAT, vous pouvez configurer plusieurs démons de serveur pour qu'ils répondent aux demandes selon les adresses IP. En d'autres termes, il s'agit de lier un démon de serveur à une adresse IP.
- Avec NAPT, vous pouvez configurer plusieurs démons (qui s'exécutent sur le même serveur physique) pour qu'ils écoutent sur différents numéros de port.

L'application fonctionne bien avec des protocoles de niveau supérieur tels que HTTP, SSL, IMAP, POP3, NNTP, SMTP, Telnet, etc.

Restrictions :

- L'implémentation de NAT/NAPT de Dispatcher est une implémentation *simple* de cette fonction. Il ne procède à l'analyse et ne traite que le contenu des en-têtes de paquets TCP/IP. Il n'analyse pas le contenu de la partie données des paquets. Pour Dispatcher, NAT/NAPT ne fonctionnera pas avec les protocoles d'application, tels que FTP, qui intègrent les adresses ou les numéros de port dans la partie données des messages. Il s'agit d'une restriction déjà identifiée des fonctions NAT/NAPT basées sur les en-têtes.
- La fonction NAT/NAPT de Dispatcher ne peut pas fonctionner si vous utilisez la fonction de cluster générique ou de port générique.

Vous avez besoin de trois adresses IP pour la machine Dispatcher : l'adresse nfa, l'adresse de cluster et l'adresse de retour. Pour implémenter NAT/NAPT, procédez comme suit (voir également «Etapes de configuration des méthodes d'acheminement nat ou cbr de Dispatcher», à la page 57) :

- Définissez le paramètre **clientgateway** à l'aide de la commande **dscontrol executor set**. Clientgateway est une adresse IP correspondant à l'adresse du routeur par lequel le trafic de retour est acheminé de Load Balancer vers les clients. Sa valeur doit correspondre à une adresse IP non nulle pour que vous puissiez utiliser NAT/NAPT. Pour plus d'informations, voir «dscontrol executor — Contrôle de l'exécuteur», à la page 357.
- Ajoutez un port à l'aide de la commande **dscontrol port add cluster:port method valeur**. La valeur de la méthode d'acheminement doit être associée à **nat**. Vous ne pouvez spécifier le paramètre method que lorsque le port est ajouté. Une fois le port ajouté, vous ne pouvez pas modifier les paramètres de la méthode d'acheminement. Pour plus d'informations, voir «dscontrol port — Configuration des ports», à la page 378.

Remarque : Si vous n'associez pas une valeur non nulle à l'adresse de passerelle, la méthode d'acheminement ne peut être que **mac** (méthode d'acheminement basée sur MAC).

- Ajoutez un serveur avec les paramètres mapport, returnaddress et router à l'aide de la commande **dscontrol**. Par exemple :

```
dscontrol server add cluster:port:serveur mapport valeur returnaddress  
adresseretur router adresseretur
```

– **mapport** (facultatif)

Mappe le numéro du port de destination de la demande client (pour Dispatcher) au numéro de port du serveur que Dispatcher utilise pour équilibrer la charge de la demande du client. Mapport permet à Load Balancer de recevoir une demande de client sur un port et de la transmettre à un autre port sur la machine serveur. Le paramètre mapport permet d'équilibrer la charge des demandes d'un client sur une machine serveur sur laquelle peuvent s'exécuter plusieurs démons serveur. La valeur par défaut du paramètre mapport est le numéro de port de destination de la demande du client.

– **returnaddress**

L'adresse retour correspond à une adresse ou à un nom d'hôte unique que vous configurez sur la machine Dispatcher. Dispatcher utilise l'adresse de retour comme adresse source lors de l'équilibrage de charge de la demande du client sur le serveur. Elle permet de garantir que le serveur renvoie le paquet à la machine Dispatcher, au lieu de l'envoyer directement au client. (Dispatcher transmettra ensuite le paquet IP au client.) Vous devez indiquer la valeur d'adresse de retour lors de l'ajout du serveur. Vous ne pouvez pas

modifier l'adresse de retour sauf si vous supprimez le serveur et que vous l'ajoutez à nouveau. L'adresse de retour ne peut pas être identique à l'adresse de cluster, de serveur ou NFA.

– **router**

Adresse du routeur vers le serveur éloigné. S'il s'agit d'un serveur rattaché en local, entrez son adresse, sauf s'il réside sur la même machine que Load Balancer. Dans ce cas, continuez à utiliser l'adresse réelle du routeur.

Pour plus d'informations sur la commande **dscontrol server** et les paramètres **mapport**, **returnaddress** et **router**, voir «dscontrol server — Configuration des serveurs», à la page 390.

Fonction CBR de Dispatcher (méthode d'acheminement cbr)

Le composant Dispatcher permet d'exécuter la fonction CBR (content-based routing) pour HTTP (avec la règle de type de contenu) et HTTPS (avec l'affinité des ID de session SSL) sans Caching Proxy. Pour le trafic HTTP et HTTPS, la méthode d'acheminement cbr du composant Dispatcher peut fournir une fonction CBR (content-based routing) plus rapide que le composant CBR qui nécessite le module Caching Proxy.

Pour HTTP : La sélection du serveur, pour fonction CBR de Dispatcher, est effectuée sur la base du contenu d'une adresse URL ou d'un en-tête HTTP. Cette option est configurée à l'aide du type de règle "Contenu". Lors de la configuration de la règle de contenu, spécifiez la chaîne de recherche "pattern" et un ensemble de serveurs pour la règle. Lors du traitement d'une nouvelle demande entrante, cette règle compare la chaîne indiquée à l'URL du client ou à l'en-tête HTTP spécifié dans la demande du client.

Si Dispatcher trouve la chaîne dans la demande du client, il transmet la demande à l'un des serveurs de la règle. Dispatcher achemine ensuite les données de la réponse du serveur vers le client (méthode d'acheminement cbr).

Si Dispatcher ne trouve pas la chaîne dans la demande du client, il ne sélectionne *pas* de serveur dans l'ensemble de serveurs de la règle.

Remarque : La règle de contenu est configurée dans le composant Dispatcher de la même façon que dans le composant CBR. Dispatcher peut utiliser la règle de contenu pour le trafic HTTP. Toutefois, le composant CBR peut utiliser la règle de contenu *à la fois* pour le trafic HTTP et HTTPS (SSL).

Pour HTTPS (SSL) : L'acheminement CBR (content-based routing) de Dispatcher basée sur la zone de session SSL ID de la demande client. Avec SSL, une demande client contient l'ID session SSL d'une session antérieure, et les serveurs gèrent une cache de leurs connexions SSL précédentes. L'affinité de l'ID de session SSL de Dispatcher permet au client et au serveur d'établir une nouvelle connexion à l'aide des paramètres de sécurité de la connexion précédente au serveur. En éliminant la renégociation des paramètres de sécurité SSL, comme les clés partagées et les algorithmes de chiffrement, les serveurs sauvegardent des cycles CPU et le client obtient une réponse plus rapidement. Pour activer l'affinité de l'ID de session SSL : le type de **protocole** indiqué pour le port doit être **SSL** et le **délai de maintien de routage** du port doit être associé à une valeur autre que zéro. Si le délai de maintien de routage est dépassé, le client peut être envoyé à un autre serveur.

Vous avez besoin de trois adresses IP pour la machine Dispatcher : l'adresse nfa, l'adresse de cluster et l'adresse de retour. Pour implémenter la fonction CBR de Dispatcher, procédez aux opérations ci-dessous (voir également «Étapes de configuration des méthodes d'acheminement nat ou cbr de Dispatcher», à la page 57):

- Définissez le paramètre **clientgateway** à l'aide de la commande **dscontrol executor set**. Clientgateway est une adresse IP correspondant à l'adresse du routeur par lequel le trafic de retour est acheminé de Dispatcher vers les clients. La valeur par défaut de clientgateway est zéro. Vous devez associer cette valeur à une adresse IP différente de zéro pour pouvoir ajouter une méthode d'acheminement CBR (fonction CBR (content-based routing)). Pour plus d'informations, voir «dscontrol executor — Contrôle de l'exécuteur», à la page 357.
- Ajoutez un port avec les paramètres **method** et **protocol** dans la commande **dscontrol port add**. La valeur de la méthode d'acheminement doit correspondre à **cbr**. Le type de protocole de port peut être HTTP ou SSL. Pour plus d'informations, voir «dscontrol port — Configuration des ports», à la page 378.

Remarque : Si vous n'associez pas une valeur différente de zéro à l'adresse de passerelle client, la méthode d'acheminement ne peut être que de type **MAC**.

- Ajoutez un serveur avec les paramètres **mapport**, **returnaddress** et **router**
**dscontrol server add cluster:port:serveur mapport valeur returnaddress
adresseretur router adresseretur**

Remarque : Pour plus d'informations sur la configuration du serveur avec les paramètres **mapport** (facultatif), **returnaddress** et **router**, reportez-vous à la page 54.

- **Pour HTTP :** procédez à la configuration à l'aide de règles basées sur le contenu de la demande client (type de règle **contenu**). Par exemple,
dscontrol rule 125.22.22.03:80:contentRule1 type content pattern motif
où *masque* indique le masque à utiliser pour une règle de type de contenu. Pour plus d'informations sur le type de règle de contenu, voir «Utilisation de règles basées sur le contenu des demandes», à la page 219. Pour plus d'informations sur les expressions valides de *masque*, voir Annexe B, «Syntaxe des règles de contenu (modèle)», à la page 477.

Remarque : La fonction de réplication des enregistrements de connexions de haute-disponibilité (qui garantit que la connexion d'un client ne sera pas annulée lorsqu'une machine Dispatcher de sauvegarde remplace la machine principale) n'est *pas* pris en charge avec la fonction **cbr** de Dispatcher.

Etapes de configuration des méthodes d'acheminement nat ou cbr de Dispatcher

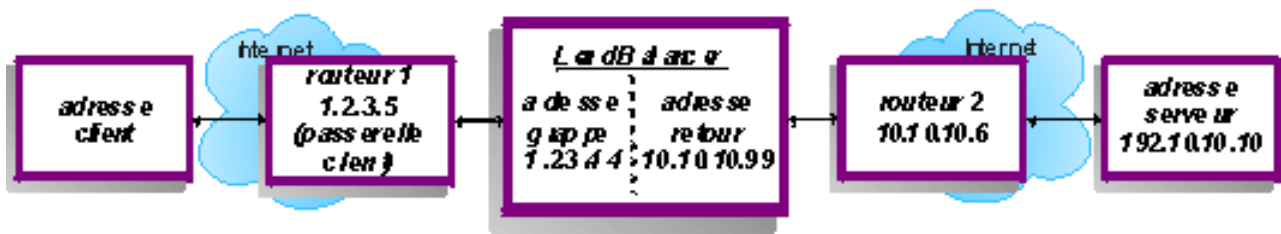


Figure 12. Exemple d'utilisation des méthodes d'acheminement nat ou cbr de Dispatcher

Vous avez besoin d'au moins trois adresses IP pour la machine Dispatcher. Pour la figure 12, les étapes minimales de configuration des méthodes nat ou cbr de Dispatcher sont les suivantes :

1. Démarrage de l'exécuteur
`dscontrol executor start`
2. Définition de la passerelle client
`dscontrol executor set clientgateway 1.2.3.5`
 REMARQUE : si votre sous-réseau ne dispose pas de routeur local, vous devez configurer une machine pour l'acheminement des adresses IP et l'utiliser comme passerelle client. Reportez-vous à la documentation de votre système d'exploitation pour déterminer la manière d'activer l'acheminement des adresses IP.
3. Définition de l'adresse de cluster
`dscontrol cluster add 1.2.3.44`
4. Configuration de l'adresse de cluster
`dscontrol executor configure 1.2.3.44`
5. Définition du port avec une méthode nat ou cbr
`dscontrol port add 1.2.3.44:80 method nat`
 ou
`dscontrol port add 1.2.3.44:80 method cbr protocol http`
6. Configuration d'une adresse de retour alias sur Load Balancer (à l'aide de la carte ethernet 0)
 NOTE : Sur les systèmes Linux, vous n'avez pas besoin de créer un alias pour l'adresse de retour si vous utilisez la méthode de transfert nat sur une machine co-implantée.
`dscontrol executor configure 10.10.10.99`
 ou utilisez la commande `ifconfig` (pour Linux ou UNIX uniquement) :
 sous AIX : `ifconfig en0 alias 10.10.10.99 netmask 255.255.255.0`
 HP-UX : `ifconfig lan0:1 10.10.10.99 netmask 255.255.255.0 up`
 sous Linux : `ifconfig eth0:1 10.10.10.99 netmask 255.255.255.0 up`
 Solaris : `ifconfig eri0 addif 10.10.10.99 netmask 255.255.255.0 up`
7. Définition des serveurs dorsaux
`dscontrol server add 1.2.3.4:80:192.10.10.10`
`router 10.10.10.6 returnaddress 10.10.10.99`

La passerelle client (1.2.3.5) correspond à l'adresse du routeur 1 entre Load Balancer et le client. Router (10.10.10.6) correspond à l'adresse du routeur 2 entre Load Balancer et le serveur dorsal. Si vous n'êtes pas sûr de l'adresse de la passerelle client ou du routeur 2, vous pouvez utiliser un programme `tracert` avec l'adresse du client (ou du serveur) pour déterminer l'adresse du routeur. La

syntaxe d'appel de ce programme varie en fonction du système d'exploitation utilisé. Pour plus d'informations sur ce programme, consultez la documentation afférente à votre programme d'exploitation.

Si le serveur se trouve sur le même sous-réseau que Load Balancer (c'est-à-dire si traceroute ne désigne aucun routeur), entrez l'adresse du serveur comme adresse de routeur. Cependant, si le serveur réside sur la même machine que Load Balancer, entrez plutôt l'adresse du routeur que celle du serveur. L'adresse du routeur est celle utilisée dans la commande "server add", sur la machine Load Balancer, à l'étape 7.

Partitionnement du serveur : serveurs logiques configurés pour un serveur physique (adresse IP)

Avec le partitionnement du serveur, vous pouvez effectuer une distinction plus avancée entre des URL particulières et leurs applications spécifiques. Par exemple, un serveur Web permet de gérer des pages JSP, des pages HTML, des fichiers GIF, des requêtes de base de données, etc. Load Balancer permet maintenant de partitionner un cluster et un serveur spécifiques d'un port en plusieurs serveurs logiques. Ainsi, vous pouvez appliquer le conseiller sur un service particulier de la machine afin de détecter si un moteur de servlet ou une demande de base de données s'exécute très rapidement ou s'il ne s'exécute pas du tout.

Le partitionnement de serveur permet à Load Balancer de détecter, par exemple, que le service HTML traite les pages rapidement mais que la connexion à la base de données a été interrompue. Ainsi vous pouvez distribuer la charge en fonction de la charge de travail de chaque service plus granulaire et non en fonction uniquement de la pondération serveur.

Partitionnement de serveur à l'aide des conseillers HTTP ou HTTPS

Le partitionnement de serveur peut se révéler utile s'il est associé aux conseillers HTTP et HTTPS. Par exemple, lorsque vous disposez d'un serveur HTML qui gère des pages HTML, GIF et JSP, si vous définissez le serveur (par ajout) une seule fois sur le port 80, vous ne recevez qu'une valeur de charge pour la totalité du serveur HTTP. Ceci peut vous induire en erreur car il est possible que le service GIF ne fonctionne pas sur le serveur. Dispatcher continue d'acheminer les pages GIF vers le serveur, mais le client n'obtient qu'un message de dépassement de délai ou d'erreur.

Si vous définissez le serveur trois fois (par exemple, ServerHTML, ServerGIF et ServerJSP) sur le port et que vous définissez le paramètre **advisorrequest** du serveur avec une chaîne différente pour chaque serveur logique, vous pouvez demander des informations concernant l'état d'un service particulier sur le serveur. ServerHTML, ServerGIF et ServerJSP correspondent à trois serveurs logiques partitionnés à partir d'un serveur physique. Pour ServerJSP, vous pouvez définir la chaîne **advisorrequest** afin d'interroger le service sur la machine qui gère les pages JSP. Pour ServerGIF, vous pouvez définir la chaîne **advisorrequest** afin d'interroger le service GIF. Pour ServerHTML, vous pouvez définir la chaîne **advisorrequest** afin d'interroger le service HTML. Ainsi, lorsque le client n'obtient pas de réponse de l'interrogation **advisorrequest** du service GIF, Dispatcher marque ce serveur logique (ServerGIF) comme inactif tandis que les deux autres serveurs logiques peuvent parfaitement fonctionner. Dispatcher n'achemine plus de pages GIF vers le serveur physique, mais peut encore envoyer des requêtes JSP et HTML au serveur.

Pour plus d'informations sur le paramètre **advisorrequest**, voir «Configuration du conseiller HTTP ou HTTPS à l'aide de l'option de demande ou de réponse (URL)», à la page 190.

Exemple de configuration d'un serveur physique en plusieurs serveurs logiques

Dans la configuration de Dispatcher, vous pouvez représenter un serveur physique ou un serveur logique à l'aide de la hiérarchie *cluster:port:serveur*. Le serveur peut être une adresse IP unique de la machine (serveur physique) sous la forme d'un nom symbolique ou d'une adresse IP. Ou, si vous définissez le serveur afin qu'il représente un serveur partitionné, vous devez alors fournir une adresse de serveur pouvant être résolue pour le serveur physique dans le paramètre **address** de la commande **dscontrol server add**. Pour plus d'informations, voir «dscontrol server — Configuration des serveurs», à la page 390.

Voici un exemple de partitionnement de serveurs physiques en serveurs logiques afin de gérer différents types de demandes.

```
Cluster: 1.1.1.1
  Port: 80
    Server: A (IP address 1.1.1.2)
              HTML server
    Server: B (IP address 1.1.1.2)
              GIF server
    Server: C (IP address 1.1.1.3)
              HTML server
    Server: D (IP address 1.1.1.3)
              JSP server
    Server: E (IP address 1.1.1.4)
              GIF server
    Server: F (IP address 1.1.1.4)
              JSP server
  Rule1: /*.htm
    Server: A
    Server: C
  Rule2: /*.jsp
    Server: D
    Server: F
  Rule3: /*.gif
    Server: B
    Server: E
```

Dans cet exemple, le serveur 1.1.1.2 est divisé en deux serveurs logiques : "A" (gérant les demandes HTML) et "B" (gérant les demandes GIF). Le serveur 1.1.1.3 est divisé en deux serveurs logiques : "C" (gérant les demandes HTML) et "D" (gérant les demandes JSP). Le serveur 1.1.1.4 est partitionné en deux serveurs logiques : "E" (gérant les demandes GIF) et "F" (gérant les demandes JSP).

Haute disponibilité

Haute disponibilité simple

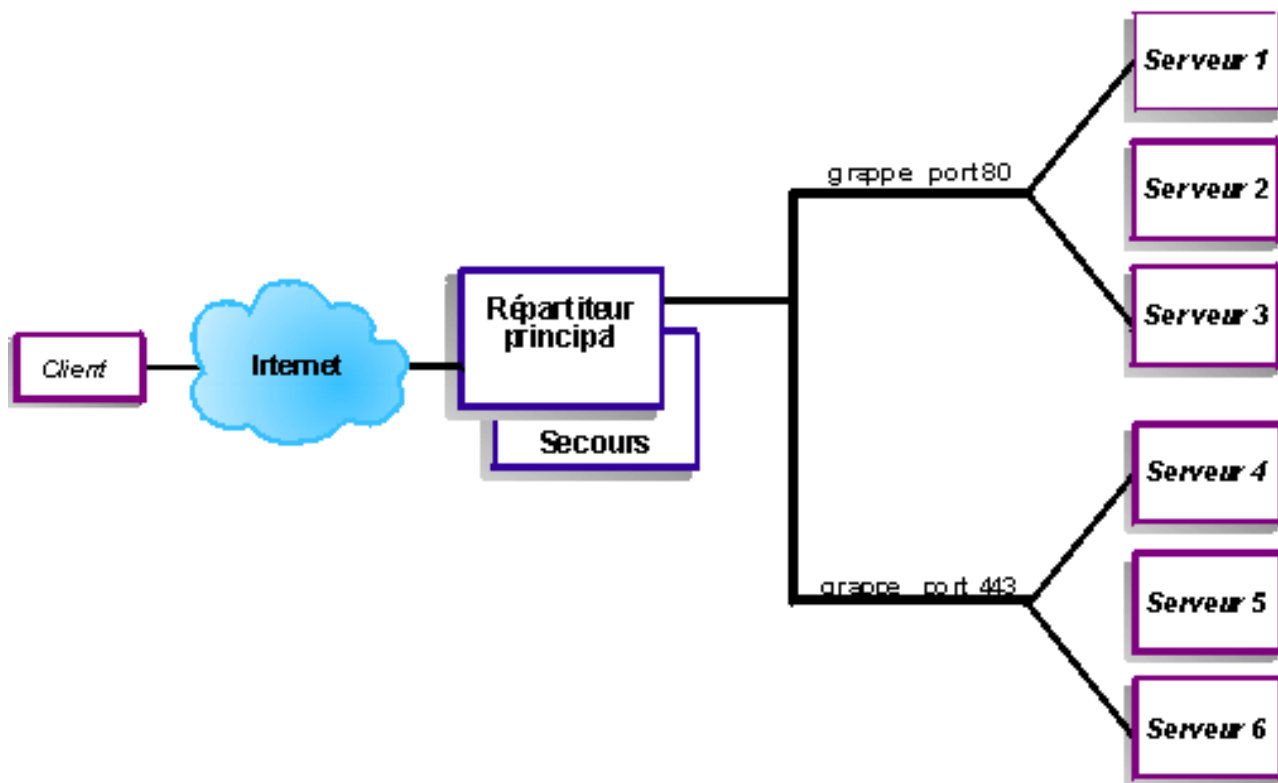


Figure 13. Exemple de Dispatcher utilisant la haute disponibilité

La fonctionnalité de haute disponibilité requiert une deuxième machine. La première se charge de l'équilibrage de charge pour la totalité du trafic client, comme dans une configuration à une seule machine. La seconde machine surveille le bon fonctionnement de la première et reprend l'équilibrage de charge si elle détecte un échec de la première machine.

Chacune des deux machines se voit affecter un rôle spécifique, *principal* ou de *sauvegarde*. La machine principale envoie régulièrement les données de connexion à la machine de secours. Pendant que la machine principale est *active* (équilibrage de charge), la machine de sauvegarde est en état d'*attente* et ses données s'actualisent en permanence, ce qui lui permet de prendre le relais des opérations en cas de besoin.

Les sessions de communication entre les deux machines sont désignées par le terme *signal de présence*. Ces signaux permettent à chaque machine de contrôler l'état de l'autre.

Si la machine de sauvegarde détecte que la machine principale est défaillante, elle prend en charge l'équilibrage de charge. A cette étape, les *états* respectifs des deux machines s'inversent : la machine de secours devient *active* et la machine principale passe en *attente*.

Dans la configuration à haute disponibilité, les deux machines doivent se trouver sur le même sous-réseau et leur configuration doit être identique.

Pour plus d'informations sur la fonction de haute disponibilité, voir «Haute disponibilité», à la page 204.

Haute disponibilité réciproque

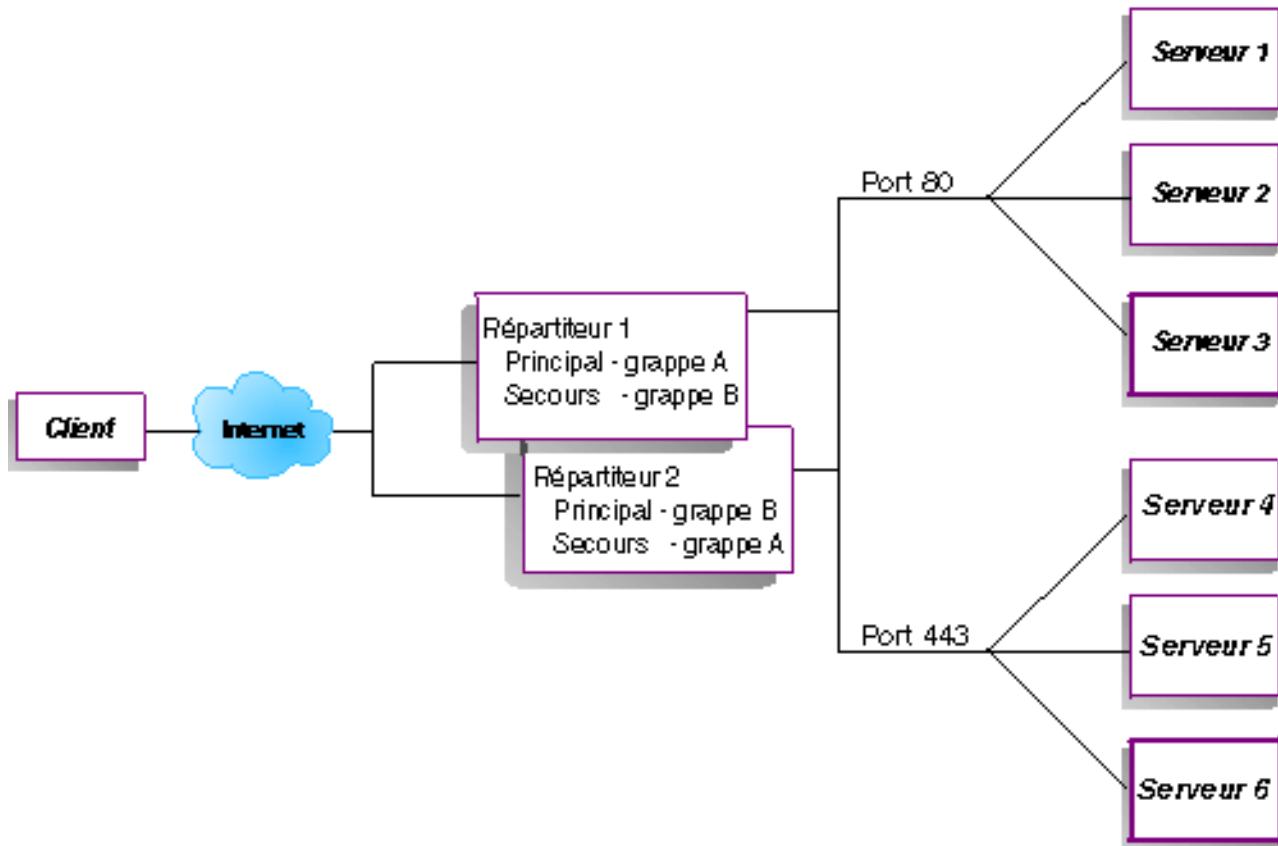


Figure 14. Exemple de Dispatcher utilisant la haute disponibilité réciproque

La fonctionnalité à haute disponibilité réciproque implique l'utilisation de deux machines. Les deux machines effectuent l'équilibrage de la charge du trafic client de manière active et assurent réciproquement la sauvegarde l'une de l'autre. Dans une configuration à haute disponibilité, une seule machine effectue l'équilibrage de charge. Dans une configuration à haute disponibilité réciproque, les deux machines assument l'équilibrage de charge d'une partie du trafic du client.

Pour la haute disponibilité réciproque, le trafic client est affecté à chaque machine sur la base d'une adresse de cluster. Chaque cluster peut être configuré avec l'adresse de non-acheminement (NFA) de sa machine principale. La machine du Dispatcher principal effectue normalement l'équilibrage de charge pour ce cluster. En cas de panne, l'autre machine assume l'équilibrage de charge pour son propre cluster et pour celui du Dispatcher qui est en panne.

La figure 14 illustre une configuration de haute disponibilité réciproque avec "cluster partagé A" et "cluster partagé B". Chaque répartiteur peut acheminer activement des paquets pour son cluster *principal*. Si l'un des répartiteurs venait à

échouer et ne pouvait plus activement acheminer les paquets pour son cluster principal, l'autre répartiteur pourrait le remplacer et acheminerait les paquets pour son cluster de *sauvegarde*.

Remarque : Les deux machines doivent configurer de la même façon leur ensembles de clusters partagés. C'est-à-dire que les ports utilisés et les serveurs définis pour chaque port doivent être identiques dans les deux configurations.

Pour plus d'informations sur la fonction de haute disponibilité, voir «Haute disponibilité», à la page 204.

Chapitre 7. Configuration de Dispatcher

Avant d'effectuer les opérations décrites dans le présent chapitre, voir Chapitre 6, «Planification de Dispatcher», à la page 51. Ce chapitre décrit comment créer une configuration de base pour le composant Dispatcher de Load Balancer.

- Voir Chapitre 8, «Déploiement de Dispatcher sur Load Balancer pour IPv4 et IPv6», à la page 81 si vous utilisez Load Balancer pour IPv4 et IPv6.
- Voir Chapitre 21, «Gestionnaire, conseillers et système Metric Server (des composants Dispatcher, CBR et Site Selector)», à la page 179 et Chapitre 22, «Fonctions avancées de Dispatcher, CBR et Site Selector», à la page 201 pour plus d'informations sur des configurations plus complexes de Load Balancer.
- Voir Chapitre 24, «Exploitation et gestion de Load Balancer», à la page 261, pour obtenir des informations sur l'administration authentifiée à distance, les fichiers journaux Load Balancer et sur l'utilisation des composants Load Balancer.

Remarque : Dans les versions antérieures où le produit se nommait Network Dispatcher, la commande de contrôle de Dispatcher était `ndcontrol`. Elle s'intitule désormais `dscontrol`.

Présentation générale des tâches de configuration

IMPORTANT : Si vous utilisez Load Balancer pour IPv4 et IPv6, voir Chapitre 8, «Déploiement de Dispatcher sur Load Balancer pour IPv4 et IPv6», à la page 81.

Avant de suivre les étapes de configuration détaillées dans ce tableau, assurez-vous que la machine Dispatcher et toutes les machines serveurs sont connectées au réseau, ont des adresses IP valides et peuvent communiquer entre elles par la triangulation ping.

Tableau 4. Tâches de configuration pour la fonction Dispatcher

Tâche	Description	Informations connexes
Configuration de la machine Dispatcher.	Définition de la configuration pour l'équilibrage de charge	«Configuration de la machine Dispatcher», à la page 66
Configuration des machines en vue de l'équilibrage de charge	Affectation d'un alias à l'unité de bouclage, recherche et suppression de la route supplémentaire	«Configuration des serveurs pour l'équilibrage de la charge», à la page 72

Méthodes de configuration

Quatre méthodes permettent une configuration de base de Dispatcher :

- Ligne de commande
- Scripts
- Interface graphique
- Assistant de configuration

Ligne de commande

C'est la méthode la plus directe pour la configuration de Dispatcher. Les valeurs des paramètres de commande doivent être saisies en anglais. Les seules exceptions

s'appliquent aux noms d'hôte (utilisés dans les commandes cluster, server et highavailability) et aux noms de fichiers (utilisés dans les commandes file).

Pour démarrer Dispatcher à partir de la ligne de commande, procédez comme suit :

1. Émettez la commande **dsserver** à partir de l'invite. Pour arrêter le service, tapez **dsserver stop**

Remarque : Pour les systèmes Windows, cliquez sur **Démarrer > Paramètres** (pour Windows 2000) > **Panneau de configuration > Outils d'administration > Services**. Cliquez à l'aide du bouton droit de la souris sur **IBM Dispatcher**, puis sélectionnez **Démarrer**. Pour arrêter le service, suivez la même procédure en sélectionnant **Arrêter**.

2. Ensuite, émettez les commandes de contrôle Dispatcher souhaitées pour définir votre configuration. Les procédures décrites dans ce manuel reposent sur l'utilisation de la ligne de commande. La commande est **dscontrol**. Pour plus de détails sur les commandes, voir Chapitre 27, «Guide des commandes Dispatcher et CBR», à la page 345.

Vous pouvez utiliser une version abrégée des paramètres de la commande dscontrol en entrant simplement la ou les quelques lettres d'identification des paramètres. Par exemple, pour obtenir l'aide correspondant à la commande file save, vous pouvez entrer **dscontrol he f** au lieu de **dscontrol help file**.

Pour démarrer l'interface de ligne de commande, entrez **dscontrol** pour ouvrir une invite dscontrol.

Pour fermer l'interface de ligne de commande, entrez **exit** ou **quit**.

Scripts

Vous pouvez entrer les commandes de configuration Dispatcher dans un fichier script de configuration pour les exécuter simultanément. Voir «Exemples de fichiers de configuration Load Balancer», à la page 481.

Remarque : Pour exécuter rapidement le contenu d'un fichier script (par exemple, mon_script), utilisez l'une des commandes suivantes :

- Pour mettre à jour la configuration actuelle, soumettez les commandes exécutables suivantes à partir du fichier script :
dscontrol file appendload mon_script
- Pour remplacer la configuration actuelle, soumettez les commandes exécutables suivantes à partir du fichier script :
dscontrol file newload mon_script

Pour sauvegarder la configuration en cours dans un fichier script (par exemple, savescript), exécutez la commande suivante :

dscontrol file save savescript

Cette commande enregistre le fichier script de configuration dans le répertoire **...ibm/edge/lb/servers/configurations/dispatcher**.

Interface graphique

Pour des instructions générales et un exemple de l'interface graphique, voir figure 41, à la page 470.

Pour démarrer l'interface graphique, procédez comme suit :

1. Vérifiez que dsserver est en cours d'exécution.
 - Pour les systèmes AIX, HP-UX, Linux ou Solaris, exécutez la commande suivante en tant que superutilisateur :
dsserver
 - Pour les systèmes Windows, dsserver est exécuté en tant que service qui démarre automatiquement.
2. Exécutez l'une des actions suivantes, selon votre système d'exploitation :
 - Pour les systèmes AIX, HP-UX, Linux ou Solaris : entrez **lbadmin**
 - Pour les systèmes Windows : cliquez sur **Démarrer > Programmes > IBM WebSphere > Edge Components > IBM Load Balancer > Load Balancer**

Pour configurer le composant Dispatcher à partir de l'interface graphique, vous devez d'abord sélectionner **Dispatcher** dans l'arborescence. Vous pouvez lancer l'exécuteur et le gestionnaire une fois que vous vous êtes connecté à un hôte. Vous pouvez également créer des clusters contenant des ports et des serveurs, puis lancer des conseillers pour le gestionnaire.

Vous pouvez utiliser l'interface graphique pour toute opération normalement exécutée par la commande **dscontrol**. Par exemple, pour définir un cluster à l'aide de la ligne de commande, vous devez entrer la commande **dscontrol cluster add cluster**. Pour définir un cluster à partir de l'interface graphique, cliquez sur Exécuteur à l'aide du bouton droit de la souris, puis dans le menu en incrustation qui apparaît, cliquez sur le bouton **Ajout d'un cluster** à l'aide du bouton gauche de la souris. Entrez l'adresse du cluster dans la fenêtre en incrustation, puis cliquez sur **OK**.

Les fichiers de configuration Dispatcher existants peuvent être chargés à l'aide des options **Chargement de la nouvelle configuration** (pour remplacer intégralement la configuration en cours) et **Ajout à la configuration en cours** (pour mettre à jour la configuration en cours) du menu en incrustation **Hôte**. Vous devez sauvegarder régulièrement votre configuration Dispatcher dans un fichier en utilisant l'option **Sauvegarder le fichier de configuration sous...** du menu en incrustation **Hôte**. Le menu **Fichier** situé en haut de l'interface graphique permet de sauvegarder les connexions à l'hôte en cours dans un fichier ou de restaurer les connexions dans des fichiers existants sur tous les composants Load Balancer.

Les commandes de configuration peuvent également être exécutées à distance. Pour plus de détails, voir «RMI (Remote Method Invocation)», à la page 262.

Pour exécuter une commande à partir de l'interface graphique : mettez le noeud Hôte en surbrillance dans l'arborescence de l'interface graphique, puis sélectionnez **Envoyer la commande...** dans le menu en incrustation Hôte. Dans la zone d'entrée de commande, entrez la commande à exécuter, par exemple **executor report**. Les résultats et l'historique des commandes sont exécutés lors de la session courante et s'affichent dans la fenêtre ouverte.

Vous pouvez accéder à l'**Aide** en cliquant sur le point d'interrogation situé dans l'angle supérieur droit de la fenêtre Load Balancer.

- **Aide sur les zones** — décrit les valeurs par défaut de chaque zone.
- **Procédures** — affiche la liste des tâches pouvant être effectuées dans cet écran.
- **Centre de documentation** — fournit un accès centralisé aux informations relatives au produit

Pour plus de détails sur l'utilisation de l'interface graphique, voir Annexe A, «Interface graphique utilisateur : Instructions générales», à la page 469.

Configuration à l'aide de l'assistant de configuration

Si vous utilisez l'assistant de configuration, suivez la procédure ci-dessous.

1. Démarrez dsserver sur Dispatcher :
 - Pour les systèmes AIX, HP-UX, Linux ou Solaris, exécutez la commande suivante en tant que superutilisateur :
dsserver
 - Pour les systèmes Windows, dsserver est exécuté en tant que service qui démarre automatiquement.
2. Démarrez la fonction Assistant de Dispatcher, **dswizard**.

Cet assistant vous guide dans les étapes requises pour la création d'une configuration de base pour le composant Dispatcher. Vous devez répondre à quelques questions concernant votre réseau. Vous serez guidé dans la configuration d'un cluster pour permettre à Dispatcher d'équilibrer le trafic dans un groupe de serveurs.

Configuration de la machine Dispatcher

La configuration de la machine Dispatcher ne peut être effectuée que par le superutilisateur root (pour les systèmes AIX, HP-UX, Linux ou Solaris) ou l'administrateur (pour les systèmes Windows).

Sur toutes les plateformes prises en charge, Load Balancer peut avoir un serveur **co-implanté**. La co-implantation implique que Load Balancer peut être implanté physiquement sur le serveur dont il assure l'équilibrage de charge.

Pour la machine Dispatcher, lorsque vous utilisez la méthode d'acheminement mac, vous avez besoin d'au moins deux adresses IP valides. Pour la méthode d'acheminement cbr ou nat, vous avez besoin d'au moins trois adresses IP valides :

- Une adresse IP spécifiquement associée à la machine Dispatcher.
Cette adresse constitue l'adresse IP principale de la machine Dispatcher et est appelée l'adresse de non-réacheminement (NFA). Il s'agit par défaut de l'adresse renvoyée par la commande **hostname**. Utilisez cette adresse pour vous connecter à la machine en vue de tâches administratives, telles que la configuration à distance avec Telnet ou l'accès au sous-agent SNMP. Si la machine Dispatcher peut déjà renvoyer des demandes vers d'autres machines du réseau (par la technique de la triangulation ping), il n'y a rien de plus à faire pour définir l'adresse de non-réacheminement.
- Une adresse IP par cluster.
Une adresse de cluster est une adresse associée à un nom de système hôte (par exemple **www.société_X.com**). Cette adresse IP est utilisée par un client pour se connecter aux serveurs du cluster en question. Dispatcher assure l'équilibrage de charge pour cette adresse.

- Pour la méthode d'acheminement cbr ou nat, une adresse IP pour l'adresse de retour

Dispatcher utilise l'adresse de retour comme adresse source lors de l'équilibrage de charge de la demande du client sur le serveur. Elle permet de garantir que le serveur renvoie le paquet à la machine Dispatcher, au lieu de l'envoyer directement au client. (Dispatcher transmettra ensuite le paquet IP au client.) Vous devez indiquer la valeur d'adresse de retour lors de l'ajout du serveur. Vous ne pouvez pas modifier l'adresse de retour sauf si vous supprimez le serveur et que vous l'ajoutez à nouveau.

Systemes Solaris uniquement :

- Par défaut, Dispatcher est configuré pour assurer l'équilibrage de charge avec des cartes d'interface réseau Ethernet 100 Mbps. La carte Ethernet 100 Mbps par défaut est désignée dans le fichier `ibmlb.conf` par `eri`. Toutefois, la prise en charge d'autres cartes d'interface (`le`, `ce`, `ge`, `hme`, `eri`, `bge`, `vge`, `qfe`, `dfme`, `fjgi` et `fjge`) est également assurée.

Par exemple, pour modifier le paramètre par défaut, éditez le fichier `/opt/ibm/edge/lb/servers/ibmlb.conf` comme suit :

- Pour utiliser une carte Ethernet 10 Mbps, remplacez `eri` par `le`.
- Pour utiliser une carte Ethernet 1 Go/s, remplacez `eri` par `ge`.
- Pour utiliser une carte multi-port, remplacez `eri` par `qfe`.

Pour prendre en charge plusieurs types de cartes, dupliquez la ligne du fichier `ibmlb.conf`, puis modifiez chaque ligne en fonction du type de carte dont vous disposez.

Par exemple, pour utiliser deux cartes Ethernet 100 Mbps, le fichier `ibmlb.conf` doit comporter une seule ligne indiquant l'unité `eri`.

Pour utiliser une carte Ethernet 10 Mbps et une carte Ethernet 100 Mbps, le fichier `ibmlb.conf` doit comporter deux lignes, l'une indiquant l'unité `le` et l'autre l'unité `eri`.

Remarque : Le fichier `ibmlb.conf` fournit des données à la commande Solaris `autopush` et doit être compatible avec cette dernière.

- Pour déterminer le type d'interface réseau Ethernet installé sur votre machine, lancez la commande suivante à partir de l'invite de commande Solaris :

```
ifconfig -a
```

Si vous obtenez le résultat suivant :

```
lo0: flags=2001000849<UP,LOOPBACK,RUNNING,MULTICAST,IPv4,VIRTUAL>
      mtu 8232 index 1 inet 127.0.0.1 netmask ffffffff
eri0: flags=1000843<UP,BROADCAST,RUNNING,MULTICAST,IPv4>
      mtu 1500 index 2 inet 9.42.93.208
      netmask fffffffc00 broadcast 9.42.95.255 ether 0:3:ba:2d:24:45
```

modifiez le fichier `ibmlb.conf` comme suit :

```
eri -1 0 ibmlb
```

- Le démarrage ou l'arrêt de Dispatcher Executor déconfigure tous les alias sur les cartes répertoriées dans le fichier `ibmlb.conf`. Pour reconfigurer automatiquement les alias sur ces cartes (à l'exception de ceux devant être employés par le composant Dispatcher de Load Balancer), utilisez le fichier script `goAliases`. Un exemple de script, qui se trouve dans le répertoire `...ibm/edge/lb/servers/samples`, doit être déplacé dans le répertoire `...ibm/edge/lb/servers/bin` avant exécution. Le script `goAliases` s'exécute automatiquement lors du démarrage ou de l'arrêt de Dispatcher Executor.

Par exemple, si les clusters X et Y sont configurés pour être utilisés par le composant CBR sur les cartes répertoriées dans le fichier `ibmlb.conf`, ils sont déconfigurés lors du lancement des commandes **dscontrol executor start** ou **dscontrol executor stop**. Ce résultat n'est peut-être pas souhaité. Lorsque les clusters X et Y sont configurés dans le script `goAliases`, ils sont automatiquement reconfigurés une fois Dispatcher Executor lancé ou arrêté.

Assurez-vous que la transmission Internet n'est pas activée pour le protocole TCP/IP.

La figure 15 montre un exemple de Dispatcher configuré avec une seule cluster, deux ports et trois serveurs.

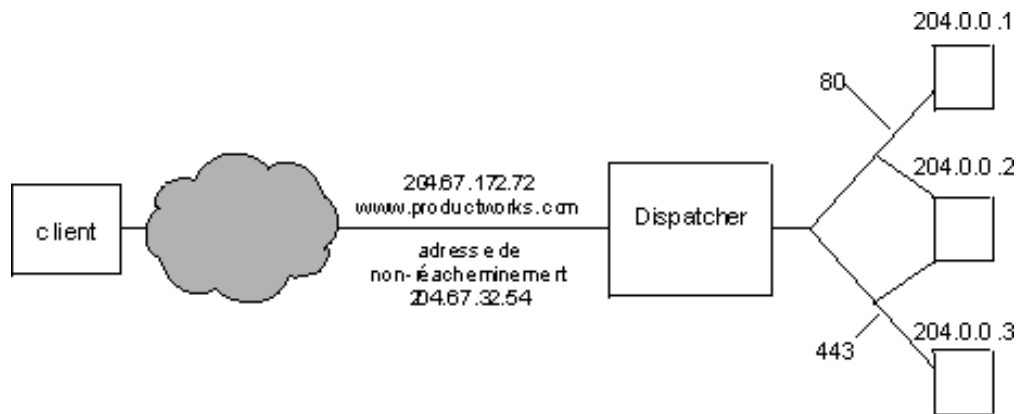


Figure 15. Exemple d'adresses IP nécessaires pour la machine Dispatcher

Pour obtenir une aide sur les commandes utilisées lors de cette procédure, voir Chapitre 27, «Guide des commandes Dispatcher et CBR», à la page 345.

Pour plus d'informations sur le fichier de configuration type, voir «Exemples de fichiers de configuration Load Balancer», à la page 481.

Etape 1. Démarrage de la fonction serveur

Systèmes AIX, HP-UX, Linux ou Solaris : Pour démarrer la fonction serveur, entrez **dsserver**.

Systèmes Windows : La fonction serveur démarre automatiquement en tant que service.

Remarque : Un fichier de configuration par défaut (`default.cfg`) est chargé automatiquement lors du démarrage de `dsserver`. Si l'utilisateur décide de sauvegarder la configuration Dispatcher dans `default.cfg`, toutes les données sauvegardées dans ce fichier sont chargées automatiquement au prochain démarrage de `dsserver`.

Etape 2. Démarrage de la fonction exécuteur

Pour démarrer la fonction exécuteur, tapez la commande **dscontrol executor start**. Notez que vous pouvez également modifier divers paramètres de l'exécuteur à cette occasion. Voir Chapitre 27, «Guide des commandes Dispatcher et CBR», à la page 345.

Etape 3. Définition de l'adresse de non-réacheminement (si différente du nom d'hôte)

Utilisez cette adresse pour vous connecter à la machine en vue de tâches administratives, comme l'utilisation de Telnet ou SMTP, par exemple. Par défaut, cette adresse correspond au nom d'hôte.

Pour définir l'adresse de non-réacheminement, entrez la commande **dscontrol executor set nfa** *adresse_IP* ou éditez le fichier de configuration type. *adresse_IP* peut être le nom symbolique ou l'adresse IP.

Etape 4. Définition et configuration des options du cluster

Dispatcher équilibrera les demandes envoyées à l'adresse du cluster entre les serveurs configurés sur les ports associés à ce cluster.

Le cluster est soit un nom symbolique, soit l'adresse en notation décimale à point, soit l'adresse spéciale 0.0.0.0 qui définit un cluster générique. Pour définir un cluster, tapez la commande **dscontrol cluster add**. Pour définir les options de cluster, tapez la commande **dscontrol cluster set** ou utilisez l'interface graphique pour lancer des commandes. Les clusters génériques peuvent être utilisés pour remplacer plusieurs adresses IP afin de permettre l'équilibrage de charge pour les paquets entrants. Pour plus de détails, voir «Utilisation d'un cluster générique pour combiner les configurations serveurs», à la page 236, «Utilisation du cluster générique pour équilibrer la charge des pare-feux», à la page 237 et «Utilisation de cluster générique avec Caching Proxy pour le proxy transparent», à la page 238.

Etape 5. Affectation d'un alias à la carte d'interface réseau

Lorsque le cluster est défini, vous devez normalement configurer son adresse sur l'une des cartes d'interface réseau de la machine Dispatcher. Pour ce faire, émettez la commande **dscontrol executor configure** *adresse_cluster*. Cette commande recherche une carte avec une adresse existante et appartenant au même sous-réseau que l'adresse du cluster. La commande de configuration de la carte système est ensuite lancée pour l'adresse du cluster en utilisant la carte trouvée et le masque de réseau de l'adresse existante figurant sur cette carte. Par exemple :

```
dscontrol executor configure 204.67.172.72
```

Vous pouvez configurer des adresses de clusters ajoutées à un serveur en attente en mode haute disponibilité ou des adresses de clusters ajoutées à un répartiteur de réseau étendu jouant le rôle de serveur éloigné. Il est également inutile d'exécuter la commande de configuration de l'exécuteur si vous utilisez le modèle de script **goldle**, en mode autonome. Pour plus d'informations sur le script goldle, voir «Utilisation de scripts», à la page 209.

Dans de rares cas, vous pouvez avoir une adresse qui ne correspond pas à une adresse de sous-réseau existante. Vous devez alors utiliser l'autre forme de la commande de configuration de l'exécuteur et fournir de manière explicite le nom et le masque de réseau de l'interface. Entrez la commande **dscontrol executor configure***adresse_cluster nom_interface sous-masque*.

Exemples :

```
dscontrol executor configure 204.67.172.72 en0 255.255.0.0
(systemes AIX)
dscontrol executor configure 204.67.172.72 eth0:1 255.255.0.0
(systemes Linux)
```

```
dscontrol executor configure 204.67.172.72 eri0 255.255.0.0
(systemes Solaris)
dscontrol executor configure 204.67.172.72 en1 255.255.0.0
(systemes Windows)
```

Systèmes Windows

Pour vous servir de l'autre forme de la commande de configuration de l'exécuteur sous Windows, vous devez déterminer le nom de l'interface à utiliser. Si votre machine comporte une seule carte Ethernet, l'interface porte le nom en0. Si vous ne disposez que d'une seule carte en anneau à jeton (Token Ring), l'interface porte le nom tr0. Si la machine comporte plusieurs cartes de l'un ou l'autre type, il est nécessaire de déterminer le mappage des cartes. Procédez comme suit :

1. A partir de la ligne de commande, lancez l'exécuteur : `dscontrol executor start`
2. Exécutez la commande : `dscontrol executor xm 1`

Le résultat apparaît à l'écran. Pour connaître le nom de l'interface à utiliser pour la configuration Load Balancer, recherchez l'adresse IP de votre machine Load Balancer dans les lignes suivant `Number of NIC records`.

L'adresse IP de votre machine Load Balancer apparaît sous la forme : `ia->adr_ai`. Le nom d'interface associé apparaît sous la forme : `ifp->nom_if`.

les noms d'interface attribués par la commande `executor configure` correspondent aux noms d'interface listés dans cette commande.

Après avoir accédé à ces informations de mappage, vous pouvez créer un alias reliant l'interface réseau à l'adresse du cluster.

Utilisation des commandes `ifconfig` pour configurer des alias de cluster

Sous Linux ou UNIX, la commande de configuration de l'exécuteur exécute des commandes `ifconfig`.

Systèmes Solaris et HP-UX : Lorsque vous utilisez des applications serveur de liaison, qui opèrent une liaison à une liste d'adresses IP ne contenant pas celle du serveur, faites appel à la commande `arp publish` plutôt qu'à `ifconfig` pour définir dynamiquement une adresse IP sur la machine Load Balancer. Par exemple :

```
arp -s <cluster> <adresse MAC Load Balancer> pub
```

Etape 6. Définition des ports et de leurs options

Pour définir un port, entrez la commande `dscontrol port add cluster:port`, éditez le fichier de configuration `type` ou utilisez l'interface graphique. La valeur de `cluster` peut être le nom symbolique ou l'adresse IP. `Port` représente le numéro du port utilisé pour ce protocole. A ce stade, vous avez également la possibilité de modifier divers paramètres de ports. Vous devez définir et configurer tous les serveurs pour un port. Voir Chapitre 27, «Guide des commandes Dispatcher et CBR», à la page 345.

Le numéro de port 0 (zéro) est utilisé pour spécifier un port générique. Ce port acceptera le trafic vers un port non défini sur le cluster. Le port générique est utilisé pour configurer des règles et des serveurs pour n'importe quel port. Vous pouvez également utiliser cette fonction en cas de configuration de serveur et de règle identique pour plusieurs ports. Le trafic sur un port peut influencer les décisions d'équilibrage de charge pour le trafic sur les autres ports. Pour plus de

détails sur les cas d'utilisation d'un port générique, voir «Utilisation du port générique pour acheminer le trafic destiné à un port non configuré», à la page 238.

Etape 7. Définition des serveurs avec équilibrage de charge

Pour définir un serveur avec équilibrage de charge, entrez la commande **dscontrol server add cluster:port:serveur**, éditez le fichier de configuration type ou utilisez l'interface graphique. *cluster* et *serveur* peuvent correspondre à des noms symboliques ou à des adresses IP. *Port* représente le numéro du port utilisé pour ce protocole. Pour effectuer l'équilibrage de charge, vous devez définir plusieurs serveurs sur le port d'un cluster.

Serveurs de liaison : Si le composant Dispatcher équilibre la charge entre des serveurs de liaison, les serveurs *doivent* être configurés pour effectuer la liaison avec l'adresse du cluster. Etant donné que Dispatcher réachemine les paquets sans modifier l'adresse IP de destination, lorsque ceux-ci arrivent, l'adresse de cluster qu'ils contiennent indique la destination. Si un serveur a été configuré pour être lié à une adresse IP autre que l'adresse de cluster, il ne pourra pas accepter les demandes destinées au cluster.

Pour savoir s'il s'agit d'un serveur de liaison, lancez la commande **netstat -an** et recherchez *serveur:port*. S'il ne s'agit pas d'un serveur de liaison, le résultat de la commande est 0.0.0.0:80. S'il s'agit d'un serveur de liaison, une adresse du type 192.168.15.103:80 apparaît.

Remarque : Pour les systèmes Solaris et Linux : Si vous utilisez des conseillers, les serveurs de liaison ne doivent pas être co-implantés.

Co-implantation d'adresses multiples : Dans une configuration de co-implantation, l'adresse du serveur co-implanté ne doit *pas* être la même que celle de non-réacheminement (NFA). Vous avez la possibilité d'utiliser une autre adresse si votre machine a été définie avec des adresses IP multiples. En ce qui concerne le composant Dispatcher, le serveur co-implanté doit être défini comme **co-implanté** via la commande **dscontrol server**. Pour plus d'informations sur les serveurs co-implantés, voir «Utilisation de serveurs implantés au même endroit», à la page 202.

Pour plus d'informations sur la syntaxe de la commande **dscontrol server**, voir «dscontrol server — Configuration des serveurs», à la page 390.

Etape 8. Démarrage de la fonction gestionnaire (facultatif)

La fonction gestionnaire permet d'améliorer l'équilibrage de charge. Pour démarrer le gestionnaire, entrez la commande **dscontrol manager start**, éditez le fichier de configuration type ou utilisez l'interface graphique.

Etape 9. Démarrage de la fonction conseiller (facultatif)

Les conseillers transmettent au gestionnaire des informations complémentaires sur la capacité à répondre aux demandes des serveurs ayant fait l'objet d'un équilibrage de charge. Chaque conseiller est spécifique à un protocole. Par exemple, tapez la commande suivante pour lancer le conseiller HTTP :

```
dscontrol advisor start http port
```

Pour consulter la liste des conseillers et des ports par défaut correspondants, voir Chapitre 27, «Guide des commandes Dispatcher et CBR», à la page 345. Pour lire la description de chaque conseiller, voir «Liste des conseillers», à la page 188.

Etape 10. Définition du niveau d'importance des informations requis pour le cluster

Si vous démarrez des conseillers, vous pouvez modifier le niveau d'importance donné aux informations des conseillers entrant dans les décisions d'équilibrage de la charge. Pour définir les proportions du cluster, entrez la commande **dscontrol cluster set cluster proportions**. Pour plus d'informations, voir «Proportion de l'importance accordée aux données d'état», à la page 180.

Configuration des serveurs pour l'équilibrage de la charge

Exécutez ces procédures si l'une des conditions ci-dessous est remplie :

- Si vous utilisez la méthode d'acheminement MAC et le serveur est un serveur dorsal.
- Vous utilisez la méthode d'acheminement MAC et le serveur est co-implanté et configuré en tant que système haute disponibilité de secours.

Remarques :

1. Les procédures, telles que la suppression de l'affectation d'alias à l'unité de bouclage, doivent être placées dans les scripts go* au cas où le système passe à l'état actif.
2. S'il s'agit de la configuration du système actif de haute disponibilité, les procédures, telles que l'affectation d'un alias à l'unité de bouclage, doivent être placées dans les scripts go* au cas où la machine passerait à l'état de serveur de secours.)

Si vous utilisez la méthode de réacheminement MAC, Dispatcher équilibrera la charge uniquement entre des serveurs qui permettent de configurer l'unité de bouclage avec une adresse IP supplémentaire. C'est pourquoi le serveur dorsal ne répondra jamais aux demandes ARP (protocole de résolution d'adresses). Suivez les étapes indiquées dans cette section pour configurer les serveurs avec équilibrage de charge.

Etape 1. Affectation d'un alias pour l'unité de bouclage

Pour que les serveurs bénéficiant d'un équilibrage de charge fonctionnent, vous devez définir (ou de préférence affecter un alias à) l'unité de bouclage (souvent appelé lo0) en fonction de l'adresse de cluster. Si vous utilisez la méthode d'acheminement MAC, le composant Dispatcher ne modifie pas l'adresse IP de destination dans le paquet TCP/IP avant de retransmettre ce paquet au serveur TCP. Si l'unité de bouclage est définie, ou se voit affecter l'adresse de cluster comme alias, les serveurs avec équilibrage de charge accepteront les paquets envoyés à cette adresse de cluster.

Si votre système d'exploitation supporte l'attribution d'alias aux interfaces réseau (comme par exemple les systèmes AIX, HP-UX, Linux, Solaris ou Windows), vous devez affecter l'adresse de cluster comme alias à l'unité de bouclage. L'utilisation d'un système d'exploitation prenant en charge les alias à pour avantage de permettre la configuration de serveurs avec équilibrage de charge desservant plusieurs adresses de cluster.

IMPORTANT : Pour les systèmes Linux, voir «Solutions alternatives pour l'affectation d'alias à l'unité de bouclage sous Linux lors de l'utilisation de la méthode d'acheminement MAC de Load Balancer», à la page 78.

Si le système d'exploitation de votre serveur ne supporte pas les alias, vous devez définir l'adresse de cluster comme alias pour l'unité de bouclage.

Pour définir l'unité de bouclage ou lui affecter un alias, utilisez la commande requise par votre système d'exploitation comme indiqué dans le tableau 5.

Tableau 5. Commandes pour l'affectation d'un alias à l'unité de bouclage (lo0) pour Dispatcher

AIX 4.3 ou version antérieure	ifconfig lo0 alias <i>adresse_cluster</i> netmask <i>masque_réseau</i> Remarque : Utilisez le masque de réseau de l'adaptateur principal
AIX 5.x	ifconfig lo0 alias <i>adresse_cluster</i> netmask 255.255.255.255
HP-UX	ifconfig lo0:1 <i>adresse_cluster</i> up
Linux	Choisissez l'une des commandes suivantes : <ul style="list-style-type: none"> • ip -4 addr add <i>adresse_cluster/32 dev lo</i> • ifconfig lo:1 <i>adresse_cluster</i> netmask 255.255.255.255 up <p>IMPORTANT : Une fois que vous émettez une des commandes de configuration sur votre machine, utilisez-la systématiquement (ip ou ifconfig), sinon des résultats imprévisibles risquent de se produire.</p>
OS/2	ifconfig lo <i>adresse_cluster</i>
OS/390	Configuration d'un alias de bouclage sur le système OS/390 <ul style="list-style-type: none"> • L'administrateur doit créer une entrée dans la liste d'adresses d'origine du membre (fichier) de paramètres IP. Par exemple <pre>HOME ;Address Link 192.168.252.11 tr0 192.168.100.100 1tr1 192.168.252.12 loopback</pre> • Plusieurs adresses peuvent être définies pour l'unité de bouclage. • L'adresse de bouclage 127.0.0.1 est configurée par défaut.
Solaris 7	ifconfig lo0:1 <i>adresse_cluster</i> 127.0.0.1 up
Solaris 8, Solaris 9 et Solaris 10	ifconfig lo0:1 plumb <i>adresse_cluster</i> netmask <i>masque_réseau</i> up

Tableau 5. Commandes pour l'affectation d'un alias à l'unité de bouclage (lo0) pour Dispatcher (suite)

Windows Server 2003	<ol style="list-style-type: none"> 1. Cliquez sur Démarrer, puis sur Panneau de configuration. 2. Si vous ne l'avez pas encore fait, ajoutez le pilote de la carte de bouclage MS. <ol style="list-style-type: none"> a. Cliquez sur Ajout de matériel. Cela lance l'assistant correspondant à cette fonction. b. Cliquez sur Suivant c. Sélectionnez Oui, j'ai déjà connecté le matériel, puis cliquez sur Suivant. d. Si la carte de bouclage MS figure dans la liste, c'est qu'elle est déjà installée. Cliquez sur Annuler pour fermer le panneau. e. Si la carte de bouclage MS <i>ne figure pas</i> dans la liste, sélectionnez Ajouter un nouveau périphérique et cliquez sur Suivant. f. Pour sélectionner le composant matériel dans une liste, dans le panneau Trouver le nouveau matériel, cliquez sur Non, puis sur Suivant. g. Sélectionnez Cartes réseau et cliquez sur suivant. h. Dans le panneau Sélectionnez la carte réseau, sélectionnez Microsoft dans la liste des fabricants, puis Microsoft Loopback Adapter. i. Cliquez sur Suivant une première fois, puis une deuxième pour installer les paramètres par défaut (ou sélectionnez l'option de support fourni (Have Disk), puis insérez le CD-ROM et effectuez l'installation à partir de ce point). j. Cliquez sur Terminer pour achever l'installation. 3. Dans le Panneau de configuration, cliquez deux fois sur Connexions réseau et accès à distance. 4. Sélectionnez la connexion portant le nom d'unité "Microsoft Loopback Adapter". 5. Sélectionnez Propriétés dans le menu déroulant. 6. Sélectionnez Internet Protocol (TCP/IP), puis cliquez sur Propriétés. 7. Cliquez sur Utiliser l'adresse IP suivante. Pour <i>Adresse IP</i> indiquez l'adresse du cluster et pour <i>Masque de sous-réseau</i>, le masque de sous réseau du serveur dorsal. <p>Remarque : N'indiquez pas d'adresse de routeur. Utilisez le système hôte local comme serveur DNS par défaut.</p>
---------------------	--

Tableau 5. Commandes pour l'affectation d'un alias à l'unité de bouclage (lo0) pour Dispatcher (suite)

Windows 2000	<ol style="list-style-type: none"> 1. Sélectionnez Démarrer, Paramètres, puis Panneau de configuration. 2. Si vous ne l'avez pas encore fait, ajoutez le pilote de la carte de bouclage MS. <ol style="list-style-type: none"> a. Cliquez deux fois sur Ajout/Suppression de matériel. Cela lance l'assistant correspondant à cette fonction. b. Cliquez sur Suivant, sélectionnez Ajouter/Dépanner un périphérique, puis sur Suivant. c. Le panneau Sélection d'un périphérique matériel s'affiche. d. Si la carte de bouclage MS figure dans la liste, c'est qu'elle est déjà installée. Cliquez sur Annuler pour fermer le panneau. e. Si la carte de bouclage MS <i>ne figure pas</i> dans la liste, sélectionnez Ajouter un nouveau périphérique et cliquez sur Suivant. f. Pour sélectionner le composant matériel dans une liste, dans le panneau Trouver le nouveau matériel, cliquez sur Non, puis sur Suivant. g. Sélectionnez Cartes réseau et cliquez sur suivant. h. Dans le panneau Sélectionnez la carte réseau, sélectionnez Microsoft dans la liste des fabricants, puis Microsoft Loopback Adapter. i. Cliquez sur Suivant une première fois, puis une deuxième pour installer les paramètres par défaut (ou sélectionnez l'option de support fourni (Have Disk), puis insérez le CD-ROM et effectuez l'installation à partir de ce point). j. Cliquez sur Terminer pour achever l'installation. 3. Dans le Panneau de configuration, cliquez deux fois sur Connexions réseau et accès à distance. 4. Cliquez à l'aide du bouton droit de la souris sur la connexion portant le nom d'unité "Microsoft Loopback Adapter" pour la sélectionner. 5. Sélectionnez Propriétés dans le menu déroulant. 6. Sélectionnez Internet Protocol (TCP/IP), puis cliquez sur Propriétés. 7. Cliquez sur Utiliser l'adresse IP suivante. Pour <i>Adresse IP</i> indiquez l'adresse du cluster et pour <i>Masque de sous-réseau</i> le masque de sous réseau par défaut (255.0.0.0). Remarque : N'indiquez pas d'adresse de routeur. Utilisez le système hôte local comme serveur DNS par défaut.
--------------	--

Tableau 5. Commandes pour l'affectation d'un alias à l'unité de bouclage (lo0) pour Dispatcher (suite)

Windows NT	<ol style="list-style-type: none"> 1. Cliquez sur Démarrer, puis sur Paramètres. 2. Cliquez sur Panneau de configuration, puis cliquez deux fois sur Réseau. 3. Si vous ne l'avez pas encore fait, ajoutez le pilote de la carte de bouclage MS. <ol style="list-style-type: none"> a. Dans la fenêtre Réseau, cliquez sur Adaptateurs. b. Sélectionnez Adaptateur de bouclage MS, puis cliquez sur OK. c. A l'invite, insérez le CD ou les disques d'installation. d. Dans la fenêtre Réseau, cliquez sur Protocoles. e. Sélectionnez Protocole TCP/IP, puis cliquez sur Propriétés. f. Sélectionnez Adaptateur de bouclage MS, puis cliquez sur OK. 4. Attribuez à l'adresse de bouclage votre adresse de cluster. Acceptez le masque de sous-réseau par défaut proposé (255.0.0.0) et n'entrez pas d'adresse de passerelle. <p>Remarque : Vous devrez peut-être quitter la fenêtre Paramètres réseau puis y revenir pour que le pilote de l'unité de bouclage MS s'affiche sous Configuration TCP/IP.</p>
------------	--

Etape 2. Vérification de l'existence d'une route supplémentaire

Sur certains systèmes d'exploitation, il se peut qu'une route par défaut ait été créée. Dans ce cas, elle doit être supprimée.

- Pour vérifier l'existence d'une route supplémentaire sur les systèmes d'exploitation Windows, utilisez la commande suivante :

```
route print
```

IMPORTANT : Sous Windows 2003, toute route supplémentaire doit être ignorée. En cas d'incidents avec le routage après l'établissement d'alias, supprimez l'alias, puis rajoutez-le à l'aide d'un autre masque de réseau.

- Pour vérifier l'existence d'une route supplémentaire sur tous les systèmes Linux et UNIX, utilisez la commande suivante :

```
netstat -nr
```

Exemple pour Windows :

1. Une fois la commande **route print** soumise, un tableau semblable à l'exemple suivant s'affiche. (Cet exemple illustre la recherche et la suppression d'une route supplémentaire vers le cluster 9.67.133.158, avec le masque de sous-réseau par défaut 255.0.0.0.)

Active Routes:

Network Address	Netmask	Gateway Address	Interface	Metric
0.0.0.0	0.0.0.0	9.67.128.1	9.67.133.67	1
9.0.0.0	255.0.0.0	9.67.133.158	9.67.133.158	1
9.67.128.0	255.255.248.0	9.67.133.67	9.67.133.67	1
9.67.133.67	255.255.255.255	127.0.0.1	127.0.0.1	1
9.67.133.158	255.255.255.255	127.0.0.1	127.0.0.1	1
9.255.255.255	255.255.255.255	9.67.133.67	9.67.133.67	1
127.0.0.0	255.0.0.0	127.0.0.1	127.0.0.1	1
224.0.0.0	224.0.0.0	9.67.133.158	9.67.133.158	1
224.0.0.0	224.0.0.0	9.67.133.67	9.67.133.67	1
255.255.255.255	255.255.255.255	9.67.133.67	9.67.133.67	1

2. L'adresse du cluster figure dans la colonne de l'adresse de passerelle. Si une route supplémentaire existe, l'adresse du cluster apparaîtra deux fois. Dans l'exemple, l'adresse du cluster (9.67.133.158) apparaît sur les lignes 2 et 8.
3. L'adresse du réseau figure sur toutes les lignes où apparaît l'adresse du cluster. Vous avez uniquement besoin de l'une de ces routes. La route en trop doit être supprimée. La route supplémentaire à supprimer est celle dont l'adresse de réseau commence par le premier chiffre de l'adresse du cluster, suivi de trois zéros. Dans l'exemple, la route supplémentaire à supprimer est celle qui se trouve sur la ligne 2, avec l'adresse de réseau **9.0.0.0** :

```
9.0.0.0    255.0.0.0    9.67.133.158    9.67.133.158    1
```

Etape 3. Suppression d'une route supplémentaire

Vous devez supprimer la route supplémentaire. Pour cela, utilisez la commande correspondant à votre système d'exploitation fournie dans le tableau 6.

Exemple : Pour supprimer la route supplémentaire comme indiqué pour l'exemple "Routes actives" de l'étape 2, entrez :

```
route delete 9.0.0.0 9.67.133.158
```

Tableau 6. Commandes de suppression d'une route supplémentaire pour Dispatcher

HP-UX	route delete <i>adresse_cluster</i> <i>adresse_cluster</i>
Windows	<p>route delete <i>adresse_réseau</i> (dans une invite MS-DOS)</p> <p>Remarque : Vous devez supprimer la route supplémentaire chaque fois que vous réamorcez le serveur.</p> <p>Sous Windows 2003, il n'est pas possible de supprimer des routes. Sous Windows 2003, toute route supplémentaire doit être ignorée. En cas d'incidents avec le routage après l'établissement d'alias, supprimez l'alias, puis rajoutez-le à l'aide d'un autre masque de réseau.</p>

A l'aide de l'exemple fourni dans la figure 15, à la page 68, et en configurant un serveur sous AIX, la commande serait :

```
route delete -net 204.0.0.0 204.67.172.72
```

Etape 4. Vérification de la configuration du serveur

Pour vérifier la configuration d'un serveur dorsal, effectuez les étapes suivantes à partir d'une autre machine du même sous-réseau lorsque Load Balancer n'est pas en cours d'exécution et la *cluster* non configurée.

1. Emettez la commande :

```
arp -d cluster
```

2. Emettez la commande :

```
ping cluster
```

La commande ping doit rester sans réponse. Si une réponse est renvoyée, assurez-vous que vous n'avez pas attribué l'adresse du cluster à l'interface à l'aide de la commande ifconfig. Vérifiez qu'aucune machine n'a une entrée ARP publiée pour l'adresse du cluster.

3. Soumettez une commande ping pour le serveur dorsal, puis émettez immédiatement la commande suivante :

```
arp -a
```

La sortie de la commande doit contenir l'adresse MAC de votre serveur.

Emettez la commande :

```
arp -s cluster adresse_mac_serveur
```

4. Soumettez une commande ping pour le cluster. Cette commande doit renvoyer une réponse. Soumettez une demande http, telnet ou d'un autre type, adressée au cluster que vous voulez voir géré par votre serveur dorsal. Vérifiez que le cluster fonctionne correctement.
5. Emettez la commande :

```
arp -d cluster
```
6. Soumettez une commande ping pour le cluster. Cette commande doit rester sans réponse.

Remarque : Si une réponse est renvoyée, émettez une instruction `arp cluster` pour obtenir l'adresse MAC de la machine incorrectement configurée. Répétez les étapes 1 à 6.

Solutions alternatives pour l'affectation d'alias à l'unité de bouclage sous Linux lors de l'utilisation de la méthode d'acheminement MAC de Load Balancer

Certaines versions de Linux émettent des réponses ARP pour toute adresse IP configurée sur la machine, quelle que soit l'interface installée. Il choisit également une adresse IP de source ARP pour les requêtes ARP who-has en se basant sur toutes les adresses IP définies sur la machine, quelle que soit l'interface sur laquelle ces adresses sont configurées. L'ensemble du trafic d'un cluster est dirigé indistinctement vers un seul serveur.

Si vous utilisez la méthode d'acheminement MAC de Dispatcher, un mécanisme doit être mis en oeuvre pour s'assurer que le trafic destiné au cluster peut être accepté par les piles des serveurs dorsaux, y compris la machine de secours haute disponibilité co-implantée, lorsque la haute disponibilité et la co-implantation sont utilisées conjointement.

Dans la plupart des cas, vous devez affecter l'adresse du cluster en tant qu'alias à l'unité de bouclage. Pour les serveurs dorsaux, le cluster doit être associé à un alias sur l'unité de bouclage. Si vous utilisez la haute disponibilité et la co-implantation, des clusters doivent être associés à un alias sur l'unité de bouclage pour les serveurs d'équilibrage de charge de secours.

Pour s'assurer que les systèmes Linux n'affichent pas les adresses dans l'unité de bouclage, vous devez les rendre compatibles avec l'acheminement MAC de Dispatcher.

1. Utilisez un noyau qui n'affiche pas les adresses. Cette option doit être privilégiée car elle ne requiert pas de temps système pour chaque paquet et ne nécessite pas une reconfiguration pour chaque noyau.
 - United Linux 1 / SLES8 avec SP2(x86) ou SP3 (toutes les autres architectures) et supérieure contient le correctif Julian de masquage ARP. Assurez-vous que ce correctif est toujours actif avant d'affecter l'adresse du cluster en tant qu'alias à l'aide de la commande :

```
# sysctl -w net.ipv4.conf.all.hidden=1 net.ipv4.conf.lo.hidden=1
```

Il est ensuite possible d'affecter des alias aux clusters selon la méthode normale, comme dans l'exemple suivant :

```
# ifconfig lo:1 $CLUSTER_ADDRESS netmask 255.255.255.255 up
```

- Utilisez la commande `arp_ignore` `sysctl` disponible dans les versions 2.4.25 et 2.6.5 et supérieure mais sachez qu'au cours des distributions, les fonctions peuvent faire l'objet d'un rétro-portage (backport). Vérifiez qu'elle est activée avant d'affecter des adresses de cluster en tant qu'alias à l'aide des commandes suivantes :

```
# sysctl -w net.ipv4.conf.all.arp_ignore=3
net.ipv4.conf.all.arp_announce=2
```

Pour affecter des alias aux clusters, utilisez la commande suivante :

```
# ip addr add $CLUSTER_ADDRESS/32 scope host dev lo
```

Une commande similaire doit se trouver dans les scripts `go*` pour les configurations de haute disponibilité et de co-implantation.

- Remarque : Lorsque vous utilisez `sysctl`, vérifiez que ces paramètres survivent au réamorçage en ajoutant les paramètres à `/etc/sysctl.conf`.
2. Utilisez des IP tables pour rediriger l'ensemble du trafic entrant du cluster vers l'hôte local. Si vous employez cette méthode, ne configurez pas l'unité de bouclage avec un alias. Utilisez plutôt la commande suivante :

```
# iptables -t nat -A PREROUTING -d $CLUSTER_ADDRESS -j REDIRECT
```

Les systèmes Linux effectuent alors une conversion NAT de la destination sur chaque paquet, en convertissant l'adresse de cluster en adresse d'interface. Cette méthode entraîne une baisse de débit d'environ 6,4 % en terme de nombre de connexions par seconde. Elle est compatible avec n'importe quelle distribution de stock prise en charge ; aucun module de noyau ou correctif+compilation+installation de noyau n'est requis.

3. Appliquez la version 1.2.0 ou supérieure du module `noarp`. La source du noyau doit être disponible et correctement configurée, et les outils de développement (`gcc`, `gnu make`, etc.) doivent être disponibles. Vous devez compiler et installer le module à chaque fois que le noyau est mis à niveau. Ce module est disponible à l'adresse <http://www.masarlabs.com/noarp/>. Etant donné que le code du noyau lui-même n'est pas modifié, cette solution est bien plus anodine que la solution n°4 (présentée plus loin) et moins sujette à erreur. Ce module doit également être configuré avant qu'une adresse de cluster soit définie en tant qu'alias sur l'unité de bouclage. Par exemple :

```
# modprobe noarp
# noarpctl add $CLUSTER_ADDRESS adresse-principale-nic
```

où *adresse-principale-nic* est une adresse appartenant au même sous-réseau que l'adresse du cluster. Il est ensuite possible d'affecter des alias aux clusters selon la méthode normale, comme dans l'exemple suivant :

```
# ifconfig lo:1 cluster address netmask 255.255.255.255 up
```

Remarque : Dans les configurations de co-implantation à haute disponibilité, placez `noarpctl adds` et `dels` dans les scripts `go*`. Cette opération garantit que le composant Load Balancer actif est capable de traiter l'adresse du cluster via le protocole ARP et que le composant Load Balancer de secours, agissant en tant que serveur, ne reçoit pas accidentellement (c'est-à-dire de manière indéterminée) l'ensemble du trafic du cluster.

4. Le correctif Julian est disponible sur le site Web suivant : <http://www.ssi.bg/~ja/#hidden>. Suivez les instructions de distribution pour l'installation de correctifs et la compilation d'un noyau pouvant être utilisé pour cette distribution. Si le composant Load Balancer est configuré pour la haute disponibilité et la

co-implantation, vérifiez que `uname -r` correspond au noyau fourni par distribution et que vous démarrez avec le fichier `.config` du noyau de distribution. Après avoir compilé, installé et exécuté le noyau avec le correctif de masquage Julian, activez ce dernier en suivant les instructions fournies pour la première solution.

Remarque : L'exécution d'un script personnalisé peut avoir des implications sur la prise en charge de la distribution.

Chapitre 8. Déploiement de Dispatcher sur Load Balancer pour IPv4 et IPv6

La prise en charge du schéma d'adressage IP étendu d'IPv6 est disponible avec Load Balancer pour IPv4 et IPv6. Load Balancer pour IPv4 et IPv6 constitue une image d'installation distincte ne comprenant que le composant Dispatcher. Ce type d'installation permet l'équilibrage de charge du trafic IPv4 et IPv6 vers les serveurs configurés sur votre réseau à l'aide de la méthode de transfert de paquets MAC.

Le présent chapitre décrit les limitations et les différences de configuration de Dispatcher sur l'installation Load Balancer pour IPv4 et IPv6 de ce produit et comprend les sections suivantes :

- «Plateformes prises en charge pour Load Balancer pour IPv4 et IPv6», à la page 82
- «Installation de Load Balancer pour IPv4 et IPv6», à la page 83
- «Limitations et remarques spéciales pour Load Balancer pour IPv4 et IPv6», à la page 84
- «Activation du traitement des paquets IPv6 dans Load Balancer pour IPv4 et IPv6», à la page 88
- «Création d'un alias pour le périphérique d'interface dans Load Balancer pour IPv4 et IPv6», à la page 88
- «Etapas de configuration de cluster requises pour Linux sur zSeries», à la page 91
- «Commandes Dispatcher (dscontrol) pour Load Balancer pour IPv4 et IPv6», à la page 92

Pour des informations générales sur le composant Dispatcher, reportez-vous aux chapitres suivants :

- Voir «Fonctions du composant Dispatcher», à la page 19 pour une présentation des fonctions de Dispatcher permettant de gérer votre réseau.
- Voir Chapitre 6, «Planification de Dispatcher», à la page 51 pour obtenir des informations sur la planification des paramètres d'équilibrage de charge de Dispatcher.
- Voir Chapitre 7, «Configuration de Dispatcher», à la page 63 pour obtenir des informations sur la configuration des paramètres d'équilibrage de charge de Dispatcher.
- Voir Chapitre 22, «Fonctions avancées de Dispatcher, CBR et Site Selector», à la page 201 pour obtenir des informations sur la configuration de Load Balancer pour les fonctions avancées.
- Voir Chapitre 24, «Exploitation et gestion de Load Balancer», à la page 261 pour obtenir des informations sur les journaux de Load Balancer et la syntaxe des composants de Load Balancer.

Il est important de noter qu'avec l'installation de Load Balancer pour IPv4 et IPv6, la syntaxe de la commande Dispatcher (dscontrol) est identique à une exception près. Le symbole at (@) a remplacé le signe deux-points (:) comme délimiteur des commandes dscontrol, lorsque Load Balancer pour IPv4 et IPv6 est utilisé. Lorsque vous faites référence à des commandes dans les autres chapitres du

présent document, n'oubliez pas de remplacer le signe deux-points (:) par (@), comme délimiteur dans les commandes dscontrol.

Plateformes prises en charge pour Load Balancer pour IPv4 et IPv6

Les installations Load Balancer pour IPv4 et IPv6 sont disponibles pour toutes les plateformes prises en charge à l'exception de Windows 2000.

Pour plus d'informations sur les conditions matérielles et logicielles requises, accédez à la page Web suivante : <http://www.ibm.com/support/docview.wss?rs=180&uid=swg27006921>

Plateformes prises en charge pour l'équilibrage de charge dans l'espace utilisateur

Sur certaines plateformes prises en charge, comme toutes les architectures Linux, les installations Load Balancer pour IPv4 et IPv6 peuvent équilibrer les charges de processus dans l'espace utilisateur, plutôt que dans l'espace noyau. Ces systèmes perdent alors toute dépendance vis à vis du module du noyau.

Pour les informations les plus récentes sur les types de plateforme acceptant l'équilibrage de charge dans l'espace utilisateur (hors noyau), voir le site Web suivant : <http://www.ibm.com/support/docview.wss?rs=180&uid=swg27006921>.

Les systèmes pris en charge exécutant les processus d'équilibrage de charge dans l'espace utilisateur comportent des procédures de configuration différentes de celles des systèmes exécutant les processus d'équilibrage de charge dans l'espace noyau. Ces différences sont traitées dans la présente section sur Load Balancer pour IPv4 et IPv6.

Remarques sur la plateforme Linux

Systèmes Linux sur zSeries

- **Les systèmes Linux sur zSeries exigent libstdc++.so.5** : Le package rpm libstdc++.so.5 est un prérequis à la bonne installation des systèmes Linux sur zSeries.
- **Restriction d'utilisation de l'interface qeth/OSA** : Pour les systèmes Linux sur zSeries, il existe une restriction d'utilisation de l'interface qeth/OSA. L'acheminement à partir d'une interface qeth/OSA en mode natif n'est pas pris en charge. Il existe toutefois une solution palliative car les systèmes Linux s'exécutent dans l'espace utilisateur et peuvent accepter l'établissement de tunnels Linux.

Prise en charge de l'établissement de tunnels Linux

Sur les systèmes Linux, les installations Load Balancer pour IPv4 et IPv6 peuvent transiter par des tunnels comme IPIP et IPGRE. Lorsque vous utilisez des machines Linux sur zSeries avec une interface qeth/OSA, vous pouvez définir un tunnel Linux qui passe par l'interface qeth/OSA. Les systèmes Linux peuvent effectuer un acheminement entre des machines situées sur les mêmes périphériques qeth/OSA, ou sur d'autres périphériques qeth/OSA, ou encore n'importe où sur le réseau.

Restrictions de serveur dorsal

Systèmes Solaris : L'équilibrage de charge du trafic IPv6 n'est pas pris en charge sur des serveurs dorsaux Solaris 5.8. Sous Solaris 5.8, il existe une incompatibilité

entre un paquet IPv6 acheminé par MAC et la pile Solaris IPv6. Lorsque le cluster est configuré sur un serveur dorsal Solaris 5.8 à l'aide de la commande `ifconfig lo0` (unité de bouclage), le paquet arrive sur le noeud Solaris 5.8, mais n'est pas accepté. Vous pouvez toutefois utiliser les installations Load Balancer pour IPv4 et IPv6 pour équilibrer la charge du trafic IPv4 sur les serveurs dorsaux Solaris 5.8.

Systèmes z/OS : L'équilibrage de charge du trafic IPv6 n'est pas pris en charge sur les serveurs dorsaux z/OS. Vous pouvez toutefois utiliser les installations Load Balancer pour IPv4 et IPv6 pour équilibrer la charge du trafic IPv4 sur les serveurs dorsaux z/OS.

Installation de Load Balancer pour IPv4 et IPv6

Les noms de package et les étapes d'installation de Load Balancer pour IPv4 et IPv6 sont identiques à ceux de Load Balancer prenant en charge uniquement les adresses de serveur IPv4. Toutefois, le nombre de packages du composant Load Balancer fournis est inférieur car seul le composant Dispatcher est disponible.

Lorsque vous utilisez les outils de création de packages du système, l'ordre d'installation recommandé des packages est légèrement différent pour les installations de Load Balancer pour IPv4 et IPv6. Le package du composant d'administration doit être installé après celui du composant Dispatcher. L'ordre d'installation recommandé pour Load Balancer pour IPv4 et IPv6 avec ces outils est le suivant : produit de base, licence, composant Dispatcher, administration, documentation, Metric Server.

Par exemple, pour les systèmes AIX, voici la liste de packages Load Balancer pour IPv4 et IPv6 dans l'ordre d'installation recommandé :

- `ibmlb.base.rte` (package du produit de base)
- `ibmlb.lb.license` (package de licence, en cas d'installation à partir d'un CD)
- `ibmlb.lb.driver` (package de pilote de périphérique, qui est un package unique pour AIX uniquement)
- `ibmlb.disp.rte` et `ibmlb.msg.lang.lb` (package du composant Dispatcher avec package des messages)
- `ibmlb.admin.rte` et `ibmlb.msg.lang.admin` (package d'administration avec package des messages)
- `ibmlb.doc.rte` et `ibmlb.msg.en_US.doc` (package de la documentation avec package des messages)
- `ibmlb.ms.rte` (package Metric Server)

Notez que les versions précédentes de Load Balancer doivent être désinstallées avant l'installation de Load Balancer pour IPv4 et IPv6. Il n'est pas possible d'installer deux instances de Load Balancer sur une même machine.

Pour les instructions d'installation du produit, voir Chapitre 4, «Installation de Load Balancer», à la page 31.

Limitations et remarques spéciales pour Load Balancer pour IPv4 et IPv6

Le composant Dispatcher offre, à défaut de toutes les fonctions, de nombreuses fonctions disponibles avec le composant Dispatcher sur les installations Load Balancer ne prenant en charge qu'IPv4. Les rubriques ci-après abordent les limitations et les différences de configuration spéciales de Dispatcher dans Load Balancer pour IPv4 et IPv6.

Configuration de l'adresse lien-local IPv6

Avec l'adressage IPv6, chaque machine de la configuration Load Balancer doit disposer d'une adresse lien-local.

L'adresse lien-local est l'adresse utilisée pour le trafic de reconnaissance dans le voisinage pour IPv6. Sans cette adresse sur la machine Load Balancer et sur les serveurs dorsaux, la reconnaissance dans le voisinage n'a pas lieu et les machines ne se reconnaissent pas entre elles. Load Balancer pour IPv6 ne peut pas acheminer le trafic sans adresse IPv6 lien-local configurée sur une interface de chaque machine de la configuration Load Balancer.

Paires cluster/serveur homogènes

Lors de la configuration de Load Balancer pour IPv4 et IPv6, tous les serveurs doivent être homogènes à l'intérieur du cluster. Par exemple, si Cluster1 est défini avec une adresse IPv4, tous les serveurs qui se trouvent sous Cluster1 doivent posséder une adresse IPv4. Si Cluster2 est défini avec une adresse IPv6, tous les serveurs définis sous Cluster2 doivent posséder une adresse IPv6. En outre, le protocole utilisé par le client pour envoyer des paquets IP doit correspondre au format IP du cluster.

La prise en charge d'un environnement mixte de clients IPv4 et IPv6 requiert que pour chaque définition de cluster logique, deux définitions de cluster réel soient définies : un cluster IPv4 et un cluster IPv6. Les clients envoyant des paquets IPv4 sont acheminés par Load Balancer au cluster logique à l'aide des adresses IPv4 configurées pour le cluster. Les clients envoyant des paquets IPv6 sont acheminés par Load Balancer au cluster logique à l'aide des adresses IPv6 configurées pour le cluster.

Fonctions Dispatcher non prises en charge

De nombreuses fonctions de Dispatcher décrites dans Chapitre 6, «Planification de Dispatcher», à la page 51 et les fonctions de Dispatcher décrites dans Chapitre 22, «Fonctions avancées de Dispatcher, CBR et Site Selector», à la page 201 sont disponibles dans Load Balancer pour IPv4 et IPv6.

La liste suivante est un récapitulatif des fonctions de Dispatcher *non* prises en charge dans Load Balancer pour IPv4 et IPv6 :

- méthode d'acheminement cbr
- méthode d'acheminement nat
- administration à distance
- équilibrage basé sur des règles
- sous-agent SNMP
- équilibrage de charge de réseau étendu
- prise en charge du protocole UDP

Voir «Fonctions du composant Dispatcher», à la page 19 pour une description détaillée des fonctions de Dispatcher permettant de gérer votre réseau.

Configuration de conseillers

Si vous utilisez le protocole IPv6 sur votre machine et que vous souhaitez utiliser des conseillers, vous devez vous assurer que la ligne suivante est incluse dans le fichier du **protocole**.

```
ipv6-icmp 58 IPv6-ICMP
```

Pour les systèmes Linux et UNIX, le fichier du protocole se trouve dans le répertoire `/etc/protocols`. Pour les systèmes Windows, le fichier du protocole se trouve dans le répertoire `C:\windows\system32\drivers\etc\`.

Limitation en cas d'utilisation de conseillers : Si Load Balancer s'exécute sur un ordinateur doté de plusieurs cartes réseau et que vous voulez que le trafic du conseiller passe par une carte particulière, vous ne pouvez pas imposer une adresse spécifique comme adresse IP source du paquet lorsque vous voulez que le trafic du conseiller passe par une carte donnée. (La propriété `-DLB_ADV_SRC_ADDR` n'est pas disponible dans les installations Load Balancer pour IPv4 et IPv6.)

Pour plus d'informations sur les conseillers, voir «Conseillers», à la page 248.

Configuration de la haute disponibilité

Si vous utilisez le protocole IPv6 sur votre machine et que vous souhaitez utiliser la haute disponibilité, vous devez vérifier que `protocol 58` est défini comme ICMPv6 dans le fichier du **protocole**. Pour plus d'informations sur la modification du fichier du protocole, voir «Configuration de conseillers».

Dans les installations Load Balancer pour IPv4 et IPv6, la configuration d'une machine Dispatcher à haute disponibilité est acceptée avec les restrictions suivantes :

- La haute disponibilité réciproque n'est pas prise en charge.
- Les paires de signaux de présence (mécanisme entre les machines Dispatcher principale et de secours permettant de détecter un échec de Dispatcher) doivent être toutes deux au format IPv4 ou toutes deux au format IPv6.
- Pour les systèmes s'exécutant dans l'espace utilisateur, comme les systèmes Linux : Dans un environnement à haute disponibilité ou autonome, vous ne devez pas créer d'alias entre l'adresse du cluster et la carte réseau.
- Pour les systèmes s'exécutant dans l'espace utilisateur, comme les systèmes Linux : Vous pouvez transférer les scripts `go*` et `highavailChange` du répertoire `.../ibm/edge/lb/servers/samples` au répertoire `.../ibm/edge/lb/servers/bin` pour consigner les modifications de l'état de haute disponibilité pour la machine Dispatcher, sans avoir besoin de les modifier.
- Pour les systèmes Linux sur zSeries utilisant une interface qeth/OSA : Pour ce type d'interface réseau uniquement, l'interdiction générale d'utilisation d'alias d'interface pour les adresses de cluster ne s'applique pas. Utilisez plutôt la procédure suivante pour garantir que le trafic du cluster est fourni à l'hôte Linux via un processus OSA :
 - Les scripts `go*` sont requis et doivent être modifiés comme suit à l'aide des commandes spécifiées dans «Étapes de configuration de cluster requises pour Linux sur zSeries», à la page 91:
 - `goActive` : Ajoutez les commandes `ip` et `iptables/ip6tables` pour configurer l'adresse du cluster et ajouter la règle `iptables`.

- goStandby : Ajoutez les commandes ip et iptables/ip6tables pour annuler la configuration de l'adresse du cluster et supprimer la règle iptables.
- goInOp : Ajoutez les commandes ip et iptables/ip6tables pour annuler la configuration de l'adresse du cluster et supprimer la règle iptables.
- goIdle : Ce script ne doit pas être créé.

Pour plus d'informations sur la fonction de haute disponibilité, voir «Haute disponibilité», à la page 204.

Co-implantation de serveurs

La co-implantation est une configuration dans laquelle Load Balancer peut se trouver sur la même machine que le serveur pour lequel il équilibre la charge des demandes.

Lors de l'utilisation d'installations Load Balancer pour IPv4 et IPv6, la fonction de co-implantation est disponible sur tous les systèmes d'exploitation pris en charge, excepté les systèmes Windows et ceux qui s'exécutent dans l'espace utilisateur, comme les systèmes Linux.

Pour plus d'informations sur la co-implantation de serveurs, voir «Utilisation de serveurs implantés au même endroit», à la page 202.

Fonction d'affinité pour les systèmes qui s'exécutent dans l'espace utilisateur (Linux)

La fonction d'affinité de Load Balancer pour les systèmes s'exécutant dans l'espace utilisateur, comme les systèmes Linux, fonctionne différemment de la fonction d'affinité pour les autres systèmes d'exploitation s'exécutant dans l'espace noyau.

Pour les systèmes s'exécutant dans l'espace utilisateur, Load Balancer mappe l'adresse IP du client avec un serveur dorsal. L'affinité est établie dès que l'adresse IP de destination d'un paquet correspond au cluster, que le port de destination correspond à celui de Load Balancer et que l'adresse IP source correspond.

Lorsque l'affinité est établie, les paquets sont envoyés par la suite au même serveur dorsal. Lorsqu'elle est interrompue, suite à une panne ou à un retrait de serveur, toute affinité, et par conséquent toutes les connexions à ce serveur, sont interrompues.

En outre, aucune information de "connexion" n'apparaît sur la ligne de commande ou l'interface graphique des clients. Seul le nombre d'enregistrements d'affinité actifs est utilisé.

Cette approche présente l'avantage de fournir une forte affinité et d'assurer une meilleure efficacité à Load Balancer.

Pour les systèmes traitant l'équilibrage de charge au niveau du noyau, l'aspect négatif de l'utilisation de l'affinité d'IP est la surcharge ajoutée au niveau de la CPU et de la mémoire dans le mécanisme d'acheminement des connexions. Sur les systèmes gérant l'équilibrage de charge dans l'espace utilisateur, la méthode d'affinité utilisée est moins consommatrice de mémoire et de CPU par rapport à l'acheminement des connexions.

De plus, à cause de ce modèle à un seul enregistrement dans les systèmes s'exécutant dans l'espace utilisateur, les valeurs stickytime et staletimeout associées

à l'affinité se sont fondues en une seule valeur : `staletimeout`. Comme la suppression d'un enregistrement d'affinité interrompt également les connexions, lors de la migration d'un système à traitement dans l'espace noyau vers un système à traitement dans l'espace utilisateur, la valeur `staletimeout` et `stickytime` maximale doit être utilisée comme nouvelle valeur `staletimeout` pour le processus Load Balancer s'exécutant dans le système à espace utilisateur.

Pour des informations générales sur la fonction d'affinité pour les systèmes à traitement dans l'espace noyau, par opposition à l'espace utilisateur, voir «Fonctionnement de la fonction d'affinité pour Load Balancer», à la page 221.

Configuration de Metric Server

Si vous utilisez le protocole IPv6 sur votre machine et que vous souhaitez utiliser Metric Server, vous devez vérifier que `protocol 58` est défini comme ICMPv6 dans le fichier du **protocole**. Pour plus d'informations sur la modification du fichier du protocole, voir «Configuration de conseillers», à la page 85.

Dans une configuration Load Balancer prenant en charge les clusters IPv4 et IPv6, les serveurs exécutant la fonction Metric Server peuvent être configurés soit comme serveurs IPv4 uniquement, soit comme serveurs IPv6 uniquement. Pour forcer le serveur de mesures à utiliser un protocole particulier, IPv4 ou IPv6, indiquez la propriété Java `java.rmi.server.hostname` dans le script `metricserver`.

IMPORTANT : Le nom d'hôte spécifié dans la propriété Java doit être l'adresse IP physique du serveur de mesures.

Sur les systèmes UNIX ou Linux : Pour permettre à Metric Server de communiquer via l'adresse IPv6 `2002:92a:8f7a:162:9:42:92:67`, indiquez la propriété Java après `$LB_CLASSPATH` dans le script de démarrage `metricserver` (dans le répertoire `/usr/bin`) comme suit :

```
/opt/ibm/edge/lb/java/jre/bin/java ..... $LB_CLASSPATH
-Djava.rmi.server.hostname=2002:92a:8f7a:162:9:42:92:67
com.ibm.internet.nd.sma.SMA_Agent
$LB_RMIPORT $LOG_LEVEL $LOG_SIZE $LOG_DIRECTORY $KEYS_DIRECTORY
$SCRIPT_DIRECTORY &
```

Sur les systèmes Windows : Pour permettre à Metric Server de communiquer via l'adresse IPv6 `2002:92a:8f7a:162:9:42:92:67`, modifiez le fichier `metricserver.cmd` (dans le répertoire `C:\winnt\system32`) comme suit :

```
start
/min /wait %IBMLBPATH%\java\jre\bin\java
-Djava.rmi.server.hostname=2002:92a:8f7a:162:9:42:92:67
-Djava.net.preferIPv4Stack=false
-Djava.net.preferIPv6Stack=true
-Xrs -cp
%LB_CLASSPATH% com.ibm.internet.nd.sma.SMA_Agent
%RMI_PORT% %LOG_LEVEL% %LOG_SIZE% %LOG_DIRECTORY% %KEYS_DIRECTORY%
%SCRIPT_DIRECTORY%
goto done

:stop
%IBMLBPATH%\java\jre\bin\java
-Djava.rmi.server.hostname=2002:92a:8f7a:162:9:42:92:67
-Djava.net.preferIPv4Stack=false
-Djava.net.preferIPv6Stack=true
-cp %LB_CLASSPATH% com.ibm.internet.nd.sma.SMA_AgentStop %RMI_PORT%
:done
```

Pour plus d'informations, voir «Metric Server», à la page 196.

Activation du traitement des paquets IPv6 dans Load Balancer pour IPv4 et IPv6

Sous AIX, Linux et Windows : Avant de démarrer l'exécuteur (`dscontrol executor start`), vous devez exécuter les commandes suivantes à partir de la ligne de commande, en tant que root :

- Pour les systèmes AIX : `autoconf6`
Pour activer le traitement continu des paquets IPv6 (même après un redémarrage du système), modifiez le fichier `/etc/rc.tcpip` en supprimant la mise en commentaire de la ligne suivante et en ajoutant l'option `-A : start`
`/usr/sbin/autoconf6 " " -A`
- Pour les systèmes Linux : `modprobe ipv6`
- Pour les systèmes Windows : `netsh interface ipv6 install`

Ces commandes activent le traitement des paquets IPv6 dans leurs systèmes d'exploitation respectifs. La commande ne doit être exécutée qu'une seule fois. Vous pouvez ensuite démarrer et arrêter l'exécuteur aussi souvent que nécessaire.

Si vous n'exécutez pas la commande permettant d'activer le traitement des paquets IPv6 sur ces systèmes, l'exécuteur ne démarre pas.

Sur les systèmes HP-UX et Solaris : Les adresses IPv6 doivent être sondées et une interface configurée, à l'aide de la commande `ifconfig`, pour que Dispatcher contrôle les paquets IPv6. Avant de démarrer l'exécuteur (`dscontrol executor start`), exécutez la commande suivante à partir de la ligne de commande, en tant que root :

- Pour les systèmes HP-UX :
`ifconfig device inet6 up`
- Pour les systèmes Solaris :
`ifconfig device inet6 plumb`
`ifconfig device inet6 adresse/préfixe up`

Si vous n'exécutez pas ces commandes, l'exécuteur démarre, mais aucun paquet IPv6 ne peut être affiché.

Création d'un alias pour le périphérique d'interface dans Load Balancer pour IPv4 et IPv6

Pour configurer l'adresse du cluster sur une carte NIC (network interface card) de la machine Dispatcher, vous pouvez exécuter la commande `dscontrol executor configure adresse_cluster`. La commande `dscontrol executor configure` exécute les commandes de configuration des adaptateurs du système d'exploitation (par exemple, les commandes `ifconfig`, `dsconfig` (IPv6 uniquement) ou `ip`). Pour créer un alias pour la carte NIC de la machine Dispatcher, vous pouvez également choisir d'exécuter directement les commandes de configuration des adaptateurs du système d'exploitation, au lieu d'utiliser la commande `executor configure`.

Remarque : Pour les systèmes s'exécutant dans l'espace utilisateur, comme les systèmes Linux : Vous ne devez pas configurer l'adresse de cluster à l'aide de la commande `dscontrol executor configure`, `ip` ou `ifconfig`. Load Balancer annonce en mode natif l'adresse du cluster sur le réseau. En outre, l'adresse du cluster n'apparaît pas comme un alias d'une interface. Il s'agit de procédures normales.

Toutefois, ces affirmations **ne s'appliquent pas** à Linux sur zSeries utilisant une interface qeth/OSA. Pour cette plateforme, vous configurez l'adresse du cluster. Pour plus d'informations, voir «Étapes de configuration de cluster requises pour Linux sur zSeries», à la page 91.

Pour créer un alias pour l'unité de bouclage (lo0) sur les serveurs dont la charge est équilibrée, vous devez utiliser les commandes de configuration des adaptateurs du système d'exploitation.

Pour l'installation Load Balancer pour IPv4 et IPv6, les commandes ci-après permettent de créer un alias pour l'interface réseau et l'unité de bouclage (*nom_interface*).

Sous AIX (5.x) :

- Pour les adresses IPv6 :

```
ifconfig nom_interface inet6 adresse_cluster/longueur_préfixe alias
```

Par exemple, pour créer un alias de l'unité de bouclage sur les serveurs dont la charge est équilibrée :

```
ifconfig lo0 inet6 2002:4a::541:56/128 alias
```

- Pour les adresses IPv4 : Inchangé. Voir tableau 5, à la page 73 pour créer un alias de l'unité de bouclage sur les serveurs dont la charge est équilibrée.

Sous HP-UX :

- Pour les adresses IPv6 :

```
ifconfig nom_interface:alias inet6 adresse_cluster up prefix longueur_préfixe
```

Par exemple, pour créer un alias de l'unité de bouclage sur les serveurs dont la charge est équilibrée :

```
ifconfig lo0:1 inet6 3ffe:34::24:45 up prefix 128
```

- Pour les adresses IPv4 : Inchangé. Voir tableau 5, à la page 73 pour créer un alias de l'unité de bouclage sur les serveurs dont la charge est équilibrée.

Sous Linux :

- Pour les adresses IPv6 ou IPv4 :

```
ip -version addr add adresse_cluster/longueur_préfixe dev lo
```

Par exemple, pour créer un alias de l'unité de bouclage sur les serveurs dont la charge est équilibrée :

```
ip -6 addr add 3ffe:34::24:45/128 dev lo
```

```
ip -4 addr add 12.42.38.125/32 dev lo
```

Remarque : Vous pouvez également avoir recours à la commande `ifconfig`. Voir tableau 5, à la page 73 pour créer un alias de l'unité de bouclage à l'aide de la commande `ifconfig`.

Une fois que vous avez utilisé une des commandes de configuration sur votre machine, reprenez-la systématiquement (**ip** ou **ifconfig**), sinon des résultats imprévisibles risquent de se produire.

Sur les systèmes Solaris 8, 9 et 10 :

- Pour les adresses IPv6 :


```
ifconfig nom_interface inet6 addif adresse_cluster/longueur_préfixe up
```

Par exemple, pour créer un alias de l'unité de bouclage sur les serveurs dont la charge est équilibrée :

```
ifconfig lo0 inet6 addif 3ffe:34::24:45/128 up
```

- Pour les adresses IPv4 : Inchangé. Voir tableau 5, à la page 73 pour créer un alias de l'unité de bouclage sur les serveurs dont la charge est équilibrée.

Sous Windows 2003 (Windows 2000 et Windows NT ne prennent pas en charge IPv6) :

- Pour les adresses IPv4 : Inchangé. Voir tableau 5, à la page 73 pour créer un alias de l'unité de bouclage sur les serveurs dont la charge est équilibrée.

- Pour les adresses IPv6 :

1. Utilisez la commande `ipconfig /all` pour déterminer le nom d'interface de l'unité de bouclage. Cette commande localise la connexion avec une description de l'unité Microsoft Loopback Adapter (carte de bouclage Microsoft). Voici un exemple de résultat de la commande `ipconfig /all`, où la carte de bouclage Microsoft est la carte de réseau Ethernet Local Area Connection 2, et la connexion est Local Area Connection 2 :

Configuration d'IP Windows

```
Nom d'hôte . . . . . : ndserv10
Suffixe DNS principal . . . . . : rtp.raleigh.ibm.com
Type de noeud . . . . . : inconnu
routage IP activé. . . . . : Non
Proxy WINS activé. . . . . : Non
Liste de recherche des suffixes DNS. . . . . : rtp.raleigh.ibm.com
```

Carte Ethernet Local Area Connection 2 :

```
Suffixe DNS spécifique de la connexion . :
Description . . . . . : Microsoft Loopback Adapter
Adresse physique. . . . . : 02-00-4C-4F-4F-50
DHCP activé. . . . . : Non
Adresse IP. . . . . : 9.42.92.158
Masque de sous-réseau . . . . . : 255.255.252.0
Adresse IP. . . . . : 9.42.92.159
Masque de sous-réseau . . . . . : 255.255.252.0
Adresse IP. . . . . : 2002:92a:8f7a:162:9:42:92:160
Adresse IP. . . . . : 2002:92a:8f7a:162:9:42:92:159
Adresse IP. . . . . : fe80::4cff:fe4f:4f50%4
Passerelle par défaut. . . . . :
Serveurs DNS. . . . . : 127.0.0.1
                        fec0:0:0:ffff::1%1
                        fec0:0:0:ffff::2%1
                        fec0:0:0:ffff::3%1
```

2. Ajoutez l'adresse du cluster à l'unité de bouclage à l'aide de la commande `netsh`. Par exemple :

```
netsh interface ipv6 add address "Local Area Connection 2"
2002:92a:8f7a:162:9:42:92:161
```

3. Lancez à nouveau la commande `ipconfig /all` : l'adresse doit à présent apparaître dans la carte de bouclage. Par exemple :

Carte Ethernet Local Area Connection 2 :

```
Suffixe DNS spécifique de la connexion . :
Description . . . . . : Microsoft Loopback Adapter
Adresse physique. . . . . : 02-00-4C-4F-4F-50
DHCP activé. . . . . : Non
Adresse IP. . . . . : 9.42.92.158
Masque de sous-réseau . . . . . : 255.255.252.0
```



```

Adresse IP. . . . . : 9.42.92.159
Masque de sous-réseau . . . . . : 255.255.252.0
Adresse IP. . . . . : 2002:92a:8f7a:162:9:42:92:161
Adresse IP. . . . . : 2002:92a:8f7a:162:9:42:92:160
Adresse IP. . . . . : 2002:92a:8f7a:162:9:42:92:159
Adresse IP. . . . . : fe80::4cff:fe4f:4f50%4
Passerelle par défaut. . . . . :
Serveurs DNS. . . . . : 127.0.0.1
                        fec0:0:0:ffff::1%1
                        fec0:0:0:ffff::2%1
                        fec0:0:0:ffff::3%1

```

4. Activez l'acheminement pour toutes les interfaces de la machine avec la commande `netsh interface ipv6 show interface`. L'acheminement IP doit être activé pour toute interface répertoriée sous le nom Local Area Connection. Par exemple :

```

netsh interface ipv6>show interface
Querying active state...

```

Idx	Met	MTU	State	Name
6	2	1280	Disconnected	Teredo Tunneling Pseudo-Interface
5	0	1500	Connected	Local Area Connection
4	0	1500	Connected	Local Area Connection 2
3	1	1280	Connected	6to4 Pseudo-Interface
2	1	1280	Connected	Automatic Tunneling Pseudo-Interface
1	0	1500	Connected	Loopback Pseudo-Interface

```

netsh interface ipv6>set interface "Local Area Connection"
forwarding=enabled
Ok.

```

```

netsh interface ipv6>set interface "Local Area Connection 2"
forwarding=enabled
Ok.

```

Sous OS/2 :

- Pour les adresses IPv6 et IPv4 : Inchangé. Voir tableau 5, à la page 73 pour créer un alias de l'unité de bouclage sur les serveurs dont la charge est équilibrée.

Etapas de configuration de cluster requises pour Linux sur zSeries

Pour Linux sur zSeries, certaines étapes de configuration supplémentaires sont nécessaires pour configurer Load Balancer :

1. Configurez l'adresse de cluster avec la commande `ip` ou `ifconfig`.

Pour les adresses IPv6 ou IPv4 :

```
ip -version addr add adresse_cluster/longueur_préfixe dev périphérique
```

Par exemple :

```
ip -4 addr add 12.42.38.125/24 dev eth0
ip -6 addr add 3ffe:34::24:45/64 dev eth0
```

2. Ajoutez une règle iptables pour supprimer les paquets entrants destinés à l'adresse du cluster :

Pour les adresses IPv4 :

```
iptables -t filter -A INPUT -d adresse_cluster -j DROP
```

Pour les adresses IPv6 :

```
ip6tables -t filter -A INPUT -d adresse_cluster -j DROP
```

Par exemple :

```
iptables -t filter -A INPUT -d 12.42.38.125 -j DROP
ip6tables -t filter -A INPUT -d 3ffe:34::24:45 -j DROP
```

Pour annuler la configuration ci-dessus, utilisez les commandes suivantes :

```
ip -version addr del adresse_cluster/longueur_préfixe dev périphérique
iptables -t filter -D INPUT -d adresse_cluster -j DROP
ip6tables -t filter -D INPUT -d adresse_cluster -j DROP
```

Commandes Dispatcher (dscontrol) pour Load Balancer pour IPv4 et IPv6

Load Balancer pour IPv4 et IPv6 ne prenant pas en charge toutes les fonctions du composant Dispatcher, les commandes `dscontrol` valides de cette installation sont un sous-ensemble des commandes `dscontrol` des installations Load Balancer qui ne prennent en charge qu'IPv4. La présente section aborde les différences de syntaxe des commandes et répertorie toutes les commandes `dscontrol` prises en charge pour le composant Dispatcher dans Load Balancer pour IPv4 et IPv6.

Différences entre les syntaxes des commandes

Avec l'installation de Load Balancer pour IPv4 et IPv6, la syntaxe de la commande Dispatcher (`dscontrol`) est identique à une exception importante près. Le délimiteur des commandes `dscontrol` est le symbole at (@) au lieu du signe deux-points (:), si Load Balancer pour IPv4 et IPv6 est utilisé.

Il était nécessaire de définir un délimiteur autre que le signe deux-points (:) car le format IPv6 l'utilise dans son schéma d'adressage.

L'exemple suivant illustre la commande `dscontrol` avec un délimiteur at (@) :

- pour ajouter un serveur IPv6 (30::200) sur le port 80, sous un cluster IPv6 (30::100)
`dscontrol server add 30::100@80@30::200`
- pour ajouter un serveur IPv4 (192.4.40.35) sur le port 80, sous un cluster IPv4 (192.4.40.30)
`dscontrol server add 192.4.40.30@80@192.4.20.35`

IMPORTANT : Si vous voulez faire référence à des commandes dans le présent document, n'oubliez pas de remplacer le signe deux-points (:) par (@), comme délimiteur dans les commandes `dscontrol`.

Commandes dscontrol prises en charge

Pour des exemples et des informations détaillées sur la syntaxe de toutes les commandes `dscontrol`, voir Chapitre 27, «Guide des commandes Dispatcher et CBR», à la page 345.

Vous trouverez ci-après un récapitulatif de toutes les commandes prises en charge pour Dispatcher dans l'installation Load Balancer pour IPv4 et IPv6 :

- `dscontrol advisor`
 - Tous les arguments et leurs valeurs de clé sont valides.
 - Pour une description détaillée de la syntaxe des commandes, voir : «`dscontrol advisor` — Contrôle du conseiller», à la page 347.
- `dscontrol binlog`
 - Tous les arguments et leurs valeurs de clé sont valides.

- Pour une description détaillée de la syntaxe des commandes, voir : «dscontrol binlog — Contrôle du fichier journal binaire», à la page 352.
 - `dscontrol cluster`
 - Tous les arguments sont valides. Les seules valeurs de clé valides sont `address` et `proportions`.
 - Pour une description détaillée de la syntaxe des commandes, voir : «dscontrol cluster — Configuration des clusters», à la page 353.
 - `dscontrol executor`
 - Tous les arguments sont valides. Pour l'argument `set`, les seules valeurs de clé valides sont `nfa`, `hatimeout` et `hasynctimeout`.
 - Pour les systèmes s'exécutant dans l'espace utilisateur, comme les systèmes Linux :**
 - Tous les arguments sont valides sauf `configure` et `unconfigure`. Il est important de noter qu'un alias des adresses du cluster ne doit jamais être créé dans la pile système.
 - Pour l'argument `set`, les seules valeurs de clé valides sont `nfa` et `hatimeout`.
 - Pour l'argument `configure`, vous devez substituer *longueur_préfixe* à *masque_de_réseau*.
 Pour IPv6, la longueur du préfixe représente le nombre de bits de la partie réseau de l'adresse IPv6. La longueur du préfixe représente l'adresse réseau dans l'adresse de l'hôte.
 Pour IPv4, déterminez la longueur du préfixe comme suit : si le masque de sous-réseau est 255.255.252.0, l'équivalent hexadécimal est FF.FF.FC.0. En mode binaire, la valeur est 11111111 11111111 11111100 00000000. Le nombre de 1 dans le masque de sous-réseau détermine la longueur du préfixe. Si le masque de sous-réseau compte 22 "uns", le préfixe est 22.
 La syntaxe de la commande `executor configure` est la suivante :
`dscontrol executor configure adresse_interface nom_interface longueur_préfixe`
- Exemple d'adressage IPv6 :
- ```
dscontrol executor configure 2002:092a:8f7a:4226:9:37:240:99 en0 112
```
- Exemple d'adressage IPv4, si le masque de sous-réseau est 255.255.252.0:
- ```
dscontrol e config 191.60.20.20 en1 22
```
- Il est important de remarquer que la commande `executor configure` n'est pas utilisée pour les systèmes s'exécutant dans l'espace utilisateur, comme les systèmes Linux, sur les installations Load Balancer pour IPv4 et IPv6.
- Pour une description détaillée de la syntaxe des commandes, voir : «dscontrol executor — Contrôle de l'exécuteur», à la page 357.
 - `dscontrol file`
 - Tous les arguments et leurs valeurs de clé sont valides.
 - Pour une description détaillée de la syntaxe des commandes, voir : «dscontrol file — Gestion des fichiers de configuration», à la page 362.
 - `dscontrol help`
 - Tous les arguments sont valides, sauf `host` (configuration d'une machine éloignée), `rule` (configuration de règles) et `subagent` (configuration du sous-agent SNMP). Les commandes `host`, `rule` et `subagent` ne sont pas prises en charge.

- Pour une description détaillée de la syntaxe des commandes, voir : «dscontrol help — Affichage ou impression de l’aide relative à cette commande», à la page 364.
- dscontrol highavailability
 - Tous les arguments sont valides. Toutes les valeurs de clé sont valides sauf both car la haute disponibilité réciproque n’est pas prise en charge.
 - Pour une description détaillée de la syntaxe des commandes, voir : «dscontrol highavailability — Contrôle de la haute disponibilité», à la page 365.
- dscontrol logstatus
 - Tous les arguments et leurs valeurs de clé sont valides.
 - Pour une description détaillée de la syntaxe des commandes, voir : «dscontrol logstatus — Affichage des paramètres du journal du serveur», à la page 370.
- dscontrol manager
 - Tous les arguments sont valides sauf version. Toutes les valeurs de clé sont valides.
 - Pour une description détaillée de la syntaxe des commandes, voir : «dscontrol manager — Contrôle du gestionnaire», à la page 371.
- dscontrol metric
 - Tous les arguments et leurs valeurs de clé sont valides.
 - Pour une description détaillée de la syntaxe des commandes, voir : «dscontrol metric — Configuration des mesures du système», à la page 377.
- dscontrol port
 - Tous les arguments sont valides à l’exception de halfopenaddressreport, qui n’est pas pris en charge.

Les valeurs de clé suivantes sont les seules valeurs valides pour les arguments add et set de la commande dscontrol port :

 - staletimeout
 - weightbound
 - stickymask

Pour les systèmes s’exécutant dans l’espace utilisateur comme les systèmes Linux : Les valeurs de clé suivantes sont les seules valeurs valides pour les arguments add et set de la commande dscontrol port :

 - staletimeout
 - weightbound
 - selectionalgorithm

Les options de selectionalgorithm (algorithme de sélection de serveur) sont :

 - connection – la sélection de serveur est basée sur une technique simple de permutation circulaire (par défaut)
 - affinity – la sélection de serveur est basée sur l’affinité client.

Par exemple :

```
dscontrol port add cluster@port selectionalgorithm affinity
```

- Pour une description détaillée de la syntaxe des commandes, voir : «dscontrol port — Configuration des ports», à la page 378.
- dscontrol server
 - Tous les arguments sont valides.

Les valeurs de clé suivantes sont valides pour l’argument add de la commande dscontrol server :

- adresse
- advisorrequest
- advisorresponse
- collocated

Le mot clé `collocated` est disponible tous les systèmes d'exploitation pris en charge, excepté les systèmes Windows et ceux qui s'exécutent dans l'espace utilisateur, comme les systèmes Linux.

- fixedweight
- weight

Les valeurs de clé suivantes sont valides pour l'argument `set` de la commande `dscontrol server` :

- advisorrequest
- advisorresponse
- collocated

Le mot clé `collocated` est disponible sur tous les systèmes d'exploitation pris en charge, excepté les systèmes Windows et ceux qui s'exécutent dans l'espace utilisateur, comme les systèmes Linux.

- fixedweight
- weight

- Pour une description détaillée de la syntaxe des commandes, voir : «`dscontrol server` — Configuration des serveurs», à la page 390.
- `dscontrol set`
 - Tous les arguments et leurs valeurs de clé sont valides.
 - Pour une description détaillée de la syntaxe des commandes, voir : «`dscontrol set` — Configuration du journal du serveur», à la page 396.
- `dscontrol status`
 - Tous les arguments et leurs valeurs de clé sont valides.
 - Pour une description détaillée de la syntaxe des commandes, voir : «`dscontrol status` — Indique par affichage si le gestionnaire et les conseillers sont en cours d'exécution», à la page 397.

Commandes `dscontrol` non prises en charge

Les commandes suivantes *ne sont pas* disponibles pour Dispatcher dans l'installation Load Balancer pour IPv4 et IPv6 :

- `dscontrol host` (configuration d'une machine éloignée)
- `dscontrol rule` (configuration de règles)
- `dscontrol subagent` (configuration du sous-agent SNMP)

Partie 3. Composant CBR (Content Based Routing)

Cette section contient des informations pour la configuration d'un démarrage rapide ainsi que des remarques relatives à la planification, et présente les diverses méthodes de configuration du composant CBR de Load Balancer. Elle se compose des chapitres suivants :

- Chapitre 9, «Configuration de démarrage rapide», à la page 99
- Chapitre 10, «Planification de CBR (Content Based Routing)», à la page 105
- Chapitre 11, «Configuration de CBR (Content Based Routing)», à la page 109

Chapitre 9. Configuration de démarrage rapide

Cet exemple de démarrage rapide indique comment configurer trois postes de travail connectés en local en associant le composant CBR au module Caching Proxy pour équilibrer la charge du trafic Web entre deux serveurs Web. (Par souci de simplicité, cet exemple se base sur des serveurs résidant sur le même segment de réseau local, alors que CBR ne l'impose pas.)

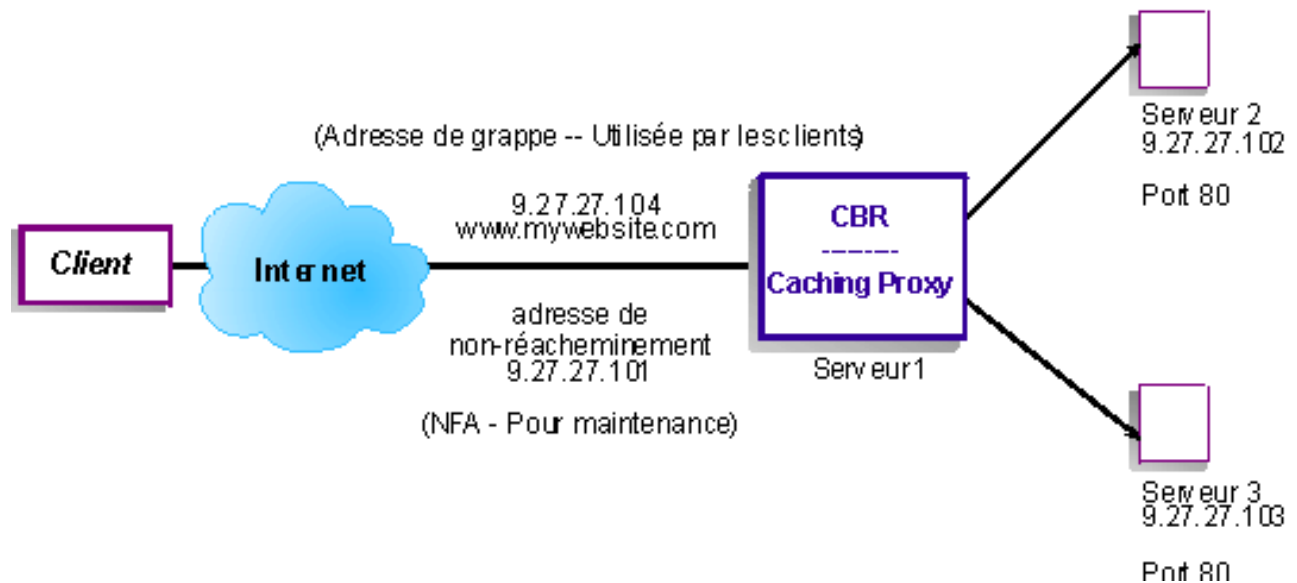


Figure 16. Configuration CBR locale simple

Matériel requis

Pour l'exemple à démarrage rapide, vous devez disposer de trois postes de travail et de quatre adresses IP. L'un des postes de travail est utilisé comme machine CBR et les deux autres comme serveurs Web. Chaque serveur Web requiert une adresse IP. Le poste CBR requiert une adresse réelle et une adresse pour l'équilibrage de charge.

Remarque : Le composant Content Based Routing (CBR) est disponible sur toutes les plateformes prises en charge à l'exception de celles exécutant une JVM 64 bits. Vous avez également la possibilité d'employer la méthode d'acheminement `cbr` du composant Dispatcher de Load Balancer pour permettre le routage CBR sans faire appel à Caching Proxy. Pour plus de détails, voir «Fonction CBR de Dispatcher (méthode d'acheminement `cbr`)», à la page 55.

Pour pouvoir utiliser CBR, vous devez installer module Caching Proxy sur le même serveur. Pour configurer Caching Proxy pour CBR, voir «Etape 1. Configuration de Caching Proxy pour utiliser CBR», à la page 114.

Préparation

1. Pour cet exemple, configurez les postes de travail sur le même segment de réseau local. Vérifiez que le trafic réseau entre les trois machines n'a pas à traverser de routeurs ou de ponts.
2. Configurez les cartes réseau de ces trois postes de travail. Dans cet exemple, nous supposons que vous disposez de la configuration réseau suivante :

Poste de travail	Nom	Adresse IP
1	server1.monsiteweb.com	9.27.27.101
2	server2.monsiteweb.com	9.27.27.102
3	server3.monsiteweb.com	9.27.27.103
Masque réseau = 255.255.255.0		

Chaque poste de travail ne contient qu'une carte d'interface réseau Ethernet standard.

3. Vérifiez que server1.monsiteweb.com peut contacter server2.monsiteweb.com et server3.monsiteweb.com (avec la commande ping).
4. Vérifiez que server2.monsiteweb.com et server3.monsiteweb.com peuvent contacter server1.monsiteweb.com (avec la commande ping).
5. Vérifiez que les serveurs Web de server2.monsiteweb.com et server3.monsiteweb.com sont opérationnels. Utilisez un navigateur Web pour accéder directement aux pages à partir de **http://server2.monsiteweb.com** (par exemple, .../member/index.html) et **http://server3.monsiteweb.com** (par exemple, .../guest/index.html).
6. Cherchez une autre adresse IP valide pour ce segment de réseau local. Il s'agit de l'adresse de cluster que vous fournirez aux clients qui souhaitent accéder à votre site. Dans cet exemple, nous utiliserons :

Nom= www.monsiteweb.com
IP=9.27.27.104

Configuration du composant CBR

A l'aide de CBR, vous pouvez créer une configuration à l'aide de la ligne de commande, de l'assistant de configuration ou de l'interface graphique. Pour cet exemple de démarrage rapide, les étapes de configuration s'effectuent via la ligne de commande.

Remarque : Les valeurs des paramètres doivent être saisies à l'aide de caractères anglais. Les seules exceptions sont les valeurs des paramètres des noms d'hôte et des noms de fichiers.

Configuration à partir de la ligne de commande

A partir d'une invite, effectuez les opérations ci-dessous.

1. Démarrez cbrserver. Exécutez la commande suivante en tant que superutilisateur ou administrateur : **cbrserver**

Remarque : Pour la plateforme Windows : Démarrez cbrserver (Content Based Routing) à partir du panneau Services : **Démarrer > Paramètres** (pour Windows 2000) > **Panneau de configuration > Outils d'administration > Services**.

2. Lancez la fonction exécuteur (executor) de CBR :
cbrcontrol executor start

3. Démarrez le module Caching Proxy (ce module peut être démarré à tout moment après démarrage de la fonction exécuteur) :

ibmproxy

Remarque : Pour Windows : Vous pouvez également démarrer Caching Proxy à partir du panneau Services : **Démarrer > Paramètres** (pour Windows 2000)> **Panneau de configuration > Outils d'administration > Services.**

4. Ajoutez le cluster (nom d'hôte, site Web, auquel les clients se connectent) à la configuration CBR :

cbrcontrol cluster add www.monsiteweb.com

5. Ajoutez l'adresse de cluster (9.27.27.104) du site Web à la carte d'interface réseau sur la machine CBR. Pour plus d'informations, voir «Etape 5. Affectation d'un alias à la carte d'interface réseau (facultatif)», à la page 116.

6. Ajoutez le port du protocole http à la configuration CBR :

cbrcontrol port add www.monsiteweb.com:80

7. Ajoutez chaque serveur Web à la configuration CBR :

cbrcontrol server add www.monsiteweb.com:80:server2.monsiteweb.com

cbrcontrol server add www.monsiteweb.com:80:server3.monsiteweb.com

8. Ajoutez des règles de contenu à la configuration CBR. (Une règle de contenu définit la manière dont une requête d'URL sera reconnue et envoyée à l'un des serveurs ou des ensembles de serveurs) :

cbrcontrol rule add www.monsiteweb.com:80:memberRule type content pattern uri=*/member/*

cbrcontrol rule add www.monsiteweb.com:80:guestRule type content pattern uri=*/guest/*

Dans cet exemple, l'utilisation de la règle de contenu permet d'envoyer les demandes des clients adressées au site Web www.monsiteweb.com vers un autre serveur en fonction d'un répertoire désigné dans leur chemin de requête d'URL. Pour plus d'informations, voir Annexe B, «Syntaxe des règles de contenu (modèle)», à la page 477.

9. Ajoutez des serveurs à vos règles :

cbrcontrol rule useserver www.monsiteweb.com:80:memberRule server2.monsiteweb.com

cbrcontrol rule useserver www.monsiteweb.com:80:guestRule server3.monsiteweb.com

CBR procède maintenant à l'équilibrage de charge en fonction d'une règle de contenu. Un client dont la demande d'URL contient **/member/** sera dirigé vers server2.monsiteweb.com. Un client dont la demande d'URL contient **/guest/** sera dirigé vers server3.monsiteweb.com.

10. Démarrez la fonction gestionnaire (manager) de CBR :

cbrcontrol manager start

11. Démarrez la fonction conseiller (advisor) de CBR :

cbrcontrol advisor start http 80

CBR vérifie désormais que les demandes des clients ne sont pas envoyées vers un serveur Web arrêté.

La configuration de base comportant des serveurs liés en local est maintenant terminée.

Test de vérification de la configuration

Vérifiez que la configuration fonctionne :

1. A l'aide d'un navigateur Web, accédez à <http://www.monsiteweb.com/member/index.htm> . Si une page s'affiche, la configuration fonctionne.
2. Rechargez la page dans le navigateur Web.
3. Observez les résultats de la commande suivante :

```
cbrcontrol server report  
www.monsiteweb.com:80:
```

La colonne du nombre total de connexions des deux serveurs doit contenir la valeur "2."

Configuration à l'aide de l'interface graphique

Pour plus d'informations sur l'utilisation de l'interface graphique de CBR, voir «Interface graphique», à la page 112 et à l'Annexe A, «Interface graphique utilisateur : Instructions générales», à la page 469.

Configuration à l'aide de l'assistant de configuration

Pour plus d'informations sur l'utilisation de l'assistant de CBR, voir «Assistant de configuration», à la page 113.

Types de configurations de cluster, de port et de serveur

La configuration de CBR pour assurer le support de votre site peut s'effectuer de plusieurs manières. Si votre site ne comprend qu'un seul nom de système hôte auquel tous vos clients se connectent, vous pouvez ne définir qu'un seul cluster de serveurs. Pour chaque serveur, configurez un port par l'intermédiaire duquel CBR communique. Voir figure 9, à la page 48.

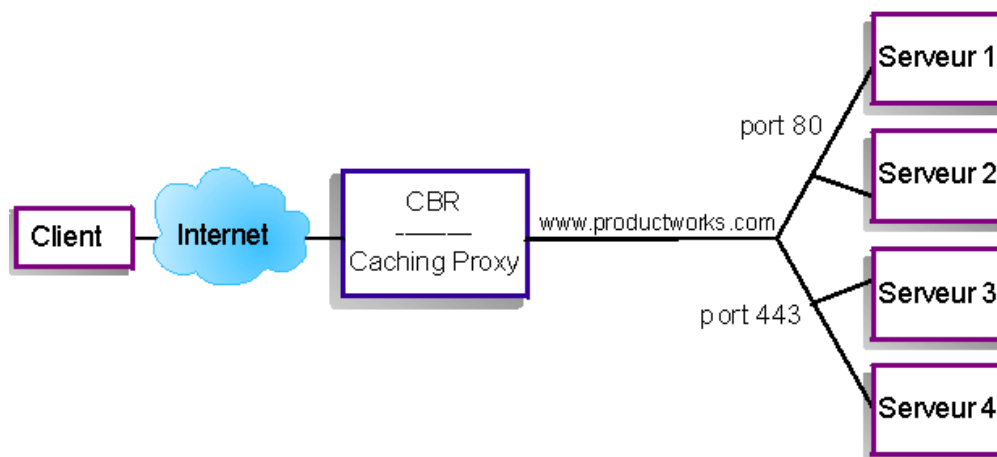


Figure 17. Exemple de composant CBR configuré avec un cluster et 2 ports

Dans cet exemple de composant CBR, un cluster est défini sur www.productworks.com. Il dispose de deux ports : le port 80 pour HTTP et le port 443 pour SSL. Un client adressant une requête à l'adresse <http://www.productworks.com> (port 80) accédera à un autre serveur qu'un client s'adressant à <https://www.productworks.com> (port 443).

Si le site est très étendu et qu'il comporte un grand nombre de serveurs, chacun étant dédié à un protocole en particulier, CBR doit être configuré selon une autre méthode. Dans ce dernier cas, il est souhaitable de définir un cluster pour chaque protocole, avec un seul port mais plusieurs serveurs, comme illustré à la figure 10, à la page 49.

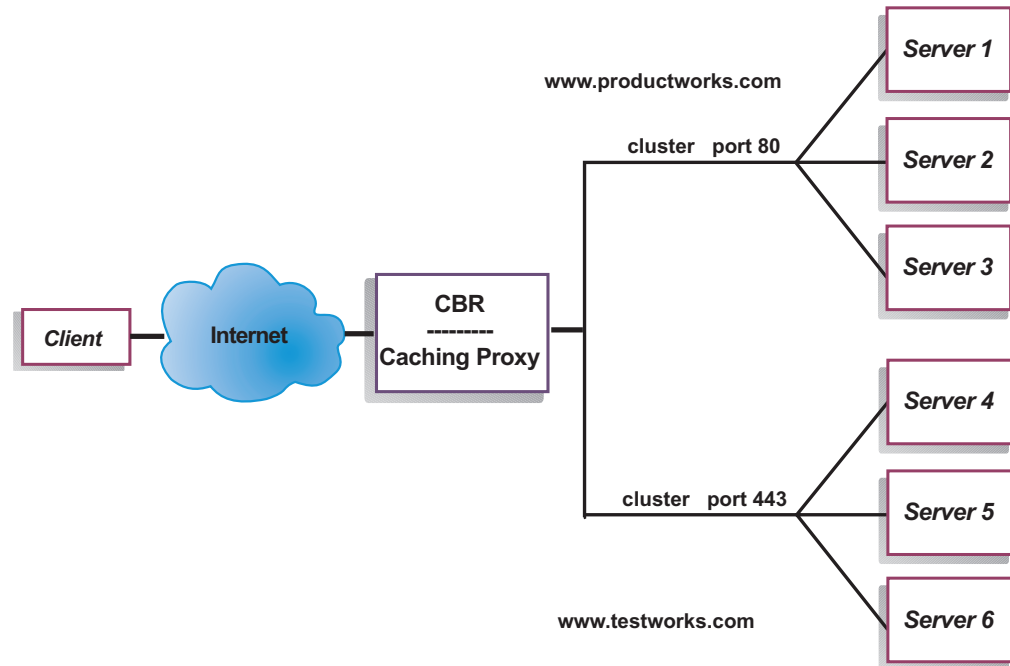


Figure 18. Exemple de composant CBR configuré avec deux clusters, chacun étant associé à un port

Dans cet exemple de composant CBR, deux clusters sont définis : `www.productworks.com` pour le port 80 (HTTP) et `www.testworks.com` pour le port 443 (SSL).

Une troisième configuration de CBR est nécessaire si votre site abrite plusieurs sociétés ou services, chacun accédant à votre site par une adresse URL distincte. Dans ce cas, vous pouvez définir un cluster pour chaque société ou service ainsi qu'un nombre de ports variable pour réceptionner les connexions de cette URL, comme illustré par la figure 11, à la page 50.

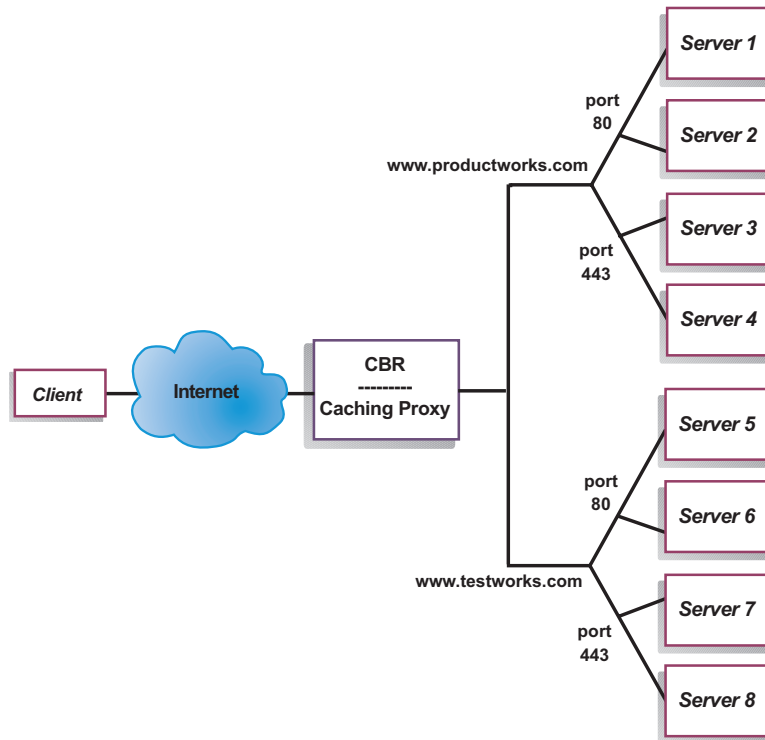


Figure 19. Exemple de composant CBR configuré avec 2 clusters, chacun étant associé à 2 ports

Dans cet exemple de composant CBR, deux clusters sont définis avec le port 80 (HTTP) et le port 443 (SSL) pour chacun des sites www.productworks.com et www.testworks.com.

Chapitre 10. Planification de CBR (Content Based Routing)

Le présent chapitre décrit les aspects que l'administrateur réseau doit prendre en compte avant d'installer et de configurer le composant CBR avec Caching Proxy.

- Voir Chapitre 3, «Gestion du réseau : Fonctions Load Balancer requises», à la page 19 pour une présentation des fonctions permettant de gérer votre réseau.
- Voir Chapitre 11, «Configuration de CBR (Content Based Routing)», à la page 109 pour obtenir des informations sur la configuration des paramètres d'équilibrage de charge du composant CBR.
- Voir Chapitre 22, «Fonctions avancées de Dispatcher, CBR et Site Selector», à la page 201 pour obtenir des informations sur la configuration de Load Balancer pour les fonctions avancées.
- Voir Chapitre 24, «Exploitation et gestion de Load Balancer», à la page 261, pour obtenir des informations sur l'administration authentifiée à distance, les fichiers journaux Load Balancer et sur l'utilisation des composants Load Balancer.

Le présent chapitre contient la section suivante :

- «Remarques relatives à la planification»
- «Équilibrage de charge basé sur des règles avec CBR», à la page 107
- «Équilibrage de charge sur les connexions sécurisées (SSL)», à la page 107
- «Équilibrage de charge client-proxy dans SSL et proxy-serveur dans HTTP», à la page 108

Remarques relatives à la planification

Le composant CBR permet d'équilibrer la charge du trafic HTTP et SSL à l'aide de Caching Proxy qui permet de transmettre la demande par un serveur proxy. CBR permet d'équilibrer la charge des serveurs configurés à partir de votre fichier de configuration CBR à l'aide des commandes cbrcontrol.

Remarque : Le composant Content Based Routing (CBR) est disponible sur toutes les plateformes prises en charge à l'exception de celles exécutant une JVM 64 bits. Vous avez également la possibilité d'employer la méthode d'acheminement cbr du composant Dispatcher de Load Balancer pour permettre le routage CBR sans faire appel à Caching Proxy. Pour plus de détails, voir «Fonction CBR de Dispatcher (méthode d'acheminement cbr)», à la page 55.

La structure de CBR ressemble beaucoup à celle de Dispatcher. CBR comprend les fonctions suivantes :

- **cbrserver** traite les demandes à partir de la ligne de commande adressées à l'exécuteur, au gestionnaire et aux conseillers.
- L'**exécuteur** prend en charge l'équilibrage de charge des demandes client. Vous devez démarrer l'exécuteur pour pouvoir utiliser le composant CBR.
- Le **gestionnaire** définit les mesures utilisées par l'exécuteur en fonction de plusieurs facteurs :
 - les décomptes internes de l'exécuteur,
 - le retour d'informations sur les serveurs fourni par les conseillers,

- le retour d’informations émanant d’un programme de contrôle système, tel que Metric Server.

L’utilisation du gestionnaire n’est que facultative. Toutefois, s’il n’est pas utilisé, l’équilibrage de charge se fait sur la base d’une planification circulaire pondérée, elle-même basée sur les mesures de charge des serveurs et les conseillers ne seront pas disponibles.

- Les **conseillers** interrogent les serveurs et analysent les résultats par protocole avant d’appeler le gestionnaire pour définir les pondérations comme il convient. L’utilisation de certains de ces conseillers n’est peut-être pas utile dans une configuration typique. Vous avez également la possibilité de développer vos propres conseillers. L’utilisation des conseillers est facultative mais recommandée. Load Balancer fournit un conseiller Caching Proxy (cachingproxy). Pour plus d’informations, voir «Conseillers», à la page 185.
- Pour configurer et gérer l’exécuteur, les conseillers et le gestionnaire, utilisez la ligne de commande (**cbrcontrol**) ou l’interface utilisateur graphique (**lbadmin**).

Les trois fonctions clés de CBR (l’exécuteur, le gestionnaire et les conseillers) agissent en collaboration pour équilibrer et répartir entre les serveurs les requêtes réceptionnées. Outre la gestion des requêtes d’équilibrage de charge, l’exécuteur contrôle le nombre de nouvelles connexions et de connexions actives, et transmet ces informations au gestionnaire.

Equilibrage de la charge des requêtes pour différents types de contenus

CBR vous permet de spécifier un ensemble de serveurs devant prendre en charge une demande client en fonction de son contenu. Le composant CBR vous permet de compartimenter votre site en plusieurs parties, chacune pouvant être traitée par des ensembles de serveurs différents. Ce partitionnement est transparent pour les clients accédant à votre site.

Division du contenu de votre site pour améliorer le temps de réponse

Vous pouvez répartir votre site en affectant à certains serveurs le traitement de requêtes cgi uniquement, et en affectant à un autre ensemble de serveurs le traitement de toutes les autres requêtes. Ceci mettrait fin au ralentissement de l’activité des serveurs dû au calcul d’énormes scripts cgi au cours d’un trafic HTML normal, et permettrait ainsi aux clients d’obtenir de meilleurs temps de réponse. Avec cette méthode, vous pouvez également utiliser des postes de travail plus puissants pour des requêtes normales. Ainsi, les clients obtiendraient un meilleur temps de réponse sans pour autant occasionner des frais de mise à niveau de tous vos serveurs. Vous pouvez également affecter des postes de travail plus puissants pour des requêtes cgi.

Vous pouvez également partitionner votre site en dirigeant vers un ensemble de serveurs les clients qui accèdent à des pages nécessitant une opération d’enregistrement, et en acheminant toutes les autres requêtes vers un deuxième ensemble de serveurs. Ainsi, les navigateurs occasionnels qui accèdent à votre site n’accapareront plus les ressources qui pourraient être utilisées par des clients devant effectuer des opérations d’enregistrement sur votre site. Cela vous permettrait également d’utiliser des postes de travail plus puissants pour traiter les clients qui se sont enregistrés.

Il est possible de combiner les deux pour plus de souplesse et pour un meilleur service.

Copie de sauvegarde du contenu du serveur Web

CBR vous permet d'indiquer plusieurs serveurs pour chaque type de requête. Par conséquent, les requêtes peuvent être équilibrées pour obtenir une réponse optimale du client. L'affectation de plusieurs serveurs à chaque partie de votre site vous permet de vous protéger en cas de défaillance d'un poste de travail ou d'un serveur. CBR reconnaîtra la défaillance et continuera d'équilibrer la charge des requêtes client aux autres serveurs du groupe.

Utilisation de plusieurs processus Caching Proxy pour optimiser l'utilisation de la CPU

Caching Proxy communique avec un processus CBR via son interface de plug-in. Le processus CBR doit s'exécuter sur la machine locale pour ce travail. Ces deux processus étant distincts, plusieurs instances Caching Proxy peuvent s'exécuter et travailler avec une seule instance de processus CBR. Vous pouvez adopter ce type de configuration pour isoler des adresses ou des fonctions entre les divers processus Caching Proxy ou pour optimiser l'utilisation des ressources de la machine en définissant plusieurs processus Caching Proxy en charge du trafic client. Les instances proxy sont à l'écoute sur différents ports ou en liaison avec des adresses IP uniques sur le même port, selon les besoins du trafic.

Équilibrage de charge basé sur des règles avec CBR

CBR et Caching Proxy examinent les requêtes HTTP à l'aide de types de règle indiqués. Pendant l'exécution, Caching Proxy accepte les demandes client et interroge le composant CBR pour savoir quel est le meilleur serveur. Lorsqu'il reçoit cette demande, CBR la compare à un ensemble de règles prioritaires. Dès qu'il en trouve une qui correspond, un serveur approprié est sélectionné dans un ensemble de serveurs préconfigurés. Enfin, CBR indique à Caching Proxy le serveur sélectionné, et les demandes sont transmises à ce dernier.

Une fois que vous avez défini un cluster pour la répartition de charge, assurez-vous que toutes les requêtes envoyées à ce cluster ont une règle qui choisira un serveur. Si aucune règle correspondant à une requête spécifique n'est trouvée, Caching Proxy enverra une page d'erreur au client. Le moyen le plus facile pour s'assurer que toutes les demandes correspondront à une règle est de créer une règle "toujours vraie" avec un niveau de priorité élevé. Vérifiez que les serveurs auxquels se réfère cette règle peuvent traiter toutes les demandes non gérées explicitement par les règles ayant des niveaux de priorité moins élevés. (Remarque : Les règles de priorité inférieure sont évaluées en premier.)

Pour plus d'informations, voir «Configuration de l'équilibrage de charge basé sur des règles», à la page 212.

Équilibrage de charge sur les connexions sécurisées (SSL)

CBR et Caching Proxy peuvent recevoir une transmission SSL d'un client vers le proxy (côté client-serveur) ainsi que prendre en charge une transmission d'un proxy vers un serveur SSL (côté proxy-serveur). Si vous définissez un port SSL sur un serveur dans la configuration CBR pour qu'il reçoive la demande SSL provenant d'un client, vous pouvez gérer un site complètement sécurisé, en utilisant CBR pour équilibrer la charge entre les serveurs sécurisés SSL.

En plus des autres modifications du fichier `ibmproxy.conf` pour CBR, une instruction de configuration doit être ajoutée au fichier `ibmproxy.conf` pour que Caching Proxy active le chiffrement SSL du proxy vers le serveur. Le format est le suivant :

```
proxy uri_structure url_structure adresse
```

où *uri_structure* correspond à la structure à respecter (par exemple : `/secure/*`), *url_structure* à un URL de remplacement (par exemple : `https://clusterA/secure/*`) et *adresse* à l'adresse du cluster (par exemple : `clusterA`).

Equilibrage de charge client-proxy dans SSL et proxy-serveur dans HTTP

CBR et Caching Proxy peuvent également recevoir une transmission SSL d'un client et déchiffrer la demande SSL avant d'acheminer la demande par proxy à un serveur HTTP. Pour que CBR prenne en charge la transmission client-proxy pour SSL et proxy-client pour HTTP, utilisez le mot clé facultatif **mapport** dans la commande `cbrcontrol server`. Il permet d'indiquer si le port du serveur est différent du port d'entrée du client. Voici un exemple d'ajout de port avec le mot clé **mapport**, dans lequel le port du client est 443 (SSL) et le port du serveur est 80 (HTTP) :

```
cbrcontrol server add cluster:443 mapport 80
```

Le numéro de port de **mapport** peut correspondre à n'importe quel entier positif. La valeur par défaut correspond au numéro de port entrant du client.

Etant donné que CBR doit être capable de traiter une demande HTTP pour un serveur configuré sur le port 443 (SSL), un conseiller spécial *ssl2http* est fourni. Il démarre sur le port 443 (le port entrant du client) et opère sur le ou les serveurs configurés pour ce port. Si deux clusters sont configurés et que pour chacun d'entre eux, le port 443 et les serveurs sont configurés avec un paramètre **mapport** différent, une seule instance du conseiller peut ouvrir le port approprié. Voici un exemple de cette configuration :

```
Executor
  Cluster1
    Port:443
      Server1 mapport 80
      Server2 mapport 8080
  Cluster2
    Port:443
      Server3 mapport 80
      Server4 mapport 8080
  Manager
    Advisor ssl2http 443
```

Chapitre 11. Configuration de CBR (Content Based Routing)

Avant d'effectuer les opérations décrites dans le présent chapitre, voir Chapitre 10, «Planification de CBR (Content Based Routing)», à la page 105. Ce chapitre décrit comment créer une configuration de base pour le composant CBR de Load Balancer.

- Voir Chapitre 21, «Gestionnaire, conseillers et système Metric Server (des composants Dispatcher, CBR et Site Selector)», à la page 179 et Chapitre 22, «Fonctions avancées de Dispatcher, CBR et Site Selector», à la page 201 pour plus d'informations sur des configurations plus complexes de Load Balancer.
- Voir Chapitre 24, «Exploitation et gestion de Load Balancer», à la page 261, pour obtenir des informations sur l'administration authentifiée à distance, les fichiers journaux Load Balancer et sur l'utilisation des composants Load Balancer.

Présentation générale des tâches de configuration

Avant de suivre les étapes de configuration détaillées dans ce tableau, assurez-vous que le poste CBR et tous les postes serveurs sont connectés au réseau, que leurs adresses IP sont valides et qu'ils peuvent communiquer entre eux par ping.

Remarque : Le composant Content Based Routing (CBR) est disponible sur toutes les plateformes prises en charge à l'exception de celles exécutant une JVM 64 bits. Vous avez également la possibilité d'employer la méthode d'acheminement cbr du composant Dispatcher de Load Balancer pour permettre le routage CBR sans faire appel à Caching Proxy. Pour plus de détails, voir «Fonction CBR de Dispatcher (méthode d'acheminement cbr)», à la page 55.

Tableau 7. Tâches de configuration pour le composant CBR

Tâche	Description	Informations connexes
Configurer le poste CBR	Conditions requises	«Configuration du poste CBR», à la page 113
Configuration des machines en vue de l'équilibrage de charge.	Définition de la configuration de l'équilibrage de charge.	«Etape 7. Définition des serveurs avec équilibrage de charge», à la page 117

Méthodes de configuration

Quatre méthodes permettent de créer une configuration de base du composant CBR de Load Balancer :

- Ligne de commande
- Scripts
- Interface graphique
- Assistant de configuration

Pour utiliser le composant CBR, Caching Proxy doit être installé.

Remarque : Caching Proxy est un service qui, par défaut, démarre automatiquement après l'installation. Vous devez l'arrêter avant de

lancer la fonction serveur CBR (cbrserver) et modifier le service Caching Proxy pour le démarrer manuellement et non automatiquement.

- Pour les systèmes Linux ou UNIX : Pour arrêter Caching Proxy, recherchez l'identificateur de processus correspondant à l'aide de la commande `ps -ef | grep ibmproxy`, puis mettez fin à ce processus à l'aide de la commande `kill id_processus`.
- Pour les systèmes Windows : Arrêtez Caching Proxy à partir du panneau Services.

Ligne de commande

C'est la méthode de configuration de CBR la plus directe. Les valeurs des paramètres de commandes doivent être saisies à l'aide de caractères anglais. Les seules exceptions s'appliquent aux noms d'hôte (utilisés dans les commandes cluster et server) et aux noms de fichiers.

Pour démarrer CBR à partir de la ligne de commande, procédez aux opérations ci-dessous.

- Sur les systèmes Linux ou UNIX : En tant que superutilisateur, exécutez la commande **cbrserver** à partir de l'invite. (Pour arrêter le service, entrez la commande **cbrserver stop**.)

Sur les systèmes Windows : Cliquez sur **Démarrer** > **Paramètres** (pour Windows 2000) > **Panneau de configuration** > **Outils d'administration** > **Services**.

Cliquez à l'aide du bouton droit de la souris sur **IBM Content Based Routing** et sélectionnez **Démarrer**. Pour arrêter le service, suivez la même procédure en sélectionnant **Arrêter**.

- Ensuite, émettez les commandes de contrôles CBR souhaitées pour définir votre configuration. Les procédures décrites dans ce manuel reposent sur l'utilisation de la ligne de commande. La commande est **cbrcontrol**. Pour plus de détails sur les commandes, voir Chapitre 27, «Guide des commandes Dispatcher et CBR», à la page 345.
- Démarrez Caching Proxy. Entrez la commande **ibmproxy** à partir de l'invite. (Vous devez lancer l'exécuteur avant Caching Proxy.)

Remarque : Pour les plateformes Windows : Démarrez Caching Proxy à partir du panneau Services : **Démarrer** > **Paramètres** (pour Windows 2000) > **Panneau de configuration** > **Outils d'administration** > **Services**.

Vous pouvez entrer une version abrégée des paramètres de contrôle **cbrcontrol**. Il suffit d'entrer les lettres spécifiques des paramètres. Par exemple, pour obtenir l'aide correspondant à la commande `file save`, vous pouvez entrer **cbrcontrol he f** au lieu de **cbrcontrol help file**.

Pour démarrer l'interface de ligne de commande, entrez **cbrcontrol** pour ouvrir une invite **cbrcontrol**.

Pour fermer l'interface de ligne de commande, entrez **exit** ou **quit**.

Remarques :

1. Sous Windows, le service dsserver du composant Dispatcher démarre automatiquement. Si vous utilisez uniquement CBR et non le composant Dispatcher, vous pouvez empêcher dsserver de démarrer automatiquement de la manière suivante :

- a. Dans la fenêtre Services, cliquez à l'aide du bouton droit de la souris sur IBM Dispatcher.
 - b. Sélectionnez Propriétés.
 - c. Dans la zone **Type de démarrage**, sélectionnez Manuel.
 - d. Cliquez sur OK et fermez la fenêtre Services.
2. Lorsque vous configurez CBR (Content Based Routing) à partir de l'invite du système d'exploitation et non à partir de l'invite `cbrcontrol>>`, prenez soin d'utiliser les caractères suivants :

() parenthèses ouvrante et fermante
 & perluète
 | barre
 ! point d'exclamation
 * astérisque

Le shell du système d'exploitation peut interpréter ces caractères comme des caractères spéciaux et les convertir en texte de remplacement avant leur évaluation par `cbrcontrol`.

Les caractères spéciaux de la liste précédente sont facultatifs dans la commande **`cbrcontrol rule add`**. Ils sont employés lors de l'indication d'un motif pour une règle de contenu. Par exemple, la commande suivante ne peut être valide qu'avec l'invite `cbrcontrol>>`.

```
rule add 10.1.203.4:80:cbr_prod_rule_ek type content
pattern uri=/nipoek/*
```

Pour que cette commande fonctionne à partir de l'invite du système d'exploitation, placez le motif entre guillemets (" ") comme suit :

```
cbrcontrol rule add 10.1.203.4:80:cbr_prod_rule_ek type content
pattern "uri=/nipoek/*"
```

Si vous omettez les guillemets, le motif sera peut-être tronqué lors de la sauvegarde de la règle dans CBR. Les guillemets ne sont pas pris en charge avec l'invite `cbrcontrol>>`.

Scripts

Vous pouvez entrer les commandes de configuration CBR dans un fichier script de configuration pour les exécuter simultanément.

Remarque : Pour exécuter rapidement le contenu d'un fichier script (par exemple, `mon_script`), utilisez l'une des commandes suivantes :

- Pour mettre à jour la configuration actuelle, soumettez les commandes exécutables suivantes à partir du fichier script :
`cbrcontrol file appendload mon_script`
- Pour remplacer la configuration actuelle, soumettez les commandes exécutables suivantes à partir du fichier script :
`cbrcontrol file newload mon_script`

Pour sauvegarder la configuration en cours dans un fichier script (par exemple, `savescript`), exécutez la commande suivante :

```
cbrcontrol file save savescript
```

Cette commande enregistre le fichier script de configuration dans le répertoire **`...ibm/edge/lb/servers/configurations/cbr`**.

Interface graphique

Pour des instructions générales et un exemple de l'interface graphique, voir figure 41, à la page 470.

Pour démarrer l'interface graphique, procédez comme suit :

1. Vérifiez que `cbrserver` fonctionne. En tant que superutilisateur ou administrateur, entrez la commande **`cbrserver`** à partir d'une invite.
2. Exécutez l'une des actions suivantes, selon votre système d'exploitation :
 - Pour les systèmes AIX, HP-UX, Linux ou Solaris : entrez **`lbadmin`**
 - Pour les systèmes Windows : cliquez sur **Démarrer > Programmes > IBM WebSphere > Edge Components > IBM Load Balancer > Load Balancer**
3. Démarrez Caching Proxy. (A partir de l'interface graphique, vous devez d'abord vous connecter à l'hôte et lancer l'exécuteur pour le composant CBR avant de démarrer Caching Proxy.) Effectuez l'une des opérations suivantes :
 - Pour les systèmes AIX, HP-UX, Linux ou Solaris : Pour lancer Caching Proxy, entrez **`ibmproxy`**
 - Pour les systèmes Windows : Pour lancer Caching Proxy, accédez au panneau Services : **Démarrer > Paramètres** (pour Windows 2000) > **Panneau de configuration > Outils d'administration > Services**

Pour pouvoir configurer le composant CBR à partir de l'interface graphique, vous devez d'abord sélectionner **Content Based Routing** dans l'arborescence. Vous pouvez lancer le gestionnaire une fois que vous vous êtes connecté à un hôte. Vous pouvez également créer des clusters contenant des ports et des serveurs, puis lancer des conseillers pour le gestionnaire.

Vous pouvez utiliser l'interface graphique pour toute opération exécutée habituellement par la commande **`cbrcontrol`**. Par exemple, pour définir un cluster à l'aide de la ligne de commande, entrez la commande **`cbrcontrol cluster add cluster`**. Pour définir un cluster à partir de l'interface graphique, cliquez à l'aide du bouton droit de la souris sur Exécuteur, puis dans le menu en incrustation, sélectionnez **Ajout d'un cluster**. Entrez l'adresse du cluster dans la fenêtre en incrustation, puis cliquez sur **OK**.

Les fichiers de configuration CBR existants peuvent être chargés à l'aide des options **Chargement de la nouvelle configuration** (pour remplacer intégralement la configuration en cours) et **Ajout à la configuration en cours** (pour mettre à jour la configuration en cours) du menu en incrustation **Hôte**. Vous devez sauvegarder régulièrement votre configuration CBR dans un fichier, en utilisant l'option **Sauvegarder le fichier de configuration en...** du menu en incrustation **Hôte**. Le menu **Fichier** situé en haut de l'interface graphique permet de sauvegarder les connexions à l'hôte en cours dans un fichier ou de restaurer les connexions dans des fichiers existants sur tous les composants Load Balancer.

Vous pouvez accéder à l'**Aide** en cliquant sur le point d'interrogation situé dans l'angle supérieur droit de la fenêtre Load Balancer.

- **Aide sur les zones** — décrit les valeurs par défaut de chaque zone.
- **Procédures** — affiche la liste des tâches pouvant être effectuées dans cet écran.
- **Centre de documentation** — fournit un accès centralisé aux informations relatives au produit

Pour exécuter une commande à partir de l'interface graphique : mettez le noeud Hôte en surbrillance dans l'arborescence de l'interface graphique, puis sélectionnez

Envoyer la commande... dans le menu en incrustation Hôte. Dans la zone d'entrée de commande, entrez la commande à exécuter, par exemple **executor report**. Les résultats et l'historique des commandes exécutées lors de la session courante s'affichent dans la fenêtre ouverte.

Pour plus de détails sur l'utilisation de l'interface graphique, voir Annexe A, «Interface graphique utilisateur : Instructions générales», à la page 469.

Assistant de configuration

Si vous utilisez l'assistant de configuration, suivez la procédure ci-dessous.

1. Démarrez cbrserver : entrez la commande **cbrserver** à partir de l'invite en tant que superutilisateur ou administrateur.

2. Démarrez la fonction Assistant de CBR.

Pour ce faire, démarrez l'assistant à partir de l'invite en entrant **cbrwizard**. Ou alors, sélectionnez l'assistant de configuration dans le menu des composants CBR proposé par l'interface graphique.

3. Démarrez Caching Proxy pour équilibrer la charge du trafic HTTP ou HTTPS (SSL).

Pour les systèmes AIX, HP-UX, Linux ou Solaris : Pour lancer Caching Proxy, entrez **ibmproxy**

Pour les systèmes Windows : Pour lancer Caching Proxy, accédez au panneau Services : **Démarrer > Paramètres** (pour Windows 2000) > **Panneau de configuration > Outils d'administration > Services**

Cet assistant vous guide, pas à pas, pendant la création d'une configuration de base du composant CBR. Il vous demande des renseignements sur votre réseau et vous guide pendant l'installation d'un cluster permettant à CBR d'équilibrer la charge du trafic d'un groupe de serveurs.

Configuration du poste CBR

La configuration de la machine CBR ne peut être effectuée que par le superutilisateur root (pour les systèmes AIX, HP-UX, Linux ou Solaris) ou l'administrateur (pour les systèmes Windows).

Vous aurez besoin d'une adresse IP pour chaque cluster de serveurs configuré. Une adresse de cluster est une adresse associée à un nom de système hôte (par exemple `www.société_X.com`). Cette adresse IP est utilisée par un client pour se connecter aux serveurs du cluster en question. Cette adresse se trouve dans la requête URL du client. Toutes les requêtes envoyées à la même adresse de cluster font l'objet d'un équilibrage de charge par CBR.

Pour les systèmes Solaris uniquement : Pour pouvoir utiliser le composant CBR, vous devez modifier les valeurs système par défaut attribuées aux communications IPC (Inter-process Communication). Vous devez augmenter la taille maximale du segment de mémoire partagée et le nombre d'identificateurs de sémaphores. Pour configurer la prise en charge de CBR, ajoutez les instructions suivantes dans le fichier `/etc/system`, puis réamorcer le système :

```
set shmsys:shminfo_shmmax=0x02000000
set semsys:seminfo_semmap=750
set semsys:seminfo_semmni=30
set semsys:seminfo_semmns=750
set semsys:seminfo_semmnu=30
set semsys:seminfo_semtime=30
```


Si vous n'attribuez pas au segment de mémoire partagée les valeurs ci-dessus, la commande **cbrcontrol executor start** échouera.

Etape 1. Configuration de Caching Proxy pour utiliser CBR

Pour utiliser le composant CBR, Caching Proxy doit être installé.

Remarque : Caching Proxy est un service qui, par défaut, démarre automatiquement après l'installation. Vous devez l'arrêter avant de lancer la fonction serveur CBR et modifier le service Caching Proxy pour le démarrer manuellement et non automatiquement.

- Pour les systèmes AIX, HP-UX, Linux et Solaris : Pour arrêter Caching Proxy, recherchez l'identificateur de processus correspondant à l'aide de la commande `ps -ef | grep ibmproxy`, puis mettez fin à ce processus à l'aide de la commande `kill id_processus`.
- Pour les systèmes Windows : Arrêtez Caching Proxy à partir du panneau Services.

Apportez les modifications ci-dessous au fichier de configuration Caching Proxy (ibmproxy.conf) :

Vérifiez que la directive d'URL entrante **CacheByIncomingUrl** a la valeur "off" (valeur par défaut).

Dans la section des règles de mappage du fichier de configuration, ajoutez pour chaque cluster une règle de mappage du type suivant :

```
Proxy /* http://cluster.domain.com/* cluster.domain.com
```

Remarque : CBR définit le protocole, le serveur et le port cible ultérieurement.

Quatre entrées doivent être modifiées pour le plug-in CBR :

- ServerInit
- PostAuth
- PostExit
- ServerTerm

Chaque entrée doit figurer sur une ligne. Le fichier ibmproxy.conf comporte plusieurs instances de "ServerInit" (une par plug-in). Les entrées relatives au plug-in CBR doivent être modifiées et ne doivent comporter aucun commentaire.

Les ajouts spécifiques au fichier de configuration de chaque système d'exploitation sont répertoriés ci-dessous.

Figure 20. Fichier de configuration CBR pour les systèmes AIX, Linux et Solaris

```
ServerInit /opt/ibm/edge/lb/servers/lib/liblbcbcr.so:ndServerInit
PostAuth /opt/ibm/edge/lb/servers/lib/liblbcbcr.so:ndPostAuth
PostExit /opt/ibm/edge/lb/servers/lib/liblbcbcr.so:ndPostExit
ServerTerm /opt/ibm/edge/lb/servers/lib/liblbcbcr.so:ndServerTerm
```


Figure 21. Fichier de configuration CBR pour les systèmes HP-UX

```
ServerInit /opt/ibm/edge/lb/servers/lib/liblbcbr.sl:ndServerInit
PostAuth /opt/ibm/edge/lb/servers/lib/liblbcbr.sl:ndPostAuth
PostExit /opt/ibm/edge/lb/servers/lib/liblbcbr.sl:ndPostExit
ServerTerm /opt/ibm/edge/lb/servers/lib/liblbcbr.sl:ndServerTerm
```

Figure 22. Fichier de configuration CBR pour les systèmes Windows

```
ServerInit C:\Program Files\IBM\edge\lb\servers\lib\liblbcbr.dll:ndServerInit
PostAuth C:\Program Files\IBM\edge\lb\servers\lib\liblbcbr.dll:ndPostAuth
PostExit C:\Program Files\IBM\edge\lb\servers\lib\liblbcbr.dll:ndPostExit
ServerTerm C:\Program Files\IBM\edge\lb\servers\lib\liblbcbr.dll:ndServerTerm
```

Etape 2. Démarrage de la fonction serveur

Pour démarrer la fonction serveur CBR, entrez **cbrserver** sur la ligne de commande.

Un fichier de configuration par défaut (default.cfg) est chargé automatiquement lors du démarrage de cbrserver. Si vous sauvegardez la configuration CBR dans default.cfg, toutes les données enregistrées dans ce fichier sont automatiquement chargées au prochain démarrage de cbrserver.

Etape 3. Démarrage de la fonction exécuteur

Pour démarrer la fonction exécuteur, entrez la commande **cbrcontrol executor start**. Notez que vous pouvez également modifier divers paramètres de l'exécuteur à cette occasion. Voir «dscontrol executor — Contrôle de l'exécuteur», à la page 357.

Etape 4. Définition et configuration des options du cluster

CBR équilibrera les requêtes envoyées au cluster entre les serveurs correspondants configurés sur les ports de ce cluster.

Le cluster est le nom symbolique situé sur la portion hôte de l'URL et qui doit correspondre au nom utilisé dans l'instruction Proxy du fichier ibmproxy.conf.

Les clusters définis dans CBR doivent correspondre à la demande entrante. Un cluster doit être défini à l'aide du nom d'hôte ou de l'adresse IP contenue dans la demande entrante. Par exemple, si la demande est entrée sous forme d'adresse IP, le cluster doit être défini en tant que cette adresse IP. Si plusieurs noms d'hôte se résolvent en une seule adresse IP (et que les demandes peuvent arriver avec l'un de ces noms d'hôte), tous les noms d'hôte doivent être définis en tant que clusters.

Pour définir un cluster, tapez la commande suivante :

```
cbrcontrol cluster add cluster
```

Pour définir les options de cluster, tapez la commande suivante :

```
cbrcontrol cluster set valeur d'option de cluster
```

Pour plus d'informations, voir Chapitre 27, «Guide des commandes Dispatcher et CBR», à la page 345.

Etape 5. Affectation d'un alias à la carte d'interface réseau (facultatif)

Si vous exécutez Caching Proxy dans une configuration de proxy inverse, lors de l'équilibrage de charge pour plusieurs sites Web, vous devez ajouter l'adresse de cluster de chaque site Web au moins à l'une des cartes d'interface réseau du système Load Balancer. Sinon, vous pouvez ignorer cette étape.

Pour les systèmes **AIX, HP-UX, Linux ou Solaris** : Pour ajouter l'adresse de cluster à l'interface réseau, servez-vous de la commande `ifconfig`. Utilisez la commande adaptée au système d'exploitation (voir tableau 8).

Tableau 8. Commandes pour l'affectation d'un alias à la carte d'interface réseau

AIX	<code>ifconfig nom_interface alias adresse_cluster netmask masque_réseau</code>
HP-UX	<code>ifconfig nom_interface adresse_cluster netmask masque_réseau up</code>
Linux	<code>ifconfig nom_interface adresse_cluster netmask masque_réseau up</code>
Solaris 8, Solaris 9 et Solaris 10	<code>ifconfig nom_interface addif adresse_cluster netmask masque_réseau up</code>

Remarque : Pour les systèmes Linux et HP-UX, *nom_interface* doit comporter un numéro unique à chaque adresse de cluster ajoutée. Par exemple : `eth0:1`, `eth0:2`, etc.

Sous **Windows 2000** : Pour ajouter l'adresse de cluster à l'interface réseau, procédez comme suit :

1. Cliquez sur **Démarrer > Paramètres > Panneau de configuration**.
2. Cliquez deux fois sur **Connexions réseau et accès à distance**.
3. Cliquez à l'aide du bouton droit de la souris sur **Connexion au réseau local**.
4. Sélectionnez **Propriétés**.
5. Sélectionnez **Protocole Internet (TCP/IP)**, puis cliquez sur **Propriétés**.
6. Sélectionnez **Utiliser l'adresse IP suivante**, puis cliquez sur **Avancé**.
7. Cliquez sur **Ajouter**, puis entrez l'adresse IP et le masque de sous-réseau du cluster.

Sous **Windows 2003** : Pour ajouter l'adresse de cluster à l'interface réseau, procédez comme suit :

1. Cliquez sur **Démarrer > Panneau de configuration > Connexions réseau > Connexion au réseau local**.
2. Cliquez sur **Propriétés**.
3. Sélectionnez **Protocole Internet (TCP/IP)**, puis cliquez sur **Propriétés**.
4. Sélectionnez **Utiliser l'adresse IP suivante**, puis cliquez sur **Avancé**.
5. Cliquez sur **Ajouter**, puis entrez l'adresse IP et le masque de sous-réseau du cluster.

Etape 6. Définition des ports et de leurs options

Le numéro de port est le port à partir duquel les applications serveur sont à l'écoute. Pour le composant CBR avec Caching Proxy exécutant le trafic HTTP, il s'agit en général du port 80.

Pour définir le port du cluster défini à l'étape précédente, entrez la commande suivante :

```
cbrcontrol port add cluster:port
```

Pour définir les options de port, entrez la commande suivante :

```
cbrcontrol port set cluster:port option value
```

Pour plus d'informations, voir Chapitre 27, «Guide des commandes Dispatcher et CBR», à la page 345.

Etape 7. Définition des serveurs avec équilibrage de charge

Les serveurs sont les postes qui exécutent les applications dont vous souhaitez équilibrer la charge. Le *serveur* est l'adresse à nom symbolique ou notation décimale de la machine serveur. Pour définir un serveur dans le cluster et le port, tapez la commande suivante :

```
cbrcontrol server add cluster:port:serveur
```

Vous devez définir un ou plusieurs serveurs par port sur un cluster pour pouvoir procéder à l'équilibrage des charges.

Etape 8. Ajout de règles à la configuration

Il s'agit de l'étape clé de la configuration CBR avec Caching Proxy. Une règle définit la manière dont une requête URL sera reconnue et envoyée à l'un des ensembles de serveurs appropriés. Le type de règle spéciale utilisé par CBR est appelé règle de contenu. Pour définir une règle de contenu, tapez la commande suivante :

```
cbrcontrol rule add cluster:port:règle type content pattern motif
```

La valeur *pattern* est l'expression régulière qui est comparée à l'URL de chaque requête client. Pour plus d'informations sur la configuration de la structure, voir Annexe B, «Syntaxe des règles de contenu (modèle)», à la page 477.

Certains autres types de règles définis dans Dispatcher peuvent également être utilisés dans CBR. Pour plus d'informations, voir «Configuration de l'équilibrage de charge basé sur des règles», à la page 212.

Etape 9. Ajout de serveurs à vos règles

Lorsqu'une règle correspond à une requête client, l'ensemble de serveurs de la règle est interrogé pour déterminer le meilleur serveur. L'ensemble de serveurs de la règle est un sous-ensemble de serveurs définis dans le port. Pour ajouter des serveurs à un ensemble de serveurs de la règle, émettez la commande suivante :

```
cbrcontrol rule useserver cluster:port:rule server
```

Etape 10. Démarrage de la fonction gestionnaire (facultatif)

La fonction gestionnaire permet d'améliorer l'équilibrage de charge. Pour démarrer le gestionnaire, tapez la commande suivante :

```
cbrcontrol manager start
```

Etape 11. Démarrage de la fonction conseiller (facultatif)

Les conseillers transmettent au gestionnaire des informations complémentaires sur la capacité à répondre aux demandes des serveurs ayant fait l'objet d'un équilibrage de charge. Chaque conseiller est spécifique à un protocole. Par exemple, tapez la commande suivante pour lancer le conseiller HTTP :

```
cbrcontrol advisor start http port
```

Etape 12. Définition du niveau d'importance des informations requis pour le cluster

Si vous démarrez des conseillers, vous pouvez modifier le niveau d'importance donné aux informations des conseillers entrant dans les décisions d'équilibrage de la charge. Pour définir le niveau d'importance des informations pour le cluster, entrez la commande **cbrcontrol cluster set cluster NiveauImportance**. Pour plus d'informations, voir «Proportion de l'importance accordée aux données d'état», à la page 180.

Etape 13. Démarrage de Caching Proxy

- Systèmes AIX : Ajoutez la ligne suivante à votre variable d'environnement LIBPATH :
`/opt/ibm/edge/lb/servers/lib`
- Systèmes Linux, HP-UX ou Solaris : Ajoutez la ligne suivante à la variable d'environnement LD_LIBRARY_PATH :
`/opt/ibm/edge/lb/servers/lib`
- Systèmes Windows : Ajoutez la ligne suivante à votre variable d'environnement :
`C:\Program Files\IBM\edge\lb\servers\lib`

Dans le nouvel environnement, démarrez Caching Proxy, en entrant **ibmproxy** à partir de l'invite.

Remarque : Pour les systèmes Windows : Démarrez Caching Proxy à partir du panneau Services : **Démarrer-> Paramètres**-(pour Windows 2000) > **Panneau de configuration -> Outils d'administration -> Services**.

Exemple de configuration CBR

Pour configurer CBR, procédez aux opérations ci-dessous.

1. Démarrez CBR : entrez la commande **cbrserver**.
2. Démarrez l'interface de ligne de commande : émettez la commande **cbrcontrol**.
3. L'invite **cbrcontrol** s'affiche. Emettez les commandes suivantes :
`(cluster(c),port(p),rule(r),server(s))`
 - `executor start`
 - `cluster add c`
 - `port add c:p`
 - `server add c:p:s`
 - `rule add c:p:r type content pattern uri=*`
 - `rule use server c:p:r s`
4. Démarrez Caching Proxy : entrez la commande **ibmproxy**. Sous Windows, démarrez Caching Proxy à partir du panneau Services.
5. Supprimez toutes les configurations de proxy à partir du navigateur.

6. Chargez `http://c/` dans votre navigateur, où `"c"` est le cluster que vous avez précédemment configuré.
 - Le serveur `"s"` est appelé.
 - La page Web suivante s'affiche : `http://s/`

Partie 4. Composant Site Selector

Cette section contient des informations pour la configuration d'un démarrage rapide ainsi que des remarques relatives à la planification, et présente les diverses méthodes de configuration du composant Site Selector de Load Balancer. Elle se compose des chapitres suivants :

- Chapitre 12, «Configuration de démarrage rapide», à la page 123
- Chapitre 13, «Planification de Site Selector», à la page 127
- Chapitre 14, «Configuration de Site Selector», à la page 131

Chapitre 12. Configuration de démarrage rapide

Cet exemple de démarrage rapide montre comment créer une configuration de nom de site à l'aide de Site Selector pour équilibrer la charge sur un ensemble de serveurs sur la base du nom de domaine utilisé dans la demande d'un client.

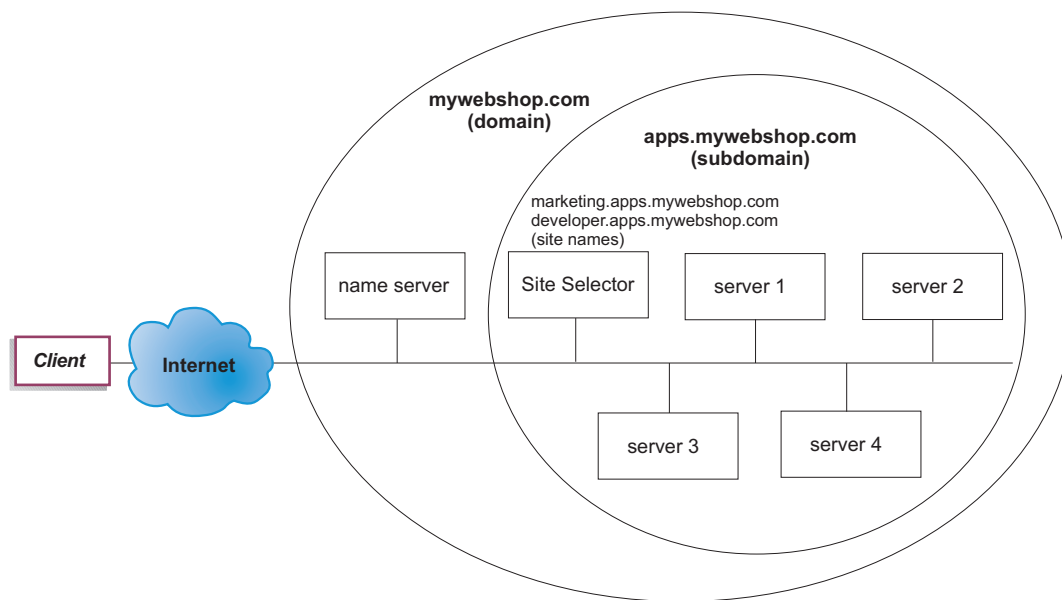


Figure 23. Configuration Site Selector simple

Matériel requis

Cet exemple de configuration de démarrage rapide nécessite :

- un accès administrateur au serveur de noms du site,
- quatre serveurs (serveur1, serveur2, serveur3, serveur4) configurés sur le réseau et un serveur supplémentaire sur lequel le composant Site Selector est installé,

Remarque : si Site Selector est co-implanté sur l'un des serveurs dont la charge est équilibrée, vous n'aurez besoin que de quatre serveurs au lieu de cinq. Toutefois, la co-implantation nuira aux performances des serveurs avec équilibrage de charge.

Préparation

Pour cet exemple de démarrage rapide, le domaine du site de la compagnie est `ma_boutique_web.com`. Site Selector est responsable du sous-domaine `ma_boutique_web.com`. Vous devez donc définir un sous-domaine dans `ma_boutique_web.com`. Par exemple : `apps.ma_boutique_web.com`. Site Selector n'est pas un système DNS totalement implémenté, tel que BIND, et agit en tant que noeud terminal dans une hiérarchie DNS. Site Selector a autorité sur le sous-domaine `apps.ma_boutique_web.com`. Le sous-domaine `apps.ma_boutique_web.com` contiendra les noms de site suivants : `marketing.apps.ma_boutique_web.com` et `développeur.apps.ma_boutique_web.com`.

1. Mettez à jour le serveur de noms de domaine du site de la compagnie (voir la figure 23, à la page 123). Créez comme un enregistrement de serveur de noms dans le fichier named.data pour le sous-domaine (apps.ma_boutique_web.com) dans lequel Site Selector est le serveur de noms faisant autorité :
apps.ma_boutique_web.com. IN NS siteselector.ma_boutique_web.com
2. Vérifiez que le nom d'hôte complet ou le site n'est pas résolu dans le système de nom de domaine en cours.
3. Installez Metric Server sur les serveurs (serveur1, serveur2, serveur3, serveur4) dont vous souhaitez que Site Selector équilibre la charge. Pour plus d'informations, voir «Metric Server», à la page 196.

Configuration du composant Site Selector

A l'aide de Site Selector, vous pouvez créer une configuration à l'aide de la ligne de commande, de l'assistant de configuration ou de l'interface graphique. Pour cet exemple de démarrage rapide, les étapes de configuration s'effectuent via la ligne de commande.

Remarque : Les valeurs des paramètres doivent être saisies à l'aide de caractères anglais. Les seules exceptions sont les valeurs des paramètres des noms d'hôte et des noms de fichiers.

Configuration à partir de la ligne de commande

A partir d'une invite, effectuez les opérations ci-dessous.

1. Démarrez le serveur ssserver sur la machine hébergeant Site Selector. En tant que superutilisateur ou administrateur, entrez la commande **ssserver** à partir d'une invite.

Remarque : Pour la plateforme Windows : Démarrez le serveur ssserver (IBM Site Selector) à partir du panneau Services : **Démarrer** > **Paramètres** (pour Windows 2000) > **Panneau de configuration** > **Outils d'administration** > **Services**.

2. Démarrez le serveur de noms sur la configuration Site Selector :
sscontrol nameserver start
3. Configurez les noms de site (marketing.apps.ma_boutique_web.com et développeur.apps.ma_boutique_web.com) sur Site Selector comme suit :
sscontrol sitename add marketing.apps.ma_boutique_web.com
sscontrol sitename add développeur.apps.ma_boutique_web.com
4. Ajoutez les serveurs à la configuration Site Selector. Pour ce faire, configurez serveur1 et serveur2 avec le nom de site marketing.apps.ma_boutique_web.com, puis configurez serveur3 et serveur4 avec le nom de site développeur.apps.ma_boutique_web.com comme suit :
sscontrol server add
marketing.apps.ma_boutique_web.com:serveur1+serveur2
sscontrol server add
développeur.apps.ma_boutique_web.com:serveur3+serveur4
5. Démarrez la fonction gestionnaire (manager) de Site Selector :
sscontrol manager start
6. Démarrez la fonction conseiller (advisor) de Site Selector (conseiller HTTP pour marketing.apps.ma_boutique_web.com et conseiller FTP pour développeur.apps.ma_boutique_web) :

```
sscontrol advisor start http marketing.apps.ma_boutique_web.com:80
sscontrol advisor start ftp developpeur.apps.ma_boutique_web.com:21
```

Site Selector vérifie désormais que les demandes des clients ne sont pas envoyées vers un serveur arrêté.

7. Vérifiez que Metric Server a été démarré sur chacun des serveurs avec équilibrage de charge.

La configuration de base de Site Selector est maintenant terminée.

Test de vérification de la configuration

Vérifiez que la configuration fonctionne :

1. A partir d'un client, doté d'un système DNS principal configuré comme étant le serveur de noms responsable de ma_boutique_web.com, lancez une commande ping sur l'un des noms de site configurés.
2. Connectez-vous à l'application. Par exemple :
 - Ouvrez un navigateur, demandez marketing.apps.ma_boutique_web.com ; une page correcte doit s'afficher.
 - Ouvrez un client FTP sur developpeur.apps.ma_boutique_web.com, puis entrez un nom d'utilisateur et un mot de passe correct.
3. Observez les résultats de la commande suivante :

```
sscontrol server status marketing.apps.ma_boutique_web.com:
```

```
sscontrol server status developpeur.apps.ma_boutique_web.com:
```

Le nombre total de réussites d'accès de chaque serveur doit s'ajouter aux demandes de commande ping et d'application.

Configuration à l'aide de l'interface graphique

Pour plus d'informations sur l'utilisation de l'interface graphique de Site Selector, voir «Interface graphique», à la page 132 et Annexe A, «Interface graphique utilisateur : Instructions générales», à la page 469.

Configuration à l'aide de l'assistant de configuration

Pour plus d'informations sur l'utilisation de l'assistant de Site Selector, voir «Assistant de configuration», à la page 133.

Chapitre 13. Planification de Site Selector

Le présent chapitre décrit les aspects que l'administrateur de réseau doit prendre en compte avant d'installer et de configurer le composant Site Selector.

- Voir Chapitre 3, «Gestion du réseau : Fonctions Load Balancer requises», à la page 19 pour une présentation des fonctions permettant de gérer votre réseau.
- Voir Chapitre 14, «Configuration de Site Selector», à la page 131 pour obtenir des informations sur la configuration des paramètres d'équilibrage de charge de Site Selector.
- Voir Chapitre 22, «Fonctions avancées de Dispatcher, CBR et Site Selector», à la page 201 pour obtenir des informations sur la configuration de Load Balancer pour les fonctions avancées.
- Voir au Chapitre 24, «Exploitation et gestion de Load Balancer», à la page 261, pour obtenir des informations sur l'administration authentifiée à distance, les fichiers journaux Load Balancer et sur l'utilisation des composants Load Balancer.

Le présent chapitre se compose des sections suivantes :

- «Remarques relatives à la planification»
- «Considérations relatives à la durée de vie (TTL)», à la page 129
- «Utilisation de la fonction de proximité réseau (Network Proximity)», à la page 129

Remarques relatives à la planification

Site Selector fonctionne avec un serveur de noms de domaine pour équilibrer la charge sur un groupe de serveurs à l'aide des mesures et des pondérations recueillies. Vous pouvez créer une configuration de site pour assurer l'équilibrage de charge sur un groupe de serveurs sur la base du nom de domaine utilisé pour la demande d'un client.

Limitations : Les seules requêtes DNS prises en charge par Site Selector sont celles de type A. Tous les autres types de requête génèrent le code retour NOTIMPL (Not Implemented - non implémenté). Si tout un domaine est attribué à Site Selector, assurez-vous qu'il ne reçoive que des requêtes de type A.

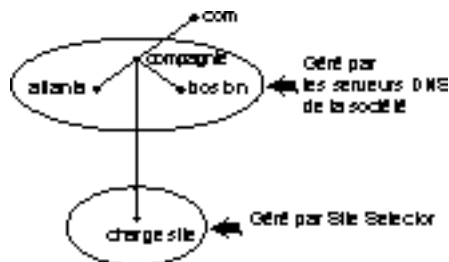


Figure 24. Exemple d'environnement DNS

Lors de la définition d'un sous-domaine de Site Selector dans l'environnement DNS, Site Selector doit disposer de droits d'accès à ce sous-domaine. Par exemple (voir la figure 24), votre entreprise dispose de droits d'accès au domaine **entreprise.com**. Elle dispose de plusieurs sous-domaines. Site Selector doit disposer

de droits d'accès à **siteload.entreprise.com** et les serveurs DNS gardent leurs droits d'accès à **atlanta.entreprise.com** et à **boston.entreprise.com**.

Pour permettre au serveur de noms de l'entreprise de reconnaître les droits d'accès de Site Selector au sous-domaine siteload, il est nécessaire d'ajouter une entrée dans le fichier de données du serveur de noms. Par exemple, sur les systèmes AIX, une entrée de serveur de noms a l'apparence suivante :

```
siteload.entreprise.com. IN NS siteselector.entreprise.com.
```

Où **siteselector.entreprise.com** correspond au nom d'hôte de la machine Site Selector. Des entrées équivalentes doivent être insérées dans les autres fichiers de base de données utilisés par les serveurs DNS.

Un client envoie une demande de résolution de nom de domaine à un serveur de noms du réseau. Le serveur de noms achemine la demande au poste Site Selector. Ce dernier résout le nom de domaine en adresse IP de l'un des serveurs qui a été configuré sous le nom du site. Site Selector renvoie l'adresse IP du serveur sélectionné au serveur de noms. Le serveur de noms renvoie l'adresse IP au client. (Site Selector joue le rôle de serveur de noms non récurrents (noeud feuille) et renvoie une erreur s'il ne résout pas la demande de nom de domaine.

Voir figure 5, à la page 15 qui montre un site dans lequel Site Selector est utilisé avec un système DNS pour équilibrer la charge entre des serveurs locaux et éloignés.

Site Selector se compose des fonctions suivantes :

- **ssserver** traite les demandes à partir de la ligne de commande adressées au serveur de noms, au gestionnaire et aux conseillers.
- La fonction de **serveur de noms** prend en charge l'équilibrage de charge des demandes de serveur de noms entrantes. Vous devez démarrer la fonction de serveur de noms pour que Site Selector puisse commencer la résolution DNS. Site Selector écoute sur le port 53 les requêtes DNS entrantes. Si le nom du site qui émet la demande est configuré, Site Selector renvoie une adresse de serveur unique (à partir d'un ensemble d'adresses de serveurs) associée au nom de site.
- Le **gestionnaire** définit les pondérations utilisées par le serveur de noms en fonction de plusieurs facteurs :
 - le retour d'informations sur les serveurs fourni par les conseillers,
 - le retour d'informations émanant d'un programme de contrôle système, tel que Metric Server.

L'utilisation du gestionnaire n'est que facultative. Toutefois, s'il n'est pas utilisé, l'équilibrage de charge se fait sur la base d'une planification circulaire pondérée, elle-même basée sur les mesures de charge des serveurs et les conseillers ne seront pas disponibles.

- **Metric Server** est un composant de surveillance système de Load Balancer installé sur la machine serveur dorsal. (Si vous co-implantez Load Balancer sur une machine serveur dont la charge est en cours d'équilibrage, installez Metric Server sur la machine Load Balancer.)

Metric Server permet à Site Selector de surveiller le niveau d'activité d'un serveur, de détecter le moment où un serveur est le moins chargé et de détecter un serveur défaillant. Par charge, on entend le travail effectivement fourni par le serveur. L'administrateur système Site Selector contrôle le type de mesure employé pour évaluer la charge. Site Selector peut être configuré en fonction de chaque environnement, en tenant compte de facteurs tels que la fréquence des

accès, le nombre total d'utilisateurs et les différents types d'accès (requêtes courtes, longues, à forte ou faible consommation de ressources CPU).

L'équilibrage de charge est basée sur les pondérations de serveur. Pour Site Selector, il existe quatre niveaux d'importance des informations que le gestionnaire utilise pour déterminer les pondérations :

- CPU
- mémoire
- port
- système

Les valeurs CPU et mémoire sont fournies par Metric Server. Par conséquent, l'utilisation de Metric Server est *recommandée* avec le composant Site Selector.

Pour plus d'informations, voir «Metric Server», à la page 196.

- Les **conseillers** interrogent les serveurs et analysent les résultats par protocole avant d'appeler le gestionnaire pour définir les pondérations comme il convient. L'utilisation de certains de ces conseillers n'est peut-être pas utile dans une configuration typique. Vous avez également la possibilité de développer vos propres conseillers. L'utilisation des conseillers est facultative mais recommandée. Pour plus d'informations, voir «Conseillers», à la page 185.
- Pour configurer et gérer le serveur de noms, les conseillers, Metric Server et le gestionnaire, utilisez la ligne de commande (**sscontrol**) ou l'interface utilisateur graphique (**lbadm**).

Les quatre fonctions clés de Site Selector (serveur de noms, gestionnaire, Metric Server et conseillers) interagissent afin d'équilibrer les demandes entrantes entre les serveurs et de les résoudre.

Considérations relatives à la durée de vie (TTL)

L'équilibrage de charge utilisant le système DNS nécessite la désactivation de l'enregistrement en mémoire cache de la résolution des noms. La valeur TTL (time to live) détermine l'efficacité de ce type d'équilibrage de charge. Elle détermine la période pendant laquelle la réponse résolue reste en mémoire cache sur un autre serveur de noms. Les valeurs TTL peu élevées permettent d'effectuer plus rapidement les modifications subtiles de la charge du serveur ou du réseau. La désactivation de l'enregistrement en mémoire cache oblige toutefois les clients à contacter le serveur de noms autorisé pour chaque demande de résolution de nom, augmentant potentiellement le temps d'attente des clients. Tenez compte de l'impact sur l'environnement de la désactivation de l'enregistrement en mémoire cache lorsque vous choisissez une valeur TTL. Vous devez en outre savoir que l'équilibrage de charge DNS peut être limité par l'enregistrement en mémoire cache côté client de la résolution des noms.

Vous pouvez configurer la durée de vie (TTL) à l'aide de la commande **sscontrol sitename [add | set]** . Pour plus d'informations, voir «sscontrol sitename — Configuration d'un nom de site», à la page 424.

Utilisation de la fonction de proximité réseau (Network Proximity)

Network proximity correspond au calcul de la position de chaque serveur par rapport au client émettant la demande. Pour déterminer la proximité réseau, l'agent Metric Server (qui doit se trouver sur chaque serveur dont la charge est équilibrée) envoie une commande ping à l'adresse IP client et renvoie le temps de réponse à Site Selector. Site Selector utilise la réponse de proximité dans la décision

relative à l'équilibrage de charge. Il combine la valeur de la réponse de proximité réseau avec la pondération provenant du gestionnaire pour créer une valeur de pondération finale pour le serveur.

L'utilisation de la fonction de proximité réseau (Network Proximity) avec Site Selector est facultative.

Site Selector fournit les options de proximité réseau suivantes pouvant être définies par nom de site :

- Durée de stockage en cache : Période pendant laquelle une réponse de proximité reste valide et sera enregistrée dans la mémoire cache.
- Pourcentage de proximité : Importance de la réponse de proximité par rapport à l'état du serveur (en tant qu'entrée de la pondération du gestionnaire).
- Attente de toutes les réponses : Détermine s'il est nécessaire d'attendre toutes les réponses de proximité (ping) des serveurs avant de répondre à la demande du client.

Si cette option est associée à la valeur **oui**, Metric Server envoie une commande ping au client pour obtenir le temps de réponse de proximité. Le serveur de noms attend que tous les serveurs Metric répondent ou que le délai d'expiration se termine. Ensuite, pour chaque serveur, le serveur de noms combine le temps de réponse de proximité avec la pondération que le gestionnaire a calculée pour créer une valeur de "pondération combinée". Site Selector fournira au client l'adresse IP du serveur associée à la meilleure pondération combinée.

(Normalement, la plupart des serveurs de noms client observent un dépassement de délai de 5 secondes. Site Selector tente de répondre avant la fin de ce délai.)

Si la valeur est **non**, une résolution de nom est fournie au client en fonction des pondérations de gestionnaire actuelles. Ensuite, Metric Server envoie une commande ping au client pour obtenir le temps de réponse de proximité. Le serveur de noms met en cache le temps de réponse qu'il reçoit de Metric Server. Lorsque le client renvoie une deuxième requête, le serveur de noms combine la pondération du gestionnaire actuelle avec la valeur de réponse ping mise en cache pour chaque serveur afin d'obtenir le serveur associé à la meilleure "pondération combinée". Site Selector renvoie l'adresse IP de ce serveur au client pour la deuxième requête.

Les options de proximité réseau peuvent être définies dans la commande **sscontrol sitename [add | set]** . Pour plus d'informations, voir Chapitre 28, «Guide des commandes Site Selector», à la page 401.

Chapitre 14. Configuration de Site Selector

Avant d'effectuer les opérations décrites dans le présent chapitre, voir Chapitre 13, «Planification de Site Selector», à la page 127. Ce chapitre décrit comment créer une configuration de base pour le composant Site Selector de Load Balancer.

- Voir Chapitre 21, «Gestionnaire, conseillers et système Metric Server (des composants Dispatcher, CBR et Site Selector)», à la page 179 et Chapitre 22, «Fonctions avancées de Dispatcher, CBR et Site Selector», à la page 201 pour plus d'informations sur des configurations plus complexes de Load Balancer.
- Voir Chapitre 24, «Exploitation et gestion de Load Balancer», à la page 261, pour obtenir des informations sur l'administration authentifiée à distance, les fichiers journaux Load Balancer et sur l'utilisation des composants Load Balancer.

Présentation générale des tâches de configuration

Remarque : Avant de suivre les étapes de configuration détaillées dans ce tableau, assurez-vous que la machine Site Selector et toutes les machines serveurs sont connectées au réseau, ont des adresses IP valides et peuvent communiquer entre elles par la triangulation ping.

Tableau 9. Tâches de configuration pour le composant Site Selector

Tâche	Description	Informations connexes
Configuration de la machine Site Selector.	Conditions requises	«Installation de la machine Site Selector», à la page 134
Configuration des machines en vue de l'équilibrage de charge.	Définition de la configuration de l'équilibrage de charge.	«Étape 4. Définition de serveurs avec équilibrage de charge», à la page 135

Méthodes de configuration

Quatre méthodes permettent de créer une configuration de base pour composant Site Selector de Load Balancer :

- Ligne de commande
- Scripts
- Interface graphique
- Assistant de configuration

Ligne de commande

C'est la méthode la plus directe pour la configuration de Site Selector. Les valeurs des paramètres de commandes doivent être saisies à l'aide de caractères anglais. Les seules exceptions s'appliquent aux noms d'hôte (utilisés dans les commandes site name et server) et aux noms de fichiers.

Pour démarrer Site Selector à partir de la ligne de commande :

1. Emettez la commande **sssserver** à partir de l'invite. Pour arrêter le service, tapez **sssserver stop**

Remarque : Pour les systèmes Windows, cliquez sur **Démarrer > Paramètres** (pour Windows 2000) > **Panneau de configuration > Outils**

d'administration > Services. Cliquez à l'aide du bouton droit de la souris sur **IBM Site Selector**, puis sélectionnez **Démarrer**. Pour arrêter le service, suivez la même procédure en sélectionnant **Arrêter**.

2. Ensuite, émettez les commandes de contrôle Site Selector voulues pour installer votre configuration. Les procédures décrites dans ce manuel reposent sur l'utilisation de la ligne de commande. La commande est **sscontrol**. Pour plus de détails sur les commandes, voir Chapitre 28, «Guide des commandes Site Selector», à la page 401.

Vous pouvez entrer une version abrégée des paramètres de commandes **sscontrol**. Il suffit d'entrer les lettres spécifiques des paramètres. Par exemple, pour obtenir l'aide correspondant à la commande **file save**, entrez **sscontrol he f** à la place de **sscontrol help file**.

Pour démarrer l'interface de ligne de commande, entrez **sscontrol** pour ouvrir une invite **sscontrol**.

Pour fermer l'interface de ligne de commande, entrez **exit** ou **quit**.

Remarque : Sous Windows, le service **dsserver** du composant Dispatcher démarre automatiquement. Si vous utilisez uniquement Site Selector et non le composant Dispatcher, vous pouvez empêcher **dsserver** de démarrer automatiquement de la manière suivante :

1. Dans le panneau Services de Windows, cliquez à l'aide du bouton droit de la souris sur **IBM Dispatcher**.
2. Sélectionnez Propriétés.
3. Dans la zone **Type de démarrage**, sélectionnez Manuel.
4. Cliquez sur OK et fermez la fenêtre Services.

Scripts

Les commandes permettant de configurer Site Selector peuvent être entrées dans un fichier script de configuration, puis exécutées ensemble.

Remarque : Pour exécuter rapidement le contenu d'un fichier script (par exemple, *mon_script*), utilisez l'une des commandes suivantes :

- Pour mettre à jour la configuration actuelle, soumettez les commandes exécutables à partir du fichier script, en entrant —
sscontrol file appendload mon_script
- Pour remplacer la configuration actuelle, soumettez les commandes exécutables à partir du fichier script en entrant —
sscontrol file newload mon_script

Pour sauvegarder la configuration en cours dans un fichier script (par exemple, *savescript*), exécutez la commande suivante :

sscontrol file save savescript

Cette commande enregistre le fichier script de configuration dans le répertoire **...ibm/edge/lb/servers/configurations/ss**.

Interface graphique

Pour des instructions générales et un exemple de l'interface graphique, voir la figure 41, à la page 470.

Pour démarrer l'interface graphique, procédez comme suit :

1. Vérifiez que ssserver fonctionne. En tant que superutilisateur ou administrateur, entrez la commande **ssserver** à partir d'une invite.
2. Ensuite, effectuez l'une des opérations suivantes :
 - Pour les systèmes AIX, HP-UX, Linux ou Solaris : entrez **lbadmin**
 - Pour les systèmes Windows : cliquez sur **Démarrer > Programmes IBM WebSphere > Edge Components > IBM Load Balancer > Load Balancer**

Pour pouvoir configurer le composant Site Selector à partir de l'interface graphique, vous devez d'abord sélectionner **Site Selector** dans l'arborescence. Une fois connecté à un hôte sur lequel le serveur ssserver est exécuté, vous pouvez créer des noms de site contenant des serveurs, démarrer le gestionnaire et lancer des conseillers.

Vous pouvez utiliser l'interface graphique pour toute opération normalement exécutée par la commande **sscontrol**. Par exemple, pour définir un nom de site à partir de la ligne de commande, vous devez entrer la commande **sscontrol sitename add nom_site**. Pour définir un nom de site à partir de l'interface graphique, cliquez à l'aide du bouton droit de la souris sur Serveur de noms, puis dans le menu en incrustation, sélectionnez **Ajouter un nom de site**. Entrez le nom du site dans le menu en incrustation, puis cliquez sur **OK**.

Les fichiers de configuration Site Selector existants peuvent être chargés à l'aide des options **Chargement de la nouvelle configuration** (pour remplacer intégralement la configuration en cours) et **Ajout à la configuration en cours** (pour mettre à jour la configuration en cours) du menu en incrustation **Hôte**. Vous devez sauvegarder votre configuration Site Selector dans un fichier en utilisant l'option **Sauvegarder le fichier de configuration en** du menu en incrustation **Hôte**. Le menu **Fichier** situé en haut de l'interface graphique permet de sauvegarder les connexions à l'hôte en cours dans un fichier ou de restaurer les connexions dans des fichiers existants sur tous les composants Load Balancer.

Pour exécuter une commande à partir de l'interface graphique : mettez le noeud Hôte en surbrillance dans l'arborescence de l'interface graphique, puis sélectionnez **Envoyer la commande...** dans le menu en incrustation Hôte. Dans la zone d'entrée de commande, entrez la commande à exécuter, par exemple **nameserver status**. Les résultats et l'historique des commandes exécutées lors de la session courante s'affichent dans la fenêtre ouverte.

Vous pouvez accéder à l'**Aide** en cliquant sur le point d'interrogation situé dans l'angle supérieur droit de la fenêtre Load Balancer.

- **Aide sur les zones** — décrit les valeurs par défaut de chaque zone.
- **Procédures** — affiche la liste des tâches pouvant être effectuées dans cet écran.
- **Centre de documentation** — fournit un accès centralisé aux informations relatives au produit

Pour plus de détails sur l'utilisation de l'interface graphique, voir Annexe A, «Interface graphique utilisateur : Instructions générales», à la page 469.

Assistant de configuration

Si vous utilisez l'assistant de configuration, suivez la procédure ci-dessous.

1. Démarrez ssserver sur Site Selector :

- Exécutez la commande suivante en tant que superutilisateur ou administrateur :

ssserver

2. Lancez la fonction d'assistant de Site Selector, **sswizard**.

Vous pouvez lancer cet assistant à partir de l'invite en entrant la commande **sswizard** ou sélectionner l'assistant de configuration à partir du menu du composant Site Selector présenté dans l'interface graphique.

L'assistant Site Selector vous guide pas à pas dans le processus de création d'une configuration de base pour le composant Site Selector. Il vous demande des renseignements sur votre réseau et vous guide pour la configuration d'un nom de site permettant à Site Selector d'équilibrer le trafic entre un groupe de serveurs.

Installation de la machine Site Selector

La configuration de la machine Site Selector ne peut être effectuée que par le superutilisateur (pour les systèmes AIX, HP-UX, Linux ou Solaris) ou l'administrateur (pour les systèmes Windows).

Vous aurez besoin d'un nom d'hôte complet ne pouvant être résolu comme nom de site pour le groupe de serveurs que vous configurez. Le nom de site est celui utilisé pour les clients pour accéder à votre site (par exemple, www.yourcompany.com). Site Selector équilibrera la charge du trafic pour ce nom de site entre les serveurs du groupe auquel le nom DNS a été attribué.

Etape 1. Démarrage de la fonction serveur

Pour démarrer la fonction serveur Site Selector, entrez **ssserver** sur la ligne de commande.

Remarque : Un fichier de configuration par défaut (default.cfg) est chargé automatiquement pendant le démarrage de **ssserver**. Si vous décidez de sauvegarder la configuration dans default.cfg, toutes les données sauvegardées dans ce fichier sont chargées automatiquement au prochain démarrage de **ssserver**.

Etape 2. Démarrage du serveur de noms

Pour démarrer le serveur de noms, entrez la commande **sscontrol nameserver start**.

Vous pouvez également lancer le serveur de noms à l'aide du mot clé bindaddress pour établir un lien uniquement avec l'adresse indiquée.

Etape 3. Définition d'un nom de site et définition des options du nom de site

Site Selector équilibrera les demandes envoyées pour le nom de site aux serveurs correspondants configurés pour cela.

Le nom de site est un nom d'hôte ne pouvant être résolu qui sera demandé par le client. Le nom de site doit être un nom de domaine complet (par exemple, www.dnsdownload.com). Lorsqu'un client demande ce nom de site, l'une des adresses IP de serveur associées au nom de site est renvoyée.

Pour définir un nom de site, émettez la commande suivante :

```
sscontrol sitename add nom_site
```

Pour définir les options du nom de site, émettez la commande suivante :

```
sscontrol sitename set valeur_option_nom_site
```

Pour plus d'informations, voir Chapitre 28, «Guide des commandes Site Selector», à la page 401.

Etape 4. Définition de serveurs avec équilibrage de charge

Les serveurs sont les postes qui exécutent les applications dont vous souhaitez équilibrer la charge. Le *serveur* est l'adresse à nom symbolique ou notation décimale de la machine serveur. Pour définir un serveur sur le nom de site défini à l'étape 3, émettez la commande suivante :

```
sscontrol server add  
nom_site:serveur
```

Vous devez définir plusieurs serveurs sous un nom de site afin de permettre l'équilibrage de charge.

Etape 5. Démarrage de la fonction gestionnaire (facultatif)

La fonction gestionnaire permet d'étendre la fonction d'équilibrage de charge. Avant de lancer la fonction gestionnaire, vérifiez que le système Metric Server est installé sur toutes les machines dont la charge est équilibrée.

Pour démarrer le gestionnaire, tapez la commande suivante :

```
sscontrol manager start
```

Etape 6. Démarrage de la fonction conseiller (facultatif)

Les conseillers transmettent au gestionnaire des informations complémentaires sur la capacité à répondre aux demandes des serveurs ayant fait l'objet d'un équilibrage de charge. Chaque conseiller est spécifique à un protocole. Load Balancer fournit de nombreux conseillers. Par exemple, pour lancer le conseiller HTTP pour un nom de site particulier, entrez la commande suivante :

```
sscontrol advisor start http nom_site:port
```

Etape 7. Définition des mesures du système (facultatif)

Pour plus d'informations sur l'utilisation des mesures du système et de Metric Server, voir «Metric Server», à la page 196.

Etape 8. Définition du niveau d'importance des informations pour le nom de site

Si vous démarrez des conseillers, vous pouvez modifier le niveau d'importance donné aux informations fournies par ces derniers (port) et entrant dans les décisions d'équilibrage de la charge. Pour définir le niveau d'importance pour le nom de site, émettez la commande **sscontrol sitename set *nom_site* proportions**. Pour plus d'informations, voir «Proportion de l'importance accordée aux données d'état», à la page 180.

Configuration des serveurs pour l'équilibrage de la charge

Il est recommandé d'utiliser Metric Server avec le composant Site Selector. Pour plus d'informations sur la configuration de Metric Server sur tous les serveurs dont Site Selector assure l'équilibrage de charge, voir «Metric Server», à la page 196.

Partie 5. Composant Cisco CSS Controller

Cette contient des informations pour la configuration d'un démarrage rapide ainsi que des remarques relatives à la planification, et présente les diverses méthodes de configuration du composant Cisco CSS Controller de Load Balancer. Elle se compose des chapitres suivants :

- Chapitre 15, «Configuration de démarrage rapide», à la page 139
- Chapitre 16, «Planification de Cisco CSS Controller», à la page 143
- Chapitre 17, «Configuration de Cisco CSS Controller», à la page 149

Chapitre 15. Configuration de démarrage rapide

Cet exemple de démarrage rapide montre comment créer une configuration à l'aide du composant Cisco CSS Controller. Cisco CSS Controller fournit des informations de pondération de serveur qui permettent à Cisco CSS Switch d'optimiser la sélection des serveurs lors des décisions d'équilibrage de la charge.

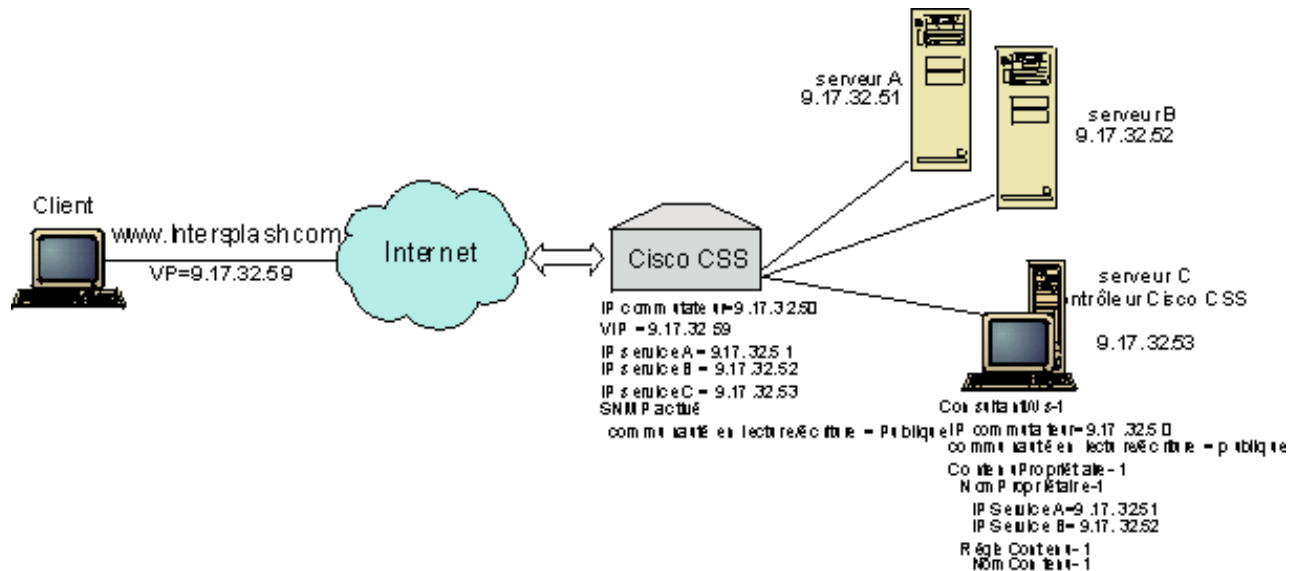


Figure 25. Configuration Cisco CSS Controller simple

Matériel requis

Cet exemple de configuration de démarrage rapide nécessite :

- un commutateur Cisco CSS ;
- un serveur doté d'un contrôleur Cisco CSS ;
- deux serveurs Web ;
- cinq adresses IP :
 - l'adresse IP qui permet aux clients d'accéder à votre site Web site, www.Intersplashx.com (9.17.32.59),
 - l'adresse IP d'une interface (passerelle) d'accès à Cisco CSS Switch (9.17.32.50),
 - l'adresse IP d'un serveur A (9.17.32.51),
 - l'adresse IP d'un serveur B (9.17.32.52),
 - l'adresse IP du serveur Cisco CSS Controller C (9.17.32.53).

Préparation

Avant de d'effectuer la configuration de cet exemple, procédez aux vérifications suivantes :

- Assurez-vous que Cisco CSS est correctement configuré. Pour plus d'informations sur la configuration, reportez-vous au manuel *Cisco Content Services Switch Getting Started Guide*.
- Assurez-vous que la machine Cisco CSS Controller peut contacter le Cisco CSS Switch (9.17.32.50), le serveur A (9.17.32.51) et le serveur B (9.17.32.52).
- Assurez-vous que la machine client peut contacter le VIP (9.17.32.59).

Configuration du composant Cisco CSS Controller

Avec Cisco CSS Controller, vous pouvez créer une configuration à l'aide de la ligne de commande ou de l'interface graphique. Pour cet exemple de démarrage rapide, les étapes de configuration s'effectuent via la ligne de commande.

Remarque : Les valeurs des paramètres doivent être saisies à l'aide de caractères anglais. Les seules exceptions sont les valeurs des paramètres des noms d'hôte et des noms de fichiers.

Configuration à partir de la ligne de commande

A partir d'une invite, effectuez les opérations ci-dessous.

1. Démarrez ccoserver sur Load Balancer. En tant que superutilisateur ou administrateur, entrez la commande **ccoserver** à partir d'une invite.
2. Ajoutez un consultant de commutateur à la configuration Cisco CSS Controller, en précisant l'adresse IP de l'interface Cisco CSS Switch et le nom de communauté en lecture/écriture. Ces valeurs doivent correspondre aux attributs équivalents sur Cisco CSS Switch :

```
cococontrol consultant add SwConsultant-1 address 9.17.32.50 community public
```

Cette opération vérifie la connectivité au Cisco CSS Switch vérifie que le nom de communauté en lecture/écriture SNMP fonctionne correctement.

3. Ajoutez le contenu de propriétaire (OwnerContent-1) au consultant du commutateur, en précisant le nom du propriétaire (OwnerName-1) et la règle de contenu (ContentRule-1) :

```
cococontrol ownercontent add SwConsultant-1:OwnerContent-1 ownername OwnerName-1 contentrule ContentRule-1
```

Ces valeurs doivent correspondre aux attributs équivalents sur Cisco CSS Switch.

Cisco CSS Controller peut maintenant communiquer avec le commutateur via SNMP et obtenir les informations de configuration nécessaires. Une fois cette procédure effectuée, Cisco CSS Controller contiendra les informations relatives aux services configurés sur Cisco CSS Switch pour le contenu de propriétaire indiqué.

4. Configurez le type de mesures à collecter (connexion active, débit de la connexion, HTTP) et les proportions de chaque mesure du contenu de propriétaire :

```
cococontrol ownercontent metrics SwConsultant-1:OwnerContent-1 activeconn 45 connrate 45 http 10
```

Cette commande configure les informations de mesure et les proportions à collecter auprès des services pour le calcul de pondération. La somme des proportions de toutes les mesures doit être égale à 100.

5. Démarrez la fonction consultant de commutateur de Cisco CSS Controller :
cococontrol consultant start SwConsultant-1

Cette commande lance tous les collecteurs de mesures et démarre les calculs de pondération de service. Cisco CSS Controller communique les résultats de ses calculs de pondération de service à Cisco CSS Switch à l'aide de SNMP.

La configuration de base de Cisco CSS Controller est maintenant terminée.

Test de vérification de la configuration

Vérifiez que la configuration fonctionne :

1. A l'aide du navigateur Web du client, accédez à **<http://www.Intersplashx.com>** . Si une page s'affiche, la configuration fonctionne.
2. Rechargez la page dans le navigateur Web.
3. Vérifiez les résultats de la commande suivante : **cococontrol service report SwConsultant-1:OwnerContent-1:Service-1**. La colonne du nombre total de connexions des deux serveurs Web doit contenir la valeur "2".

Configuration à l'aide de l'interface graphique

Pour plus d'informations sur l'utilisation de l'interface graphique du contrôleur Cisco CSS, voir «Interface graphique», à la page 151 et Annexe A, «Interface graphique utilisateur : Instructions générales», à la page 469.

Chapitre 16. Planification de Cisco CSS Controller

Le présent chapitre décrit les aspects qu'un administrateur de réseau doit prendre en compte avant d'installer et de configurer le composant Cisco CSS Controller.

- Pour des informations sur la configuration des paramètres d'équilibrage de charge du composant Cisco CSS Controller, voir Chapitre 17, «Configuration de Cisco CSS Controller», à la page 149.
- Voir Chapitre 23, «Fonctions avancées de Cisco CSS Controller et Nortel Alteon Controller», à la page 243 pour obtenir des informations sur la configuration de Load Balancer pour les fonctions avancées.
- Pour des informations sur l'administration authentifiée à distance, les fichiers journaux Load Balancer et l'utilisation des composants Load Balancer, voir Chapitre 24, «Exploitation et gestion de Load Balancer», à la page 261.

Le présent chapitre se compose des sections suivantes :

- «Configuration requise»
- «Remarques relatives à la planification»
 - «Positionnement du consultant dans le réseau», à la page 144
 - «Haute disponibilité», à la page 146
 - «Calcul des pondérations», à la page 146
 - «Identification des incidents», à la page 147

Configuration requise

Pour plus d'informations sur les conditions matérielles et logicielles requises, accédez à la page Web suivante : <http://www.ibm.com/support/docview.wss?rs=180&uid=swg27006921>

Vous avez également besoin des éléments suivants :

- un système sur lequel exécuter le Cisco CSS Controller,
- un commutateur Cisco CSS 11000 de services de contenu installé et configuré.

Remarques relatives à la planification

Cisco CSS Controller gère une série de consultants de commutateur. Chaque consultant détermine les pondérations des services dont la charge est équilibrée par un seul commutateur. Le commutateur auquel le consultant fournit les pondérations est configuré pour l'équilibrage de charge de contenu. Le consultant utilise le protocole SNMP pour transmettre au commutateur les pondérations calculées. Le commutateur utilise les pondérations reçues pour sélectionner un service pour la règle de contenu dont il équilibre la charge lorsque l'algorithme d'équilibrage de charge est pondéré par permutation circulaire (round-robin). Pour déterminer les pondérations, le consultant utilise un ou plusieurs des éléments d'information suivants :

- Disponibilité et temps de réponse, déterminés à l'aide des **conseillers** de l'application qui communiquent avec les applications qui s'exécutent sur le service.
- Informations relatives à la charge système, déterminée par extraction d'une valeur de mesure des **agents de mesure** qui s'exécutent sur le service.
- Informations de connexion relatives au service, provenant du commutateur.

- Informations d'accessibilité, extraites en contactant le service.

Pour une description de l'équilibrage de charge par contenu et des informations détaillées sur la configuration du commutateur, reportez-vous au manuel *Cisco Content Services Switch Getting Started Guide*.

Pour qu'un consultant obtienne les informations dont il a besoin afin de déterminer les pondérations d'un service, les éléments suivants sont nécessaires :

- Connectivité IP entre le consultant et les services pour lesquels les pondérations sont calculées.
- Connectivité IP entre le consultant et le commutateur qui équilibre la charge des serveurs pour lesquels les pondérations sont calculées.
- SNMP activé sur le commutateur. Les fonctions de lecture et d'écriture doivent être activées.

Positionnement du consultant dans le réseau

Comme indiqué à la figure 26, à la page 145, il est possible de connecter le consultant au réseau derrière le ou les commutateurs pour lesquels il fournit des pondérations. Certains paramètres doivent être configurés sur le commutateur, d'autres sur le contrôleur, pour activer la connectivité entre le contrôleur, le commutateur et les services.

Dans la figure 26, à la page 145 :

- Un consultant est connecté au réseau derrière les commutateurs pour lesquels il fournit des pondérations.
- Le réseau est constitué de deux réseaux VLAN.
- Pour que le consultant communique avec les services des deux réseaux VLAN, la transmission IP doit être activée sur les interfaces par lesquelles les services sont connectés et sur celle par laquelle le consultant est connecté.
- L'adresse IP du commutateur doit être configurée comme passerelle par défaut sur le consultant et sur les systèmes des services.

Pour des informations détaillées sur la configuration des réseaux VLAN et du routage IP sur le commutateur, reportez-vous au manuel *Cisco Content Services Switch Getting Started Guide*.

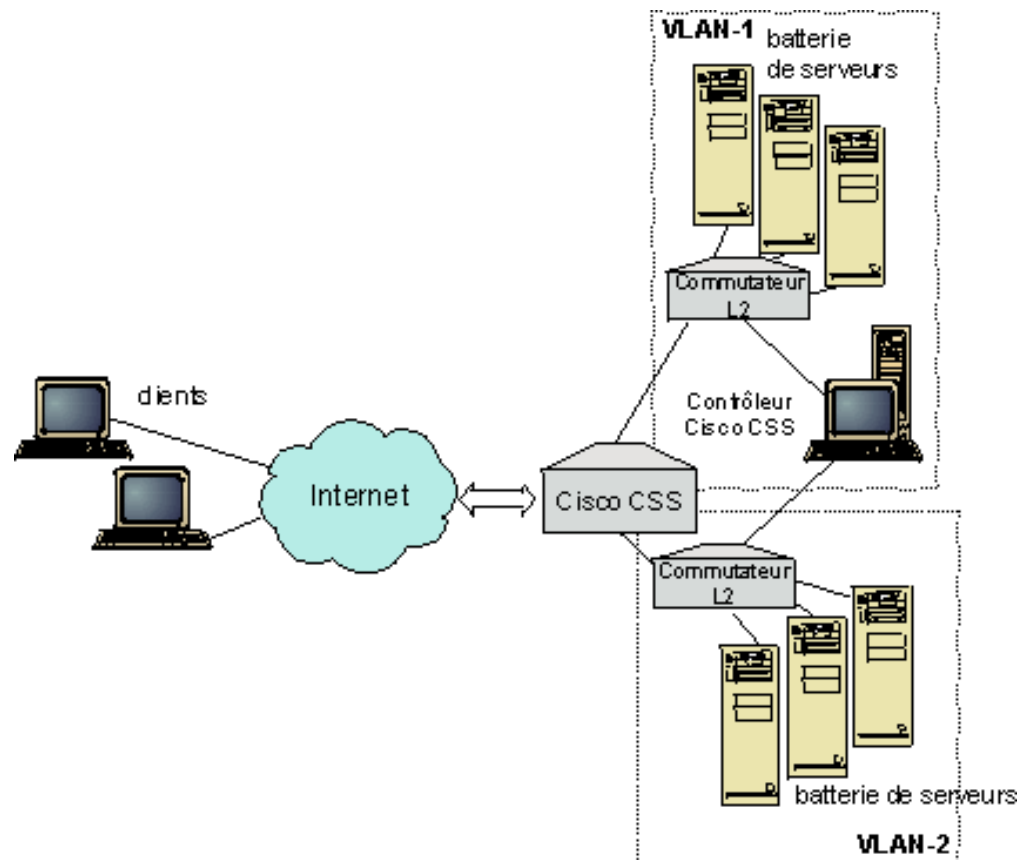


Figure 26. Exemple de consultant connecté derrière les commutateurs

Pour la gestion de Cisco CSS Controller vous pouvez utiliser l'une des interfaces suivantes :

- Un navigateur
- Une interface graphique utilisateur (éloignée ou locale)
- Une ligne de commande (éloignée ou locale)

Pour la gestion à distance, dans la figure 27, à la page 146 :

- Le consultant est connecté derrière le commutateur pour lequel il fournit des pondérations.
- L'interface graphique s'exécute sur un système éloigné devant le commutateur.
- Le commutateur doit être configuré de sorte que le système éloigné puisse, par son intermédiaire, communiquer avec le système contrôleur.

Pour des informations détaillées, reportez-vous au manuel *Cisco Content Services Switch Getting Started Guide*.

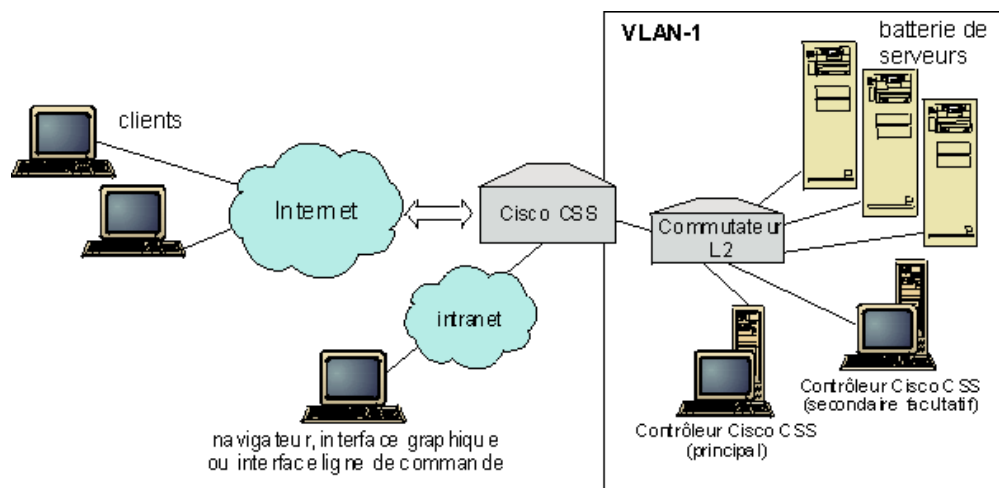


Figure 27. Exemple de consultant (avec partenaire haute disponibilité optionnel), configuré derrière le commutateur avec une interface graphique devant le commutateur

Haute disponibilité

La haute disponibilité du contrôleur augmente la tolérance aux pannes de Load Balancer. Conçue avec un souci de haute disponibilité dans la transmission des paquets, la haute disponibilité du contrôleur implique l'exécution simultanée de deux contrôleurs, l'un assurant le rôle de contrôleur principal, l'autre celui de contrôleur secondaire.

Chaque contrôleur est configuré avec les mêmes informations de commutateur et un seul contrôleur est actif à la fois. Ainsi, du fait de la logique de haute disponibilité, seule le contrôleur actif calcule et met à jour les pondérations.

La haute disponibilité du contrôleur communique avec ses partenaires à l'aide de simples paquets UDP (user datagram protocol) transmis via une adresse et un port que l'utilisateur configure. Ces paquets sont utilisés pour l'échange d'informations entre les contrôleurs dans le cadre de la haute disponibilité (accès aux informations) et pour déterminer la disponibilité du contrôleur des partenaires (signaux de présence). Si le contrôleur secondaire détecte que le contrôleur actif est en erreur pour une raison ou une autre, il prend le relais du contrôleur actif défaillant. Le contrôleur secondaire devient alors le contrôleur actif, et commence à calculer les nouvelles pondérations et à mettre à jour le commutateur avec ces nouvelles valeurs.

Outre sur les partenaires, la haute disponibilité peut être configurée sur les cibles accédées. La haute disponibilité des contrôleurs utilise les informations d'accès pour déterminer le contrôleur actif et le contrôleur secondaire. Le contrôleur actif est celui qui peut contacter (par test ping) le plus de cibles et qui est accessible depuis son partenaire.

Pour plus d'informations, voir «Haute disponibilité», à la page 243.

Calcul des pondérations

Lorsque le consultant détecte qu'un service n'est pas disponible, il met ce service en suspens sur le commutateur pour que ce dernier prenne pas le service en compte lors de l'équilibrage de la charge des demandes. Une fois le service de

nouveau disponible, le consultant active le service sur le commutateur pour qu'il soit de nouveau pris en compte dans l'équilibrage de la charge des demandes.

Identification des incidents

Cisco CSS Controller enregistre des entrées dans les journaux suivants :

- server.log
- consultant.log
- highavailability.log
- metriccollector.log
- binary.log

Ces journaux se trouvent dans les répertoires suivants :

- Pour les systèmes AIX, HP-UX, Linux et Solaris : ...ibm/edge/lb/servers/logs/cco/*nom_consultant*
- Pour les systèmes Windows : ...ibm\edge\lb\servers\logs\cco*nom_consultant*

Vous pouvez définir la taille et le niveau de consignation de chaque journal. Pour plus d'informations, voir «Utilisation des journaux Load Balancer», à la page 265.

Chapitre 17. Configuration de Cisco CSS Controller

Avant d'effectuer les opérations décrites dans le présent chapitre, voir Chapitre 16, «Planification de Cisco CSS Controller», à la page 143. Ce chapitre explique comment créer une configuration de base pour le composant Cisco CSS Controller de Load Balancer.

- Pour des configurations plus complexes, voir Chapitre 23, «Fonctions avancées de Cisco CSS Controller et Nortel Alteon Controller», à la page 243.
- Pour obtenir des informations sur l'administration authentifiée à distance, les fichiers journaux et l'utilisation du composant Cisco CSS Controller, voir Chapitre 24, «Exploitation et gestion de Load Balancer», à la page 261.

Présentation générale des tâches de configuration

Avant de suivre une des méthodes de configuration décrites dans ce chapitre :

1. Assurez-vous que Cisco CSS Switch et que tous les serveurs sont correctement configurés.
2. Configurez Cisco CSS Controller, en vérifiant que l'adresse de Cisco CSS Switch et le nom de la communauté SNMP correspondent aux attributs équivalents de Cisco CSS Switch. Pour plus d'informations sur la configuration du consultant, voir «ccocontrol consultant — Configuration et contrôle d'un consultant», à la page 430.

Tableau 10. Tâches de configuration du composant Cisco CSS Controller

Tâche	Description	Informations connexes
Configuration de la machine Cisco CSS Controller	Conditions requises	«Installation de la machine Contrôleur pour commutateurs Cisco CSS», à la page 152
Test de la configuration	Confirmation du bon fonctionnement de la configuration	«Test de vérification de la configuration», à la page 154

Méthodes de configuration

Trois méthodes permettent de créer une configuration de base pour le composant Cisco CSS Controller de Load Balancer :

- Ligne de commande
- Fichier XML
- Interface graphique

Ligne de commande

Il s'agit de la méthode de configuration de Cisco CSS Controller la plus directe. Les procédures décrites dans ce manuel reposent sur l'utilisation de la ligne de commande. Les valeurs des paramètres de commande doivent être saisies en anglais. Les seules exceptions s'appliquent aux noms d'hôte (utilisés, par exemple, dans la commande **consultant add**) et aux noms de fichiers.

Pour démarrer Cisco CSS Controller à partir de la ligne de commande :

1. Emettez la commande **ccoserver** à partir de l'invite. Pour arrêter le serveur, tapez **ccoserver stop**

Remarques :

- a. Pour les systèmes Windows, cliquez sur **Démarrer > Paramètres** (pour Windows 2000) > **Panneau de configuration > Outils d'administration > Services**. Cliquez à l'aide du bouton droit de la souris sur **IBM Cisco CSS Controller**, puis sélectionnez **Démarrer**. Pour arrêter le service, suivez la même procédure en sélectionnant **Arrêter**.
 - b. Pour les systèmes Windows, vous pouvez démarrer automatiquement **ccoserver** à l'amorçage, comme suit :
 - 1) Cliquez sur **Démarrer > Paramètres > Panneau de configuration > Outils d'administration > Services**.
 - 2) Cliquez à l'aide du bouton droit de la souris sur **IBM Cisco CSS Controller**, puis sélectionnez **Propriétés**.
 - 3) Cliquez sur la flèche de la zone **Type de démarrage**, puis sélectionnez **Automatique**.
 - 4) Cliquez sur **OK**.
2. Emettez ensuite les commandes de contrôle Cisco CSS Controller voulues pour définir votre configuration. Les procédures décrites dans ce manuel reposent sur l'utilisation de la ligne de commande. La commande est **ccocontrol**. Pour plus de détails sur les commandes, voir Chapitre 29, «Guide des commandes Cisco CSS Controller», à la page 429.

Vous pouvez entrer une version abrégée des paramètres de contrôle **ccocontrol**. Il suffit d'entrer les lettres spécifiques des paramètres. Ainsi, pour obtenir de l'aide sur la commande **file save**, vous pouvez entrer **ccocontrol he f** au lieu de **ccocontrol help file**.

Pour démarrer l'interface de ligne de commande, entrez **ccocontrol** afin d'ouvrir une invite **ccocontrol**.

Pour fermer l'interface de ligne de commande, entrez **exit** ou **quit**.

Remarque : Sur les plateformes Windows, le service **dsserver** du composant Dispatcher démarre automatiquement. Si vous utilisez uniquement Cisco CSS Controller et non le composant Dispatcher, vous pouvez empêcher **dsserver** de démarrer automatiquement de la manière suivante :

1. Dans le panneau Services de Windows, cliquez à l'aide du bouton droit de la souris sur **IBM Dispatcher**.
2. Sélectionnez **Propriétés**.
3. Dans la zone **Type de démarrage**, sélectionnez **Manuel**.
4. Cliquez sur **OK** et fermez la fenêtre Services.

XML

La configuration définie peut-être sauvegardée dans un fichier XML. La configuration peut ainsi être chargée ultérieurement lorsque vous voulez la recréer rapidement.

Pour exécuter le contenu d'un fichier XML (par exemple, **monscript.xml**), utilisez l'une ou l'autre des commandes suivantes :

- Pour sauvegarder la configuration courante dans un fichier XML, entrez la commande suivante :
ccocontrol file save *NomFichierXML*
- Pour charger une configuration sauvegardée, entrez la commande suivante :
ccocontrol file load *NomFichierXML*
La commande de chargement (load) n'est utilisable qu'après exécution d'une commande **file save**.

Les fichiers XML sont sauvegardés dans le répertoire **...ibm/edge/lb/servers/configurations/cco/**.

Interface graphique

Pour des instructions générales et un exemple de l'interface graphique, voir figure 41, à la page 470.

Pour démarrer l'interface graphique, procédez comme suit :

1. Si ccoserver n'est pas déjà en cours d'exécution, lancez-le maintenant avec la commande émise en tant que superutilisateur :
ccoserver.
2. Ensuite, procédez de l'une des manières suivantes :
 - Pour les systèmes AIX, HP-UX, Linux ou Solaris : entrez **lbadmin**
 - Pour les systèmes Windows : cliquez sur **Démarrer > Programmes > IBM WebSphere > Edge Components > IBM Load Balancer > Load Balancer**

Pour configurer le composant Cisco CSS Controller à partir de l'interface graphique :

1. Cliquez à l'aide du bouton droit de la souris sur Cisco CSS Controller dans l'arborescence.
2. Connectez-vous à un hôte.
3. Créez un ou plusieurs consultants de commutateur contenant les contenus de propriétaires souhaités et leurs mesures associées.
4. Démarrez le consultant.

Vous pouvez utiliser l'interface graphique pour toute opération effectuée via la commande **ccocontrol**. Par exemple :

- pour définir un consultant à l'aide de la ligne de commande, entrez **ccocontrol consultant add** *IDconsultant* **address** *AdresseIP* **community** *nom*.
- Pour définir un consultant à partir de l'interface graphique, cliquez à l'aide du bouton droit de la souris sur le noeud Hôte, puis sur **Ajouter un consultant de commutateur**. Entrez l'adresse du commutateur et le nom de la communauté dans la fenêtre en incrustation, puis cliquez sur OK.
- Utilisez l'option de **chargement de configuration** du menu en incrustation Hôte pour charger les fichiers de configuration Cisco CSS Controller existants et les annexer à la configuration courante.
- Sélectionnez **Sauvegarder le fichier de configuration en** pour sauvegarder régulièrement la configuration Cisco CSS Controller dans un fichier.
- Cliquez sur **Fichier** dans la barre de menus afin de sauvegarder les connexions à l'hôte en cours dans un fichier ou de restaurer les connexions dans des fichiers existants sur tous les composants Load Balancer.

Pour exécuter une commande à partir de l'interface graphique, procédez comme suit :

1. Cliquez sur le noeud **Hôte** à l'aide du bouton droit de la souris, puis sélectionnez **Envoyer la commande...**
2. Dans la zone d'entrée de commande, entrez la commande à exécuter, par exemple **consultant report**.
3. Cliquez sur Envoyer.

Les résultats et l'historique des commandes exécutées lors de la session courante s'affichent dans la fenêtre Résultats.

Pour accéder à l'**aide**, cliquez sur le point d'interrogation situé dans l'angle supérieur droit de la fenêtre Load Balancer.

- **Aide sur les zones** — décrit les valeurs par défaut de chaque zone.
- **Procédures** — affiche la liste des tâches pouvant être effectuées dans cet écran.
- **Centre de documentation** — fournit un accès centralisé aux informations relatives au produit

Pour plus de détails sur l'utilisation de l'interface graphique, voir Annexe A, «Interface graphique utilisateur : Instructions générales», à la page 469.

Installation de la machine Contrôleur pour commutateurs Cisco CSS

La configuration de la machine Cisco CSS Controller ne peut être effectuée que par le superutilisateur root (pour les systèmes AIX, HP-UX, Linux ou Solaris) ou l'administrateur (pour les systèmes Windows).

Consultant doit pouvoir se connecter à Cisco CSS Switch en tant qu'administrateur Cisco CSS Switch.

Lors de la configuration du consultant, vous devez configurer une adresse et un nom de communauté SNMP qui correspondent aux attributs équivalents de Cisco CSS Switch.

Pour obtenir une aide sur les commandes utilisées lors de cette procédure, voir Chapitre 29, «Guide des commandes Cisco CSS Controller», à la page 429.

Etape 1. Démarrage de la fonction serveur

Si ccoserver ne s'exécute pas déjà, entrez **ccoserver** en tant que superutilisateur pour le démarrer.

Remarque : Pour les systèmes Windows, cliquez sur **Démarrer > Paramètres** (pour Windows 2000) > **Panneau de configuration > Outils d'administration > Services**. Cliquez à l'aide du bouton droit de la souris sur IBM Cisco Controller, puis sélectionnez Démarrer.

Etape 2. Démarrage de l'interface de ligne de commande

Entrez **ccocontrol** pour démarrer l'interface de ligne de commande.

Etape 3. Configuration du consultant

Vous devez configurer le nom de communauté SNMP et l'adresse du commutateur. Ces valeurs doivent correspondre aux attributs équivalents sur Cisco CSS Switch.

Pour ajouter un consultant, entrez :

```
consultant add ID_consultant_commutateur address adresse_IP_commutateur  
community nom_communauté
```

Etape 3. Configuration d'un contenu de propriétaire

Un contenu de propriétaire est une représentation d'une règle de contenu pour un propriétaire, défini sur Cisco CSS Switch. Le nom du propriétaire et le nom de la règle de contenu doivent être définis de la même manière que sur le commutateur.

Pour définir un contenu de propriétaire, entrez :

```
ownercontent add  
ID_consultant_commutateur:ID_contenu_propriétaire ownername  
nom_propriétaire  
contentrule nom_règle_contenu
```

Etape 4. Vérification de la définition des services

Une fois le contenu de propriétaire défini, le consultant complète la configuration en récupérant les services configurés sur le commutateur. Comparez la configuration sur le commutateur et sur le consultant pour vous assurer que les services correspondent.

Etape 5. Configuration des mesures

Les mesures permettent de déterminer les pondérations des services et les proportions associées (importance d'une mesure par rapport à une autre), et peuvent être toute combinaison de mesures de données de connexion, de mesures de conseiller d'application et de mesures de serveur de mesures. La somme des proportions doit toujours être égale à 100.

Lorsque le contenu de propriétaire est configuré, les mesures par défaut définies sont **activeconn** et **connrate**. Si vous voulez des mesures supplémentaires ou différentes des mesures par défaut, entrez :

```
ownercontent metrics ID_consultant_commutateur:ID_contenu_propriétaire  
mesure1 NiveauImportance1  
mesure2 NiveauImportance2...mesureN NiveauImportanceN
```

Etape 6. Lancement du consultant

Pour démarrer le consultant, entrez :

```
consultant start  
ID_consultant_commutateur
```

Les collecteurs de mesure démarrent et le calcul des pondération commence.

Etape 7. Lancement du système Metric Server (facultatif)

Si les mesures système sont définies à l'étape 5, le serveur de mesures doit être démarré sur les machines de service. Pour plus d'informations sur le serveur de mesures, voir «Metric Server», à la page 196.

Etape 8. Configuration de la haute disponibilité (facultatif)

Pour configurer la haute disponibilité, entrez :

```
highavailability add address adresse_IP partneraddress adresse_IP port 80  
role principal
```

Dans un environnement à haute disponibilité, vous pouvez configurer plusieurs commutateurs. Pour garantir que les informations de pondération sont encore disponibles lorsqu'un commutateur prend le relais d'un autre, Cisco CSS Controller doit être configuré de manière à fournir les pondérations de tous les commutateurs et de leurs homologues de secours.

Pour des informations détaillées sur l'emploi et la configuration de la haute disponibilité des composants Controller, voir Chapitre 23, «Fonctions avancées de Cisco CSS Controller et Nortel Alteon Controller», à la page 243.

Test de vérification de la configuration

Vérifiez que la configuration fonctionne :

1. Attribuez la valeur 4 au niveau de consignation du consultant.
2. Déconnectez un serveur de Cisco CSS Switch pendant une minute ou arrêtez le serveur d'applications pendant une minute.
3. Reconnectez le serveur ou démarrez à nouveau le serveur d'applications.
4. Attribuez à nouveau le niveau désiré (1) au niveau de consignation du consultant.
5. Affichez le fichier consultant.log des répertoires suivants et cherchez le **service de définition setServerWeights** :
 - Pour les systèmes AIX, HP-UX, Linux et Solaris : ...ibm/edge/lb/servers/logs/cco/nom_consultant
 - Pour les systèmes Windows : ...ibm\edge\lb\servers\logs\cco\nom_consultant

Partie 6. Composant Nortel Alteon Controller

Cette section contient des informations pour la configuration d'un démarrage rapide ainsi que des remarques relatives à la planification, et présente les diverses méthodes de configuration du composant Nortel Alteon Controller de Load Balancer. Elle se compose des chapitres suivants :

- Chapitre 18, «Configuration de démarrage rapide», à la page 157
- Chapitre 19, «Planification de Nortel Alteon Controller», à la page 161
- Chapitre 20, «Configuration de Nortel Alteon Controller», à la page 171

Chapitre 18. Configuration de démarrage rapide

Cet exemple de démarrage rapide montre comment créer une configuration à l'aide du composant Nortel Alteon Controller. Nortel Alteon Controller fournit des pondérations de serveur à Nortel Alteon Web Switch. Ces pondérations sont utilisées afin de sélectionner des serveurs pour des services dont le commutateur équilibre la charge.

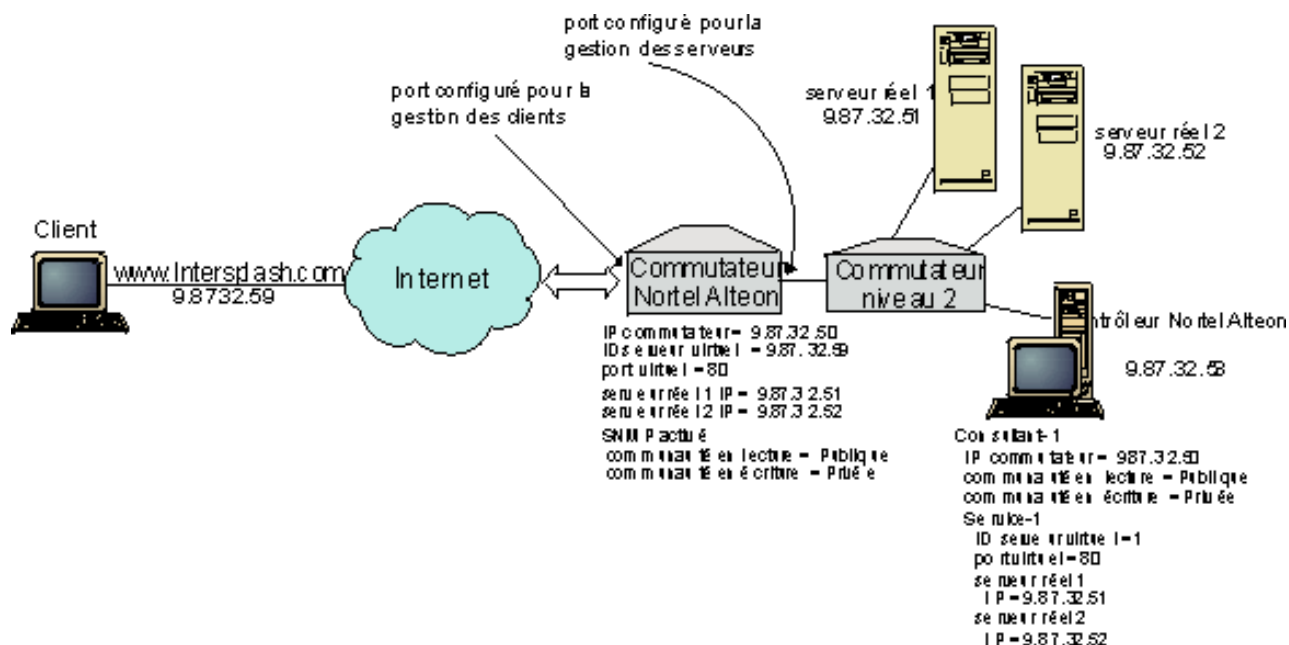


Figure 28. Configuration Nortel Alteon Controller simple

Matériel requis

Cet exemple de configuration de démarrage rapide nécessite :

- Un Nortel Alteon Web Switch exécutant le système d'exploitation Web version 9.0 ou 10.0
- Un serveur doté du composant Nortel Alteon Controller
- Deux serveurs Web
- Un commutateur niveau 2 connecté à un port du Nortel Alteon Web Switch

Remarque : Si vous n'utilisez pas de commutateur niveau 2, vous pouvez connecter la machine Nortel Alteon Controller et les serveurs Web directement aux ports du Nortel Alteon Web Switch.

- Cet exemple de configuration nécessite cinq types d'adresses IP :
 - l'adresse IP qui permet aux clients d'accéder à votre site Web site, `www.Intersplashx.com` (`9.87.32.59`),
 - l'adresse IP d'une interface configurée sur le Nortel Alteon Web Switch (`9.87.32.50`),
 - l'adresse IP du serveur réel 1 (`9.87.32.51`),
 - l'adresse IP du serveur réel 2 (`9.87.32.52`),

- l'adresse IP du serveur Nortel Alteon Controller (9.87.32.53).

Préparation

Avant de d'effectuer la configuration de cet exemple, procédez aux vérifications suivantes :

- Assurez-vous comme suit que Nortel Alteon Web Switch est correctement configuré. (Pour plus d'informations, reportez-vous au manuel Nortel Alteon Web OS Application Guide) :
 - Activez l'équilibrage de charge du serveur niveau 4 sur le commutateur.
 - Configurez une interface IP (9.87.32.50) sur Nortel Alteon Web Switch.
 - Activez SNMP sur Nortel Alteon Web Switch.
 - Activez le traitement du client d'équilibrage de charge du serveur sur le port de Nortel Alteon Web Switch qui reçoit les demandes des clients.
 - Activez le traitement du serveur d'équilibrage de charge du serveur sur le port de Nortel Alteon Web Switch auquel les serveurs se connectent.
 - Configurez la passerelle par défaut comme interface IP du commutateur (9.87.32.50) sur le serveur réel 1, le serveur réel 2 et Nortel Alteon Controller.
 - Configurez Nortel Alteon Web Switch avec le serveur réel 1 et le serveur réel 2.
 - Configurez Nortel Alteon Web Switch avec un groupe de serveurs constitué du serveur réel 1 et du serveur réel 2, et attribuez à ce groupe l'ID 1.
 - Configurez Nortel Alteon Web Switch avec un serveur virtuel dont l'adresse IP est 9.87.32.59. Attribuez l'ID 1 à ce serveur virtuel.
 - Configurez Nortel Alteon Web Switch avec un service utilisant le port virtuel 80 et pris en charge par le groupe 1.
- Assurez-vous que la machine client peut contacter l'adresse IP du serveur virtuel, 9.87.32.59.
- Assurez-vous que la machine Nortel Alteon Controller peut contacter l'interface IP du Nortel Alteon Web Switch (9.87.32.50), le serveur réel 1 (9.87.32.51) et le serveur réel 2 (9.87.32.52).

Configuration du composant Nortel Alteon Controller

Avec Nortel Alteon Controller, vous pouvez créer une configuration à l'aide de la ligne de commande ou de l'interface graphique. Pour cet exemple de démarrage rapide, les étapes de configuration s'effectuent via la ligne de commande.

Remarque : Les valeurs des paramètres doivent être saisies à l'aide de caractères anglais. Les seules exceptions sont les valeurs des paramètres des noms d'hôte et des noms de fichiers.

Configuration à partir de la ligne de commande

A partir d'une invite, effectuez les opérations ci-dessous.

1. Démarrez nalservice sur Nortel Alteon Controller. En tant que superutilisateur ou administrateur, entrez la commande **nalservice** à partir d'une invite.
2. Ajoutez un consultant à la configuration Nortel Alteon Controller, en précisant l'adresse IP de l'interface Nortel Alteon Web Switch. (N'indiquez les communautés en lecture et en écriture que si elles diffèrent de celle par défaut (publique, privée)) :

nalcontrol consultant add Consultant-1 address 9.87.32.50

Cette opération vérifie la connectivité au Nortel Alteon Web Switch vérifie que les noms de communauté SNMP fonctionnent correctement.

3. Ajoutez un service (Service-1) au consultant (Consultant-1), en précisant l'identificateur de serveur virtuel (1) et le numéro de port virtuel (80) du service :

nalcontrol service add Consultant-1:Service-1 vsid 1 vport 80

Nortel Alteon Controller communiquera avec le commutateur via SNMP et obtiendra de celui-ci les informations de configuration nécessaires. Une fois cette procédure effectuée, Nortel Alteon Controller contiendra les informations relatives aux serveurs configurés sur Nortel Alteon Web Switch pour le service.

4. Configurez les mesures à collecter pour l'ensemble de serveurs associé au service :

nalcontrol service metrics Consultant-1:Service-1 http 40 activeconn 30 connrate 30

Cette commande configure les informations de mesure à collecter auprès des serveurs et l'importance relative de ces mesures lors du calcul des pondérations.

5. Démarrez la fonction consultant de Nortel Alteon Controller :

nalcontrol consultant start Consultant-1

Cette commande lance tous les collecteurs de mesures et démarre les calculs de pondération de serveur. Nortel Alteon Controller communique les résultats de ses calculs de pondération de serveur à Nortel Alteon Web Switch à l'aide de SNMP.

La configuration de base de Nortel Alteon Controller est maintenant terminée.

Test de vérification de la configuration

Vérifiez que la configuration fonctionne :

1. A l'aide du navigateur Web du client, accédez à **http://www.Intersplashx.com** . Si une page s'affiche, la configuration fonctionne.
2. Rechargez la page dans le navigateur Web.
3. Vérifiez les résultats de la commande suivante : **nalcontrol service report Consultant-1:Service-1**. La colonne du nombre total de connexions des deux serveurs Web doit contenir la valeur "2".

Configuration à l'aide de l'interface graphique

Pour plus d'informations sur l'utilisation de l'interface graphique de Nortel Alteon Controller, voir «Interface graphique», à la page 173 et à l'Annexe A, «Interface graphique utilisateur : Instructions générales», à la page 469.

Chapitre 19. Planification de Nortel Alteon Controller

Le présent chapitre décrit les aspects qu'un administrateur de réseau doit prendre en compte avant d'installer et de configurer le composant Nortel Alteon Controller.

- Pour obtenir des informations sur la configuration des paramètres d'équilibrage de charge du composant Nortel Alteon Controller, voir Chapitre 20, «Configuration de Nortel Alteon Controller», à la page 171.
- Pour la procédure de configuration des conseillers et des serveurs de mesures, voir Chapitre 23, «Fonctions avancées de Cisco CSS Controller et Nortel Alteon Controller», à la page 243.
- Pour des informations sur l'administration authentifiée à distance, les fichiers journaux Load Balancer et l'utilisation des composants Load Balancer, voir Chapitre 24, «Exploitation et gestion de Load Balancer», à la page 261.

Le présent chapitre se compose des sections suivantes :

- «Configuration requise»
- «Remarques relatives à la planification»
 - «Positionnement du consultant dans le réseau», à la page 162
 - «Attributs de serveur sur le commutateur (définis par le contrôleur)», à la page 164
 - «Configuration de serveurs de secours», à la page 165
 - «Configuration de groupes», à la page 166
 - «Haute disponibilité», à la page 166
 - «Optimisation», à la page 168
 - «Identification des incidents», à la page 169

Configuration requise

Pour plus d'informations sur les conditions matérielles et logicielles requises, accédez à la page Web suivante : <http://www.ibm.com/support/docview.wss?rs=180&uid=swg27006921>

Vous avez également besoin des éléments suivants :

- un système sur lequel exécuter le Nortel Alteon Controller,
- un composant Nortel Alteon Web Switch installé et configuré. Les plateformes matérielles de commutateur Web sont AD3, AD4, 180e 184 et le routeur niveau 4/7 pour Passport 8600.

Remarques relatives à la planification

Nortel Alteon Controller gère une série de consultants de commutateur. Chaque consultant détermine les pondérations des serveurs dont la charge est équilibrée par un seul commutateur. Le commutateur auquel le consultant fournit les pondérations est configuré pour l'équilibrage de charge de contenu. Le consultant utilise le protocole SNMP pour transmettre au commutateur les pondérations calculées. Le commutateur utilise les pondérations reçues pour sélectionner un serveur pour le service dont il équilibre la charge. Pour déterminer les pondérations, le consultant utilise un ou plusieurs des éléments d'information suivants :

- Disponibilité et temps de réponse, déterminés à l'aide de **conseillers** qui communiquent avec les applications qui s'exécutent sur les serveurs.
- Informations relatives à la charge système, déterminée par extraction d'une valeur de mesure des **agents de mesure** qui s'exécutent sur les serveurs.
- Informations de connexion relatives aux serveurs, provenant du commutateur.
- Informations d'accessibilité, extraites en contactant les serveurs.

Pour une description de l'équilibrage de la charge des serveurs et des informations détaillées sur la configuration du commutateur, reportez-vous au manuel "Nortel Alteon Web OS Application Guide".

Pour qu'un consultant obtienne les informations dont il a besoin afin de déterminer les pondérations d'un serveur, les éléments suivants sont nécessaires :

- Connectivité IP entre le consultant et les serveurs pour lesquels les pondérations sont calculées.
- Connectivité IP entre le consultant et le commutateur qui équilibre la charge des serveurs pour lesquels les pondérations sont calculées.
- SNMP activé sur le commutateur. Les fonctions de lecture et d'écriture doivent être activées.

Positionnement du consultant dans le réseau

Le consultant peut être connecté au réseau devant ou derrière le ou les commutateurs pour lesquels il fournit des pondérations. Certains paramètres doivent être configurés sur le commutateur, d'autres sur le contrôleur, pour activer la connectivité entre le contrôleur, le commutateur et les serveurs.

Dans la figure 29, à la page 163 :

- Un consultant est connecté au réseau derrière les commutateurs pour lesquels il fournit des pondérations.
- Le réseau est constitué de deux réseaux VLAN.
- Pour que le consultant communique avec les serveurs des deux réseaux VLAN, la transmission IP doit être activée sur les interfaces par lesquelles les serveurs sont connectés et sur celle par laquelle le consultant est connecté.
- L'adresse IP du commutateur doit être configurée comme passerelle par défaut sur le consultant et sur les systèmes serveurs.

Pour des informations détaillées sur la configuration des réseaux VLAN et du routage IP sur le commutateur, reportez-vous au manuel "Nortel Alteon Web OS Application Guide" ou au Guide des commandes.

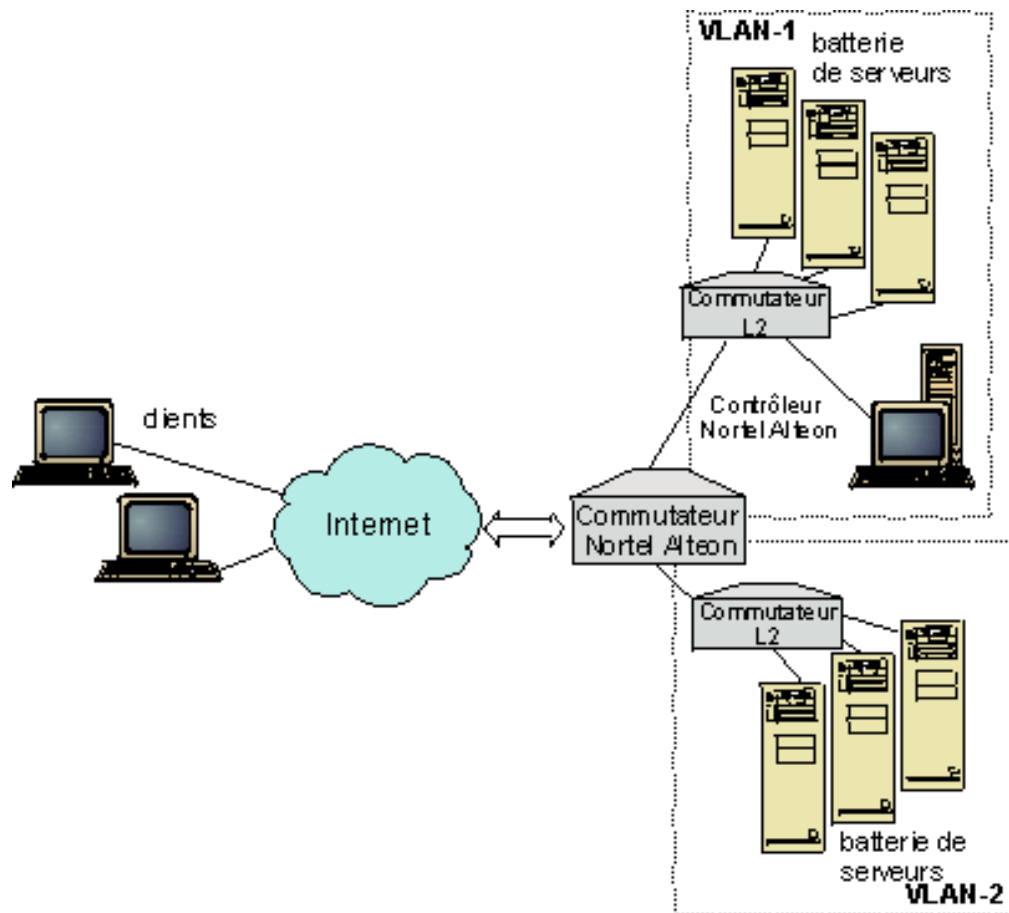


Figure 29. Exemple de consultant connecté derrière le commutateur

Dans la figure 30, à la page 164 :

- Le consultant est connecté au commutateur via un intranet situé devant le commutateur.
- Le mode d'accès direct d'équilibrage de la charge des serveurs doit être activé sur le commutateur pour que le consultant puisse communiquer avec le commutateur et les serveurs.
- Une fois ce mode activé, tous les clients peuvent envoyer des demandes directement à n'importe quel serveur. Pour réserver l'accès direct aux serveurs au consultant, vous pouvez indiquer l'équilibrage de charge *mnet* et *mmask* au commutateur. Pour des informations détaillées sur la configuration de l'équilibrage de la charge des serveurs et sur l'interaction directe des serveurs, reportez-vous au manuel "Nortel Alteon Web OS Application Guide" ou au Guide des commandes.

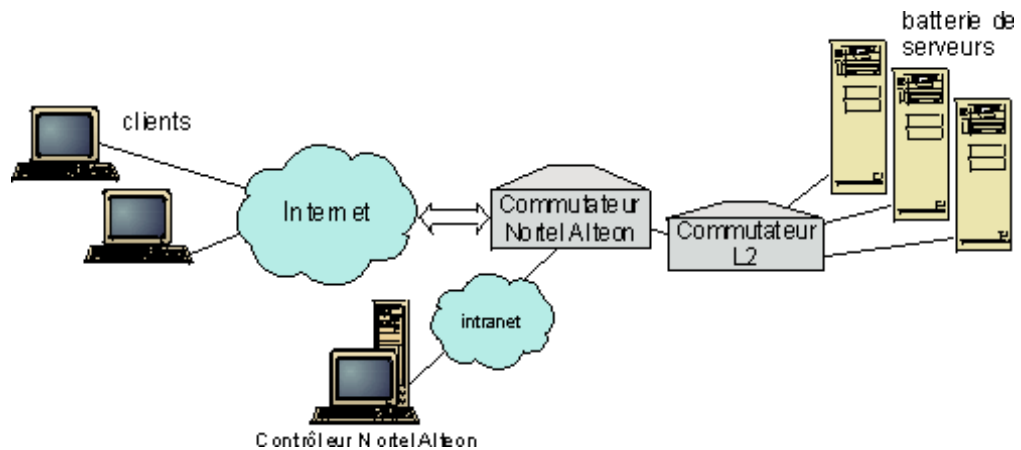


Figure 30. Exemple de consultant connecté via un intranet situé devant le commutateur

Pour la gestion de Nortel Alteon Controller vous pouvez utiliser l'une des interfaces suivantes :

- Un navigateur
- Une interface graphique utilisateur
- Une ligne de commande éloignée

Dans la figure 31 :

- Le consultant est connecté derrière le commutateur pour lequel il fournit des pondérations.
- L'interface graphique s'exécute sur un système éloigné devant le commutateur.
- Le réseau doit être configuré de sorte que l'interface graphique puisse communiquer avec le contrôleur.

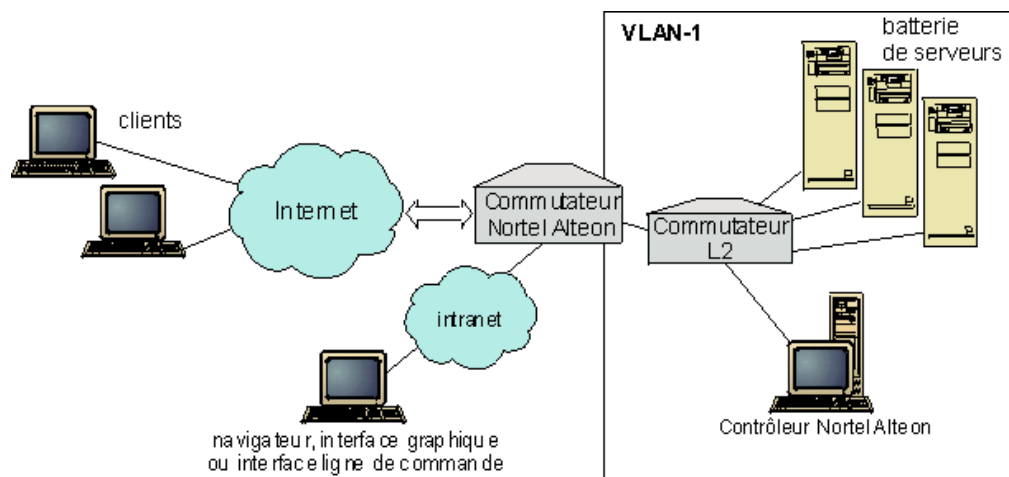


Figure 31. Exemple de consultant derrière le commutateur et d'interface graphique devant le commutateur

Attributs de serveur sur le commutateur (définis par le contrôleur)

Lorsqu'un consultant calcule les pondérations des serveurs qui fournissent un service dont la charge est équilibrée par un commutateur, ce consultant désactive la vérification normale de l'état des serveurs sur le commutateur afin de réduire

tout trafic inutile vers les serveurs. Le consultant réactive la vérification d'état lorsqu'il arrête de fournir des pondérations pour le service. L'intervalle de vérification de l'état d'un serveur est défini par la variable MIB `slbNewCgRealServerPingInterval`.

Lorsque le consultant détecte qu'un serveur n'est pas disponible, il fixe le nombre maximum de connexions du serveur à zéro pour que le commutateur ne prenne pas le serveur en compte lorsqu'il équilibre la charge des demandes. Lorsque le serveur est de nouveau disponible, le nombre maximum de connexions revient à sa valeur d'origine. Le nombre maximum de connexions d'un serveur est défini par la variable MIB `slbNewCgRealServerMaxCons`.

Lorsqu'une pondération est calculée pour un serveur réel, elle est définie pour le serveur. La valeur de pondération d'un serveur est définie par la variable MIB `slbNewCgRealServerWeight`.

Configuration de serveurs de secours

le commutateur permet de configurer certains serveurs en tant que serveurs de secours pour les autres. Lorsque le commutateur détecte qu'un serveur doté d'un serveur de secours n'est pas disponible, il peut commencer à envoyer les demandes à ce serveur de secours. Lorsque le consultant calcule les pondérations pour un service doté d'un serveur de secours, il effectue ce calcul à la fois pour les serveurs principaux et pour les serveurs de secours de sorte qu'il dispose des pondérations requises lorsque la sélection d'un serveur de secours s'avère nécessaire.

La pondération d'un serveur de secours peut être plus élevée que celle d'un serveur principal. En effet, aucune demande ne lui étant transmise, sa charge est faible tant que le commutateur ne l'utilise pas.

Pour éviter les serveurs inactifs, les serveurs d'un service sont généralement les serveurs de secours des serveurs d'un autre service. Lors de la mise en oeuvre d'une configuration de ce type, évitez d'affecter les mêmes serveurs réels à plusieurs services simultanément actifs. Si cela se produit, le consultant remplace la pondération du serveur pour chaque service dont le serveur fait partie.

Chaque serveur réel est identifié par un entier et dispose d'un attribut de pondération et d'adresse IP. Deux serveurs réels peuvent avoir la même adresse IP. Auquel cas, les deux serveurs réels sont associés au même serveur physique. Les serveurs réels identifiés comme serveurs de secours ne peuvent être configurés comme tels que pour un seul service. Si les mêmes serveurs physiques assurent la sauvegarde de serveurs affectés à plusieurs services, ils doivent être configurés une fois pour chaque service et recevoir une identification de serveur propre à chaque service. Les serveurs de secours ont ainsi une pondération unique affectée pour chaque service dont ils assurent la sauvegarde.

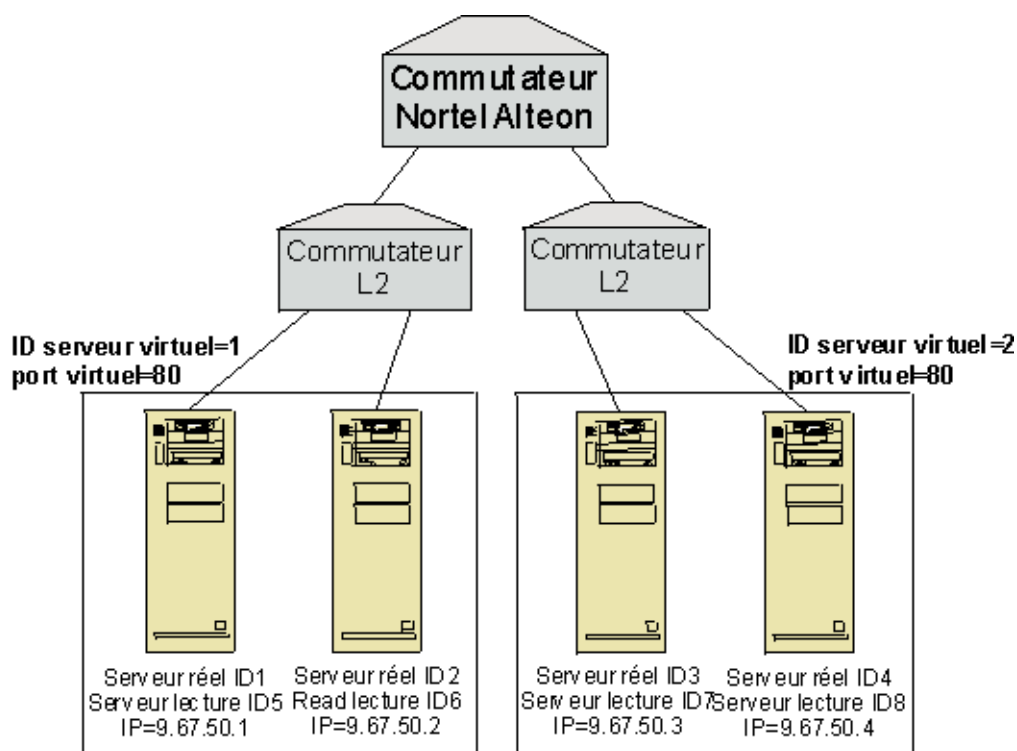


Figure 32. Exemple de consultant configuré avec des serveurs de secours

Configuration de groupes

Les serveurs d'un commutateur peuvent être configurés comme appartenant à plusieurs groupes et les groupes du commutateur configurés pour prendre en charge plusieurs services.

Comme il est possible de configurer le même serveur pour plusieurs services, la pondération est calculée pour chaque service auquel le serveur participe. La pondération peut donc parfois être incorrecte car le service pour lequel elle est prévue n'est pas connu en permanence.

De plus, si le consultant détermine les pondérations pour un service et pas pour un autre, il est possible que la vérification de l'état des serveurs soit désactivée sur le service pour lequel le consultant ne calcule pas les pondérations. Dans ce cas, le commutateur risque de ne pas correctement équilibrer la charge de ce service.

En conséquence, vous devez vous assurer qu'un serveur réel n'a pas été affecté à plusieurs services dont la charge est équilibrée. Cela ne signifie pas qu'un même serveur ne peut pas traiter les demandes de plusieurs services, mais qu'un serveur réel ayant un seul identificateur doit être configuré sur le commutateur de chaque service dont le serveur gère les demandes.

Haute disponibilité

Nortel Alteon Controller et Nortel Alteon Web Switch disposent tous deux de fonctions de haute disponibilité.

Vous pouvez configurer deux contrôleurs pour s'exécuter sur différents systèmes dans une configuration de secours automatique.

Deux commutateurs ou plus peuvent se servir mutuellement de serveur de secours lorsque vous les configurez en tant que routeur d'interface IP virtuelle (VIR) ou de routeur de serveur IP virtuel (VSR).

Un consultant (géré par le contrôleur) fournit des pondérations pour un commutateur uniquement. Un commutateur de secours pouvant prendre le relais du commutateur principal, vous devez configurer le contrôleur avec un consultant pour chaque commutateur ayant la possibilité de devenir le commutateur principal. De cette manière, lorsqu'un commutateur devient le principal, vous avez la garantie qu'il reçoit les pondérations requises.

En outre, lorsque les contrôleurs sont connectés à un VIR, ils peuvent communiquer avec les serveurs, les commutateurs et le contrôleur de secours, même s'ils perdent la connectivité à l'un des commutateurs.

Pour des informations détaillées sur la haute disponibilité au niveau du commutateur, reportez-vous au manuel "Nortel Alteon Web OS Application Guide".

La haute disponibilité du contrôleur augmente la tolérance aux pannes de Load Balancer. Conçue avec un souci de haute disponibilité dans la transmission des paquets, la haute disponibilité du contrôleur implique l'exécution simultanée de deux contrôleurs, l'un assurant le rôle de contrôleur principal, l'autre celui de contrôleur secondaire.

Chaque contrôleur est configuré avec les mêmes informations de commutateur. Comme avec la haute disponibilité classique, un seul contrôleur est actif à la fois. Ainsi, du fait de la logique de haute disponibilité, seule le contrôleur actif calcule et met à jour les pondérations.

La haute disponibilité du contrôleur communique avec ses partenaires à l'aide de simples paquets UDP (user datagram protocol) transmis via une adresse et un port que l'utilisateur configure. Ces paquets sont utilisés pour l'échange d'informations entre les contrôleurs dans le cadre de la haute disponibilité (accès aux informations) et pour déterminer la disponibilité du contrôleur des partenaires (signaux de présence). Si le contrôleur secondaire détecte que le contrôleur actif est en erreur pour une raison ou une autre, il prend le relais du contrôleur actif défaillant. Le contrôleur secondaire devient alors le contrôleur actif, et commence à calculer les nouvelles pondérations et à mettre à jour le commutateur avec ces nouvelles valeurs.

Outre sur les partenaires, la haute disponibilité peut être configurée sur les cibles accédées. Comme avec la haute disponibilité classique, la haute disponibilité des contrôleurs utilise les informations d'accès pour déterminer le contrôleur actif et le contrôleur secondaire. Le contrôleur actif est celui qui peut contacter (par test ping) le plus de cibles et qui est accessible depuis son partenaire.

Pour plus d'informations, voir «Haute disponibilité», à la page 243.

Dans la figure 33, à la page 168 :

- Deux contrôleurs Nortel Alteon Controller sont connectés derrière les commutateurs.
- L'un est le contrôleur principal qui alimente les commutateurs en pondérations de serveur, l'autre est le contrôleur de secours.

- Les contrôleurs doivent disposer d'une liaison TCP/IP pour que le contrôleur de secours sache quand il doit prendre le relais du contrôleur principal.
- Deux commutateurs Nortel Alteon Web Switch sont configurés, l'un comme VIR l'autre comme VSR.
- Le routeur VIR fournit la haute disponibilité pour les connexions aux serveurs.
- Le routeur VSR fournit la haute disponibilité pour l'accès aux serveurs virtuels configurés sur les commutateurs.
- L'un des commutateurs est le commutateur principal, l'autre le commutateur de secours.
- Le contrôleur principal fournit des pondérations aux deux commutateurs.
- Le contrôleur de secours envoie des signaux de présence au contrôleur principal pour déterminer quand il doit prendre le relais.

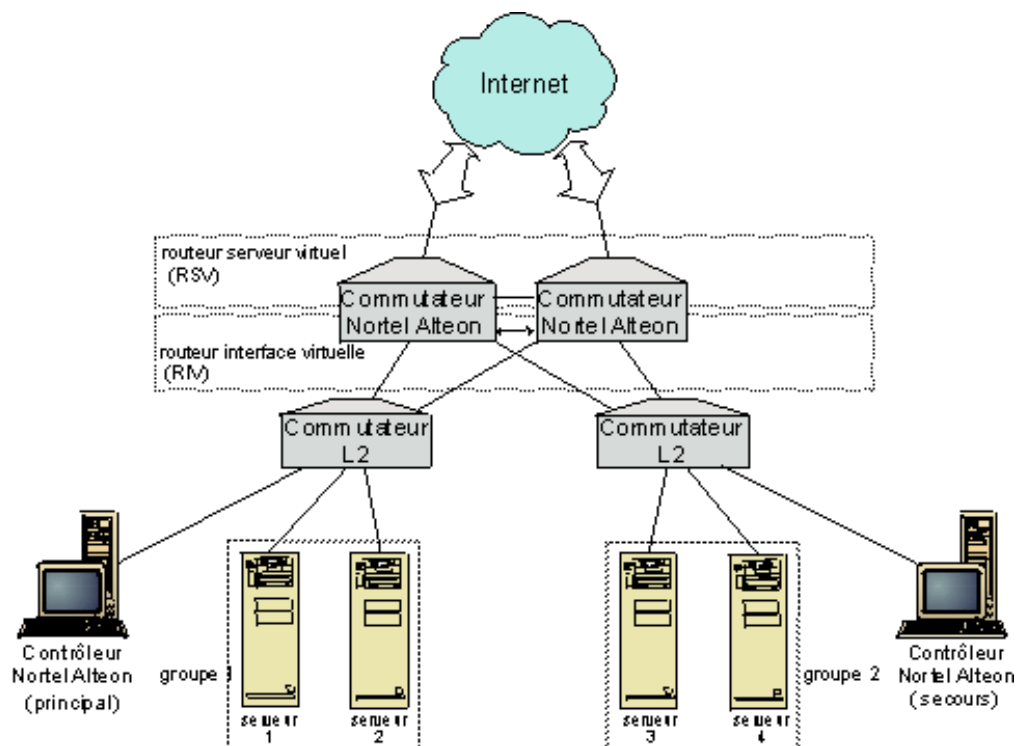


Figure 33. Exemple de Nortel Alteon Controller et de Nortel Alteon Web Switch haute disponibilité

Optimisation

Pour ne pas avoir à modifier trop souvent les pondérations, vous pouvez configurer un seuil de sensibilité pour le consultant. Le seuil de sensibilité indique la quantité de modifications requises entre la nouvelle et l'ancienne pondérations pour que la pondération soit changée. Pour plus d'informations, voir «Seuil de sensibilité», à la page 248.

Si le commutateur est trop occupé à mettre à jour des pondérations, vous pouvez augmenter la durée d'inactivité du consultant pour réduire le trafic entre le contrôleur et les serveurs plus le commutateur. La durée d'inactivité fixe le nombre de secondes entre chaque cycle de définition des pondérations.

Si les serveurs gèrent trop de demandes de surveillance provenant du consultant, vous pouvez modifier la durée d'inactivité des collecteurs de mesures. Pour plus d'informations, voir «Délai d'inactivité dans le calcul des pondérations», à la page 247.

Identification des incidents

Cisco CSS Controller enregistre des entrées dans les journaux suivants :

- server.log
- consultant.log
- highavailability.log
- metriccollector.log
- binary.log

Ces journaux se trouvent dans les répertoires suivants :

- Pour les systèmes AIX, HP-UX, Linux et Solaris : `...ibm/edge/lb/servers/logs/nal/nom_consultant`
- Pour les systèmes Windows : `...ibm\edge\lb\servers\logs\nal\nom_consultant`

Vous pouvez définir la taille et le niveau de consignation de chaque journal. Pour plus d'informations, voir «Utilisation des journaux Load Balancer», à la page 265.

Chapitre 20. Configuration de Nortel Alteon Controller

Avant d'effectuer les opérations décrites dans le présent chapitre, voir Chapitre 19, «Planification de Nortel Alteon Controller», à la page 161. Ce chapitre décrit comment créer une configuration de base pour le composant Nortel Alteon Controller de Load Balancer.

- Pour plus d'informations, voir Chapitre 23, «Fonctions avancées de Cisco CSS Controller et Nortel Alteon Controller», à la page 243.
- Pour plus d'informations sur l'administration authentifiée à distance, les fichiers journaux et l'utilisation du composant Nortel Alteon Controller, voir Chapitre 24, «Exploitation et gestion de Load Balancer», à la page 261.

Présentation générale des tâches de configuration

Avant de suivre une des méthodes de configuration décrites dans ce chapitre, vérifiez que Nortel Alteon Web Switch et tous les serveurs sont correctement configurés.

Tableau 11. Tâches de configuration pour le composant Nortel Alteon Controller

Tâche	Description	Informations connexes
Configurer Nortel Alteon Web Switch et les serveurs	Configuration du commutateur.	Configuration du commutateur, page 174
Configurer la machine Nortel Alteon Controller	Configuration du contrôleur.	«Etape 1. Démarrage de la fonction serveur», à la page 174
Test de la configuration	Confirmation du bon fonctionnement de la configuration	«Test de vérification de la configuration», à la page 176

Méthodes de configuration

Trois méthodes permettent de créer une configuration de base pour le composant Nortel Alteon Controller de Load Balancer :

- Ligne de commande
- Fichier XML
- Interface graphique

Ligne de commande

Il s'agit de la méthode la plus directe pour configurer Nortel Alteon Controller. Les procédures décrites dans ce manuel reposent sur l'utilisation de la ligne de commande.

Pour démarrer Nortel Alteon Controller à partir de la ligne de commande :

1. Emettez la commande **nalserver** à partir de l'invite. Pour arrêter le service, tapez **nalserver stop**.

Remarques :

- a. Pour les systèmes Windows, cliquez sur **Démarrer > Paramètres** (pour Windows 2000) > **Panneau de configuration > Outils d'administration > Services**. Cliquez à l'aide du bouton droit de la souris sur IBM Nortel

Alteon Controller, puis sélectionnez Démarrer. Pour arrêter le service, suivez la même procédure en sélectionnant Arrêter.

- b. Pour les systèmes Windows, vous pouvez démarrer automatiquement nalservice à l'amorçage, comme suit :
 - 1) Cliquez sur **Démarrer > Paramètres** (pour Windows 2000) > **Panneau de configuration > Outils d'administration > Services**.
 - 2) Cliquez à l'aide du bouton droit de la souris sur IBM Nortel Alteon Controller, puis sélectionnez Propriétés.
 - 3) Cliquez sur la flèche de la zone Type de démarrage, puis sélectionnez Automatique.
 - 4) Cliquez sur OK.
2. Emettez ensuite les commandes de contrôle Nortel Alteon Controller voulues pour définir votre configuration. Les procédures décrites dans ce manuel reposent sur l'utilisation de la ligne de commande. La commande est **nalcontrol**. Pour plus de détails sur les commandes, voir Chapitre 30, «Guide des commandes Nortel Alteon Controller», à la page 449.

Vous pouvez utiliser une version abrégée des paramètres de la commande **nalcontrol** en entrant simplement la ou les quelques lettres d'identification des paramètres. Par exemple, pour obtenir l'aide correspondant à la commande **file save**, vous pouvez entrer **nalcontrol he f** au lieu de **nalcontrol help file**.

Pour fermer l'interface de ligne de commande, entrez **exit** or **quit**.

Remarques :

1. Utilisez les lettres de l'anglais pour toutes les valeurs des paramètres des commandes. Les seules exceptions s'appliquent aux noms d'hôte (utilisés dans les commandes **server**) et aux noms de fichiers (utilisés dans les commandes **file**).
2. Pour les systèmes Windows, le service **dsservice** du composant Dispatcher démarre automatiquement. Si vous utilisez uniquement Nortel Alteon Controller et non le composant Dispatcher, vous pouvez empêcher **ndservice** de démarrer automatiquement de la manière suivante :
 - a. Dans le panneau Services de Windows, cliquez à l'aide du bouton droit de la souris sur IBM Dispatcher.
 - b. Sélectionnez Propriétés.
 - c. Dans la zone **Type de démarrage**, sélectionnez Manuel.
 - d. Cliquez sur OK et fermez la fenêtre Services.

XML

La configuration définie peut-être sauvegardée dans un fichier XML. La configuration peut ainsi être chargée ultérieurement lorsque vous voulez la recréer rapidement.

Pour exécuter le contenu d'un fichier XML (par exemple, **monscript.xml**), utilisez les commandes suivantes :

- Pour sauvegarder la configuration courante dans un fichier XML, entrez la commande suivante :

nalcontrol file save XMLFilename

La commande de chargement (load) n'est utilisable qu'après exécution d'une commande **file save**.

- Pour charger une configuration sauvegardée, entrez la commande suivante :

```
nalcontrol file  
load XMLFileName
```

La commande de chargement (load) n'est utilisable qu'après exécution d'une commande **file save**.

Les fichiers XML sont sauvegardés dans le répertoire **...ibm/edge/lb/servers/configurations/nal/**.

Interface graphique

Pour avoir un exemple de l'interface graphique, voir figure 41, à la page 470.

Pour démarrer l'interface graphique, procédez comme suit :

1. Si nalservice n'est pas déjà en cours d'exécution, lancez-le maintenant avec la commande **nalservice** émise en tant que superutilisateur.
2. Ensuite, procédez de l'une des manières suivantes :
 - Pour les systèmes AIX, HP-UX, Linux ou Solaris : entrez **ladmin**
 - Pour les systèmes Windows : cliquez sur **Démarrer > Programmes > IBM WebSphere > Edge Components > IBM Load Balancer > Load Balancer**

Pour configurer le composant Nortel Alteon Controller à partir de l'interface graphique :

1. Cliquez à l'aide du bouton droit de la souris sur Nortel Alteon Controller dans l'arborescence.
2. Connectez-vous à un hôte.
3. Créez un ou plusieurs consultants de commutateur contenant les services souhaités et leurs mesures associées.
4. Démarrez le consultant.

Vous pouvez utiliser l'interface graphique pour toute opération effectuée via la commande **nalcontrol**. Par exemple :

- Pour définir une cible à contacter à l'aide de la ligne de commande, entrez **nalcontrol highavailability usereach adresse**. Pour définir une cible à contacter à partir de l'interface graphique, cliquez à l'aide du bouton droit de la souris sur Haute disponibilité > Ajouter une cible à contacter... Entrez l'adresse à contacter dans la fenêtre en incrustation, puis cliquez sur OK.
- Utilisez l'option de **chargement de configuration** du menu en incrustation Hôte pour annexer la configuration stockée dans un fichier à la configuration courante. Si vous voulez charger une *nouvelle* configuration, vous devez arrêter puis redémarrer le serveur avant de charger le fichier de la nouvelle configuration.
- Cliquez sur le noeud Hôte à l'aide du bouton droit de la souris, puis sélectionnez **Sauvegarder le fichier de configuration en** pour sauvegarder de façon régulière la configuration Nortel Alteon Controller dans un fichier.
- Cliquez sur **Fichier** dans la barre de menus afin de sauvegarder les connexions à l'hôte en cours dans un fichier ou de restaurer les connexions dans des fichiers existants sur tous les composants Load Balancer.

Pour exécuter une commande à partir de l'interface graphique, procédez comme suit :

1. Cliquez sur le noeud **Hôte** à l'aide du bouton droit de la souris, puis sélectionnez **Envoyer la commande...**

2. Dans la zone d'entrée de commande, entrez la commande à exécuter, par exemple **consultant report**.
3. Cliquez sur Envoyer.

Les résultats et l'historique des commandes exécutées lors de la session courante s'affichent dans la fenêtre Résultats.

Pour accéder à l'aide, cliquez sur le point d'interrogation situé dans l'angle supérieur droit de la fenêtre Load Balancer.

- **Aide sur les zones** — décrit les valeurs par défaut de chaque zone.
- **Procédures** — affiche la liste des tâches pouvant être effectuées dans cet écran.
- **Centre de documentation** — fournit un accès centralisé aux informations relatives au produit

Pour plus de détails sur l'utilisation de l'interface graphique, voir Annexe A, «Interface graphique utilisateur : Instructions générales», à la page 469.

Installation de Nortel Alteon Controller

Pour obtenir une aide sur les commandes utilisées lors de cette procédure, voir Chapitre 30, «Guide des commandes Nortel Alteon Controller», à la page 449.

Avant de configurer la machine Nortel Alteon Controller :

- Vous devez posséder le rôle d'utilisateur root (sur les systèmes AIX, HP-UX, Linux et Solaris) ou le rôle d'administrateur (sur les systèmes Windows).
- Nortel Alteon Controller doit bénéficier de la connectivité IP à Nortel Alteon Web Switch et à tous les serveurs pour lesquels des pondérations sont calculées.
- Nortel Alteon Web Switch doit être configuré comme suit :
 1. Activez l'équilibrage de charge du serveur niveau 4 sur le commutateur.
 2. Configurez une interface IP.
 3. Activez SNMP.
 4. Activez le traitement du client d'équilibrage de charge du serveur sur le port qui reçoit les demandes des clients.
 5. Activez le traitement du serveur d'équilibrage de charge du serveur sur le port via lequel les serveurs réels se connectent.
 6. Configurez des serveurs réels pour les postes serveur Web.
 7. Configurez un groupe de serveurs réels constitué des serveurs réels qui exécutent le serveur d'applications.
 8. Configurez un serveur virtuel.
 9. Configurez un service sur un port virtuel et affectez le groupe de serveurs réels à son service.

Etape 1. Démarrage de la fonction serveur

Si nalservice ne s'exécute pas déjà, entrez **nalservice** en tant que superutilisateur pour le démarrer.

Remarque : Pour les systèmes Windows, cliquez sur **Démarrer > Paramètres** (pour Windows 2000) > **Panneau de configuration > Outils d'administration > Services**. Cliquez à l'aide du bouton droit de la souris sur IBM Nortel Alteon Controller, puis sélectionnez Démarrer.

Etape 2. Démarrage de l'interface de ligne de commande

Entrez **nalcontrol** pour démarrer l'interface de ligne de commande.

Etape 3. Définition d'un consultant de Nortel Alteon Web Switch

Pour ajouter un consultant de commutateur, entrez :

```
consultant add  
ID_consultant_commutateur address adresse_IP_commutateur
```

Etape 4. Ajout d'un service au consultant de commutateur

Pour ajouter un service, entrez :

```
service add ID_consultant_commutateur:ID_service vsid ID_serveur_virtuel  
vport numéro_port_virtuel
```

Un service est identifié par un identificateur de serveur virtuel (VSID) et un numéro de port virtuel (VPORT), tous deux associés à un serveur virtuel précédemment configuré sur le commutateur.

Etape 5. Configuration des mesures

Les mesures sont les informations permettant de déterminer les pondérations des serveurs. A chaque mesure est affectée un niveau d'importance indiquant son importance par rapport aux autres mesures. Vous pouvez configurer toute combinaison de mesures : mesures de données de connexion, mesures de conseiller d'application et mesures de serveur de mesures. Les proportions doivent toujours évaluer 100.

Lorsqu'un service est configuré, les mesures par défaut définies sont **activeconn** et **connrate**. Si vous voulez des mesures supplémentaires ou différentes des mesures par défaut, entrez :

```
service metrics ID_consultant_commutateur:ID_service nom_mesure 50  
nom_mesure2 50
```

Etape 6. Lancement du consultant

Pour démarrer le consultant, entrez :

```
consultant start  
ID_consultant_commutateur
```

Les collecteurs de mesure démarrent et le calcul des pondération commence.

Etape 7. Configuration de la haute disponibilité (facultatif)

Pour configurer la haute disponibilité, entrez :

```
highavailability add address adresse_IP partneraddress adresse_IP port 80  
role principal
```

Pour des informations détaillées sur l'emploi et la configuration de la haute disponibilité des composants Controller, voir Chapitre 23, «Fonctions avancées de Cisco CSS Controller et Nortel Alteon Controller», à la page 243.

Etape 8. Lancement du système Metric Server (facultatif)

Si les mesures système sont définies à l'étape 5, le serveur de mesures doit être démarré sur les machines de service. Pour plus d'informations sur le serveur de mesures, voir «Système Metric Server», à la page 253.

Etape 9. Régénération de la configuration de Nortel Alteon Controller

Modifier la configuration sur le Nortel Alteon Web Switch permet de régénérer la configuration du contrôleur. Entrez :

service refresh

Avant de régénérer la configuration, arrêtez le consultant. Une fois la configuration mise à jour, redémarrez le consultant.

Test de vérification de la configuration

Vérifiez que la configuration fonctionne :

1. Attribuez la valeur 4 au niveau de consignation du consultant.
2. Déconnectez un serveur de Nortel Alteon Web Switch pendant une minute ou arrêtez le serveur d'applications pendant une minute.
3. Reconnectez le serveur ou démarrez à nouveau le serveur d'applications.
4. Attribuez à nouveau le niveau désiré (1) au niveau de consignation du consultant.
5. Affichez le fichier consultant.log des répertoires ci-après et cherchez le **service de définition setServerWeights**. Ceci implique qu'une tentative d'envoi de pondérations au commutateur a été effectuée.
 - Pour les systèmes AIX, HP-UX, Linux et Solaris : ...ibm/edge/lb/servers/logs/cco/nom_consultant
 - Pour les systèmes Windows : ...ibm\edge\lb\servers\logs\cco\nom_consultant
6. Affichez les pondérations de serveur sur le commutateur et vérifiez qu'elles correspondent aux pondérations répertoriées sur le rapport du contrôleur.

Partie 7. Fonctions et fonctions avancées de Load Balancer

Cette section contient des informations relatives aux fonctions et fonctions avancées disponibles de configuration de Load Balancer. Elle se compose des chapitres suivants :

- Chapitre 21, «Gestionnaire, conseillers et système Metric Server (des composants Dispatcher, CBR et Site Selector)», à la page 179
- Chapitre 22, «Fonctions avancées de Dispatcher, CBR et Site Selector», à la page 201
- Chapitre 23, «Fonctions avancées de Cisco CSS Controller et Nortel Alteon Controller», à la page 243

Chapitre 21. Gestionnaire, conseillers et système Metric Server (des composants Dispatcher, CBR et Site Selector)

Le présent chapitre explique comment configurer les paramètres d'équilibrage de charge et comment installer les fonctions gestionnaire, conseiller et Metric Server de Load Balancer.

Remarque : Lors de la lecture de ce chapitre, si vous n'utilisez *pas* le composant Dispatcher, remplacez "dscontrol" par l'élément suivant :

- Pour CBR, utilisez **cbrcontrol**
- Pour Site Selector, utilisez **sscontrol** (Voir Chapitre 28, «Guide des commandes Site Selector», à la page 401)

IMPORTANT : Si vous utilisez l'installation Load Balancer pour IPv4 et IPv6, voir Chapitre 8, «Déploiement de Dispatcher sur Load Balancer pour IPv4 et IPv6», à la page 81 pour connaître les limitations et les différences de configuration avant d'afficher le contenu de cette section.

Tableau 12. Tâches de configuration avancées pour Load Balancer

Tâche	Description	Informations connexes
Modification des paramètres de l'équilibrage de charge	Les paramètres d'équilibrage de charge suivants peuvent être modifiés : <ul style="list-style-type: none">• Proportion de l'importance accordée aux données d'état. Le rapport par défaut est 50-50-0-0. Si vous utilisez la valeur par défaut, les informations fournies par les conseillers, par le système Metric Server et par VLM ne sont pas exploitées.• Pondérations• Pondérations fixées par l'administrateur• Intervalles gestionnaire• Seuil de sensibilité• Indice de lissage	«Optimisation de la fonction d'équilibrage de charge fournie par Load Balancer», à la page 180
Utilisation des scripts pour la génération d'une alerte ou d'une erreur du serveur d'enregistrement lorsque le gestionnaire indique si le serveur est actif ou non	Load Balancer fournit des exits utilisateur qui déclenchent des scripts que vous pouvez personnaliser lorsque le gestionnaire indique si le serveur est actif ou non	«Utilisation de scripts pour la génération d'une alerte ou d'une erreur du serveur d'enregistrement», à la page 184
Utilisation des conseillers	Décrit et répertorie les conseillers, qui signalent les états spécifiques des serveurs	«Conseillers», à la page 185
Utilisation de l'option de demande ou réponse (URL) de conseiller HTTP ou HTTPS	Définit une chaîne HTTP URL client unique propre à un service que vous voulez demander sur la machine	«Configuration du conseiller HTTP ou HTTPS à l'aide de l'option de demande ou de réponse (URL)», à la page 190
Utilisation d'un auto conseiller	Fournit au serveur principal l'état de chargement d'une configuration WAN Load Balancer à deux niveaux	«Utilisation d'un conseiller Self dans une configuration WAN à deux niveaux», à la page 191

Tableau 12. Tâches de configuration avancées pour Load Balancer (suite)

Tâche	Description	Informations connexes
Création de conseillers personnalisés	Explique comment écrire vos propres conseillers	«Création de conseillers personnalisés», à la page 192
Utilisation de l'agent Metric Server	Le système Metric Server fournit des informations de chargement à Load Balancer	«Metric Server», à la page 196
Utilisation du conseiller WLM (Workload Manager)	Le conseiller WLM fournit des informations de chargement à Load Balancer	«Conseiller Workload Manager», à la page 198

Optimisation de la fonction d'équilibrage de charge fournie par Load Balancer

La fonction gestionnaire de Load Balancer effectue l'équilibrage de charge en fonction des paramètres suivants :

- «Proportion de l'importance accordée aux données d'état»
- «Pondérations», à la page 181
- «Intervalles gestionnaire», à la page 183
- «Intervalles conseiller», à la page 187
- «Délai de rapport du conseiller», à la page 187
- «Seuil de sensibilité», à la page 183
- «Indice de lissage», à la page 184

Tous ces paramètres peuvent être modifiés en vue d'optimiser l'équilibrage de la charge du réseau.

Proportion de l'importance accordée aux données d'état

Le gestionnaire peut utiliser certains ou l'ensembles de facteurs externes suivants pour les décisions de pondération :

- *Connexions actives* : Nombre de connexions actives sur chaque serveur d'équilibrage de charge (indiqué par l'exécuteur). Cette proportion ne s'applique pas à Site Selector.

ou —

Unité centrale : Pourcentage de l'unité centrale utilisé sur chaque serveur d'équilibrage de charge (entré à partir de l'agent Metric Server). Pour Site Selector uniquement, cette proportion apparaît à la place de la colonne de la proportion des connexions actives.

- *Nouvelles connexions* : Nombre de nouvelles connexions sur chaque serveur d'équilibrage de charge (indiqué par l'exécuteur). Cette proportion ne s'applique pas à Site Selector.

ou —

Mémoire : Pourcentage de mémoire utilisé (entrée à partir de l'agent Metric Server) sur chaque serveur d'équilibrage de charge. Pour Site Selector uniquement, cette proportion apparaît à la place de la colonne de la proportion des nouvelles connexions.

- *Spécifique au port* : Entrée effectuée par les conseillers écoutant au niveau de ce port.
- *Mesure système* : Entrées provenant des outils de contrôle système, tels que Metric Server ou WLM.

Outre la charge courante de chaque serveur et d'autres données nécessaires à ses calculs, le gestionnaire obtient les deux premières valeurs (nouvelles connexions et connexions actives) de la part de l'exécuteur. Ces valeurs dépendent de données générées et stockées en interne par l'exécuteur.

Remarque : Pour Site Selector, le gestionnaire extrait les deux premières valeurs (unité centrale et mémoire) de Metric Server.

Vous pouvez modifier la proportion d'importance relative des quatre valeurs sur la base d'un cluster (ou nom de site). Les proportions correspondent à des pourcentages ; la somme des proportions relatives est égale à 100%. Le ratio par défaut est 50/50/0/0, ce qui revient à ignorer les informations système et celles transmises par les conseillers. Dans votre environnement, vous serez amené à essayer différentes combinaisons de proportions pour déterminer celle qui offre les meilleures performances.

Remarque : Lorsque vous ajoutez un conseiller (autre que WLM), si la **proportion du port** est égale à zéro, le gestionnaire ajoute 1 à cette valeur. Etant donné que la somme des proportions relatives doit être égale à 100, la valeur 1 est soustraite de la valeur la plus élevée.

Lorsque vous ajoutez le conseiller WLM, si la **proportion de la mesure système** est égale à zéro, le gestionnaire ajoute alors 1 à cette valeur. Etant donné que la somme des proportions relatives doit être égale à 1, la valeur 1 est soustraite de la valeur la plus élevée.

Le nombre de connexions actives dépend du nombre de clients ainsi que du délai nécessaire pour accéder aux services offerts par les machines serveurs d'équilibrage de charge. Si les connexions client sont rapides (comme dans le cas de courtes pages Web obtenues par HTTP GET), le nombre de connexions actives est faible. Si les connexions client sont lentes (comme dans le cas de requêtes de base de données), le nombre de connexions actives est plus élevé.

Il est recommandé de ne pas attribuer des valeurs trop basses aux nouvelles connexions et aux connexions actives. Si la valeur 20 (au moins) n'est pas attribuée aux deux première valeurs, vous désactivez l'équilibrage de charge et le lissage.

Pour définir la proportion des valeurs d'importance, utilisez la commande **dscontrol cluster set *cluster* proportions**. Pour plus d'informations, voir «dscontrol cluster — Configuration des clusters», à la page 353.

Pondérations

Les pondérations sont définies par le gestionnaire en fonction des décomptes internes de l'exécuteur, du retour d'informations des conseillers et du retour d'informations procuré par un programme de contrôle système, tel que Metric Server. Si vous voulez définir des pondérations manuellement lors de l'exécution du gestionnaire, indiquez l'option **fixedweight** lors de l'exécution de la commande **dscontrol server**. Pour obtenir une description de l'option **fixedweight**, voir «Pondérations fixées par le gestionnaire», à la page 182.

Les pondérations définies s'appliquent à tous les serveurs connectés sur un même port. Pour chaque port, les demandes sont réparties entre les serveurs selon la pondération relative de chacun. Par exemple, si un serveur a une pondération (paramètre **Weight**) de 10 et un autre de 5, le premier recevra deux fois plus de demandes que le second.

Pour définir la limite de pondération maximale d'un serveur, entrez la commande **dscontrol port set port weightbound weight**. Cette commande intervient sur l'écart existant entre les serveurs au niveau du nombre de demandes reçues par chacun. Si la limite de pondération maximale est de 1, tous les serveurs peuvent avoir une pondération égale à 1, 0 en veille, ou -1 si désactivé. A mesure que cette valeur augmente, l'écart entre les pondérations des serveurs augmente également. Avec une limite de pondération de 2, un serveur donné pourra recevoir deux fois plus de demandes qu'un autre. Avec une limite de pondération de 10, un serveur pourra recevoir dix fois plus de demandes qu'un autre. La limite de pondération maximale par défaut est de 20.

Si un conseiller détecte la défaillance d'un serveur, il en informe le gestionnaire qui attribue au serveur une pondération de zéro. Ainsi, l'exécuteur n'enverra pas de nouvelles connexions à ce serveur tant que cette pondération restera égale à zéro. Si ce serveur disposait d'une ou plusieurs connexions avant la modification de sa pondération, les connexions pourront toutefois s'achever normalement.

Si tous les serveurs sont arrêtés, le gestionnaire définit la pondération à une valeur correspondant à la moitié de la limite de pondération maximale.

Pondérations fixées par le gestionnaire

Sans le gestionnaire, les conseillers ne peuvent pas être lancés ni détecter les pannes de serveur. Si vous choisissez de lancer les conseillers mais ne voulez *pas* que le gestionnaire mette à jour la pondération que vous fixée pour un serveur particulier, utilisez l'option **fixedweight** de la commande **dscontrol server**. Par exemple :

```
dscontrol server set  
cluster:port:serveur fixedweight yes
```

Une fois la valeur **yes** attribuée à l'option **fixedweight**, utilisez la commande **dscontrol server set weight** pour attribuer la valeur souhaitée à la pondération. La valeur de pondération du serveur reste fixe tant que le gestionnaire est en activité à moins que vous n'émettiez une commande **dscontrol** en attribuant la valeur **no** à l'option **fixedweight**. Pour plus de détails, voir «**dscontrol server** — Configuration des serveurs», à la page 390.

Envoie d'une réinitialisation TCP à un serveur arrêté (composant Dispatcher uniquement)

Si la **réinitialisation TCP** est activée, Dispatcher envoie une réinitialisation TCP au client lorsque celui-ci est connecté à un serveur de pondération 0. La pondération d'un serveur est égale à zéro si elle est ainsi configurée ou si un conseiller l'a déclaré arrêté. Une réinitialisation TCP provoque la fermeture immédiate de la connexion. Cette fonction est utile pour les connexions longues durées où elle donne au client la possibilité de renégocier plus vite une connexion refusée. Activez la réinitialisation TCP à l'aide de la commande **dscontrol port add | set port reset yes**. La valeur par défaut est **no**.

Remarque : La réinitialisation TCP s'applique à toutes les méthodes de réacheminement de Dispatcher. Toutefois, pour pouvoir utiliser la fonction de réinitialisation TCP, vous devez définir le paramètre **clientgateway** de la commande **dscontrol executor** avec une adresse de routeur.

Associée à la réinitialisation TCP, la fonction **tentative du conseiller** est utile à configurer. Avec cette fonctionnalité, un conseiller peut renouveler une tentative de

connexion avant de déclarer un serveur arrêté. Ainsi, le conseiller ne déclare prématurément pas un serveur arrêté au risque de provoquer des incidents de réinitialisation de connexion. En clair, le fait que le conseiller échoue à la première tentative ne signifie pas nécessairement que la connexion existante est coupée. Pour plus d'informations, voir «Tentative du conseiller», à la page 188.

Intervalles gestionnaire

Pour optimiser les performances générales du réseau, la fréquence des interactions entre le gestionnaire et l'exécuteur est limitée. Pour modifier cet intervalle d'interaction, entrez les commandes **dscontrol manager interval** et **dscontrol manager refresh**.

L'intervalle gestionnaire indique la fréquence selon laquelle le gestionnaire réactualise les pondérations des serveurs utilisés par l'exécuteur pour acheminer les connexions. Si l'intervalle gestionnaire est trop court, le gestionnaire interrompra l'exécuteur constamment et les performances déclineraient. Dans le cas contraire, le routage des demandes assuré par l'exécuteur reposera sur des informations anciennes et incertaines.

Par exemple, pour définir un intervalle gestionnaire d'une seconde, entrez la commande suivante :

```
dscontrol manager interval 1
```

Le seuil de régénération du gestionnaire détermine la fréquence selon laquelle le gestionnaire demande des données d'état à l'exécuteur. Le seuil de régénération dépend de la durée de l'intervalle.

Par exemple, pour fixer à 3 intervalles le seuil de régénération du gestionnaire, entrez la commande suivante :

```
dscontrol manager refresh 3
```

Après cette commande, le gestionnaire devra patienter 3 intervalles avant de demander des données d'état à l'exécuteur.

Seuil de sensibilité

D'autres méthodes d'optimisation de l'équilibrage de charge des serveurs sont disponibles. Pour fonctionner en vitesse maximale, les pondérations des serveurs ne sont actualisées que si les pondérations ont évolué de manière significative. La mise à jour constante des pondérations pour un écart mineur de l'état des serveurs induirait un surcroît d'activité injustifié. Lorsque, pour tous les serveurs d'un port donné, l'écart en pourcentage de la pondération totale dépasse le seuil de sensibilité, le gestionnaire réactualise les pondérations des serveurs utilisés par l'exécuteur pour répartir les connexions. Supposons par exemple que la pondération totale passe de 100 à 105. L'écart est de 5%. Avec un seuil de sensibilité par défaut de 5, le gestionnaire ne met pas à jour les pondérations utilisées par l'exécuteur, car l'écart en pourcentage n'est pas **supérieur** au seuil. Si, en revanche la pondération totale passe de 100 à 106, le gestionnaire met à jour les pondérations. Pour attribuer au seuil de sensibilité du gestionnaire une valeur autre que la valeur par défaut (par exemple, 6), entrez la commande suivante :

```
dscontrol manager sensitivity 6
```

Dans la plupart des cas, vous n'aurez pas besoin de modifier cette valeur.

Indice de lissage

Le gestionnaire calcule dynamiquement les pondérations des serveurs. Il en découle qu'une fois mise à jour, une nouvelle pondération peut être très différente de l'ancienne. Dans la plupart des cas, cela ne porte pas à conséquence. Cependant, cela peut parfois induire de fortes variations dans la manière dont l'équilibrage de charge est effectué pour les demandes. Par exemple, l'un des serveurs peut finir par réceptionner la plupart des demandes du fait d'une pondération élevée. Le gestionnaire s'apercevra alors que le serveur en question traite un nombre élevé de connexions et répond lentement. Il transposera alors la pondération sur des serveurs moins encombrés et le même phénomène se reproduira, induisant une exploitation improductive des ressources.

Pour corriger ce dysfonctionnement, le gestionnaire utilise un indice de lissage. L'indice de lissage limite l'écart de pondération d'un serveur, filtrant et uniformisant effectivement la variation dans la répartition des demandes. Plus l'indice de lissage sera élevé, moins les pondérations des serveurs varieront. Plus l'indice de lissage sera faible, plus les pondérations des serveurs changeront. La valeur par défaut de l'indice de lissage est de 1,5. Avec un index de 1,5, les pondérations des serveurs seront plutôt fluctuantes. Pour un index de 4 ou 5, ces pondérations seront plus constantes. Par exemple, pour fixer l'indice de lissage à 4, entrez la commande suivante :

```
dscontrol manager smoothing 4
```

Dans la plupart des cas, vous n'aurez pas besoin de modifier cette valeur.

Utilisation de scripts pour la génération d'une alerte ou d'une erreur du serveur d'enregistrement

Load Balancer fournit des exits utilisateur qui déclenchent des scripts que vous pouvez personnaliser. Vous pouvez créer des scripts afin d'effectuer des actions automatisées. Il est, par exemple, possible de prévenir un administrateur lorsque le gestionnaire indique qu'un serveur est inactif ou simplement d'enregistrer l'erreur. Le répertoire d'installation, `...ibm/edge/lb/servers/samples`, contient des exemples de script que vous pouvez personnaliser. Pour pouvoir exécuter les fichiers, vous devez les déplacer dans le répertoire `...ibm/edge/lb/servers/bin` et supprimer l'extension de fichier `".sample"`. Les scripts exemples suivants sont fournis :

- **serverDown** — le gestionnaire indique qu'un serveur est inactif.
- **serverUp** — le gestionnaire indique qu'un serveur est à nouveau actif.
- **managerAlert** — il est indiqué que tous les serveurs d'un port particulier sont inactifs.
- **managerClear** — au moins un serveur est à nouveau disponible actif après qu'il a été indiqué que tous les serveurs étaient inactifs pour un port particulier.

Si tous les serveurs d'un cluster sont marqués comme étant arrêtés (par l'utilisateur ou par les conseillers), la fonction `managerAlert` (si elle est configurée) démarre et le composant Load Balancer tente de router le trafic vers les serveurs en utilisant une technique de permutation circulaire. Le script `serverDown` ne démarre pas lorsque le dernier serveur du cluster est détecté comme étant hors ligne.

De par sa conception, Load Balancer continue à router le trafic au cas où un serveur redeviendrait actif et répondrait à la demande. Si ce composant interrompait le trafic, le client ne recevrait pas de réponse.

Lorsque Load Balancer détecte que le premier serveur d'un cluster est redevenu actif, le script managerClear (s'il est configuré) démarre mais le script serverUp (s'il est configuré) ne s'exécute pas tant qu'un autre serveur n'est pas réactivé.

Considérations à prendre en compte lors de l'utilisation des scripts **serverUp** et **serverDown** :

- Si vous définissez un cycle de gestionnaire inférieur à 25 % de la durée du conseiller, les états des serveurs démarrés ou arrêtés générés peuvent être erronés. Par défaut, le gestionnaire est exécuté toutes les 2 secondes, mais le conseiller l'est toutes les 7 secondes. Par conséquent, le gestionnaire attend de nouvelles informations du conseiller tous les 4 cycles. Toutefois, si vous supprimez cette restriction (en définissant un cycle de gestionnaire supérieur à 25 % de la durée du conseiller), les performances seront considérablement réduites car un même serveur pourra alors être conseillé par plusieurs conseillers.
- Lorsqu'un serveur est arrêté, le script serverDown démarre. Toutefois, si vous exécutez une commande serverUp, le serveur est considéré comme actif jusqu'à ce que le gestionnaire obtienne de nouvelles informations du cycle du conseiller. Si le serveur est toujours arrêté, le script serverDown est à nouveau exécuté.

Conseillers

Les conseillers sont des agents de Load Balancer. Ils ont pour rôle d'évaluer l'état et la charge des serveurs. Ils effectuent cette tâche via un échange proactif de type client/serveur. Les conseillers peuvent être considérés comme des clients des serveurs d'application.

Le produit fournit plusieurs conseillers pour les protocoles les plus couramment utilisés. Cependant, l'utilisation de tous les conseillers fournis avec chaque composant de Load Balancer ne présente aucun intérêt. (Par exemple, n'utilisez pas le conseiller Telnet avec le composant CBR.) Load Balancer prend également en charge le concept de «conseiller personnalisé» permettant aux utilisateurs d'écrire leurs propres conseillers.

Restrictions d'utilisation des applications de serveur de liaison : Pour utiliser des conseillers sur des serveurs de liaison, démarrez deux instances du serveur : une instance pour une liaison sur cluster:port et l'autre pour une liaison sur serveur:port. Pour savoir s'il s'agit d'un serveur de liaison, lancez la commande `netstat -an` et recherchez `serveur:port`. S'il ne s'agit pas d'un serveur de liaison, le résultat de la commande est `0.0.0.0:80`. S'il s'agit d'un serveur de liaison, une adresse du type `192.168.15.103:80` apparaît.

Sur les systèmes HP-UX et Solaris, restriction d'utilisation de serveurs de liaison : Si vous utilisez la commande `arp publish` au lieu de la commande `ifconfig alias`, Load Balancer *accepte* l'utilisation de conseillers lors de l'équilibrage de la charge des serveurs de liaison (autres composants Load Balancer tels que CBR Locator ou Site Selector compris) lorsqu'ils sont reliés à l'adresse IP du cluster. Toutefois, si vous utilisez des conseillers sur des serveurs de liaison, ne co-implantez pas Load Balancer sur la même machine qu'un serveur de liaison.

Remarque : Si Load Balancer s'exécute sur un ordinateur doté de plusieurs cartes réseau et que vous voulez que le trafic du conseiller passe par une carte particulière, vous pouvez imposer une adresse spécifique comme adresse IP source des paquets. Pour ce faire, ajoutez à la ligne

```
java...SRV_XXXConfigServer... du fichier script de démarrage  
approprié de Load Balancer (dsserver, cbrserver ou ssserver) les  
éléments suivants :  
-DLB_ADV_SRC_ADDR=adresse_IP
```

Fonctionnement des conseillers

Les conseillers ouvrent régulièrement une connexion TCP avec chaque serveur et envoient un message de demande au serveur. Le contenu du message dépend du protocole exécuté sur le serveur. Par exemple, le conseiller HTTP envoie une demande HTTP «HEAD» au serveur.

Les conseillers attendent ensuite une réponse du serveur. Une fois la réponse obtenue, le conseiller évalue l'état du serveur. Pour calculer la valeur de la «charge», la plupart des conseillers mesurent le délai de réponse du serveur, puis ils utilisent cette valeur (en millisecondes) comme valeur de charge.

Le conseiller reporte cette valeur au gestionnaire. Elle apparaît dans le rapport du gestionnaire, dans la colonne «Port». Le gestionnaire calcule ensuite un ensemble de valeurs de pondération à partir de toutes ses sources, selon les proportions, et définit ces valeurs de pondération dans la fonction exécuteur. L'exécuteur utilise ces pondérations pour équilibrer la charge des nouvelles connexions client entrantes.

Si le conseiller détermine que le serveur est actif et que son état est correct, il renvoie au gestionnaire une valeur de charge positive non nulle. Si le conseiller détermine que le serveur n'est pas actif, il renvoie une valeur de charge spéciale négative (-1). Le gestionnaire et l'exécuteur n'enverront plus aucune connexion en direction de ce serveur tant qu'il ne sera pas de nouveau actif.

Remarque : Avant d'envoyer le message de demande initial, le conseiller lance une commande ping au serveur. Cela a pour but d'obtenir rapidement l'état de la machine pour déterminer si elle est en ligne. Une fois que le serveur a répondu au ping, aucun autre ping n'est envoyé. Pour désactiver les pings, ajoutez -DLB_ADV_NB_PING dans le fichier script de lancement de Load Balancer.

Démarrage et arrêt d'un conseiller

Vous pouvez lancer un conseiller pour un port particulier de tous les clusters (conseiller de groupe). Vous pouvez également choisir d'exécuter différents conseillers sur le même port mais sur des clusters différents (conseiller spécifique cluster/site). Par exemple, si Load Balancer est défini avec trois clusters (*cluster A*, *cluster B*, *cluster C*), pour chaque cluster le port 80 a une fonction différente.

- Conseiller spécifique cluster/site : Pour démarrer un conseiller sur le port 80 pour *cluster A*, indiquez à la fois le cluster et le port :
dscontrol advisor start http *clusterA*:80

Cette commande lance le conseiller HTTP sur le port 80 pour le *cluster A*. Le conseiller HTTP fonctionne sur tous les serveurs connectés au port 80 pour le cluster A.

- Conseiller de groupe : Pour démarrer un conseiller personnalisé sur le port 80 pour tous les autres clusters, indiquez simplement le port :
dscontrol advisor start *ADV_personnalisé* 80

Cette commande lance le conseiller *ADV_personnalisé* sur le port 80 pour la *cluster B* et la *cluster C*. Le conseiller personnalisé fonctionne sur tous les serveurs connectés au port 80 pour la *cluster B* et la *cluster C*. (Pour obtenir plus d'informations sur les conseillers personnalisés, voir «Création de conseillers personnalisés», à la page 192.)

Remarque : Le conseiller de groupes fonctionne sur tous les clusters/sites ne disposant pas d'un conseiller spécifique.

Lorsque vous utilisez la configuration exemple précédente, vous pouvez choisir d'arrêter le conseiller personnalisé *ADV_custom* pour le port 80 sur un cluster uniquement ou pour les deux clusters (*cluster B* et *cluster C*).

- Pour arrêter le conseiller personnalisé pour le port 80 uniquement pour la *cluster B*, indiquez le cluster et le port :
`dscontrol advisor stop ADV_personnalisé clusterB:80`
- Pour arrêter le conseiller personnalisé pour le port 80 sur la *cluster B* et la *cluster C*, indiquez uniquement le port :
`dscontrol advisor stop ADV_personnalisé 80`

Intervalles conseiller

Remarque : Les valeurs par défaut du conseiller doivent être correctes pour la plupart des scénarios possibles. Soyez prudent lorsque vous entrez des valeurs autres que celles fournies par défaut.

L'intervalle conseiller détermine la fréquence selon laquelle un conseiller demande des données d'état aux serveurs associés au port dont il a la charge, puis transmet ces données au gestionnaire. Si l'intervalle conseiller est trop court, le conseiller interrompra les serveurs constamment et les performances déclineront. Dans le cas contraire, les décisions d'allocation de pondérations prises par le gestionnaire reposeront sur des informations anciennes et incertaines.

Par exemple, pour fixer à 3 secondes l'intervalle du conseiller HTTP sur le port 80, entrez la commande suivante :

```
dscontrol advisor interval http 80 3
```

Notez qu'il n'est pas logique de spécifier un intervalle conseiller inférieur à l'intervalle gestionnaire. L'intervalle conseiller par défaut est sept secondes.

Délai de rapport du conseiller

Pour s'assurer que le gestionnaire n'utilise pas d'informations périmées pour ses décisions d'équilibrage de charge, le gestionnaire n'utilisera pas les informations d'un conseiller dont l'horodatage sera antérieur à celui défini dans le délai de rapport du conseiller. Le délai de rapport du conseiller doit être supérieur à l'intervalle de sondage du conseiller. Si le délai est inférieur, le gestionnaire ignore les états qu'il est censé exploiter. Par défaut, les rapports des conseillers n'ont pas de délai d'expiration — la valeur par défaut est Unlimited (illimité).

Par exemple, pour fixer à 30 secondes l'intervalle du conseiller HTTP sur le port 80, entrez la commande suivante :

```
dscontrol advisor timeout http 80 30
```

Pour obtenir plus d'informations sur la définition du délai de rapport du conseiller, voir «dscontrol advisor — Contrôle du conseiller», à la page 347.

Délai de connexion du conseiller et délai de réception pour les serveurs

Pour Load Balancer, vous pouvez définir les valeurs de délai du conseiller lorsqu'une erreur au niveau d'un port particulier du serveur (service) est détectée. Les valeurs de délai d'erreur serveur (connecttimeout et receivetimeout) déterminent la durée attendue par un conseiller avant de signaler qu'une connexion ou une réception n'a pas abouti.

Pour obtenir une détection d'erreur serveur très rapide, attribuez la valeur la plus basse (une seconde) aux délais de connexion et de réception du conseiller et attribuez la valeur la plus basse (une seconde) à l'intervalle du gestionnaire et du conseiller.

Remarque : Si le trafic de votre environnement atteint un volume modéré voire élevé et que le temps de réponse du serveur augmente, vérifiez que vous n'avez pas attribué des valeurs trop faibles à connecttimeout et à receivetimeout. Sinon, le conseiller peut indiquer de manière prématuré une erreur réseau lorsqu'un serveur est occupé.

Par exemple, pour attribuer la valeur 9 secondes à connecttimeout et à receivetimeout pour le conseiller HTTP sur le port 80, entrez la commande suivante :

```
dscontrol advisor connecttimeout http 80 9
dscontrol advisor receivetimeout http 80 9
```

La valeur par défaut de connexion et de réception est trois fois supérieure à la valeur indiquée pour l'intervalle du conseiller.

Tentative du conseiller

Les conseillers peuvent essayer de nouveau d'établir une connexion avant de marquer un serveur comme arrêté. Le conseiller ne déclare un serveur comme étant arrêté qu'après avoir effectué le nombre de tentatives de connexion fixé plus une. Il est préférable que le nombre de **tentatives** ne dépasse pas 3. La commande ci-après fixe 2 tentatives pour le conseiller LDAP du port 389 :

```
dscontrol advisor retry ldap 389 2
```

Liste des conseillers

- Le conseiller **HTTP** ouvre une connexion, envoie une demande HEAD par défaut attend la connexion et renvoie le temps écoulé comme chargement. Pour obtenir plus d'informations sur le mode de modification du type de demande envoyée par le conseiller HTTP, voir «Configuration du conseiller HTTP ou HTTPS à l'aide de l'option de demande ou de réponse (URL)», à la page 190.
- Le conseiller **HTTPS** est un conseiller "lourd" des connexions SSL. Il établit une connexion SSL complète avec le serveur. Le conseiller HTTPS ouvre une connexion SSL, envoie une demande HTTPS, attend une réponse, ferme la connexion et renvoie le temps écoulé en tant que chargement. (Voir aussi le conseiller SSL, qui est un conseiller "léger" des connexions SSL.)

Remarque : Le conseiller HTTPS n'a aucun rapport avec les clés ou les certificats du serveur, mais ils ne doivent pas être expirés.

- Le conseiller **SIP** ouvre une connexion, envoie une demande OPTIONS, attend une réponse, ferme la connexion et renvoie le temps écoulé en tant que

chargement. Le conseiller SIP qui est pris en charge s'exécute sur TCP uniquement et nécessite l'installation d'une application sur un serveur répondant à une requête OPTIONS.

- Le conseiller **FTP** ouvre une connexion, envoie une demande SYST, attend une réponse, ferme la connexion et renvoie le temps écoulé en tant que chargement.
- Le conseiller **LDAP** ouvre une connexion, envoie une demande BIND anonyme, attend une réponse, ferme la connexion et renvoie le temps écoulé en tant que chargement.
- Le conseiller **Telnet** ouvre une connexion, attend le premier message du serveur, ferme la connexion et renvoie le temps écoulé en tant que chargement.
- Le conseiller **NNTP** ouvre une connexion, attend le premier message du serveur, envoie une commande quit, ferme la connexion et renvoie le temps écoulé comme chargement.
- Le conseiller **IMAP** ouvre une connexion, attend la première réponse du serveur, envoie une commande quit, ferme la connexion et renvoie le temps écoulé comme chargement.
- Le conseiller **POP3** ouvre une connexion, attend la première réponse du serveur, envoie une commande quit, ferme la connexion et renvoie le temps écoulé comme chargement.
- Le conseiller **SMTP** ouvre une connexion, attend la première réponse du serveur, envoie une commande quit, ferme la connexion et renvoie le temps écoulé comme chargement.
- Le conseiller **SSL** est un conseiller "léger" des connexions SSL. Il n'établit pas de connexion SSL complète avec le serveur. Le conseiller SSL ouvre une connexion, envoie une demande CLIENT_HELLO, attend une réponse, ferme la connexion et renvoie le temps écoulé comme chargement. (Voir aussi le conseiller HTTPS, qui est un conseiller "lourd" des connexions SSL.)

Remarque : Le conseiller SSL n'a aucun rapport avec la gestion des clés ou des certificats.

- Le conseiller **ssl2http** est lancé et fonctionne sur les serveurs répertoriés sous le port 443 mais le conseiller ouvre une connexion au "portdemappage" pour les demandes HTTP. Utilisez uniquement le conseiller ssl2http pour CBR si le protocole client vers proxy est de type SSL et si le protocole proxy vers serveur est de type HTTP. Pour plus d'informations, voir «Équilibrage de charge client-proxy dans SSL et proxy-serveur dans HTTP», à la page 108.
- Le conseiller Caching Proxy (cachingproxy) ouvre une connexion, envoie une demande HTTP GET spécifique de Caching Proxy et interprète la réponse en tant que chargement Caching Proxy.

Remarque : Lors de l'utilisation d'un conseiller Caching Proxy, il est nécessaire d'exécuter Caching Proxy sur tous les serveurs faisant l'objet d'un équilibrage de charge. Caching Proxy n'a pas à être installé sur la machine qui héberge Load Balancer à moins d'être co-implanté sur la machine dont il assure l'équilibrage de charge.

- Le conseiller **DNS** ouvre une connexion, envoie une demande de pointeur pour DNS, attend une réponse, ferme la connexion et renvoie le temps écoulé comme chargement.
- Le conseiller **Connect** n'échange aucune donnée spécifique du protocole avec le serveur. Il mesure simplement le temps que durent l'ouverture et la fermeture d'une connexion TCP avec le serveur. Ce conseiller est utile pour les applications

de serveur qui utilisent TCP, mais ces applications doivent disposer d'un protocole de haut niveau pour lequel aucun conseiller IBM ou personnalisé n'est disponible.

- Le conseiller **ping** n'ouvre pas de connexion TCP avec les serveurs, mais indique si le serveur répond ou non à une commande ping. Bien que le conseiller ping soit utilisable avec tout type de port, il est également conçu pour les configurations utilisant le port générique, capable d'acheminer un trafic multiprotocole. Il est également utile pour les configurations utilisant des protocoles non TCP avec leurs serveurs, tels que UDP.
- Le conseiller **reach** émet des commandes ping vers les machines cibles. Il est également conçu pour permettre aux composants haute disponibilité de Dispatcher de déterminer l'accessibilité des cibles à atteindre. Ses résultats sont transmis au composant de haute disponibilité et ne figurent *pas* dans le rapport du gestionnaire. Contrairement aux autres conseillers, le conseiller reach démarre automatiquement par la fonction gestionnaire du composant Dispatcher.
- Le conseiller **DB2** fonctionne en association avec les serveurs DB2. Dispatcher comporte une fonction intégrée permettant de vérifier l'état des serveurs DB2 sans que les clients aient besoin d'écrire leurs propres conseillers personnalisés. Le conseiller DB2 communique avec le port de connexion DB2 uniquement et non avec le port de connexion Java.
- Le conseiller **self** rassemble des informations sur le statut du chargement des serveurs dorsaux. Vous pouvez utiliser le conseiller lors de l'utilisation de Dispatcher dans une configuration à deux niveaux, dans laquelle Dispatcher fournit les informations provenant du conseiller self au composant Load Balancer de niveau supérieur. Le conseiller self mesure de manière spécifique les connexions par seconde sur les serveurs dorsaux du système Dispatcher se trouvant au niveau de l'exécutant. Pour plus d'informations, voir «Utilisation d'un conseiller Self dans une configuration WAN à deux niveaux», à la page 191.
- Le conseiller **WLM** (Workload Manager) est conçu pour fonctionner avec les serveurs sur les gros systèmes OS/390 exécutant le composant MVS Workload Manager (WLM). Pour plus d'informations, voir «Conseiller Workload Manager», à la page 198.
- Dispatcher permet à un utilisateur d'écrire un conseiller *personnalisé*. Cette opération permet la prise en charge de protocoles propriétaires (en plus de TCP) pour lesquels IBM n'a pas développé de conseiller spécifique. Pour plus d'informations, voir «Création de conseillers personnalisés», à la page 192.
- Le conseiller **WAS** (WebSphere Application Server) fonctionne en association avec les serveurs WebSphere Application Server. Des fichiers modèle pour ce conseiller sont fournis dans le répertoire d'installation. Pour plus d'informations, voir «Conseiller WAS», à la page 193.

Configuration du conseiller HTTP ou HTTPS à l'aide de l'option de demande ou de réponse (URL)

L'option d'URL du conseiller HTTP ou HTTPS est disponible pour les composants Dispatcher et CBR.

Après avoir lancé un conseiller HTTP ou HTTPS, vous pouvez définir une chaîne d'URL HTTP client unique, propre au service auquel vous souhaitez accéder sur le serveur. Le conseiller peut ainsi contrôler l'état des services d'un serveur. Vous pouvez effectuer cette opération en définissant des serveurs logiques avec des noms de serveurs uniques ayant la même adresse IP physique. Pour plus d'informations, voir «Partitionnement du serveur : serveurs logiques configurés pour un serveur physique (adresse IP)», à la page 58.

Pour chaque serveur logique défini sous le port HTTP, vous pouvez indiquer une chaîne HTTP URL client unique, spécifique du service pour lequel vous voulez interroger le serveur. Le conseiller HTTP ou HTTPS utilise la chaîne **advisorrequest** pour vérifier l'état des serveurs. La valeur par défaut est `HEAD / HTTP/1.0`. La chaîne **advisorresponse** est la réponse que le conseiller recherche dans la réponse HTTP. Le conseiller utilise la chaîne **advisorresponse** pour effectuer une comparaison par rapport à la réponse réelle reçue du serveur. La valeur par défaut est null.

Important : Si l'URL HTTP contient un espace :

- Lorsque vous lancez la commande à partir de l'invite du shell **dscontrol>>**, vous devez mettre la chaîne contenant un espace entre guillemets. Par exemple :

```
server
set cluster:port:serveur advisorrequest "head / http/1.0"
server set cluster:port:serveur
advisorresponse "HTTP 200 OK"
```

- Lorsque vous lancez la commande **dscontrol** à partir de l'invite du système d'exploitation, vous devez placer les caractères `"\"` et `\""` respectivement avant et après le texte. Par exemple :

```
dscontrol server set cluster:port:serveur
advisorrequest "\"head / http/1.0\""
```

```
dscontrol server set cluster:port:serveur advisorresponse "\"HTTP 200 OK\""
```

Lorsque vous créez la demande que le conseiller HTTP ou HTTPS envoie aux serveurs dorsaux pour vérifier s'ils fonctionnent, vous tapez le début de la demande HTTP et Load Balancer la complète en spécifiant les éléments suivants :

```
\r\nAccept:
*/*\r\nUser-Agent:IBM_Network_Dispatcher_HTTP_Advisor\r\n\r\n
```

Si vous souhaitez ajouter d'autres zones d'en-tête HTTP avant que Load Balancer ajoute cette chaîne en fin de la demande, insérez votre propre chaîne `\r\n` dans la demande. Voici un exemple de ce que vous devez taper pour ajouter la zone d'en-tête d'hôte HTTP à votre demande :

```
GET /pub/WWW/TheProject.html HTTP/1.0 \r\nHôte: www.w3.org
```

Remarque : Après le démarrage d'un conseiller HTTP ou HTTPS pour un numéro de port HTTP spécifié, la valeur de demande et réponse du conseiller est activée pour les serveurs sous ce port HTTP.

Pour plus d'informations, voir «dscontrol server — Configuration des serveurs», à la page 390.

Utilisation d'un conseiller Self dans une configuration WAN à deux niveaux

Le conseiller self est disponible dans le composant Dispatcher.

Lorsque Load Balancer se trouve dans une configuration WAN (wide area network) à deux niveaux, un conseiller *self* est fourni qui rassemble des informations de statut de chargement sur les serveurs dorsaux.

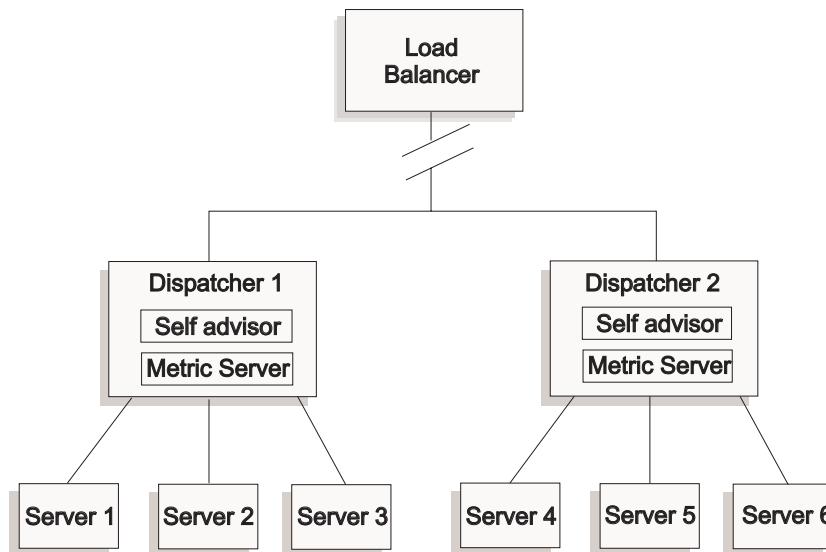


Figure 34. Exemple de configuration WAN à deux niveaux utilisant le conseiller self

Dans cet exemple, le conseiller self ainsi que le système Metric Server se trouvent sur deux machines Dispatcher dont l'équilibrage de charge est assuré par le système Load Balancer de niveau supérieur. Le conseiller self mesure de manière spécifique les connexions par seconde sur les serveurs dorsaux du système Dispatcher se trouvant au niveau de l'exécutant.

Le conseiller self inscrit les résultats dans le fichier dsloadstat. Load Balancer fournit également une mesure externe appelée dsload. L'agent du système Metric Server de chaque machine Dispatcher exécute son fichier de configuration qui appelle le script dsload de mesure externe. Le script dsload extrait une chaîne du fichier dsloadstat et le renvoie à l'agent du système Metric Server. Ensuite, chaque agent du système Metric Server (de chaque élément Dispatcher) renvoie la valeur de statut de chargement à l'élément Load Balancer se trouvant au niveau supérieur. Cette valeur sera utilisée pour déterminer le système Dispatcher qui transmettra les demandes client.

L'exécutable dsload se trouve dans le répertoire `...ibm/edge/lb/ms/script` pour Load Balancer.

Pour plus d'informations sur l'utilisation de Dispatcher dans des configurations WAN, voir «Configuration du support de réseau étendu pour Dispatcher», à la page 228. Pour plus d'informations sur le système Metric Server, voir «Metric Server», à la page 196.

Création de conseillers personnalisés

Le conseiller personnalisé est un petit programme Java, que vous fournissez sous forme de fichier classe, appelé par le code de base. Le code de base fournit tous les services administratifs, tels que le démarrage et l'arrêt d'une instance du conseiller personnalisé, la génération d'états et de rapports et l'enregistrement des informations de l'historique dans un fichier journal. Il renvoie également les résultats au composant gestionnaire. Régulièrement, le code de base lance un cycle de conseiller au cours duquel il évalue individuellement tous les serveurs de sa configuration. Il commence par ouvrir une connexion avec la machine serveur. Si la connexion s'ouvre, le code de base appelle la méthode « getLoad » (fonction) dans

le conseiller personnalisé. Ce dernier effectue la procédure nécessaire à l'évaluation du serveur. Généralement, il envoie au serveur un message défini par l'utilisateur et attend la réponse. L'accès à la connexion ouverte est fourni au conseiller personnalisé. Le code de base ferme ensuite la connexion au serveur et envoie au gestionnaire les informations relatives à la charge.

Le code de base et le conseiller personnalisé peuvent opérer en mode normal ou en mode replace. Le choix du mode de fonctionnement est indiqué dans le fichier du conseiller personnalisé en tant que paramètre dans la méthode du constructeur.

En mode normal, le conseiller personnalisé échange des données avec le serveur et le code du conseiller de base évalue la durée de l'échange et calcule la valeur de la charge. Le code de base renvoie cette valeur au gestionnaire. Le conseiller personnalisé doit simplement retourner un zéro (succès) ou une valeur négative (échec). Lorsque dans le fichier du constructeur, la valeur false est attribuée à l'indicateur replace, le mode normal est défini.

En mode replace, le code de base n'effectue aucune mesure de temps. Le code du conseiller personnalisé effectue toutes les opérations nécessaires, puis renvoie une valeur de charge. Le code de base accepte la valeur et la retourne au gestionnaire. Pour obtenir de meilleurs résultats, situez votre valeur de charge entre 10 et 1000, 10 représentant un serveur rapide et 1000 un serveur plus lent. Lorsque dans le fichier du constructeur, la valeur true est attribuée à l'indicateur replace, le mode replace est défini.

Avec cette fonctionnalité, vous pouvez développer vos propres conseillers qui fourniront les informations sur les serveurs dont vous avez besoin. Un exemple de conseiller personnalisé, **ADV_exemple.java**, est fourni avec le produit Load Balancer. Une fois Load Balancer installé, vous pouvez trouver le code exemple dans le répertoire d'installation
...<rép_install>/servers/samples/CustomAdvisors.

Le répertoire d'installation par défaut est :

- Pour les systèmes AIX, HP-UX, Linux, Solaris : /opt/ibm/edge/lb
- Pour les systèmes Windows : C:\Program Files\IBM\edge\lb

Remarque : Si vous ajoutez un conseiller personnalisé à Dispatcher, ou tout autre composant compatible avec Load Balancer, vous devez arrêter, puis redémarrer **dsserver** (ou le service pour les systèmes Windows) pour que le processus Java puisse lire les nouveaux fichiers de classes du conseiller personnalisé. Les fichiers de classes du conseiller personnalisé ne sont chargés qu'au démarrage. Il n'est pas nécessaire d'arrêter l'exécuteur. Ce dernier continue à s'exécuter après arrêt de dsserver, ou du service.

Si le conseiller personnalisé fait référence à d'autres classes Java, le chemin d'accès aux classes du fichier script de lancement de Load Balancer (dsserver, cbrserver, ssserver) doit être mis à jour pour inclure l'emplacement.

Conseiller WAS

Le répertoire d'installation de Load Balancer contient des fichiers exemple de conseiller personnalisé spécifiques au conseiller WebSphere Application Server (WAS).

- Le fichier à compiler et à exécuter sur la machine Load Balancer se nomme ADV_was.java.
- LBAAdvisor.java.servlet (à renommer LBAAdvisor.java) est le fichier à compiler et à exécuter sur la machine WebSphere Application Server.

Les fichiers exemple de conseiller WebSphere Application Server se trouvent dans le même répertoire que le fichier ADV_exemple.java.

Convention d'attribution de nom

Le nom de fichier de votre conseiller personnalisé doit avoir le format «ADV_monconseiller.java.» Il doit être précédé du préfixe «ADV_» en majuscules. Tous les caractères suivants doivent être en minuscules.

Conformément aux conventions Java, le nom de la classe définie dans le fichier doit correspondre au nom du fichier. Si vous copiez le code exemple, veuillez à remplacer toutes les occurrences de «ADV_exemple» dans le fichier par le nom de votre nouvelle classe.

Compilation

Les conseillers personnalisés sont écrits en langage Java. Utilisez le compilateur Java qui est installé avec Load Balancer. Les fichiers suivants sont référencés pendant la compilation :

- le fichier du conseiller personnalisé,
- le fichier de classes de base, ibmlb.jar, qui se trouve dans le répertoire d'installation **...ibm/edge/lb/servers/lib**.

Le chemin d'accès aux classes doit désigner à la fois le fichier du conseiller personnalisé et le fichier de classes de base lors de la compilation.

Pour les systèmes Windows, exemple de commande de compilation :

```
rep_install/java/bin/javac -classpath
    rep_install\lb\servers\lib\ibmlb.jar ADV_fred.java
```

où :

- Votre fichier conseiller s'appelle ADV_fred.java.
- Votre fichier conseiller se trouve dans le répertoire courant.

Le résultat de la compilation est un fichier .class, par exemple :

ADV_fred.class

Avant de lancer le conseiller, copiez le fichier .class dans le répertoire d'installation **...ibm/edge/lb/servers/lib/CustomAdvisors**.

Remarque : Si vous le souhaitez, vous pouvez compiler les conseillers personnalisés sur un système d'exploitation et l'exécuter sur un autre. Par exemple, vous pouvez compiler le conseiller sur des systèmes Windows, copier le fichier .class (en binaire) sur une machine AIX à partir de laquelle vous exécutez le conseiller personnalisé.

Pour les systèmes AIX, HP-UX, Linux et Solaris, la syntaxe est similaire.

Exécution

Pour exécuter le conseiller personnalisé, vous devez tout d'abord copier le fichier .class dans le répertoire d'installation approprié :


```
...ibm/edge/lb/servers/lib/CustomAdvisors/ADV_fred.class
```

Configurez le composant, démarrez la fonction gestionnaire, puis exécutez la commande permettant de lancer le conseiller personnalisé.

```
dscontrol advisor start fred 123
```

où :

- fred est le nom de votre conseiller, comme dans ADV_fred.java,
- 123 est le port sur lequel votre conseiller opérera.

Si le conseiller personnalisé fait référence à d'autres classes Java, le chemin d'accès aux classes du fichier script de lancement de Load Balancer (dsserver, cbrserver, ssserver) doit être mis à jour pour inclure l'emplacement.

Sous-programmes requis

Comme tous les conseillers, un conseiller personnalisé étend la fonction de la base du conseiller, intitulée ADV_Base. En fait, c'est la base du conseiller qui effectue la plupart des fonctions du conseiller, telles que la communication des charges au gestionnaire afin que ces dernières soient utilisées dans l'algorithme de pondération du gestionnaire. La base du conseiller effectue également les opérations de connexion et de fermeture de la connexion et fournit des méthodes d'envoi et de réception qui seront utilisées par le conseiller. Le conseiller n'est lui-même utilisé que pour l'envoi de données vers le port du serveur conseillé et pour la réception de données sur ce dernier. Les méthodes TCP de la base du conseiller sont programmées pour calculer la charge. Un indicateur du constructeur de ADV_base remplace, si vous le souhaitez, la charge existante par la nouvelle charge renvoyée par le conseiller.

Remarque : En fonction d'une valeur définie dans le constructeur, la base du conseiller fournit la charge à l'algorithme de pondération à un intervalle donné. Si le véritable conseiller n'a pas terminé ses opérations afin de renvoyer une charge valide, la base du conseiller utilise la charge précédente.

Ci-dessous, sont énumérées les méthodes de classe de base.

- Sous-programme **constructeur**. Le constructeur appelle le constructeur de la classe de base (reportez-vous au fichier type de conseiller).
- Méthode **ADV_AdvisorInitialize**. Cette méthode fournit un point d'ancrage au cas où des procédures supplémentaires doivent être suivies une fois l'initialisation de la classe de base terminée.
- Sous-programme **getload**. La classe de base du conseiller se charge de l'ouverture de la connexion ; getload ne doit qu'émettre les demandes d'envoi et de réception appropriées pour terminer le cycle de conseil.

Ordre de recherche

Load Balancer consulte tout d'abord la liste des conseillers en langage naturel. S'il ne trouve pas un conseiller donné, Load Balancer consulte la liste des conseillers personnalisés du clients.

Affectation du nom et du chemin

- La classe de conseiller personnalisé doit se trouver dans le sous-répertoire de **...ibm/edge/lb/servers/lib/CustomAdvisors/** dans le répertoire de base de Load Balancer. Le répertoire par défaut dépend du système d'exploitation :

- Systèmes AIX, HP-UX, Linux et Solaris
/opt/ibm/edge/lb/servers/lib/CustomAdvisors/
- Systèmes Windows
C:\Program Files\IBM\edge\lb\servers\lib\CustomAdvisors
- Seuls les caractères alphabétiques minuscules sont autorisés. Cela permet d'éliminer la distinction entre majuscules et minuscules lorsqu'un opérateur entre des commandes sur la ligne de commande. Le nom de fichier du conseiller doit être précédé de **ADV_**.

Conseiller type

Un programme permettant de créer un conseiller type est présenté à la section «Conseiller type», à la page 487. Après installation, ce conseiller exemple se trouve dans le répertoire **...ibm/edge/lb/servers/samples/CustomAdvisors**.

Metric Server

Cette fonction est disponible pour tous les composants Load Balancer.

Metric Server fournit à Load Balancer les informations de téléchargement sous la forme de données numériques système, relatives à l'état du serveur. Le gestionnaire Load Balancer adresse des demandes aux agents du système Metric Server situés sur chacun des serveurs, leur attribuant des pondérations destinées au processus d'équilibrage de charge à l'aide des données rassemblées par les agents. Les résultats sont regroupés dans le rapport du gestionnaire.

Remarque : Des erreurs d'arrondi peuvent se produire lorsque plusieurs données numériques sont rassemblées et normalisées pour chaque serveur en une seule valeur de charge.

Pour plus d'informations sur le fonctionnement de Metric Server (démarrage et arrêt) et sur l'utilisation des journaux Metric Server, voir «Utilisation du composant Metric Server», à la page 279.

Pour obtenir un exemple de configuration, voir figure 5, à la page 15.

Restrictions relatives à WLM

Comme le conseiller WLM, le rapport du Metric Server concerne l'ensemble des systèmes de serveurs et non chaque démon de serveur associé à un protocole. WLM et Metric Server placent leurs résultats dans la colonne relative au système du rapport du gestionnaire. Par conséquent, il n'est pas possible d'exécuter simultanément le conseiller WLM et Metric Server.

Conditions préalables

L'agent Metric Server doit être installé et en cours d'exécution sur tous les serveurs dont la charge est équilibrée.

Conditions d'utilisation de Metric Server

La procédure ci-après permet de configurer Metric Server pour Network Dispatcher. Vous pouvez configurer Metric Server pour les autres composants de Load Balancer à l'aide d'une procédure similaire.

- Gestionnaire Load Balancer (côté Load Balancer)
 1. Démarrez **dsserver**.

2. Emettez la commande **dscontrol manager start** *manager.log port*
port correspond au port RMI sur lequel fonctionnent tous les agents du système Metric Server. La valeur par défaut du port RMI est 10004, cette valeur est définie dans le fichier *metricserver.cmd*.
3. Emettez la commande **dscontrol metric add** *cluster:systemMetric*
systemMetric correspond au nom du script (se trouvant sur le serveur dorsal) qui doit s'exécuter sur chacun des serveurs de la configuration sous le cluster indiqué (ou nom de site). Deux scripts sont fournis au client : **cpuload** et **memload**. Vous pouvez également créer des scripts de mesure système personnalisés. Le script contient une commande qui renvoie une valeur numérique comprise entre 0 et 100 ou la valeur -1 si le serveur est arrêté. Cette valeur numérique doit représenter une mesure de charge et non une valeur de disponibilité.

Remarque : Pour Site Selector, **cpuload** et **memload** sont exécutés automatiquement.

Restriction : Sous Windows, si le nom du script System Metric comporte une extension autre que ".exe", vous devez indiquer le nom complet du fichier (par exemple, "monscriptsys.bat"). Cette restriction est due à une limitation Java.

4. Ajoutez à la configuration uniquement les serveurs contenant un agent du système Metric Server s'exécutant sur le port indiqué dans le fichier *metricserver.cmd*. Le port doit correspondre à la valeur de port indiquée dans la commande **manager start**.

Remarque : Garantie de la sécurité —

- Sur la machine Load Balancer, créez un fichier de clés (à l'aide de la commande **lbkeys create**). Pour plus d'informations sur **lbkeys**, voir «RMI (Remote Method Invocation)», à la page 262.
- Sur le serveur dorsal, copiez le fichier de clés obtenu, pour le composant que vous utilisez, dans le répertoire **...ibm/edge/lb/admin/keys**. Vérifiez que le superutilisateur dispose de droits lui permettant de lire le fichier de clés.

• Agent Metric Server (côté serveur)

1. Lors de l'installation de Load Balancer, installez l'ensemble Metric Server.
2. Vérifiez le script **metricserver** dans le répertoire **/usr/bin** afin de contrôler que le port RMI souhaité est utilisé. (Pour Windows 2003, le répertoire est C:\WINDOWS\system32.) Le port RMI par défaut est 10004.

Remarque : La valeur du port RMI indiquée doit être identique à la valeur du port RMI du système Metric Server sur la machine Load Balancer.

3. Les deux scripts suivants sont déjà fournis au client : **cpuload** (renvoie le pourcentage de cpu utilisé, compris entre 0 et 100) et **memload** (donne le pourcentage de mémoire utilisée compris entre 0 et 100). Ces scripts se trouvent dans le répertoire **...ibm/edge/lb/ms/script**.

Les clients peuvent éventuellement écrire leurs propres fichiers scripts personnalisés qui définiront la commande passée par Metric Server sur les serveurs. Vérifiez que tous les scripts personnalisés sont exécutables et se trouvent dans le répertoire **...ibm/edge/lb/ms/script**. Les scripts personnalisés **doivent** renvoyer une valeur de charge comprise entre 0 et 100.

Remarque : Un script de mesure personnalisé doit être un programme valide ou un script ayant l'extension ".bat" ou ".cmd". De manière plus spécifique, pour les systèmes Linux et UNIX, les scripts doivent commencer par la déclaration de shell, sinon ils risquent de ne pas s'exécuter correctement.

4. Démarrez l'agent en émettant la commande **metricserver**.
5. Pour arrêter l'agent Metric Server, émettez la commande **metricserver stop**.

Pour exécuter le système Metric Server ailleurs que sur l'hôte local, vous devez modifier le fichier `metricserver` sur le serveur ayant fait l'objet d'un équilibrage de charge. Insérez la ligne suivante après "java" dans le fichier `metricserver` :

```
-Djava.rmi.server.hostname=AUTRE_ADRESSE
```

Ajoutez en outre la ligne suivante avant les instructions "if" dans le fichier `metricserver` : `hostname AUTRE_ADRESSE`.

Pour la plateforme Windows : Vous devez également affecter un alias à `AUTRE_ADRESSE` dans la pile Microsoft de la machine Metric Server. Par exemple :

```
call netsh interface ip add  
address "Local Area Connection"  
addr=9.37.51.28 mask=255.255.240.0
```

Lorsque vous collectez des mesures de domaines différents, vous devez affecter de manière explicite au paramètre `java.rmi.server.hostname` du script serveur (`dsrserver`, `cbrserver`, etc.) le nom de domaine complet (FQDN) de la machine qui demande les mesures. Cela est nécessaire car, suivant votre configuration et votre système d'exploitation, `InetAddress.getLocalHost.getHostName()` risque de ne pas renvoyer la valeur FQDN.

Conseiller Workload Manager

Le code de WLM ne s'exécute que sur des grands systèmes MVS. Il peut être utilisé pour demander la charge sur la machine MVS.

Si MVS Workload Management a été configuré sur votre système OS/390, Dispatcher peut accepter de WLM des informations relatives à la charge et les utiliser dans le processus d'équilibrage de charge. Grâce au conseiller WLM, Dispatcher ouvre régulièrement des connexions via le port WLM sur chaque serveur de la table d'hôte Dispatcher et accepte les chiffres relatifs à la capacité renvoyés. Comme ces chiffres représentent la capacité encore disponible et que Dispatcher attend des valeurs représentant la charge sur chaque machine, le conseiller inverse et normalise les chiffres relatifs à la capacité pour obtenir des valeurs de charge (ainsi, des chiffres de capacité élevés correspondent à des valeurs de charge faibles et représentent un serveur en bon état). Les valeurs de charge obtenues sont placées dans la colonne relative au système du rapport du gestionnaire.

Il existe plusieurs différences importantes entre le conseiller WLM et les autres conseillers Dispatcher :

1. Les autres conseillers ouvrent des connexions aux serveurs en utilisant le même port que pour le trafic client normal. Le conseiller WLM ouvre des connexions aux serveurs en utilisant un port différent de celui utilisé pour le trafic normal.

Sur chaque machine serveur, l'agent WLM doit être configuré pour effectuer l'écoute sur le port sur lequel le conseiller Dispatcher WLM a été lancé. Le port WLM par défaut est 10007.

2. Les autres conseillers évaluent uniquement les serveurs définis dans la configuration cluster:port:serveur de Dispatcher pour laquelle le port serveur correspond au port conseiller. Le conseiller WLM fournit des conseils sur *chaque* serveur de la configuration de Dispatcher (quel que soit l'élément cluster:port). Vous ne devez donc pas définir de serveurs non WLM lorsque vous utilisez le conseiller WLM.
3. Les autres conseillers placent les informations relatives à la charge dans la colonne «Port» du rapport du gestionnaire. Le conseiller WLM place les informations sur la charge dans la colonne System du rapport du gestionnaire.
4. Il est possible d'utiliser les conseillers de protocole avec le conseiller WLM. Les conseillers de protocole évaluent la charge des serveurs sur le port utilisé pour le trafic normal et le conseiller WLM évalue la charge du système sur le port WLM.

Restrictions relatives à Metric Server

Comme l'agent Metric Server, le rapport de l'agent WLM concerne les systèmes de serveur dans leur ensemble et non chacun des démons de serveur associés à un protocole. Metric Server et WLM placent leurs résultats dans la colonne relative au système du rapport du gestionnaire. Par conséquent, il n'est pas possible d'exécuter simultanément le conseiller WLM et Metric Server.

Chapitre 22. Fonctions avancées de Dispatcher, CBR et Site Selector

Le présent chapitre explique comment configurer les paramètres d'équilibrage de charge et comment installer les fonctions avancées de Load Balancer.

Remarque : Lors de la lecture de ce chapitre, si vous n'utilisez *pas* le composant Dispatcher, remplacez "dscontrol" par l'élément suivant :

- Pour CBR, utilisez **cbrcontrol**
- Pour Site Selector, utilisez **sscontrol** (voir Chapitre 28, «Guide des commandes Site Selector», à la page 401)

IMPORTANT : Si vous utilisez l'installation Load Balancer pour IPv4 et IPv6, voir Chapitre 8, «Déploiement de Dispatcher sur Load Balancer pour IPv4 et IPv6», à la page 81 pour connaître les limitations et les différences de configuration avant d'afficher le contenu du présent chapitre.

Tableau 13. Tâches de configuration avancées pour Load Balancer

Tâche	Description	Informations connexes
Placement de Load Balancer sur une machine dont il équilibre la charge	Installation d'une machine Load Balancer co-implantée.	«Utilisation de serveurs implantés au même endroit», à la page 202
Configuration de la haute disponibilité ou de la haute disponibilité réciproque	Installe une deuxième machine Dispatcher comme répartiteur de secours.	«Haute disponibilité», à la page 204
Configuration d'un équilibrage de charge basé sur des règles	Définition des conditions sous lesquelles un sous-ensemble donné de serveurs est utilisé.	«Configuration de l'équilibrage de charge basé sur des règles», à la page 212
Utilisation de la substitution d'affinité de port pour fournir au serveur un dispositif de remplacement de la fonction de maintien de routage de port.	Permet à un serveur de remplacer sur son port les paramètres de maintien de routage.	«Substitution d'affinité de port», à la page 219
Utilisation de la fonction de maintien de routage (affinité) pour fidéliser un port de cluster.	Permet aux demandes clients d'être acheminées vers le même serveur.	«Fonctionnement de la fonction d'affinité pour Load Balancer», à la page 221
Utilisation de l'affinité de ports croisés pour étendre la fonction de maintien de routage (affinité) aux autres ports.	Permet aux demandes clients reçues par des ports différents d'être dirigées vers le même serveur.	«Affinité de ports croisés», à la page 222
Utilisation du masque d'adresse d'affinité pour désigner une adresse de sous-réseau IP commune.	Permet aux demandes clients reçues par le même sous-réseau d'être dirigées vers un même serveur.	«Masque d'adresse de l'affinité (masque de maintien de routage)», à la page 222
Utilisation de l'affinité de cookie actif pour l'équilibrage de charge des serveurs pour le composant CBR	Option de règle permettant de conserver l'affinité pour un serveur particulier.	«Affinité de cookie actif», à la page 225

Tableau 13. Tâches de configuration avancées pour Load Balancer (suite)

Tâche	Description	Informations connexes
Utilisation de l'affinité de cookie pour le routage en fonction du contenu de Dispatcher et le composant CBR	Option de règle permettant de conserver l'affinité pour un serveur particulier en fonction de la valeur nom du cookie/cookie.	«Affinité de cookie passif», à la page 226
Utilisation de l'affinité d'URI pour effectuer l'équilibrage de charge au sein des serveurs Caching avec un contenu unique à placer en mémoire cache sur chaque serveur	Option de règle permettant de conserver l'affinité pour un serveur particulier en fonction de l'URI.	«Affinité d'URI», à la page 227
Configuration du support de réseau étendu pour Dispatcher	Installe une machine Dispatcher éloignée pour l'équilibrage de charge sur un réseau étendu. Ou effectue l'équilibrage de charge dans un réseau étendu (sans Dispatcher éloigné) à l'aide d'une plateforme de serveur prenant en charge GRE.	«Configuration du support de réseau étendu pour Dispatcher», à la page 228
Utilisation de liens explicites	Evitez d'ignorer Dispatcher dans vos liens.	«Utilisation de liens explicites», à la page 235
Utilisation d'un réseau privé	Configurez le répartiteur (Dispatcher) pour qu'il assure l'équilibrage de charge des serveurs sur un réseau privé.	«Utilisation d'une configuration réseau privée», à la page 235
Utilisation d'un cluster générique pour combiner des configurations serveur communes	Les adresses non explicitement configurées utiliseront le cluster générique pour équilibrer le trafic.	«Utilisation d'un cluster générique pour combiner les configurations serveurs», à la page 236
Utilisation d'un cluster générique pour équilibrer la charge des pare-feux.	La totalité du trafic sera équilibré sur les pare-feux.	«Utilisation du cluster générique pour équilibrer la charge des pare-feux», à la page 237
Utilisation d'un cluster générique avec Caching Proxy pour le proxy transparent	Permet d'utiliser Dispatcher pour activer un proxy transparent.	«Utilisation de cluster générique avec Caching Proxy pour le proxy transparent», à la page 238
Utilisation du port générique pour acheminer le trafic destiné à un port non configuré	Prend en charge le trafic qui n'est configuré pour aucun port particulier.	«Utilisation du port générique pour acheminer le trafic destiné à un port non configuré», à la page 238
Utilisation de la détection de "refus de service" pour indiquer aux administrateurs (via une alerte) des attaques éventuelles	Dispatcher analyse les demandes entrantes d'un certain nombre de connexions partielles sur les serveurs.	«Détection d'attaque de refus de service», à la page 239
Utilisation de la consignment binaire pour analyser les statistiques des serveurs.	Permet d'enregistrer les informations sur les serveurs dans des fichiers binaires et d'extraire ces informations.	«Utilisation de la consignment binaire pour analyser les statistiques des serveurs», à la page 240
Utilisation d'une configuration de client co-implanté	Autoriser Load Balancer à résider sur la même machine qu'un client	«Utilisation d'un client co-implanté», à la page 241

Utilisation de serveurs implantés au même endroit

Load Balancer peut se trouver sur la même machine qu'un serveur pour lequel il équilibre la charge des demandes. On parle alors de *co-implantation* d'un serveur. La co-implantation s'applique aux composants Dispatcher et Site Selector. Elle est également prise en charge pour le composant CBR, mais uniquement avec des serveurs Web et un serveur Caching Proxy de type serveur de liaison.

Remarque : Un serveur co-implanté est en concurrence avec Load Balancer pour les ressources aux moments de fort trafic. Toutefois, en l'absence de machines surchargées, l'utilisation d'un serveur co-implanté permet de réduire le nombre total de machines nécessaires pour configurer un site avec équilibrage de charge.

Pour le composant Dispatcher

Systèmes Linux : Pour configurer simultanément la co-implantation et la haute disponibilité lors de l'exécution du composant Dispatcher avec la méthode d'acheminement MAC, voir «Solutions alternatives pour l'affectation d'alias à l'unité de bouclage sous Linux lors de l'utilisation de la méthode d'acheminement MAC de Load Balancer», à la page 78.

Systèmes Windows : Pour configurer simultanément la co-implantation et la haute disponibilité lors de l'exécution du composant Dispatcher avec la méthode d'acheminement MAC, voir «Configuration de la co-implantation et de la haute disponibilité (systèmes Windows)», à la page 211.

Systèmes Solaris : Dans cet environnement, vous ne pouvez pas configurer de conseillers WAN lorsque le point d'entrée est co-implanté. Voir «Utilisation de conseillers éloignés avec le support de réseau étendu de Dispatcher», à la page 230.

Dans les versions précédentes, il était nécessaire de préciser que l'adresse du serveur co-implanté devait être la même que l'adresse de non-acheminement (NFA) dans la configuration. Cette restriction a maintenant été éliminée.

Pour configurer un serveur afin qu'il soit co-implanté, la commande **dscontrol server** propose l'option **collocated** qui peut être *oui* ou *non*. La valeur par défaut est non. L'adresse du serveur doit être une adresse IP valide d'une carte d'interface réseau sur la machine. Le paramètre **collocated** ne doit pas être défini pour les serveurs co-implantés à l'aide de la méthode d'acheminement NAT ou CBR de Dispatcher.

Vous pouvez configurer un serveur co-implanté de l'une des manières suivantes :

- Si vous utilisez une adresse de non-réacheminement (NFA) en tant qu'adresse de serveur co-implanté, définissez l'adresse NFA à l'aide de la commande **dscontrol executor set nfa adresse_IP**. Ajoutez ensuite le serveur en utilisant l'adresse NFA avec la commande **dscontrol server add cluster:port:serveur**.
- Si vous utilisez une adresse autre que l'adresse NFA, ajoutez le serveur avec l'adresse IP voulue tout en attribuant la valeur *oui* au paramètre **collocated** de la manière suivante : **dscontrol server add cluster:port:serveur collocated oui**.

Pour l'acheminement NAT ou CBR de Dispatcher, vous devez configurer (alias) une adresse NFA de carte réseau inutilisée. Le serveur doit être configuré pour écouter cette adresse. Configurez le serveur en utilisant la syntaxe de commande suivante :

```
dscontrol server add  
cluster:port:nouvel_alias address nouvel_alias router  
ip_routeur  
returnaddress adresse_retour
```

Si vous n'effectuez pas cette configuration, des erreurs système risquent de se produire et le serveur peut ne pas répondre.

Configuration de la co-implantation de serveur avec l'acheminement NAT de Dispatcher

Lorsque vous configurez un serveur co-implanté avec la méthode d'acheminement NAT de Dispatcher, le routeur spécifié dans la commande `dscontrol server add` doit correspondre à une adresse de routeur réelle et non à l'adresse IP du serveur.

Vous pouvez désormais configurer la prise en charge de la co-implantation sur tous les systèmes d'exploitation lors de la configuration de la méthode d'acheminement NAT de Dispatcher si vous exécutez les étapes suivantes sur la machine Dispatcher :

- **Sur les systèmes AIX**, le serveur co-implanté est configuré comme n'importe quel serveur. Aucune modification de la configuration n'est nécessaire.
- **Sur les systèmes Linux**, le serveur co-implanté est configuré comme n'importe quel serveur. Aucune modification de la configuration n'est nécessaire.
- **Sur les systèmes Solaris et HP-UX**, un alias est attribué au cluster en utilisant `ifconfig`, et l'adresse de retour ne doit pas se voir attribuer un alias mais appliquer la commande `arp publish`. Pour ce faire, exécutez la commande suivante :

```
arp -s nom_hôte ether_addr pub
```

en utilisant l'adresse MAC locale pour `ether_addr`. L'application peut ainsi envoyer le trafic à l'adresse de retour du noyau.

- **Sur les plateformes Windows**, le cluster et l'adresse de retour doivent être configurés à l'aide de la commande `dscontrol executor configure` et ne doivent pas être placés dans l'architecture en réseau Windows. Pour l'application locale, vous devez ajouter un nouvel alias IP à la carte d'extension locale de l'architecture en réseau Windows. Dans les paramètres TCP/IP, localisez l'option Avancés qui permet d'ajouter des IP à une carte. Cette deuxième adresse IP est utilisée comme définition de serveur dans la configuration de Dispatcher.

Composant CBR

Le composant CBR prend en charge la co-implantation sur toutes les plateformes sans qu'il soit nécessaire de procéder à des configurations supplémentaires. Vous devez toutefois utiliser des serveurs Web et Caching Proxy de liaison.

Pour le composant Site Selector

Site Selector prend en charge la co-implantation sur toutes les plateformes sans qu'aucune configuration supplémentaire ne soit requise.

Haute disponibilité

La fonction de haute disponibilité (configurable avec la commande `dscontrol highavailability`) est disponible pour le composant Dispatcher (mais pas pour les composants CBR et Site Selector).

Pour améliorer la disponibilité de Dispatcher, la fonction de haute disponibilité de Dispatcher utilise les mécanismes suivants :

- Deux répartiteurs connectés aux mêmes clients et au même cluster de serveurs, ainsi qu'une connectivité entre les répartiteurs. Les deux machines Dispatcher doivent utiliser le même type de système d'exploitation et la même plateforme.

- Un mécanisme de «signal de présence» entre les deux répartiteurs, afin de détecter les incidents au niveau de Dispatcher. Au moins, une des paires de signaux de présence doit disposer des NFA de la paire en tant qu'adresse source et de destination.
Au moins une des paires de signaux de présence doit utiliser si possible un sous-réseau différent du trafic classique du cluster. Un trafic distinct de signaux de présence permet d'éviter les faux relais lors des fortes charges réseau et d'améliorer les temps de reprise totale.
- Une liste de cibles à atteindre, adresses que les deux machines Dispatcher doivent pouvoir contacter pour équilibrer leur trafic correctement. Pour plus de détails, voir «Détections des incidents en utilisant signal de présence et cible à atteindre», à la page 208.
- La synchronisation des informations de Dispatcher (c'est-à-dire les tables de connexion, les tables d'accessibilité et d'autres informations).
- Une logique de sélection du répartiteur actif responsable d'un cluster de serveurs déterminé, et du répartiteur de secours qui est synchronisé en continu pour ce cluster de serveurs.
- Un mécanisme permettant une prise de contrôle IP plus rapide lorsque la logique ou un opérateur choisit l'état actif ou l'état de secours.

Remarque : Pour obtenir une description de la configuration *haute disponibilité réciproque* dans laquelle deux machines Dispatcher, qui partagent deux ensembles de cluster, se fournissent une sauvegarde mutuelle, voir «Haute disponibilité réciproque», à la page 61. La haute disponibilité réciproque est semblable à la haute disponibilité mais se base sur l'adresse de cluster en particulier et non sur l'ensemble de la machine Dispatcher. Les deux machines doivent configurer de la même façon leur ensembles de cluster partagés.

Configuration de la haute disponibilité

La syntaxe complète de la commande **dscontrol highavailability** est fournie dans la section «dscontrol highavailability — Contrôle de la haute disponibilité», à la page 365.

Pour plus d'informations sur les tâches détaillées ci-dessous, voir «Configuration de la machine Dispatcher», à la page 66.

1. Créez des fichiers script d'alias dans chacune des 2 machines Dispatcher. Voir «Utilisation de scripts», à la page 209.
2. Démarrez le serveur sur les deux machines serveurs Dispatcher.
3. Démarrez l'exécuteur sur les deux machines.
4. Assurez-vous que l'adresse de non-réacheminement (NFA) de chaque machine Dispatcher est configurée et qu'il s'agit d'une adresse IP valide pour le sous-réseau des machines Dispatcher.
5. Configurez le système de "signal de présence" (heartbeat) sur les deux machines.

```
dscontrol highavailability heartbeat add adresse_source adresse_destination
```

Remarque : Les variables *adresse_source* et *adresse_destination* ont pour valeur les adresses IP (noms DNS ou adresses IP) des machines Dispatcher. Ces valeurs seront inversées pour chaque machine. Par exemple :

```
Machine principale (primary) -  
highavailability heartbeat add 9.67.111.3 9.67.186.8  
machine de secours (backup) - highavailability heartbeat  
add 9.67.186.8 9.67.111.3
```

Au moins, une des paires de signaux de présence doit disposer des NFA de la paire en tant d'adresse source et de destination.

Au moins une des paires de signaux de présence doit utiliser si possible un sous-réseau différent du trafic classique du cluster. Un trafic distinct de signaux de présence permet d'éviter les faux relais lors des fortes charges réseau et d'améliorer les temps de reprise totale.

Définit le nombre de secondes nécessaires à l'exécuteur pour arrêter les signaux de présence de disponibilité pour dépassement du délai d'expiration. Par exemple :

```
dscontrol executor set hatimeout 3
```

La valeur par défaut est 2 secondes.

6. Sur les deux machines, utilisez la commande **reach add** pour configurer la liste des adresses IP auxquelles Dispatcher doit pouvoir accéder pour assurer un service complet. Par exemple :

```
dscontrol  
highavailability reach add 9.67.125.18
```

Les cibles à atteindre sont recommandées mais pas obligatoires. Pour plus d'informations, voir «Détections des incidents en utilisant signal de présence et cible à atteindre», à la page 208.

7. Ajoute les données de sauvegarde à chaque machine :

- Pour la machine **principale** :

```
dscontrol highavailability backup add primary [auto  
| manual] port
```

- Pour la machine de **sauvegarde** :

```
dscontrol highavailability backup add backup  
[auto | manual] port
```

- En cas de haute disponibilité réciproque, chaque machine Dispatcher joue **les deux** rôles de machine principale et de secours :

```
dscontrol highavailability backup add both [auto  
| manual] port
```

Remarque : Sélectionnez un port non utilisé sur les machines en tant que *port*. Le numéro de port entré sert de clé pour garantir que le destinataire du paquet est l'hôte correct.

8. Contrôlez l'état de la fonction de haute disponibilité pour chaque machine.

```
dscontrol highavailability status
```

Chacune des machines doit avoir le rôle (principal, secondaire ou les deux), les états et sous-états qui conviennent. La machine principale doit être active et synchronisée ; la machine de secours doit être en mode veille et se synchroniser rapidement avec l'autre. Les deux stratégies doivent être identiques.

9. Définissez les paramètres des clusters, des ports et des serveurs pour les deux machines.

Remarque : Pour la configuration de haute disponibilité réciproque (figure 14, à la page 61), Par exemple, configurez les ensembles de cluster partagés entre les deux machines Dispatcher de la manière suivante :

- Pour le Dispatcher 1, utilisez la commande :
`dscontrol cluster set clusterA hôte_principal NFAdispatcher1`
`dscontrol cluster set clusterB hôte_principal dispatcherNFA2`
- Pour le Dispatcher 2, utilisez la commande :
`dscontrol cluster set clusterB hôte_principal NFAdispatcher2`
`dscontrol cluster set clusterA hôte_principal dispatcherNFA1`

10. Démarrez le gestionnaire et les conseillers sur les deux machines.

Remarques :

1. Pour configurer une machine Dispatcher unique pour acheminer les paquets sans machine de secours, n'émettez aucune commande de haute disponibilité au moment de l'initialisation.
2. Pour passer de deux machines Dispatcher configurées pour la haute disponibilité à une seule machine autonome, arrêtez l'exécuteur sur l'une des machines, puis supprimez les fonctions de haute disponibilité (les signaux de présence, les seuils et la sauvegarde) sur l'autre machine.
3. Dans les deux cas ci-dessus, la carte d'interface réseau doit être reliée aux adresses de cluster par un alias, selon la procédure adaptée.
4. Lorsque deux machines Dispatcher sont exécutées en mode haute disponibilité et sont synchronisées, entrez toutes les commandes dscontrol (afin de mettre à jour la configuration) sur la machine de secours puis sur la machine active.
5. Lorsque deux machines Dispatcher fonctionnent en mode haute disponibilité, des résultats imprévus peuvent se produire si l'on affecte des valeurs différentes à l'un ou l'autre des paramètres de l'exécuteur, des clusters, des ports ou des serveurs (par exemple, port délai de maintien de routage).
6. Dans le cas de la haute disponibilité réciproque, tenez compte du cas où l'une des machine Dispatcher doit acheminer des paquets de données à son cluster principal tout en prenant en charge des paquets destinés au cluster de sauvegarde. Assurez-vous que le débit n'excède pas les capacités de la machine.
7. Pour les systèmes Linux, si vous souhaitez configurer simultanément la haute disponibilité et la co-implantation à l'aide de la méthode d'acheminement de port MAC du composant Dispatcher, voir «Solutions alternatives pour l'affectation d'alias à l'unité de bouclage sous Linux lors de l'utilisation de la méthode d'acheminement MAC de Load Balancer», à la page 78.
8. Pour les systèmes Windows, si vous souhaitez configurer simultanément la haute disponibilité et la co-implantation, voir «Configuration de la co-implantation et de la haute disponibilité (systèmes Windows)», à la page 211.
9. Pour tout conseil sur la réponse à des incidents potentiels de configuration de la haute disponibilité, comme :
 - suppression des connexions après la reprise,
 - synchronisation impossible des machines partenaires,
 - demandes dirigées à tort vers la machine partenaire de secours,voir «Incident : Conseils sur la configuration de la haute disponibilité», à la page 318.

Détections des incidents en utilisant signal de présence et cible à atteindre

Outre les critères de détection d'incidents de base (perte de connectivité entre le système Dispatcher de secours et le système Dispatcher actif, détectée via les messages de signal de présence), un autre mécanisme de détection d'incidents appelé *critères d'accessibilité* est disponible. Lorsque vous configurez Dispatcher, vous pouvez indiquer une liste d'hôtes auxquels chaque système Dispatcher doit pouvoir accéder pour fonctionner correctement. Les deux partenaires en haute disponibilité communiquent en continu par l'intermédiaire de signaux de présence et ils s'indiquent le nombre de cibles à contacter auxquelles ils peuvent chacun accéder (via la commande ping). Si le serveur en attente parvient à accéder à un plus grand nombre de cibles à contacter (via la commande ping) que le serveur actif, une reprise survient.

Les signaux de présence sont envoyés par la machine Dispatcher active et doivent être reçus par la machine Dispatcher en veille toutes les demi secondes. Si la machine Dispatcher en veille ne parvient pas à recevoir un signal de présence dans les 2 secondes, une reprise survient. Une reprise n'est effectuée par la machine Dispatcher de secours que si tous les signaux de présence échouent. En d'autres termes, lorsque deux paires de signaux de présence sont configurées, les deux signaux de présence doivent échouer. Pour stabiliser un environnement en haute disponibilité et éviter les reprises, ajoutez plusieurs paires de signaux de présence.

Pour les cibles à contacter, vous devez choisir au moins un hôte pour chaque sous-réseau que la machine Dispatcher utilise. Il peut s'agir de systèmes hôtes tels que les routeurs, les serveurs IP ou d'autres types d'hôtes. L'accessibilité de l'hôte est obtenue grâce au conseiller de contact, qui lance un ping à l'hôte. La reprise a lieu si les messages de signal de présence ne peuvent pas être transmis ou si les critères d'accessibilité sont mieux respectés par la machine Dispatcher de secours que par la machine Dispatcher principale. Pour prendre la décision sur la base de toutes les informations disponibles, le répartiteur actif envoie régulièrement au répartiteur de secours ses données d'accessibilité. La machine Dispatcher de secours compare ensuite ces données aux siennes, puis décide de procéder ou non au basculement.

Remarque : Lorsque vous configurez la cible d'accessibilité, vous devez également lancer le *conseiller d'accessibilité*. Il est démarré automatiquement lorsque vous démarrez la fonction gestionnaire. Pour plus d'informations sur le conseiller de contact, voir 190.

Stratégie de reprise

Deux machines Dispatcher sont configurées : la machine principale et une deuxième machine appelée machine de *sauvegarde* (ou de secours). Au démarrage, la machine principale transmet toutes les données de connexion à la machine de secours afin d'obtenir une parfaite synchronisation. La machine principale devient *active*, c'est-à-dire qu'elle commence l'équilibrage de charge. Parallèlement, la machine de secours contrôle l'état de la machine principale et conserve l'état de *veille*.

Si, à tout instant, la machine de secours décèle une défaillance de la machine principale, elle prend le *relais* des fonctions d'équilibrage de charge de la machine principale et devient, à son tour, la machine active. Une fois la machine principale redevenue opérationnelle, les machines se comportent selon la *stratégie* de reprise après incident définie par l'utilisateur. Il existe deux types de stratégie :

Automatique

La machine principale reprend le routage des paquets de données dès qu'elle redevient opérationnelle.

Manuelle

La machine de secours continue le routage des paquets de données, même après que la machine principale soit redevenue opérationnelle. Une intervention manuelle est nécessaire pour rendre de nouveau active la machine principale et replacer la machine de secours en état de veille.

La stratégie définie doit être identique pour les deux machines.

La stratégie de reprise manuelle oblige l'utilisateur à forcer le routage des paquets vers une machine spécifique, à l'aide de la commande "takeover". La reprise manuelle s'avère utile lorsque l'autre machine doit subir des opérations de maintenance. La stratégie de reprise automatique est conçue pour un fonctionnement normal sans surveillance.

Pour une configuration de haute disponibilité réciproque, il n'existe pas de défaillance par cluster. Si l'une des machines est victime d'une défaillance, même si celle-ci ne concerne qu'un des clusters, l'autre machine prendra le relais pour chacun des deux clusters.

Remarque : Des informations de connexion peuvent se perdre pendant le relais. Il peut en résulter une interruption des connexions longues existantes (telnet, par exemple) utilisées au moment du relais.

Utilisation de scripts

Pour que Dispatcher puisse acheminer les paquets de données, chaque adresse de cluster doit posséder un alias la reliant à une interface réseau.

- Dans le cadre d'une configuration Dispatcher autonome, chaque adresse de cluster doit posséder un alias la reliant à une carte d'interface réseau (par exemple, en0, tr0).
- Dans une configuration de haute disponibilité :
 - Sur la machine active, chaque adresse de cluster doit posséder un alias la reliant à une carte d'interface réseau (par exemple, en0, tr0).
 - Sur la machine de secours, chaque adresse de cluster doit être reliée par un alias à une unité de bouclage (par exemple, lo0) si vous utilisez la méthode d'acheminement MAC avec des serveurs co-implantés.
- Sur toute machine où l'exécuteur aura été arrêté, tous les alias doivent être supprimés pour éviter les conflits avec une machine venant d'être initialisée.

Pour plus d'informations sur l'association d'alias à la carte d'interface réseau, voir «Étape 5. Affectation d'un alias à la carte d'interface réseau», à la page 69.

Dans la mesure où les machines Dispatcher changent d'état lorsqu'une défaillance est détectée, les commandes citées plus haut doivent être lancées automatiquement. Dispatcher exécutera des scripts créés par l'utilisateur pour le faire. Le répertoire **...ibm/edge/lb/servers/samples** contient des scripts exemple qui *doivent* être déplacés dans le répertoire **...ibm/edge/lb/servers/bin** pour s'exécuter. Les scripts s'exécutent automatiquement uniquement si dsserver est en cours d'exécution.

Remarques :

1. Pour une configuration de haute disponibilité réciproque chacun des scripts "go" est appelé par le Dispatcher à l'aide d'un paramètre d'identification de

l'adresse principale du Dispatcher. Le script doit rechercher ce paramètre et exécuter les commandes **executor configure** des adresses de cluster associées à ce Dispatcher principal.

2. Pour configurer la haute disponibilité pour la méthode d'acheminement NAT de Dispatcher, vous devez ajouter des adresses de retour dans les fichiers script.

Les scripts exemples suivants peuvent être utilisés :

goActive

Le script goActive s'exécute lorsque l'un des systèmes Dispatcher devient actif et commence à acheminer les paquets de données.

- Si vous exécutez Dispatcher dans le cadre d'une configuration de haute disponibilité, vous devez créer ce script. Ce script doit supprimer les alias des unités de bouclage et en ajouter pour les autres périphériques.
- Si vous utilisez Dispatcher dans le cadre d'une configuration autonome, ce script n'est pas nécessaire.

goStandby

Le script goStandby s'exécute lorsqu'un système Dispatcher passe en mode veille. Il contrôle alors l'état de la machine active mais n'achemine pas de paquets.

- Si vous exécutez Dispatcher dans le cadre d'une configuration de haute disponibilité, vous devez créer ce script. Ce script doit supprimer les alias des périphériques et en ajouter pour les unités de bouclage.
- Si vous utilisez Dispatcher dans le cadre d'une configuration autonome, ce script n'est pas nécessaire.

goInOp

Le script goInOp s'exécute lorsqu'un exécuteur Dispatcher est arrêté.

- Si vous exécutez Dispatcher généralement dans le cadre d'une configuration de haute disponibilité, il peut être nécessaire de créer ce script. Ce script supprime tous les alias de périphériques et de bouclage.
- Si vous utilisez généralement Dispatcher dans une configuration autonome, ce script est facultatif. Vous pouvez le créer afin qu'il supprime les alias de périphériques, ou les supprimer manuellement.

goIdle Le script goIdle s'exécute lorsqu'un des systèmes Dispatcher devient inactif et commence à acheminer les paquets de données. Cela se produit lorsque les fonctions de haute disponibilité n'ont pas été définies, comme dans le cas d'une configuration autonome. Cela peut également arriver dans une configuration de haute disponibilité, avant que ces fonctions n'aient été définies ou après qu'elles aient été supprimées.

- Si vous exécutez Dispatcher généralement dans le cadre d'une configuration de haute disponibilité, *ne créez pas* ce script.
- Si vous utilisez généralement Dispatcher dans une configuration autonome, ce script est facultatif. Vous pouvez le créer de sorte qu'il ajoute les alias de périphériques, ou les ajouter manuellement. Si vous ne créez pas ce script dans une configuration autonome, vous devez utiliser la commande **dscontrol executor configure** ou vous devez configurer manuellement les alias à chaque lancement de l'exécuteur.

highavailChange

Le script highavailChange s'exécute lorsque l'état de haute disponibilité est modifié dans le système Dispatcher, de telle sorte qu'un des scripts "go" est appelé. Le paramètre transmis à ce script correspond au nom du script

"go" exécuté par Dispatcher. Vous pouvez créer ce script pour utiliser les informations de changement d'état, par exemple, pour alerter un administrateur ou simplement pour enregistrer l'événement.

Sur les systèmes Windows : Si, dans votre configuration, Site Selector équilibre la charge de deux machines Dispatcher fonctionnant en environnement à haute disponibilité, vous devrez définir un alias pour les systèmes Metric Server dans la pile Microsoft des systèmes Metric Server. Insérez cet alias dans le script goActive. Par exemple :

```
call netsh interface ip add
address "Local Area Connection"
addr=9.37.51.28 mask=255.255.240.0
```

Supprimez les alias des scripts goStandby et goInOp. Par exemple :

```
call netsh
interface ip delete address "Local Area Connection"
addr=9.37.51.28
```

Si la machine est équipée de plusieurs cartes d'interface réseau, vérifiez dans un premier temps l'interface à utiliser en entrant la commande suivante au niveau de l'invite : netsh interface ip show address. Elle renvoie la liste des interfaces configurées et le numéro de la connexion au réseau local (par exemple, "Local Area Connection 2") permettant de déterminer celle à utiliser.

Sous Linux pour S/390 : Dispatcher génère automatiquement un protocole de résolution d'adresse ATM pour transférer les adresses IP d'une machine Dispatcher vers une autre. Ce mécanisme est donc lié au type de réseau sous-jacent. Lors de l'exécution de Linux pour S/390, Dispatcher ne peut effectuer des reprises en haute disponibilité de manière native (complètes avec transferts d'adresse IP) que sur les interfaces qui peuvent générer automatiquement un protocole de résolution d'adresse ATM et configurer l'adresse sur l'interface locale. Ce mécanisme ne fonctionne pas sur les interfaces point à point telles qu'TUCV et CTC et ne fonctionne pas correctement dans certaines configurations de qeth/QDIO.

Pour les interfaces et configurations dans lesquelles la fonction de reprise IP native de Dispatcher ne fonctionne pas correctement, le client peut insérer les commandes appropriées dans les scripts go pour transférer manuellement les adresses. Ces topologies réseau peuvent ainsi également bénéficier de la haute disponibilité.

Configuration de la co-implantation et de la haute disponibilité (systèmes Windows)

Il est possible de configurer à la fois la haute disponibilité et la co-implantation sur les serveurs Windows. Des étapes supplémentaires sont toutefois nécessaires pour configurer ensemble ces fonctionnalités Load Balancer sur des systèmes Windows.

Sur les systèmes Windows, lorsque vous utilisez la co-implantation avec la haute disponibilité, vous avez besoin d'une adresse IP supplémentaire (espèce d'adresse IP factice) qui peut être ajoutée à l'adaptateur de bouclage sur le système Windows. L'adaptateur de bouclage doit être installé sur les machines principale et de secours. Pour installer le périphérique de bouclage sur les systèmes Windows, suivez les étapes décrites dans «Configuration des serveurs pour l'équilibrage de la charge», à la page 72.

Lorsque les étapes vous invitent à ajouter l'adresse IP de cluster au bouclage, vous devez ajouter une adresse IP factice, et non l'adresse du cluster. En effet, les scripts

go* à haute disponibilité pour les systèmes Windows doivent supprimer, puis ajouter l'adresse du cluster au périphérique de bouclage, selon que la machine Load Balancer est active ou de secours.

Les systèmes Windows ne permettent pas de supprimer la dernière adresse IP configurée du périphérique de bouclage car ce dernier ne fonctionne pas en mode DHCP. L'adresse factice permet à Load Balancer de supprimer à tout moment son adresse de cluster. L'adresse IP factice ne sera pas utilisée pour n'importe quel type de trafic et peut servir sur les machines actives ou de secours.

Mettez à jour et déplacez les scripts go* de Load Balancer sur les machines actives et de secours, puis démarrez Dispatcher. L'adresse du cluster est ajoutée, puis supprimée de l'interface réseau et du périphérique de bouclage aux moments opportuns.

Configuration de l'équilibrage de charge basé sur des règles

Vous pouvez utiliser un équilibrage basé sur des règles pour déterminer de manière précise quand et pourquoi des paquets sont envoyés à des serveurs et quels sont ces serveurs. Load Balancer parcourt toute les règles que vous ajoutez, de la première à la dernière priorité et s'arrête sur la première règle vérifiée avant d'équilibrer la charge en fonction du contenu entre les serveurs associés à cette règle. Ils équilibrent déjà la charge en fonction de la destination et du port, mais l'utilisation de règles permet d'étendre votre capacité à répartir les connexions.

Dans la plupart des cas lors de la configuration de règles, vous devez configurer une règle par défaut **toujours vraie** afin d'intercepter les demandes provenant des autres règles de priorité élevée. Il peut s'agir d'une réponse du type "Désolé, ce site est actuellement inaccessible. Faites une nouvelle tentative ultérieurement" lorsque tous les autres serveurs ne peuvent pas traiter la demande client.

Vous devez utiliser l'équilibrage de charge dépendant des règles avec Dispatcher et Site Selector lorsque vous voulez utiliser un sous-ensemble de serveurs pour une raison quelconque. Vous *devez* toujours utiliser des règles pour le composant CBR.

Vous pouvez sélectionner les types de règles suivants :

- Pour Dispatcher :
 - Adresse IP du client
 - Port du client
 - Heure
 - Type de service (TOS)
 - Nombre de connexions par seconde
 - Nombre total de connexions actives
 - Largeur de bande réservée
 - Largeur de bande partagée
 - Toujours vraie
 - Contenu d'une demande
- Pour CBR :
 - Adresse IP du client
 - Heure
 - Nombre de connexions par seconde
 - Nombre total de connexions actives

- Toujours vraie
- Contenu d’une demande
- Pour Site Selector :
 - Adresse IP du client
 - Heure
 - Mesure de tous les serveurs
 - Moyenne des mesures
 - Toujours vraie

Planifiez la logique à suivre par les règles avant de commencer à ajouter des règles à votre configuration.

Evaluation des règles

Toutes les règles possèdent un nom, un type, une priorité et peuvent avoir une valeur de début et une valeur de fin ainsi qu’un ensemble de serveurs. En outre, à la règle de type contenu du composant CBR est associée une structure d’expression standard. (Pour obtenir des exemples et des scénarios sur le mode d’utilisation de la règle de contenu et la syntaxe de motif valide pour la règle de contenu, voir Annexe B, «Syntaxe des règles de contenu (modèle)», à la page 477.)

Les règles sont évaluées en fonction de leur priorité : en d’autres termes, une règle de priorité 1 (nombre le moins élevé) avant une règle de priorité 2 (nombre plus élevé). La première règle vérifiée est utilisée. Lorsqu’une règle est vérifiée, aucune autre règle n’est évaluée.

Pour qu’une règle soit vérifiée, elle doit remplir deux conditions :

1. Le prédicat de cette règle doit être vrai. C’est-à-dire que la valeur évaluée doit être comprise entre la valeur de début et la valeur de fin, ou que le contenu doit correspondre à l’expression standard spécifiée dans la structure de la règle de type contenu. Pour les règles de type “vraie”, le prédicat est toujours respecté, quelles que soient les valeurs de début et de fin.
2. Si des serveurs sont associés à cette règle, l’un d’eux au moins doit présenter une pondération supérieure à 0 pour recevoir des paquets.

Si aucun serveur n’est associé à une règle, cette dernière ne doit remplir que la première condition pour être vérifiée. Dans ce cas, Dispatcher abandonne la demande de connexion, Site Selector renvoie la demande de serveur de nom avec une erreur et CBR provoque une page d’erreur Caching Proxy.

Si aucune règle n’est vérifiée, Dispatcher sélectionne un serveur parmi l’ensemble des serveurs disponibles du port, Site Selector sélectionne un serveur parmi l’ensemble des serveurs disponibles sur le nom du site et CBR provoque l’affichage d’une page d’erreur par Caching Proxy.

Utilisation de règles basées sur l’adresse IP des clients

Ce type de règle est disponible dans le composant Dispatcher, CBR ou Site Selector.

Vous pouvez souhaiter utiliser des règles basées sur l’adresse IP des clients pour trier les clients et leur affecter des ressources en fonction de leur provenance.

Par exemple, vous constatez la présence sur le réseau d'un nombre important de transmissions impayées et donc indésirables en provenance de clients appartenant à un ensemble spécifique d'adresses IP. Vous créez donc une règle à l'aide de la commande **dscontrol rule** , par exemple :

```
dscontrol rule add 9.67.131.153:80:ni type ip  
beginrange 9.0.0.0 endrange 9.255.255.255
```

Cette règle "ni" permet de trier les connexions des clients indésirables. Ajoutez ensuite à la règle les serveurs qui doivent être accessibles, ou si vous n'ajoutez pas de serveurs à la règle, les demandes provenant des adresses 9.x.x.x ne sont pas transmises par l'un de vos serveurs.

Utilisation de règles basées sur le port du client

Ce type de règle est disponible uniquement avec le composant Dispatcher.

Vous pouvez souhaiter utiliser des règles basées sur le port client lorsque vos clients utilisent un type de logiciel nécessitant un port spécifique de TCP/IP lors des demandes.

Vous pouvez, par exemple, créer une règle spécifiant que toute demande dont le port client est 10002, doit utiliser un ensemble de serveurs rapides spéciaux car vous savez que les demandes client associées à ce port proviennent d'un groupe de clients privilégiés.

Utilisation de règles basées sur l'heure

Ce type de règle est disponible dans le composant Dispatcher, CBR ou Site Selector.

Vous pouvez souhaiter utiliser des règles basées sur l'heure en vue de la planification des pondérations. Si par exemple votre site Web est surchargé chaque jour pendant les mêmes créneaux horaires, il est préférable de dédier cinq serveurs supplémentaires aux heures de pointe.

Ce type de règle est également intéressant lorsque vous voulez arrêter certains serveurs chaque jour à minuit, pour des raisons de maintenance. Dans ce cas, vous pouvez définir une règle qui exclut ces serveurs pendant la période de maintenance nécessaire.

Utilisation de règles basées sur le type de services (TOS)

Ce type de règle est disponible uniquement avec le composant Dispatcher.

Vous pouvez souhaiter utiliser des règles fondées sur le contenu du champ "type de service" (TOS) de l'en-tête IP. Par exemple, si la valeur TOS d'une demande client entrante indique un service normal, cette demande peut être routée vers un ensemble de serveurs. Si une autre demande arrive, munie cette fois d'une valeur TOS différente indiquant une priorité de service élevée, elle peut être dirigée vers un autre ensemble de serveurs.

La règle TOS permet de configurer entièrement chacun des bits de l'octet TOS en utilisant la commande **dscontrol rule**. Utilisez 0 ou 1 pour les bits importants que vous souhaitez apparier dans l'octet TOS. La valeur x est utilisée par défaut. Voici un exemple d'ajout d'une règle TOST :

```
dscontrol rule add 9.67.131.153:80:tsr type service tos 0xx1010x
```

Utilisation de règles basées sur le nombre de connexions par seconde

Ce type de règle est disponible dans les composants Dispatcher et CBR.

Remarque : Le gestionnaire doit être lancé pour exécuter les opérations suivantes.

Vous pouvez souhaiter utiliser des règles basées sur le nombre de connexions par seconde lorsque vous devez partager certains serveurs avec d'autres applications. Vous pouvez, par exemple, définir deux règles :

1. Si le nombre de connexions par seconde du port 80 est compris entre 0 et 2000, utiliser ces 2 serveurs
2. Si le nombre de connexions par seconde du port 80 est supérieur à 2000, utiliser ces 10 serveurs

Vous pouvez aussi utiliser Telnet et vouloir lui réserver deux des cinq serveurs, sauf lorsque le nombre de connexions par seconde dépasse un certain niveau. Ainsi, Dispatcher équilibre la charge entre les cinq serveurs, pendant les heures de pointe.

Définition de l'option d'évaluation de règle "upserversonrule" avec la règle de type "connexion" : Si, lorsque vous utilisez la règle de type connexions et définissez l'option **upserversonrule**, certains serveurs de l'ensemble de serveurs sont inactifs, vous pouvez préserver les serveurs actifs d'une surcharge. Pour plus d'informations, voir «Option d'évaluation de serveur», à la page 220.

Utilisation de règles basées sur le nombre total de connexions actives

Ce type de règle est disponible dans le composant Dispatcher ou CBR.

Remarque : Le gestionnaire doit être lancé pour exécuter les opérations suivantes.

Vous pouvez souhaiter utiliser des règles basées sur le nombre total de connexions actives d'un port lorsque les serveurs sont surchargés et commencent à ignorer certains paquets. Dans ce cas, certains serveurs Web continuent d'accepter les connexions même s'ils ne disposent pas d'un nombre suffisant d'unités d'exécution pour répondre à la demande. Il en résulte que le poste client réclame un certain délai de temporisation et que le client accédant à votre site Web n'est pas servi. Vous pouvez utiliser des règles basées sur le nombre de connexions actives pour équilibrer les pondérations d'un pool de serveurs.

Par exemple, vous savez par expérience que les serveurs arrêteront leur service après avoir accepté 250 connexions. Vous pouvez créer une règle à l'aide de la commande **dscontrol rule** ou **cbrcontrol rule**, par exemple :

```
dscontrol rule add 130.40.52.153:80:pool2 type  
active  
beginrange 250 endrange 500
```

ou

```
cbrcontrol rule add 130.40.52.153:80:pool2 type active  
beginrange 250 endrange 500
```

Vous pouvez ensuite ajouter à la règle vos serveurs en cours plus d'autres serveurs qui, autrement seraient utilisés pour d'autres processus.

Utilisation de règles basées sur la largeur de bande réservée et sur la largeur de bande partagée

Les règles de largeur de bande réservée et partagée sont disponibles uniquement dans le composant Dispatcher.

Pour les règles de largeur de bande, Dispatcher calcule la largeur de bande en tant que débit auquel les données sont délivrées aux clients par un ensemble de serveurs spécifique. Dispatcher effectue le suivi de la capacité aux niveaux du serveur, de la règle, du port, du cluster et de l'exécuteur. Pour chacun de ces niveaux, il existe une zone compteur d'octets : les kilo-octets transmis par seconde. Dispatcher calcule ces débits sur un intervalle de 60 secondes. Vous pouvez consulter ces valeurs de débit dans l'interface ou dans la sortie d'un rapport de ligne de commande.

Règle de largeur de bande réservée

La règle de largeur de bande réservée permet de contrôler le nombre de kilo-octets délivrés par seconde à un ensemble de serveurs. En définissant un seuil (définition d'une plage de largeur de bande précise) pour chaque ensemble de serveurs dans la configuration, vous pouvez contrôler et garantir le montant de la largeur de bande utilisé par chaque combinaison cluster-port.

Voici un exemple d'ajout de règle de largeur de bande réservée :

```
dscontrol rule add 9.67.131.153:80:rbw type reservedbandwidth  
beginrange 0 endrange 300
```

Les valeurs de début et de fin sont indiquées en kilo-octets par seconde.

Règle de largeur de bande partagée

Avant de configurer la règle de largeur de bande partagée, vous devez indiquer la quantité maximale de largeur de bande (kilo-octets par seconde) pouvant être partagée au niveau de l'exécuteur ou du cluster à l'aide de la commande **dscontrol executor** ou **dscontrol cluster** avec l'option `sharedbandwidth`. La valeur `sharebandwidth` ne doit pas dépasser la largeur totale de bande (capacité réseau totale) disponible. L'utilisation de la commande **dscontrol** pour définir la largeur de bande partagée ne fournit qu'une limite maximale pour la règle.

Voici des exemples de la syntaxe de la commande.

```
dscontrol executor set sharedbandwidth taille  
dscontrol cluster [add | set] 9.12.32.9 sharedbandwidth taille
```

La *taille* de l'option `sharedbandwidth` correspond à une valeur entière (kilo-octets par seconde). La valeur par défaut est zéro. Si la valeur est zéro, la bande passante ne peut pas être partagée.

Le partage de la largeur de bande au niveau du cluster permet au cluster d'utiliser une largeur de bande maximale indiquée. Tant que la largeur de bande utilisée par le cluster est inférieure à la quantité indiquée, cette règle est respectée (true). Lorsque la largeur totale de bande utilisée dépasse la quantité indiquée, cette règle n'est plus respectée (false).

Le partage de la largeur de bande au niveau de l'exécuteur permet à toute la configuration Dispatcher de partager une quantité maximale de largeur de bande. Tant que la largeur de bande utilisée au niveau de l'exécuteur est inférieure à la quantité indiquée, cette règle est respectée (true). Lorsque la largeur totale de bande utilisée dépasse la quantité définie, cette règle n'est plus respectée (false).

Ci-dessous, se trouvent des exemples d'ajout ou de définition d'une règle `sharedbandwidth`.

```
dscontrol rule add 9.20.30.4:80:shbw type sharedbandwidth sharelevel valeur  
dscontrol rule set 9.20.34.11:80:shrule sharelevel valeur
```

La *valeur* de l'option `sharelevel` est soit `exécuteur`, soit `cluster`. `Sharelevel` est un paramètre requis dans la règle `sharebandwidth`

Utilisation de règles de largeur de bande réservée et partagée

Dispatcher permet d'attribuer une largeur de bande indiquée à des ensembles de serveurs dans votre configuration à l'aide de la règle *largeur de bande réservée*. En précisant un début et une fin de plage, vous pouvez contrôler la plage de kilo-octets livrée aux clients par un ensemble de serveurs. Dès que la règle n'est plus vraie (limite de plage dépassée), la règle de priorité inférieure suivante est appliquée. S'il s'agit d'une règle "toujours vraie", un serveur peut être sélectionné pour envoyer une réponse "site occupé" au client.

Prenons, par exemple, un groupe de trois serveurs sur le port 2222. Si la largeur de bande réservée est fixée à 300, le nombre maximal de kilo-octets par seconde est 300, sur une durée de 60 secondes. Lorsque ce débit est excédé, la règle n'est plus respectée. Si cette règle est la seule, Dispatcher sélectionne l'un des trois serveurs pour traiter la demande. S'il existe une règle "toujours vraie" de priorité inférieure, la demande est dirigée vers un autre serveur et reçoit la réponse "site occupé".

La règle de largeur de bande partagée offre aux clients l'accès à des serveurs supplémentaires. Plus précisément, lorsqu'elle est utilisée comme règle de priorité inférieure faisant suite à une règle de largeur de bande réservée, un client peut encore accéder à un serveur lorsque la largeur de bande réservée a été dépassée.

Par exemple, si vous placez une règle de largeur de bande partagée après une règle de largeur de bande réservée, vous permettez aux clients d'accéder à trois serveurs de manière contrôlée. Tant qu'il reste de la largeur de bande réservée à utiliser, la règle est vraie et l'accès accordé. Lorsqu'il ne reste plus de largeur de bande réservée disponible, cette règle n'est plus vraie et la suivante est appliquée. Si la règle suivante est une règle "toujours vraie", la demande est au besoin dirigée vers un autre serveur.

En utilisant les règles de largeur de bande réservée et partagée comme indiqué dans l'exemple ci-dessus, permet plus de souplesse et de contrôle au niveau des accords (ou refus) d'accès aux serveurs. Les serveurs d'un port particulier peuvent se voir attribuer une largeur de bande limitée, tandis que d'autres peuvent utiliser autant de largeur de bande que disponible.

Remarque : Dispatcher surveille la largeur de bande utilisée en évaluant le trafic client, par exemple les accusés de réception de données (acks), transmis à un serveur. Si, pour une raison quelconque, Dispatcher ne "visualise" pas ce trafic, les résultats obtenus avec les règles de largeur de bande sont imprévisibles.

Règle Mesure de tous les serveurs

Ce type de règle est disponible uniquement dans le composant Site Selector.

Pour la règle Mesure de tous les serveurs, vous choisissez une mesure système (`cpuload`, `memload` ou votre propre script de mesure système personnalisé) et Site Selector compare la valeur de mesure système (renvoyée par l'agent du système Metric Server se trouvant dans chaque serveur d'équilibrage de charge) avec les

valeurs de début et de fin indiquées dans la règle. La valeur de mesure de tous les serveurs de l'ensemble de serveurs doit être définie dans la plage afin que la règle puisse s'exécuter.

Remarque : Le script de mesure système choisi doit se trouver sur chaque serveur d'équilibrage de charge.

Ci-dessous, se trouve un exemple d'ajout de mesure à toutes les règles de la configuration.

```
sscontrol rule add dnsload.com:allrule1 type metricall  
metricname cpuload beginrange 0 endrange 100
```

Règle Moyenne des mesures

Ce type de règle est disponible uniquement dans le composant Site Selector.

Pour la règle Moyenne des mesures, vous choisissez une mesure système (cpuload, memload ou votre propre script de mesure système personnalisé) et Site Selector compare la valeur de mesure système (renvoyée par l'agent du système Metric Server se trouvant dans chaque serveur d'équilibrage de charge) avec les valeurs de début et de fin indiquées dans la règle. La *moyenne* des valeurs des mesures système de tous les serveurs de l'ensemble de serveurs doit être définie dans la plage pour que la règle puisse s'exécuter.

Remarque : Le script de mesure système choisi doit se trouver sur chaque serveur d'équilibrage de charge.

Ci-dessous, se trouve un exemple d'ajout de règle de moyenne des mesures à votre configuration.

```
sscontrol rule add dnsload.com:avgrule1 type metricavg  
metricname cpuload beginrange 0 endrange 100
```

Utilisation de règles toujours vraies

Ce type de règle est disponible dans le composant Dispatcher, CBR ou Site Selector.

Une règle peut être créée comme règle "toujours vraie." Ces règles seront toujours sélectionnées, sauf si tous les serveurs associés sont arrêtés. Pour cette raison, leur niveau de priorité est généralement inférieur à celui de autres règles.

Vous pouvez même avoir plusieurs règles "toujours vraies", chacune d'elles associée à un ensemble de serveurs. La première règle vérifiée pour laquelle un serveur est disponible est choisie. Supposons par exemple que vous disposiez de six serveurs. Vous voulez que deux d'entre eux traitent le trafic quelles que soient les circonstances, à moins qu'ils soient tous les deux arrêtés. Si les deux premiers serveurs sont arrêtés, vous voulez qu'un deuxième jeu de serveurs traite le trafic. Enfin, si les quatre serveurs sont arrêtés, vous utiliserez les deux derniers pour traiter le trafic. Vous pouvez définir trois règles "toujours vraie". Le premier jeu de serveurs est toujours choisi, tant que l'un d'eux est actif. Si les deux serveurs sont arrêtés, l'un des serveurs du deuxième jeu est choisi, etc.

Autre exemple : il se peut que vous vouliez une règle "toujours vraie" pour garantir que les clients entrants qui ne remplissent aucune des règles définies ne sont pas pris en charge. Il vous faut donc créer, à l'aide de la commande **dscontrol rule**, la règle suivante :

```
dscontrol rule add 130.40.52.153:80:jamais type true priority 100
```


Vous n'ajoutez alors pas de serveur à cette règle, ce qui provoque l'abandon sans réponse des paquets des clients.

Remarque : Il n'est pas nécessaire de définir de valeur de début et de fin lors de la création d'un règle Toujours vraie.

Vous pouvez définir plusieurs règles "toujours vraies", puis choisir ensuite celle qui doit être exécutée en modifiant les niveaux de priorité.

Utilisation de règles basées sur le contenu des demandes

Ce type de règle est disponible dans le composant CBR ou Dispatcher (lorsque vous utilisez la méthode d'acheminement CBR de Dispatcher).

Vous pouvez utiliser des règles basées sur le contenu pour envoyer des demandes à des ensembles de serveurs spécialement configurés pour prendre en charge une partie du trafic de votre site. Par exemple, vous pouvez utiliser un ensemble de serveurs pour la prise en charge de toutes les demandes *cgi-bin*, un autre pour les demandes audio, et un troisième pour toutes les autres demandes. Vous pouvez ajouter une règle dont la structure correspond au chemin d'accès du répertoire *cgi-bin*, une autre correspondant au type de vos fichiers audio, et une troisième règle "toujours vraie" pour prendre en charge le reste du trafic. Vous ajouterez ensuite les serveurs appropriés à chaque type de règle.

Important : Pour obtenir des exemples et des scénarios sur le mode d'utilisation de la règle de contenu et la syntaxe de modèle valide pour la règle de contenu, voir Annexe B, «Syntaxe des règles de contenu (modèle)», à la page 477.

Substitution d'affinité de port

La substitution d'affinité de port permet le remplacement du maintien de routage du port d'un serveur spécifique. Par exemple, dans le cas où vous utilisez une règle pour limiter le nombre de connexions alloué à chaque serveur d'application et disposez d'un serveur de débordement à règle fixe qui annonce "please try again later" à propos de cette application. Le délai du maintien de routage du port est de 25 minutes et vous ne souhaitez pas que le client soit maintenu sur ce serveur. La substitution d'affinité de port vous permet alors de changer de serveur de débordement afin de remplacer l'affinité qui est habituellement associée à ce port. Ainsi, les demandes ultérieures du client destinées au cluster seront transmises au serveur d'applications ayant le plus de disponibilités et non pas au serveur de débordement, afin d'équilibrer la charge.

Pour plus d'informations sur la syntaxe de commande de la substitution d'affinité de port, en utilisant l'option de serveur **maintien de routage**, voir «dscontrol server — Configuration des serveurs», à la page 390. .

Ajout de règles à la configuration

Vous pouvez ajouter des règles à l'aide de la commande **dscontrol rule add** , en modifiant le fichier de configuration exemple ou en utilisant l'interface graphique. Vous pouvez ajouter une ou plusieurs règles à chaque port que vous avez défini.

Il s'agit d'une procédure en deux étapes : vous devez ajouter la règle, puis définir les serveurs sur lesquels les services doivent être effectués si la règle est vraie. Supposons par exemple que l'administrateur système veuille déterminer le taux d'utilisation des serveurs proxy pour chaque division du site. Une adresse IP est

octroyée à chaque division. Créez le premier jeu de règles en fonction des adresses IP des clients pour séparer les charges individuelles de chaque division :

```
dscontrol rule add 130.40.52.153:80:div1 type ip b 9.1.0.0 e 9.1.255.255
dscontrol rule add 130.40.52.153:80:div2 type ip b 9.2.0.0 e 9.2.255.255
dscontrol rule add 130.40.52.153:80:div3 type ip b 9.3.0.0 e 9.3.255.255
```

Ajoutez ensuite un serveur distinct à chaque règle, puis mesurez la charge de chaque serveur pour facturer correctement les divisions en fonction des services qu'elles utilisent. Par exemple :

```
dscontrol rule useserver
130.40.52.153:80:div1 207.72.33.45
dscontrol rule useserver 130.40.52.153:80:div2 207.72.33.63
dscontrol rule useserver 130.40.52.153:80:div3 207.72.33.47
```

Option d'évaluation de serveur

L'option d'évaluation de serveur est disponible uniquement dans le composant Dispatcher.

La commande **dscontrol rule** a une option d'évaluation de serveur pour les règles. L'option *evaluate* permet de choisir d'évaluer les conditions de la règle parmi tous les serveurs du port ou d'évaluer les conditions de la règle des serveurs faisant partie de la règle. (Dans les versions précédentes de Load Balancer, il n'était possible de mesurer que la condition de chaque règle parmi tous les serveurs du port.)

Remarques :

1. L'option d'évaluation de serveur est valide uniquement pour les règles prenant leurs propres décisions en fonction des caractéristiques des serveurs : règle du nombre total de connexions (par seconde), règle des connexions actives et règle de largeur de bande réservée.
2. La règle de type "connexion" dispose d'une option d'évaluation supplémentaire, **upserversonrule**. Pour plus d'informations, voir «Utilisation de règles basées sur le nombre de connexions par seconde», à la page 215.

Ci-dessous, se trouvent des exemples d'ajout ou de définition de l'option d'évaluation au niveau d'une règle de largeur de bande réservée.

```
dscontrol rule add 9.22.21.3:80:rbweval type reservedbandwidth evaluate niveau
dscontrol rule set 9.22.21.3:80:rbweval evaluate niveau
```

Vous pouvez attribuer la valeur port, règle ou upserversonrule au *niveau* d'évaluation. La valeur par défaut est port.

Evaluation des serveurs dans la règle

L'option permettant de mesurer les conditions de la règle dans les serveurs de la règle permet de configurer deux règles ayant les caractéristiques suivantes :

- La première règle évaluée contient tous les serveurs gérant le contenu du site Web et la valeur *règle* est attribuée à l'option d'évaluation (évaluation des conditions de la règle dans les serveurs de la règle).
- La deuxième règle est une règle Toujours vraie qui contient un seul serveur qui répond par une réponse de type "site occupé".

Par conséquent, lorsque le trafic dépasse le seuil des serveurs dans la première règle, le trafic est envoyé au serveur "site occupé" dans la deuxième règle. Lorsque le trafic est inférieur au seuil des serveurs de la première règle, le trafic reprend pour les serveurs de la première règle.

Evaluation des serveurs sur le port

Lors de l'utilisation des deux règles décrites dans l'exemple précédent, si vous attribuez l'option *port* à l'option d'évaluation pour la première règle (évaluation des conditions de la règle dans tous les serveurs du port), lorsque le trafic dépasse la limite de cette règle, le trafic est envoyé au serveur "site occupé" associé à la deuxième règle.

La première règle mesure l'ensemble du trafic du serveur (incluant le serveur "site occupé") sur le port afin de déterminer si le trafic a dépassé la limite. Alors que la surcharge diminue pour les serveurs associés à la première règle, le trafic se poursuit sur le serveur "site occupé" étant donné que le trafic sur ce port dépasse toujours la limite de la première règle.

Fonctionnement de la fonction d'affinité pour Load Balancer

Pour les composants Dispatcher et CBR : Vous activez la fonction d'affinité lorsque vous faites en sorte que le maintien de routage d'un port de clusters soit conservé. Lorsque vous configurez un port de cluster de telle sorte que le routage soit conservé, les demandes clients peuvent être dirigées vers le même serveur. Cette opération peut être effectuée en attribuant un nombre de secondes à l'option **délai de maintien de routage** au niveau de l'exécuteur, du cluster ou du port. Lorsque vous attribuez la valeur zéro au délai de maintien de routage, cette fonction est désactivée.

Lors de l'activation de l'affinité de ports croisés, les valeurs de délai de maintien de routage des ports partagés doit avoir la même valeur (différente de zéro). Pour plus d'informations, voir «Affinité de ports croisés», à la page 222.

Pour le composant Site Selector : Vous activez la fonction d'affinité lorsque vous faites en sorte que le maintien de routage d'un nom de site soit conservé. La configuration du maintien de routage pour un nom de site permet au client d'utiliser le même serveur pour plusieurs requêtes de service annuaire. Cette opération peut être effectuée en attribuant un nombre de secondes à l'option **délai de maintien de routage** sur le nom de site. Lorsque vous attribuez la valeur zéro au délai de maintien de routage, cette fonction est désactivée.

Le délai de maintien de routage pour un serveur est le délai entre la fermeture d'une connexion et l'ouverture d'une nouvelle connexion au cours de laquelle un client est renvoyé au même serveur utilisé lors de la première connexion. Passé le délai de maintien de routage, le client peut être envoyé à un serveur autre que le premier. Ce délai est configuré à l'aide de la commande `dscontrol executor, port ou cluster`. Lorsqu'un serveur est arrêté à l'aide de la commande `server down` (`dscontrol server down`), si le délai de maintien de routage a une valeur différente de zéro pour ce serveur, les clients existants continuent à être servis par ce serveur jusqu'à expiration du délai. Le serveur n'est arrêté qu'après expiration du délai de maintien de routage.

Comportement lorsque l'affinité est désactivée

Lorsque l'affinité est désactivée, dès qu'une connexion TCP est reçue d'un client, Load Balancer utilise le serveur correct et lui transmet les paquets. Si une autre connexion arrive du même client, Load Balancer la traite comme si les deux connexions n'étaient pas liées et utilise à nouveau le serveur correct.

Comportement lorsque l'affinité est activée

Lorsque l'affinité est activée, si une autre demande est reçue du même client, la demande est acheminée vers le même serveur.

Progressivement, le client arrête d'envoyer des transactions et l'enregistrement d'affinité disparaît. D'où l'importance du "délai" de maintien de routage. La durée de vie de chaque enregistrement d'affinité est déterminée en secondes par le "délai de maintien de routage". Lorsque des demandes suivantes sont reçues lors du délai de maintien de routage, l'enregistrement d'affinité est toujours valide et la demande est toujours acheminée vers le même serveur. Si aucune connexion supplémentaire n'est reçue lors du délai de maintien de routage, l'enregistrement est vidé. Un nouveau serveur sera sélectionné pour toute connexion reçue une fois ce délai écoulé.

La commande `server down (dscontrol server down)` est utilisée pour arrêter un serveur. Ce dernier ne s'arrête pas tant que le délai de maintien de routage n'arrive pas à expiration.

Affinité de ports croisés

L'affinité de ports croisés s'applique uniquement aux méthodes d'acheminement MAC et NAT/NATP du composant Dispatcher.

L'affinité de ports croisés se définit comme l'extension à plusieurs ports de la fonction maintien de routage. Par exemple, si la demande d'un client est d'abord reçue par un port puis une deuxième demande de ce client par un autre port, l'affinité de ports croisés permet au répartiteur d'envoyer cette demande au même serveur. Pour utiliser cette fonction, les ports doivent :

- partager la même adresse de cluster
- partager les mêmes serveurs
- avoir le même (non nul) **délai de maintien de routage** valeur
- avoir le même **masque de maintien de routage** valeur

Il est possible de relier plusieurs ports au même **trans ports**. Quand un même client se connectera, à l'avenir, sur le même port ou sur un port partagé, ses connexions seront traitées par le même serveur. Voici un exemple de configuration de ports multiples munis d'une affinité de ports croisés au port 10 :

```
dscontrol port set cluster:20 crossport 10
dscontrol port set cluster:30 crossport 10
dscontrol port set cluster:40 crossport 10
```

Après l'établissement de l'affinité de ports croisés, vous disposez de la possibilité de modifier le délai de maintien de routage du port. Il est cependant recommandé de choisir la même valeur pour tous les délais de maintien de routage des ports partagés. Dans le cas contraire, des résultats inattendus peuvent se produire.

Pour supprimer l'affinité de ports croisés, remplacez la valeur **trans ports** sur son propre numéro de port. Voir «`dscontrol port` — Configuration des ports», à la page 378, pour plus d'informations sur la syntaxe de commande de l'option **trans ports**.

Masque d'adresse de l'affinité (masque de maintien de routage)

Le masque d'adresse de l'affinité ne s'applique qu'au composant Dispatcher.

Le masque d'adresse de l'affinité est une amélioration de la fonction de maintien de routage, destinée aux clients de groupe situés à des adresses de sous-réseau communes. La sélection du **masque de maintien de routage** de la commande **dscontrol port** permet de masquer les bits communs à poids fort de l'adresse IP sur 32 bits. Si cette fonction est configurée lors de la première connexion d'un client à un port, alors toutes les connexions suivantes des clients ayant la même adresse de sous-réseau (indiquée par la partie masquée de l'adresse) seront dirigées vers le même serveur.

Remarque : Pour activer le masque de maintien de routage, le **maintien de routage** du port doit être défini à zéro.

Par exemple, si vous souhaitez que toutes les demandes client disposant d'une adresse de réseau classe A identique soient envoyées au même serveur, fixez à 8 (bits) la valeur du port du masque de maintien de routage. En ce qui concerne les demandes de clients de groupe possédant la même adresse de réseau classe B, fixez la valeur du masque de maintien de routage à 16 (bits). Pour les demandes de clients de groupe disposant de la même adresse de réseau classe C, fixez la valeur du masque de maintien de routage à 24 (bits).

Pour obtenir de meilleurs résultats, fixez la valeur du masque de maintien de routage dès le lancement de Load Balancer. Si vous modifiez cette valeur, les résultats seront imprévisibles.

Interaction avec l'affinité de ports croisés : Lors de l'activation de l'affinité de ports croisés, les valeurs du masque de maintien de routage des ports partagés doivent être identiques. Pour plus d'informations, voir «Affinité de ports croisés», à la page 222.

Pour activer le masque d'adresse d'affinité, émettez une commande **dscontrol port** du type :

```
dscontrol port set cluster:port stickytime 10 stickymask 8
```

Les valeurs possibles des masques de maintien de routage sont 8, 16, 24 et 32. Une valeur de 8 indique que les 8 premiers bits à poids fort de l'adresse IP (adresse de réseau classe A) seront masqués. Une valeur de 16 indique que les 16 premiers bits à poids fort de l'adresse IP (adresse de réseau classe B) seront masqués. Une valeur de 24 indique que les 24 premiers bits à poids fort de l'adresse IP (adresse de réseau classe C) seront masqués. Si vous spécifiez une valeur de 32, l'adresse IP toute entière sera masquée, ce qui entraînera la désactivation de la fonction de masque d'adresse d'affinité. La valeur par défaut du masque de maintien de routage est 32.

Voir «dscontrol port — Configuration des ports», à la page 378, pour plus d'informations sur la syntaxe de commande du masque de maintien de routage(fonction de masque d'adresse d'affinité).

Mise au repos de la gestion des connexions serveur

La mise au repos de la gestion des connexions s'applique aux composants Dispatcher et CBR.

Pour retirer un serveur de la configuration Load Balancer quelle qu'en soit la raison (mises à jour, mises à niveau, service, etc.), vous pouvez utiliser la commande **dscontrol manager quiesce** . La sous-commande **quiesce** permet aux connexions de s'achever (sans avoir été traitées) et transmet les connexions

ultérieures du client vers le serveur mis au repos si la connexion est associée à un délai de maintien de routage et que ce dernier n'est pas arrivé à expiration. La sous-commande `quiesce` désactive toute nouvelle connexion au serveur.

Mise au repos de la gestion des connexions avec maintien de routage

Utilisez l'option "Mettre au repos maintenant" si le délai de maintien de routage est défini et que vous voulez que les nouvelles connexions soient envoyées à un autre serveur (et non au serveur mis au repos) avant que le délai de maintien de routage n'expire. Voici un exemple d'utilisation de cette option pour mettre le serveur 9.40.25.67 au repos :

```
dscontrol manager quiesce 9.40.25.67 now
```

L'option `now` détermine comment les options avec maintien de routage seront gérées.

- Si vous n'indiquez *pas* "now," les connexions existantes sont terminées et les connexions ultérieures sont transmises au serveur mis au repos à partir des clients ayant des connexions associées à un délai de routage tant que le serveur mis au repos reçoit la nouvelle demande avant expiration du délai de routage. Cependant, si vous n'avez pas activé la fonction de maintien de routage (affinité), le serveur mis au repos ne peut pas recevoir de nouvelles connexions. Il s'agit de la manière la moins brusque de placer des serveurs au repos. Par exemple, vous pouvez mettre au repos un serveur puis attendre le moment où le trafic est faible (peut-être le matin) pour retirer complètement le serveur de la configuration.
- En indiquant l'option "now," vous mettez le serveur au repos. Il permet dorénavant aux connexions existantes de se terminer mais il ne permet pas les nouvelles connexions, incluant les connexions provenant de clients ayant des connexions existantes avec délai de maintien de routage. Il s'agit de la manière la plus brusque de mettre des serveurs au repos, qui était seule disponible dans les versions précédentes de Load Balancer.

Option d'affinité de la règle basé sur le contenu de la demande du client

Vous pouvez définir les types d'affinité suivants dans la commande `dscontrol rule` :

- Active cookie — Permet l'équilibrage de charge du trafic Web et une affinité avec le même serveur en fonction des cookies générés par Load Balancer.
L'affinité de cookie actif ne s'applique qu'au composant CBR.
- Passive cookie — Permet l'équilibrage de la charge du trafic Web et une affinité avec le même serveur en fonction des cookies d'auto-identification générés par les serveurs. Vous devez définir le paramètre `cookieName` conjointement à l'affinité de cookie passif dans la commande `rule`.
L'affinité de cookie passif s'applique au composant CBR et à la méthode d'acheminement CBR du composant Dispatcher.
- URI — Permet l'équilibrage de charge du trafic Web vers des serveurs Caching Proxy dans le but d'augmenter la capacité de la mémoire cache.
L'affinité d'URI s'applique au composant CBR et à la méthode d'acheminement CBR du composant Dispatcher.

L'option d'affinité par défaut est "none". La valeur zéro (inactif) doit être associée à l'option `stickytime` de la commande `port` pour affecter la valeur de cookie actif ou

URI à l'option d'**affinité** de la commande rule. Lorsque l'affinité de cette dernière est définie, il devient impossible d'activer l'option stickytime de la commande port.

Affinité de cookie actif

L'affinité de cookie actif ne s'applique qu'au composant CBR.

Elle permet de "fidéliser" les clients à un serveur donné. Pour l'activer, attribuez un nombre positif au paramètre **stickytime** (délai de maintien de routage) d'une règle et optez pour l'affinité de cookie actif ("activecookie"), lors de l'ajout de la règle ou à l'aide de la commande rule set. Pour obtenir des informations sur la syntaxe de cette commande, voir «dscontrol rule — Configuration des règles», à la page 384.

Lorsqu'une règle a été activée pour l'affinité de cookie actif, l'équilibrage de charge des nouvelles demandes client est effectué à l'aide d'algorithmes CBR standard, tandis que les demandes suivantes du même client sont envoyées au serveur initialement choisi. Le serveur choisi est stocké en tant que cookie dans la réponse au client. Tant que les demandes suivantes du client contiennent ce cookie et qu'elles arrivent dans le délai de maintien de routage, le client conserve l'affinité pour le serveur initial.

L'affinité de cookie actif permet d'assurer qu'un client fait l'objet d'un équilibrage de charge vers le même serveur pendant une période déterminée. A cet effet, un cookie est envoyé pour être stocké par le navigateur des clients. Ce cookie indique les cluster:port:règle adoptés pour la prise de décision, le serveur cible de l'équilibrage de charge et la date d'expiration de l'affinité. Il est au format suivant : **IBMCBR=cluster:port:règle+serveur-heure!** Les informations *cluster:port:règle* et *serveur* sont codées pour ne révéler aucune information sur la configuration CBR.

Fonctionnement de l'affinité de cookie actif

Lorsqu'une règle est déclenchée et que l'affinité de cookie actif est activée, le cookie envoyé par le client est examiné.

- En cas de détection d'un cookie contenant l'identificateur des cluster:port:règle déclenchés, le serveur cible de l'équilibrage de charge et la date d'expiration sont extraits du cookie.
- Si le serveur figure toujours dans l'ensemble utilisé par la règle, que son poids est positif ou qu'il s'agit d'un serveur au repos, et que la date d'expiration est ultérieure à la date du jour, le serveur indiqué dans le cookie est choisi comme serveur cible pour l'équilibrage de charge.
- Si l'une de ces trois conditions n'est pas remplie, le choix d'un serveur s'effectue avec un algorithme normal.
- Lorsqu'un serveur a été choisi (à l'aide de l'une des deux méthodes), un nouveau cookie est constitué avec IBMCBR, cluster:port:règle, serveur_choisi et un horodatage. Ce dernier est la date d'expiration de l'affinité. La valeur "cluster:port:règle et serveur_choisi" est codée pour ne révéler aucune information sur la configuration CBR.
- Un paramètre d'"expiration" est également inclus dans le cookie. Le format de ce paramètre est lisible par le navigateur et entraîne la non-validité du cookie sept jours après la date d'expiration. Ainsi, la base de cookies du client n'est pas encombrée.

Le nouveau cookie est ensuite inséré dans les en-têtes qui reviennent au client. Le navigateur du client renvoie les demandes suivantes lorsqu'il est configuré pour accepter les cookies.

Chaque instance d'affinité du cookie a une longueur de 65 octets et se termine au point d'exclamation. Ainsi, un cookie de 4096 octets peut contenir environ 60 règles de cookie actif individuel par domaine. Une fois le cookie entièrement plein, les instances d'affinité expirées sont supprimées. Si toutes les instances sont encore valides, les plus anciennes sont supprimées et les nouvelles pour la règle en cours ajoutées.

Remarque : CBR remplace toutes les occurrences de cookies à l'ancien format IBM CBR détectées sur le proxy.

Pour la commande `rule`, vous ne pouvez attribuer que la valeur `activecookie` (cookie actif) à l'option d'affinité de cookie actif lorsque le délai de maintien de routage est zéro (non activé). Lorsque l'affinité de cookie actif est active au niveau d'une règle, vous ne pouvez pas activer le maintien de routage sur le port.

Activation de l'affinité du cookie actif

Pour activer l'affinité de cookie actif pour une règle donnée, utilisez la commande `rule set` :

```
rule set cluster:port:règle stickytime 60
rule set cluster:port:règle affinity activecookie
```

Objectifs de l'affinité de cookie actif

Le maintien de routage d'une règle est normalement utilisé pour les interfaces CGI ou les servlets qui enregistrent l'état du client sur le serveur. Cet état est identifié par un ID cookie (cookie serveur). L'état du client figure uniquement sur le serveur sélectionné. Pour maintenir cet état d'une demande à l'autre, le client a besoin du cookie du serveur.

Remplacement du délai d'expiration de l'affinité de cookie actif

Le délai d'expiration par défaut de l'affinité de cookie actif correspond au délai du serveur auquel s'ajoute le délai de maintien de routage plus vingt-quatre heures. Si les heures sont inexactes sur les systèmes de vos clients (ceux qui envoient des requêtes à la machine CBR), par exemple, s'ils dépassent de plus de 24 heures le délai du serveur), ces systèmes ignorent les cookies provenant de la CBR car ils considèrent que ces cookies ont déjà expiré. Pour rallonger le délai d'expiration, modifiez le script `cbrserver`. Dans le fichier script, modifiez la ligne `javaw` en y ajoutant le paramètre suivant après `LB_SERVER_KEYS` : `-DCOOKIEEXPIREINTERVAL=X` où `X` correspond au nombre de jours à ajouter au délai d'expiration.

Sur les systèmes AIX, Solaris et Linux, le fichier `cbrserver` se trouve dans le répertoire `/usr/bin`.

Sur les systèmes Windows, il se trouve dans le répertoire `\winnt\system32`.

Affinité de cookie passif

L'affinité de cookie passif s'applique à la méthode d'acheminement CBR (content-based routing) du composant Dispatcher et au composant CBR. Pour obtenir la procédure de configuration de la méthode d'acheminement CBR de Dispatcher, voir «Fonction CBR de Dispatcher (méthode d'acheminement `cbr`)», à la page 55.

L'affinité de cookie passif offre une façon de fidéliser les clients à un serveur donné. Lorsque vous attribuez la valeur "cookiepassif" à l'affinité d'une règle, l'affinité de cookie passif permet d'équilibrer la charge du trafic Web destiné au même serveur, en fonction des cookies d'auto-identification générés par les serveurs. Vous configurez l'affinité de cookie passif au niveau de la règle.

Lors du déclenchement de la règle, si l'affinité de cookie passif est activée, Load Balancer choisit le serveur en fonction du nom du cookie se trouvant dans l'en-tête HTTP de la demande client. Load Balancer commence à comparer le nom du cookie dans l'en-tête HTTP du client à la valeur du cookie configurée pour chaque serveur.

La première fois que Load Balancer détecte un serveur dont la valeur de cookie *contient* le nom de cookie du client, Load Balancer choisit ce serveur pour la demande.

Remarque : Load Balancer offre cette flexibilité pour traiter les cas où le serveur risque de générer une valeur de cookie dont une partie variable a été ajoutée à la partie statique. Par exemple, la valeur de cookie du serveur peut correspondre au nom du serveur (valeur statique) auquel un horodatage a été ajouté (valeur variable).

Si le nom du cookie est introuvable dans la demande client ou s'il ne correspond à aucun contenu des valeurs de cookie des serveurs, le serveur est choisi à l'aide de la méthode de sélection existante ou de la technique de permutation circulaire pondérée.

Pour configurer l'**affinité de cookie passif** :

- Pour Dispatcher, configurez d'abord la méthode d'acheminement CBR de Dispatcher. (Voir «Fonction CBR de Dispatcher (méthode d'acheminement cbr)», à la page 55.) Cette étape est ignorée pour le composant CBR.
- Attribuez la valeur "passivecookie" au paramètre **affinity** (affinité) dans la commande **dscontrol rule [add|set]**. De plus, le nom du cookie devant être recherché par Load Balancer dans la demande d'en-tête HTTP doit être attribué au paramètre **cookieName** (nom du cookie).
- Dans la commande **dscontrol server [add|set]**, définissez le paramètre **cookievalue** (valeur de cookie) pour chaque serveur de l'ensemble de serveurs associé à la règle.

Pour la commande rule, vous ne pouvez attribuer que la valeur "passivecookie" (cookie passif) à l'option d'affinité de cookie passif lorsque le délai de maintien de routage est zéro (non activé). Lorsque l'affinité de cookie passif est active au niveau d'une règle, il n'est pas possible d'activer le maintien de routage sur le port.

Affinité d'URI

L'affinité d'URI s'applique à la méthode d'acheminement CBR de Dispatcher et au composant CBR. Pour obtenir la procédure de configuration de la méthode d'acheminement CBR de Dispatcher, voir «Fonction CBR de Dispatcher (méthode d'acheminement cbr)», à la page 55.

L'affinité URI permet d'équilibrer le trafic Web sur des serveurs Caching Proxy, ce qui permet de mettre en mémoire cache un contenu unique sur un serveur spécifique. Ainsi, vous augmentez la capacité de la mémoire cache du site en éliminant les éléments superflus placés en mémoire cache sur plusieurs machines.

Configurez l'affinité d'URI au niveau de la règle. Une fois que la règle est déclenchée, si l'affinité d'URI est activée et que le même ensemble de serveurs est actif et répond, Load Balancer transmet les nouvelles demandes client entrantes ayant le même URI au même serveur.

Généralement, Load Balancer peut distribuer des demandes à plusieurs serveurs gérant un contenu identique. Lorsque vous utilisez Load Balancer avec un groupe de serveurs de mise en mémoire cache, le contenu dont l'accès est souvent demandé peut être placé dans la mémoire cache de tous les serveurs. En répliquant le contenu placé en mémoire cache identique, vous pouvez prendre en charge un grand nombre de clients. Cela est particulièrement utile pour les sites Web dont le volume est important.

Cependant, si le site Web prend en charge un trafic client modéré vers un contenu très divers et que vous préférez une mémoire cache répartie sur plusieurs serveur, votre site sera plus performant si chaque serveur de mise en cache comporte un contenu unique et que Load Balancer distribue la demande uniquement au serveur de mise en cache avec ce contenu.

Avec l'affinité d'URI, Load Balancer vous permet de distribuer le contenu mis en cache vers des serveurs individuels, éliminant la mise en cache superflue de contenu sur plusieurs machines. Grâce à cette fonction, les performances des sites ayant un contenu divers utilisant les serveurs Caching Proxy sont améliorées. Les demandes identiques sont envoyées au même serveur, plaçant ainsi en mémoire cache le contenu uniquement sur les serveurs individuels. La taille de la mémoire cache s'accroît avec chaque nouveau serveur ajouté au pool.

Pour configurer l'**affinité d'URI** :

- Pour Dispatcher, configurez d'abord la méthode d'acheminement CBR de Dispatcher. (Voir «Fonction CBR de Dispatcher (méthode d'acheminement cbr)», à la page 55.) Cette étape est ignorée pour le composant CBR.
- Attribuez la valeur "uri" au paramètre **affinité** dans la commande **dscontrol rule [add|set]** ou **cbrcontrol rule [add|set]** .

Pour la commande rule, vous ne pouvez attribuer que la valeur "URI" à l'option d'affinité d'URI lorsque le délai de maintien de routage est zéro (non activé). Lorsque l'affinité d'URI est active au niveau d'une règle, il n'est pas possible d'activer le maintien de routage sur le port.

Configuration du support de réseau étendu pour Dispatcher

Cette fonction est disponible uniquement pour le composant Dispatcher.

Si vous n'utilisez ni le support de réseau étendu de Dispatcher ni la méthode d'acheminement CBR de Dispatcher, la configuration de Dispatcher requiert que la machine Dispatcher et ses serveurs soient tous connectés au même segment de réseau local (voir figure 35, à la page 229). Une demande client arrive à la machine Dispatcher et est envoyée au serveur. Du serveur, la réponse est renvoyée directement au client.

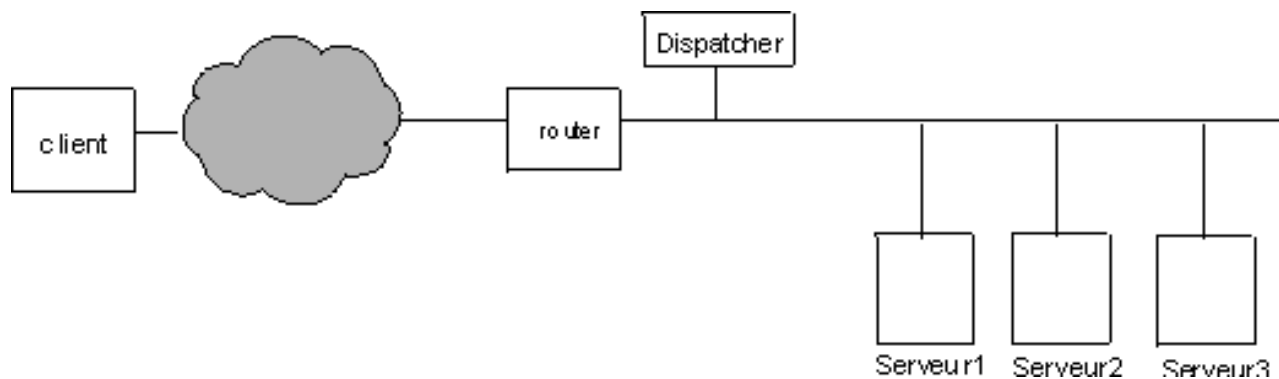


Figure 35. Exemple de configuration consistant en un seul segment de réseau local

La fonction de répartiteur étendu permet la prise en charge des serveurs hors site, appelés *serveurs éloignés* (voir la figure 36). Si GRE n'est pas pris en charge sur le site distant et que la méthode d'acheminement NAT de Dispatcher n'est pas utilisée, le site distant doit correspondre à une machine Dispatcher éloignée (Dispatcher 2) et aux serveurs associés localement (Serveur G, Serveur H et Serveur I). Le paquet d'un client passe d'Internet à la machine Dispatcher initiale. Il passe ensuite à un système Dispatcher éloigné géographiquement et enfin à l'un de ses serveurs locaux.

Tous les systèmes Dispatcher (locaux ou éloignés) doivent exécuter le même type de système d'exploitation et de plateforme pour exécuter des configurations de réseau étendu.

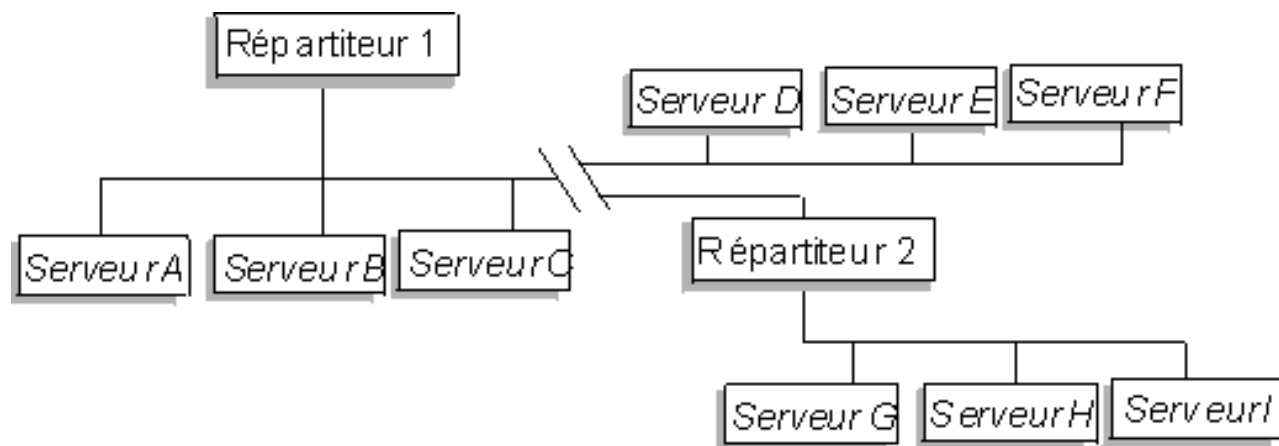


Figure 36. Exemple de configuration utilisant des serveurs locaux et éloignés

Cela permet à une adresse de cluster de supporter l'ensemble des demandes client du monde entier, tout en répartissant la charge entre l'ensemble des serveurs.

Le système Dispatcher qui reçoit le paquet en premier peut tout de même être connecté à des serveurs locaux et répartir la charge entre ses serveurs locaux et les serveurs éloignés.

Syntaxe des commandes

Pour configurer un support de réseau étendu :

1. Ajoutez les serveurs. Lorsque vous ajoutez un serveur à un répartiteur (Dispatcher), vous devez définir si ce serveur est local ou éloigné (voir plus

haut). Pour ajouter un serveur et le définir comme serveur local, entrez la commande **dscontrol server add** sans spécifier de routeur. Il s'agit de la valeur par défaut. Pour définir ce serveur comme serveur éloigné, vous devez spécifier le routeur par l'intermédiaire duquel le répartiteur doit envoyer le paquet afin d'atteindre le serveur éloigné. Le serveur doit être un autre répartiteur et son adresse doit être l'adresse de non-réacheminement du répartiteur. Par exemple, dans figure 37, à la page 232, si vous ajoutez *LB 2* en tant que serveur éloigné dans *LB 1*, vous devez définir *router 1* comme l'adresse de routage. Syntaxe générale :

```
dscontrol server add cluster:port:serveur router adresse
```

Pour obtenir plus d'informations sur le mot clé *router*, voir «dscontrol server — Configuration des serveurs», à la page 390.

2. Configurez les alias. Sur la première machine Dispatcher (à laquelle la demande client est transmise depuis Internet), un alias doit être affecté à l'adresse du cluster à l'aide de la commande **executor configure**. (Pour les systèmes Linux ou UNIX, vous pouvez utiliser la commande **executor configure** ou **ifconfig**.) Toutefois, sur les machines Dispatcher éloignées, l'adresse du cluster *n'a pas* d'alias défini sur la carte d'interface réseau.

Utilisation de conseillers éloignés avec le support de réseau étendu de Dispatcher

Sur les machines Dispatcher servant de point d'entrée :

Les machines Dispatcher servant de point d'entrée et fonctionnant sous AIX, Linux (à l'aide de GRE) ou Solaris affichent correctement les charges des conseillers. Les autres plateformes doivent compter sur l'équilibrage de la charge à l'aide de la technique de permutation circulaire ou utiliser les méthodes de transfert nat/cbr de Dispatcher au lieu d'un réseau étendu.

Systèmes AIX

- Aucune configuration spéciale n'est requise.

Systèmes HP-UX

- Dans une configuration de réseau étendu, l'emploi de conseillers éloignés est soumis à certaines restrictions lorsqu'un système Dispatcher servant de point d'entrée s'exécute sur une plateforme HP-UX. Avec la méthode d'acheminement MAC de Dispatcher, les conseillers HP-UX visent toujours directement l'adresse du serveur au lieu du cluster. Comme ils ne visent pas le cluster, le Dispatcher éloigné n'équilibre pas la charge de la demande du conseiller sur les serveurs éloignés. Les conseillers éloignés fonctionnent toutefois parfaitement avec l'acheminement CBR ou NAT de Dispatcher.

Systèmes Linux

- Dans une configuration de réseau étendu, l'emploi de conseillers éloignés est soumis à certaines restrictions lorsqu'un système Dispatcher servant de point d'entrée s'exécute sur une plateforme Linux. Avec la méthode d'acheminement MAC de Dispatcher, les conseillers Linux visent toujours directement l'adresse du serveur au lieu du cluster. Comme ils ne visent pas le cluster, le Dispatcher éloigné n'équilibre pas la charge de la demande du conseiller sur les serveurs éloignés. Les conseillers éloignés fonctionnent toutefois parfaitement avec l'acheminement CBR ou NAT de Dispatcher.
- Si vous utilisez GRE (generic routing encapsulation) pour envoyer le trafic à un serveur éloigné sans Dispatcher éloigné dans votre configuration, aucune

restriction ne s'applique à l'emploi de conseillers lors de l'exécution de la méthode d'acheminement MAC, NAT ou CBR de Dispatcher sur une plateforme Linux. Pour plus d'informations sur GRE, voir «Support GRE (Generic Routing Encapsulation)», à la page 234.

Systèmes Solaris

- Dans une configuration de réseau étendu, si vous utilisez une machine Dispatcher servant de point d'entrée et fonctionnant sur une plateforme Solaris, vous devez utiliser la méthode de configuration arp, au lieu des méthodes de configuration de l'exécuteur ifconfig ou dscontrol. Par exemple :

```
arp -s mon_adresse_cluster mon_adresse_mac pub
```

- Les restrictions applicables à la plateforme Solaris sont les suivantes :
 - Les conseillers de réseau étendu fonctionnent uniquement avec la méthode arp de la configuration de cluster.
 - Les conseillers des serveurs de liaison fonctionnent uniquement avec la méthode arp de la configuration de cluster.
 - Les conseillers des serveurs de liaison fonctionnent uniquement avec la méthode arp de la configuration de cluster. Lorsque vous utilisez des conseillers de serveurs de liaison, ne co-implantez pas Load Balancer sur le même serveur que l'application de liaison.

Systèmes Windows

- Dans une configuration de réseau étendu, l'emploi de conseillers éloignés est soumis à certaines restrictions lorsqu'un point d'entrée s'exécute sur une plateforme Windows. Avec la méthode d'acheminement MAC de Dispatcher, les conseillers Windows visent toujours directement l'adresse du serveur au lieu du cluster. Comme ils ne visent pas le cluster, le Dispatcher éloigné n'équilibre pas la charge de la demande du conseiller sur les serveurs éloignés. Les conseillers éloignés fonctionnent toutefois parfaitement avec l'acheminement CBR ou NAT de Dispatcher.

Sur des répartiteurs éloignés : Exécutez les étapes de configuration suivantes pour chaque adresse de cluster éloignée. Pour définir une configuration de haute disponibilité à l'emplacement éloigné du composant Dispatcher, vous devez effectuer l'opération sur les deux systèmes.

Systèmes AIX

- Pour que les conseillers fonctionnent correctement, Dispatcher doit faire en sorte que chaque cluster soit configuré dans l'interface avec un masque de réseau 255.255.255.255. Utilisez l'une des syntaxes suivantes pour configurer un cluster :
 - `ifconfig nom_interface alias adresse_cluster netmask 255.255.255.255.`
Par exemple,
`ifconfig en0 alias 10.10.10.99 netmask 255.255.255.255`
 - `dscontrol executor configure adresse_interface nom_interface masque_reseau.` Par exemple,
`dscontrol executor configure 204.67.172.72 en0 255.255.255.255`

Remarque : Il est nécessaire de disposer de conseillers qui fonctionnent à la fois sur des machines Dispatcher locales et éloignées.

Systèmes HP-UX, Linux, Solaris et Windows

- Aucune configuration supplémentaire n'est requise.

Exemple de configuration

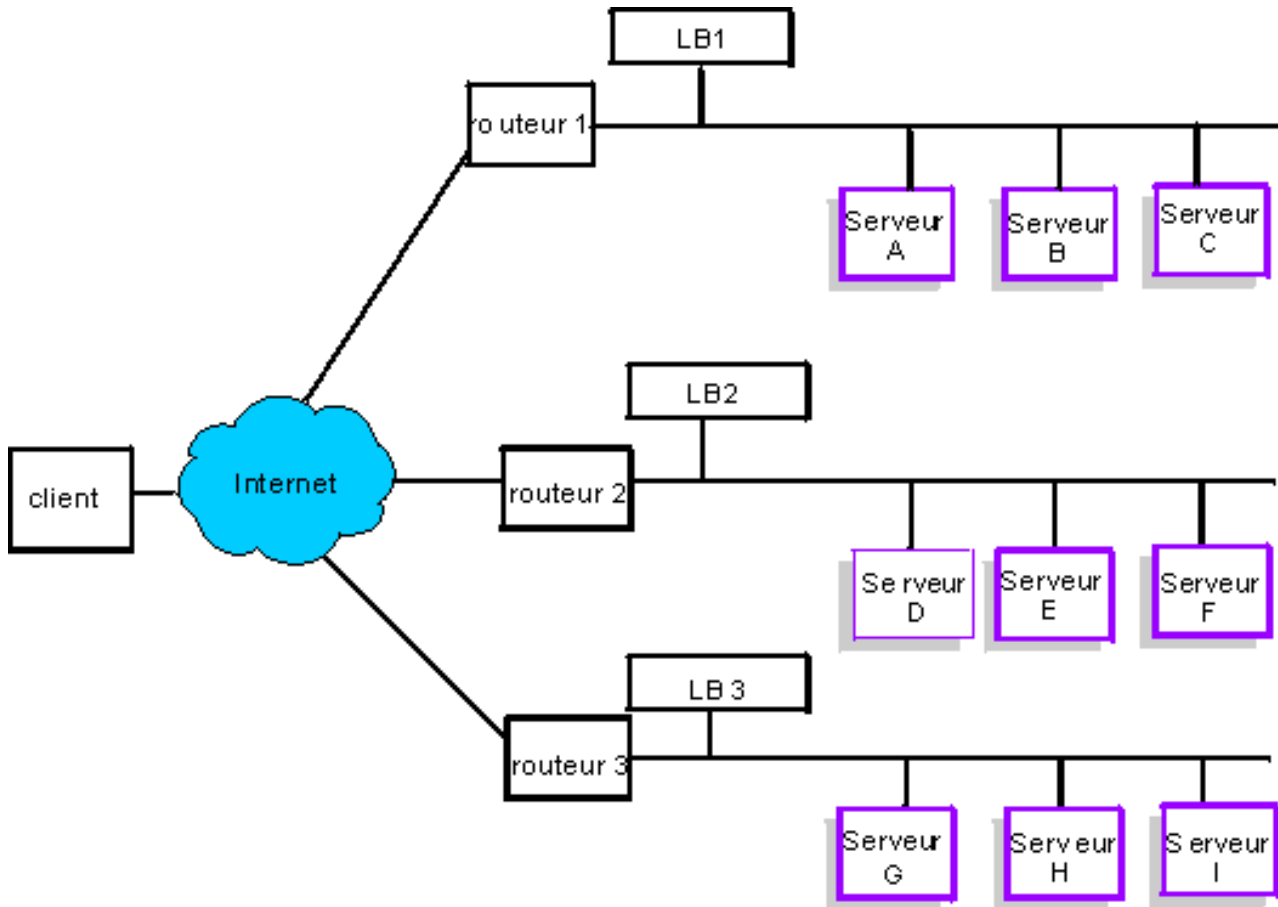


Figure 37. Exemple de configuration en réseau étendu avec des composants Load Balancer éloignés

Cet exemple s'applique à la configuration illustrée à la figure 37.

Vous trouverez ci-après la méthode à utiliser pour configurer les machines Dispatcher afin qu'elles supportent l'adresse de cluster xebec sur le port 80. LB1 est défini comme «point d'entrée». Il est supposé qu'une connexion Ethernet est utilisée. LB1 comporte cinq serveurs définis : trois serveurs locaux (ServeurA, ServeurB, ServeurC) et deux serveurs éloignés (LB2 et LB3). Par ailleurs, trois serveurs locaux ont été définis pour chacun des serveurs éloignés LB2 et LB3.

Sur la console de la première machine Dispatcher (LB1), procédez comme suit :

1. Lance l'exécuteur.
dscontrol executor start
2. Définissez l'adresse de non-réacheminement de la machine Dispatcher.
dscontrol executor set nfa LB1
3. Définissez le cluster.
dscontrol cluster add xebec
4. Définissez le port.
dscontrol port add xebec:80
5. Définissez les serveurs.
 - a. **dscontrol server add xebec:80:ServeurA**

- b. **dscontrol server add xebec:80:ServeurB**
 - c. **dscontrol server add xebec:80:ServeurC**
 - d. **dscontrol server add xebec:80:LB2 router Routeur1**
 - e. **dscontrol server add xebec:80:LB3 router Routeur1**
6. Configurez l'adresse de cluster.
dscontrol executor configure xebec

Sur la console de la deuxième machine Dispatcher (LB2) :

- 1. Lance l'exécuteur.
dscontrol executor start
- 2. Définissez l'adresse de non-réacheminement de la machine Dispatcher.
dscontrol executor set nfa LB2
- 3. Définissez le cluster.
dscontrol cluster add xebec
- 4. Définissez le port.
dscontrol port add xebec:80
- 5. Définissez les serveurs.
 - a. **dscontrol server add xebec:80:ServeurD**
 - b. **dscontrol server add xebec:80:ServeurE**
 - c. **dscontrol server add xebec:80:ServeurF**

Sur la console de la troisième machine Dispatcher (LB3) :

- 1. Lance l'exécuteur.
dscontrol executor start
- 2. Définissez l'adresse de non-réacheminement de la machine Dispatcher.
dscontrol executor set nfa LB3
- 3. Définissez le cluster.
dscontrol cluster add xebec
- 4. Définissez le port.
dscontrol port add xebec:80
- 5. Définissez les serveurs.
 - a. **dscontrol server add xebec:80:ServeurG**
 - b. **dscontrol server add xebec:80:ServeurH**
 - c. **dscontrol server add xebec:80:ServeurI**

Remarques

- 1. Sur tous les serveurs (A-I) reliez l'adresse de cluster par un alias à une unité de bouclage.
- 2. Les clusters et les ports sont ajoutés avec la commande **dscontrol** sur toutes les machines Dispatcher concernées : la machine définie comme point d'entrée et tous les serveurs éloignés.
- 3. Voir «Utilisation de conseillers éloignés avec le support de réseau étendu de Dispatcher», à la page 230 pour plus d'informations sur l'utilisation des conseillers éloignés avec le support de réseau étendu.
- 4. Le support de réseau étendu ne permet pas les boucles de routage infinies. (Si une machine Dispatcher reçoit un paquet d'une autre machine Dispatcher, elle ne le transmet pas à une troisième machine Dispatcher.) Un réseau étendu n'accepte qu'un niveau de serveurs éloignés.

5. Un réseau étendu accepte les protocoles UDP et TCP.
6. Un réseau étendu fonctionne avec la haute disponibilité : chaque machine Dispatcher peut être assistée d'une machine de secours adjacente (sur le même segment de réseau local).
7. Le gestionnaire et les conseillers fonctionnent en réseau étendu et, s'ils sont utilisés, doivent être lancés sur l'ensemble des machines Dispatcher concernées.
8. Load Balancer ne prend en charge la fonction WAN que sous des systèmes d'exploitation analogues.

Support GRE (Generic Routing Encapsulation)

GRE (Generic Routing Encapsulation) est un protocole Internet défini dans RFC 1701 et RFC 1702. Lorsque vous utilisez GRE, Load Balancer peut encapsuler les paquets IP de clients dans des paquets IP/GRE et les transmettre aux plateformes de serveur telles qu'OS/390 qui prend en charge GRE. Le support GRE permet au composant Dispatcher d'équilibrer la charge des paquets sur plusieurs adresses de serveurs associées à une adresse MAC.

Load Balancer implémente GRE en tant qu'élément de sa fonction de réseau étendu. Ainsi, Load Balancer peut fournir l'équilibrage de charge de réseau étendu directement aux systèmes de serveur pouvant désencapsuler les paquets GRE. Il n'est pas nécessaire que Load Balancer soit installé sur le site éloigné si les serveurs éloignés prennent en charge les paquets GRE encapsulés. Load Balancer encapsule les paquets WAN, la valeur décimale 3735928559 étant attribuée à l'ensemble de zones de clés GRE.

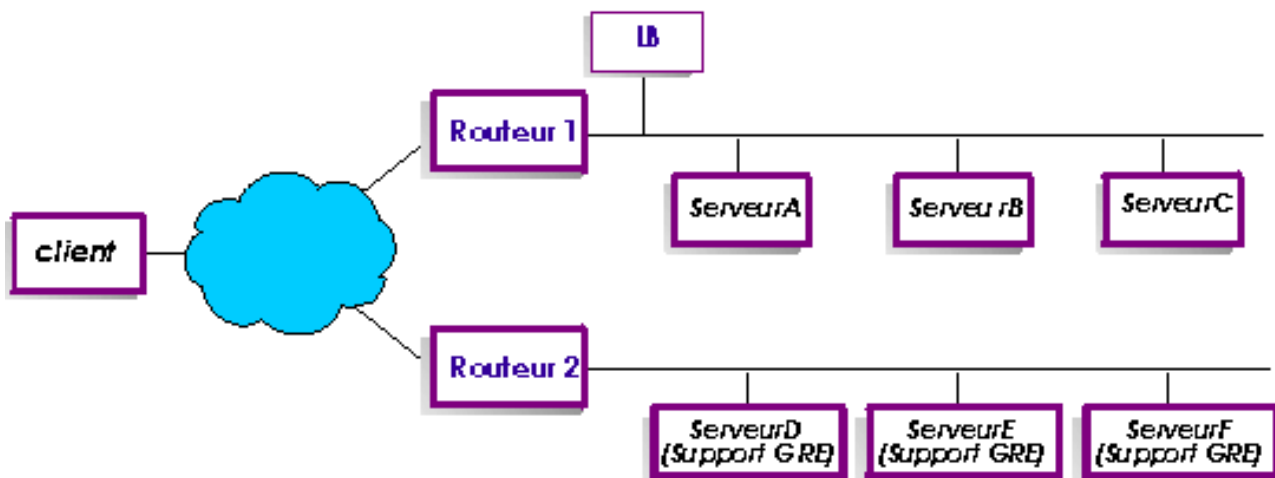


Figure 38. Exemple de configuration en réseau étendu avec une plateforme serveur prenant en charge GRE

Dans cet exemple, (figure 38), afin d'ajouter le serveurD éloigné, qui prend en charge GRE, définissez-le dans la configuration Load Balancer comme si vous définissiez un serveur WAN dans la hiérarchie cluster:port:serveur :

```
dscontrol server add cluster:port:ServeurD router Routeur1
```

Pour les systèmes Linux, configuration d'excapsulation GRE pour réseau étendu (WAN)

Les systèmes Linux disposent en natif d'une possibilité d'excapsulation GRE permettant à Load Balancer d'équilibrer la charge d'images de serveur Linux pour S/390, lorsque de nombreuses images de serveur se partagent une adresse MAC. Ainsi, le point d'entrée Load Balancer peut équilibrer la charge directement sur des

serveurs Linux WAN sans passer par Load Balancer sur le site éloigné. Les conseillers du point d'entrée Load Balancer peuvent alors traiter directement chaque serveur éloigné.

Sur le point d'entrée Load Balancer, procédez à la configuration décrite pour WAN.

Pour configurer chaque serveur dorsal Linux, entrez les commandes ci-après en tant que superutilisateur. (Ces commandes peuvent être ajoutées à la fonction de démarrage du système de sorte que les modifications sont conservées entre les réamorçages.)

```
# modprobe ip_gre
# ip tunnel add gre-nd mode gre ikey 3735928559
# ip link set gre-nd up
# ip addr add adresse_cluster dev gre-nd
```

Remarque : Le serveur Linux configuré à l'aide de ces instructions *ne doit pas* se trouver sur le même segment physique que le point d'entrée Load Balancer. En effet, le serveur Linux répond aux requêtes "ARP who-has" pour l'adresse de cluster, provoquant ainsi une condition d'indétermination qui peut conduire à un "short-circuit" dans lequel tout trafic vers l'adresse de cluster n'est acheminé que vers le gagnant de l'indétermination ARP.

Utilisation de liens explicites

En règle générale, les fonctions d'équilibrage de charge de Dispatcher fonctionnent indépendamment de la physionomie des sites sur lesquels le produit est installé. Il existe une zone, cependant, dans laquelle le contenu du site peut s'avérer important, et dans laquelle les décisions prises au sujet de ce contenu peuvent avoir une influence non négligeable sur l'efficacité de Dispatcher. Il s'agit de la zone d'adressage de liens.

Lorsque vos pages indiquent des liens pointant sur des serveurs individuels de votre site, un client est en réalité forcé à s'adresser à une machine déterminée, détournant de ce fait toute fonction d'équilibrage de charge éventuellement mise en oeuvre. Pour cette raison, utilisez systématiquement l'adresse de Dispatcher dans tous les liens figurant sur vos pages. Il est à noter que le type d'adressage utilisé n'est pas toujours apparent, notamment si votre site applique une procédure de programmation automatique permettant la création dynamique de HTML. Pour optimiser l'équilibrage de charge, identifiez les éventuelles occurrences d'adressage explicite et évitez-les autant que possible.

Utilisation d'une configuration réseau privée

Vous pouvez configurer Dispatcher et les machines serveurs TCP de sorte qu'ils utilisent un réseau privé. Cette configuration peut réduire l'encombrement des accès utilisateurs ou du réseau externe, susceptible d'affecter les performances.

Pour les systèmes AIX, cette configuration peut également tirer parti des vitesses élevées du commutateur SP High Performance Switch, si vous utilisez Dispatcher et les machines serveurs TCP comme noeuds sur un cadre SP.

Pour créer un réseau privé, chaque machine doit être équipée d'au moins deux cartes de réseau local, l'une d'elles étant reliée au réseau privé. La deuxième carte

de réseau local doit être également configurée sur un sous-réseau différent. La machine Dispatcher transmettra alors les demandes aux machines serveurs TCP par l'intermédiaire du réseau privé.

Systèmes Windows : Configurez également chaque adresse NFA avec la commande `executor configure`.

Les serveurs ajoutés à l'aide de la commande **`dscontrol server add`** doivent être ajoutés avec les adresses de réseau privé. Par exemple, reprenant le cas du serveur A de la figure 39, la syntaxe de la commande sera la suivante :

`dscontrol server add adresse_cluster:80:10.0.0.1`

et non

`dscontrol server add adresse_cluster:80:9.67.131.18`

Si vous utilisez Site Selector pour fournir des données de charge à Dispatcher, Site Selector doit être configuré pour acheminer ces états vers les adresses privées.

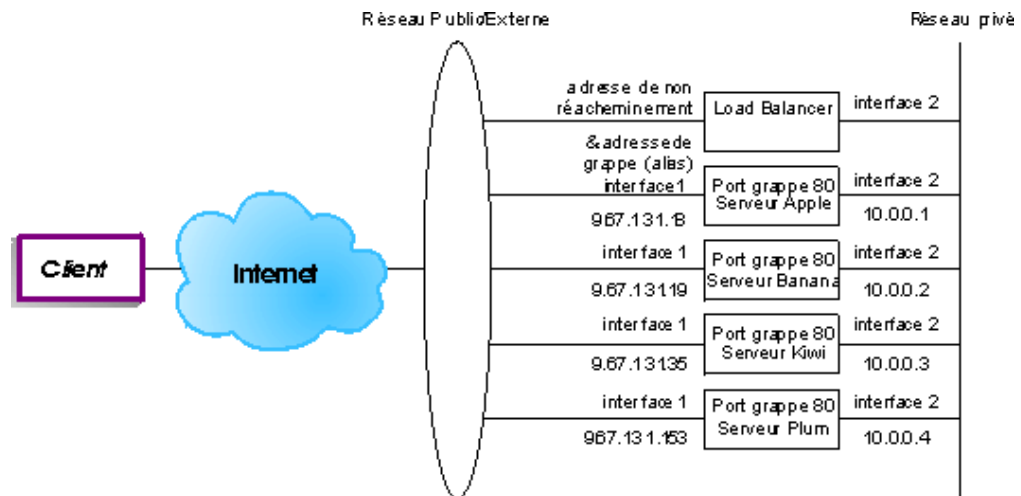


Figure 39. Exemple de réseau privé utilisant Dispatcher

L'utilisation d'une configuration de réseau privé ne s'applique qu'au composant Dispatcher.

Utilisation d'un cluster générique pour combiner les configurations serveurs

L'utilisation d'un cluster générique pour combiner les configurations serveurs ne s'applique qu'au composant Dispatcher.

Le terme "générique" fait référence à l'aptitude du cluster à s'adapter à plusieurs adresses IP (c'est-à-dire à fonctionner comme un "joker"). L'adresse 0.0.0.0 permet d'indiquer un cluster générique.

Si vous devez assurer l'équilibrage de plusieurs adresses de cluster ayant des configurations port/serveur identiques, vous pouvez combiner tous les clusters dans une seule configuration de cluster générique.

Vous devez toujours configurer de manière explicite chaque adresse de cluster sur les cartes réseau de votre poste Dispatcher. Toutefois, vous ne devez ajouter aucune des adresses de cluster à la configuration de Dispatcher à l'aide de la commande `dscontrol cluster add`.

Ajoutez simplement le cluster générique (adresse 0.0.0.0), et configurez les ports et les serveurs correctement pour l'équilibrage de charge. Tout trafic à destination des adresses configurées sur les cartes est équilibré en utilisant la configuration du cluster générique.

L'avantage de cette approche réside dans le fait que le trafic vers toutes les adresses de clusters est pris en compte lors du choix du meilleur serveur. Si un cluster est particulièrement chargé et qu'il a créé de nombreuses connexions sur l'un des serveurs, le trafic vers les autres adresses de cluster est équilibré en tenant compte de cette information.

Vous pouvez combiner le cluster générique avec des clusters réels si certaines adresses de cluster ont une configuration port/serveur unique alors que d'autres partagent la même configuration. Les configurations uniques doivent être attribuées à une adresse de cluster réelle. Toutes les configurations communes peuvent être attribuée au cluster générique.

Utilisation du cluster générique pour équilibrer la charge des pare-feux

L'utilisation du cluster générique pour équilibrer la charge des pare-feux ne s'applique qu'au composant Dispatcher. L'adresse 0.0.0.0 permet d'indiquer un cluster générique.

Vous pouvez utiliser le cluster générique pour équilibrer le trafic vers des adresses qui ne sont pas explicitement configurées sur une carte réseau du poste Dispatcher. Pour que cela fonctionne, le répartiteur doit au moins être en mesure de voir la totalité du trafic qu'il est supposé équilibrer. Le poste répartiteur ne verra pas le trafic vers des adresses non explicitement configurées sur l'une de ses cartes réseau, excepté s'il est configuré en tant que route par défaut pour certains trafic.

Une fois Dispatcher configuré comme route par défaut, le trafic TCP ou UDP passant par la machine Dispatcher est équilibré en utilisant la configuration du cluster générique.

C'est ce principe qui est utilisé pour équilibrer les pare-feux. Les pare-feux peuvent traiter des paquets à destination de n'importe quelle adresse et de n'importe quel port. Pour cette raison, vous devez être en mesure d'équilibrer le trafic indépendamment de l'adresse ou du port cible.

Les pare-feux permettent de gérer le trafic de clients non sécurisés vers des serveurs sécurisés et les réponses de ces serveurs, ainsi que le trafic de clients sécurisés vers des serveurs non sécurisés et les réponses de ces derniers.

Vous devez configurer deux machines Dispatcher : l'une pour envoyer le trafic non-sécurisé vers les adresses de pare-feux non sécurisés et l'autre le trafic sécurisé vers les adresses de pare-feux sécurisés. Comme ces deux machines Dispatcher doivent utiliser le cluster générique et le port générique avec des adresses de serveur différentes, les deux répartiteurs doivent se trouver sur deux machines distinctes.

Utilisation de cluster générique avec Caching Proxy pour le proxy transparent

L'utilisation du cluster générique avec Caching Proxy pour le proxy transparent ne s'applique qu'au composant Dispatcher. L'adresse 0.0.0.0 permet d'indiquer un cluster générique.

La fonction de cluster générique permet également d'utiliser Dispatcher pour activer une fonction de proxy transparent pour un serveur Caching Proxy se trouvant sur le même système que Dispatcher. Cette fonction est disponible sous AIX uniquement car il doit y avoir communication entre le composant dispatcher et le composant TCP du système d'exploitation.

Pour activer cette fonction, vous devez lancer Caching Proxy écoutant les demandes client sur le port 80. Vous configurez ensuite un cluster générique (0.0.0.0). Dans le cluster générique, vous configurez le port 80. Dans le port 80, vous configurez l'adresse NFA de la machine Dispatcher en tant que serveur unique. Désormais, tout trafic client vers une adresse du port 80 est acheminé vers le serveur Caching Proxy exécuté sur le poste de travail Dispatcher. La demande client est ensuite traitée par un proxy comme d'habitude et la réponse est envoyée de Caching Proxy au client. Dans ce mode, le composant Dispatcher n'effectue pas l'équilibrage de charge.

Utilisation du port générique pour acheminer le trafic destiné à un port non configuré

Le port générique permet de gérer le trafic destiné à un port non explicitement configuré. Ce principe est utilisé pour équilibrer la charge des pare-feux. Il est également utilisé pour assurer la bonne gestion du trafic destiné à un port non configuré. En définissant un port générique sans serveur, vous garantisiez que toutes les demandes en direction de ce port non configuré sont supprimées et non renvoyées au système d'exploitation. Le numéro de port 0 (zéro) permet d'indiquer un port générique, par exemple :

```
dscontrol port add cluster:0
```

Port générique pour le traitement du trafic FTP

Lors de la configuration d'un cluster pour le traitement du port FTP passif et du port générique, le port FTP passif utilise par défaut l'intégralité de la fourchette de ports TCP non privilégiés pour les connexions aux données. Cela signifie que pour un client connecté via un cluster d'équilibrage de charge à un port de contrôle FTP, toutes les connexions de contrôle ultérieures et les connexions aux ports dont le numéro est élevé (port > 1023) à ce même cluster seront automatiquement acheminées par Load Balancer vers le même serveur que la connexion de contrôle FTP.

Si le port générique et le port FTP d'un même cluster ne possèdent pas le même jeu de serveurs, les applications dont le numéro de port est élevé (port > 1023) peuvent échouer lorsqu'il existe une connexion de contrôle FTP pour un client. Par conséquent, la configuration de jeux de serveurs différents pour le port FTP et le port générique sur un même cluster n'est pas recommandée. Si vous optez pour ce scénario, la fourchette de ports passifs du démon FTP doit être définie dans la configuration de Load Balancer.

Détection d'attaque de refus de service

Cette fonction est disponible uniquement pour le composant Dispatcher.

Dispatcher permet de détecter les attaques de "refus de service" possible et d'en alerter l'administrateur. Pour cela, Dispatcher analyse les demandes entrantes d'un certain nombre de connexions partielles TCP sur les serveurs, point commun des attaques de refus de service. Dans une attaque de refus de service, un site reçoit un grand nombre de paquets SYN d'un grand nombre d'adresses IP source et de numéros de port source, mais le site ne reçoit pas les paquets suivants de ces connexions TCP. De cette manière, vous avez un grand nombre de connexion partielles TCP sur les serveurs. Les serveurs peuvent devenir très lents et peuvent ne plus accepter de nouvelles connexions entrantes.

Remarque : Dispatcher ne détermine la fin d'une attaque de refus de service qu'en cas de trafic entrant via le cluster et le port. Dispatcher n'est pas capable de détecter que l'attaque est terminée tant que le trafic n'a pas repris.

Load Balancer fournit des exit utilisateur qui déclenchent des scripts que vous pouvez personnaliser. Ces scripts avertissent l'administrateur d'une attaque de refus de service possible. Dispatcher met à votre disposition les fichiers script exemple ci-après dans le répertoire **...ibm/edge/lb/servers/samples**.

- halfOpenAlert — une attaque de refus de service (DoS) probable a été détectée
- halfOpenAlertDone — l'attaque DoS est terminée

Pour pouvoir exécuter les fichiers, vous devez les déplacer dans le répertoire **...ibm/edge/lb/servers/bin** et supprimer l'extension de fichier ".sample".

Pour implémenter la détection d'attaque DoS, définissez le paramètre **maxhalfopen** dans la commande **dscontrol port** de la manière suivante :

```
dscontrol port set 127.40.56.1:80 maxhalfopen 1000
```

Dans l'exemple ci-dessus, Dispatcher compare le nombre total de connexions partielles (pour tous les serveurs se trouvant sur le cluster 127.40.56.1 du port 80) avec la valeur maximale de 100 (indiqué par le paramètre maxhalfopen). Si le nombre de connexions partielles dépasse la limite, un script d'alerte (halfOpenAlert) est appelé. Lorsque le nombre de connexions partielles est inférieur à la limite, un autre script d'alerte (halfOpenAlertDone) est effectué pour indiquer que l'attaque est terminée.

Pour déterminer comment définir la valeur maxhalfopen : Régulièrement (toutes les 10 minutes, par exemple), effectuez un rapport de connexion partielle (**dscontrol port halfopenaddressreport cluster:port**) lorsque le trafic de votre site est normal ou élevé. Le rapport de connexion partielle renvoie un message "Nombre total de connexions partielles reçues". Vous devez attribuer au paramètre maxhalfopen une valeur supérieure de 50 à 200% au nombre le plus élevée de connexions partielles rencontrées par votre site.

Le paramètre halfopenaddressport permet d'effectuer un rapport de données statistiques ainsi que de générer des entrées dans le journal (**..ibm/edge/lb/servers/logs/dispatcher/halfOpen.log**) pour toutes les adresses client (jusqu'à environ 8000 paires d'adresses) qui ont accédé à des serveurs disposant de connexions partielles.

Remarque : Il existe une alarme SNMP correspondant aux scripts halfOpenAlert et halfOpenAlertDone. Si le sous-agent SNMP est configuré et en cours d'exécution, les alarmes correspondantes sont envoyées dans les conditions de déclenchement de scripts. Pour plus d'informations sur le sous-agent SNMP, voir «Utilisation du protocole SNMP (Simple Network Management Protocol, protocole simplifié de gestion de réseau) avec le composant Dispatcher», à la page 269.

Pour renforcer la protection des serveurs dorsaux contre les attaques de refus de service, vous pouvez configurer des clusters et des ports génériques. En particulier, ajoutez sous chaque cluster configuré un port générique sans serveur. Ajoutez également un cluster générique doté d'un port générique, mais sans serveur. Ces actions ont pour résultat le rejet des paquets qui ne sont pas adressés à un cluster ni à un port non génériques. Pour obtenir des informations sur les clusters et les ports génériques, voir «Utilisation d'un cluster générique pour combiner les configurations serveurs», à la page 236 et «Utilisation du port générique pour acheminer le trafic destiné à un port non configuré», à la page 238.

Utilisation de la consignation binaire pour analyser les statistiques des serveurs

Remarque : La fonction de consignation binaire s'applique à Dispatcher et au composant CBR.

La fonction de consignation binaire permet de stocker les informations du serveur dans des fichiers binaires. Ces fichiers peuvent ensuite être traités pour analyser les informations relatives aux serveurs qui ont été rassemblées.

Les informations suivantes sont stockées dans le journal binaire pour chaque serveur défini dans la configuration.

- Adresse de cluster
- Numéro de port
- ID serveur
- Adresse du serveur
- Pondération du serveur
- Nombre total de connexions sur le serveur
- Connexions actives sur le serveur
- Charge du port du serveur
- Charge système du serveur

Certaines de ces informations sont extraites de l'exécuteur comme faisant partie du cycle du gestionnaire. C'est pourquoi, le gestionnaire doit être en cours d'exécution afin que les informations puissent être consignées dans les journaux binaires.

Utilisez l'ensemble de commandes **dscontrol binlog** pour configurer la consignation binaire.

- binlog start
- binlog stop
- binlog set interval <seconde>
- binlog set retention <heures>
- binlog status

L'option `start` commence à consigner les informations relatives au serveur dans les journaux binaires du répertoire logs. Un journal est créé au début de chaque heure, la date et l'heure constituant le nom du fichier.

L'option `stop` arrête la consignation des informations relatives au serveur dans les journaux binaires. Le service de consignation est arrêté par défaut.

L'option `set interval` contrôle la fréquence d'inscription des informations dans les journaux. Le gestionnaire enverra les informations du serveur au serveur de consignation à chaque intervalle défini pour le gestionnaire. Les informations sont écrites dans les journaux uniquement si l'intervalle de consignation indiqué a expiré depuis l'écriture du dernier enregistrement dans le journal. Par défaut, la valeur de l'intervalle de consignation est 60 secondes. Il y a interaction entre les paramètres relatifs à l'intervalle défini pour le gestionnaire et l'intervalle de consignation. Comme les informations ne sont pas fournies au serveur de consignation plus fréquemment que l'intervalle défini pour le gestionnaire, l'indication d'un intervalle de consignation inférieur à l'intervalle du gestionnaire, entraîne en réalité la définition d'un intervalle de consignation identique à l'intervalle du gestionnaire. Cette technique de consignation permet d'accéder aux informations du serveur quel que soit le niveau de granularité. Vous pouvez connaître toutes les modifications apportées au serveur qui sont vues par le gestionnaire pour le calcul des pondérations du serveur. Cependant, ces informations peuvent ne pas être requises pour analyser l'utilisation et les tendances du serveur. La consignation des informations du serveur toutes les 60 secondes permet d'obtenir un aperçu de la progression des informations du serveur. La définition d'un intervalle de consignation très faible peut générer un nombre de données très important.

L'option `set retention` permet de contrôler la durée de conservation des fichiers journaux. Les journaux dont la durée de vie a dépassé la durée définie sont supprimés par le serveur de consignation. Cela se produit uniquement si le serveur de consignation est appelé par le gestionnaire. Par conséquent, si le gestionnaire est arrêté, les fichiers journaux plus anciens ne sont pas supprimés.

L'option `status` renvoie les paramètres courants de la fonction de consignation, c'est-à-dire l'état actif ou inactif du service, l'intervalle de consignation et la durée de conservation.

Un exemple de programme Java et un fichier de commandes ont été fournis dans le répertoire `...ibm/edge/lb/servers/samples/BinaryLog`. Ce modèle indique comment rappeler toutes les informations contenues dans les fichiers journaux pour les afficher à l'écran. Il peut être personnalisé pour effectuer n'importe quel type d'analyse. Par exemple (à l'aide du script et du programme fournis) :

```
dslogreport 2001/05/01 8:00 2001/05/01 17:00
```

Cet exemple permet d'obtenir un rapport sur les informations du serveur Dispatcher de 8 à 17 heures le premier mai 2001. (Pour CBR, utilisez `cbrlogreport`.)

Utilisation d'un client co-implanté

Seuls les systèmes Linux acceptent des configurations dans lesquelles le client réside sur la même machine que Load Balancer.

Les configurations avec client co-implanté risquent de ne pas fonctionner correctement sur les autres plateformes car Load Balancer utilise des techniques différentes pour examiner les paquets entrants selon les systèmes d'exploitation

pris en charge. La plupart du temps, sur les systèmes autres que Linux, Load Balancer ne reçoit pas de paquets de la machine locale. Il reçoit les paquets provenant uniquement du réseau. C'est la raison pour laquelle les demandes envoyées à l'adresse du cluster à partir de la machine locale ne sont pas reçues par Load Balancer et ne peuvent pas être traitées.

Chapitre 23. Fonctions avancées de Cisco CSS Controller et Nortel Alteon Controller

Le présent chapitre contient les sections suivantes :

- «Co-implantation»
- «Haute disponibilité»
- «Optimisation de la fonction d'équilibrage de charge fournie par Load Balancer», à la page 246
- «Conseillers», à la page 248
- «Système Metric Server», à la page 253
- «Utilisation de la consignment binaire pour analyser les statistiques des serveurs», à la page 255
- «Utilisation de scripts pour la génération d'une alerte ou d'une erreur du serveur d'enregistrement», à la page 257

Remarque : Dans ce chapitre, **xxxcontrol** correspond à **ccocontrol** pour Cisco CSS Controller et à **nalcontrol** pour Nortel Alteon Controller.

Co-implantation

Cisco CSS Controller ou Nortel Alteon Controller peut se trouver sur la même machine qu'un serveur pour lequel vous équilibrez la charge des demande. On parle alors de *co-implantation* d'un serveur. Aucune configuration supplémentaire n'est requise.

Remarque : Un serveur co-implanté est en concurrence avec Load Balancer pour les ressources aux moments de fort trafic. Toutefois, en l'absence de machines surchargées, l'utilisation d'un serveur co-implanté permet de réduire le nombre total de machines nécessaires pour configurer un site avec équilibrage de charge.

Haute disponibilité

La fonction haute disponibilité est désormais disponible pour Cisco CSS Controller et pour Nortel Alteon Controller.

Pour une meilleure tolérance aux pannes du contrôleur, la haute disponibilité intègre les fonctions suivantes :

- Un mécanisme de signal de présence permettant de déterminer la disponibilité des contrôleurs partenaires. Les signaux de présence s'échangent entre les adresses configurées avec la commande **xxxcontrol highavailability add**. Vous pouvez fixer l'intervalle au cours duquel les signaux sont échangés et l'intervalle pendant lequel un contrôleur prend le relais sur son partenaire.
- Une liste des cibles que chaque contrôleur doit être à même de contacter pour calculer les pondérations et mettre le commutateur à jour. Pour plus d'informations, voir «Détection des incidents», à la page 245.
- La logique de choix du contrôleur actif en fonction d'informations de disponibilité et d'accessibilité.
- Une stratégie de relais configurable utilisée pour déterminer comment un contrôleur prend le relais de son partenaire.

- Un mécanisme de relais manuel pour la maintenance des contrôleurs actifs.
- Des rapports qui affichent le rôle, l'état, la synchronisation, etc., en cours du contrôleur.

Configuration

Pour la syntaxe complète de **xxxcontrol highavailability**, voir «ccocontrol highavailability — Contrôle de la haute disponibilité», à la page 438 et «nalcontrol highavailability — Contrôle de la haute disponibilité», à la page 458.

Pour configurer la haute disponibilité du contrôleur, procédez comme suit :

1. Démarrez le serveur du contrôleur sur les deux machines contrôleurs.
2. Configurez chaque contrôleur avec les mêmes paramètres de configuration.
3. Configurez comme suit le rôle, l'adresse et l'adresse du partenaire de haute disponibilité locale :

```
xxxcontrol highavailability add address 10.10.10.10  
partneraddress 10.10.10.20 port 143 role primary
```

4. Configurez comme suit le rôle, l'adresse et l'adresse du partenaire de haute disponibilité du partenaire :

```
xxxcontrol highavailability add address 10.10.10.20  
partneraddress 10.10.10.10 port 143 role secondary
```

Les paramètres address (adresse) et partneraddress (adresse du partenaire) sont inversés entre les machines principale et secondaire.

5. Configurez éventuellement les paramètres de haute disponibilité sur les contrôleurs local et partenaire, comme suit :

```
xxxcontrol highavailability set beatinterval 1000
```

6. Configurez éventuellement les cibles à contacter sur les contrôleurs local et partenaire, comme suit :

```
xxxcontrol highavailability usereach 10.20.20.20
```

Vous devez configurer le même nombre de cibles à contacter sur le contrôleur local et sur le contrôleur partenaire.

7. Démarrez le composant haute disponibilité et définissez une stratégie de récupération sur les contrôleurs local et partenaire, comme suit :

```
xxxcontrol highavailability start auto
```

8. Affichez éventuellement les informations de haute disponibilité sur les contrôleurs local et partenaire, comme suit :

```
xxxcontrol highavailability report
```

9. Définissez éventuellement un relais sur le contrôleur de secours pour que ce dernier prenne le relais du contrôleur actif, comme suit :

```
xxxcontrol highavailability takeover
```

Cette option n'est requise que pour la maintenance.

Remarques :

1. Pour configurer un contrôleur sans haute disponibilité, n'émettez aucune commande de haute disponibilité.
2. Pour convertir deux contrôleurs d'une configuration en haute disponibilité en un unique contrôleur, commencez par arrêter la haute disponibilité sur le contrôleur de secours, puis arrêtez-la éventuellement sur le contrôleur actif.
3. Lorsque vous exécutez deux contrôleurs dans une configuration en haute disponibilité, vous risquez d'obtenir des résultats imprévisibles si les propriétés

des contrôleurs diffèrent d'un commutateur à l'autre ; par exemple, ID du consultant du commutateur, adresse du commutateur, etc. Vous risquez également d'obtenir des résultats imprévisibles si les propriétés de haute disponibilité des contrôleurs ne correspondent pas ; par exemple, port, rôle, cibles à contacter, intervalle entre les signaux de présence, intervalle de prise de relais et stratégie de récupération.

Détection des incidents

Outre la perte de connectivité entre le contrôleur de secours et le contrôleur actif, détectée via les messages de signal de présence, un autre mécanisme de détection d'incidents appelé *critères d'accessibilité* est disponible.

Lorsque vous configurez la haute disponibilité des contrôleurs, vous pouvez indiquer une liste d'hôtes que chaque contrôleur doit pouvoir contacter pour fonctionner correctement. Vous devez choisir au moins un hôte pour chaque sous-réseau que la machine contrôleur utilise. Il peut s'agir de systèmes hôtes tels que des routeurs, des serveurs IP ou d'autres types d'hôtes.

L'accessibilité de l'hôte est obtenue grâce au conseiller de contact, qui lance un ping à l'hôte. Le basculement a lieu si les messages de signal de présence ne peuvent pas être transmis ou si les critères d'accessibilité sont mieux respectés par la machine contrôleur de secours que par la machine contrôleur principale. Pour prendre la décision sur la base de toutes les informations disponibles, le contrôleur actif envoie régulièrement au contrôleur de secours ses données d'accessibilité, et inversement. Les contrôleurs comparent alors leurs informations d'accessibilité avec celles de leur partenaire et décident lequel doit être actif.

Stratégie de récupération

Les deux machines contrôleurs sont configurées avec le rôle principale ou secondaire. Au démarrage, les contrôleurs s'échangent des informations jusqu'à ce que chaque machine soit synchronisée. A ce stade, le contrôleur principal passe à l'état actif et commence à calculer les pondérations et à mettre à jour le commutateur, tandis que la machine secondaire passe à l'état de veille (machine de secours) et surveille la disponibilité de la machine principale.

Si, à tout instant, la machine de secours décèle une défaillance de la machine principale, elle prend le relais des fonctions d'équilibrage de charge de la machine active (défaillante) et devient, à son tour, la machine active. Lorsque la machine principale redevient opérationnelle, les deux machines déterminent quel sera le contrôleur actif en fonction de la stratégie de récupération configurée.

Il existe deux types de stratégie de récupération :

Récupération automatique

Le contrôleur principal passe à l'état actif (calcule et met à jour les pondérations) dès qu'il redevient opérationnel. La machine secondaire se met en veille dès que la principale est active.

Récupération manuelle

Le contrôleur secondaire actif reste actif même lorsque le contrôleur principal est redevenu opérationnel.

Le contrôleur principal passe à l'état de veille et requiert une intervention manuelle pour revenir à l'état actif.

La stratégie définie doit être identique pour les deux machines.

Exemples

Pour des exemple de configuration haute disponibilité Cisco CSS Controller, voir «Exemples», à la page 440.

Pour des exemple de configuration haute disponibilité Nortel Alteon Controller, voir «Exemples», à la page 460.

Optimisation de la fonction d'équilibrage de charge fournie par Load Balancer

La fonction contrôleur de Load Balancer effectue l'équilibrage de charge en fonction des paramètres suivants :

- «Importance accordée aux informations de mesure»
- «Pondérations», à la page 247
- «Délai d'inactivité dans le calcul des pondérations», à la page 247
- «Délai d'inactivité du conseiller», à la page 249
- «Seuil de sensibilité», à la page 248

Tous ces paramètres peuvent être modifiés en vue d'optimiser l'équilibrage de la charge du réseau.

Importance accordée aux informations de mesure

Le contrôleur peut utiliser certains ou l'ensembles de collecteurs de mesures suivants pour les décisions de pondération :

- *Connexions actives* : Nombre de connexions actives sur chaque serveur d'équilibrage de charge, extraites du commutateur.
- *Débit de la connexion* : Nombre de nouvelles connexions sur chaque serveur d'équilibrage de charge depuis la dernière demande, extraites du commutateur.
- *Unité centrale* : Pourcentage de l'unité centrale utilisé sur chaque serveur d'équilibrage de charge (entré à partir de l'agent Metric Server).
- *Mémoire* : Pourcentage de mémoire utilisée (entré à partir de l'agent Metric Server) sur chaque serveur d'équilibrage de charge.
- *Mesure système* : Entrées provenant des outils de contrôle système, tels que Metric Server ou WLM.
- *Spécifique à l'application* : Entrée effectuée par les conseillers écoutant au niveau de ce port.

Les mesures par défaut sont activeconn (connexions actives) et connrate (débit de la connexion).

Vous pouvez modifier le niveau d'importance relatif des valeurs de mesure. Les proportions sont des pourcentages ; la somme des proportions est égale à 100%. Les valeurs relatives aux connexions actives et au débit de la connexion sont utilisées par défaut et leurs proportions fixées à 50/50. Dans votre environnement, vous serez amené à essayer différentes combinaisons de proportions pour déterminer celui qui offre les meilleures performances.

Pour définir les valeurs de niveau d'importance, procédez comme suit :

Pour Cisco CSS Controller

cococontrol ownercontent metrics *NomMetric1 NiveauImportance1*
NomMetric2 NiveauImportance2

Pour Nortel Alteon Controller

nalcontrol service metrics *NomMetric1 NiveauImportance1 NomMetric2*
NiveauImportance2

Pondérations

Les pondérations sont définies en fonction du temps de réponse et de la disponibilité de l'application, du retour d'informations des conseillers et du retour d'informations procuré par un programme de contrôle système, tel que Metric Server. Si vous voulez définir des pondérations manuellement, indiquez l'option **fixedweight** pour le serveur. Pour obtenir une description de l'option **fixedweight**, voir «Pondérations fixées par le contrôleur».

Les pondérations définies s'appliquent à tous les serveurs fournissant un service. Pour chaque service, les demandes sont réparties entre les serveurs selon la pondération relative de chacun. Par exemple, si un serveur a une pondération (paramètre **Weight**) de 10 et un autre de 5, le premier recevra deux fois plus de demandes que le second.

Si un conseiller détecte la défaillance d'un serveur, il attribue au serveur une pondération de -1. Pour Cisco CSS Controller et Nortel Alteon Controller le commutateur est informé que le serveur n'est pas disponible et le commutateur n'affecte plus de connexions au serveur.

Pondérations fixées par le contrôleur

Sans le contrôleur, les conseillers ne peuvent s'exécuter ni détecter les pannes de serveur. Si vous choisissez de lancer les conseillers mais ne voulez *pas* que le contrôleur mette à jour la pondération que vous avez fixée pour un serveur particulier, utilisez l'option **fixedweight** de la commande **cococontrol service** pour Cisco CSS Controller ou de la commande **nalcontrol server** pour Nortel Alteon Controller.

La commande **fixedweight** vous permet d'affecter à la pondération la valeur souhaitée. La valeur de pondération du serveur reste fixe tant que le contrôleur est en activité à moins que vous n'émettiez une autre commande en attribuant la valeur **no** à l'option **fixedweight**.

Délai d'inactivité dans le calcul des pondérations

Pour optimiser les performances globales, vous pouvez réduire la fréquence de collecte des mesures.

Le délai d'inactivité du consultant indique à quelle fréquence le consultant met à jour les pondérations des serveurs. Si le délai d'inactivité du consultant est trop court, le consultant interrompra constamment le commutateur et les performances déclineront. En revanche, s'il est trop long, l'équilibrage de charge du commutateur reposera sur des informations anciennes et incertaines.

Pour fixer le délai d'inactivité du consultant à 1 seconde, par exemple, entrez la commande suivante :

xxxcontrol consultant set *IDconsultant sleeptime*
intervalle

Seuil de sensibilité

D'autres méthodes d'optimisation de l'équilibrage de charge des serveurs sont disponibles. Pour fonctionner en vitesse maximale, les pondérations des serveurs ne sont actualisées que si les pondérations ont évolué de manière significative. La mise à jour constante des pondérations pour un écart mineur de l'état des serveurs induirait un surcroît d'activité injustifié. Lorsque, pour tous les serveurs d'un service donné, l'écart en pourcentage de la pondération totale dépasse le seuil de sensibilité, les pondérations qu'utilise Load Balancer pour répartir les connexions sont réactualisées. Supposons par exemple que la pondération totale passe de 100 à 105. L'écart est de 5%. Avec un seuil de sensibilité par défaut de 5, les pondérations qu'utilise Load Balancer ne sont pas mises à jour, car l'écart en pourcentage n'est pas **supérieur** au seuil. Si, en revanche la pondération totale passe de 100 à 106, les pondérations sont mises à jour. Pour attribuer au seuil de sensibilité du consultant une valeur autre que la valeur par défaut, entrez la commande suivante :

```
xxxcontrol consultant set IDconsultant sensitivity pourcentageChangement
```

Dans la plupart des cas, vous n'aurez pas besoin de modifier cette valeur.

Conseillers

Les conseillers sont des agents de Load Balancer. Ils ont pour rôle d'évaluer l'état et la charge des serveurs. Ils effectuent cette tâche via un échange proactif de type client/serveur. Considérez les conseillers comme des clients des serveurs d'application.

Remarque : Pour connaître la liste complète des conseillers, voir «Liste des conseillers», à la page 188.

Fonctionnement des conseillers

Les conseillers ouvrent régulièrement une connexion TCP avec chaque serveur et envoient un message de demande au serveur. Le contenu du message dépend du protocole exécuté sur le serveur. Par exemple, le conseiller HTTP envoie une demande HTTP «HEAD» au serveur.

Les conseillers attendent ensuite une réponse du serveur. Une fois la réponse obtenue, le conseiller évalue l'état du serveur. Pour calculer la valeur de la *charge*, la plupart des conseillers mesurent le délai de réponse du serveur, puis ils utilisent cette valeur (en millisecondes) comme valeur de charge.

Le conseiller reporte cette valeur au consultant. Elle apparaît dans le rapport du consultant. Le consultant calcule ensuite un ensemble de valeurs de pondération à partir de toutes ses sources, selon les proportions, puis transmet ces valeurs de pondération au commutateur. Le commutateur utilise ces pondérations pour équilibrer la charge des nouvelles connexions client entrantes.

Si le conseiller détermine que le serveur est actif et que son état est correct, il renvoie au consultant une valeur de charge positive non nulle. Si le conseiller détermine que le serveur n'est pas actif, il renvoie une valeur de charge spéciale négative (-1) pour informer le commutateur de l'inactivité du serveur. Le commutateur n'enverra donc plus aucune connexion en direction de ce serveur tant qu'il ne sera pas de nouveau actif.

Délai d'inactivité du conseiller

Remarque : Les valeurs par défaut du conseiller sont correctes pour la plupart des scénarios possibles. Soyez prudent lorsque vous entrez des valeurs autres que celles fournies par défaut.

Le délai d'inactivité du conseiller détermine la fréquence à laquelle un conseiller demande des données d'état aux serveurs associés au port dont il a la charge, puis transmet ces données au consultant. Si le délai d'inactivité du conseiller est trop court, le conseiller interrompra les serveurs constamment et les performances déclineront. En revanche, s'il est trop long, l'équilibrage de charge du consultant reposera sur des informations anciennes et incertaines.

Par exemple, pour fixer à 3 secondes l'intervalle du conseiller HTTP, entrez la commande suivante :

```
xxxcontrol metriccollector set IDconsultant:HTTP sleeptime 3
```

Délai de connexion du conseiller et délai de réception pour les serveurs

Vous pouvez définir le temps nécessaire à un conseiller pour détecter qu'un port particulier d'un serveur ou d'un service est défaillant. Les valeurs de délai d'erreur serveur (connecttimeout et receivetimeout) déterminent la durée attendue par un conseiller avant de signaler qu'une connexion ou une réception n'a pas abouti.

Pour obtenir une détection d'erreur serveur très rapide, attribuez la valeur la plus basse (une seconde) aux délais de connexion et de réception du conseiller et attribuez la valeur la plus basse (une seconde) au délai d'inactivité du conseiller et du consultant.

Remarque : Si le trafic dans votre environnement va de modéré à fort et que le temps de réponse du serveur augmente, n'attribuez pas une valeur trop faible à timeoutconnect et timeoutreceive. Sinon, le conseiller risque de déclarer prématurément un serveur occupé comme étant en erreur.

Pour attribuer, par exemple, la valeur 9 secondes à timeoutconnect pour le conseiller HTTP, entrez la commande suivante :

```
xxxcontrol metriccollector set IDconsultant:HTTP  
timeoutconnect 9
```

La valeur par défaut de connexion et de réception est trois fois supérieure à la valeur indiquée pour le délai d'inactivité du conseiller.

Tentative du conseiller

Les conseillers peuvent essayer de nouveau d'établir une connexion avant de marquer un serveur comme arrêté. Le serveur ne signale un serveur comme étant arrêté qu'après avoir effectué le nombre de tentatives de connexion fixé plus une. Si ce nombre n'a pas été défini, il est par défaut égal à zéro.

Pour le contrôleur CSS Cisco, fixez le nombre de **tentatives** à l'aide de la commande **ccocontrol ownercontent set**. Pour plus de détails, voir «ccocontrol ownercontent — Contrôle du nom de propriétaire et de la règle de contenu», à la page 443.

Pour le contrôleur Nortel Alteon, fixez le nombre de **tentatives** à l'aide de la commande **nalcontrol service set**. Pour plus de détails, voir «nalcontrol service — Configuration d'un service», à la page 465.

Création de conseillers personnalisés

Remarque : Dans cette section, **serveur** est le terme générique qui désigne un service pour Cisco CSS Controller ou un serveur pour Nortel Alteon Controller.

Le conseiller personnalisé est un petit programme Java, que vous fournissez sous forme de fichier classe, appelé par le code de base. Le code de base fournit tous les services d'administration, tels que :

- Démarrage et arrêt d'une instance du conseiller personnalisé
- Génération d'états et de rapports
- Enregistrement des informations d'historique dans un fichier journal

Il renvoie également les résultats au consultant. Régulièrement, le code de base lance un cycle de conseiller au cours duquel il évalue individuellement tous les serveurs de sa configuration. Il commence par ouvrir une connexion avec la machine serveur. Si la connexion s'ouvre, le code de base appelle la méthode (fonction) `getLoad` dans le conseiller personnalisé. Ce dernier effectue la procédure nécessaire à l'évaluation du serveur. Généralement, il envoie au serveur un message défini par l'utilisateur, puis attend une réponse. L'accès à la connexion ouverte est fourni au conseiller personnalisé. Le code de base ferme ensuite la connexion au serveur et envoie au consultant les informations relatives à la charge.

Le code de base et le conseiller personnalisé peuvent opérer en mode normal ou en mode replace. Le choix du mode de fonctionnement est indiqué dans le fichier du conseiller personnalisé en tant que paramètre dans la méthode du constructeur.

En mode normal, le conseiller personnalisé échange des données avec le serveur et le code du conseiller de base évalue la durée de l'échange et calcule la valeur de la charge. Le code de base renvoie cette valeur au consultant. Le conseiller personnalisé doit simplement retourner un zéro (succès) ou une valeur négative (échec). Lorsque dans le fichier du constructeur, la valeur `false` est attribuée à l'indicateur `replace`, le mode normal est défini.

En mode `replace`, le code de base n'effectue aucune mesure de temps. Le code du conseiller personnalisé effectue toutes les opérations nécessaires, puis renvoie une valeur de charge. Le code de base accepte la valeur et la retourne au consultant. Pour obtenir de meilleurs résultats, situez votre valeur de charge entre 10 et 1000, 10 représentant un serveur rapide et 1000 un serveur plus lent. Lorsque dans le fichier du constructeur, la valeur `true` est attribuée à l'indicateur `replace`, le mode `replace` est défini.

Avec cette fonctionnalité, vous pouvez développer vos propres conseillers pour fournir les informations sur les serveurs dont vous avez besoin. Un exemple de conseiller personnalisé, **ADV_ctlrsample.java**, est fourni pour les contrôleurs. Une fois Load Balancer installé, le code exemple se trouve dans le répertoire d'installation **.../ibm/edge/lb/servers/samples/CustomAdvisors** .

Les répertoires d'installation par défaut sont :

- Systèmes AIX, HP-UX, Linux, Solaris : `/opt/ibm/edge/lb`

- Systèmes Windows : C:\Program Files\IBM\ibm\edge\lb

Remarque : Si vous ajoutez un conseiller personnalisé à Cisco CSS Controller ou à Nortel Alteon Controller, vous devez arrêter, puis redémarrer **ccoserver** ou **nalserver** (pour les systèmes Windows, utilisez Services) pour que le processus Java puisse lire les nouveaux fichiers de classes du conseiller personnalisé. Les fichiers de classes du conseiller personnalisé ne sont chargés qu'au démarrage.

Convention d'attribution de nom

Le nom de fichier de votre conseiller personnalisé doit être au format *ADV_monconseiller.java*. Il doit être précédé du préfixe **ADV_** en majuscules. Tous les caractères suivants doivent être en minuscules.

Conformément aux conventions Java, le nom de la classe définie dans le fichier doit correspondre au nom du fichier. Si vous copiez le code exemple, veuillez à remplacer toutes les occurrences de **ADV_ctrlsample** dans le fichier par le nom de votre nouvelle classe.

Compilation

Les conseillers personnalisés sont écrits en langage Java. Utilisez le compilateur Java qui est installé avec Load Balancer. Les fichiers suivants sont référencés pendant la compilation :

- le fichier du conseiller personnalisé,
- le fichier de classes de base, **ibmlb.jar**, qui se trouve dans le répertoire d'installation **...ibm/edge/lb/servers/lib**.

Le chemin d'accès aux classes doit désigner à la fois le fichier du conseiller personnalisé et le fichier de classes de base lors de la compilation.

Pour Windows, une commande de compilation peut avoir l'aspect suivant :

```
rep_install/java/bin/javac -classpath  
rep_install\lb\servers\lib\ibmlb.jar ADV_pam.java
```

où :

- Votre fichier conseiller s'appelle **ADV_pam.java**.
- Votre fichier conseiller se trouve dans le répertoire courant.

Le résultat de la compilation est un fichier **.class**, par exemple :

ADV_pam.class

Avant de lancer le conseiller, copiez le fichier **.class** dans le répertoire d'installation **...ibm/edge/lb/servers/lib/CustomAdvisors**.

Remarque : Si vous le souhaitez, vous pouvez compiler les conseillers personnalisés sur un système d'exploitation et l'exécuter sur un autre. Par exemple, vous pouvez compiler le conseiller sur des systèmes Windows, copier le fichier **.class** (en binaire) sur une machine AIX à partir de laquelle vous exécutez le conseiller personnalisé.

Pour les systèmes AIX, HP-UX, Linux et Solaris, la syntaxe est similaire.

Exécution

Pour exécuter le conseiller personnalisé, vous devez tout d'abord copier le fichier .class dans le répertoire d'installation approprié :

```
...ibm/edge/lb/servers/lib/CustomAdvisors/ADV_pam.class
```

Démarrez le consultant, puis entrez la commande suivante pour démarrer le conseiller personnalisé :

Pour Cisco CSS Controller

```
cococontrol ownercontent metrics IDconsultant:IDcontenupropriétaire pam 100
```

Pour Nortel Alteon Controller

```
nalcontrol service metrics IDconsultant:IDservice pam 100
```

où :

- pam est le nom de votre conseiller, comme dans ADV_pam.java,
- 100 est le niveau d'importance de pondération accordée à ce conseiller

Sous-programmes requis

Comme tous les conseillers, un conseiller personnalisé étend la fonction de la base du conseiller, intitulée ADV_Base. En fait, c'est la base du conseiller qui effectue la plupart des fonctions du conseiller, telles que la communication des charges au consultant afin que ces dernières soient utilisées dans l'algorithme de pondération du consultant. La base du conseiller effectue également les opérations de connexion et de fermeture de la connexion et fournit des méthodes d'envoi et de réception qui seront utilisées par le conseiller. Le conseiller n'est lui-même utilisé que pour l'envoi de données vers le port du serveur conseillé et pour la réception de données sur ce dernier. Les méthodes TCP de la base du conseiller sont programmées pour calculer la charge. Un indicateur du constructeur de ADV_base remplace, si vous le souhaitez, la charge existante par la nouvelle charge renvoyée par le conseiller.

Remarque : En fonction d'une valeur définie dans le constructeur, la base du conseiller fournit la charge à l'algorithme de pondération à un intervalle donné. Si le véritable conseiller n'a pas terminé ses opérations afin de renvoyer une charge valide, la base du conseiller utilise la charge précédente.

Ci-dessous, sont énumérées les méthodes de classe de base.

- Sous-programme **constructeur**. Le constructeur appelle le constructeur de la classe de base (reportez-vous au fichier type de conseiller).
- Méthode **ADV_AdvisorInitialize**. Cette méthode fournit un point d'ancrage au cas où des procédures supplémentaires doivent être suivies une fois l'initialisation de la classe de base terminée.
- Sous-programme **getLoad**. La classe de base du conseiller se charge de l'ouverture de la connexion ; getLoad ne doit qu'émettre les demandes d'envoi et de réception appropriées pour terminer le cycle de conseil.

Ordre de recherche

Les contrôleurs consultent d'abord la liste fournie de conseillers natifs. S'ils n'y trouvent pas le conseiller recherché, ils consultent la liste des conseillers personnalisés.

Affectation du nom et du chemin

- La classe de conseiller personnalisé doit se trouver dans le sous-répertoire de **...ibm/edge/lb/servers/lib/CustomAdvisors/** dans le répertoire de base de Load Balancer. Le répertoire par défaut dépend du système d'exploitation :
 - Systèmes AIX, HP-UX, Linux ou Solaris
`/opt/ibm/edge/lb/servers/lib/CustomAdvisors/`
 - Systèmes Windows
`C:\Program Files\IBM\edge\lb\servers\lib\CustomAdvisors`
- Seuls les caractères alphabétiques minuscules sont autorisés. Cela permet d'éliminer la distinction entre majuscules et minuscules lorsqu'un opérateur entre des commandes sur la ligne de commande. Le nom de fichier du conseiller doit être précédé de **ADV_**.

Conseiller type

Un programme permettant de créer un conseiller de contrôleur type est présenté à la section «Conseiller type», à la page 487. Après installation, ce conseiller exemple se trouve dans le répertoire **...ibm/edge/lb/servers/samples/CustomAdvisors**.

Système Metric Server

Metric Server fournit à Load Balancer les informations de téléchargement sous la forme de données numériques-système, relatives à l'état du serveur. Le consultant Load Balancer adresse des demandes aux agents du système Metric Server situés sur chacun des serveurs, leur attribuant des pondérations destinées au processus d'équilibrage de charge à l'aide des données rassemblées par les agents. Les résultats sont regroupés dans le rapport du service pour Cisco CSS Controller ou le rapport du serveur pour Nortel Alteon Controller.

Conditions préalables

L'agent Metric Server doit être installé et en cours d'exécution sur tous les serveurs dont la charge est équilibrée.

Conditions d'utilisation de Metric Server

La procédure ci-après permet de configurer Metric Server pour les contrôleurs.

- Côté contrôleur
 1. Démarrez **ccoserver** ou **nalserver**.
 2. Pour Cisco CSS Controller, ajoutez un consultant de commutateur, puis un contenu de propriétaire.
Pour Nortel Alteon Controller, ajoutez un consultant de commutateur, puis un service.
 3. Indiquez le port sur lequel l'agent du serveur de mesures écoute. Cette information doit correspondre à celle du fichier `metricserver.cmd`. Le port par défaut est 10004. Entrez la commande suivante :

Pour Cisco CSS Controller

```
ccocontrol service set IDconsultant:IDcontenupropriétaire:IDserveur  
metricserverport numéroport
```

Pour Nortel Alteon Controller

```
nalcontrol server set IDconsultant:IDservice:IDserveur metricserverport  
numéroport
```

4. Entrez la commande de mesures système suivante :

Pour Cisco CSS Controller

cococontrol ownercontent metrics IDconsultant:IDcontenupropriétaire
NomMetric importance

Pour Nortel Alteon Controller

nalcontrol service metrics IDconsultant:IDservice nomMetric importance

où *NomMetric* est le nom du script System Metric.

Le script System Metric se trouve sur le serveur dorsal et s'exécute sur chacun des serveurs de la configuration sous le contenu de propriétaire ou le service indiqué. Deux scripts, **cpuload** et **memload** sont fournis, mais vous pouvez créer vos propres scripts System Metric personnalisés. Le script contient une commande de renvoi d'une valeur numérique. Cette valeur numérique représente une mesure de charge, et non un indicateur de disponibilité.

Restriction : Pour les systèmes Windows, si le nom du script system metric comporte une extension autre que ".exe", vous devez indiquer le nom complet du fichier (par exemple, monScriptSystemMetric.bat). Il s'agit d'une limitation du code Java.

5. Emettez la commande pour votre contrôleur, comme suit :

Pour Cisco CSS Controller

cococontrol consultant start

Pour Nortel Alteon Controller

nalcontrol consultant start

Remarque : Garantie de la sécurité —

- Sur la machine du contrôleur, créez des fichiers de clés à l'aide de la commande **lbkeys create #** suivant comme commande de configuration de cluster. Pour plus d'informations sur lbkeys, voir «RMI (Remote Method Invocation)», à la page 262.
 - Sur le serveur, copiez le fichier de clés obtenu dans le répertoire **...ibm/edge/lb/admin/key**. Vérifiez que le superutilisateur dispose de droits lui permettant de lire le fichier de clés.
- Agent Metric Server (côté serveur)
 1. Lors de l'installation de Load Balancer, installez l'ensemble Metric Server.
 2. Vérifiez le script **metricserver** dans le répertoire **/usr/bin** afin de contrôler que le port RMI souhaité est utilisé. (Pour les systèmes Windows, le répertoire est C:\WINNT\SYSTEM32.) Le port RMI par défaut est 10004.

Remarque : La valeur du port RMI indiquée doit être identique à la valeur du port RMI du système Metric Server sur le contrôleur.

3. Les deux scripts suivants sont fournis : **cpuload** (renvoie le pourcentage de cpu utilisé, compris entre 0 et 100) et **memload** (donne le pourcentage de mémoire utilisée compris entre 0 et 100). Ces scripts se trouvent dans le répertoire **...ibm/edge/lb/ms/script**.

Vous pouvez éventuellement écrire vos propres fichiers scripts personnalisés qui définiront la commande passée par Metric Server sur les serveurs. Vérifiez que tous les scripts personnalisés sont exécutables et se trouvent dans le répertoire **...ibm/edge/lb/ms/script**. Les scripts personnalisés **doivent** renvoyer une valeur de charge.

Remarque : Un script de mesure personnalisé doit être un programme valide ou un script ayant l'extension .bat ou .cmd. De manière plus

spécifique, pour les systèmes Linux et UNIX, les scripts doivent commencer par la déclaration de shell, sinon ils risquent de ne pas s'exécuter correctement.

4. Démarrez l'agent en émettant la commande **metricserver**.
5. Pour arrêter l'agent Metric Server, tapez **metricserver stop**.

Pour exécuter le système Metric Server ailleurs que sur l'hôte local, vous devez modifier le fichier `metricserver` sur le serveur ayant fait l'objet d'un équilibrage de charge. Insérez la ligne suivante après **java** dans le fichier `metricserver` :

```
-Djava.rmi.server.hostname=AUTRE_ADRESSE
```

Ajoutez en outre la ligne suivante avant les instructions "if" dans le fichier `metricserver` : `hostname AUTRE_ADRESSE`.

Pour les systèmes Windows : Affectez un alias à `AUTRE_ADRESSE` dans la pile Microsoft. Pour ce faire, voir à la page 211.

Conseiller Workload manager

Le code de WLM ne s'exécute que sur des grands systèmes MVS. Il peut être utilisé pour demander la charge sur la machine MVS.

Si MVS Workload Management a été configuré sur votre système OS/390, les contrôleurs peuvent accepter de WLM des informations relatives à la charge et les utiliser dans le processus d'équilibrage de charge. Grâce au conseiller WLM, les contrôleurs ouvrent régulièrement des connexions via le port WLM sur chaque serveur de la table d'hôte consultant et acceptent les chiffres relatifs à la capacité renvoyés. Comme ces chiffres représentent la capacité encore disponible et que le consultant attend des valeurs représentant la charge sur chaque machine, le conseiller inverse et normalise les chiffres relatifs à la capacité pour obtenir des valeurs de charge (ainsi, des chiffres de capacité élevés correspondent à des valeurs de charge faibles et représentent un serveur en bon état). Il existe plusieurs différences importantes entre le conseiller WLM et les autres conseillers contrôleur :

1. Les autres conseillers ouvrent des connexions aux serveurs en utilisant le même port que pour le trafic client normal. Le conseiller WLM ouvre des connexions aux serveurs en utilisant un port différent de celui utilisé pour le trafic normal. Sur chaque machine serveur, l'agent WLM doit être configuré pour effectuer l'écoute sur le port sur lequel le conseiller de contrôleur WLM a été lancé. Le port WLM par défaut est 10007.
2. Il est possible d'utiliser les conseillers de protocole avec le conseiller WLM. Les conseillers de protocole évaluent la charge des serveurs sur le port utilisé pour le trafic normal et le conseiller WLM évalue la charge du système sur le port WLM.

Utilisation de la consignment binaire pour analyser les statistiques des serveurs

La fonction de consignment binaire permet de stocker les informations du serveur dans des fichiers binaires. Ces fichiers peuvent ensuite être traités pour analyser les informations relatives aux serveurs qui ont été rassemblées.

Les informations suivantes sont stockées dans le journal binaire pour chaque serveur défini dans la configuration.

- Parent (IDcontenupropriétaire pour Cisco CSS Controller; IDservice pour Nortel Alteon Controller)
- ID du serveur
- Adresse du serveur
- Port du serveur
- Pondération du serveur
- Nombre de mesures configurées pour ce serveur
- Liste des valeurs de mesure

Le consultant doit s'exécuter pour consigner des informations dans les journaux binaires.

Utilisez l'ensemble de commandes **xxxcontrol consultant binarylog** pour configurer la consignation binaire.

- `binarylog start`
- `binarylog stop`
- `binarylog report`
- `binarylog set interval <secondes>`
- `binarylog set retention <heures>`

L'option `start` commence à consigner les informations relatives au serveur dans les journaux binaires du répertoire `logs`. Un journal est créé au début de chaque heure, la date et l'heure constituant le nom du fichier.

L'option `stop` arrête la consignation des informations relatives au serveur dans les journaux binaires. Le service de consignation est arrêté par défaut.

L'option `set interval` contrôle la fréquence d'inscription des informations dans les journaux. Le consultant enverra les informations du serveur au serveur de consignation à chaque intervalle défini pour le consultant. Les informations sont écrites dans les journaux uniquement si l'intervalle de consignation indiqué a expiré depuis l'écriture du dernier enregistrement dans le journal. Par défaut, la valeur de l'intervalle de consignation est 60 secondes.

Il y a interaction entre les paramètres relatifs à l'intervalle défini pour le consultant et l'intervalle de consignation. Comme les informations ne sont pas fournies au serveur de consignation plus fréquemment que l'intervalle défini pour le consultant, l'indication d'un intervalle de consignation inférieur à l'intervalle du consultant, entraîne en réalité la définition d'un intervalle de consignation identique à l'intervalle du consultant.

Cette technique de consignation permet d'accéder aux informations du serveur quel que soit le niveau de granularité. Vous pouvez connaître toutes les modifications apportées au serveur qui sont vues par le consultant pour le calcul des pondérations du serveur. Cependant, ces informations peuvent ne pas être requises pour analyser l'utilisation et les tendances du serveur. La consignation des informations du serveur toutes les 60 secondes permet d'obtenir un aperçu de la progression des informations du serveur. La définition d'un intervalle de consignation très faible peut générer un nombre de données très important.

L'option `set retention` permet de contrôler la durée de conservation des fichiers journaux. Les journaux dont la durée de vie a dépassé la durée définie sont supprimés par le serveur de consignation. Ceci ne se produit que si le consultant

appelle le serveur de consignation, de sorte que si vous arrêtez le consultant, les anciens fichiers journaux ne sont pas supprimés.

Un exemple de programme Java et un fichier de commandes sont fournis dans le répertoire **...ibm/edge/lb/servers/samples/BinaryLog**. Ce modèle indique comment rappeler toutes les informations contenues dans les fichiers journaux pour les afficher à l'écran. Il peut être personnalisé pour effectuer n'importe quel type d'analyse.

Par exemple (à l'aide du script et du programme fournis) :

```
xxxlogreport  
2002/05/01 8:00 2002/05/01 17:00
```

Cet exemple permet d'obtenir un rapport sur les informations du serveur du contrôleur de 8 à 17 heures le premier mai 2002.

Utilisation de scripts pour la génération d'une alerte ou d'une erreur du serveur d'enregistrement

Load Balancer fournit des exits utilisateur qui déclenchent des scripts que vous pouvez personnaliser. Vous pouvez créer des scripts afin d'effectuer des actions automatisées. Il est, par exemple, possible de prévenir un administrateur lorsqu'un serveur est inactif ou simplement d'enregistrer l'erreur. Le répertoire d'installation, **...ibm/edge/lb/servers/samples**, contient des exemples de script que vous pouvez personnaliser. Pour exécuter les fichiers, copiez-les dans le répertoire **...ibm/edge/lb/servers/bin**, puis renommez chaque fichier en fonction des directions indiquées dans le script.

Les exemples de scripts suivants, dans lesquels **xxx** est **cco** pour Cisco CSS Controller, et **nal** pour Nortel Alteon Controller sont fournis :

- **xxxserverdown** — le contrôleur indique qu'un serveur est inactif.
- **xxxserverUp** — le contrôleur indique qu'un serveur est à nouveau actif.
- **xxxallserversdown** — tous les serveurs d'un service particulier sont signalés inactifs.

Partie 8. Administration et identification des incidents de Load Balancer

Cette section contient des informations relatives à l'administration et à l'identification des incidents liés à Load Balancer. Elle se compose des chapitres suivants :

- Chapitre 24, «Exploitation et gestion de Load Balancer», à la page 261
- Chapitre 25, «Résolution des incidents», à la page 281

Chapitre 24. Exploitation et gestion de Load Balancer

Remarque : Lors de la lecture de ce chapitre, dans les sections générales qui ne concernent pas particulièrement un composant, si vous n'utilisez *pas* le composant Dispatcher, remplacez "dscontrol" et "dsserver" par les éléments suivants :

- Pour CBR, utilisez **cbrcontrol** et **cbrserver**
- Pour Site Selector, utilisez **sscontrol** et **ssserver**
- Pour Cisco CSS Controller, utilisez **ccocontrol** et **ccoserver**
- Pour Nortel Alteon Controller, utilisez **nalcontrol** et **nalserver**

IMPORTANT : Si vous utilisez l'installation Load Balancer pour IPv4 et IPv6, voir Chapitre 8, «Déploiement de Dispatcher sur Load Balancer pour IPv4 et IPv6», à la page 81 pour connaître les limitations et les différences de configuration avant d'afficher le contenu du présent chapitre.

Le présent chapitre explique comment exploiter et gérer Load Balancer et inclut les sections suivantes :

- «Administration à distance de Load Balancer»
 - «RMI (Remote Method Invocation)», à la page 262
 - «administration basée sur le Web», à la page 263
- «Utilisation des journaux Load Balancer», à la page 265
 - «Pour Dispatcher, CBR et Site Selector», à la page 265
 - «Pour Cisco CSS Controller et Nortel Alteon Controller», à la page 267
- «Utilisation du composant Dispatcher», à la page 268
 - «Utilisation du protocole SNMP (Simple Network Management Protocol, protocole simplifié de gestion de réseau) avec le composant Dispatcher», à la page 269
- «Utilisation du composant CBR (Content Based Routing)», à la page 276
- «Utilisation du composant Site Selector», à la page 277
- «Utilisation du composant Cisco CSS Controller», à la page 278
- «Utilisation du composant Nortel Alteon Controller», à la page 278

Administration à distance de Load Balancer

Load Balancer offre deux manières différentes d'exécuter ses programmes de configuration sur une machine autre que celle sur laquelle se trouve Load Balancer. La communication entre les programmes de configuration (dscontrol, cbrcontrol, sscontrol, ccocontrol, nalcontrol) et le serveur (dsserver, cbrserver, etc.) peut s'établir de l'une des manières suivantes :

- Java RMI (Remote Method Invocation)
- administration basée sur le Web

L'administration à distance via RMI est plus rapide que l'administration basée sur le Web.

L'administration basée sur le Web, outre qu'elle s'effectue à distance, présente l'avantage d'être une méthode sécurisée et authentifiée, capable de communiquer avec la machine Load Balancer même en présence d'un pare-feu. De plus, cette

méthode d'administration *ne requiert pas* d'installation particulière et utilise des clés d'authentification (lbkeys) sur la machine client éloignée qui communique avec la machine Load Balancer.

RMI (Remote Method Invocation)

Pour RMI, la commande permettant de connecter une machine Load Balancer pour l'administration à distance est **dscontrol host:hôte_éloigné**.

Si l'appel RMI vient d'une machine autre que la machine locale, une séquence d'authentification clé publique/clé privée se produit avant l'acceptation de la commande de configuration. une séquence d'authentification doit se produire avant que la commande configuration soit acceptée.

Les communications entre les programmes de contrôle exécutés sur la même machine que les serveurs du composant ne sont pas authentifiées.

La commande suivante permet de générer des clés publiques et privées à utiliser pour l'authentification à distance :

lbkeys [create | delete]

Cette commande s'exécute uniquement sur la même machine que Load Balancer.

L'option **create** permet de créer une clé privée dans le répertoire key des serveurs (...**ibm/edge/lb/servers/key/**) et de créer des clés publiques dans le répertoire keys d'administration (...**ibm/edge/lb/admin/keys/**) pour chacun des composants Load Balancer. Le nom de fichier de la clé publique est : *composant-AdresseServeur-PortRMI*. Ces clés publiques doivent ensuite être transmises aux clients éloignés et placés dans le répertoire keys d'administration.

Pour une machine Load Balancer avec une adresse de nom d'hôte 10.0.0.25 utilisant le port RMI par défaut pour chaque composant, la commande **lbkeys create** génère les fichiers suivants :

- La clé privée : ...**ibm/edge/lb/servers/key/authorization.key**
- Les clés publiques :
 - ...**ibm/edge/lb/admin/keys/dispatcher-10.0.0.25-10099.key**
 - ...**ibm/edge/lb/admin/keys/cbr-10.0.0.25-11099.key**
 - ...**ibm/edge/lb/admin/keys/ss-10.0.0.25-12099.key**
 - ...**ibm/edge/lb/admin/keys/cco-10.0.0.25-13099.key**
 - ...**ibm/edge/lb/admin/keys/na1-10.0.0.25-14099.key**

Le jeu de fichiers d'administration a été installé sur une autre machine. Les fichiers de clés publiques doivent être placés dans le répertoire ...**ibm/edge/lb/admin/keys** sur la machine client éloignée.

Le client éloigné sera désormais autorisé à configurer Load Balancer sur 10.0.0.25.

Ces mêmes clés doivent être utilisées sur tous les clients éloignés que vous souhaitez autoriser à configurer Load Balancer sur 10.0.0.25.

Si vous devez de nouveau exécuter la commande **lbkeys create**, un nouveau jeu de clés publiques/privées sera généré. Ceci signifie que tous les clients éloignés qui

ont tenté de se connecter à l'aide des clés précédentes ne seront plus autorisés. La nouvelle clé doit être placée dans le répertoire adéquat sur les clients auxquels vous voulez attribuer des autorisations.

La commande **lbkeys delete** permet de supprimer les clés privées et publiques sur la machine serveur. Si ces clés sont supprimées, aucun client éloigné ne sera autorisé à se connecter aux serveurs.

L'option **force** peut être associée à la commande lbkeys et à la commande lbkeys delete. Elle permet de supprimer les invites demandant si vous voulez remplacer ou supprimer les clés existantes.

Après avoir établi la connexion RMI, vous pouvez naviguer entre les programmes de configuration à l'aide des commandes dscontrol, cbrcontrol, sscontrol, cocontrol, nalcontrol, dswizard, cbrwizard et sswizard, émises à partir d'une invite de commande. Vous pouvez également configurer Load Balancer à l'aide de l'interface graphique en entrant la commande lbadmin à partir d'une invite de commande.

Remarque : En raison des modifications apportées aux packages de sécurité de la version Java, les clés Load Balancer générées pour les éditions antérieures à 5.1.1 risquent de ne pas être compatibles avec celles de l'édition actuelle ; vous devez donc régénérer vos clés lorsque vous installez une nouvelle édition.

administration basée sur le Web

Conditions requises

La **machine client** sur laquelle s'effectue l'administration à distance requiert les éléments suivants pour l'administration basée sur le Web :

- JRE version 1.3.0 (ou supérieure)
- Pour plus d'informations sur les navigateurs pris en charge, accédez à la page Web suivante : <http://www.ibm.com/support/docview.wss?rs=180&uid=swg27006921>

Remarque : Si vous utilisez Netscape, ne modifiez pas la taille (Réduire, Agrandir, Restaurer en bas, etc.) du navigateur Netscape dans lequel s'affiche l'interface graphique de Load Balancer. En effet, étant donné que Netscape recharge une page à chaque redimensionnement de la fenêtre du navigateur, une déconnexion de l'hôte en découle. Vous devez donc vous reconnecter à l'hôte après chaque modification de la taille de la fenêtre.

La **machine hôte** à laquelle vous accédez pour l'administration à distance basée sur le Web requiert les éléments suivants :

- Caching Proxy version 6
- Perl version 5.5 (ou supérieure)

Configuration de Caching Proxy

- Avec Caching Proxy, l'utilitaire IBM Key Management (iKeyman) ou un autre utilitaire de gestion des clés est nécessaire pour créer des certificats de serveur SSL. (Pour plus d'informations sur la création des certificats, reportez-vous au manuel *Caching Proxy - Guide d'administration*.)

- Dans la section "Load Balancer Web-based Administration" (administration basée sur le Web de Load Balancer) du fichier de configuration de Caching Proxy (ibmproxy.conf), ajoutez les instructions suivantes après définition des domaines de protection, mais avant mappage des règles :

Pour les systèmes Windows —

```
Protect /lb-admin/lbwebaccess PROT-ADMIN
Exec /lb-admin/lbwebaccess C:\PROGRA~1\IBM\edge\lb\admin\lbwebaccess.pl
Pass /lb-admin/help/* C:\PROGRA~1\IBM\edge\lb\admin\help\*
Pass /lb-admin/*.jar C:\PROGRA~1\IBM\edge\lb\admin\lib\*.jar
Pass /lb-admin/* C:\PROGRA~1\IBM\edge\lb\admin\*
Pass /documentation/lang/*
C:\PROGRA~1\IBM\edge\lb\documentation\lang*
```

où *lang* désigne le sous-répertoire de votre langue de travail (par exemple, en_US)

Pour les systèmes Linux et UNIX —

```
Protect /lb-admin/lbwebaccess PROT-ADMIN
Exec /lb-admin/lbwebaccess /opt/ibm/edge/lb/admin/lbwebaccess.pl
Pass /lb-admin/help/* /opt/ibm/edge/lb/admin/help/*
Pass /lb-admin/*.jar /opt/ibm/edge/lb/admin/lib/*.jar
Pass /lb-admin/* /opt/ibm/edge/lb/admin/*
Pass /documentation/lang/*
/opt/ibm/edge/lb/documentation/lang*
```

Remarque : Sous HP-UX, le script lbwebaccess.pl suppose que le programme binaire Perl se trouve dans le répertoire /usr/bin/. (La première ligne du script contient #!/usr/bin/perl.) Entrez le chemin d'accès du répertoire de l'application Perl. Vous pouvez également créer un lien symbolique. Par exemple, si Perl est installé dans le répertoire /opt/perl/bin/perl, exécutez la commande suivante :

```
ln -s /opt/perl/bin/perl /usr/bin/perl
```

Exécution et accès à l'administration basée sur le Web

Pour s'exécuter, l'administration basée sur le Web doit être lancée sur la machine hôte Load Balancer en émettant la commande **lbwebaccess** à partir de l'invite de commande de l'hôte.

Vous devez également indiquer l'ID utilisateur et le mot de passe d'accès éloigné à l'hôte. Cet ID utilisateur et ce mot de passe sont identiques à l'ID utilisateur et au mot de passe d'administration Caching Proxy.

Pour afficher l'administration basée sur le Web de Load Balancer, accédez à l'URL suivante du navigateur Web à partir de l'emplacement éloigné :

```
http:// nom_hôte/lb-admin/lbadmin.html
```

Où *nom_hôte* est le nom de la machine à laquelle vous accédez pour communiquer avec Load Balancer.

Une fois la page Web chargée, l'interface graphique Load Balancer, nécessaire à l'administration à distance basée sur le Web, s'affiche dans la fenêtre du navigateur.

A partir de l'interface graphique Load Balancer, vous pouvez également exécuter des commandes de contrôle de la configuration. Pour émettre une commande à partir de l'interface graphique, procédez comme suit :

1. Sélectionnez le noeud Hôte dans l'arborescence de l'interface graphique.

2. Sélectionnez **Envoyer la commande...** dans le menu en incrustation Hôte.
3. Dans la zone d'entrée de commande, entrez la commande à exécuter. Par exemple : **executor report**. Les résultats et l'historique des commandes exécutées lors de la session courante s'affichent dans la fenêtre ouverte.

Régénération à distance de la configuration

Avec l'administration à distance basée sur le Web, lorsque plusieurs administrateurs mettent à jour la configuration Load Balancer à partir de postes éloignés, vous devez régénérer la configuration pour, par exemple, visualiser le cluster, le port ou le serveur ajouté (ou supprimé) par un autre administrateur. L'interface graphique de l'administration à distance basée sur le Web propose les fonctions **Régénérer la configuration** et **Régénérer toutes les configurations**.

Pour régénérer la configuration à partir de l'interface graphique basée sur le Web :

- Pour un seul hôte : cliquez à l'aide du bouton droit de la souris sur un noeud **Hôte** dans l'arborescence de l'interface graphique, puis sélectionnez **Régénérer la configuration**.
- Pour tous les hôtes : sélectionnez **Fichier** dans le menu, puis sélectionner **Régénérer toutes les configurations**.

Utilisation des journaux Load Balancer

Pour Dispatcher, CBR et Site Selector

Load Balancer enregistre les entrées dans un journal du serveur, un journal du gestionnaire, un journal du contrôleur de mesures (consignation des communications avec les agents Metric Server) et dans un journal pour chaque conseiller utilisé.

Remarque : De plus, pour le composant Dispatcher uniquement, les entrées peuvent être ajoutées dans un journal de sous-agent (SNMP).

Remarque : Le composant Content Based Routing (CBR) est disponible sur toutes les plateformes prises en charge à l'exception de celles exécutant une JVM 64 bits. Vous avez également la possibilité d'employer la méthode d'acheminement cbr du composant Dispatcher de Load Balancer pour permettre le routage CBR sans faire appel à Caching Proxy. Pour plus de détails, voir «Fonction CBR de Dispatcher (méthode d'acheminement cbr)», à la page 55.

Vous pouvez définir le niveau de consignation pour déterminer le détail des messages consignés dans les journaux. Au niveau 0, les erreurs sont enregistrées dans un fichier journal et Load Balancer consigne également les en-têtes et les enregistrements des événements survenus une seule fois (par exemple, un message indiquant que le lancement d'un conseiller sera enregistré dans le journal du gestionnaire). Le niveau 1 inclut les données en circulation, et ainsi de suite jusqu'au niveau 5, qui inclut tous les messages émis susceptibles d'aider à résoudre un incident lorsque cela s'avère nécessaire. La valeur par défaut du journal du gestionnaire, du conseiller, du serveur ou du sous-agent est 1.

Vous pouvez également fixer la taille maximale d'un fichier journal. Lorsque vous attribuez une taille maximale au fichier journal, ce dernier fonctionne en boucle. Lorsque le fichier atteint la taille indiquée, les entrées suivantes sont écrites en haut du fichier et remplacent les entrées existantes. Lorsque vous spécifiez une

nouvelle taille pour un fichier journal, elle ne doit pas être inférieure à sa taille courante. Les entrées de fichier journal sont horodatées, ce qui permet de déterminer l'ordre dans lequel elles ont été créées.

Plus le niveau de consignation choisi est élevé, plus la taille du fichier journal doit être définie judicieusement. Au niveau 0, il sera probablement sage de laisser la taille du fichier journal à sa valeur par défaut (1 Mo). Par contre, à partir du niveau 3, limitez la taille du fichier journal sans trop d'excès pour lui garder son utilité.

- Pour configurer le niveau de consignation ou la taille maximale du journal, utilisez la commande **dscontrol set**. (Pour afficher les paramètres du journal du serveur, utilisez la commande **dscontrol logstatus**.)
- Pour configurer le niveau de consignation ou la taille maximale du journal pour un journal de gestionnaire, utilisez la commande **dscontrol manager**.
- Pour configurer le niveau de consignation ou la taille maximale du journal du contrôleur de mesures qui consigne les événements de communication avec les agents Metric Server, utilisez la commande **dscontrol manager metric set**.
- Pour configurer le niveau de consignation ou la taille maximale du journal pour un journal de conseiller, utilisez la commande **dscontrol advisor**.
- Pour configurer le niveau de consignation ou la taille maximale du journal pour un journal de sous-agent, utilisez la commande **dscontrol subagent**. (Seul le composant Dispatcher utilise le sous-agent SNMP.)

Modification des chemins des fichiers journaux

Par défaut, les journaux générés par Load Balancer sont stockés dans le répertoire des journaux de l'installation de Load Balancer. Pour modifier ce chemin, définissez la variable *lb_logdir* dans le script dsserver.

Systèmes AIX, HP-UX, Linux et Solaris : Le script dsserver se trouve dans le répertoire /usr/bin. Dans ce script, la variable *lb_logdir* indique le répertoire par défaut. Vous pouvez modifier cette variable pour indiquer le répertoire de fichiers journaux de votre choix. Exemple :

```
LB_LOGDIR=/chemin\de\mes\fichiers\journaux/
```

Systèmes Windows : Le fichier dsserver se trouve dans le répertoire système Windows C:\WINNT\SYSTEM32, pour Windows 2003. Dans le fichier dsserver, la variable *lb_logdir* indique le répertoire par défaut. Vous pouvez modifier cette variable pour indiquer le répertoire de fichiers journaux de votre choix. Exemple :

```
set LB_LOGDIR=c:\chemin\de\mes\fichiers\journaux\
```

Quel que soit le système d'exploitation utilisé, assurez-vous qu'il n'y a pas d'espace avant ou après le signe égal et que le chemin se termine par une barre oblique ("/" ou "\") selon le cas).

Consignation binaire

Remarque : La fonction de consignation binaire ne s'applique pas au composant Site Selector.

La fonction de consignation binaire de Load Balancer utilise le répertoire contenant les autres fichiers journaux. Voir «Utilisation de la consignation binaire pour analyser les statistiques des serveurs», à la page 240.

Pour Cisco CSS Controller et Nortel Alteon Controller

Vous pouvez définir le niveau de consignation pour déterminer le détail des messages consignés dans les journaux. Au niveau 0, les erreurs sont enregistrées dans un fichier journal et Load Balancer consigne également les en-têtes et les enregistrements des événements survenus une seule fois (par exemple, un message indiquant que le lancement d'un conseiller sera enregistré dans le journal du consultant). Le niveau 1 inclut les données en circulation, et ainsi de suite jusqu'au niveau 5, qui inclut tous les messages émis susceptibles d'aider à résoudre un incident lorsque cela s'avère nécessaire. Le niveau par défaut des journaux est 1.

Vous pouvez également fixer la taille maximale d'un fichier journal. Dans ce cas, le fichier se bouclera ; une fois sa taille maximale atteinte, les nouvelles entrées seront consignées au début du fichier, écrasant les entrées les plus anciennes. Lorsque vous spécifiez une nouvelle taille pour un fichier journal, elle ne doit pas être inférieure à sa taille courante. Les entrées de fichier journal sont horodatées, ce qui permet de déterminer l'ordre dans lequel elles ont été créées.

Plus le niveau de consignation choisi est élevé, plus la taille du fichier journal doit être définie judicieusement. Au niveau 0, il sera probablement sage de laisser la taille du fichier journal à sa valeur par défaut (1 Mo). Par contre, à partir du niveau 3, limitez la taille du fichier journal sans trop d'excès pour lui garder son utilité.

Journaux des contrôleurs

Cisco CSS Controller et Nortel Alteon Controller ont les journaux suivants :

- journal du contrôleur (commande **controller set**)
- journal du consultant (commande **consultant set**)
- journal de haute disponibilité (commande **highavailability set**)
- journal du collecteur de mesures (commande **metriccollector set**)
- journal binaire (commande **consultant binarylog**)

Voici un exemple de configuration du niveau de consignation ou la taille maximale du journal du contrôleur de mesures qui consigne les événements de communication avec les agents Metric Server :

```
xxxcontrol metriccollector set IDconsultant:IDservice:NomSystemMetric  
loglevel x logsize y
```

Modification des chemins des fichiers journaux

Par défaut, les journaux générés par les contrôleurs sont stockés dans le répertoire des journaux de l'installation du contrôleur. Pour modifier ce chemin, définissez la variable *xxx_logdir* dans le script xxxserver.

Systèmes AIX, HP-UX, Linux et Solaris : Le script xxxserver se trouve dans le répertoire /usr/bin. Dans ce script, la variable *xxx_logdir* indique le répertoire par défaut. Vous pouvez modifier cette variable pour indiquer le répertoire de fichiers journaux de votre choix. Exemple :

```
xxx_LOGDIR=/chemin\de\mes\fichiers\journaux/
```

Systèmes Windows : Le fichier xxxserver se trouve dans le répertoire système Windows, généralement C:\WINNT\SYSTEM32. Dans le fichier xxxserver, la variable *xxx_logdir* indique le répertoire par défaut. Vous pouvez modifier cette variable pour indiquer le répertoire de fichiers journaux de votre choix. Exemple :

```
set xxx_LOGDIR=c:\chemin\de\mes\fichiers\journaux\
```

Quel que soit le système d'exploitation utilisé, assurez-vous qu'il n'y a pas d'espace avant ou après le signe égal et que le chemin se termine par une barre oblique ("/" ou "\" selon le cas).

Consignation binaire

La fonction de consignation binaire de Load Balancer utilise le répertoire contenant les autres fichiers journaux. Voir «Utilisation de la consignation binaire pour analyser les statistiques des serveurs», à la page 240.

Utilisation du composant Dispatcher

La présente section explique comment utiliser et gérer le composant Dispatcher.

Démarrage et arrêt de Dispatcher

- A partir d'une ligne de commande, entrez **dsserver** pour démarrer Dispatcher.
- A partir d'une ligne de commande, entrez **dsserver stop** pour arrêter Dispatcher.

Utilisation de la valeur du délai d'attente

Pour Load Balancer, les connexions sont considérées comme périmées lorsqu'aucune activité ne s'est produite sur cette connexion pendant le nombre de secondes indiquées dans le délai d'attente. Lorsque ce nombre de secondes est dépassé et qu'aucune activité n'a eu lieu, Load Balancer supprime cet enregistrement de connexion de ces tables et le trafic à venir pour cette connexion est ignoré.

Au niveau du port, par exemple, vous pouvez indiquer la valeur du délai d'attente à partir de la commande **dscontrol port set staletimeout**.

Le délai d'attente peut être défini au niveau de l'exécuteur, du cluster et du port. Au niveau de l'exécuteur et du cluster, la valeur par défaut est 300 secondes et le filtrage est effectué jusqu'au port. Au niveau du port, la valeur par défaut dépend du port. Certains ports bien définis ont des valeurs de délai d'attente différentes. Par exemple, le port telnet 23 a une valeur par défaut de 259,200 secondes.

Certains services ont également leurs propres valeurs de délai d'attente. Par exemple, LDAP (Lightweight Directory Access Protocol) dispose d'un paramètre de configuration appelé **idletimeout**. Lorsque le nombre de secondes indiqué à l'aide de ce paramètre est dépassé, la fermeture d'une connexion de client inactif est provoquée. Vous pouvez attribuer la valeur 0 à ce paramètre, ce qui signifie que la fermeture de la connexion ne sera jamais provoquée.

Des problèmes de connectivité peuvent se produire lorsque la valeur du délai d'attente de Load Balancer est inférieure à la valeur du délai du service. Dans le cas de LDAP, la valeur par défaut du paramètre **staletimeout** (délai d'attente) de Load Balancer est 300 secondes. Si aucune activité ne se produit sur la connexion pendant 300 secondes, Load Balancer supprime l'enregistrement de connexion de ses tables. Si la valeur **idletimeout** est supérieure à 300 secondes (ou si elle est égale à 0), le client peut encore croire qu'une connexion au serveur est établie. Lorsque le client transmet des paquets, ces derniers seront ignorés par Load Balancer. De cette façon, LDAP est bloqué lorsqu'une demande au serveur est

effectuée. Pour éviter ce problème, attribuez une valeur différente de zéro au paramètre, identique ou inférieure à la valeur du paramètre `staletimeout` de Load Balancer.

Contrôle du nettoyage des enregistrements de connexions à l'aide des paramètres `fintimeout` et `staletimeout`

Une fois tous les paquets de données transmis, un client envoie un paquet FIN pour informer le serveur que la transaction est terminée. Lorsque Dispatcher réceptionne le paquet FIN, il remplace l'état de la transaction, active, par FIN. Lorsqu'une transaction est à l'état FIN, la mémoire réservée à la connexion est libérée.

Pour améliorer les performances de l'affectation et de la réutilisation des enregistrements de connexions, utilisez la commande **`executor set fintimeout`** pour contrôler la période pendant laquelle Dispatcher doit conserver les connexions à l'état FIN, c'est-à-dire actives, dans les tables Dispatcher et en état d'accepter le trafic. Lorsqu'une connexion à l'état FIN dépasse la valeur **`fintimeout`**, elle est supprimée des tables Dispatcher et prête à être utilisée. Vous pouvez modifier la valeur du délai d'expiration FIN à l'aide de la commande **`dscontrol executor set fincount`**.

Utilisez la commande **`dscontrol executor set staletimeout`** pour contrôler la période pendant laquelle Dispatcher doit conserver les connexions à l'état Established lorsqu'aucun trafic n'a été détecté comme étant actif dans les tables Dispatcher, et en état d'accepter le trafic. Pour plus d'informations, voir «Utilisation de la valeur du délai d'attente», à la page 268.

Interface graphique — option de menu Contrôler

Divers diagrammes peuvent être affichés en fonction des informations visualisées par l'exécuteur et transmises au gestionnaire. (Le gestionnaire doit s'exécuter pour pouvoir utiliser l'option de menu Contrôler de l'interface graphique).

- Nombre de connexions par seconde pour chaque serveur (plusieurs serveurs peuvent être affichés sur un même graphique)
- Pondérations relatives pour chaque serveur d'un port donné
- Durée moyenne de connexion pour chaque serveur d'un port donné

Utilisation du protocole SNMP (Simple Network Management Protocol, protocole simplifié de gestion de réseau) avec le composant Dispatcher

Un système de gestion de réseau est un programme qui s'exécute en continu et qui sert à surveiller et à refléter l'état d'un réseau et à contrôler ce dernier. SNMP (Simple Network Management Protocol), protocole courant permettant de communiquer avec des périphériques d'un réseau, est la norme de gestion de réseau en cours. Les périphériques de réseau sont généralement dotés d'un *agent* SNMP et d'un ou de plusieurs sous-agents. L'agent SNMP communique avec le *poste de gestion de réseau* ou répond aux requêtes SNMP de la ligne de commande. Le *sous-agent* SNMP extrait et met à jour des données et transmet ces dernières à l'agent SNMP de sorte que celui-ci communique en retour avec le demandeur.

Dispatcher donne une *Bibliothèque d'informations de gestion* SNMP (`ibmNetDispatcherMIB`) et un sous-agent SNMP. Cela permet d'utiliser un système de gestion de réseau (tel que Tivoli NetView, Tivoli Distributed Monitoring ou HP

OpenView) pour surveiller l'état, le débit et l'activité de Dispatcher. Les données MIB décrivent la gestion de Dispatcher et reflètent l'état en cours de ce dernier. Elles sont installées dans le sous-répertoire **..lb/admin/MIB**.

Remarque : Les données MIB, **ibmNetDispatcherMIB.02**, ne seront pas chargées à l'aide du programme **xnmloadmib2** de Tivoli NetView. Pour résoudre ce problème, mettez en commentaire la section NOTIFICATION-GROUP des données MIB. En d'autres termes, insérez **"- "** au début de la ligne **"indMibNotifications Group NOTIFICATION-GROUP"** et au début des 6 lignes suivantes.

Le système de gestion de réseau utilise des commandes SNMP GET pour consulter les valeurs MIB des autres machines. Il peut ensuite vous envoyer une notification en cas de dépassement des valeurs seuil indiquées. Vous pouvez ensuite changer les performances de Dispatcher en modifiant les données de configuration de Dispatcher, afin d'effectuer une mise au point proactive ou de résoudre les incidents liés à Dispatcher avant qu'ils se transforment en pannes de serveur Web ou Dispatcher.

Commandes et protocole SNMP

Le système fournit généralement un agent SNMP pour chaque poste de gestion de réseau. L'utilisateur adresse une commande GET à l'agent SNMP. En retour, ce dernier émet une commande GET pour extraire les valeurs de variables MIB indiquées à partir d'un sous-agent responsable de ces dernières.

Dispatcher fournit un sous-agent qui permet la mise à jour et l'extraction de données MIB. Le sous-agent répond aux données MIB appropriées lorsque l'agent SNMP émet une commande GET. L'agent SNMP communique les données au poste de gestion de réseau. Celui-ci peut vous envoyer une notification en cas de dépassement des valeurs seuil indiquées.

Le support SNMP de Dispatcher comporte un sous-agent SNMP qui utilise la fonction DPI (Distributed Program Interface). Il s'agit d'une interface entre un agent SNMP et les sous-agents de ce dernier. Windows utilise l'agent d'extension Windows en tant qu'interface entre un agent SNMP et les sous-agents de ce dernier.

Activation de SNMP sur les systèmes AIX, HP-UX, Linux et Solaris

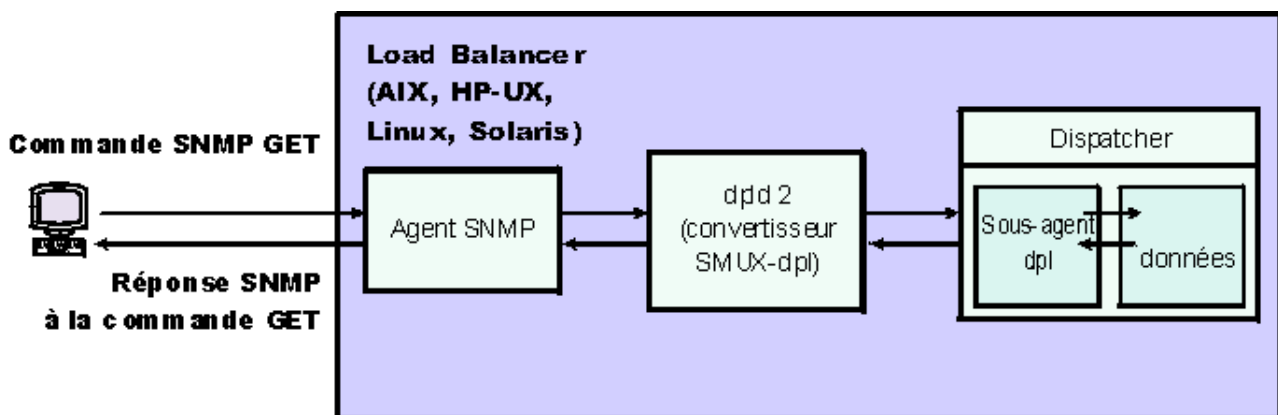


Figure 40. Commandes SNMP pour les systèmes Linux et UNIX

Les systèmes AIX fournissent un agent SNMP qui utilise le protocole SNMP Multiplexer (SMUX) et fournissent DPID2, qui est un exécutable supplémentaire fonctionnant comme traducteur entre DPI et SMUX.

Pour les systèmes HP-UX, vous devez obtenir un agent SNMP fonctionnant avec SMUX car HP-UX n'en fournit pas. Load Balancer fournit DPID2 pour les systèmes HP-UX.

Les systèmes Linux fournissent un agent SNMP qui utilise SMUX. La plupart des versions Linux (Red Hat, par exemple) sont livrées avec un package UCD SNMP. UCD SNMP version 4.1 ou ultérieure dispose d'agents SMUX actifs. Load Balancer fournit DPID2 pour les systèmes Linux.

Remarque : Pour les systèmes SuSE Linux, vous devez obtenir un agent SNMP configuré pour prendre en charge SMUX car SuSE n'en fournit pas.

Pour les systèmes Solaris, vous devez obtenir un agent SNMP fonctionnant avec SMUX car Solaris n'en fournit pas. Load Balancer fournit DPID2 pour les systèmes Solaris dans le répertoire `/opt/ibm/edge/lb/servers/samples/SNMP`.

L'agent DPI doit fonctionner comme un utilisateur root. Avant d'exécuter le démon DPID2, mettez à jour les fichiers `/etc/snmpd.peers` et `/etc/snmpd.conf` comme suit :

Pour les systèmes AIX et Solaris :

- Dans le fichier `/etc/snmpd.peers`, ajoutez l'entrée suivante pour dpid :
"dpid2" 1.3.6.1.4.1.2.3.1.2.2.1.1.2 "mot_de_passe_dpid"
- Dans `/etc/snmpd.conf`, ajoutez l'entrée suivante pour dpid :
smux 1.3.6.1.4.1.2.3.1.2.2.1.1.2 mot_de_passe_dpid #dpid

Pour les systèmes Linux :

- Dans `/etc/snmpd.peers` (s'il n'existe pas, créez-le), ajoutez l'entrée suivante pour dpid :
"dpid2" 1.3.6.1.4.1.2.3.1.2.2.1.1.2 "mot_de_passe_dpid"
- Dans le fichier `/etc/snmp/snmpd.conf`, ajoutez l'entrée suivante pour dpid :
smuxpeer .1.3.6.1.4.1.2.3.1.2.2.1.1.2 mot_de_passe_dpid

Vous devez également associer un commentaire à toutes les lignes du fichier `snmpd.conf` qui commencent par les mots suivants : `com2sec`, `group`, `view` ou `access`.

Activation de SNMP sur les systèmes HP-UX

Pour installer le support SNMP de HP-UX, procédez comme suit :

1. Si aucune version de GNU SED n'est installée, procurez-la vous à partir du site Web suivant : <http://www.hp.com>.
2. Récupérez le fichier `ucd-snmp-4.2.4.tar.gz` de la page Web suivante : http://sourceforge.net/project/showfiles.php?group_id=12694.
3. Vérifiez que "gcc" et "gmake ou make" sont installés sur votre machine. Si ce n'est pas le cas, installez-les.
4. Décompressez le fichier `ucd-snmp-4.2.4.tar.gz` ainsi que tous les fichiers source du répertoire.
5. Allez dans le répertoire où se trouvent les fichiers source, puis exécutez les commandes suivantes :

- a. `run ./configure --with-mib-modules=smux`
- b. `make`
- c. Exécutez les deux commandes suivantes en tant que superutilisateur :
 - 1) `umask 022`
 - 2) `make install`
- d. `export SNMPCONFPATH=/etc/snmp`
- e. `start /usr/local/sbin/snmpd -s` (Démarre l'agent SNMP)
- f. `start dpid2` (Démarre le convertisseur DPI)
- g. `dscontrol subagent start` (Démarre le sous-agent Dispatcher)

Activation de SNMP sur les systèmes SuSE Linux

Pour utiliser Load Balancer SNMP avec SuSE Linux, procédez comme suit :

1. Supprimez le module `ucd-snmp` rpm installé du système SuSE.
2. Extrayez `ucd-snmp-4.2.4.tar.gz` de http://sourceforge.net/project/showfiles.php?group_id=12694.
3. Vérifiez que "`gcc`" et "`gmake` ou `make`" sont installés sur le système SuSE (s'ils ne sont pas installés, effectuez l'opération).
4. Décompressez le fichier `ucd-snmp-4.2.4.tar.gz` ainsi que tous les fichiers source du répertoire.
5. Allez dans le répertoire où se trouvent les fichiers source, puis exécutez les commandes suivantes :
 - a. `run ./configure --with-mib-modules=smux`
 - b. `make`
 - c. Exécutez les deux commandes suivantes en tant que superutilisateur :
 - 1) `umask 022 #`
 - 2) `make install`
 - d. `export SNMPCONFPATH=/etc/snmp`
 - e. `start /usr/local/sbin/snmpd -s`
 - f. `start dpid2`

Régénérez `snmpd` (s'il s'exécute déjà) de sorte qu'il relise le fichier `snmpd.conf` :
`refresh -s snmpd`

Démarrez l'homologue DPID SMUX :
`dpid2`

Les démons doivent être lancés dans l'ordre suivant :

1. Agent SNMP
2. Programme de traduction DPI
3. Sous-agent Dispatcher

Activation de SNMP sur les systèmes Solaris

Pour installer le support SNMP de Solaris, procédez aux opérations ci-dessous.

1. Tuez le démon SNMP de Solaris qui s'exécute (`snmpdx` et `snmpXdmi`).
2. Renommez les fichiers comme suit :
 - `/etc/rc3.d/S76snmpdx` en `/etc/rc3.d/K76snmpdx`
 - `/etc/rc3.d/S77dmi` en `/etc/rc3.d/K77dmi`
3. Téléchargez les modules suivants à partir de <http://www.sunfreeware.com/> :

- libgcc-3.0.3-sol8-sparc-local (SMClibgcc)
 - openssl-0.9.6c-sol8-sparc-local (SMCssl)
 - popt-1.6.3-sol8-sparc-local (SMCpopt)
4. Installez les modules téléchargés à l'aide de la commande `pkgadd`.
 5. Téléchargez `ucd-snmp-4.2.3-solaris8.tar.gz` à partir de http://sourceforge.net/project/showfiles.php?group_id=12694
 6. Décompressez `ucd-snmp-4.2.3-solaris8.tar.gz` (avec `gunzip` ou `untar`) sur le répertoire racine (`/`)
 7. Emettez les commandes suivantes :


```
export LD_LIBRARY_PATH=$LD_LIBRARY_PATH: /usr/local/lib:/usr/local/ssl/lib:/usr/lib
export PATH=/usr/local/sbin:/usr/local/bin:$PATH
export SNMPCONFPATH=/etc/snmp
export MIBDIRS=/usr/local/share/snmp/mibs
cp /opt/ibm/edge/lb/servers/samples/SNMP/dpid2
   /usr/local/sbin/dpid2
```
 8. S'il n'existe pas, créez `/etc/snmpd.peers`. Insérez dans `snmpd.peers` les éléments suivants :


```
"dpid2"
1.3.6.1.4.1.2.3.1.2.2.1.1.2      "mot_de_passe_dpid"
```
 9. S'il n'existe pas, créez `/etc/snmp/snmpd.conf`. Insérez dans `snmpd.conf` les éléments suivants :


```
smuxpeer
1.3.6.1.4.1.2.3.1.2.2.1.1.2      mot_de_passe_dpid
```
 10. Démarrez `/usr/local/sbin/snmpd`.
 11. Démarrez `/usr/local/sbin/dpid2`.

Remarques :

1. Les modules suivants sont au format de module.
 - libgcc-3.0.3-sol8-sparc-local (SMClibgcc)
 - openssl-0.9.6c-sol8-sparc-local (SMCssl)
 - popt-1.6.3-sol8-sparc-local (SMCpopt)
 Sur le site Web <http://sunfreeware.com/>, les noms portent l'extension `.gz`, aussi ne leur appliquez pas la commande de décompression `gunzip` ou `untar`. Utilisez plutôt la commande `pkgadd NomModule`.
2. Lorsque vous ajoutez `smuxpeer` à `/etc/snmp/snmpd.conf`, assurez-vous qu'aucun espace n'est entré dans la chaîne **mot_de_passe_dpid**.
3. La fonction SNMP de Load Balancer est testée avec `ucd-snmp` version 4.2.3 et prise en charge de `smux`. Les futures éditions de `ucd-snmp` avec `smux` devraient fonctionner avec une configuration similaire.

Activation de SNMP sous Windows

Pour installer le support SNMP de Windows, procédez comme suit :

1. Cliquez sur Démarrer > Paramètres (Windows 2000) > Panneau de configuration > Ajout/Suppression de programmes.
2. Cliquez sur **Ajouter/Supprimer des composants Windows**.
3. Dans l'assistant Composants de Windows, cliquez sur **Outils de gestion et d'analyse** (sans activer ou désactiver la case à cocher correspondante), puis cliquez sur **Détails**

4. Cochez la case **SNMP (Protocole simplifié de gestion de réseau)**, puis cliquez sur OK.
5. Cliquez sur Suivant.

Définition d'un nom de communauté pour SNMP

L'exécuteur étant en cours de fonctionnement, lancez la commande **dscontrol subagent start [nom_communauté]** pour définir le nom de communauté utilisé entre l'agent d'extension Windows et l'agent SNMP.

IMPORTANT : Sous Windows 2003, par défaut SNMP ne répond à aucun nom de communauté présenté. Dans ce cas, le sous-agent SNMP ne répond à aucune demande SNMP. Pour vous assurer que le sous-agent SNMP réponde au nom de communauté, vous devez affecter aux propriétés du service SNMP le nom de communauté et les hôtes de destination appropriés. Configurez les propriétés de la sécurité SNMP de la manière suivante :

1. Ouvrez Gestion de l'ordinateur
2. Dans l'arborescence de la console, cliquez sur **Services**
3. Dans la sous-fenêtre des détails, cliquez sur **Service SNMP**
4. Dans le menu d'action, cliquez sur **Propriétés**
5. Dans la page Sécurité, sous les noms de communauté acceptés, cliquez sur **Ajouter**
6. Sous Droits de communauté, sélectionnez le niveau d'autorisation de cet hôte pour le traitement des demandes à partir de la communauté sélectionnée (droit de **lecture** au minimum)
7. Dans la zone Nom de la communauté, entrez le nom fourni au sous-agent Load Balancer, en respectant la casse (nom de communauté par défaut : public), puis cliquez sur **Ajouter**
8. Spécifiez si les paquets d'un hôte SNMP doivent être acceptés. Choisissez l'une des options suivantes :
 - Pour accepter les demandes SNMP d'un hôte du réseau, quelle que soit son identité, cliquez sur **Accepter les paquets SNMP provenant de n'importe quel hôte**. (Lorsque cette option est sélectionnée, les personnes et les entités doivent être authentifiées, en fonction de critères tels qu'un mot de passe ou un certificat.)
 - Pour limiter l'acceptation des paquets SNMP, cliquez sur **Limiter l'acceptation des paquets SNMP**, sur **Accepter les paquets SNMP provenant de ces hôtes**, puis sur **Ajouter**. Entrez le nom d'hôte ou l'adresse IP ou IPX approprié, puis cliquez sur **Ajouter**, après chaque entrée.
9. Redémarrez le service SNMP pour que la modification soit appliquée

Interruptions

SNMP communique en envoyant et en recevant des *interruptions*, messages envoyés par des périphériques gérés afin de signaler des conditions d'erreur ou la survenue d'événements importants, par exemple, le dépassement d'un seuil.

Le sous-agent utilise les interruptions suivantes :

- indHighAvailStatus
- indSrvrGoneDown
- indDOSAttack
- indDOSAttackDone

L'interruption **indHighAvailStatus** annonce que la valeur de la variable (hasState) correspondant à l'état de la haute disponibilité a changé. Les valeurs possibles de hasState sont :

- idle** cette machine effectue un équilibrage de charge et n'essaie pas d'établir un contact avec son répartiteur associé.
- listen** La haute disponibilité vient de démarrer et Dispatcher est à l'écoute de son partenaire.
- active** cette machine effectue l'équilibrage de charge.
- standby**
cette machine contrôle la machine active.
- preempt**
cette machine est dans un état transitoire pendant le passage de la machine principale à la machine de secours.
- elect** Le répartiteur choisit, avec son partenaire, la machine principale et la machine de secours.
- no_exec**
L'exécuteur n'est pas lancé.

L'interruption **indSrvrGoneDown** annonce que la pondération du serveur spécifié par les segments csID (ID cluster), psNum (numéro port) et ssID (ID serveur) de l'identificateur d'objet est passée à zéro. Le dernier nombre total de connexions actives du serveur est envoyé avec l'interruption. Cette interruption indique que, pour le répartiteur, le serveur spécifié a été arrêté."

L'interruption **indDOSAttack** indique que numhalfopen (nombre de connexions partielles constituées uniquement de paquets SYN) a dépassé le seuil maxhalfopen pour le port précisé par les segments csID (ID de cluster) et psNum (numéro de port) de l'identificateur d'objet. Le nombre de serveurs configurés sur le port est transmis dans l'interruption. Cette interruption indique que Load Balancer peut faire l'objet d'une attaque de refus de service.

L'interruption **indDOSAttackDone** indique que numhalfopen (nombre de connexions partielles constituées uniquement de paquets SYN) est en deçà du seuil maxhalfopen pour le port précisé par les segments csID et psNum de l'identificateur d'objet. Le nombre de serveurs configurés sur le port est transmis dans l'interruption. Lorsque Load Balancer détermine que l'attaque de refus de service est terminée, cette interruption est envoyée après l'interruption indDOSAttack.

Pour les systèmes Linux et UNIX, en raison d'une restriction au niveau de l'interface API SMUX, il se peut que l'ID entreprise signalé dans des interruptions provenant du sous-agent ibmNetDispatcher corresponde à l'ID entreprise de dpid2 et non à celui d'ibmNetDispatcher, 1.3.6.1.4.1.2.6.144. Cependant, les utilitaires de gestion SNMP peuvent déterminer la source de l'interruption car les données contiennent un ID objet provenant du MIB ibmNetDispatcher.

Activation et désactivation du support SNMP à partir de la commande dscontrol

Le support SNMP est activé à l'aide de la commande **dscontrol subagent start**. Il est désactivé à l'aide de la commande **dscontrol subagent stop**.

Pour plus d'informations sur la commande dscontrol, voir «dscontrol subagent — Configuration du sous-agent SNMP», à la page 398.

Rejet de l'ensemble du trafic vers Load Balancer avec la fonction ipchains ou iptables (systèmes Linux)

Le pare-feu ipchains est intégré au noyau Linux. Lors de l'exécution simultanée de Load Balancer et de ipchains, Load Balancer voit le paquets avant ipchains. Vous pouvez ainsi utiliser ipchains pour renforcer un système Load Balancer Linux, tel qu'un système Load Balancer permettant de charger des pare-feux d'équilibrage de charge.

Lorsque ipchains ou iptables est configuré pour une utilisation restreinte (aucun trafic entrant ou sortant autorisé), la partie dédiée à l'acheminement des paquets de Load Balancer continue de fonctionner normalement.

Notez que ipchains et iptables *ne permettent pas* le filtrage du trafic entrant avant l'équilibrage de la charge.

Le fonctionnement correct de l'ensemble de Load Balancer nécessite l'autorisation de trafic supplémentaire. Voici quelques exemples de communication :

- Les conseillers communiquent entre le système Load Balancer et les serveurs dorsaux.
- Load Balancer envoie une commande ping aux serveurs dorsaux, aux cibles à atteindre et aux systèmes Load Balancer à haute disponibilité partenaires.
- Les interfaces utilisateur (interface graphique, ligne de commande et assistants) utilisent les appels RMI.
- Les serveurs dorsaux doivent répondre aux commandes ping provenant du système Load Balancer.

En règle générale, la stratégie ipchains adaptée aux systèmes Load Balancer consiste à refuser tout type de trafic, sauf celui à destination et provenant des serveurs dorsaux, du Load Balancer haute disponibilité partenaire, des cibles à atteindre ou des hôtes de configuration.

N'activez pas iptables lorsque Load Balancer s'exécute avec le noyau Linux version 2.4.10.x. En effet, l'activation sous cette version de noyau Linux peut provoquer à terme une dégradation des performances.

Pour désactiver iptables, listez les modules (`lsmod`) pour savoir lesquels utilisent `ip_tables` et `ip_conntrack`, puis supprimez ceux-ci à l'aide des commandes `rmmod ip_tables` et `rmmod ip_conntrack`. Lorsque vous réamorcez la machine, ces modules sont de nouveau ajoutés de sorte que vous devez répéter cette procédure après chaque réamorçage.

Pour plus d'informations, voir «Incident : Les iptables de Linux peuvent interférer avec le routage de paquets», à la page 324.

Utilisation du composant CBR (Content Based Routing)

La présente section explique comment utiliser le composant CBR de Load Balancer.

Remarque : Le composant Content Based Routing (CBR) est disponible sur toutes les plateformes prises en charge à l'exception de celles exécutant une JVM 64 bits. Vous avez également la possibilité d'employer la méthode d'acheminement `cbr` du composant Dispatcher de Load Balancer pour permettre le routage CBR sans faire appel à Caching

Proxy. Pour plus de détails, voir «Fonction CBR de Dispatcher (méthode d'acheminement cbr)», à la page 55.

Démarrage et arrêt de CBR

- A partir d'une ligne de commande, entrez **cbrserver** pour lancer CBR.
- A partir d'une ligne de commande, entrez **cbrserver stop** pour arrêter CBR.

CBR et Caching Proxy gèrent conjointement les demandes HTTP et HTTPS (SSL) via l'interface API du plug-in de Caching Proxy. Caching Proxy doit être en cours d'exécution sur la même machine pour que CBR puisse commencer à effectuer l'équilibrage de charge des serveurs. Configurez CBR et Caching Proxy en respectant les instructions de la section «Exemple de configuration CBR», à la page 118.

Contrôle de CBR

Après avoir lancé CBR, vous pouvez le contrôler en utilisant une des méthodes suivantes :

- Configurez CBR à l'aide de la commande **cbrcontrol**. La syntaxe complète de cette commande est décrite dans Chapitre 27, «Guide des commandes Dispatcher et CBR», à la page 345. Voici quelques exemples d'utilisation de cette commande.
- Configurez CBR à l'aide de l'interface utilisateur graphique. A partir de la ligne de commande, entrez **lbadmin** pour ouvrir l'interface graphique. Pour obtenir plus d'informations sur le mode de configuration de CBR à l'aide de l'interface graphique, voir «Interface graphique», à la page 112.

Utilisation des journaux de CBR

Les journaux utilisés par CBR sont similaire à ceux de Dispatcher. Pour plus d'informations, voir «Utilisation des journaux Load Balancer», à la page 265.

Remarque :

Dans les versions précédentes, pour CBR, il était possible de modifier le chemin d'accès au répertoire log dans le fichier de configuration Caching Proxy. Vous pouvez maintenant changer le chemin du répertoire dans lequel le journal est stocké dans le fichier `cbrserver`. Pour plus de détails, voir «Modification des chemins des fichiers journaux», à la page 267.

Utilisation du composant Site Selector

Démarrage et arrêt de Site Selector

- A partir d'une ligne de commande, entrez **sssserver** pour lancer Site Selector.
- A partir d'une ligne de commande, entrez **sssserver stop** pour arrêter Site Selector.

Contrôle d'Site Selector

Après avoir lancé Site Selector, vous pouvez le contrôler en utilisant une des méthodes suivantes :

- Configurez Site Selector à l'aide de la commande **sscontrol**. La syntaxe complète de cette commande est décrite dans Chapitre 28, «Guide des commandes Site Selector», à la page 401. Voici quelques exemples d'utilisation de cette commande.

- Configurez Site Selector à l'aide de l'interface utilisateur graphique. A partir de la ligne de commande, entrez **lbadm** pour ouvrir l'interface graphique. Pour obtenir plus d'informations sur le mode de configuration de Site Selector à l'aide de l'interface graphique, voir «Interface graphique», à la page 132.

Utilisation des journaux Site Selector

Les journaux employés par Site Selector sont similaires à ceux utilisés dans Dispatcher. Pour plus d'information, voir «Utilisation des journaux Load Balancer», à la page 265.

Utilisation du composant Cisco CSS Controller

Démarrage et arrêt de Cisco CSS Controller

1. A partir d'une ligne de commande, entrez **ccoserver** pour lancer Cisco CSS Controller.
2. A partir d'une ligne de commande, entrez **ccoserver stop** pour arrêter Cisco CSS Controller.

Contrôle de Cisco CSS Controller

Après avoir lancé Cisco CSS Controller, vous pouvez le contrôler en utilisant une des méthodes suivantes :

- Configurez Cisco CSS Controller à l'aide de la commande **cococontrol**. La syntaxe complète de cette commande est décrite dans Chapitre 29, «Guide des commandes Cisco CSS Controller», à la page 429. Voici quelques exemples d'utilisation de cette commande.
- Configurez Cisco CSS Controller à l'aide de l'interface utilisateur graphique. A partir de la ligne de commande, entrez **lbadm** pour ouvrir l'interface graphique. Pour plus d'informations sur la configuration de Cisco CSS Controller à l'aide de l'interface graphique, voir «Interface graphique», à la page 151.

Utilisation des journaux Cisco CSS Controller

Les journaux employés par Cisco CSS Controller sont similaires à ceux utilisés dans Dispatcher. Pour plus d'information, voir «Utilisation des journaux Load Balancer», à la page 265.

Utilisation du composant Nortel Alteon Controller

Démarrage et arrêt de Nortel Alteon Controller

1. A partir d'une ligne de commande, entrez **nalserver** pour démarrer Nortel Alteon Controller.
2. A partir d'une ligne de commande, entrez **nalserver stop** pour arrêter Nortel Alteon Controller.

Contrôle de Nortel Alteon Controller

Une fois Nortel Alteon Controller démarré, vous pouvez le contrôler en utilisant une des méthodes suivantes :

- Configurez Nortel Alteon Controller à l'aide de la commande **nalcontrol**. La syntaxe complète de cette commande est décrite dans Chapitre 30, «Guide des commandes Nortel Alteon Controller», à la page 449. Voici quelques exemples d'utilisation de cette commande.
- Configurez Nortel Alteon Controller à l'aide de l'interface utilisateur graphique. A partir de la ligne de commande, entrez **lbadmin** pour ouvrir l'interface graphique. Pour plus d'informations sur la configuration de Nortel Alteon Controller à l'aide de l'interface graphique, voir «Interface graphique», à la page 173.

Utilisation des journaux Nortel Alteon Controller

Les journaux employés de Nortel Alteon Controller sont similaires à ceux de Dispatcher. Pour plus d'information, voir «Utilisation des journaux Load Balancer», à la page 265.

Utilisation du composant Metric Server

Démarrage et arrêt de Metric Server

Metric Server fournit à Load Balancer des informations relatives à la charge des serveurs. Il réside sur chaque serveur soumis à l'équilibrage de charge.

Systemes Linux et UNIX :

- A partir de la ligne de commande de chaque serveur hébergeant Metric Server, entrez **metricserver start** pour lancer Metric Server.
- A partir de la ligne de commande de chaque serveur hébergeant Metric Server, entrez **metricserver stop** pour arrêter Metric Server.

Systemes Windows :

Cliquez sur Démarrer > Paramètres (pour Windows 2000) > Panneau de configuration > Outils d'administration > Services. Cliquez à l'aide du bouton droit de la souris sur IBM Metric Server, puis sélectionnez Démarrer. Pour arrêter le service, suivez la même procédure en sélectionnant Arrêter.

Utilisation des journaux Metric Server

Modifiez le niveau de consignation dans le script de démarrage de Metric Server. Vous pouvez indiquer un niveau de consignation compris entre 0 et 5, à l'instar de la plage admise pour les journaux de Load Balancer. Cette action génère un journal des agents dans le répertoire **...ms/logs**.

Chapitre 25. Résolution des incidents

Ce chapitre permet la détection et la résolution des incidents associés à Load Balancer.

- Avant d'appeler le service d'assistance IBM voir «Collecte des informations de résolution des incidents».
- Recherchez le symptôme rencontré dans le «Tableaux de résolution des incidents», à la page 285.

Collecte des informations de résolution des incidents

Utilisez les informations de cette section pour rassembler les données nécessaires au service d'assistance IBM. Les informations sont classées sous les rubriques suivantes :

- «Informations générales (obligatoires)»
- «Incidents liés à la haute disponibilité», à la page 282
- «Incidents liés aux conseillers», à la page 283
- «Incident liés au routage par contenu (CBR)», à la page 284
- «Impossibilité d'accéder au cluster», à la page 284
- «Echec de toutes les tentatives de résolution des incidents», à la page 285
- «Mises à niveau», à la page 285
- «Liens utiles», à la page 285

Informations générales (obligatoires)

Le composant Dispatcher (et lui seul), dispose d'un outil d'identification des incidents qui collecte automatiquement les données propres au système d'exploitation et les fichiers de configuration de composants donnés. Pour lancer cet outil, entrez **lbpd** à partir du répertoire approprié, c'est-à-dire :

Pour les systèmes Linux et UNIX : `/opt/ibm/edge/lb/servers/bin/`

Pour les systèmes Windows : `C:\Program Files\IBM\edge\lb\servers\bin`

Cet outil d'identification des incidents regroupe les données collectées dans des fichiers comme suit :

Pour les systèmes Linux et UNIX : `/opt/ibm/edge/lb/lbpmr.tar.Z`

Pour les systèmes Windows : `C:\Program Files\IBM\edge\lb\lbpmr.zip`

Remarque : Vous devez disposer d'un utilitaire de compression de ligne de commande pour les systèmes Windows.

Avant d'appeler le service d'assistance IBM, regroupez les informations ci-après.

- Pour Dispatcher uniquement, le fichier `lbpmr` généré par l'outil d'identification des incidents précédemment cité.
- Dans un environnement à haute disponibilité, les fichiers de configuration des deux machines Load Balancer. Sur tous les systèmes d'exploitation, utilisez le script permettant de charger la configuration ou entrez la commande suivante :
`dscontrol file save primary.cfg`

Cette commande place le fichier de configuration dans le répertoire `.../ibm/edge/lb/servers/configuration/composant/`.

- Le nom et la version du système d'exploitation que vous utilisez.
- La version de Load Balancer.
 - Si Load Balancer s'exécute, entrez les commandes suivantes :
 - Pour le composant Dispatcher : `dscontrol executor report`
 - Pour CBR : `cbrcontrol executor status`
 - Pour Site Selector, vérifiez le début du fichier `server.log` du répertoire `.../ibm/edge/lb/servers/logs/ss/`.
 - Pour Cisco CSS Controller et Nortel Alteon Controller : `xxxcontrol controller report`
 - Lancez les commandes suivantes pour vérifier que Load Balancer est installé et obtenir le niveau en cours de Load Balancer :
 - Sur les systèmes AIX : `lspp -l | grep ibmlb`
 - Sur les systèmes HP-UX : `swlist | grep ibmlb`
 - Sur les systèmes Linux : `rpm -qa | grep ibmlb`
 - Sur les systèmes Solaris : `pkginfo | grep ibm`

Sur les systèmes Windows, pour vérifier que Load Balancer est installé : Sélectionnez Démarrer > Paramètres > Panneau de configuration > Ajout/Suppression de programmes.
- Pour obtenir le niveau Java en cours, entrez la commande suivante :
`java -fullversion`
- Si vous utilisez un anneau à jeton ou Ethernet
- Entrez l'une des commandes suivantes pour extraire les statistiques de protocole et les informations de connexion TCP/IP :
 - Sur les systèmes AIX, HP-UX, Linux et Solaris : `netstat -ni`
 - Sur les systèmes Windows : `ipconfig /all`

Commande requise à partir de tous les serveurs et de Load Balancer.
- Pour extraire les informations de table de routage, entrez les commandes suivantes :
 - Sur les systèmes AIX, HP-UX, Linux et Solaris : `netstat -nr`
 - Sur les systèmes Windows : `route print`

Commande requise à partir de tous les serveurs et de Load Balancer.

Incidents liés à la haute disponibilité

En cas d'incident dans un environnement de haute disponibilité, regroupez les informations requises ci-après.

- Affectez à `hamon.log` le niveau de consignation 5, comme suit : `dscontrol set loglevel 5`.
- Affectez à `reach.log` le niveau de consignation 5, comme suit : `dscontrol manager reach set loglevel 5`.
- Extrayez les scripts des répertoires suivants :
 - Systèmes AIX, HP-UX, Linux et Solaris : `/opt/ibm/edge/lb/servers/bin`
 - Systèmes Windows : `C:\Program Files\ibm\edge\lb\servers\bin`

Les scripts à extraire sont les suivants :

```
goActive
goStandby
goIdle (s'il existe)
goInOp (s'il existe)
```


Extrayez aussi les fichiers de configuration. Pour plus de détails, voir «Informations générales (obligatoires)», à la page 281.

Incidents liés aux conseillers

En cas d'incident lié aux conseillers (par exemple lorsque des conseillers déclarent par erreur des serveurs inactifs), regroupez les informations requises ci-après.

- Attribuez au journal du conseiller le niveau de consignation 5, comme suit :

```
dscontrol advisor loglevel  
http 80 5
```

ou

```
dscontrol advisor loglevel NomConseiller port niveau_journal
```

ou

```
dscontrol advisor loglevel NomConseiller cluster:port niveau_journal
```

ou

```
nalcontrol metriccollector set IDconsultant:IDservice:NomSystemMetric  
niveau_journalvaleur
```

Cette ligne de commande crée un journal nommé *ADV_NomConseiller*, par exemple, *ADV_http.log*. Ce journal se trouve dans les répertoires suivants :

Plateformes AIX, HP-UX, Linux et Solaris : */opt/ibm/edge/lb/servers/logs/composant*

Plateformes Windows : *C:\Program Files\ibm\edge\lb\servers\logs\composant*

Où *composant* correspond à :

dispatcher = Dispatcher

cbr = CBR (Content Based Routing)

cco = Cisco CSS Controller

nal = Nortel Alteon Controller

ss = Site Selector

Remarque : Lors de la création d'un conseiller personnalisé, il est recommandé d'utiliser le journal *ADVLOG(niveaujournal,message)* pour vérifier que le conseiller fonctionne correctement.

L'appel *ADVLOG* génère des instructions dans le fichier journal des conseillers lorsque le niveau de consignation est inférieur à celui associé aux conseillers. Si le niveau de consignation est 0, l'instruction est toujours générée. Vous ne pouvez pas utiliser *ADVLOG* à partir du constructeur. Le fichier journal n'est créé qu'une fois que l'exécution du constructeur du conseiller personnalisé est terminée car le nom du fichier journal dépend des informations définies dans le constructeur.

Il existe cependant une autre manière de déboguer votre conseiller personnalisé qui permet d'éviter cette restriction. Vous pouvez utiliser des instructions *System.out.println(message)* pour imprimer les messages à l'écran. Editez le script *dsserver* et remplacez *javaw* par *java* pour que les instructions d'impression apparaissent dans la fenêtre. La fenêtre utilisée pour démarrer *dsserver* ne doit pas être fermée, pour que les impressions soient affichées. Si vous utilisez une plateforme Windows, vous devez arrêter le service Dispatcher, puis le démarrer manuellement à partir d'une fenêtre pour apercevoir les messages.

Pour plus d'informations sur ADVLOG, reportez-vous au document *Programming Guide for Edge Components*.

Incident liés au routage par contenu (CBR)

En cas d'incident lié au routage par contenu (CBR), regroupez les informations requises ci-après.

- Entrez la commande suivante pour afficher la version : `cbrcontrol executor status`.
- Extrayez les fichiers suivants :
 - `ibmproxy.conf`, des répertoires suivants :
 - systèmes Linux et UNIX : `/etc/`
 - Systèmes Windows : `C:\Program Files\IBM\edge\cp\etc\en_US\`
 - Le fichier de configuration CBR, situé dans les répertoires suivants :
 - systèmes Linux et UNIX : `/opt/ibm/edge/lb/servers/configurations/cbr`
 - Systèmes Windows : `C:\Program Files\IBM\edge\lb\servers\configurations\cbr`
 - Vérifiez que le fichier `ibmproxy.conf` contient les entrées appropriées. Pour plus de détails, voir «Etape 1. Configuration de Caching Proxy pour utiliser CBR», à la page 114.

Impossibilité d'accéder au cluster

Si vous n'arrivez pas à accéder au cluster, il est possible que l'une des machines Load Balancer ou les deux ont attribué un alias au cluster. Pour déterminer laquelle détient le cluster, procédez comme suit :

1. Sur le même sous-réseau et *non* sur une machine Load Balancer ou serveur :

```
ping
cluster
arp -a
```

Si vous utilisez des méthodes d'acheminement NAT ou CBR de Dispatcher, lancez également une commande ping vers l'adresse de retour.

2. Observez le résultat de `arp` et faites correspondre l'adresse MAC (adresse hexadécimale en 16 chiffres) avec l'un des résultats de `netstat -ni` pour déterminer à quelle machine appartient physiquement le cluster.
3. Utilisez les commandes suivantes pour interpréter le résultat des deux machines et déterminer si elles détiennent toutes deux l'adresse du cluster.

Sur les systèmes AIX et HP-UX : `netstat -ni`

Sur les systèmes Linux et Solaris : `ifconfig -a`

Sur les systèmes Windows : `ipconfig /all`

Si vous n'obtenez pas de réponse à la commande ping et que vous n'utilisez pas l'ULB, il est possible qu'aucune des machine n'ait attribué d'alias à l'adresse IP du cluster sur son interface, par exemple `en0`, `tr0` et ainsi de suite.

Remarque : Sur les systèmes Linux s'exécutant sur une installation Load Balancer pour IPv4 et IPv6, si vous n'obtenez pas de réponse à la commande ping, un serveur dorsal est indisponible ; toutefois, l'entrée arp doit quand même être mise à jour. Une autre possibilité consiste à utiliser éventuellement `arping`.

Echec de toutes les tentatives de résolution des incidents

Si vous n'arrivez pas à résoudre des incidents de routage et que toutes vos tentatives ont échoué, émettez la commande suivante pour lancer une trace du trafic réseau :

- Sur les systèmes AIX, à partir de la machine Load Balancer :
`iptrace -a -s adresse_IP_Client_EnEchec -d adresse_IP_cluster -b iptrace.trc`

Exécutez la trace, recréez l'incident, puis tuez le processus.

- Sous HP-UX :
`tcpdump -i lan0 host cluster et host client`

Vous devrez peut-être télécharger tcpdump de l'un des sites d'archivage de logiciels HP-UX GNU.

- Sur les systèmes Linux :
`tcpdump -i eth0 host cluster et host client`

Exécutez la trace, recréez l'incident, puis tuez le processus.

- Sous Solaris :
`snoop -v adresse_IP_client adresse_IP_destination > snooptrace.out`
- Sur les systèmes Windows, un "renifleur" est nécessaire. Utilisez les mêmes entrées que pour un filtre.

Vous pouvez également augmenter les niveaux de différents journaux (par exemple, journal du gestionnaire, journal du conseiller, etc.) et analyser les informations qu'ils contiennent.

Mises à niveau

Pour identifier un incident déjà résolu dans un correctif Service Release, recherchez les mises à niveau disponibles. Pour obtenir la liste des défauts Edge Components corrigés, reportez-vous à la page Web d'assistance de WebSphere Application Server : <http://www.ibm.com/software/webservers/appserv/was/support/>. À partir de cette page, cliquez sur le lien permettant d'accéder au site de téléchargement des correctifs.

Code Java

La version correcte du code Java est installée lors de l'installation de Load Balancer.

Liens utiles

Pour accéder aux liens des pages Web du support et de la bibliothèque, voir «Informations connexes», à la page xvii. La page de support Web contient un lien à l'aide sous la forme de notes techniques.

Tableaux de résolution des incidents

Pour les opérations répertoriées, reportez-vous au tableau indiqué.

- Informations de résolution des incidents liés à Dispatcher — tableau 14, à la page 286
- Informations de résolution des incidents liés à CBR — tableau 15, à la page 292
- Informations de résolution des incidents liés à Site Selector — tableau 16, à la page 293

- Informations de résolution des incidents liés à Cisco CSS Controller — tableau 17, à la page 295
- Informations de résolution des incidents liés à Nortel Alteon Controller — tableau 18, à la page 296
- Informations de résolution des incidents liés à Metric Server — tableau 19, à la page 297

Tableau 14. Tableau de résolution des incidents de Dispatcher

Symptôme	Cause possible	Voir...
Dispatcher ne fonctionne pas correctement	Conflit de numéros de port	«Vérification des numéros de port Dispatcher», à la page 299
Le serveur configuré ne répond pas aux requêtes d'équilibrage de charge	Conflit d'adresses ou adresse erronée	«Incident : Le répartiteur et le serveur ne répondent pas», à la page 302
Absence de prise en charge des connexions des machines client ou dépassement de délai des connexions	<ul style="list-style-type: none"> • Mauvaise configuration de réacheminement • NIC sans alias avec l'adresse de cluster • Le serveur n'a pas d'unité de bouclage ayant un alias pour l'adresse de cluster • Le chemin supplémentaire n'est pas supprimé • Le port n'est pas défini pour chaque cluster 	«Incident : Les requêtes Dispatcher ne sont pas équilibrées», à la page 302
Les machines client ne sont pas prises en charge ou le délai imparti à ces connexions est dépassé	La fonction haute disponibilité est inopérante	«Incident : La fonction haute disponibilité de Dispatcher est inopérante», à la page 303
Impossible d'ajouter un signal de présence (plateforme Windows)	L'adresse source n'est pas configurée sur un adaptateur	«Incident : Impossible d'ajouter un signal de présence (plateforme Windows)», à la page 303
Le serveur ne livre pas les requêtes (plateforme Windows)	Une route supplémentaire a été créée dans la table de routage	«Incident : Routes supplémentaires (Windows 2000)», à la page 303
Les conseillers ne fonctionnent pas correctement en réseau étendu	Les conseillers ne fonctionnent pas sur les machines éloignées	«Incident : Les conseillers ne fonctionnent pas correctement», à la page 304
Dispatcher, Microsoft IIS et SSL ne fonctionnent pas ou risquent de s'arrêter	Impossible d'envoyer des données codées via les protocoles	«Incident : Dispatcher, Microsoft IIS et SSL ne fonctionnent pas (plateforme Windows)», à la page 304
Connexion à une machine distante refusée	Une ancienne version des clés est encore utilisée	«Incident : Connexion du répartiteur à une machine éloignée», à la page 304

Tableau 14. Tableau de résolution des incidents de Dispatcher (suite)

Symptôme	Cause possible	Voir...
La commande dscontrol ou lbadmin n'a pas abouti, le message 'Le serveur ne répond pas' ou 'Impossible d'accéder au serveur RMI' s'affiche	<ol style="list-style-type: none"> 1. Echec des commandes en raison d'une pile mise sur "sock". Ou les commandes n'ont pas abouti car dscontrol n'a pas été lancé. 2. La définition des ports RMI est incorrecte. 3. Un hôte local est incorrect dans le fichier hôte 	«Incident : La commande dscontrol ou lbadmin n'a pas abouti», à la page 304
Message d'erreur "Impossible de trouver le fichier...", lors de l'exécution de Netscape en tant que navigateur par défaut pour visualiser l'aide en ligne (plateforme Windows)	Paramétrage incorrect pour l'association de fichier HTML	«Incident : Affichage du message d'erreur "Fichier introuvable..." lorsque vous tentez de visualiser l'aide en ligne (plateforme Windows)», à la page 305
L'interface graphique ne démarre pas correctement	Espace de pagination insuffisant	«Incident : L'interface graphique ne démarre pas correctement», à la page 305
Erreur lors de l'exécution de Dispatcher lorsque Caching Proxy est installé	Dépendance de fichiers Caching Proxy	«Incident : Erreur lors de l'exécution de Dispatcher lorsque Caching Proxy est installé», à la page 305
L'interface utilisateur graphique ne s'affiche pas correctement.	La résolution est incorrecte.	«Incident : L'interface graphique ne s'affiche pas correctement», à la page 306
Les panneaux d'aide apparaissent parfois sous d'autres fenêtres	Restriction Java	«Incident : Sous Windows, les fenêtres d'aide disparaissent parfois sous d'autres fenêtres ouvertes», à la page 306
Load Balancer ne peut pas traiter et transmettre de cadre	Une adresse MAC unique est nécessaire pour chaque carte NIC	«Incident: Load Balancer ne peut pas traiter et transmettre un cadre», à la page 306
Un écran bleu apparaît	Aucune carte réseau n'est installée et configurée	«Incident : Un écran bleu s'affiche lors du démarrage de l'exécuteur Load Balancer», à la page 306
La fonction Path MTU Discovery permet d'éviter le trafic retour	Le cluster est associé à un alias sur l'unité de bouclage	«Incident : La fonction Path MTU Discovery permet d'éviter le trafic retour avec Load Balancer», à la page 306
La fonction haute disponibilité de Load Balancer en mode réseau étendu est inopérante	La machine Dispatcher éloignée doit être définie en tant que serveur d'un cluster sur la machine Dispatcher locale	«Incident : La fonction haute disponibilité de Load Balancer en mode réseau étendu est inopérante», à la page 307

Tableau 14. Tableau de résolution des incidents de Dispatcher (suite)

Symptôme	Cause possible	Voir...
Arrêt (ou comportement imprévu) de l'interface graphique lors de la tentative de chargement d'un fichier de configuration volumineux.	La mémoire est insuffisante pour permettre à Java de traiter une modification de l'interface graphique de cette ampleur	«Incident : Arrêt (ou comportement imprévu) de l'interface graphique lors de la tentative de chargement d'un fichier de configuration volumineux», à la page 308
Adresses IP non résolues correctement sur la connexion éloignée	Lors de l'utilisation d'un client éloigné sur une implémentation SSL, des noms d'hôtes ou des noms de domaines complets ne sont pas correctement convertis en adresses IP	«Incident : Adresses IP non résolues correctement sur la connexion éloignée», à la page 309
L'interface coréenne de Load Balancer affiche sous AIX et Linux des polices non souhaitées ou qui se chevauchent	Vous devez modifier les polices de caractères par défaut	«Incident : L'interface coréenne de Load Balancer affiche sous AIX et Linux des polices non souhaitées ou qui se chevauchent», à la page 309
Sous Windows, lorsqu'un alias a été attribué à l'unité de bouclage de MS, le système d'exploitation ne répond pas correctement à l'adresse d'alias lors de l'émission d'une commande telle que hostname	Dans la liste des connexions réseau, le nouvel alias ne doit pas se trouver au-dessus de l'adresse locale	«Incident : Sous Windows, adresse d'alias renvoyée au lieu de l'adresse locale lors de l'émission de commandes telles que hostname», à la page 310
Comportement inattendu de l'interface graphique lors de l'utilisation de Windows avec une carte vidéo Matrox AGP	Incident lors de l'utilisation de cartes Matrox AGP en cours d'exécution de l'interface graphique de Load Balancer	«Incident : sous Windows, comportement inattendu de l'interface graphique lors de l'utilisation de cartes vidéo Matrox AGP», à la page 310
Comportement inattendu, tel un arrêt du système, lors de l'exécution de "rmmod ibmlb" sous Linux	Incident lors du retrait manuel du noyau du module Load Balancer (ibmlb).	«Incident : Comportement inattendu lors de l'exécution de "rmmod ibmlb" (systèmes Linux)», à la page 310
Temps de réponse important lors de l'exécution de commandes sur la machine Dispatcher	Un temps de réponse important peut être dû à une surcharge de la machine liée à un volume élevé de trafic client	«Incident : Temps de réponse important lors de l'exécution de commandes sur la machine Dispatcher», à la page 310
Pour la méthode d'acheminement MAC de Dispatcher, le conseiller SSL ou HTTPS n'enregistre pas les charges des serveurs	Incident lié au fait que l'application serveur SSL n'est pas configurée avec l'adresse IP du cluster	«Incident : Le conseiller SSL ou HTTPS n'enregistre pas les charges des serveurs (avec l'acheminement MAC)», à la page 311
Déconnexion de l'hôte lors de l'administration Web à distance via Netscape	Cette déconnexion se produit lors du redimensionnement de la fenêtre du navigateur	«Incident : Déconnexion de l'hôte lors du redimensionnement de la fenêtre du navigateur Netscape en cours d'administration Web», à la page 311

Tableau 14. Tableau de résolution des incidents de Dispatcher (suite)

Symptôme	Cause possible	Voir...
Regroupement de connexions activé et serveur Web établissant une liaison à 0.0.0.0	Configurez le serveur Microsoft IIS en tant que serveur de liaison	«Incident : Regroupement de connexions activé et serveur Web établissant une liaison à 0.0.0.0», à la page 311
Sur la plateforme Windows, des caractères nationaux Latin-1 endommagés apparaissent sur la ligne de commande	Modifiez les propriétés des polices de la fenêtre de la ligne de commande	«Incident : Sous Windows, des caractères nationaux Latin-1 endommagés apparaissent dans la fenêtre de la ligne de commande», à la page 312
Sur la plateforme HP-UX, le message suivant est généré : java.lang.OutOfMemoryError unable to create new native thread	Certaines installations HP-UX autorisent par défaut 64 unités d'exécution par processus. Cela est suffisant.	«Incident : Sous HP-UX, la mémoire est insuffisante pour Java et une erreur d'unité d'exécution est générée», à la page 312
Sur la plateforme Windows, les conseillers et les cibles à contacter marquent tous les serveurs comme étant arrêtés	Le déchargement des tâches n'est pas désactivé ou il doit activer ICMP.	«Incident : Sous Windows, les conseillers et les cibles à contacter marquent tous les serveurs comme étant arrêtés», à la page 313
Sur la plateforme Windows, la résolution de l'adresse IP en nom d'hôte n'est pas possible lorsque plusieurs adresses sont configurées sur un adaptateur	L'adresse IP que vous voulez comme nom d'hôte doit d'abord apparaître dans le registre.	«Incident : Sous Windows, la résolution de l'adresse IP en nom d'hôte n'est pas possible lorsque plusieurs adresses sont configurées sur un adaptateur», à la page 313
Sur la plateforme Windows, les conseillers ne fonctionnent pas dans une configuration en haute disponibilité après une panne réseau	Lorsque le système détecte une panne réseau, il efface sa mémoire cache ARP (Address Resolution Protocol)	«Incident : Sous Windows, les conseillers ne fonctionnent pas dans une configuration en haute disponibilité après une panne réseau», à la page 314
Sur les systèmes Linux, la commande "IP address add" et les alias de bouclage de cluster multiples sont incompatibles	Lorsque vous attribuez des alias à plusieurs adresses sur l'unité de bouclage, vous devez utiliser la commande ifconfig et non la commande ip address add	«Incident : Sous Linux, n'utilisez pas la commande "IP address add" lors de l'affectation d'alias à plusieurs clusters de l'unité de bouclage», à la page 315
Message d'erreur : "Adresse de routeur non spécifiée ou non valide pour la méthode port" lors de la tentative d'ajout d'un serveur	Liste de contrôle permettant d'identifier l'incident qui s'est produit lors de l'ajout d'un serveur	«Incident : Message d'erreur "Adresse de routeur non spécifiée ou non valide pour la méthode port", à la page 315
Sur les systèmes Solaris, les processus Load Balancer s'arrêtent lorsque vous quittez la fenêtre de session de terminal à partir de laquelle ils ont été lancés	Utilisez la commande nohup afin que les processus lancés ne reçoivent pas un signal d'arrêt lorsque vous quittez la session de terminal.	«Incident : Sur les systèmes Solaris, les processus Load Balancer s'arrêtent lorsque vous quittez la fenêtre de terminal à partir de laquelle ils ont été lancés», à la page 316

Tableau 14. Tableau de résolution des incidents de Dispatcher (suite)

Symptôme	Cause possible	Voir...
Un ralentissement se produit lors du chargement des configurations Load Balancer	Ce délai peut provenir des appels DNS effectués pour résoudre et vérifier l'adresse du serveur.	«Incident : Délai lors du chargement d'une configuration Load Balancer», à la page 316
Sur les systèmes Windows, un message d'erreur s'affiche pour indiquer qu'il existe un conflit d'adresses IP avec un autre système du réseau	Si la fonction de haute disponibilité est configurée, il est possible que des adresses de cluster soient définies sur les deux systèmes pendant une courte période et qu'elles génèrent ce message d'erreur.	«Incident : Sur les systèmes Windows, un message d'erreur lié à un conflit d'adresses IP apparaît à l'écran», à la page 316
Les machines principale et de secours sont toutes deux activées en mode haute disponibilité.	Cet incident peut survenir lorsque les scripts go ne sont pas exécutés sur la machine principale ou la machine de secours.	«Incident : les machines principale et de secours sont toutes deux activées en mode haute disponibilité», à la page 317
Les demandes client échouent lorsque Dispatcher tente de renvoyer des réponses de grande page	Les demandes client générant des réponses de grande page arrivent à expiration si la taille maximale en transmission (MTU) n'est pas définie correctement sur la machine Dispatcher lors de l'utilisation de la méthode de transfert nat ou cbr.	«Incident : Les demandes client échouent lors de la tentative de renvoi de réponses de grande page», à la page 317
Sous Windows, l'erreur "Le serveur ne répond pas" survient lors de l'exécution d'une commande dscontrol ou lbadmin	Lorsqu'il existe plusieurs adresses IP sur un système Windows et que le fichier hôte ne spécifie pas l'adresse à associer au nom d'hôte.	«Incident : Sous Windows, l'erreur "Le serveur ne répond pas" survient lors de l'exécution d'une commande dscontrol ou lbadmin», à la page 318
Les machines Dispatcher à haute disponibilité risquent de ne pas être synchronisées sous Linux pour S/390 avec des périphériques qeth	Lors de l'utilisation de la haute disponibilité sur des systèmes Linux S/390 avec le pilote réseau qeth, la synchronisation des machines Dispatcher active et de secours risque d'échouer.	«Incident : Les machines Dispatcher à haute disponibilité risquent de ne pas être synchronisées sur les systèmes Linux pour S/390 avec des pilotes qeth», à la page 318
Conseils pour la configuration de la fonction de haute disponibilité pour Load Balancer	Ces conseils visent à réduire les incidents liés à la haute disponibilité tels que : <ul style="list-style-type: none"> • suppression des connexions après la reprise, • synchronisation impossible des machines partenaires, • demandes dirigées à tort vers la machine partenaire de secours. 	«Incident : Conseils sur la configuration de la haute disponibilité», à la page 318

Tableau 14. Tableau de résolution des incidents de Dispatcher (suite)

Symptôme	Cause possible	Voir...
Limitations de la configuration de l'acheminement MAC pour Dispatcher avec les plateformes zSeries et S/390	Sous Linux, il existe des limitations lorsque vous utilisez des serveurs zSeries ou S/390 dotés de cartes OSA (Open System Adapter). Des solutions alternatives sont fournies.	«Incident : Sous Linux, limitations de la configuration Dispatcher lors de l'utilisation de serveurs zSeries ou S/390 dotés de cartes OSA (Open System Adapter)», à la page 320
Sur certaines versions de Red Hat Linux, des fuites de mémoire se produisent en cas d'exécution de Load Balancer configuré avec le gestionnaire et les conseillers	Les versions IBM Java SDK de JVM et la bibliothèque NPTL (Native POSIX Thread Library) livrée avec certaines distributions de Linux, comme Red Hat Enterprise Linux 3.0, peuvent être à l'origine de la fuite de mémoire.	«Incident : Sur certaines versions Linux, une fuite de mémoire se produit lors de l'exécution de Dispatcher configuré avec le gestionnaire et les conseillers», à la page 322
Sur SUSE Linux Enterprise Server 9, un rapport Dispatcher report indique que les paquets sont acheminés (le nombre de paquets augmente), alors que les paquets n'atteignent jamais les serveur dorsal	Le module NAT iptables est chargé. Cette version d'iptables contient une erreur possible, bien que non confirmée, entraînant un comportement étrange lors des interactions avec Dispatcher.	«Incident : Sur SUSE Linux Enterprise Server 9, Dispatcher achemine les paquets, mais ceux-ci n'arrivent pas jusqu'au serveur dorsal», à la page 322
Sur les systèmes Windows, lors de l'utilisation de la fonction de haute disponibilité de Dispatcher, des incidents peuvent se produire lors du relais.	Si le goScript configurant l'adresse IP de cluster sur la machine active s'exécute avant celui qui annule l'adresse IP de cluster sur la machine de secours, des incidents risquent de se produire.	«Incident : Sur le système Windows, un message de conflit d'adresses IP apparaît pendant la reprise de la haute disponibilité», à la page 323
Sur les systèmes Linux, des iptables peuvent interférer avec le routage de paquets	Les iptables Linux peuvent interférer avec l'équilibrage de charge du trafic et doivent être désactivées sur la machine Load Balancer.	«Incident : Les iptables de Linux peuvent interférer avec le routage de paquets», à la page 324
Sur les systèmes Solaris, lorsque vous essayez de configurer un serveur IPv6 sur la machine Dispatcher, le message "Impossible d'ajouter le serveur" s'affiche	Cette erreur peut provenir de la gestion de la requête ping sur une adresse IPv6 par le système d'exploitation Solaris.	«Incident : Impossible d'ajouter un serveur IPv6 à la configuration Load Balancer sur les systèmes Solaris», à la page 324

Tableau 14. Tableau de résolution des incidents de Dispatcher (suite)

Symptôme	Cause possible	Voir...
Un message d'avertissement sur un ensemble de fichiers Java s'affiche lors de l'installation de correctifs de service ou lors de l'installation en mode natif, à l'aide des outils de création de packages du système.	L'installation du produit comporte plusieurs packages qu'il n'est pas nécessaire d'installer sur la même machine, donc chacun de ces packages installe un ensemble de fichiers Java. En cas d'installation sur la même machine, un message d'avertissement indique que l'ensemble de fichiers Java appartient également à un autre ensemble de fichiers.	«Un message d'avertissement Java s'affiche lors de l'installation de correctifs de service», à la page 324
Mise à niveau de l'ensemble de fichiers Java fourni avec les installations Load Balancer	En cas d'incident au niveau de l'ensemble de fichiers Java, contactez le service d'assistance IBM pour recevoir une mise à niveau de l'ensemble de fichiers Java qui a été fourni avec l'installation Load Balancer.	«Mise à niveau de l'ensemble de fichiers Java fourni avec l'installation Load Balancer», à la page 325

Tableau 15. Tableau de résolution des incidents de CBR

Symptôme	Cause possible	Voir.
CBR ne fonctionne pas correctement	Conflit de numéros de port	«Vérification des numéros de port CBR», à la page 299
La commande cbrcontrol ou lbadmin n'a pas abouti, le message 'Le serveur ne répond pas' ou 'Impossible d'accéder au serveur RMI' s'affiche	Echec des commandes en raison d'une pile mise sur "sock". Ou les commandes n'ont pas abouti car cbrserver n'a pas été lancé	«Incident : La commande cbrcontrol ou lbadmin n'a pas abouti», à la page 325
La charge des demandes n'est pas équilibrée	Caching Proxy a été lancé avant l'exécuteur	«Incident : Les requêtes ne sont pas équilibrées», à la page 326
Sous Solaris, la commande cbrcontrol executor start renvoie le message suivant : 'Erreur : l'exécuteur n'a pas été lancé'.	La commande peut échouer lorsqu'une modification des valeurs IPC système par défaut est nécessaire ou que le lien permettant d'accéder à la bibliothèque est incorrect.	«Incident : Sur les systèmes Solaris, la commande cbrcontrol executor start n'aboutit pas», à la page 326
La règle d'URL ne fonctionne pas	Erreur de syntaxe ou de configuration	«Incident : erreur de syntaxe ou de configuration», à la page 326
Comportement inattendu de l'interface graphique lors de l'utilisation de systèmes Windows avec une carte vidéo Matrox AGP	Incident lors de l'utilisation de cartes Matrox AGP en cours d'exécution de l'interface graphique de Load Balancer	«Incident : sous Windows, comportement inattendu de l'interface graphique lors de l'utilisation de cartes vidéo Matrox AGP», à la page 326

Tableau 15. Tableau de résolution des incidents de CBR (suite)

Arrêt (ou comportement imprévu) de l'interface graphique lors de la tentative de chargement d'un fichier de configuration volumineux.	La mémoire est insuffisante pour permettre à Java de traiter une modification de l'interface graphique de cette ampleur	«Incident : Arrêt (ou comportement imprévu) de l'interface graphique lors de la tentative de chargement d'un fichier de configuration volumineux», à la page 308
Déconnexion de l'hôte lors de l'administration Web à distance via Netscape	Cette déconnexion se produit lors du redimensionnement de la fenêtre du navigateur	«Incident : Déconnexion de l'hôte lors du redimensionnement de la fenêtre du navigateur Netscape en cours d'administration Web», à la page 327
Sur la plateforme Windows, des caractères nationaux Latin-1 endommagés apparaissent sur la ligne de commande	Modifiez les propriétés des polices de la fenêtre de la ligne de commande	«Incident : Sous Windows, des caractères nationaux Latin-1 endommagés apparaissent dans la fenêtre de la ligne de commande», à la page 327
Sur la plateforme HP-UX, le message suivant est généré : <code>java.lang.OutOfMemoryError</code> unable to create new native thread	Certaines installations HP-UX autorisent par défaut 64 unités d'exécution par processus. Cela est suffisant.	«Incident : Sous HP-UX, la mémoire est insuffisante pour Java et une erreur d'unité d'exécution est générée», à la page 327
Sur la plateforme Windows, les conseillers et les cibles à contacter marquent tous les serveurs comme étant arrêtés	Le déchargement des tâches n'est pas désactivé ou il doit activer icmp.	«Incident : Sous Windows, les conseillers et les cibles à contacter marquent tous les serveurs comme étant arrêtés», à la page 327
Sur la plateforme Windows, la résolution de l'adresse IP en nom d'hôte n'est pas possible lorsque plusieurs adresses sont configurées sur un adaptateur	L'adresse IP que vous voulez comme nom d'hôte doit d'abord apparaître dans le registre.	«Incident : Sur les systèmes Windows, la résolution de l'adresse IP en nom d'hôte n'est pas possible lorsque plusieurs adresses sont configurées sur un adaptateur», à la page 328
Sur les systèmes Solaris, les processus Load Balancer s'arrêtent lorsque vous quittez la fenêtre de session de terminal à partir de laquelle ils ont été lancés	Utilisez la commande nohup afin que les processus lancés ne reçoivent pas un signal d'arrêt lorsque vous quittez la session de terminal.	«Incident : Sur les systèmes Solaris, les processus Load Balancer s'arrêtent lorsque vous quittez la fenêtre de terminal à partir de laquelle ils ont été lancés», à la page 316

Tableau 16. Tableau de résolution des incidents de Site Selector

Symptôme	Cause possible	Voir...
Site Selector ne s'exécute pas correctement	Conflit de numéros de port	«Vérification des numéros de port Site Selector», à la page 300
Site Selector n'effectue pas de demandes entrantes permutées de façon circulaire à partir des clients Solaris	Les système Solaris exécutent un "démon de mémoire cache de service annuaire"	«Incident : Site Selector ne permet pas le trafic à permutation circulaire à partir des clients Solaris», à la page 328

Tableau 16. Tableau de résolution des incidents de Site Selector (suite)

Symptôme	Cause possible	Voir...
La commande sscontrol ou lbadmin n'a pas abouti, le message 'Le serveur ne répond pas' ou 'Impossible d'accéder au serveur RMI' s'affiche	Echec des commandes en raison d'une pile mise sur "sock". Ou les commandes n'ont pas abouti car sserver n'a pas été lancé.	«Incident : la commande sscontrol ou lbadmin n'a pas abouti», à la page 328
Echec du démarrage de sserver sous Windows	Les systèmes Windows ne nécessitent pas toujours la présence du nom d'hôte dans le système DNS.	«Incident : Echec du démarrage de sserver sous Windows», à la page 329
Machine ayant des chemins en double pour lequel l'équilibrage de charge ne s'effectue pas correctement — la résolution de noms semble ne pas aboutir	Machine Site Selector ayant plusieurs cartes associées au même sous-réseau	«Incident : Site Selector ayant des chemins en double pour lequel l'équilibrage de charge ne s'effectue pas correctement», à la page 329
Comportement inattendu de l'interface graphique lors de l'utilisation de Windows avec une carte vidéo Matrox AGP	Incident lors de l'utilisation de cartes Matrox AGP en cours d'exécution de l'interface graphique de Load Balancer	«Incident : sous Windows, comportement inattendu de l'interface graphique lors de l'utilisation de cartes vidéo Matrox AGP», à la page 329
Arrêt (ou comportement imprévu) de l'interface graphique lors de la tentative de chargement d'un fichier de configuration volumineux.	La mémoire est insuffisante pour permettre à Java de traiter une modification de l'interface graphique de cette ampleur	«Incident : Arrêt (ou comportement imprévu) de l'interface graphique lors de la tentative de chargement d'un fichier de configuration volumineux», à la page 308
Déconnexion de l'hôte lors de l'administration Web à distance via Netscape	Cette déconnexion se produit lors du redimensionnement de la fenêtre du navigateur	«Incident : Déconnexion de l'hôte lors du redimensionnement de la fenêtre du navigateur Netscape en cours d'administration Web», à la page 330
Sur la plateforme Windows, des caractères nationaux Latin-1 endommagés apparaissent sur la ligne de commande	Modifiez les propriétés des polices de la fenêtre de la ligne de commande	«Incident : Sous Windows, des caractères nationaux Latin-1 endommagés apparaissent dans la fenêtre de la ligne de commande», à la page 330
Sur la plateforme HP-UX, le message suivant est généré : java.lang.OutOfMemoryError unable to create new native thread	Certaines installations HP-UX autorisent par défaut 64 unités d'exécution par processus. Cela est suffisant.	«Incident : Sous HP-UX, la mémoire est insuffisante pour Java et une erreur d'unité d'exécution est générée», à la page 330
Sur la plateforme Windows, les conseillers et les cibles à contacter marquent tous les serveurs comme étant arrêtés	Le téléchargement des tâches n'est pas désactivé ou il doit activer icmp.	«Incident : Sous Windows, les conseillers et les cibles à contacter marquent tous les serveurs comme étant arrêtés», à la page 330

Tableau 16. Tableau de résolution des incidents de Site Selector (suite)

Symptôme	Cause possible	Voir...
Sur les systèmes Solaris, les processus Load Balancer s'arrêtent lorsque vous quittez la fenêtre de session de terminal à partir de laquelle ils ont été lancés	Utilisez la commande nohup afin que les processus lancés ne reçoivent pas un signal d'arrêt lorsque vous quittez la session de terminal.	«Incident : Sur les systèmes Solaris, les processus Load Balancer s'arrêtent lorsque vous quittez la fenêtre de terminal à partir de laquelle ils ont été lancés», à la page 316

Tableau 17. Tableau de résolution des incidents de Contrôleur pour commutateurs Cisco CSS

Symptôme	Cause possible	Voir...
échec du lancement de ccoserver	Conflit de numéros de port	«Vérification des numéros de port Cisco CSS Controller», à la page 301
La commande ccocontrol ou lbadmin n'a pas abouti, le message 'Le serveur ne répond pas' ou 'Impossible d'accéder au serveur RMI' s'affiche	Echec des commandes en raison d'une pile mise sur "sock". Ou les commandes n'aboutissent pas car ccoserver n'a pas été lancé.	«Incident : La commande ccocontrol ou lbadmin n'a pas abouti», à la page 331
Erreur de réception : Impossible de créer un registre sur le port 13099	Licence du produit expirée	«Incident : Impossible de créer un registre sur le port 13099», à la page 331
Comportement inattendu de l'interface graphique lors de l'utilisation de Windows avec une carte vidéo Matrox AGP	Incident lors de l'utilisation de cartes Matrox AGP en cours d'exécution de l'interface graphique de Load Balancer	«Incident : sous Windows, comportement inattendu de l'interface graphique lors de l'utilisation de cartes vidéo Matrox AGP», à la page 332
Réception d'une erreur de connexion lors de l'ajout d'un consultant	Paramètres de configuration incorrects sur le commutateur ou le contrôleur	«Incident : Réception d'une erreur de connexion lors de l'ajout d'un consultant», à la page 332
Pondérations non actualisées sur le commutateur	Communication entre le contrôleur et le commutateur impossible ou interrompue	«Incident : Pondérations non actualisées sur le commutateur», à la page 332
La commande de régénération n'a pas actualisé la configuration du consultant	Communication entre le commutateur et le contrôleur impossible ou interrompue	«Incident : La commande de régénération n'a pas actualisé la configuration du consultant», à la page 332
Arrêt (ou comportement imprévu) de l'interface graphique lors de la tentative de chargement d'un fichier de configuration volumineux.	La mémoire est insuffisante pour permettre à Java de traiter une modification de l'interface graphique de cette ampleur	«Incident : Arrêt (ou comportement imprévu) de l'interface graphique lors de la tentative de chargement d'un fichier de configuration volumineux», à la page 308
Déconnexion de l'hôte lors de l'administration Web à distance via Netscape	Cette déconnexion se produit lors du redimensionnement de la fenêtre du navigateur	«Incident : Déconnexion de l'hôte lors du redimensionnement de la fenêtre du navigateur Netscape en cours d'administration Web», à la page 333

Tableau 17. Tableau de résolution des incidents de Contrôleur pour commutateurs Cisco CSS (suite)

Symptôme	Cause possible	Voir...
Sur la plateforme Windows, des caractères nationaux Latin-1 endommagés apparaissent sur la ligne de commande	Modifiez les propriétés des polices de la fenêtre de la ligne de commande	«Incident : Sous Windows, des caractères nationaux Latin-1 endommagés apparaissent dans la fenêtre de la ligne de commande», à la page 333
Sur la plateforme HP-UX, le message suivant est généré : java.lang.OutOfMemoryError unable to create new native thread	Certaines installations HP-UX autorisent par défaut 64 unités d'exécution par processus. Cela est suffisant.	«Incident : Sous HP-UX, la mémoire est insuffisante pour Java et une erreur d'unité d'exécution est générée», à la page 333
Sur les systèmes Solaris, les processus Load Balancer s'arrêtent lorsque vous quittez la fenêtre de session de terminal à partir de laquelle ils ont été lancés	Utilisez la commande nohup afin que les processus lancés ne reçoivent pas un signal d'arrêt lorsque vous quittez la session de terminal.	«Incident : Sur les systèmes Solaris, les processus Load Balancer s'arrêtent lorsque vous quittez la fenêtre de terminal à partir de laquelle ils ont été lancés», à la page 316

Tableau 18. Tableau de résolution des incidents de Nortel Alteon Controller

Symptôme	Cause possible	Voir...
échec du lancement de nalserver	Conflit de numéros de port	«Vérification des numéros de port Nortel Alteon Controller», à la page 301
La commande nalcontrol ou lbadm n'a pas abouti, le message 'Le serveur ne répond pas' ou 'Impossible d'accéder au serveur RMI' s'affiche	Echec des commandes en raison d'une pile mise sur "sock". Ou les commandes n'ont pas abouti car nalserver n'a pas été lancé.	«Incident : la commande nalcontrol ou lbadm n'a pas abouti», à la page 333
Erreur de réception : Impossible de créer un registre sur le port 14099	Licence du produit expirée	«Incident : Impossible de créer un registre sur le port 14099», à la page 334
Comportement inattendu de l'interface graphique lors de l'utilisation de Windows avec une carte vidéo Matrox AGP	Incident lors de l'utilisation de cartes Matrox AGP en cours d'exécution de l'interface graphique de Load Balancer	«Incident : sous Windows, comportement inattendu de l'interface graphique lors de l'utilisation de cartes vidéo Matrox AGP», à la page 334
Arrêt (ou comportement imprévu) de l'interface graphique lors de la tentative de chargement d'un fichier de configuration volumineux.	La mémoire est insuffisante pour permettre à Java de traiter une modification de l'interface graphique de cette ampleur	«Incident : Arrêt (ou comportement imprévu) de l'interface graphique lors de la tentative de chargement d'un fichier de configuration volumineux», à la page 308
Déconnexion de l'hôte lors de l'administration Web à distance via Netscape	Cette déconnexion se produit lors du redimensionnement de la fenêtre du navigateur	«Incident : Déconnexion de l'hôte lors du redimensionnement de la fenêtre du navigateur Netscape en cours d'administration Web», à la page 335

Tableau 18. Tableau de résolution des incidents de Nortel Alteon Controller (suite)

Symptôme	Cause possible	Voir...
Réception d'une erreur de connexion lors de l'ajout d'un consultant	Paramètres de configuration incorrects sur le commutateur ou le contrôleur	«Incident : Réception d'une erreur de connexion lors de l'ajout d'un consultant», à la page 335
Pondérations non actualisées sur le commutateur	Communication entre le contrôleur et le commutateur impossible ou interrompue	«Incident : Pondérations non actualisées sur le commutateur», à la page 335
La commande de régénération n'a pas actualisé la configuration du consultant	Communication entre le commutateur et le contrôleur impossible ou interrompue	«Incident : La commande de régénération n'a pas actualisé la configuration du consultant», à la page 335
Sur la plateforme Windows, des caractères nationaux Latin-1 endommagés apparaissent sur la ligne de commande	Modifiez les propriétés des polices de la fenêtre de la ligne de commande	«Incident : Sous Windows, des caractères nationaux Latin-1 endommagés apparaissent dans la fenêtre de la ligne de commande», à la page 335
Sur la plateforme HP-UX, le message suivant est généré : java.lang.OutOfMemoryError unable to create new native thread	Certaines installations HP-UX autorisent par défaut 64 unités d'exécution par processus. Cela est suffisant.	«Incident : Sous HP-UX, la mémoire est insuffisante pour Java et une erreur d'unité d'exécution est générée», à la page 336
Sur les systèmes Solaris, les processus Load Balancer s'arrêtent lorsque vous quittez la fenêtre de session de terminal à partir de laquelle ils ont été lancés	Utilisez la commande nohup afin que les processus lancés ne reçoivent pas un signal d'arrêt lorsque vous quittez la session de terminal.	«Incident : Sur les systèmes Solaris, les processus Load Balancer s'arrêtent lorsque vous quittez la fenêtre de terminal à partir de laquelle ils ont été lancés», à la page 316

Tableau 19. Tableau de dépannage du système Metric Server

Symptôme	Cause possible	Voir...
IOException Metric Server sous Windows lors de l'exécution de fichiers de mesures utilisateur de format BAT ou CMD	Le nom complet des mesures est obligatoire	«Incident : IOException Metric Server sous Windows lors de l'exécution de fichiers de mesures utilisateur de format .bat or .cmd», à la page 336
Le système Metric Server ne fournit pas à la machine Load Balancer les informations relatives à la charge.	Causes possibles : <ul style="list-style-type: none"> absence de fichier de clés sur la machine Metric Server nom d'hôte de la machine Metric Server non enregistré sur le serveur de noms local nom d'hôte local du fichier /etc/hosts converti à l'adresse de bouclage 127.0.0.1 	«Incident : Metric Server n'indique pas la charge à la machine Load Balancer», à la page 336

Tableau 19. Tableau de dépannage du système Metric Server (suite)

Symptôme	Cause possible	Voir...
L'entrée "Signature is necessary for access to agent" (signature nécessaire pour accéder à l'agent) apparaît dans le journal de la machine Metric Server lors du transfert des fichiers de clés vers le serveur	Echec de l'autorisation du fichier de clés en raison d'une altération.	«Incident : Le journal de la machine Metric Server indique qu'une signature est nécessaire pour accéder à l'agent», à la page 337
Sur les systèmes AIX, lorsque Metric Server s'exécute dans des conditions difficiles sur un système multiprocesseur (AIX 4.3.3 ou AIX 5.1), il est possible que le résultat de la commande ps -vg soit altéré	Le correctif APAR IY33804 rectifie cet incident AIX identifié	«Incident : Sur les systèmes AIX, lorsque Metric Server s'exécute dans des conditions difficiles, il est possible que le résultat de la commande ps -vg soit altéré», à la page 337
Configuration de Metric Server dans une configuration de second niveau avec équilibrage de la charge entre des machines Dispatcher haute disponibilité par Site Selector	Metric Server (dans une configuration de second niveau) n'est pas configuré pour écouter une nouvelle adresse IP.	«Incident : Configuration de Metric Server dans une configuration de second niveau avec équilibrage de la charge entre des machines Dispatcher haute disponibilité par Site Selector», à la page 337
Les scripts (metricserver, cpuload, memload) exécutés sur des machines Solaris dotées de plusieurs CPU génèrent des messages de console non souhaités	Ce comportement est dû à l'utilisation de la commande système VMSTAT pour collecter des statistiques sur la CPU et la mémoire à partir du noyau.	«Incident : Les scripts exécutés sur des machines Solaris dotées de plusieurs CPU génèrent des messages de console non souhaités», à la page 339
Sur les systèmes Solaris, les processus Load Balancer s'arrêtent lorsque vous quittez la fenêtre de session de terminal à partir de laquelle ils ont été lancés	Utilisez la commande nohup afin que les processus lancés ne reçoivent pas un signal d'arrêt lorsque vous quittez la session de terminal.	«Incident : Sur les systèmes Solaris, les processus Load Balancer s'arrêtent lorsque vous quittez la fenêtre de terminal à partir de laquelle ils ont été lancés», à la page 316
Sur les systèmes Linux, extraction impossible de valeurs de Metric Server lors de l'exécution de Load Balancer pour IPv6	Lors de l'exécution sur les plateformes Linux, il existe une incompatibilité dans la sélection des adresses IPv6 source. Metric Monitor tente alors de communiquer avec Metric Server via l'adresse IP source erronée.	«Incident : Avec Load Balancer pour IPv6, extraction impossible de valeurs de Metric Server sur des systèmes Linux», à la page 339
La valeur de mesure renvoie -1 après le démarrage de Metric Server	Cet incident peut provenir de la perte d'intégrité des fichiers de clés lors de leur transfert au client.	«Incident : Après le démarrage de Metric Server, la valeur de mesure renvoie -1», à la page 340

Vérification des numéros de port Dispatcher

En cas d'incidents lors de l'exécution de Dispatcher, il se peut que l'une des applications utilise un numéro de port généralement utilisé par Dispatcher. Notez que le serveur Dispatcher utilise les numéros de port suivants :

- 10099 pour recevoir des commandes de dscontrol
- 10004 pour envoyer des demandes de mesure au système Metric Server
- 10199 pour le port du serveur RMI

Si une autre application utilise l'un des numéros de port Dispatcher, vous pouvez modifier les numéros de port de Dispatcher *ou* modifier le numéro de port de l'application.

Pour modifier les numéros de port Dispatcher, procédez comme suit :

- Pour modifier le port permettant de recevoir des commandes
 - Remplacez la variable LB_RMIPORT figurant au début du script du fichier dsserver par le port que Dispatcher doit utiliser pour recevoir les commandes.
- Pour modifier le port permettant de recevoir les rapports de mesure du système Metric Server
 - Remplacez la variable RMI_PORT du fichier metricserver par le port devant être utilisé pour les communications avec le système Metric Server.
 - Fournissez l'argument port_mesure lors du démarrage du gestionnaire. Pour la syntaxe de la commande **dscontrol manager start**, voir «dscontrol manager — Contrôle du gestionnaire», à la page 371.

Modifiez le numéro du port RMI de l'application, comme suit :

- Pour modifier le port qu'utilise l'application
 - Remplacez la variable LB_RMISERVERPORT du fichier dsserver par le port que doit utiliser l'application. (Le port RMI qu'utilise par défaut l'application est 10199.)

Remarque : Pour Windows, les fichiers dsserver et metricserver se trouvent dans le répertoire C:\winnt\system32. Pour les autres plateformes, ces fichiers se trouvent dans le répertoire /usr/bin/.

Vérification des numéros de port CBR

En cas d'incidents lors de l'exécution de CBR, il se peut que l'une des applications utilise un numéro de port généralement utilisé par CBR. Prenez en compte que CBR utilise le numéro de port suivant :

- 11099 pour recevoir des commandes de cbrcontrol
- 10004 pour envoyer des demandes de mesure au système Metric Server
- 11199 pour le port du serveur RMI

Remarque : Le composant Content Based Routing (CBR) est disponible sur toutes les plateformes prises en charge à l'exception de celles exécutant une JVM 64 bits. Vous avez également la possibilité d'employer la méthode d'acheminement cbr du composant Dispatcher de Load Balancer pour permettre le routage CBR sans faire appel à Caching Proxy. Pour plus de détails, voir «Fonction CBR de Dispatcher (méthode d'acheminement cbr)», à la page 55.

Si une autre application utilise l'un des numéros de port CBR, vous pouvez modifier les numéros de port de CBR *ou* modifier le numéro de port de l'application.

Pour modifier les numéros de port CBR, procédez comme suit :

- Pour modifier le port permettant de recevoir des commandes
 - Remplacez la variable LB_RMIPORT figurant au début du fichier cbrserver par le port que CBR doit utiliser pour recevoir des commandes.
- Pour modifier le port permettant de recevoir les rapports de mesure du système Metric Server
 - Remplacez la variable RML_PORT du fichier metricserver par le port que doit utiliser CBR pour communiquer avec le système Metric Server.
 - Fournissez l'argument port_mesure lors du démarrage du gestionnaire. Reportez-vous à la syntaxe de la commande **manager start** «dscontrol manager — Contrôle du gestionnaire», à la page 371

Modifiez le numéro du port RMI de l'application, comme suit :

- Pour modifier le port qu'utilise l'application
 - Remplacez la variable LB_RMISERVERPORT figurant au début du fichier cbrserver par le port que doit utiliser l'application. (Le port RMI qu'utilise par défaut l'application est 11199.)

Remarque : Pour Windows, les fichiers cbrserver et metricserver se trouvent dans le répertoire C:\winnt\system32. Pour les autres plateformes, ces fichiers se trouvent dans le répertoire /usr/bin/.

Vérification des numéros de port Site Selector

En cas d'incidents lors de l'exécution du composant Site Selector, il se peut que l'une des applications utilise un numéro de port généralement utilisé par Site Selector. Prenez en compte le fait que Site Selector utilise les numéros de port suivants :

- 12099 pour recevoir des commandes de sscontrol
- 10004 pour envoyer des demandes de mesure au système Metric Server
- 12199 pour le port du serveur RMI

Si une autre application utilise l'un des numéros de port Site Selector, vous pouvez modifier les numéros de port de Site Selector *ou* modifier le numéro de port de l'application.

Pour modifier les numéros de port de Site Selector, procédez comme suit :

- Pour modifier le numéro de port permettant de recevoir des commandes,
 - Remplacez la variable LB_RMIPORT figurant au début du fichier ssserver par le port que doit utiliser Site Selector pour recevoir les commandes.
- Pour modifier le port permettant de recevoir les rapports de mesure du système Metric Server
 - Remplacez la variable RML_PORT dans le fichier metricserver par le port que Site Selector doit utiliser pour communiquer avec le système Metric Server.
 - Fournissez l'argument port_mesure lors du démarrage du gestionnaire. Reportez-vous à la syntaxe de commande **manager start** «sscontrol manager — Contrôle du gestionnaire», à la page 411

Modifiez le numéro du port RMI de l'application, comme suit :

- Pour modifier le port qu'utilise l'application
 - Remplacez la variable LB_RMISERVERPORT figurant au début du fichier sserver par le port que doit utiliser l'application. (Le port RMI qu'utilise par défaut l'application est 12199.)

Remarque : Pour Windows, les fichiers sserver et metricserver se trouvent dans le répertoire C:\winnt\system32. Pour les autres plateformes, ces fichiers se trouvent dans le répertoire /usr/bin/.

Vérification des numéros de port Cisco CSS Controller

En cas d'incidents lors de l'exécution du composant Cisco CSS Controller, il se peut qu'une autre application utilise l'un des numéros de port utilisés par le serveur lbcserver de Cisco CSS Controller. Prenez en compte le fait que Cisco CSS Controller utilise les numéros de port suivants :

- 13099 pour recevoir des commandes de ccocontrol
- 10004 pour envoyer des demandes de mesure au système Metric Server
- 13199 pour le port du serveur RMI

Si une autre application utilise l'un des numéros de port Cisco CSS Controller, vous pouvez modifier les numéros de port de Cisco CSS Controller *ou* modifier le numéro de port de l'application.

Pour modifier les numéros de port Cisco CSS Controller, procédez comme suit :

- Pour modifier le port permettant de recevoir des commandes de ccocontrol, modifiez la variable CCO_RMIPORT du fichier ccoserver. Remplacez 13099 par le port sur lequel vous voulez que Cisco CSS Controller reçoive les commandes ccocontrol.
- Pour modifier le port permettant de recevoir les rapports de mesure du système Metric Server :
 1. Modifiez la variable RMI_PORT dans le fichier metricserver. Remplacez 10004 sur lequel Cisco CSS Controller doit communiquer avec le système Metric Server.
 2. Fournissez l'argument port_mesure lors du démarrage du consultant.

Modifiez le numéro du port RMI de l'application, comme suit :

- Pour modifier le port qu'utilise l'application
 - Remplacez la variable CCO_RMISERVERPORT figurant au début du fichier ccoserver par le port que doit utiliser l'application. (Le port RMI qu'utilise par défaut l'application est 13199.)

Remarque : Pour Windows, les fichiers ccoserver et metricserver se trouvent dans le répertoire C:\winnt\system32. Pour les autres plateformes, ces fichiers se trouvent dans le répertoire /usr/bin.

Vérification des numéros de port Nortel Alteon Controller

En cas d'incidents lors de l'exécution du composant Nortel Alteon Controller, il se peut qu'une autre application utilise l'un des numéros de port utilisés par le serveur nalserver de Nortel Alteon Controller. Prenez en compte le fait que Nortel Alteon Controller utilise les numéros de port suivants :

- 14099 pour recevoir les commandes de nalcontrol

10004 pour envoyer des demandes de mesure au système Metric Server
14199 pour le port du serveur RMI

Si une autre application utilise l'un des numéros de port Nortel Alteon Controller, vous pouvez modifier les numéros de port de Nortel Alteon Controller *ou* modifier le numéro de port de l'application.

Pour modifier les numéros de port Nortel Alteon Controller, , procédez comme suit :

- Pour modifier le port permettant de recevoir des commandes de nalcontrol, modifiez la variable NAL_RMI_PORT du fichier nalserver. Remplacez 14099 par le port sur lequel vous voulez que Nortel Alteon Controller reçoive les commandes nalcontrol.
- Pour modifier le port permettant de recevoir les rapports de mesure du système Metric Server :
 1. Modifiez la variable RMI_PORT dans le fichier metricserver. Remplacez 10004 sur lequel Nortel Alteon Controller doit communiquer avec le système Metric Server.
 2. Fournissez l'argument port_mesure lors du démarrage du consultant.

Modifiez le numéro du port RMI de l'application, comme suit :

- Pour modifier le port qu'utilise l'application
 - Remplacez la variable NAL_RMISERVERPORT figurant au début du fichier nalserver par le port que doit utiliser l'application. (Le port RMI qu'utilise par défaut l'application est 14199.)

Remarque : Pour Windows, les fichiers nalserver et metricserver se trouvent dans le répertoire C:\winnt\system32. Pour les autres plateformes, ces fichiers se trouvent dans le répertoire /usr/bin.

Résolution des incidents courants—Dispatcher

Incident : Dispatcher ne fonctionne pas

Cet incident peut se produire lorsqu'une autre application utilise l'un des ports utilisés par Dispatcher. Pour plus de détails, voir «Vérification des numéros de port Dispatcher», à la page 299.

Incident : Le répartiteur et le serveur ne répondent pas

Cet incident se produit lorsque qu'une adresse autre que l'adresse spécifiée est utilisée. Lorsque vous placez le Dispatcher et le serveur sur le même poste, assurez-vous que l'adresse NFA est utilisée ou est configurée comme utilisant le même poste. Vérifiez l'adresse dans le fichier hôte.

Incident : Les requêtes Dispatcher ne sont pas équilibrées

Les symptômes de cet incident sont l'absence de prise en charge des connexions des machines client ou le dépassement du délai des connexions. Effectuez les contrôles suivants pour diagnostiquer cet incident :

1. Avez-vous configuré l'adresse de non-réacheminement, les clusters, les ports et les serveurs pour l'acheminement ? Vérifiez le fichier de configuration.
2. L'alias de la carte d'interface de réseau est-il associé à l'adresse de cluster ?
Pour les systèmes Linux et UNIX, utilisez netstat -ni pour vérifier.

3. L'alias de l'unité de bouclage de chaque serveur est-il associé à l'adresse de cluster ? Pour les systèmes Linux et UNIX, utilisez `netstat -ni` pour vérifier.
4. La voie d'acheminement supplémentaire est-elle supprimée ? Pour les systèmes Linux et UNIX, utilisez `netstat -nr` pour vérifier.
5. Utilisez la commande **dscontrol cluster status** pour vérifier les informations relatives à chacun des clusters que vous avez définis. Assurez-vous qu'un port est défini pour chaque cluster.
6. Utilisez la commande **dscontrol server report ::** pour vérifier que vos serveurs ne sont pas hors service ou qu'ils n'ont pas pour valeur une pondération égale à zéro.

Pour Windows et les autres plateformes, voir également «Configuration des serveurs pour l'équilibrage de la charge», à la page 72.

Incident : La fonction haute disponibilité de Dispatcher est inopérante

Cet incident se produit lorsqu'un environnement de haute disponibilité de Dispatcher est configuré et que les connexions des machines client ne sont pas prises en charge ou que le délai imparti à ces connexions est dépassé. Effectuez les contrôles suivants pour corriger ou diagnostiquer l'incident :

- Assurez-vous que les scripts `goActive`, `goStandby` et `goInOp` ont été créés et qu'ils se trouvent dans le répertoire `bin` dans lequel Dispatcher est installé. Pour obtenir plus d'informations sur ces scripts, reportez-vous à la rubrique «Utilisation de scripts», à la page 209
- Pour les systèmes **AIX**, **HP-UX**, **Linux** et **Solaris**, assurez-vous que droit d'exécution est défini pour les scripts `goActive`, `goStandby` et `goInOp`.
- Pour les systèmes Windows, veillez à configurer l'adresse de non-réacheminement à l'aide de la commande **executor configure**.

Les étapes suivantes permettent de vérifier de manière efficace que les scripts de haute disponibilité fonctionnent correctement :

1. générez un rapport en exécutant les commandes `netstat -an` et `ifconfig -a` sur la machine
2. exécutez le script `goActive`
3. exécutez le script `goStandby`
4. générez une nouvelle fois un rapport en exécutant les commandes `netstat -an` et `ifconfig -a`

Ces deux rapports sont identiques si les scripts sont correctement configurés.

Incident : Impossible d'ajouter un signal de présence (plateforme Windows)

Cette erreur Windows se produit lorsque l'adresse source n'est pas configurée sur un adaptateur. Effectuez les contrôles suivants pour corriger ou diagnostiquer l'incident :

- Veillez à configurer l'adresse de non-réacheminement en utilisant l'interface en anneau à jeton ou Ethernet et en exécutant l'une des commandes suivantes :
`dscontrol executor configure <adresse ip>`

Incident : Routes supplémentaires (Windows 2000)

Après la configuration des serveurs, il se peut qu'une ou plusieurs routes supplémentaires aient été créées par inadvertance. Si elles ne sont pas supprimées,

ces routes empêchent le fonctionnement de Dispatcher. Voir «Configuration des serveurs pour l'équilibrage de la charge», à la page 72 pour les contrôler et les supprimer.

Incident : Les conseillers ne fonctionnent pas correctement

Si vous utilisez un support de réseau étendu et que les conseillers ne semblent pas fonctionner correctement, vérifiez qu'ils sont bien lancés sur les répartiteurs locaux et éloignés.

Une commande ping ICMP est envoyée aux serveurs avant la demande du conseiller. S'il existe un pare-feu entre Load Balancer et les serveurs, vérifiez qu'il autorise les commandes ping. Si cette configuration pose un problème de sécurité pour votre réseau, modifiez l'instruction java dans dsserver pour désactiver toutes les commandes ping vers les serveurs en ajoutant la propriété java suivante :

```
LB_ADV_NO_PING="true"
java -DLB_ADV_NO_PING="true"
```

Voir «Utilisation de conseillers éloignés avec le support de réseau étendu de Dispatcher», à la page 230.

Incident : Dispatcher, Microsoft IIS et SSL ne fonctionnent pas (plateformeWindows)

Lorsque vous utilisez Dispatcher, Microsoft IIS et SSL, s'ils ne fonctionnent pas ensemble, il se peut qu'un incident lié à la sécurité SSL se soit produit. Pour plus d'informations sur la génération d'une paire de clés, l'acquisition d'un certificat, l'installation d'un certificat avec une paire de clés et la configuration d'un répertoire pour SSL, reportez-vous à la documentation *Microsoft Information and Peer Web Services*.

Incident : Connexion du répartiteur à une machine éloignée

Dispatcher utilise des clés pour vous permettre de vous connecter à une machine éloignée et la configurer. Les clés indiquent un port RMI pour la connexion. Il est possible de changer de port RMI pour des raisons de sécurité ou de conflits. Lorsque vous changez les ports RMI, le nom de fichier ou la clé est différent(e). Si votre répertoire contient plusieurs clés pour la même machine éloignée et qu'elles indiquent des ports RMI différents, la ligne de commande essaie la première qu'elle trouve. Si elle n'est pas appropriée, la connexion est refusée. La connexion ne sera établie que si vous supprimez la clé incorrecte.

Incident : La commande dscontrol ou lbadmin n'a pas abouti

1. La commande dscontrol renvoie : **Erreur : Pas de réponse du serveur**. Ou la commande lbadmin renvoie : **Erreur : impossible d'accéder au serveur RMI**. Ces erreurs peuvent se manifester lorsque votre machine a une pile sur "sock". Pour corriger ce problème, éditez le fichier socks.cnf pour qu'il contienne les lignes suivantes :

```
EXCLUDE-MODULE java
EXCLUDE-MODULE javaw
```

2. Les consoles d'administration des interfaces Load Balancer (ligne de commande, interface graphique et assistants) communiquent avec dsserver par appels RMI (Remote Method Invocation). Par défaut, la communication utilise trois ports : chacun étant défini dans le script de démarrage de dsserver :
 - 10099 pour recevoir des commandes de dscontrol
 - 10004 pour envoyer des demandes de mesure au système Metric Server

- 10199 pour le port du serveur RMI

Ceci peut être source de problèmes lorsqu’une des consoles d’administration s’exécute sur la même machine qu’un pare-feu ou passe par un pare-feu. Par exemple, lorsque vous émettez des commandes dscontrol alors que Load Balancer s’exécute sur la même machine qu’un pare-feu, des erreurs de type **Erreur : Pas de réponse du serveur** peuvent s’afficher.

Pour éviter ce type d’incident, modifiez le fichier script ndserver afin de définir le port qu’utilise RMI pour le pare-feu (ou autre application). Remplacez la ligne `LB_RMISERVERPORT=10199` par `LB_RMISERVERPORT=votrePort`. Où *vousPort* est un autre port.

Lorsque vous avez terminé, relancez la commande dsserver et ouvrez le trafic des ports 10099, 10004, 10199 et 10100 ou du port d’adresse hôte choisi pour l’exécution de la console d’administration.

3. Ces erreurs peuvent également se produire si vous n’avez pas encore lancé **dsserver**.
4. S’il existe plusieurs adaptateurs sur la machine, vous devez désigner celui à utiliser par dsserver en ajoutant la ligne suivante dans le script dsserver : `java.rmi.server.hostname=<nom_hôte ou adresseIP>`

Par exemple : `java -Djava.rmi.server.hostname="10.1.1.1"`

Incident : Affichage du message d’erreur “Fichier introuvable...” lorsque vous tentez de visualiser l’aide en ligne (plateforme Windows)

Sur les plateformes Windows, lorsque vous utilisez Netscape comme navigateur par défaut, le message d’erreur suivant peut s’afficher : “Impossible trouver fichier ‘<nomfichier>.html’ (ou un de ses composants). Vérifiez que le chemin et le nom de fichier sont corrects et que toutes les bibliothèques nécessaires sont disponibles.

Le problème est dû à un réglage incorrect pour l’association de fichier HTML. La solution est la suivante :

1. Cliquez sur **Poste de travail**, cliquez sur **Outils**, sélectionnez **Options dossier** et cliquez sur **Types fichiers**
2. Sélectionnez “Document hypertexte Netscape”
3. Cliquez sur le bouton **Avancé**, sélectionnez **ouvrir**, cliquez sur le bouton **Editer**
4. Entrez *NSShell* dans la zone **Application** : (et non dans la zone Application utilisée pour réaliser action :), puis cliquez sur **OK**

Incident : L’interface graphique ne démarre pas correctement

Pour que l’interface graphique, lbadadmin, fonctionne correctement, vous devez disposer d’une quantité d’espace de pagination suffisante. Dans le cas contraire, l’interface graphique peut ne pas démarrer complètement. Si cela se produit, vérifiez l’espace de pagination et augmentez-la si nécessaire.

Incident : Erreur lors de l’exécution de Dispatcher lorsque Caching Proxy est installé

Si vous désinstallez Load Balancer afin de réinstaller une autre version et que vous obtenez une erreur lorsque vous tentez de lancer le composant Dispatcher, vérifiez si Caching Proxy est installé. Caching Proxy est lié à un des fichiers Dispatcher. Ce fichier sera désinstallé lors de la désinstallation de Caching Proxy.

Pour résoudre ce problème :

1. Désinstallez Caching Proxy.
2. Désinstallez Load Balancer.
3. Réinstallez Load Balancer et Caching Proxy.

Incident : L'interface graphique ne s'affiche pas correctement

Si des incidents se produisent relatifs à l'apparence de l'interface graphique de Load Balancer, vérifiez la configuration de la résolution du bureau du système d'exploitation. Pour un affichage de l'interface graphique optimal, nous vous recommandons d'utiliser une résolution de 1024x768 pixels.

Incident : Sous Windows, les fenêtre d'aide disparaissent parfois sous d'autres fenêtres ouvertes

Sous Windows, lorsque vous ouvrez des fenêtres d'aide, elles peuvent disparaître en arrière-plan sous les fenêtres existantes. Si cela se produit, cliquez sur la fenêtre pour qu'elle s'affiche à nouveau en premier plan.

Incident: Load Balancer ne peut pas traiter et transmettre un cadre

Sous Solaris, chaque carte réseau possède la même adresse MAC par défaut. Cela fonctionne correctement lorsque chaque carte se trouve sur un sous-réseau IP différent. Cependant, dans un environnement commuté, lorsque plusieurs cartes NIC ayant la même adresse MAC et la même adresse de sous-réseau IP communiquent avec le même commutateur, ce dernier envoie l'ensemble du trafic associé à l'adresse MAC (et aux adresses IP) via le même câble. Seule la carte ayant la dernière placé un cadre sur ce réseau voit les paquets IP associés aux deux cartes. Solaris peut ignorer les paquets d'une adresse IP valide arrivant à l'interface "incorrecte".

Si toutes les interfaces réseau ne sont pas conçues pour Load Balancer, conformément à la configuration définie dans le fichier `ibmlb.conf` et si la carte NIC non définie dans `ibmlb.conf` reçoit un cadre, Load Balancer ne peut pas traiter et transmettre le cadre.

Pour éviter ce problème, vous devez remplacer la valeur par défaut et définir une adresse unique pour chaque interface. Utilisez cette commande :

```
ifconfig interface ether adrMac
```

Par exemple :

```
ifconfig eri0 ether 01:02:03:04:05:06
```

Incident : Un écran bleu s'affiche lors du démarrage de l'exécuteur Load Balancer

Sous Windows, vous devez avoir une carte réseau installée et configurée avant le démarrage de l'exécuteur.

Incident : La fonction Path MTU Discovery permet d'éviter le trafic retour avec Load Balancer

Le système d'exploitation AIX contient une fonction de mise en réseau appelée "Path MTU Discovery". Lors d'une transaction avec un client, si le système d'exploitation détermine qu'une unité de transmission maximale (MTU) inférieure doit être utilisée pour les paquets sortants, la fonction Path MTU Discovery

entraîne la création par AIX d'une route de rappel de cette unité. La nouvelle route est réservée à cette adresse IP client et enregistre l'unité MTU qui permet d'atteindre celle-ci.

Lors de la création de la route, un incident peut survenir sur les serveurs en raison de l'association du cluster à un alias sur l'unité de bouclage. Si l'adresse de la passerelle pour la route entre dans le sous-réseau du cluster ou du masque réseau, les systèmes AIX créent la route sur l'unité de bouclage. Cet événement s'est produit car il s'agissait de l'alias de la dernière interface associé à ce sous-réseau.

Par exemple, si le cluster s'appelle 9.37.54.69, le masque réseau 255.255.255.0 et la passerelle prévue 9.37.54.1, les systèmes AIX utilisent l'unité de bouclage pour la route. En raison de cette action, les réponses du serveur ne sortent jamais et le client dépasse le délai d'attente. Habituellement, le client voit une réponse du cluster, puis il ne reçoit plus rien lorsque la route est créée.

Deux solutions permettent de pallier cet incident :

1. Désactivez la fonction Path MTU Discovery pour que le système AIX n'ajoute pas de route en mode dynamique. Utilisez les commandes suivantes :
no -a répertorie les paramètres de mise en réseau AIX
no -o option=valeur
définit les paramètres TCP sur les systèmes AIX
2. Associez l'adresse IP de cluster sur l'unité de bouclage au masque réseau 255.255.255.255. Cela signifie que le sous-réseau associé à un alias est uniquement l'adresse IP de cluster. Lorsque les systèmes AIX créent les routes dynamiques, l'adresse IP de la passerelle cible ne correspond pas à ce sous-réseau. En conséquence, la route est amenée à utiliser l'interface réseau incorrecte. Supprimez ensuite la nouvelle route lo0 créée lors de la définition d'alias. Pour la supprimer, recherchez-la d'abord sur l'unité de bouclage avec une destination réseau pour l'adresse IP de cluster. Cette action doit être effectuée à chaque utilisation d'alias pour le cluster.

Remarques :

1. La fonction Path MTU Discovery est désactivée par défaut dans les versions d'AIX antérieures à la version 4.3.2 et activée par défaut à partir de la version 4.3.3.
2. Les commandes ci-après permettent de désactiver la fonction Path MTU Discovery et doivent être exécutées à chaque amorçage du système. Ajoutez-les dans le fichier /etc/rc.net.
 - -o udp_pmtu_discover=0
 - -o tcp_pmtu_discover=0

Incident : La fonction haute disponibilité de Load Balancer en mode réseau étendu est inopérante

Lorsque vous configurez un mode WAND (Wide Area Load Balancer), vous devez définir la machine Dispatcher locale en tant que serveur d'un cluster sur la machine Dispatcher locale. Habituellement, vous utilisez l'adresse de non-réacheminement (NFA) de la machine Dispatcher éloignée pour l'adresse de destination du serveur éloigné. Dans ce cas, si vous configurez ensuite sur la machine Dispatcher éloignée la fonction haute disponibilité, celle-ci reste inopérante. La cause en est la suivante : la machine Dispatcher locale désigne toujours la machine principale côté éloigné lorsque vous utilisez l'adresse NFA.

Pour éviter cet incident :

1. Définissez un autre cluster sur la machine Dispatcher éloignée. Il n'est pas nécessaire de définir des ports ou des serveurs pour ce cluster.
2. Ajoutez l'adresse de ce cluster dans les scripts goActive et goStandby.
3. Sur la machine Dispatcher locale, définissez l'adresse de cluster en tant que serveur, et non en tant qu'adresse NFA de la machine Dispatcher principale éloignée.

Lorsque la machine Dispatcher principale éloignée entre en service, elle associe cette adresse à un alias sur la carte, permettant ainsi l'acceptation du trafic. En cas de défaillance, l'adresse se déplace sur la machine de secours qui continue à accepter le trafic destiné à cette adresse.

Incident : Arrêt (ou comportement imprévu) de l'interface graphique lors de la tentative de chargement d'un fichier de configuration volumineux

Lors de l'utilisation d'une administration lbadmin ou Web (lbwebaccess) pour charger un fichier de configuration volumineux (200 commandes **add** ou même plus), l'interface graphique peut s'arrêter ou avoir un comportement imprévu, comme présenter un temps de réponse excessif lorsque vous apportez des modifications à l'écran.

Cela vient du fait que la mémoire est insuffisante pour permettre à Java de traiter une configuration de cette ampleur.

L'environnement d'exécution dispose d'une option permettant d'augmenter la mémoire disponible dont peut disposer Java.

Il s'agit de l'option `-Xmxn` où *n* correspond à la taille maximale, en octets, de la mémoire. *n* doit être un multiple de 1024 et supérieur à 2 Mo. La valeur *n* peut être suivie du caractère *k* ou *K* pour indiquer qu'il s'agit de kilo-octets ou du caractère *m* ou *M* pour indiquer qu'il s'agit de méga-octets. Par exemple, `-Xmx128M` et `-Xmx81920k` sont valides. La valeur par défaut est 64M. Solaris 8 accepte une valeur maximale de 4000M.

Par exemple, pour ajouter cette option, ouvrez le fichier script lbadmin, modifiez "javaw" en "javaw -Xmxn" comme suit. (Pour les systèmes AIX, modifiez "java" en "java -Xmxn") :

- **Systèmes AIX**

```
javaw -Xmx256m -cp $LB_CLASSPATH $LB_INSTALL_PATH $LB_CLIENT_KEYS  
com.ibm.internet.nd.framework.FWK_Main 1>/dev/null 2>&1 &
```

- **Systèmes HP-UX**

```
java -Xmx256m -cp $LB_CLASSPATH $LB_INSTALL_PATH $LB_CLIENT_KEYS  
com.ibm.internet.nd.framework.FWK_Main 1>/dev/null 2>&1 &
```

- **Systèmes Linux**

```
javaw -Xmx256m -cp $LB_CLASSPATH $LB_INSTALL_PATH $LB_CLIENT_KEYS  
com.ibm.internet.nd.framework.FWK_Main 1>/dev/null 2>&1 &
```

- **Systèmes Solaris**

```
java -Xmx256m -cp $LB_CLASSPATH $LB_INSTALL_PATH $LB_CLIENT_KEYS  
com.ibm.internet.nd.framework.FWK_Main 1>/dev/null 2>&1 &
```

- **Systèmes Windows**

```
START javaw -Xmx256m -cp %LB_CLASSPATH% %LB_INSTALL_PATH%  
%LB_CLIENT_KEYS% com.ibm.internet.nd.framework.FWK_Main
```

Aucune valeur *n* particulière n'est recommandée, mais elle doit être supérieure à l'option par défaut. Pour commencer, vous pouvez doubler cette dernière.

Incident : lbadmin se déconnecte du serveur après mise à jour de la configuration

Si l'administration (lbadmin) de Load Balancer se déconnecte du serveur après mise à jour de la configuration, vérifiez quelle version de dsserver vous essayez de configurer sur le serveur et vérifiez qu'il s'agit de la même version que celle de lbadmin ou dscontrol.

Incident : Adresses IP non résolues correctement sur la connexion éloignée

Lorsqu'un client éloigné est utilisé sur une implémentation SSL, des noms d'hôtes ou des noms de domaines complets ne sont pas correctement convertis en adresses IP au format d'adresse IP. L'implémentation SSL doit ajouter à la résolution des noms de domaines des données propres à la connexion.

Si les adresses IP ne sont pas correctement résolues sur la connexion éloignée, indiquez l'adresse IP au format d'adresse IP.

Incident : L'interface coréenne de Load Balancer affiche sous AIX et Linux des polices non souhaitées ou qui se chevauchent

Pour rectifier les problèmes de chevauchement de polices ou de polices indésirables sur l'interface Load Balancer coréenne, procédez comme suit :

Sur les systèmes AIX

1. Arrêtez tous les processus Java sur le système AIX.
2. Ouvrez le fichier font.properties.ko dans un éditeur. Ce fichier se trouve dans le répertoire *principal/jre/lib* où *principal* est le répertoire principal Java.
3. Recherchez la chaîne :
`-Monotype-TimesNewRomanWT-medium-r-normal
--*-%d-75-75-*--ksc5601.1987-0`
4. Remplacez toutes les instances de cette chaîne par :
`-Monotype-SansMonoWT-medium-r-normal
--*-%d-75-75-*--ksc5601.1987-0`
5. Sauvegardez le fichier.

Sur les systèmes Linux

1. Arrêtez tous les processus Java sur le système.
2. Ouvrez le fichier font.properties.ko dans un éditeur. Ce fichier se trouve dans le répertoire *principal/jre/lib* où *principal* est le répertoire principal Java.
3. Recherchez la chaîne (sans espace) :
`-monotype-
timesnewromanwt-medium-r-normal--*-%d-75-75-p-*microsoft-symbol`
4. Remplacez toutes les instances de cette chaîne par :
`-monotype-sansmonowt-medium-r-normal--*-%d-75-75-p-*microsoft-symbol`
5. Sauvegardez le fichier.

Incident : Sous Windows, adresse d'alias renvoyée au lieu de l'adresse locale lors de l'émission de commandes telles que hostname

Sous Windows, lorsqu'un alias a été attribué à l'unité de bouclage de MS, le système d'exploitation renvoie l'adresse d'alias au lieu de l'adresse locale lors de l'émission d'une commande telle que `hostname`. Pour rectifier cette erreur, veuillez à placer le nouvel alias au-dessous de l'adresse locale dans la liste des connexions réseau. Ainsi, l'adresse locale est utilisée avant l'alias de bouclage.

Pour consulter la liste des connexions réseau, procédez comme suit :

1. Cliquez sur **Démarrer > Paramètres > Connexions réseau et accès à distance**
2. 0 partir de l'option de menu **Avancé**, sélectionnez **Paramètres avancés...**
3. Vérifiez que **Connexion locale** figure en premier dans la liste **Connexions**.
4. Au besoin, utilisez les boutons situés à droite pour déplacer les entrées vers le haut ou le bas de la liste.

Incident : sous Windows, comportement inattendu de l'interface graphique lors de l'utilisation de cartes vidéo Matrox AGP

Sous Windows, l'interface graphique de Load Balancer peut se comporter de manière inattendue lorsque vous utilisez une carte vidéo Matrox AGP. Lorsque vous cliquez sur un bouton de la souris, un espace légèrement plus large que le pointeur de la souris peut être altéré avec inversion possible de la mise en évidence ou déplacement des images sur l'écran. Les anciennes cartes Matrox n'ont pas présenté ce type de comportement. Il n'existe pas de rectificatif pour les cartes Matrox AGP.

Incident : Comportement inattendu lors de l'exécution de "rmmod ibmlb" (systèmes Linux)

Sur les systèmes Linux, si `dsserver` est encore actif lors du retrait manuel du module Load Balancer du noyau, un comportement inattendu, tel qu'un arrêt du système ou des noyaux java, peut se produire. Avant de retirer manuellement du noyau le module Load Balancer, vous devez arrêter `dsserver`.

Si la commande `"dsserver stop"` ne fonctionne pas, arrêtez le processus Java avec `SRV_KNDConfigServer`. Pour arrêter le processus, recherchez l'identifiant correspondant à l'aide de la commande `ps -ef | grep SRV_KNDConfigServer`, puis mettez fin à ce processus à l'aide de la commande `kill id_processus`.

Vous pouvez, en toute sécurité, exécuter la commande `"rmmod ibmlb"` de retrait du noyau du module Load Balancer.

Incident : Temps de réponse important lors de l'exécution de commandes sur la machine Dispatcher

Si vous exécutez le composant Dispatcher pour l'équilibrage de charge, le trafic client peut surcharger l'ordinateur. Le module Load Balancer du noyau est hautement prioritaire, et s'il gère constamment des paquets de clients, il est possible que le reste du système ne réponde plus. L'exécution de commandes dans l'espace utilisateur peut être très longue ou ne jamais se terminer.

Dans ce cas, vous devez restructurer votre configuration de sorte que le trafic ne surcharge pas la machine Load Balancer. Vous pouvez également répartir la charge sur plusieurs machines Load Balancer ou remplacer votre machine par un ordinateur plus performant et plus rapide.

Lorsque vous essayez de déterminer si la longueur des temps de réponse provient d'un volume important de trafic client, vérifiez si cela se produit aux heures de pointe. Une mauvaise configuration des systèmes entraînant des boucles de routage peut provoquer le même incident. Mais avant de modifier la configuration de Load Balancer, vérifiez si les symptômes ne sont pas liés à une charge trop importante au niveau du client.

Incident : Le conseiller SSL ou HTTPS n'enregistre pas les charges des serveurs (avec l'acheminement MAC)

Lorsque vous utilisez la méthode d'acheminement MAC, Load Balancer envoie des paquets aux serveurs en utilisant l'adresse du cluster pour laquelle un alias a été défini sur l'unité de bouclage. Certaines applications serveurs (telle SSL) exigent que les informations de configuration (tels les certificats) soient définies en fonction de l'adresse IP. L'adresse IP doit être l'adresse du cluster configurée sur l'unité de bouclage de façon à correspondre au contenu des paquets entrants. Si vous ne configurez pas l'application serveur avec l'adresse IP du cluster, la demande client n'est pas correctement acheminée vers le serveur.

Incident : Déconnexion de l'hôte lors du redimensionnement de la fenêtre du navigateur Netscape en cours d'administration Web

Lorsque vous utilisez l'administration Web à distance pour configurer Load Balancer, ne modifiez pas la taille (Réduire, Agrandir, Restaurer en bas, etc.) du navigateur Netscape dans lequel s'affiche l'interface graphique de Load Balancer. En effet, étant donné que Netscape recharge une page à chaque redimensionnement de la fenêtre du navigateur, une déconnexion de l'hôte en découle. Vous devez donc vous reconnecter à l'hôte après chaque modification de la taille de la fenêtre. Pour l'administration à distance basée sur le Web sur une plateforme Windows, utilisez Internet Explorer.

Incident : Regroupement de connexions activé et serveur Web établissant une liaison à 0.0.0.0

Lorsque vous exécutez le serveur Microsoft IIS, version 5.0 sur des serveurs d'arrière-plan Windows, vous devez configurer le serveur Microsoft IIS en tant que serveur de liaison. Sinon, le regroupement de connexions est activé par défaut, et le serveur Web établit une liaison à 0.0.0.0 et écoute la totalité du trafic, au lieu d'établir une liaison aux adresses IP virtuelles configurées en tant qu'identificateurs multiples pour le site. Si une application de l'hôte local s'arrête alors que le regroupement de connexions est activé, les conseillers des serveurs AIX ou Windows NT détectent cet arrêt. En revanche, si une application d'un hôte virtuel s'arrête alors que l'hôte local reste actif, les conseillers ne détectent pas l'incident et le serveur Microsoft IIS continue à répondre à la totalité du trafic, y compris à celui de l'application arrêtée.

Pour déterminer si le regroupement de connexions est activé et si le serveur Web établit une liaison à 0.0.0.0, émettez la commande suivante :

```
netstat -an
```

Les instructions de configuration du serveur Microsoft IIS en tant que serveur de liaison (qui désactive le regroupement de connexions), sont disponible sur le site Web de Microsoft "Product Support Services". Selon l'information recherchez, vous pouvez également vous rendre sur les sites suivants :

IIS5 : Le matériel Load Balance ne détecte pas un site Web arrêté (Q300509)

<http://support.microsoft.com/default.aspx?scid=kb;en-us;Q300509>

Désactivation du regroupement de connexions (Q238131)

<http://support.microsoft.com/default.aspx?scid=kb;en-us;Q238131>

Incident : Sous Windows, des caractères nationaux Latin-1 endommagés apparaissent dans la fenêtre de la ligne de commande

Dans une fenêtre de ligne de commande du système d'exploitation Windows, certains caractères nationaux de la famille Latin-1 sont altérés. Par exemple, la lettre "a" avec tilde s'affiche sous la forme d'un symbole pi. Pour rectifier cette erreur, vous devez modifier les propriétés de police de la fenêtre de ligne de commande. Pour modifier la police, procédez comme suit :

1. Cliquez sur l'icône située dans l'angle supérieur gauche de la fenêtre de ligne de commande.
2. Sélectionnez Propriétés, puis cliquez sur l'onglet Police.
3. La police par défaut est Raster ; remplacez-la par Lucida Console, puis cliquez sur OK.

Incident : Sous HP-UX, la mémoire est insuffisante pour Java et une erreur d'unité d'exécution est générée

Certaines installations HP-UX 11i sont préconfigurées pour n'autoriser que 64 unités d'exécution par processus. Toutefois, certaines configurations Load Balancer en requièrent davantage. Pour les systèmes HP-UX, définissez les unités d'exécution par processus sur 256 au moins. Pour augmenter cette valeur, utilisez l'utilitaire "sam" pour définir le paramètre de noyau `max_thread_proc`. Si vous pensez avoir besoin d'un nombre d'unités d'exécution plus important, vous pouvez affecter à ce paramètre une valeur supérieure à 256.

Pour augmenter la valeur du paramètre `max_thread_proc`, procédez comme suit :

1. A partir de la ligne de commande, entrez : `sam`
2. Sélectionnez **Configuration du noyau > Paramètres configurables**
3. A partir de la barre de défilement, sélectionnez **max_thread_proc**
4. Appuyez sur la barre d'espace pour sélectionner **max_thread_proc**
5. Appuyez une fois sur la touche de tabulation, puis sur la flèche de droite jusqu'à ce que vous puissiez sélectionner **Actions**
6. Appuyez sur Entrée pour afficher le menu **Actions**, puis appuyez sur **M** pour sélectionner l'option de modification des paramètres configurables. (Si vous ne trouvez pas cette option, sélectionnez **max_thread_proc**)
7. Appuyez sur la touche de tabulation jusqu'à ce que vous puissiez sélectionner la zone **Formule/Valeur**
8. Entrez une valeur égale ou supérieure à 256.
9. Cliquez sur **OK**
10. Appuyez une fois sur la touche de tabulation, puis sélectionnez **Actions**
11. Appuyez sur la touche **K** pour traiter le nouveau noyau.

12. Sélectionnez **Oui**
13. Réinitialisez votre système

Incident : Sous Windows, les conseillers et les cibles à contacter marquent tous les serveurs comme étant arrêtés

Lorsque vous configurez votre adaptateur sur une machine Load Balancer, vous devez vérifier que les deux paramètres suivants sont corrects pour que le conseiller puisse fonctionner :

- Désactivez Task Offloading, qui est généralement utilisé sur les cartes d'adaptateur 3Com.
 - Pour désactiver Task Offloading, procédez comme suit : Cliquez sur Démarrer > Paramètres > Panneau de configuration > Connexions réseau et accès à distance, puis sélectionnez l'adaptateur.
 - Dans la fenêtre en incrustation, cliquez sur Propriétés.
 - Cliquez sur Configurer, puis sélectionnez l'onglet Paramètres avancés.
 - Dans la sous-fenêtre des propriétés, cliquez sur la propriété Task Offload, puis sélectionnez désactiver dans la zone de valeur.
- Activez Protocole 1 (ICMP) pour les protocoles IP si vous activez le filtrage TCP/IP. Si ICMP n'est pas activé, le test de connexion (ping) au serveur dorsal échoue. Pour vérifier si ICMP est activé, procédez comme suit :
 - Cliquez sur Démarrer > Paramètres > Panneau de configuration > Connexions réseau et accès à distance, puis sélectionnez l'adaptateur.
 - Dans la fenêtre en incrustation, cliquez sur Propriétés.
 - Dans la sous-fenêtre des composants, sélectionnez Internet Protocol (TCP/IP), puis cliquez sur Propriétés.
 - Cliquez sur Paramètres avancés, puis sélectionnez l'onglet Options.
 - Sélectionnez le filtrage TCP/IP dans la sous-fenêtre des options, puis cliquez sur Propriétés.
 - Si vous avez sélectionné **Activer le filtrage TCP/IP** et **Autoriser seulement** pour les protocoles IP, vous devez ajouter Protocole IP 1. Ce protocole doit être ajouté en plus des ports TCP et UDP existants que vous avez activés.

Incident : Sous Windows, la résolution de l'adresse IP en nom d'hôte n'est pas possible lorsque plusieurs adresses sont configurées sur un adaptateur

Sous Windows, lorsque vous configurez un adaptateur avec plusieurs adresses IP, configurez d'abord dans le registre l'adresse IP à affilier au nom d'hôte.

Load Balancer dépendant de `InetAddress.getLocalHost()` dans de nombreuses instances (par exemple, `lbkeys create`), l'attribution comme alias de plusieurs adresses IP à un même adaptateur peut être source d'incident. Pour éviter cela, entrez l'adresse IP à utiliser pour la résolution du nom d'hôte en premier dans le registre. Par exemple :

1. Démarrez Regedit
2. Modifiez les noms de valeur ci-après de la manière suivante :
 - `HKEY_LOCAL_MACHINE -> SYSTEM -> ControlSet001 -> Services -> YourInterfaceAddress} -> Parameters -> Tcpip-> IPAddress`
 - Placez l'adresse IP à utiliser pour la résolution du nom d'hôte en premier.

- HKEY_LOCAL_MACHINE -> SYSTEM -> ControlSet001 -> Services -> Tcpip -> Parameters -> Interfaces -> *AdresseDeVotreInterface* -> IPAddress
 - Placez l'adresse IP à utiliser pour la résolution du nom d'hôte en premier.
 - HKEY_LOCAL_MACHINE -> SYSTEM -> ControlSet002 -> Services -> *AdresseDeVotreInterface* -> Parameters -> Tcpip -> IPAddress
 - Placez l'adresse IP à utiliser pour la résolution du nom d'hôte en premier.
 - HKEY_LOCAL_MACHINE -> SYSTEM -> ControlSet002 -> Services -> Tcpip -> Parameters -> Interfaces -> *AdresseDeVotreInterface* -> IPAddress
 - Placez l'adresse IP à utiliser pour la résolution du nom d'hôte en premier.
 - HKEY_LOCAL_MACHINE -> SYSTEM -> CurrentControlSet -> Services -> *AdresseDeVotreInterface* -> Parameters -> Tcpip -> IPAddress
 - Placez l'adresse IP à utiliser pour la résolution du nom d'hôte en premier.
 - HKEY_LOCAL_MACHINE -> SYSTEM -> CurrentControlSet -> Services -> Tcpip -> Parameters -> Interfaces -> *AdresseDeVotreInterface* -> IPAddress
 - Placez l'adresse IP à utiliser pour la résolution du nom d'hôte en premier.
3. Réinitialisez la machine
 4. Vérifiez que votre nom d'hôte est converti en l'adresse IP appropriée. Par exemple, ping *votrenomhôte*.

Incident : Sous Windows, les conseillers ne fonctionnent pas dans une configuration en haute disponibilité après une panne réseau

Par défaut, lorsque le système d'exploitation Windows détecte une panne réseau, il efface sa mémoire cache ARP (protocole de résolution d'adresse) et toutes les entrées statiques. Lorsque le réseau est à nouveau disponible, la mémoire cache ARP est de nouveau alimentée par les demandes ARP envoyées sur le réseau.

Dans une configuration en haute disponibilité, les deux serveurs assurent les opérations principales lorsqu'une perte de la connectivité réseau affecte l'un d'eux. Lorsque la demande ARP est envoyée pour alimenter la mémoire cache ARP, les deux serveurs répondent et la mémoire cache ARP marque alors l'entrée comme non valide. Par conséquent, les conseillers ne peuvent pas créer de socket vers les serveurs de secours.

Vous pouvez résoudre cet incident en empêchant Windows d'effacer la mémoire cache ARP lors d'une perte de connectivité. Microsoft a publié un article expliquant comment effectuer cette tâche. Cet article se trouve sur le site Web de Microsoft, dans la base de connaissances Microsoft (référence 239924), à l'adresse suivante : <http://support.microsoft.com/default.aspx?scid=kb;en-us;239924>.

Vous trouverez ci-après un récapitulatif des étapes, décrites dans l'article Microsoft, permettant d'empêcher le système d'effacer la mémoire cache ARP.

1. Utilisez l'éditeur de registre (regedit ou regedit32) pour ouvrir le registre.
2. Affichez la clé suivante du registre :
HKEY_LOCAL_MACHINE\System\CurrentControlSet\Services\Tcpip\Parameters
3. Ajoutez la valeur de registre suivante : Nom de la valeur : DisableDHCPMediaSense Type de valeur : REG_DWORD.
4. Une fois que vous avez ajouté cette clé, éditez-la et affectez-lui la valeur 1.
5. Réinitialisez la machine pour que la modification soit appliquée.

Remarque : Cette opération affecte la mémoire cache ARP quelle que soit la valeur du paramètre DHCP.

Incident : Sous Linux, n'utilisez pas la commande "IP address add" lors de l'affectation d'alias à plusieurs clusters de l'unité de bouclage

Certaines considérations doivent être prises en compte lors de l'utilisation de serveurs Linux (noyau 2.4.x) et de la méthode d'acheminement MAC de Dispatcher. Si une adresse de cluster a été configurée pour le serveur sur l'unité de bouclage à l'aide de la commande **ip address add**, vous ne pouvez attribuer un alias qu'à une seule adresse de cluster.

Lorsque vous affectez un alias à plusieurs clusters sur l'unité de bouclage, utilisez la commande **ifconfig** (par exemple,

```
ifconfig lo:num clusterAddress netmask 255.255.255.255 up)
```

En outre, les méthodes de configuration d'interface **ifconfig** et **ip** sont incompatibles. La meilleure pratique consiste, pour un site, à choisir une méthode et à ne pas en changer.

Incident : Message d'erreur "Adresse de routeur non spécifiée ou non valide pour la méthode port"

Lors de l'ajout de serveurs à la configuration Dispatcher, le message d'erreur suivant peut être généré : "Erreur : Adresse de routeur non spécifiée ou non valide pour la méthode port".

Utilisez cette liste de contrôle pour identifier l'incident :

- Vérifiez que vous avez appliqué le niveau de maintenance le plus récent.
- Vérifiez que vous utilisez une distribution IBM de Java (sauf sur les plateformes Solaris).
- Vérifiez que le système n'est pas configuré pour utiliser DHCP sous Windows.
- Si la méthode d'acheminement MAC (par défaut) est utilisée, le serveur, le cluster et une carte NIC au minimum doivent être installés dans le même sous-réseau. Par exemple, il n'est pas possible de définir un cluster à l'adresse 10.1.1.1 et un serveur à l'adresse 130.2.3.4 car ils ne se trouvent pas dans le même sous-réseau.

Remarque : Si la méthode d'acheminement NAT ou CBR est utilisée, les serveurs ne doivent pas nécessairement se trouver dans le même sous-réseau que le cluster.

- Si tous ces éléments se trouvent dans le même sous-réseau et que vous avez affecté un alias au cluster, vérifiez que l'affectation s'effectue sur une carte NIC qui route le trafic vers ce sous-réseau. Par exemple, si en0 est défini pour 13.2.3.4 et en1 pour 9.1.2.3 et que la définition du cluster est 9.5.7.3, vous devez configurer le cluster sur en1. L'interface par défaut est en0.
- Sur les plateformes Linux, vérifiez que vous avez chargé le noyau correct en recherchant le fichier `loadoutput.log` dans le répertoire `/usr/lpp/ibm/internet/nd/logs/dispatcher`. Vérifiez si ce fichier a enregistré des erreurs.

Par défaut, le paramètre `router` a la valeur 0, ce qui indique que le serveur est local. Lorsque vous définissez l'adresse du routeur du serveur à une valeur autre que 0, cela indique qu'il s'agit d'un serveur éloigné, installé dans un sous-réseau

différent. Pour plus d'informations sur le paramètre `router` dans la commande d'ajout d'un serveur, voir «`dscontrol server` — Configuration des serveurs», à la page 390.

Si le serveur ajouté est installé dans un sous-réseau différent, le paramètre `router` doit correspondre à l'adresse du routeur devant être utilisé dans le sous-réseau local pour communiquer avec le serveur éloigné.

Incident : Sur les systèmes Solaris, les processus Load Balancer s'arrêtent lorsque vous quittez la fenêtre de terminal à partir de laquelle ils ont été lancés

Sur les systèmes Solaris, si vous démarrez les scripts Load Balancer (tels que `dsserver` ou `lbadmin`) à partir d'une fenêtre de terminal et que vous quittez cette dernière, le processus Load Balancer s'arrête.

Pour résoudre cet incident, démarrez les scripts Load Balancer à l'aide de la commande **`nohup`**. Par exemple : **`nohup dsserver`**. Cette commande permet d'éviter que les processus lancés à partir de la session de terminal ne reçoivent un signal d'arrêt du terminal lorsque ce dernier est fermé. L'exécution du processus peut ainsi se poursuivre même après la fermeture de la session de terminal. Spécifiez la commande **`nohup`** avant les scripts Load Balancer dont le traitement doit continuer après la fermeture d'une session de terminal.

Incident : Délai lors du chargement d'une configuration Load Balancer

Le chargement d'une configuration Load Balancer peut prendre un certain temps en raison des appels DNS effectués pour résoudre et vérifier l'adresse du serveur.

Si le système DNS de la machine Load Balancer n'est pas correctement configuré ou s'il prend du temps, le chargement de la configuration en est ralenti à cause de l'envoi des demandes DNS sur le réseau par les processus Java.

Une solution consiste à ajouter les adresses et les noms d'hôte du serveur au fichier local `/etc/hosts`.

Incident : Sur les systèmes Windows, un message d'erreur lié à un conflit d'adresses IP apparaît à l'écran

Si la fonction de haute disponibilité est configurée, il est possible que des adresses de cluster soient définies sur les deux systèmes pendant une courte période et que le système génère l'affichage d'un message d'erreur pour indiquer un conflit d'adresses IP avec un autre système du réseau. Dans ce cas, vous pouvez ignorer ce message. Il est possible qu'une adresse de cluster soit configurée en même temps sur les deux systèmes de haute disponibilité, notamment lors du démarrage de chaque système ou lors de l'initiation d'une procédure de reprise.

Vérifiez les scripts `go*` pour vous assurer que vous avez défini ou supprimé correctement les configurations des adresses de cluster. Si vous avez exécuté un fichier de configuration et installé des scripts `go*`, vérifiez que des commandes "executor configure" n'ont pas été définies pour les adresses de cluster dans le fichier de configuration car ces commandes entrent en conflit avec les commandes de configuration et de suppression de configuration indiquées dans les scripts `go*`.

Pour plus d'informations sur les scripts go* lors de la configuration de la fonction de haute disponibilité, voir «Utilisation de scripts», à la page 209.

Incident : les machines principale et de secours sont toutes deux activées en mode haute disponibilité

Cet incident peut survenir lorsque les scripts go ne sont pas exécutés sur la machine principale ou la machine de secours. Ces scripts ne peuvent pas s'exécuter si dsserver n'est pas démarré sur les deux machines. Vérifiez sur les deux machines que dsserver est en cours d'exécution.

Incident : Les demandes client échouent lors de la tentative de renvoi de réponses de grande page

Les demandes client générant des réponses de grande page arrivent à expiration si la taille maximale en transmission (MTU) n'est pas définie correctement sur la machine Dispatcher. En effet, pour les méthodes de transfert cbr et nat du composant Dispatcher, Dispatcher utilise la valeur MTU par défaut au lieu de la négocier.

La valeur MTU est définie sur chaque système d'exploitation en fonction du type de support de communication (par exemple, Ethernet ou Token-Ring). La valeur MTU des routeurs du segment local peut être inférieure si ces derniers se connectent à un autre type de support de communication. Dans le cas d'un trafic TCP normal, une négociation MTU survient lors de la configuration de la connexion et la valeur MTU la plus faible est utilisée pour envoyer des données entre les machines.

Dispatcher ne prend pas en charge la négociation de la valeur MTU pour sa méthode de transfert cbr ou nat car il est activement impliqué comme noeud final des connexions TCP. Pour les méthodes de transfert cbr et nat, Dispatcher utilise par défaut la valeur MTU 1500. Cette valeur correspondant à la taille type de MTU pour un réseau Ethernet standard, la plupart des clients n'ont pas besoin de la modifier.

Lorsque vous utilisez la méthode de transfert cbr ou nat de Dispatcher, si vous disposez d'un routeur vers le segment local dont la valeur MTU est la plus faible, vous devez définir la même valeur MTU sur la machine Dispatcher.

Pour résoudre cet incident, définissez la taille de segment maximale (mss) à l'aide de la commande suivante : `dscontrol executor set mss nouvelle_valeur`

Par exemple :

```
dscontrol executor set mss 1400
```

La valeur par défaut de mss est 1460.

Le paramètre mss ne s'applique pas à la méthode de transfert mac de Dispatcher ou à tout composant de Load Balancer autre que Dispatcher.

Incident : Sous Windows, l'erreur "Le serveur ne répond pas" survient lors de l'exécution d'une commande dscontrol ou lbadmin

Lorsqu'il existe plusieurs adresses IP sur un système Windows et que le fichier **hôte** ne spécifie pas l'adresse à associer au nom d'hôte, le système d'exploitation choisit d'associer au nom d'hôte l'adresse la plus petite.

Pour résoudre cet incident, remplacez le fichier `c:\Windows\system32\drivers\etc\hosts` par le nom d'hôte de votre machine et l'adresse IP à lui associer.

IMPORTANT : L'adresse IP ne peut pas correspondre à une adresse de cluster.

Incident : Les machines Dispatcher à haute disponibilité risquent de ne pas être synchronisées sur les systèmes Linux pour S/390 avec des pilotes qeth

Lors de l'utilisation de la haute disponibilité sur des systèmes Linux pour S/390 avec le pilote réseau qeth, la synchronisation des machines Dispatcher active et de secours risque d'échouer. Il se peut que cet incident soit limité au noyau 2.6 de Linux.

Si cet incident survient, procédez comme suit :

Définissez un périphérique réseau CTC (channel-to-channel- canal à canal) entre les images Dispatcher active et de secours et ajoutez un signal de présence entre les adresses IP des deux noeuds CTC.

Incident : Conseils sur la configuration de la haute disponibilité

Avec la fonction de haute disponibilité pour Load Balancer, une machine partenaire peut prendre le relais de l'équilibrage de charge en cas de défaillance ou d'arrêt du partenaire principal. Pour maintenir les connexions entre les deux partenaires, des enregistrements de connexion sont transmis entre les deux machines. Lorsque le partenaire de secours prend en charge la fonction d'équilibrage de charge, l'adresse IP de cluster est supprimée de la machine de secours et ajoutée dans la nouvelle machine principale. De nombreuses considérations temporelles et de configuration peuvent avoir un impact sur cette opération de reprise.

Les conseils répertoriés dans cette section peuvent répondre à des incidents potentiels de configuration de la haute disponibilité, comme :

- suppression des connexions après la reprise,
- synchronisation impossible des machines partenaires,
- demandes dirigées à tort vers la machine partenaire de secours.

Les conseils suivants sont pratiques pour garantir le succès de la configuration de la haute disponibilité sur les machines Load Balancer.

- L'emplacement des commandes de haute disponibilité dans vos fichiers script peut faire la différence.

Exemples de commandes de haute disponibilité :

```
dscontrol highavailability heartbeat add ...
dscontrol highavailability backup add ...
dscontrol highavailability reach add ...
```

La plupart du temps, vous devez positionner les définitions de haute disponibilité à la fin du fichier. Les instructions de cluster, port et serveur doivent être placées avant celles de haute disponibilité. En effet, lors de la synchronisation de la haute disponibilité, les définitions de cluster, port et serveur sont recherchées à la réception d'un enregistrement de connexion.

Si le cluster, port et serveur n'existent pas, l'enregistrement de connexion est annulé. En cas de reprise alors que l'enregistrement de connexion n'a pas été répliqué sur la machine partenaire, la connexion échoue.

L'exception à cette règle est l'utilisation de serveurs co-implantés, configurés avec la méthode d'acheminement MAC. Dans ce cas, les instructions de haute disponibilité doivent précéder les instructions du serveur co-implanté. Si ce n'est pas le cas, Load Balancer reçoit une demande sur le serveur co-implanté, mais celle-ci ressemble à une demande entrante adressée au cluster et fait l'objet d'un équilibrage de charge. Ceci risque de conduire à un bouclage des paquets sur le réseau et à un excès de trafic. Lorsque les instructions de haute disponibilité sont placées avant le serveur co-implanté, Load Balancer est informé qu'il ne doit pas acheminer le trafic entrant sauf s'il se trouve à l'état ACTIF.

- Sur les systèmes d'exploitation z/OS ou OS/390, l'hyperviseur contrôle l'interface et multiplexe l'interface réelle parmi les systèmes d'exploitation de l'hôte. L'hyperviseur permet seulement à un hôte à la fois de s'enregistrer sur une adresse IP, et une fenêtre de mise à jour est présente. Autrement dit, lorsque l'IP de cluster est supprimée de la machine de secours, vous devrez peut-être ajouter un délai avant d'essayer d'ajouter l'IP de cluster à la machine principale ; sinon, l'ajout échoue et les connexions entrantes ne sont pas traitées.

Pour remédier à ce comportement, ajoutez un délai de veille dans le script goActive. Le délai de veille nécessaire dépend du déploiement. Un délai de veille de 10 est recommandé pour commencer.

- Les partenaires de haute disponibilité doivent pouvoir s'envoyer des commandes PING réciproques et doivent se trouver sur le même sous-réseau. Par défaut, les machines tentent de communiquer toutes les demi-secondes et détectent une défaillance après 4 échecs. Si une machine est occupée, des échecs risquent de se produire alors que le système fonctionne encore correctement. Vous pouvez augmenter le nombre de tentatives avant échec en lançant la commande :

```
dscontrol executor set hatimeout <valeur>
```

- Lors de la synchronisation des partenaires, tous les enregistrements de connexion sont envoyés de la machine active vers la machine de secours. La synchronisation doit s'achever dans l'intervalle par défaut des 50 secondes. Pour ce faire, les anciennes connexions ne doivent pas rester en mémoire pendant une période trop longue. Des incidents ont été notamment relevés au niveau des ports LDAP et de longues périodes de `staletimeout` (de plus d'une journée). La définition d'une valeur `staletimeout` élevée entraîne la mise en mémoire des anciennes connexions, ce qui pousse un plus grand nombre d'enregistrements de connexion d'être transmis en synchronisation, et également une consommation accrue de mémoire sur les deux machines.

En cas d'échec de la synchronisation avec une période `staletimeout` raisonnable, vous pouvez augmenter le délai de synchronisation à l'aide de la commande :

```
e xm 33 5 nouveau_délai
```

Cette commande n'est pas enregistrée dans le fichier de configuration lorsqu'il est sauvegardé, donc vous devez l'ajouter manuellement dans le fichier de configuration si vous voulez que ce paramètre persiste entre les arrêts.

La valeur de dépassement de délai est enregistrée en une demi-seconde ; par conséquent, la valeur par défaut de nouveau_délai est 100 (50 secondes).

- Lorsqu'une machine partenaire prend le relais de la charge de travail, elle génère automatiquement un protocole de résolution d'adresse ATM pour indiquer aux machines se trouvant sur le même sous-réseau la nouvelle adresse matérielle associée à l'adresse IP de cluster. Vous devez vous assurer que vos routeurs répondent aux protocoles de résolution d'adresse ATM automatiques et mettent à jour leur mémoire cache, ou les demandes seront envoyées au partenaire inactif.

Remarque : Pour plus d'informations sur la configuration de la fonction de haute disponibilité, voir «Haute disponibilité», à la page 204.

Incident : Sous Linux, limitations de la configuration Dispatcher lors de l'utilisation de serveurs zSeries ou S/390 dotés de cartes OSA (Open System Adapter)

En général, avec la méthode d'acheminement MAC, les serveurs de la configuration Load Balancer doivent tous se trouver sur le même segment de réseau, quelle que soit la plateforme. Les périphériques de réseau actifs comme le routeur, les passerelles et les pare-feux interfèrent avec Load Balancer. En effet, Load Balancer fonctionne comme un routeur spécialisé, modifiant uniquement les en-têtes de couche liaison de données pour son tronçon suivant et final. Toute topologie de réseau dans laquelle le tronçon suivant n'est pas final n'est pas valide pour Load Balancer.

Remarque : Les tunnels, de type CTC (canal à canal) ou IUCV, sont souvent pris en charge. Toutefois, Load Balancer doit utiliser directement le tunnel jusqu'à la destination finale pour son acheminement (il ne peut pas s'agir d'un tunnel réseau à réseau).

Il existe une limitation pour les serveurs zSeries et S/390 qui partagent la carte OSA car cette dernière a un fonctionnement différent de la plupart des cartes réseau. La carte OSA possède sa propre implémentation de couche réseau, (sans aucun rapport avec Ethernet), qui est présentée aux hôtes Linux et z/OS se trouvant derrière elle. En réalité, les cartes OSA ressemblent à des hôtes Ethernet-Ethernet (et non à des hôtes OSA) et les hôtes qui l'utilisent y répondent comme s'il s'agissait d'Ethernet.

La carte OSA exécute également certaines fonctions directement relatives à la couche IP : par exemple, la réponse aux demandes de protocole de résolution d'adresses. Ou encore, la carte OSA partagée achemine les paquets IP en fonction d'une adresse IP de destination, et non d'une adresse Ethernet comme un commutateur de couche 2. En pratique, la carte OSA est un segment de réseau avec une passerelle vers elle-même.

Load Balancer s'exécutant sur un hôte Linux S/390 ou zSeries peut acheminer le trafic vers des hôtes se trouvant sur la même carte OSA ou sur des hôtes Ethernet. Tous les hôtes se trouvant sur la même carte OSA partagée se trouvent effectivement sur le même segment.

Load Balancer peut effectuer un *acheminement* à partir d'une carte OSA partagée en raison de la nature de la passerelle OSA. Cette dernière reconnaît le port OSA qui possède l'IP de cluster. Le pont reconnaît l'adresse MAC des hôtes directement connectés au segment Ethernet. Par conséquent, Load Balancer peut effectuer un acheminement par méthode MAC sur une passerelle OSA.

Toutefois, Load Balancer ne peut pas effectuer d'acheminement vers une carte OSA partagée, notamment lorsqu'il se trouve sous une machine S/390 Linux et que le serveur dorsal se trouve sur une carte OSA différente de celle de Load Balancer. Le processus OSA du serveur dorsal annonce l'adresse MAC OSA pour l'IP de serveur, mais à l'arrivée d'un paquet avec l'adresse de destination Ethernet de l'OSA du serveur et l'IP du cluster, la carte OSA du serveur ne sait pas lequel de ses hôtes, le cas échéant, doit recevoir ce paquet. Les principes qui gouvernent l'acheminement MAC OSA-Ethernet à partir d'une carte OSA partagée ne fonctionnent pas lors des tentatives d'acheminement vers une carte OSA partagée.

Solution :

Dans les configurations de Load Balancer qui utilisent les serveurs zSeries ou S/390 dotés de cartes OSA, vous disposez de deux moyens pour remédier à l'incident décrit.

1. Utilisation des fonctionnalités de plateforme

Si les serveurs de la configuration Load Balancer se trouvent sur le même type de plateforme zSeries ou S/390, vous pouvez définir des connexions point à point (CTC ou IUCV) entre Load Balancer et chaque serveur. Configurez les noeuds finaux avec les adresses IP privées. La connexion point à point est réservée au trafic Load Balancer-serveur. Ajoutez ensuite les serveurs avec l'adresse IP du noeud final du serveur du tunnel. Dans cette configuration, le trafic du cluster passe par la carte OSA de Load Balancer et est acheminé via la connexion point à point, là où le serveur répond par sa propre route par défaut. La réponse utilise la carte OSA du serveur pour son émission, et il peut s'agir ou non de la même carte.

2. Utilisation de la fonction GRE de Load Balancer

Remarque : Remarque : La fonction GRE n'est pas disponible dans l'environnement à double protocole de Load Balancer pour IPv4 et IPv6.

Si les serveurs de la configuration Load Balancer ne se trouvent pas sur le même type de plateforme zSeries ou S/390, ou s'il n'est pas possible de définir une connexion point à point entre Load Balancer et chaque serveur, il est préférable d'utiliser la fonctionnalité GRE (Generic Routing Encapsulation) de Load Balancer, protocole permettant à Load Balancer d'effectuer un acheminement via des routeurs.

Lors de l'utilisation de GRE, le paquet client->IP cluster est reçu par Load Balancer, encapsulé et envoyé au serveur. Au niveau du serveur, le paquet client->IP cluster d'origine est excapsulé et le serveur répond directement au client. L'avantage que présente l'utilisation de GRE est que Load Balancer visualise uniquement le trafic client-serveur, et non le trafic serveur-client. L'inconvénient est qu'elle réduit la taille de segment maximal de la connexion TCP à cause de la surcharge d'encapsulation.

Pour configurer Load Balancer pour un acheminement avec encapsulation GRE, ajoutez les serveurs à l'aide de la commande suivante :

```
dscontrol server add adresse_cluster:port:routeur_serveur_dorsal serveur_dorsal
```

où routeur_serveur_dorsal est valide si Load Balancer et le serveur dorsal se trouvent sur le même sous-réseau IP. Sinon, indiquez comme routeur l'adresse IP valide du tronçon suivant.

Pour configurer les systèmes Linux afin qu'ils exécutent une excapsulation GRE native, pour chaque serveur dorsal, lancez les commandes suivantes :

```
modprobe ip_gre
ip tunnel add grelb0 mode gre ikey 3735928559
ip link set grelb0 up
ip addr add adresse_cluster dev grelb0
```

Remarque : Ne définissez pas l'adresse de cluster sur l'unité de bouclage des serveurs dorsaux. Avec les serveurs dorsaux z/OS, vous devez utiliser les commandes z/OS pour configurer les serveurs pour une excapsulation GRE.

Incident : Sur certaines versions Linux, une fuite de mémoire se produit lors de l'exécution de Dispatcher configuré avec le gestionnaire et les conseillers

Lors de l'exécution de Load Balancer configuré avec les fonctions de gestionnaire et de conseillers, des fuites de mémoire importantes peuvent se produire sur certaines versions Red Hat Linux. La fuite de mémoire Java augmente si le paramètre d'intervalle de temps défini pour le conseiller est petit.

Les versions d'IBM Java SDK de JVM et de la bibliothèque NPTL (Native POSIX Thread Library) livrée avec certaines distributions de Linux, comme Red Hat Enterprise Linux 3.0, peuvent être à l'origine de la fuite de mémoire. La bibliothèque d'unités d'exécution améliorée NPTL est livrée avec certaines distributions de systèmes Linux, comme Red Hat Enterprise Linux 3.0.

Pour les informations les plus récentes sur les systèmes Linux et le kit IBM Java SDK livré avec ces systèmes, voir <http://www.ibm.com/developerworks/java/jdk/linux/tested.html>.

Utilisez comme outil de diagnostic la commande `vmstat` ou `ps` qui permet de détecter les fuites de mémoire.

Pour remédier à une fuite de mémoire, lancez la commande suivante avant d'exécuter la machine Load Balancer pour désactiver la bibliothèque NPTL :

```
export LD_ASSUME_KERNEL=2.4.10
```

Incident : Sur SUSE Linux Enterprise Server 9, Dispatcher achemine les paquets, mais ceux-ci n'arrivent pas jusqu'au serveur dorsal

Sur Suse Linux Enterprise Server 9, avec la méthode d'acheminement MAC, il est possible que le rapport de Dispatcher indique que le paquet a été acheminé (le nombre de paquets augmente) ; en fait, le paquet n'atteint jamais le serveur dorsal.

En présence de cet incident, l'une et/ou l'autre des situations suivantes peut se produire :

- Côté machine Dispatcher, le message suivant s'affiche :
`ip_finish_output2: No header cache and no neighbour!`
- Côté client, le message suivant s'affiche :
`ICMP Destination unreachable: Fragmentation Needed`

Cet incident peut être lié au module NAT iptables qui est chargé. Sur SLES 9, cette version d'iptables contient une erreur possible, bien que non confirmée, entraînant un comportement étrange lors des interactions avec Dispatcher.

Solution :

Déchargez le module NAT iptables et le module Connection Tracking.

Par exemple :

```
# lsmod | grep ip
    iptable_filter          3072  0
    iptable_nat             22060  0
    ip_conntrack            32560  1 iptable_nat
    ip_tables               17280  2
    iptable_filter,iptable_nat
    ipv6                   236800  19
    # rmmod iptable_nat
    # rmmod ip_conntrack
```

Supprimez les modules dans leur ordre d'utilisation. Plus précisément, vous pouvez supprimer un module uniquement si le nombre de références (dernière colonne de la sortie lsmod) est égal à zéro. Si vous avez configuré des règles dans iptables, vous devez les supprimer. Par exemple : iptables -t nat -F.

Le module iptable_nat utilisant ip_conntrack, vous devez d'abord supprimer le module iptable_nat, puis le module ip_conntrack.

Remarque : Lorsque vous essayez simplement de dresser la liste des règles configurées sur une table, le module correspondant est chargé. Par exemple : iptables -t nat -L. Veillez à ne pas lancer cette commande après la suppression des modules.

Incident : Sur le système Windows, un message de conflit d'adresses IP apparaît pendant la reprise de la haute disponibilité

Sur les systèmes Windows, si vous utilisez la fonctionnalité de haute disponibilité de Load Balancer, en cas de reprise, les goScripts permettent de configurer l'IP de cluster sur le Load Balancer actif et à annuler la configuration de l'IP de cluster sur le système de secours. Si le goScript configurant l'adresse IP de cluster sur la machine active s'exécute avant celui qui annule la configuration de l'adresse IP de cluster sur la machine de secours, des incidents risquent de se produire. Un message en incrustation risque de signaler que le système a détecté un conflit d'adresses IP. Si vous lancez la commande ipconfig /all, vous risquez également de voir l'adresse IP 0.0.0.0 sur la machine.

Solution :

Lancez la commande suivante pour annuler manuellement la configuration de l'adresse IP de cluster à partir de la machine principale :

```
dscontrol executor unconfigure clusterIP
```

L'adresse 0.0.0.0 est alors supprimée de la pile IP Windows.

Après libération de l'adresse IP de cluster par le partenaire de haute disponibilité, lancez la commande suivante pour rajouter manuellement l'IP de cluster :

```
dscontrol executor configure clusterIP
```

Une fois cette commande lancée, recherchez à nouveau l'adresse IP de cluster sur la pile IP Windows en lançant la commande suivante :

```
ipconfig /all
```

Incident : Les iptables de Linux peuvent interférer avec le routage de paquets

Les iptables de Linux peuvent interférer avec l'équilibrage de charge du trafic et doivent être désactivées sur la machine Dispatcher.

Lancez la commande suivante pour déterminer si les iptables sont chargées :

```
lsmod | grep ip_tables
```

La commande précédente peut aboutir à un résultat semblable à celui-ci :

```
ip_tables          22400    3
iptables_mangle,iptable_nat,iptable_filter
```

Lancez la commande suivante pour chaque iptable répertoriée dans le résultat pour afficher les règles des tables :

```
iptables -t <nom_court> -L
```

Par exemple :

```
iptables -t mangle -L
iptables -t nat -L
iptables -t filter -L
```

Si iptable_nat est chargé, vous devez effectuer son déchargement. Comme iptable_nat est dépendant de iptable_conntrack, supprimez également iptable_conntrack. Lancez la commande suivante pour décharger ces deux iptables :

```
rmmod iptable_nat iptable_conntrack
```

Incident : Impossible d'ajouter un serveur IPv6 à la configuration Load Balancer sur les systèmes Solaris

Sur les systèmes Solaris, lorsque vous essayez de configurer un serveur IPv6 sur une installation Load Balancer pour IPv4 et IPv6, le message Impossible d'ajouter le serveur apparaît. Cette erreur peut provenir de la gestion de la requête ping sur une adresse IPv6 par le système d'exploitation Solaris.

Sur les systèmes Solaris, lors de l'ajout d'un serveur dans la configuration, Load Balancer lance une commande ping sur le serveur pour obtenir l'adresse MAC de ce dernier. La machine Solaris peut choisir une adresse de cluster configurée comme adresse source de la requête ping au lieu d'utiliser l'adresse NFA de la machine. Si l'adresse de cluster est configurée sur l'unité de bouclage du serveur, la réponse ping n'est pas reçue sur la machine Load Balancer ; le serveur n'est alors pas ajouté dans la configuration.

La solution consiste à configurer une autre adresse IPv6 sur la machine Load Balancer avant ou après configuré l'adresse de cluster IPv6. Cette adresse doit être une adresse sans alias sur l'unité de bouclage du serveur dorsal sur lequel vous essayez d'ajouter la configuration Load Balancer. Ajoutez ensuite le serveur à la configuration Load Balancer.

Un message d'avertissement Java s'affiche lors de l'installation de correctifs de service

Load Balancer fournit un ensemble de fichiers Java avec l'installation du produit. L'installation du produit comporte plusieurs packages qu'il n'est pas nécessaire

d'installer sur la même machine (par exemple, le package Metric Server, le package d'administration et le package du produit de base). Ces trois packages de codes exigent le fonctionnement d'un ensemble de fichiers Java, mais chacun peut être installé sur des machines distinctes. Chacun de ces packages installe un ensemble de fichiers Java. S'il est installé sur la même machine, l'ensemble de fichiers Java appartient à chacun de ces ensembles de fichiers. Lorsque vous installez le deuxième et le troisième ensembles de fichiers Java, vous recevez un message d'avertissement indiquant que l'ensemble de fichiers appartient également à un autre ensemble de fichiers.

Lorsque vous installez du code à l'aide de méthodes d'installation natives (par exemple, installp sur AIX), ignorez les messages d'avertissement indiquant que l'ensemble de fichiers Java appartient à un autre ensemble de fichiers.

Mise à niveau de l'ensemble de fichiers Java fourni avec l'installation Load Balancer

Pendant le processus d'installation de Load Balancer, un ensemble de fichiers Java est également installé. Load Balancer est la seule application utilisant la version Java qui s'installe avec le produit. Ne procédez pas tout seul à la mise à niveau de cette version de l'ensemble de fichiers Java. En cas d'incident nécessitant une mise à niveau de l'ensemble de fichiers Java, contactez le service d'assistance IBM pour recevoir une mise à niveau officielle de l'ensemble de fichiers Java qui a été fourni avec l'installation Load Balancer.

Résolution des incidents courants—CBR

Incident : CBR ne fonctionne pas

Cet incident peut se produire lorsqu'une autre application utilise l'un des ports utilisés par CBR. Pour plus de détails, voir «Vérification des numéros de port CBR», à la page 299.

Incident : La commande cbrcontrol ou lbadmin n'a pas abouti

1. La commande cbrcontrol renvoie : **Erreur : Pas de réponse du serveur**. Ou la commande lbadmin renvoie : **Erreur : impossible d'accéder au serveur RMI**. Ces erreurs peuvent se manifester lorsque votre machine a une pile sur "sock". Pour corriger ce problème, éditez le fichier socks.cnf pour qu'il contienne les lignes suivantes :

```
EXCLUDE-MODULE java  
EXCLUDE-MODULE javaw
```

2. Les consoles d'administration des interfaces Load Balancer (ligne de commande, interface graphique et assistants) communiquent avec cbrserver par appels RMI (Remote Method Invocation). Par défaut, la communication utilise trois ports : chacun étant défini dans le script de démarrage de cbrserver :
 - 11099 pour recevoir des commandes de cbrcontrol
 - 10004 pour envoyer des demandes de mesure au système Metric Server
 - 11199 pour le port du serveur RMI

Ceci peut être source de problèmes lorsqu'une des consoles d'administration s'exécute sur la même machine qu'un pare-feu ou passe par un pare-feu. Par exemple, lorsque vous émettez des commandes cbrcontrol alors que Load Balancer s'exécute sur la même machine qu'un pare-feu, des erreurs de type **Erreur : Pas de réponse du serveur** peuvent s'afficher.

Pour éviter ce type d'incident, modifiez le fichier script `cbrserver` afin de définir le port qu'utilise RMI pour le pare-feu (ou autre application). Remplacez la ligne `LB_RMISERVERPORT=11199` par `LB_RMISERVERPORT=votrePort`. Où *vousPort* est un autre port.

Lorsque vous avez terminé, relancez la commande `cbrserver` et ouvrez le trafic des ports 11099, 10004, 11199 et 11100 ou du port d'adresse hôte choisi pour l'exécution de la console d'administration.

3. Ces erreurs peuvent également se produire si vous n'avez pas encore lancé **cbrserver**.

Incident : Les requêtes ne sont pas équilibrées

Caching Proxy et CBR ont été lancés, mais les requêtes ne sont pas équilibrées. Cette erreur peut se produire lorsque vous démarrez Caching Proxy avant l'exécuteur. Si c'est le cas, le journal `stderr` de Caching Proxy contient le message d'erreur indiquant l'échec de la connexion à l'exécuteur (`ndServerInit`). Pour éviter cet incident, démarrez l'exécuteur avant Caching Proxy.

Incident : Sur les systèmes Solaris, la commande `cbrcontrol executor start` n'aboutit pas

Sur les systèmes Solaris, la commande **`cbrcontrol executor start`** renvoie le message suivant : "Erreur : l'exécuteur n'a pas été lancé". Cette erreur se produit si vous ne configurez pas les communications IPC (Inter-process Communication) pour le système de telle sorte que la taille maximale d'un segment de mémoire partagée et des ID sémaphore soit supérieure à la valeur par défaut du système d'exploitation. Pour augmenter la taille du segment de mémoire partagée et des ID sémaphore, vous devez modifier le fichier `/etc/system`. Pour plus d'informations sur la configuration de ce fichier, reportez-vous à la page 113.

Incident : erreur de syntaxe ou de configuration

Si la règle d'URL ne fonctionne pas, cela peut être dû à une erreur de syntaxe ou de configuration. Pour ce problème, vérifiez :

- que la règle est correctement configurée. Pour obtenir plus de détails, voir Annexe B, «Syntaxe des règles de contenu (modèle)», à la page 477.
- Lancer un **rapport de règle `cbrcontrol`** pour cette règle et vérifier la colonne 'Exécutions, s'assurer qu'elle est incrémentée en fonction du nombre de requêtes effectuées. Si tel est le cas, vérifiez à nouveau la configuration du serveur
- Si la règle n'est pas émise, ajouter une règle 'toujours vrai'. Emettre un **rapport de règle `cbrcontrol`** sur la règle 'toujours vrais' pour vérifier qu'elle est émise.

Incident : sous Windows, comportement inattendu de l'interface graphique lors de l'utilisation de cartes vidéo Matrox AGP

Sous Windows, l'interface graphique de Load Balancer peut se comporter de manière inattendue lorsque vous utilisez une carte vidéo Matrox AGP. Lorsque vous cliquez sur un bouton de la souris, un espace légèrement plus large que le pointeur de la souris peut être altéré avec inversion possible de la mise en évidence ou déplacement des images sur l'écran. Les anciennes cartes Matrox n'ont pas présenté ce type de comportement. Il n'existe pas de rectificatif pour les cartes Matrox AGP.

Incident : Déconnexion de l'hôte lors du redimensionnement de la fenêtre du navigateur Netscape en cours d'administration Web

Lorsque vous utilisez l'administration Web à distance pour configurer Load Balancer, ne modifiez pas la taille (Réduire, Agrandir, Restaurer en bas, etc.) du navigateur Netscape dans lequel s'affiche l'interface graphique de Load Balancer. En effet, étant donné que Netscape recharge une page à chaque redimensionnement de la fenêtre du navigateur, une déconnexion de l'hôte en découle. Vous devez donc vous reconnecter à l'hôte après chaque modification de la taille de la fenêtre. Pour l'administration à distance basée sur le Web sur une plateforme Windows, utilisez Internet Explorer.

Incident : Sous Windows, des caractères nationaux Latin-1 endommagés apparaissent dans la fenêtre de la ligne de commande

Dans une fenêtre de ligne de commande du système d'exploitation Windows, certains caractères nationaux de la famille Latin-1 sont altérés. Par exemple, la lettre "a" avec tilde s'affiche sous la forme d'un symbole pi. Pour rectifier cette erreur, vous devez modifier les propriétés de police de la fenêtre de ligne de commande. Pour modifier la police, procédez comme suit :

1. Cliquez sur l'icône située dans l'angle supérieur gauche de la fenêtre de ligne de commande.
2. Sélectionnez Propriétés, puis cliquez sur l'onglet Police.
3. La police par défaut est Raster ; remplacez-la par Lucida Console, puis cliquez sur OK.

Incident : Sous HP-UX, la mémoire est insuffisante pour Java et une erreur d'unité d'exécution est générée

Certaines installations HP-UX 11i sont préconfigurées pour n'autoriser que 64 unités d'exécution par processus. Toutefois, certaines configurations Load Balancer en requièrent davantage. Pour les systèmes HP-UX, définissez les unités d'exécution par processus sur 256 au moins. Pour augmenter cette valeur, utilisez l'utilitaire "sam" pour définir le paramètre de noyau `max_thread_proc`. Si vous pensez avoir besoin d'un nombre d'unités d'exécution plus important, vous pouvez affecter à ce paramètre une valeur supérieure à 256.

Pour augmenter la valeur du paramètre `max_thread_proc`, reportez-vous à la procédure de la page 312.

Incident : Sous Windows, les conseillers et les cibles à contacter marquent tous les serveurs comme étant arrêtés

Lorsque vous configurez votre adaptateur sur une machine Load Balancer, vous devez vérifier que les deux paramètres suivants sont corrects pour que le conseiller puisse fonctionner :

- Désactivez Task Offloading, qui est généralement utilisé sur les cartes d'adaptateur 3Com.
- Activez Protocole 1 (ICMP) pour les protocoles IP si vous activez le filtrage TCP/IP. Si ICMP n'est pas activé, le test de connexion (ping) au serveur dorsal échoue.

Reportez-vous à la page 313 pour des instructions sur la configuration de ces paramètres.

Incident : Sur les systèmes Windows, la résolution de l'adresse IP en nom d'hôte n'est pas possible lorsque plusieurs adresses sont configurées sur un adaptateur

Sous Windows, lorsque vous configurez un adaptateur avec plusieurs adresses IP, configurez d'abord dans le registre l'adresse IP à affilier au nom d'hôte.

Load Balancer dépendant de `InetAddress.getLocalHost()` dans de nombreuses instances (par exemple, `lbkeys create`), l'attribution comme alias de plusieurs adresses IP à un même adaptateur peut être source d'incident. Pour éviter cela, entrez l'adresse IP à utiliser pour la résolution du nom d'hôte en premier dans le registre.

Pour connaître la procédure permettant de configurer le nom d'hôte en premier dans le registre, reportez-vous à la page 313.

Résolution des incidents courants—Site Selector

Incident : Site Selector ne s'exécute pas

Cet incident peut se produire lorsqu'une autre application utilise un des ports utilisés par Site Selector. Pour plus de détails, voir «Vérification des numéros de port Site Selector», à la page 300.

Incident : Site Selector ne permet pas le trafic à permutation circulaire à partir des clients Solaris

Symptôme : Le composant Site Selector n'effectue pas de demandes entrantes permutées de façon circulaire à partir des clients Solaris.

Cause possible : Les systèmes Solaris exécutent un démon de mémoire cache de service annuaire. Si ce démon est en cours d'exécution, la demande du programme de résolution suivante est traitée à partir de cette mémoire cache et non à partir du composant Site Selector ayant effectuée la demande.

Solution : Désactivez le démon de mémoire cache du service annuaire sur la machine Solaris.

Incident : la commande `sscontrol` ou `lbadmin` n'a pas abouti

1. La commande `sscontrol` renvoie : **Erreur : Pas de réponse du serveur**. Ou la commande `lbadmin` renvoie : **Erreur : impossible d'accéder au serveur RMI**. Ces erreurs peuvent se manifester lorsque votre machine a une pile sur "sock". Pour corriger ce problème, éditez le fichier `socks.cnf` pour qu'il contienne les lignes suivantes :

```
EXCLUDE-MODULE java
EXCLUDE-MODULE javaw
```
2. Les consoles d'administration des interfaces Load Balancer (ligne de commande, interface graphique et assistants) communiquent avec `ssserver` par appels RMI (Remote Method Invocation). Par défaut, la communication utilise trois ports : chacun étant défini dans le script de démarrage de `ssserver` :
 - 12099 pour recevoir des commandes de `sscontrol`

- 10004 pour envoyer des demandes de mesure au système Metric Server
- 12199 pour le port du serveur RMI
- 53 pour l'envoi et la réception du trafic DNS

Ceci peut être source de problèmes lorsqu'une des consoles d'administration s'exécute sur la même machine qu'un pare-feu ou passe par un pare-feu. Par exemple, lorsque vous émettez des commandes `sscontrol` alors que Load Balancer s'exécute sur la même machine qu'un pare-feu, des erreurs de type **Erreur : Pas de réponse du serveur** peuvent s'afficher.

Pour éviter ce type d'incident, modifiez le fichier script `ssserver` afin de définir le port qu'utilise RMI pour le pare-feu (ou autre application). Remplacez la ligne `LB_RMISERVERPORT=10199` par `LB_RMISERVERPORT=votrePort`. Où *vousPort* est un autre port.

Lorsque vous avez terminé, relancez la commande `ssserver` et ouvrez le trafic des ports 12099, 10004, 12199 et 12100 ou du port d'adresse hôte choisi pour l'exécution de la console d'administration.

3. Ces erreurs peuvent également se produire si vous n'avez pas encore lancé `ssserver`.

Incident : Echec du démarrage de `ssserver` sous Windows

Site Selector doit pouvoir participer à un système DNS. Toutes les machines concernées par la configuration doivent également participer à ce système. Les systèmes Windows ne nécessitent pas toujours la présence du nom d'hôte configuré dans le système DNS. Avec Site Selector, un nom d'hôte doit être défini dans le système DNS pour que le démarrage s'effectue correctement.

Vérifiez que cet hôte est défini dans le système DNS. Modifiez le fichier `ssserver.cmd` et supprimez le "w" de "javaw". Vous devriez ainsi obtenir plus d'informations sur les erreurs.

Incident : Site Selector ayant des chemins en double pour lequel l'équilibrage de charge ne s'effectue pas correctement

Le serveur annuaire de Site Selector n'est associé à aucune adresse sur la machine. Il répond aux demandes destinées aux adresses IP valides de la machine. Site Selector fait confiance au système d'exploitation pour l'acheminement de la réponse au client. Si la machine Site Selector comporte plusieurs cartes et que certaines d'entre elles sont connectées au même sous-réseau, il est possible que le système d'exploitation envoie la réponse au client à partir d'une adresse différente de celle de réception. Certaines applications client n'acceptent pas de réponse provenant d'une adresse différente de celle de l'envoi. Par conséquent, la résolution de nom semble ne pas aboutir.

Incident : sous Windows, comportement inattendu de l'interface graphique lors de l'utilisation de cartes vidéo Matrox AGP

Sous Windows, l'interface graphique de Load Balancer peut se comporter de manière inattendue lorsque vous utilisez une carte vidéo Matrox AGP. Lorsque vous cliquez sur un bouton de la souris, un espace légèrement plus large que le pointeur de la souris peut être altéré avec inversion possible de la mise en évidence ou déplacement des images sur l'écran. Les anciennes cartes Matrox n'ont pas présenté ce type de comportement. Il n'existe pas de rectificatif pour les cartes Matrox AGP.

Incident : Déconnexion de l'hôte lors du redimensionnement de la fenêtre du navigateur Netscape en cours d'administration Web

Lorsque vous utilisez l'administration Web à distance pour configurer Load Balancer, ne modifiez pas la taille (Réduire, Agrandir, Restaurer en bas, etc.) du navigateur Netscape dans lequel s'affiche l'interface graphique de Load Balancer. En effet, étant donné que Netscape recharge une page à chaque redimensionnement de la fenêtre du navigateur, une déconnexion de l'hôte en découle. Vous devez donc vous reconnecter à l'hôte après chaque modification de la taille de la fenêtre. Pour l'administration à distance basée sur le Web sur une plateforme Windows, utilisez Internet Explorer.

Incident : Sous Windows, des caractères nationaux Latin-1 endommagés apparaissent dans la fenêtre de la ligne de commande

Dans une fenêtre de ligne de commande du système d'exploitation Windows, certains caractères nationaux de la famille Latin-1 sont altérés. Par exemple, la lettre "a" avec tilde s'affiche sous la forme d'un symbole pi. Pour rectifier cette erreur, vous devez modifier les propriétés de police de la fenêtre de ligne de commande. Pour modifier la police, procédez comme suit :

1. Cliquez sur l'icône située dans l'angle supérieur gauche de la fenêtre de ligne de commande.
2. Sélectionnez Propriétés, puis cliquez sur l'onglet Police.
3. La police par défaut est Raster ; remplacez-la par Lucida Console, puis cliquez sur OK.

Incident : Sous HP-UX, la mémoire est insuffisante pour Java et une erreur d'unité d'exécution est générée

Certaines installations HP-UX 11i sont préconfigurées pour n'autoriser que 64 unités d'exécution par processus. Toutefois, certaines configurations Load Balancer en requièrent davantage. Pour les systèmes HP-UX, définissez les unités d'exécution par processus sur 256 au moins. Pour augmenter cette valeur, utilisez l'utilitaire "sam" pour définir le paramètre de noyau `max_thread_proc`. Si vous pensez avoir besoin d'un nombre d'unités d'exécution plus important, vous pouvez affecter à ce paramètre une valeur supérieure à 256.

Pour augmenter la valeur du paramètre `max_thread_proc`, reportez-vous à la procédure de la page 312.

Incident : Sous Windows, les conseillers et les cibles à contacter marquent tous les serveurs comme étant arrêtés

Lorsque vous configurez votre adaptateur sur une machine Load Balancer, vous devez vérifier que les deux paramètres suivants sont corrects pour que le conseiller puisse fonctionner :

- Désactivez Task Offloading, qui est généralement utilisé sur les cartes d'adaptateur 3Com.
- Activez Protocole 1 (ICMP) pour les protocoles IP si vous activez le filtrage TCP/IP. Si ICMP n'est pas activé, le test de connexion (ping) au serveur dorsal échoue.

Reportez-vous à la page 313 pour des instructions sur la configuration de ces paramètres.

Résolution des incidents courants—Cisco CSS Controller

Incident : ccoserver ne démarre pas

Cet incident peut se produire lorsqu'une autre application utilise un des ports employés par le serveur ccoserver de Cisco CSS Controller. Pour obtenir plus d'informations, voir «Vérification des numéros de port Cisco CSS Controller», à la page 301.

Incident : La commande ccocontrol ou lbadm n'a pas abouti

1. La commande ccocontrol renvoie : **Erreur : Pas de réponse du serveur**. Ou la commande lbadm renvoie : **Erreur : impossible d'accéder au serveur RMI**. Ces erreurs peuvent se manifester lorsque votre machine a une pile sur "sock". Pour corriger ce problème, éditez le fichier socks.cnf pour qu'il contienne les lignes suivantes :

```
EXCLUDE-MODULE java
EXCLUDE-MODULE javaw
```

2. Les consoles d'administration des interfaces Load Balancer (ligne de commande et interface graphique) communiquent avec ccoserver par appels RMI (Remote Method Invocation). Par défaut, la communication utilise trois ports : chacun étant défini dans le script de démarrage de ccoserver :

- 13099 pour recevoir des commandes de ccocontrol
- 10004 pour envoyer des demandes de mesure au système Metric Server
- 13199 pour le port du serveur RMI

Ceci peut être source de problèmes lorsqu'une des consoles d'administration s'exécute sur la même machine qu'un pare-feu ou passe par un pare-feu. Par exemple, lorsque vous émettez des commandes ccocontrol alors que Load Balancer s'exécute sur la même machine qu'un pare-feu, des erreurs de type **Erreur : Pas de réponse du serveur** peuvent s'afficher.

Pour éviter ce type d'incident, modifiez le fichier script ccoserver afin de définir le port qu'utilise RMI pour le pare-feu (ou autre application). Remplacez la ligne `CCO_RMISERVERPORT=14199` par `CCO_RMISERVERPORT=votrePort`. Où *votrePort* est un autre port.

Lorsque vous avez terminé, relancez la commande ccoserver et ouvrez le trafic des ports 13099, 10004, 13199 et 13100 ou du port d'adresse hôte choisi pour l'exécution de la console d'administration.

3. Ces erreurs peuvent également se produire si vous n'avez pas encore lancé **ccoserver**.

Incident : Impossible de créer un registre sur le port 13099

Cet incident peut se produire lorsqu'il manque une licence de produit valide. Lorsque vous tentez de lancer ccoserver, le message suivant s'affiche :

Votre licence a expiré !

Adressez-vous à votre ingénieur commercial ou à votre représentant IBM.

Pour corriger ce problème :

1. Si vous avez déjà essayé de lancer ccoserver, entrez **ccoserver stop**.

2. Copiez votre licence en cours de validité dans le répertoire
...ibm/edge/lb/servers/conf.
3. Entrez `ccoserver` pour lancer le serveur.

Incident : sous Windows, comportement inattendu de l'interface graphique lors de l'utilisation de cartes vidéo Matrox AGP

Sous Windows, l'interface graphique de Load Balancer peut se comporter de manière inattendue lorsque vous utilisez une carte vidéo Matrox AGP. Lorsque vous cliquez sur un bouton de la souris, un espace légèrement plus large que le pointeur de la souris peut être altéré avec inversion possible de la mise en évidence ou déplacement des images sur l'écran. Les anciennes cartes Matrox n'ont pas présenté ce type de comportement. Il n'existe pas de rectificatif pour les cartes Matrox AGP.

Incident : Réception d'une erreur de connexion lors de l'ajout d'un consultant

Vous pouvez être confronté à une erreur de connexion, due à des paramètres de configuration incorrects, lors de l'ajout d'un consultant. Pour corriger ce problème :

- Vérifiez que l'adresse ou la communauté indiquée correspond exactement aux valeurs configurées sur le commutateur.
- Vérifiez que la communication entre le contrôleur et le commutateur est possible.
- Vérifiez que la communauté détient les droits de lecture/écriture sur le commutateur. Le contrôleur tentera d'activer la variable `ApSvcLoadEnable` (SNMP) lors du test de la connexion pour vérifier l'accès en écriture.

Incident : Pondérations non actualisées sur le commutateur

Pour corriger ce problème :

- Si vous utilisez les mesures Nombre de connexions actives et Débit de la connexion, entrez la commande `ccocontrol service SWID:OCID:serviceIO report`. Vérifiez que les valeurs des mesures varient en fonction du débit du trafic sur le commutateur.
- Augmentez le niveau de consignation du journal du consultant et recherchez les occurrences de délai d'expiration SNMP. Si des délais d'expiration se produisent, les solutions possibles sont les suivantes :
 - Réduction de la charge sur le commutateur.
 - Diminution du délai réseau entre le commutateur et le contrôleur.
- Arrêtez puis redémarrez le consultant.

Incident : La commande de régénération n'a pas actualisé la configuration du consultant

Augmentez le niveau de consignation du journal du consultant, puis relancez la commande. Si elle échoue de nouveau, recherchez dans le journal les délais d'expirations SNMP ou autres erreurs de communication SNMP.

Incident : Déconnexion de l'hôte lors du redimensionnement de la fenêtre du navigateur Netscape en cours d'administration Web

Lorsque vous utilisez l'administration Web à distance pour configurer Load Balancer, ne modifiez pas la taille (Réduire, Agrandir, Restaurer en bas, etc.) du navigateur Netscape dans lequel s'affiche l'interface graphique de Load Balancer. En effet, étant donné que Netscape recharge une page à chaque redimensionnement de la fenêtre du navigateur, une déconnexion de l'hôte en découle. Vous devez donc vous reconnecter à l'hôte après chaque modification de la taille de la fenêtre. Pour l'administration à distance basée sur le Web sur une plateforme Windows, utilisez Internet Explorer.

Incident : Sous Windows, des caractères nationaux Latin-1 endommagés apparaissent dans la fenêtre de la ligne de commande

Dans une fenêtre de ligne de commande du système d'exploitation Windows, certains caractères nationaux de la famille Latin-1 sont altérés. Par exemple, la lettre "a" avec tilde s'affiche sous la forme d'un symbole pi. Pour rectifier cette erreur, vous devez modifier les propriétés de police de la fenêtre de ligne de commande. Pour modifier la police, procédez comme suit :

1. Cliquez sur l'icône située dans l'angle supérieur gauche de la fenêtre de ligne de commande.
2. Sélectionnez Propriétés, puis cliquez sur l'onglet Police.
3. La police par défaut est Raster ; remplacez-la par Lucida Console, puis cliquez sur OK.

Incident : Sous HP-UX, la mémoire est insuffisante pour Java et une erreur d'unité d'exécution est générée

Certaines installations HP-UX 11i sont préconfigurées pour n'autoriser que 64 unités d'exécution par processus. Toutefois, certaines configurations Load Balancer en requièrent davantage. Pour les systèmes HP-UX, définissez les unités d'exécution par processus sur 256 au moins. Pour augmenter cette valeur, utilisez l'utilitaire "sam" pour définir le paramètre de noyau max_thread_proc. Si vous pensez avoir besoin d'un nombre d'unités d'exécution plus important, vous pouvez affecter à ce paramètre une valeur supérieure à 256.

Pour augmenter la valeur du paramètre max_thread_proc, reportez-vous à la procédure de la page 312.

Résolution des incidents courants—Nortel Alteon Controller

Incident : nalserver ne démarre pas

Cet incident peut se produire lorsqu'une autre application utilise un des ports employés par le serveur nalserver de Nortel Alteon Controller. Pour plus d'informations, voir «Vérification des numéros de port Nortel Alteon Controller», à la page 301.

Incident : la commande nalcontrol ou lbadmin n'a pas abouti

1. La commande nalcontrol renvoie : **Erreur : Pas de réponse du serveur**. Ou la commande lbadmin renvoie : **Erreur : impossible d'accéder au serveur RMI**.

Ces erreurs peuvent se manifester lorsque votre machine a une pile sur "sock". Pour corriger ce problème, éditez le fichier socks.cnf pour qu'il contienne les lignes suivantes :

```
EXCLUDE-MODULE java
EXCLUDE-MODULE javaw
```

2. Les consoles d'administration des interfaces Load Balancer (ligne de commande et interface graphique) communiquent avec nalserver par appels RMI (Remote Method Invocation). Par défaut, la communication utilise trois ports : chacun étant défini dans le script de démarrage de nalserver :
 - 14099 pour recevoir les commandes de nalcontrol
 - 10004 pour envoyer des demandes de mesure au système Metric Server
 - 14199 pour le port du serveur RMI

Ceci peut être source de problèmes lorsqu'une des consoles d'administration s'exécute sur la même machine qu'un pare-feu ou passe par un pare-feu. Par exemple, lorsque vous émettez des commandes nalcontrol alors que Load Balancer s'exécute sur la même machine qu'un pare-feu, des erreurs de type **Erreur : Pas de réponse du serveur** peuvent s'afficher.

Pour éviter ce type d'incident, modifiez le fichier script nalserver afin de définir le port qu'utilise RMI pour le pare-feu (ou autre application). Remplacez la ligne `NAL_RMISERVERPORT=14199` par `NAL_RMISERVERPORT=votrePort`. Où *votrePort* est un autre port.

Lorsque vous avez terminé, relancez la commande nalserver et ouvrez le trafic des ports 14099, 10004, 14199 et 14100 ou du port d'adresse hôte choisi pour l'exécution de la console d'administration.

3. Ces erreurs peuvent également se produire si vous n'avez pas encore lancé **nalserver**.

Incident : Impossible de créer un registre sur le port 14099

Cet incident peut se produire lorsqu'il manque une licence de produit valide. Lorsque vous tentez de lancer nalserver, le message suivant s'affiche :

Votre licence a expiré !

Adressez-vous à votre ingénieur commercial ou à votre représentant IBM.

Pour corriger ce problème :

1. Si vous avez déjà essayé de lancer nalserver, entrez **nalserver stop**.
2. Copiez votre licence en cours de validité dans le répertoire `...ibm/edge/lb/servers/conf`.
3. Entrez **nalserver** pour lancer le serveur.

Incident : sous Windows, comportement inattendu de l'interface graphique lors de l'utilisation de cartes vidéo Matrox AGP

Sous Windows, l'interface graphique de Load Balancer peut se comporter de manière inattendue lorsque vous utilisez une carte vidéo Matrox AGP. Lorsque vous cliquez sur un bouton de la souris, un espace légèrement plus large que le pointeur de la souris peut être altéré avec inversion possible de la mise en évidence ou déplacement des images sur l'écran. Les anciennes cartes Matrox n'ont pas présenté ce type de comportement. Il n'existe pas de rectificatif pour les cartes Matrox AGP.

Incident : Déconnexion de l'hôte lors du redimensionnement de la fenêtre du navigateur Netscape en cours d'administration Web

Lorsque vous utilisez l'administration Web à distance pour configurer Load Balancer, ne modifiez pas la taille (Réduire, Agrandir, Restaurer en bas, etc.) du navigateur Netscape dans lequel s'affiche l'interface graphique de Load Balancer. En effet, étant donné que Netscape recharge une page à chaque redimensionnement de la fenêtre du navigateur, une déconnexion de l'hôte en découle. Vous devez donc vous reconnecter à l'hôte après chaque modification de la taille de la fenêtre. Pour l'administration à distance basée sur le Web sur une plateforme Windows, utilisez Internet Explorer.

Incident : Réception d'une erreur de connexion lors de l'ajout d'un consultant

Vous pouvez être confronté à une erreur de connexion, due à des paramètres de configuration incorrects, lors de l'ajout d'un consultant. Pour corriger ce problème :

- Vérifiez que l'adresse ou la communauté indiquée correspond exactement aux valeurs configurées sur le commutateur.
- Vérifiez que la communication entre le contrôleur et le commutateur est possible.
- Vérifiez que la communauté détient les droits de lecture/écriture sur le commutateur. Le contrôleur tentera d'activer la variable ApSvcLoadEnable (SNMP) lors du test de la connexion pour vérifier l'accès en écriture.

Incident : Pondérations non actualisées sur le commutateur

Pour corriger ce problème :

- Si vous utilisez les mesures Nombre de connexions actives et Débit de la connexion, entrez la commande `ccocontrol service SWID:OCID:serviceIO report`. Vérifiez que les valeurs des mesures varient en fonction du débit du trafic sur le commutateur.
- Augmentez le niveau de consignation du journal du consultant et recherchez les occurrences de délai d'expiration SNMP. Si des délais d'expiration se produisent, les solutions possibles sont les suivantes :
 - Réduction de la charge sur le commutateur.
 - Diminution du délai réseau entre le commutateur et le contrôleur.
- Arrêtez puis redémarrez le consultant.

Incident : La commande de régénération n'a pas actualisé la configuration du consultant

Augmentez le niveau de consignation du journal du consultant, puis relancez la commande. Si elle échoue de nouveau, recherchez dans le journal les délais d'expirations SNMP ou autres erreurs de communication SNMP.

Incident : Sous Windows, des caractères nationaux Latin-1 endommagés apparaissent dans la fenêtre de la ligne de commande

Dans une fenêtre de ligne de commande du système d'exploitation Windows, certains caractères nationaux de la famille Latin-1 sont altérés. Par exemple, la lettre "a" avec tilde s'affiche sous la forme d'un symbole pi. Pour rectifier cette

erreur, vous devez modifier les propriétés de police de la fenêtre de ligne de commande. Pour modifier la police, procédez comme suit :

1. Cliquez sur l'icône située dans l'angle supérieur gauche de la fenêtre de ligne de commande.
2. Sélectionnez Propriétés, puis cliquez sur l'onglet Police.
3. La police par défaut est Raster ; remplacez-la par Lucida Console, puis cliquez sur OK.

Incident : Sous HP-UX, la mémoire est insuffisante pour Java et une erreur d'unité d'exécution est générée

Certaines installations HP-UX 11i sont préconfigurées pour n'autoriser que 64 unités d'exécution par processus. Toutefois, certaines configurations Load Balancer en requièrent davantage. Pour les systèmes HP-UX, définissez les unités d'exécution par processus sur 256 au moins. Pour augmenter cette valeur, utilisez l'utilitaire "sam" pour définir le paramètre de noyau `max_thread_proc`. Si vous pensez avoir besoin d'un nombre d'unités d'exécution plus important, vous pouvez affecter à ce paramètre une valeur supérieure à 256.

Pour augmenter la valeur du paramètre `max_thread_proc`, reportez-vous à la procédure de la page 312.

Résolution des incidents courants—Metric Server

Incident : IOException Metric Server sous Windows lors de l'exécution de fichiers de mesures utilisateur de format .bat or .cmd

Vous devez utiliser le nom complet des mesures enregistrées par l'utilisateur sur les postes Metric Server Windows. Par exemple, vous devez indiquer **usermetric.bat**, et non **usermetric**. Le nom **usermetric** est valide sur la ligne de commande, mais est inopérant lorsqu'il est employé à partir de l'environnement d'exécution. Si vous n'utilisez pas de nom complet pour les mesures, vous recevez une exception d'entrée-sortie Metric Server. Attribuez la valeur 3 à la variable `LOG_LEVEL` dans le fichier de commandes `metricserver`, puis consultez la sortie de journal. Dans cet exemple, l'exception se présente comme suit :

```
... java.io.IOException: CreateProcess: usermetric error=2
```

Incident : Metric Server n'indique pas la charge à la machine Load Balancer

Différentes raisons peuvent expliquer pourquoi le système Metric Server ne transmet pas les informations relatives à la charge à Load Balancer. Procédez aux vérifications suivantes pour en déterminer la cause :

- Vérifiez que les fichiers de clés ont bien été transférés sur Metric Server.
- Vérifiez que le nom d'hôte de la machine Metric Server est enregistré sur le serveur de noms local.

Vous pouvez également résoudre cet incident en spécifiant le nom d'hôte dans la propriété Java `java.rmi.server.hostname` dans le script `metricserver`.

- Redémarrez avec un niveau de consignation supérieur et recherchez les erreurs.
- Sur la machine Load Balancer, réduisez le niveau de consignation du journal Metric Monitor à l'aide de la commande **dscontrol manager metric set**. Recherchez les erreurs dans le fichier `MetricMonitor.log`.

Incident : Le journal de la machine Metric Server indique qu'une signature est nécessaire pour accéder à l'agent

Le journal de la machine Metric Server présente ce message d'erreur suite au transfert de fichiers de clés sur le serveur.

Cette erreur est consignée lorsque le fichier de clés ne donne pas d'autorisation avec la clé de la paire en raison d'une altération dans cette dernière. Pour remédier à ce problème, procédez comme suit :

- Téléchargez de nouveau le fichier de clés via le protocole FTP selon la méthode de transfert binaire.
- Créez une nouvelle clé et redistribuez-la.

Incident : Sur les systèmes AIX, lorsque Metric Server s'exécute dans des conditions difficiles, il est possible que le résultat de la commande ps -vg soit altéré

Lorsque Metric Server s'exécute dans des conditions difficiles sur une plateforme AIX multiprocesseur (4.3.3, 32 bits 5.1 ou 64 bits 5.1), il est possible que le résultat de la commande ps -vg soit altéré. Par exemple :

```
55742 - A 88:19 42 18014398509449680
6396 32768 22 36 2.8 1.0 java -Xms
```

Les zones SIZE et/ou RSS de la commande ps peuvent indiquer une quantité de mémoire utilisée excessive.

Il s'agit d'un problème de noyau AIX connu. Le correctif APAR IY33804 rectifie ce problème. Ce correctif est disponible auprès du support AIX à l'adresse <http://techsupport.services.ibm.com/server/fixes> ou auprès de votre revendeur AIX.

Incident : Configuration de Metric Server dans une configuration de second niveau avec équilibrage de la charge entre des machines Dispatcher haute disponibilité par Site Selector

Dans une configuration Load Balancer de second niveau, si Site Selector (premier niveau) équilibre la charge sur une paire de partenaires Dispatcher haute disponibilité (second niveau), vous devez effectuer certaines opérations pour configurer le composant Metric Server. Vous devez configurer le serveur de mesures pour qu'il écoute une nouvelle adresse IP qui lui soit réservée. Sur les deux machines Dispatcher haute disponibilité, le serveur de mesures n'est actif que sur la machine Dispatcher active.

Pour une configuration correcte, effectuez la procédure suivante :

- Configurez le serveur de mesures pour qu'il écoute la nouvelle adresse IP locale. Il ne doit pas être autorisé à répondre sur l'adresse NFA locale. Pour des informations sur la configuration, voir «Metric Server», à la page 196.
- Site Selector ne devant communiquer qu'avec la machine Dispatcher active, vous devez démarrer, puis arrêter le serveur de mesures dans les scripts go haute disponibilité. Pour démarrer ou arrêter le serveur de mesures correctement, attribuez un alias à l'adresse IP du nouveau serveur de mesures sur la machine. Modifiez les scripts go pour transférer l'adresse IP du serveur de mesures (procédure similaire au transfert des adresses de cluster) afin que le script goActive transfère l'adresse IP du serveur de mesures de l'unité de bouclage

vers un adaptateur physique et que le script goStandby fasse l'inverse. Une fois que l'adresse IP a été transférée, le script goActive doit exécuter la commande metricserver pour démarrer le serveur de mesures. Le script goStandby doit exécuter la commande **metricserver stop** pour empêcher le serveur de mesures de communiquer avec Site Selector lorsqu'il est en mode veille.

- Sous Windows, pour transférer l'adresse IP du serveur de mesures, voir «Utilisation de scripts», à la page 209.
- Les modifications apportées au script goStandby incluent des instructions spécifiques au système d'exploitation :
 - **Systèmes HP-UX, Linux et Solaris** : Dans la section du script goStandby où l'adresse de cluster est transférée vers l'unité de bouclage, insérez les commandes permettant de transférer l'adresse IP de Metric Server vers l'unité de bouclage. Insérez ensuite la commande **metricserver stop** pour empêcher le serveur de mesures de répondre à Site Selector.
 - **Systèmes AIX** : Dans la section du script goStandby où l'adresse de cluster est transférée vers l'unité de bouclage, insérez les commandes permettant de transférer l'adresse IP du serveur de mesures vers l'unité de bouclage. Ajoutez ensuite une route de sorte que vous puissiez communiquer avec l'alias de l'unité de bouclage. Exécutez la commande `route add IPserveurmesures 127.0.0.1`. Insérez ensuite la commande **metricserver stop** pour empêcher le serveur de mesures de répondre à Site Selector. Une fois que Metric Server a été arrêté, il ne reste plus qu'à supprimer la route de l'unité de bouclage. Pour empêcher toute confusion ultérieure, insérez `route deleteIPserveurmesures`.

Par exemple :

```
ifconfig en0 delete 9.27.23.61
ifconfig lo0 alias 9.27.23.61 netmask 255.255.255.0
route add 9.27.23.61 127.0.0.1
metricserver stop
# Mode veille pendant 60 secondes au maximum ou jusqu'à l'arrêt du serveur
de mesures let loopcount=0
while [[ "$loopcount" -lt "60" && 'ps -ef |grep AgentStop|
      grep -c -v gr ep' -eq "1" ]]
do
  sleep 1
  let loopcount=loopcount+1
done
route delete 9.27.23.61
```

- **Systèmes Windows** : Installez d'abord l'unité de bouclage du serveur de mesures (appelé Local Area Connection 2 dans l'exemple ci-après) sur votre machine avec une adresse IP. Ajoutez un type d'adresse réseau privée non utilisé, tel que 10.1.1.1. Après avoir configuré l'unité de bouclage, modifiez les scripts go. Le script goStandby devra inclure la commande netsh pour transférer l'adresse IP du serveur de mesures vers l'unité de bouclage de ce dernier. Exécutez ensuite la commande **metricserver stop**.

Par exemple :

```
call netsh interface ip delete address "Local Area Connection" addr=9.27.23.61
call netsh interface ip add address "Local Area Connection 2" addr=9.27.23.61
mask = 255.255.255.0
sleep 3
metricserver stop
```


Incident : Les scripts exécutés sur des machines Solaris dotées de plusieurs CPU génèrent des messages de console non souhaités

Lors de leur exécution sur des machines Solaris dotées de plusieurs CPU, les scripts `metricserver`, `cpuload` et `memload` peuvent générer des messages de console non souhaités. Ce comportement est dû à l'utilisation de la commande système `VMSTAT` pour collecter des statistiques sur la CPU et la mémoire à partir du noyau. Certains messages renvoyés par `VMSTAT` indiquent que l'état du noyau a changé. Les scripts ne peuvent pas gérer ces messages, ce qui entraîne la génération de messages de console inutiles provenant du shell.

Exemples de ces messages de la console :

```
/opt/ibm/edge/lb/ms/script/memload[29]: TOTAL=: erreur de syntaxe
/opt/ibm/edge/lb/ms/script/memload[31]: LOAD=4*100/0: diviser par zéro
/opt/ibm/edge/lb/ms/script/memload[29]: TOTAL=659664+: plus de tokens attendus
```

Ces messages peuvent être ignorés.

Incident : Avec Load Balancer pour IPv6, extraction impossible de valeurs de Metric Server sur des systèmes Linux

Lors de l'exécution sur les plateformes Linux, il existe une incompatibilité dans la sélection de l'adresse IPv6 source. Metric Monitor tente alors de communiquer avec Metric Server via l'adresse IP source erronée.

Sur les systèmes Linux, l'adresse source IPv6 sélectionnée par défaut pour une route donnée est la dernière adresse configurée correspondant à la partie réseau de la route.

Si un cluster IPv6 est la dernière interface configurée et que celle-ci correspond à la partie réseau d'une route dans la table de routage, cette interface est utilisée comme adresse IP source par défaut pour cette route. Si cette route est utilisée entre Load Balancer et Metric Server, les communications entre les deux noeuds ne sont pas établies.

La communication n'est pas établie car le noeud Load Balancer tente de communiquer avec Metric Server en utilisant l'adresse de cluster comme adresse IP source. Lorsque le cluster est configuré sur l'unité de bouclage du noeud Metric Server, la réponse de Metric Server arrive à l'unité de bouclage et la communication n'est pas établie.

Solution :

Pour savoir quelle adresse utilise le noeud Linux pour une route donnée et quelle interface est utilisée pour les communications RMI entre Metric Monitor et Metric Server, lancez la commande suivante :

```
ip -6 route get your_ipv6_route
```

Par exemple, lorsque vous lancez la commande :

```
ip -6 route get fec0::/64
```

La réponse suivante est renvoyée :

```
fec0:: via fec0:: dev eth0 src fec0::4 metric 0 cache mtu 1500 advmss 1383
```

Si fec0::4 est une adresse de cluster, une autre interface doit être ajoutée au périphérique pour empêcher l'utilisation du cluster comme source par défaut, ou une interface non cluster antérieure peut être supprimée, puis rajoutée.

Par exemple :

```
ip -6 addr add fec0::5/64 dev eth0
```

Incident : Après le démarrage de Metric Server, la valeur de mesure renvoie -1

Cet incident peut résulter de la perte d'intégrité des fichiers de clés pendant le transfert vers le client.

Si vous utilisez FTP pour transférer les fichiers de clés entre la machine Load Balancer et le serveur dorsal, assurez-vous que vous utilisez le mode binaire pour placer ou récupérer les fichiers de clés dans le serveur FTP.

Partie 9. Guide des commandes

Cette section contient des informations relatives aux commandes de tous les composants de Load Balancer. Elle se compose des chapitres suivants :

- Chapitre 26, «Lecture d'un schéma de syntaxe», à la page 343
- Chapitre 27, «Guide des commandes Dispatcher et CBR», à la page 345
- Chapitre 28, «Guide des commandes Site Selector», à la page 401
- Chapitre 29, «Guide des commandes Cisco CSS Controller», à la page 429
- Chapitre 30, «Guide des commandes Nortel Alteon Controller», à la page 449

Chapitre 26. Lecture d'un schéma de syntaxe

Le schéma de syntaxe explique comment indiquer une commande de sorte que le système d'exploitation puisse interpréter correctement les données que vous entrez. Lisez le schéma de syntaxe de gauche à droite et de haut en bas, suivant la ligne horizontale (chemin principal).

Symboles et ponctuation

Les symboles suivants sont utilisés dans les schémas de syntaxe :

Symbole

Description

- » Marque le début de la syntaxe de la commande.
- « Marque la fin de la syntaxe de la commande.

Vous devez inclure tous les signes de ponctuation tels que le deux-points, le point d'interrogation et le signe moins, qui sont indiqués dans le schéma de syntaxe.

Paramètres

Les types de paramètres suivants sont utilisés dans les schémas de syntaxe :

Paramètre

Description

Obligatoire

Les paramètres obligatoires s'affichent dans le chemin principal.

Facultatif

Les paramètres facultatifs s'affichent sous le chemin principal.

Les paramètres sont classés par mots clés ou par variables. Les mots clés s'affichent en minuscules et peuvent être entrés en minuscules. Par exemple, un nom de commande est un mot clé. Les variables apparaissent en italique et représentent des noms ou des valeurs que vous indiquez.

Exemples de syntaxe

Dans l'exemple suivant, la commande de l'utilisateur correspond à un mot clé. La variable obligatoire est *id_util* et la variable facultative *mot_de_passe*. Remplacez les variables par vos propres valeurs.

»—utilisateur—*id_util*— —«
 mot_de_passe

Mots clés obligatoires : Les mots clés et les variables obligatoires apparaissent sur la ligne du chemin principal.

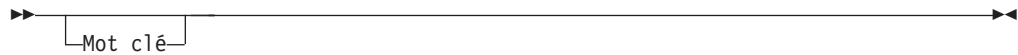
»—mot_clé_obligatoire—«

Vous devez codifier les mots clés et les valeurs obligatoires.

Choisissez un élément obligatoire dans une pile : Si vous devez effectuer votre choix parmi plusieurs mots clés ou variables obligatoires qui s'excluent, ceux-ci sont empilés verticalement dans l'ordre alphanumérique.

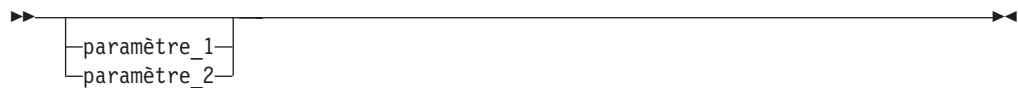


Valeurs facultatives : Les mots clés et les variables facultatifs apparaissent sous la ligne du chemin principal.



Vous pouvez choisir de ne pas codifier les mots clés et les variables facultatifs.

Choisissez un mot clé facultatif dans une pile : Si vous devez effectuer votre choix parmi plusieurs mots clés ou variables facultatifs qui s'excluent, ceux-ci sont empilés verticalement dans l'ordre alphanumérique sous la ligne du chemin principal.



Variables : Un mot apparaissant intégralement en italique correspond à une *variable*. Lorsque la syntaxe comporte une variable, celle-ci doit être remplacée par un nom ou une valeur admise, comme cela est défini dans le texte.



Caractères non alphanumériques : Si un schéma indique un caractère non alphanumérique (par exemple, deux-point, des guillemets ou le signe moins), ce dernier doit être codifié dans le cadre de la syntaxe. Dans cet exemple, vous devez codifier *cluster:port*.



Chapitre 27. Guide des commandes Dispatcher et CBR

Le présent chapitre décrit l'utilisation des commandes **dscontrol** de Dispatcher. Il traite également des commandes de CBR.

Dans les versions antérieures où le produit se nommait Network Dispatcher, la commande de contrôle de Dispatcher était **ndcontrol**. Elle s'intitule désormais **dscontrol**. Veillez à mettre à jour tous les fichiers script précédents pour utiliser la commande **dscontrol** (au lieu de **ndcontrol**) et configurer Dispatcher.

CBR utilise un sous-ensemble des commandes Dispatcher répertoriées dans ce guide des commandes. Lorsque vous utilisez ces diagrammes de syntaxe pour **CBR**, remplacez **dscontrol** par **cbrcontrol**. Pour plus d'informations, voir «Différences de configuration entre CBR et Dispatcher», à la page 346.

IMPORTANT : Si vous utilisez l'installation Load Balancer pour IPv4 et IPv6 de ce produit, seul le composant Dispatcher est disponible. Le composant Dispatcher de ce type d'installation utilise un sous-ensemble des commandes **dscontrol** répertoriées dans ce guide des commandes. Lors de l'utilisation de ces diagrammes de syntaxe, remplacez le signe deux-points (:) par le signe at (@) comme délimiteur de la commande **dscontrol**. Pour plus d'informations, voir «Différences entre les syntaxes des commandes», à la page 92 et «Commandes **dscontrol** prises en charge», à la page 92 au sujet de l'installation de Load Balancer pour IPv4 et IPv6.

La liste suivante contient les commandes relevées dans le présent chapitre :

- «**dscontrol advisor** — Contrôle du conseiller», à la page 347
- «**dscontrol binlog** — Contrôle du fichier journal binaire», à la page 352
- «**dscontrol cluster** — Configuration des clusters», à la page 353
- «**dscontrol executor** — Contrôle de l'exécuteur», à la page 357
- «**dscontrol file** — Gestion des fichiers de configuration», à la page 362
- «**dscontrol help** — Affichage ou impression de l'aide relative à cette commande», à la page 364
- «**dscontrol highavailability** — Contrôle de la haute disponibilité», à la page 365
- «**dscontrol host** — Configuration d'une machine éloignée», à la page 369
- «**dscontrol logstatus** — Affichage des paramètres du journal du serveur», à la page 370
- «**dscontrol manager** — Contrôle du gestionnaire», à la page 371
- «**dscontrol metric** — Configuration des mesures du système», à la page 377
- «**dscontrol port** — Configuration des ports», à la page 378
- «**dscontrol rule** — Configuration des règles», à la page 384
- «**dscontrol server** — Configuration des serveurs», à la page 390
- «**dscontrol set** — Configuration du journal du serveur», à la page 396
- «**dscontrol status** — Indique par affichage si le gestionnaire et les conseillers sont en cours d'exécution», à la page 397
- «**dscontrol subagent** — Configuration du sous-agent SNMP», à la page 398

Vous pouvez entrer une version abrégée des paramètres de commande `dscontrol`. Il suffit d'entrer les lettres spécifiques des paramètres. Par exemple, pour obtenir l'aide correspondant à la commande `file save`, vous pouvez entrer **`dscontrol he f`** au lieu de **`dscontrol help file`**.

Pour démarrer l'interface de ligne de commande, entrez **`dscontrol`** pour ouvrir une invite `dscontrol`.

Pour fermer l'interface de ligne de commande, entrez **`exit`** ou **`quit`**.

Les valeurs des paramètres de commandes doivent être saisies à l'aide de caractères anglais. Les seules exceptions s'appliquent aux noms d'hôte (utilisés dans les commandes `cluster`, `server` et `highavailability`) et aux noms de fichiers (utilisés dans les commandes `file`).

Différences de configuration entre CBR et Dispatcher

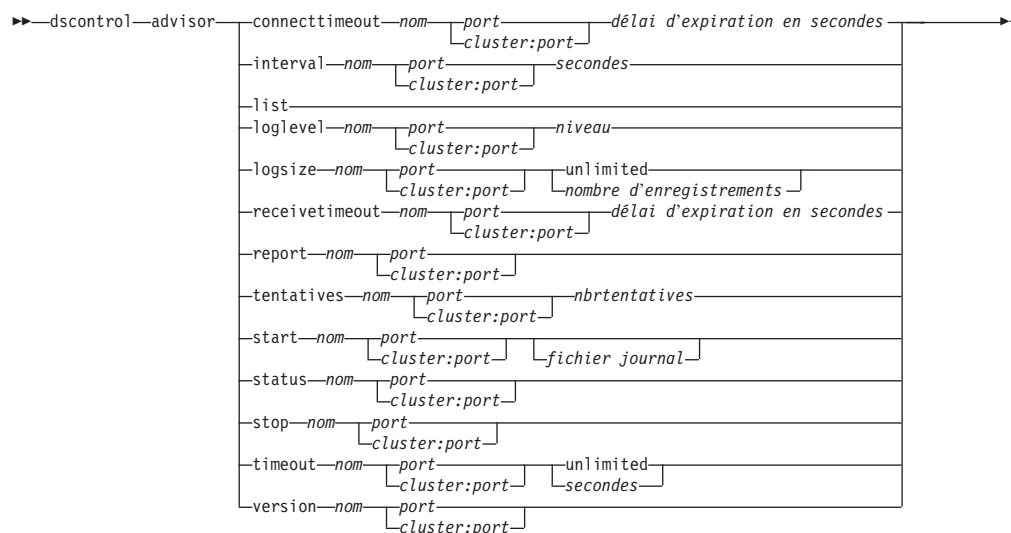
L'interface de ligne de commande de CBR est un sous-ensemble de l'interface de ligne de commande de Dispatcher. Pour CBR, remplacez la commande `dscontrol` par la commande **`cbrcontrol`** pour configurer le composant.

Remarque : Le composant Content Based Routing (CBR) est disponible sur toutes les plateformes prises en charge à l'exception de celles exécutant une JVM 64 bits. Vous avez également la possibilité d'employer la méthode d'acheminement `cbr` du composant Dispatcher de Load Balancer pour permettre le routage CBR sans faire appel à Caching Proxy. Pour plus de détails, voir «Fonction CBR de Dispatcher (méthode d'acheminement `cbr`)», à la page 55.

Certaines des commandes *inutilisées* dans CBR sont répertoriées ci-dessous.

1. `highavailability`
2. `subagent`
3. `executor`
 - `report`
 - `set nfa <valeur>`
 - `set fintimeout <valeur>`
 - `set hatimeout <valeur>`
 - `set hasynctimeout <valeur>`
 - `set porttype <valeur>`
4. `cluster`
 - `report {c}`
 - `set {c} porttype`
5. `port`
 - `add {c:p} porttype`
 - `add {c:p} protocol`
 - `set {c:p} porttype`
6. `rule add {c:p:r} type port`
7. `server`
 - `add {c:p:s} router`
 - `set {c:p:s} router`

dscontrol advisor — Contrôle du conseiller



connecttimeout

Permet de définir le délai à l'expiration duquel un conseiller signale qu'une connexion à un serveur pour un port particulier d'un serveur (d'un service) a échoué. Pour plus d'informations, voir «Délai de connexion du conseiller et délai de réception pour les serveurs», à la page 188.

nom

Nom du conseiller. Les valeurs possibles sont **connect**, **db2**, **dns**, **ftp**, **http**, **https**, **cachingproxy**, **imap**, **ldap**, **nntp**, **ping**, **pop3**, **self**, **sip**, **smtp**, **ssl**, **ssl2http**, **telnet** et **wlm**.

Pour plus d'informations sur les conseillers que fournit Load Balancer, voir «Liste des conseillers», à la page 188.

Les noms des conseillers personnalisés sont au format xxxx, ADV_xxxx étant le nom de la classe mettant en oeuvre le conseiller personnalisé. Pour plus d'informations, voir «Création de conseillers personnalisés», à la page 192.

port

Numéro du port contrôlé par le conseiller.

cluster:port

La valeur de cluster est facultative dans les commandes du conseiller, mais la valeur de port est requise. Si la valeur de cluster n'est pas indiquée, le conseiller s'exécutera sur le port de tous les clusters. Si vous indiquez un cluster, le conseiller s'exécutera sur le port uniquement pour le cluster spécifié. Pour plus d'informations, voir «Démarrage et arrêt d'un conseiller», à la page 186.

Le cluster correspond à l'adresse IP ou au nom symbolique. Le port correspond au numéro du port que le conseiller surveille.

délai d'expiration en secondes

Il s'agit d'un entier positif représentant la période en secondes pendant laquelle le conseiller attend avant de signaler qu'une connexion à un serveur a échoué. La valeur par défaut est trois fois la valeur spécifiée pour l'intervalle du conseiller.

interval

Définit la fréquence à laquelle le conseiller demande des informations aux serveurs.

secondes

Il s'agit d'un entier positif qui représente le nombre de secondes entre les demandes envoyées aux serveurs pour connaître leurs états en cours. Valeur par défaut : 7.

list

Affiche la liste des conseillers qui fournissent des informations au gestionnaire.

loglevel

Définit le niveau de consignation relatif à un journal de conseiller.

niveau

Valeur du niveau (0 à 5). La valeur par défaut est 1. Plus la valeur est élevée, plus la quantité d'informations consignée dans le journal du conseiller est importante. Les valeurs possibles sont les suivantes : 0 (Aucun), 1 (Minimal), 2 (De base), 3 (Modéré), 4 (Avancé), 5 (Prolixe).

logsize

Définit la taille maximale d'un journal de conseiller. Lorsque vous attribuez une taille maximale au fichier journal, ce dernier fonctionne en boucle. Lorsque le fichier atteint la taille indiquée, les entrées suivantes sont écrites à partir du haut du fichier et remplacent les entrées existantes. La taille du journal ne peut pas être moins élevée que la taille actuelle du journal. Les entrées du journal sont horodatées, ce qui permet de déterminer l'ordre dans lequel elles ont été enregistrées. Plus le niveau de consignation est élevé, plus la taille du journal doit être choisie attentivement car l'espace peut être saturé rapidement lors d'une consignation à des niveaux plus élevés.

nombre d'enregistrements

Taille maximale (en octets) du fichier journal du conseiller. Vous pouvez indiquer un nombre positif supérieur à zéro, ou le mot **unlimited**. Il se peut que le fichier journal n'atteigne pas la taille maximale exacte avant l'écrasement, car la taille des entrées de journal elles-mêmes varie. La valeur par défaut est 1 Mo.

receivetimeout

Permet de définir le délai à l'expiration duquel un conseiller signale que la réception d'un envoi provenant d'un port particulier d'un serveur (d'un service) a échoué. Pour plus d'informations, voir «Délai de connexion du conseiller et délai de réception pour les serveurs», à la page 188.

délai d'expiration en secondes

Il s'agit d'un entier positif qui représente la période en secondes pendant laquelle le conseiller attend avant de signaler qu'une réception d'un envoi provenant d'un serveur a échoué. La valeur par défaut est trois fois la valeur spécifiée pour l'intervalle du conseiller.

report

Affiche un rapport sur l'état du conseiller.

tentative

Nombre de tentatives accordées à un conseiller avant de déclarer un serveur arrêté.

nbrtentatives

Entier supérieur ou égal à zéro. Il est préférable que le nombre de tentatives ne dépasse pas 3. Par défaut, le nombre de tentatives est égal à zéro.

start

Lance le conseiller. Il existe des conseillers pour chaque protocole. Les ports par défaut sont les suivants :

Nom du conseiller	Protocole	Port
cachingproxy	HTTP (via Caching Proxy)	80
connect	ICMP	12345
db2	privé	50000
dns	DNS	53
ftp	FTP	21
http	HTTP	80
https	SSL	443
imap	IMAP	143
ldap	LDAP	389
nntp	NNTP	119
ping	PING	0
pop3	POP3	110
self	privé	12345
sip	SIP	5060
smtp	SMTP	25
ssl	SSL	443
ssl2http	SSL	443
telnet	Telnet	23
WLM	privé	10,007

Remarque : Le conseiller FTP doit être activé uniquement pour le port de contrôle FTP (21). Ne démarrez pas un conseiller FTP sur le port de données FTP (20).

fichier journal

Nom du fichier dans lequel les données de gestion sont consignées. Chaque enregistrement du journal est horodaté.

Le fichier par défaut se présente sous la forme *nomconseiller_port.log*, par exemple, **http_80.log**. Pour changer le répertoire dans lequel les fichiers journaux sont enregistrés, voir «Modification des chemins des fichiers journaux», à la page 267. Les fichiers journaux par défaut des conseillers spécifiques de clusters (ou de sites) sont créés avec l'adresse du cluster, comme **http_127.40.50.1_80.log**.

status

Affiche l'état en cours de toutes les valeurs d'un conseiller qui peuvent être affectées globalement, ainsi que les valeurs par défaut associées.

stop

Arrête le conseiller.

timeout

Définit le nombre de secondes pour lequel le gestionnaire considère que les informations provenant du conseiller sont valides. Si le gestionnaire considère que les informations du conseiller sont antérieures à ce délai, il n'utilise pas ces

informations pour déterminer les pondérations relatives aux serveurs sur le port contrôlé par le conseiller. Il est fait exception à ce délai lorsque le conseiller a informé le gestionnaire qu'un serveur spécifique est hors service. Le gestionnaire utilise ces informations relatives au serveur même après le dépassement du délai imparti aux informations du conseiller.

secondes

Nombre positif représentant le nombre de secondes, ou le mot **unlimited**. La valeur par défaut est unlimited.

version

Affiche la version en cours du conseiller.

Exemples

- Pour démarrer le conseiller http sur le port 80 pour le cluster 127.40.50.1, entrez :
`dscontrol advisor start http 127.40.50.1:80`
- Pour démarrer le conseiller http sur le port 88 pour tous les clusters, entrez :
`dscontrol advisor start http 88`
- Pour arrêter le conseiller http sur le port 80 pour le cluster 127.40.50.1, entrez :
`dscontrol advisor stop http 127.40.50.1:80`
- Pour définir la durée (30 secondes) pendant laquelle un conseiller HTTP du port 80 attend avant de signaler qu'une connexion à un serveur a échoué, entrez :
`dscontrol advisor connecttimeout http 80 30`
- Pour définir la durée (20 secondes) pendant laquelle un conseiller HTTP du port 80 sur le cluster 127.40.50.1 attend avant de signaler qu'une connexion à un serveur a échoué, entrez :
`dscontrol advisor connecttimeout http 127.40.50.1:80 20`
- Pour associer à l'intervalle du conseiller FTP (pour le port 21) la valeur de 6 secondes, entrez :
`dscontrol advisor interval ftp 21 6`
- Pour afficher la liste des conseillers qui fournissent des informations au gestionnaire, entrez :
`dscontrol advisor list`

Cette commande génère des résultats similaires à l'exemple suivant :

CONSEILLER	CLUSTER:PORT	DELAI	

http	127.40.50.1:80	unlimited	
ftp	21	unlimited	

- Pour remplacer le niveau de consignation du journal du conseiller par 0, afin d'optimiser les performances, entrez :
`dscontrol advisor loglevel http 80 0`
- Pour attribuer à la taille du journal du conseiller ftp pour le port 21 la valeur de 5000 octets, entrez :
`dscontrol advisor logsize ftp 21 5000`
- Pour définir la durée (60 secondes) pendant laquelle un conseiller HTTP (pour le port 80) attend avant de signaler qu'une réception d'un envoi provenant d'un serveur a échoué, entrez :
`dscontrol advisor receivetimeout http 80 60`
- Pour afficher un rapport sur l'état du conseiller ftp (pour le port 21), entrez :
`dscontrol advisor report ftp 21`

Cette commande génère des résultats similaires à l'exemple suivant :

Rapport du conseiller :

Nom du conseiller Ftp
Numéro du port 21

Adresse de cluster 9.67.131.18
Adresse du serveur 9.67.129.230
Charge 8

Adresse de cluster 9.67.131.18
Adresse du serveur 9.67.131.215
Charge -1

- Pour afficher l'état actuel des valeurs associées au conseiller http pour le port 80, entrez :

```
dscontrol advisor status  
http 80
```

Cette commande génère des résultats similaires à l'exemple suivant :

Etat du conseiller (advisor) :

Intervalle (secondes)..... 7
Délai d'expiration (secondes) Unlimited
Délai d'expiration de connexion (secondes)... 21
Délai d'expiration de réception (secondes)... 21
Nom du fichier journal du conseiller Http_80.log
Niveau de consignation 1
Taille maximale du journal (octets) Unlimited
Nombre de tentatives 0

- Pour associer la valeur de 5 secondes au délai d'attente des informations du conseiller ftp sur le port 21, entrez :

```
dscontrol advisor timeout ftp 21 5
```

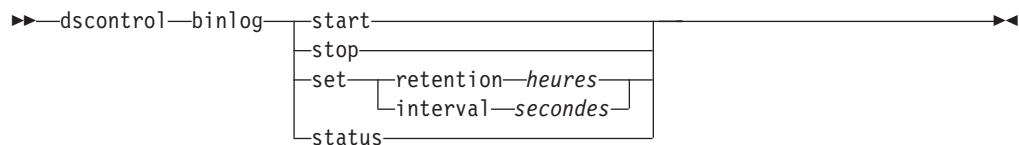
- Pour afficher le numéro de la version actuelle du conseiller ssl pour le port 443, entrez :

```
dscontrol advisor version ssl 443
```

Cette commande génère des résultats similaires à l'exemple suivant :

Version : 04.00.00.00 - 07/12/2001-10:09:56-EDT

dscontrol binlog — Contrôle du fichier journal binaire



start

Lance la consignation binaire.

stop

Arrête la consignation

set

Définit les zones de consignation. Pour plus d'informations sur la définition des zones pour la consignation binaire, voir «Utilisation de la consignation binaire pour analyser les statistiques des serveurs», à la page 240.

rétenition

Nombre d'heures pendant lesquelles les fichiers journaux sont conservés. La valeur de rétenition par défaut est 24.

heures

Nombre d'heures.

interval

Nombre de secondes qui s'écoulent entre deux entrées de journal. La valeur par défaut est 60.

secondes

Nombre de secondes.

status

Affiche la rétenition et les intervalles du journal.

dscontrol cluster — Configuration des clusters



add

Ajoute ce cluster. Vous devez définir au moins un cluster.

cluster

Nom ou adresse du cluster auquel les clients se connectent sous forme de nom symbolique ou d'adresse IP. Il est possible d'utiliser la valeur 0.0.0.0 pour spécifier un cluster générique. «Utilisation d'un cluster générique pour combiner les configurations serveurs», à la page 236.

Vous pouvez utiliser le signe deux-points (:) comme caractère générique sauf pour la commande dscontrol cluster add. Par exemple, la commande dscontrol cluster set : weightbound 80 permet de définir une pondération de 80 pour tous les clusters.

Remarque : Chaque cluster supplémentaire doit être séparé du précédent par le signe plus (+).

address

Adresse IP unique de la machine TCP sous forme de nom d'hôte ou d'adresse IP. S'il n'est pas possible de résoudre la valeur du cluster, vous devez fournir l'adresse IP de la machine physique.

Remarque : Address ne s'applique qu'au composant Dispatcher.

address

Valeur de l'adresse du cluster.

proportions

Au niveau du cluster, définit le niveau d'importance (proportion) des connexions actives (*actives*), des nouvelles connexions (*nouvelles*), des informations en provenance des conseillers (*port*) et des informations provenant d'un programme de contrôle système tel que Metric Server (*système*) utilisées par le gestionnaire pour définir les pondérations des serveurs. Chacune des valeurs décrites ci-après est exprimée en pourcentage de la valeur totale, et par conséquent, leur somme est toujours égale à 100. Pour plus d'informations, voir «Proportion de l'importance accordée aux données d'état», à la page 180.

activ.

Nombre compris entre 0 et 100 représentant la proportion de pondération à affecter aux connexions actives. Valeur par défaut : 50

new

Nombre compris entre 0 et 100 représentant la proportion de pondération à affecter aux nouvelles connexions. Valeur par défaut : 50

port

Nombre compris entre 0 et 100 représentant la proportion de pondération à affecter aux informations provenant des conseillers. La valeur par défaut est 0.

Remarque : Lorsqu'un conseiller est démarré et que la proportion du port est 0, Load Balancer définit automatiquement la valeur 1 pour que le gestionnaire puisse utiliser les informations du conseiller en tant qu'entrée pour calculer la pondération du serveur.

system

Nombre compris entre 0 et 100 représentant la proportion de pondération à affecter aux informations provenant d'une mesure système, par exemple, Metric Server. La valeur par défaut est 0.

maxports

Nombre maximal de ports. La valeur par défaut du paramètre maxports est 8.

size

Nombre de ports autorisés

maxservers

Nombre maximal par défaut de serveurs par port. Ce paramètre peut être remplacé pour des ports spécifiques, à l'aide de la commande **port maxservers**. La valeur par défaut de maxservers est 32.

size

Nombre de serveurs autorisés par port

stickytime

Délai de maintien de routage par défaut pour les ports à créer. Ce paramètre peut être remplacé pour des ports individuels, à l'aide de la commande **port stickytime**. La valeur par défaut du délai de maintien de routage est 0. La valeur par défaut du délai de maintien de routage est 0.

Remarque : Pour la méthode d'acheminement CBR de Dispatcher, si le délai de maintien de routage est différent de zéro, le délai de maintien de routage est activé pour les ports SSL (non HTTP). L'affinité SSL ID est activée si le port ajouté est un port SSL et que le délai de maintien de routage des ports à créer est différent de zéro. Pour la désactiver, vous devez associer de façon explicite un délai de maintien de routage de 0 au port.

time

Valeur du paramètre stickytime en secondes.

weightbound

Limite de pondération par défaut du port. Ce paramètre peut être remplacé pour des ports individuels, à l'aide de la commande **port weightbound**. La valeur par défaut de la pondération est de 20.

weight

Valeur de la limite de pondération

porttype

Type de port par défaut. Ce paramètre peut être remplacé pour des ports individuels, à l'aide de la commande **port porttype**.

type

La valeurs possibles sont **tcp**, **udp** et **both**.

primaryhost

Adresse NFA des machines Dispatcher, principale et de secours. Dans une configuration de haute disponibilité réciproque, le cluster est associé soit à la machine principale, soit à celle de secours.

Si vous modifiez l'hôte principal d'un cluster alors que les machines principale et de secours sont lancées et exécutées en haute disponibilité réciproque, vous devez contraindre le nouvel hôte principal à prendre le relais. Vous devrez ensuite mettre à jour les scripts puis déconfigurer et reconfigurer le cluster correctement. Pour plus d'informations, voir «Haute disponibilité réciproque», à la page 61.

address

Valeur de l'adresse du primaryhost. Elle correspond par défaut à l'adresse NFA de cette machine.

staletimeout

Nombre de secondes d'inactivité possible sur une connexion avant que cette dernière soit supprimée. La valeur par défaut pour FTP est de 900 et celle pour Telnet de 32 000 000. La valeur par défaut pour tous les autres protocoles est 300. Ce paramètre peut être remplacé pour des ports individuels, à l'aide de la commande **port staletimeout**. Pour plus d'informations, voir «Utilisation de la valeur du délai d'attente», à la page 268.

staletimout

Valeur du paramètre staletimeout.

sharedbandwidth

Quantité maximale de bande passante (en kilo-octets par seconde) pouvant être partagée au niveau du cluster. Pour plus d'informations sur la bande passante partagée, voir «Utilisation de règles basées sur la largeur de bande réservée et sur la largeur de bande partagée», à la page 216 et «Règle de largeur de bande partagée», à la page 216.

Remarque : Shared bandwidth ne s'applique qu'au composant Dispatcher.

size

La taille de la **bande passante partagée** correspond à un entier. La valeur par défaut est zéro. Si la valeur est zéro, la bande passante ne peut pas être partagée au niveau du cluster.

set

Définit les caractéristiques du cluster.

remove

Supprime ce cluster.

report

Affiche les zones internes du cluster.

Remarque : Report ne s'applique qu'au composant Dispatcher.

status

Affiche l'état en cours d'un cluster spécifique.

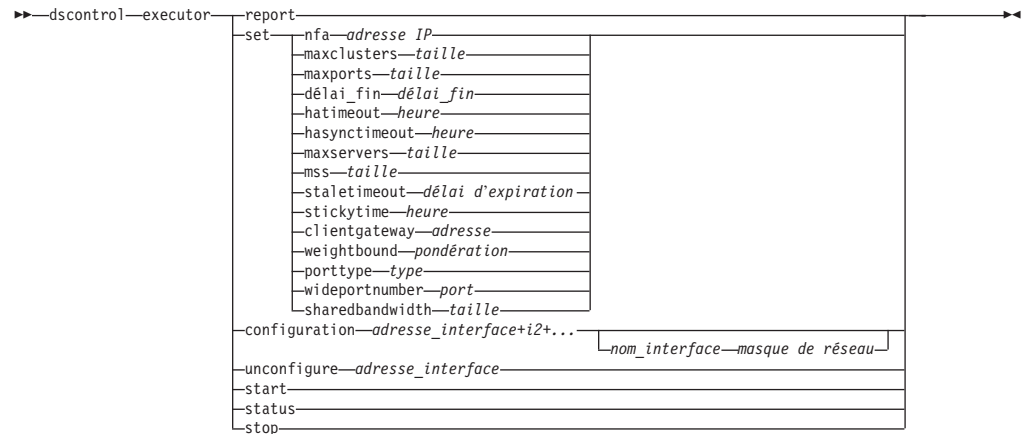
Exemples

- Pour ajouter l'adresse de cluster 130.40.52.153, entrez :
`dscontrol cluster add 130.40.52.153`
- Pour supprimer l'adresse de cluster 130.40.52.153, entrez :
`dscontrol cluster remove 130.40.52.153`
- Pour définir l'importance relative associée aux données (actives, nouvelles, de port, système) reçues par les serveurs se trouvant sur le cluster 9.6.54.12, entrez :
`dscontrol cluster set 9.6.54.12 proportions 60 35 5 0`
- Ajout d'un cluster générique
`dscontrol cluster add 0.0.0.0`
- Dans le cas d'une configuration de haute disponibilité réciproque, placez le NFA de la machine de secours (9.65.70.19) en position d'hôte principal dans l'adresse de cluster 9.6.54.12 :
`dscontrol cluster set 9.6.54.12 primaryhost 9.65.70.19`
- Pour visualiser l'état de l'adresse de cluster 9.67.131.167, entrez :
`dscontrol cluster status 9.67.131.167`

Cette commande génère des résultats similaires à l'exemple suivant :

```
Etat du cluster :
-----
Cluster ..... 9.67.131.167
Adresse ..... 9.67.131.167
Nombre de ports cibles ..... 3
Délai de maintien de routage par défaut ..... 0
Délai d'expiration par défaut ..... 30
Limite de pondération par défaut du port ..... 20
Nombre maximal de ports ..... 8
Protocole du port par défaut ..... tcp/udp
Nombre maximal de serveurs par défaut ..... 32
Proportion accordée aux connexions actives ..... 0.5
Proportion accordée aux nouvelles connexions ..... 0.5
Proportion accordée aux connexions spécifiques du port 0
Proportion accordée aux mesures du système ..... 0
Largeur de bande partagée (Ko) ..... 0
Adresse de l'hôte principal ..... 9.67.131.167
```

dscontrol executor — Contrôle de l'exécuteur



report

Affiche un rapport d'analyse sur les statistiques. Exemple : nombre total de paquets, paquets annulés, paquets transmis avec des erreurs, etc.

Remarque : Report ne s'applique qu'au composant Dispatcher.

set

Définit les zones de l'exécuteur.

nfa

Définit l'adresse de non-réacheminement. Tout paquet envoyé à cette adresse n'est pas réacheminé par la machine Dispatcher.

Remarque : NFA ne s'applique qu'au composant Dispatcher.

IP address

Adresse IP (Internet Protocol) sous forme de nom symbolique ou en notation décimale à points.

maxclusters

Nombre maximal de clusters pouvant être configurés. La valeur par défaut du paramètre `maxclusters` est 100.

size

Nombre maximal de clusters pouvant être configurés.

maxports

Valeur par défaut du nombre maximum de ports pour les clusters à créer. Ce paramètre peut être remplacé à l'aide des commandes **cluster set** ou **cluster add**. La valeur par défaut du paramètre `maxports` est 8.

size

Nombre de ports.

fintimeout

Nombre de secondes durant lequel une connexion doit être gardée en mémoire avant que cette dernière ne soit mise à l'état FIN. La valeur par défaut du paramètre `fintimeout` est 60.

fintimeout

Valeur du paramètre `fintimeout`.

Remarque : Fintimeout ne s'applique qu'au composant Dispatcher.

hatimeout

Nombre de secondes nécessaires à l'exécuteur pour arrêter les signaux de présence de disponibilité pour dépassement du délai d'expiration. La valeur par défaut est 2.

Remarque : La valeur hatimeout s'applique au composant Dispatcher.

time

Valeur du paramètre hatimeout.

hasynctimeout

Nombre de secondes nécessaires à l'exécuteur pour arrêter la réplication des enregistrements de connexion entre la machine principale et la machine de secours à cause du dépassement du délai d'expiration. La valeur par défaut est 50.

Le temporisateur permet de garantir que les machines principale et de secours tentent de se synchroniser. Toutefois, en présence d'un trop grand nombre de connexions, et lorsque la machine active continue à gérer une charge de trafic entrant importante, la synchronisation risque de ne pas se terminer avant expiration du temporisateur. Par conséquent, Load Balancer tente sans arrêt la resynchronisation et les deux machines ne sont jamais synchronisées. Dans ce type de situation, donnez une valeur supérieure à la valeur par défaut de hasynctimeout pour donner aux deux machines suffisamment de temps pour échanger des informations sur les connexions existantes. Pour définir ce temporisateur, lancez la commande hasynctimeout après la commande dscontrol executor start, mais avant les commandes de haute disponibilité (dscontrol highavailability).

Remarque : La valeur hasynctimeout s'applique au composant Dispatcher.

time

Valeur du paramètre hasynctimeout.

maxservers

Nombre maximal par défaut de serveurs par port. Ce paramètre peut être remplacé à l'aide de la commande **cluster** ou **port**. La valeur par défaut de maxservers est 32.

mss

Nombre maximal d'octets dans le segment de données de la connexion TCP/UDP. Le nombre total d'octets du segment de données et de l'en-tête doit être inférieur au nombre d'octets de l'unité de transmission maximale (MTU). La valeur par défaut de mss est 1460.

Remarque : La taille de segment maximale ne s'applique qu'à la méthode de transfert nat ou cbr du composant Dispatcher.

size

Nombre de serveurs

staletimeout

Nombre de secondes d'inactivité possible sur une connexion avant que cette dernière soit supprimée. La valeur par défaut pour FTP est de 900 et celle pour Telnet de 32 000 000. La valeur par défaut pour tous les autres ports est 300. Ce paramètre peut être remplacé à l'aide de la commande **cluster** ou **port**. Pour plus d'informations, voir «Utilisation de la valeur du délai d'attente», à la page 268.

staletimeout

Valeur du paramètre *staletimeout*.

stickytime

Valeur du délai de maintien de routage par défaut de tous les futurs clusters. Elle peut être remplacée par la commande **cluster** ou **port**. La valeur par défaut du paramètre *stickytime* est 0.

time

Valeur du paramètre *stickytime* en secondes.

clientgateway

Clientgateway est une adresse IP utilisée pour NAT/NAPT ou Fonction CBR de Dispatcher. Il s'agit de l'adresse du routeur par lequel le trafic de retour est transmis de Load Balancer vers les clients. Clientgateway doit être associé à une valeur non nulle avant l'ajout d'un port à une méthode d'acheminement NAT/NAPT ou Fonction CBR de Dispatcher. Voir «Réacheminement NAT/NAPT de Dispatcher (méthode d'acheminement nat)», à la page 53 et «Fonction CBR de Dispatcher (méthode d'acheminement cbr)», à la page 55 pour plus d'informations.

Remarque : Clientgateway ne s'applique qu'au composant Dispatcher.

address

Adresse de la passerelle client sous forme de nom symbolique ou en notation décimale à points. La valeur par défaut est 0.0.0.0.

weightbound

Valeur de la limite de pondération par défaut de tous les futurs ports. Elle peut être remplacée par la commande **cluster** ou **port**. La valeur par défaut du paramètre *weightbound* est 20.

weight

Valeur du paramètre pondération (*weightbound*).

porttype

Valeur du type de port par défaut pour tous les futurs ports. Ce paramètre peut être remplacé à l'aide des commandes **cluster** ou **port**.

Remarque : Porttype ne s'applique qu'au composant Dispatcher.

type

La valeurs possibles sont **tcp**, **udp** et **both**.

wideportnumber

Port TCP inutilisé pour chaque machine Dispatcher. Le paramètre *wideportnumber* doit être le même pour toutes les machines Dispatcher. La valeur par défaut du paramètre *wideportnumber* est 0, ce qui indique que la prise en charge du réseau étendu n'est pas utilisée.

Remarque : Wideportnumber ne s'applique qu'au composant Dispatcher.

port

Valeur de **wideportnumber**.

sharedbandwidth

Quantité maximale de bande passante (en kilo-octets par seconde) pouvant être partagée au niveau de l'exécuteur. Pour plus d'informations sur la bande passante partagée, voir «Utilisation de règles basées sur la largeur de bande réservée et sur la largeur de bande partagée», à la page 216 et «Règle de largeur de bande partagée», à la page 216.

Remarque : Shared bandwidth ne s'applique qu'au composant Dispatcher.

size

La taille de la **bande passante partagée** correspond à un entier. La valeur par défaut est zéro. Si la valeur est zéro, la bande passante ne peut pas être partagée au niveau de l'exécuteur.

configure

Configure une adresse (par exemple une adresse de cluster, une adresse de retour ou une adresse de signal de présence de haute disponibilité) sur la carte d'interface réseau de la machine Dispatcher. Cette opération est connue comme configuration d'un alias sur la machine Dispatcher.

Remarque : Configure ne s'applique qu'au composant Dispatcher.

adresse_interface

Adresse sous forme de nom symbolique ou d'adresse IP.

Remarque : Chaque adresse d'interface supplémentaire doit être séparée de la précédente par le signe plus (+).

nom_interface masque_reseau

Requis uniquement si l'adresse ne correspond pas à une adresse de sous-réseau existante. Le paramètre *nom_interface* peut être une valeur du type en0, eth1, eri0. Le *masque_reseau* est le masque de 32 bits utilisé pour identifier les bits de l'adresse de sous-réseau dans la partie d'une adresse IP réservée pour l'hôte.

unconfigure

Supprime l'adresse d'alias de la carte d'interface réseau.

Remarque : Unconfigure ne s'applique qu'au composant Dispatcher.

start

Lance l'exécuteur.

status

Affiche l'état actuel des valeurs de l'exécuteur pouvant être définies ainsi que les valeurs par défaut.

stop

Arrête l'exécuteur.

Remarque : Stop s'applique à Dispatcher et à CBR.

Exemples

- Pour afficher les compteurs internes de Dispatcher entrez :

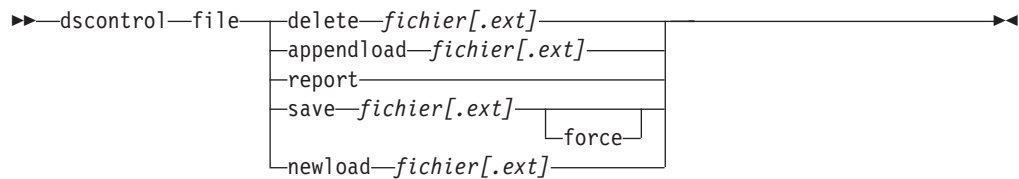
```
dscontrol executor status
```

```
Etat de l'exécuteur :
```

```
-----
Adresse de non-réacheminement ..... 9.67.131.151
Adresse de la passerelle client ..... 0.0.0.0
Délai de maintien de l'état FIN ..... 60
Numéro de port du réseau étendu ..... 0
Largeur de bande partagée (Ko) ..... 0
Nombre maximal de ports par défaut par cluster . 8
Nombre maximal de clusters ..... 100
Nombre maximal de serveurs par défaut par port 32
Délai d'attente par défaut ..... 300
Délai de maintien de routage par défaut ..... 0
Limite de pondération par défaut ..... 20
Type de port par défaut ..... tcp/udp
```

- Pour affecter à l'adresse de non réacheminement la valeur 130.40.52.167, entrez :
`dscontrol executor set nfa 130.40.52.167`
- Pour définir le nombre maximal de clusters, entrez :
`dscontrol executor set maxclusters 4096`
- Pour démarrer l'exécuteur, entrez :
`dscontrol executor start`
- Pour arrêter l'exécuteur, entrez :
`dscontrol executor
stop`

dscontrol file — Gestion des fichiers de configuration



delete

Supprime le fichier.

fichier[.ext]

Un fichier de configuration se compose de commandes dscontrol.

Vous pouvez indiquer n'importe quelle extension de fichier (.ext) ou n'en indiquer aucune.

appendload

Pour mettre à jour la configuration actuelle, la commande appendload lance les commandes exécutables de votre fichier script.

report

Génère un rapport sur les fichiers disponibles.

save

Sauvegarde la configuration en cours de Load Balancer dans le fichier.

Remarque : Les fichiers sont sauvegardés dans les répertoires suivants et chargés à partir de ces mêmes répertoires, où *composant* correspond à dispatcher ou cbr :

- systèmes Linux et UNIX : **/opt/ibm/edge/lb/servers/configurations/composant**
- Plateforme Windows : **C:\Program Files\ibm\edge\lb\servers\configurations\composant**

force

Si vous voulez sauvegarder votre fichier dans un fichier existant du même nom, utilisez **force** pour supprimer le fichier existant avant de sauvegarder le nouveau fichier. Si vous n'utilisez pas l'option force, le fichier existant n'est pas remplacé.

newload

Permet de charger et d'exécuter un nouveau fichier de configuration dans Load Balancer. Le nouveau fichier de configuration remplace la configuration actuelle.

Exemples

- Pour supprimer un fichier, entrez :
`dscontrol file delete fichier3`

Le fichier (fichier3) est supprimé.
- Pour charger un nouveau fichier de configuration afin de remplacer la configuration actuelle, entrez :
`dscontrol file newload fichier1.sv`

Le fichier (fichier1.sv) a été chargé dans Dispatcher.

- Pour charger et ajouter un fichier de configuration à la configuration actuelle, entrez :

```
dscontrol file appendload fichier2.sv
```

Le fichier (fichier2.sv) a été chargé et ajouté à la configuration actuelle.

- Pour visualiser un rapport de vos fichiers (à savoir les fichiers que vous avez sauvegardés précédemment), entrez :

```
dscontrol file report
```

RAPPORT SUR LES FICHIERS :

fichier1.sauv

fichier2.sv

fichier3

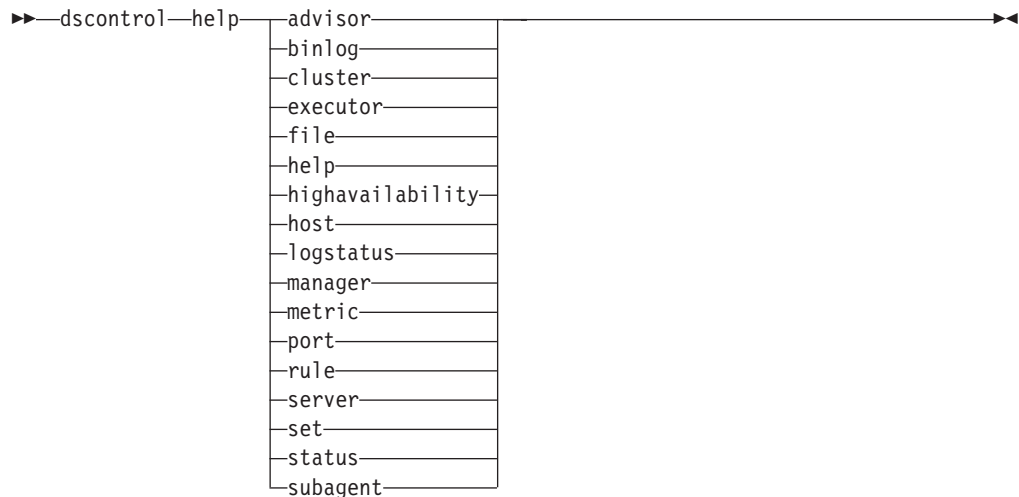
- Pour sauvegarder votre configuration dans un fichier intitulé fichier3, entrez :

```
dscontrol file save
```

```
fichier3
```

La configuration est sauvegardée dans fichier3.

dscontrol help — Affichage ou impression de l'aide relative à cette commande



Exemples

- Pour obtenir de l'aide sur la commande dscontrol, entrez :
dscontrol help

Cette commande génère des résultats similaires à l'exemple suivant :

ARGUMENTS DE LA COMMANDE HELP :

Syntaxe : help <option>

Exemple : help cluster

help	- Affichage des informations d'aide
advisor	- Aide sur la commande advisor
cluster	- Aide sur la commande cluster
port	- Aide sur la commande port
executor	- Aide sur la commande executor
file	- Aide sur la commande file
host	- Aide sur la commande host
binlog	- Aide sur la commande binary
manager	- Aide sur la commande manager
metric	- Aide sur la commande metric
rule	- Aide sur la commande rule
server	- Aide sur la commande server
subagent	- Aide sur la commande subagent
set	- Aide sur la commande set
status	- Aide sur la commande status
logstatus	- Aide sur la commande server log status
highavailability	- Aide sur la commande highavailability

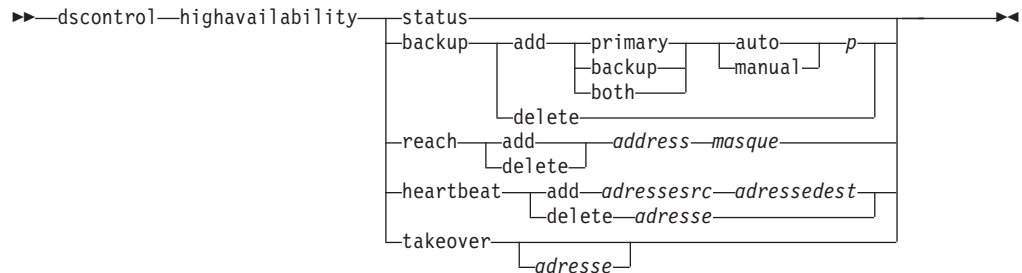
Il est à noter que les paramètres placés entre les signes <> sont des variables.

- L'aide affiche parfois des choix de variables en utilisant un | pour séparer les options disponibles :

```
fintimeout <adresse de cluster>|all <heure>
-Change FIN timeout
(Utilisez "all" pour modifier tous les clusters)
```

dscontrol highavailability — Contrôle de la haute disponibilité

Remarque : Le diagramme de la syntaxe dscontrol high availability ne s'applique qu'au composant Dispatcher.



status

Renvoie un rapport sur la haute disponibilité. Les machines sont identifiées comme étant à l'un des trois états suivants :

Active Le réacheminement de paquets par un poste déterminé (principal, de sauvegarde, ou les deux) est en cours.

Standby

Le réacheminement de paquets par un poste déterminé (principal, de sauvegarde, ou les deux) n'est pas en cours ; celui-ci contrôle l'état d'un Dispatcher **actif**.

Inactive

Le réacheminement de paquets par un poste déterminé est en cours, et celui-ci ne tente pas d'établir une connexion avec son Dispatcher partenaire.

En outre, le mot clé **status** renvoie des informations relatives à divers sous-états :

Synchronized

Un poste déterminé a établi une connexion avec un autre Dispatcher.

Autres sous-états

Ce poste essaie d'établir une connexion avec son Dispatcher partenaire, mais cette tentative n'a pas encore abouti.

backup

Permet de sauvegarder des informations relatives au poste principal (primary) ou de sauvegarde (backup).

add

Permet de définir et d'exécuter les fonctions de haute-disponibilité relatives à ce poste.

primary

Identifie le répartiteur qui sert de poste *principal*.

backup

Identifie la machine Dispatcher qui sert de poste de *sauvegarde*.

both

Identifie la machine Dispatcher qui joue le *double* rôle de poste principal et de sauvegarde. Il s'agit d'une fonction de la haute responsabilité réciproque qui

associe, sur la base des clusters, les rôles de poste principale et de sauvegarde. Pour plus d'informations, voir «Haute disponibilité réciproque», à la page 61.

auto

Spécifie une stratégie de rétablissement *automatique* permettant au poste principal (primary) de reprendre l'acheminement des paquets dès qu'il est remis en service.

manual

Spécifie une stratégie de rétablissement *manuelle* qui ne permet au poste principal (primary) de reprendre le réacheminement des paquets que lorsque l'administrateur a émis une commande **takeover**.

p[ort]

Port TCP non utilisé sur les deux postes, à utiliser par Dispatcher pour ses messages de cadence. Le *port* doit être identique pour le poste principal et pour le poste de sauvegarde.

delete

Supprime ce poste de la liste des postes à haute disponibilité, de sorte qu'il ne puisse plus servir de poste de sauvegarde (backup) ou principal (primary).

reach

Ajoute ou supprime une adresse cible pour les répartiteurs principal et de sauvegarde. Le conseiller d'accessibilité envoie des *pings* à partir des deux répartiteurs pour déterminer le niveau d'accessibilité de leurs cibles.

Remarque : Lorsque vous configurez la cible d'accessibilité, vous devez également démarrer le conseiller d'accessibilité. Le conseiller d'accessibilité démarre automatiquement par la fonction gestionnaire.

add

Ajoute une adresse cible pour le conseiller d'accessibilité.

delete

Supprime une adresse cible du conseiller d'accessibilité.

address

Adresse IP (au format symbolique ou d'adresse IP) du noeud cible.

mask

Masque de sous-réseau.

heartbeat

Définit la session de communication entre les postes Dispatcher principal et de sauvegarde.

add

Indique au Dispatcher source l'adresse de son partenaire (adresse de destination).

srcaddress

Adresse source. Adresse (IP ou symbolique) de ce poste Dispatcher.

dstaddress

Adresse de destination. Adresse (IP ou symbolique) de l'autre poste Dispatcher.

Remarque : *srcaddress* et *dstaddress* doivent correspondre aux NFA des machines pour au moins une paire de signaux de présence.

delete

Supprime la paire d'adresses des informations de cadence (heartbeat). Vous pouvez indiquer l'adresse de destination ou l'adresse source de la paire de signaux de présence.

address

Adresse (IP ou symbolique) de la destination ou de la source.

takeover

Configuration de haute disponibilité simple (Les machines Dispatcher occupent la fonction, soit de poste *principal* soit de *sauvegarde*) :

- Demande à un Dispatcher en attente de passer à l'état actif et de commencer le routage des paquets, ce qui permet de forcer le Dispatcher actif à passer en attente. Cette commande doit être émise sur le poste en attente et ne fonctionne que dans le cas d'une stratégie **manuelle**. Le sous-état doit être *synchronisé*.

Configuration de haute disponibilité réciproque (le rôle de chaque machine Dispatcher est *double*) :

- La machine Dispatcher caractérisée par la fonction haute disponibilité réciproque contient deux clusters qui correspondent à ceux de son partenaire. L'un des deux clusters joue le rôle de cluster principal (cluster de sauvegarde du partenaire), et l'autre de cluster de sauvegarde (cluster principal du partenaire). La commande takeover donne l'ordre à la machine Dispatcher de commencer le routage des paquets en direction du ou des cluster(s) de l'autre machine. Cette commande ne peut être émise que lorsque les clusters de la machine Dispatcher sont en mode *d'attente* et que le sous-état est *synchronisé*. Ceci va contraindre les clusters actifs du partenaire à passer en mode d'attente. La commande ne fonctionne que dans le cadre d'une stratégie **manuelle**. Pour plus d'informations, voir «Haute disponibilité réciproque», à la page 61.

Remarques :

1. Il est à noter que les rôles des deux postes (*principal* et *de sauvegarde*) ne changent pas. Seul leur *état* relatif (*actif* ou *en attente*) est modifié.
2. Trois *scripts* permettent de passer à l'état actif : goActive, goStandby et goInOp. Voir «Utilisation de scripts», à la page 209.

address

La valeur de l'adresse de relais est facultative. Elle ne doit être utilisée que lorsque la machine joue le *double* rôle de poste principal et de sauvegarde (configuration de haute disponibilité réciproque). L'adresse indiquée correspond au NFA de la machine Dispatcher chargée habituellement du trafic de ce cluster. Dans le cas où les deux clusters sont relayés, indiquez la propre adresse NFA du Dispatcher.

Exemples

- Pour vérifier l'état de la fonction de haute disponibilité d'une machine :

```
dscontrol highavailability status
```

Résultat :

Etat de la haute disponibilité :

Rôleprincipal

Stratégie de récupération .. manuelle

Etat Actif

Sous-état..... Synchronisé

Hôte principal..... 9.67.131.151

Port12345
Cible privilégiée..... 9.67.134.223

Etat du signal de présence :

Nombre 1
Source/destination 9.67.131.151/9.67.134.223

Etat de l'accessibilité :

Nombre 1
Adresse 9.67.131.1 accessible

- Pour ajouter les informations de sauvegarde à la machine principale via la stratégie de récupération automatique et le port 80 :
dscontrol highavailability backup add primary auto 80
- Pour ajouter une adresse à laquelle le répartiteur doit pouvoir accéder :
dscontrol
highavailability reach add 9.67.125.18
- Pour ajouter aux machines principale et de secours des informations sur le signal de présence :
Machine principale (primary) -
highavailability heartbeat add 9.67.111.3 9.67.186.8
machine de secours (backup) - highavailability heartbeat
add 9.67.186.8 9.67.111.3
- Pour indiquer au répartiteur en attente de devenir actif et obliger de la sorte la machine active à passer en attente :
dscontrol highavailability takeover

dscontrol host — Configuration d’une machine éloignée

►►—dscontrol—host:—*hôte_éloigné*—◄◄

hôte_éloigné

Nom de la machine Load Balancer en cours de configuration. Lorsque vous tapez cette commande, assurez-vous qu’il ne reste pas d’espace entre **host:** et *hôte_éloigné*, par exemple :

dscontrol host:*hôte_éloigné*

Après avoir tapé cette commande dans l’indicatif DOS, entrez toute commande dscontrol valide que vous désirez envoyer à la machine Load Balancer éloignée.

dscontrol logstatus — Affichage des paramètres du journal du serveur

►►—dscontrol—logstatus—◄◄

logstatus

Affiche les paramètres du journal du serveur (nom, niveau de consignation et taille du journal).

Exemples

Pour afficher l'état du journal, entrez :

```
dscontrol logstatus
```

Cette commande génère des résultats similaires à l'exemple suivant :

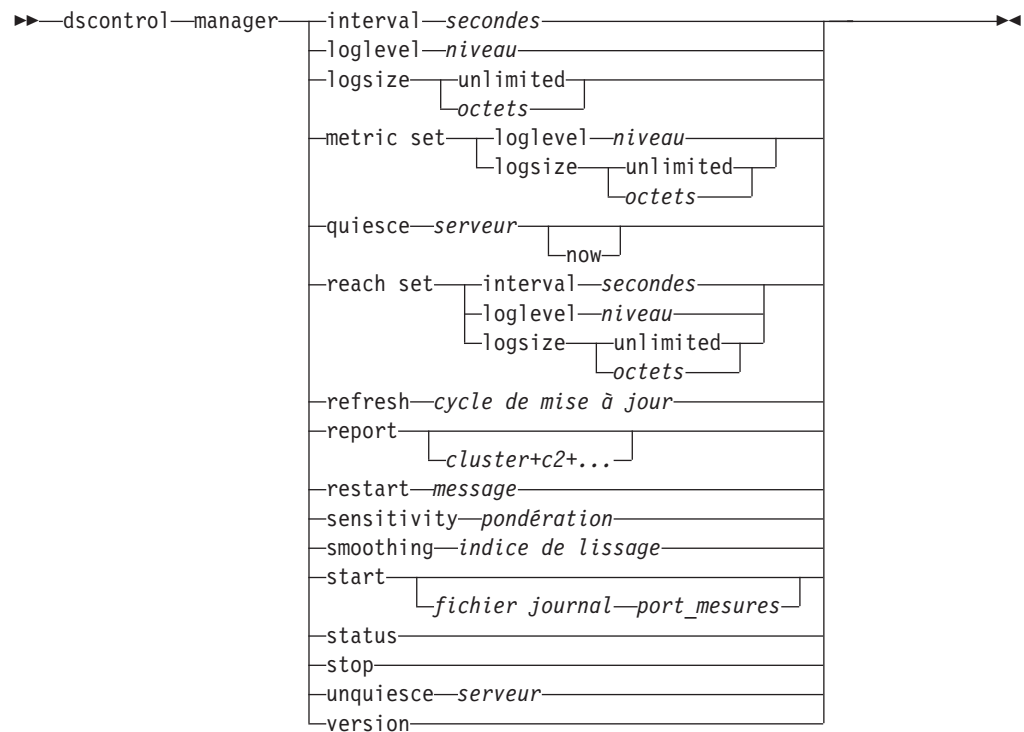
Etat du journal du Dispatcher :

Nom du fichier journal C:\PROGRA~1\IBM\edge\lb\servers\logs\
dispatcher\server.log

Niveau de consignation 1

Taille maxi du journal (octets) 1048576

dscontrol manager — Contrôle du gestionnaire



interval

Définit la fréquence de mise à jour, par le gestionnaire, des pondérations des serveurs pour l'exécuteur, grâce à la mise à jour des critères utilisés par l'exécuteur pour acheminer les requêtes client.

secondes

Nombre positif représentant la fréquence (en secondes) de mise à jour, par le gestionnaire, des pondérations pour l'exécuteur. Valeur par défaut : 2

loglevel

Permet de définir le niveau de consignation relatif au journal du gestionnaire.

niveau

Valeur du niveau (0 à 5). Plus la valeur est élevée, plus la quantité d'informations consignées dans le journal du gestionnaire est importante.

Valeur par défaut : 1. Les valeurs possibles sont les suivantes : 0 (Aucun), 1 (Minimal), 2 (De base), 3 (Modéré), 4 (Avancé), 5 (Prolixe).

logsize

Définit la taille maximale du journal du gestionnaire. Lorsque vous attribuez une taille maximale au fichier journal, ce dernier fonctionne en boucle. Lorsque le fichier atteint la taille indiquée, les entrées suivantes sont écrites à partir du haut du fichier et remplacent les entrées existantes. La taille du journal ne peut pas être moins élevée que la taille actuelle du journal. Les entrées de journal sont horodatées de sorte que vous pouvez identifier l'ordre dans lequel elles sont consignées. Plus le niveau de consignation est élevé, plus la taille du journal doit être choisie attentivement car l'espace peut être saturé rapidement lors d'une consignation à des niveaux plus élevés.

octets

Taille maximale (en octets) du fichier journal du gestionnaire. Vous pouvez

indiquer un nombre positif supérieur à zéro, ou le mot **unlimited**. Il se peut que le fichier journal n'atteigne pas la taille maximale exacte avant l'écrasement, car la taille des entrées de journal elles-mêmes varie. La valeur par défaut est 1 Mo.

metric set

Définit le **niveau de consignation** et la **taille** du journal du contrôleur de mesures. Les niveaux de consignation admis sont les suivants : 0 (Aucun), 1 (Minimal), 2 (De base), 3 (Modéré), 4 (Avancé), 5 (Prolixe). Le niveau par défaut est 1. La taille du journal définit le nombre maximum d'octets pouvant être consignés dans le journal du contrôleur de mesures. Vous pouvez indiquer un nombre positif supérieur à zéro ou la valeur "Unlimited". La taille par défaut est 1 Mo.

mettre au repos

N'indiquez plus de connexions à envoyer à un serveur sauf les nouvelles connexions ultérieures du client vers le serveur mis au repos si la connexion est associée à un délai de maintien de routage et que ce dernier n'est pas arrivé à expiration. Le gestionnaire affecte la valeur 0 à la pondération de ce serveur, pour chaque port pour lequel celui-ci est défini. Utilisez cette commande si vous voulez effectuer une intervention de maintenance rapide sur un serveur puis le réactiver. Si vous supprimez de la configuration un serveur mis au repos, puis que vous l'ajoutez de nouveau, son état ne sera plus celui dans lequel il se trouvait avant d'être mis au repos. Pour plus d'informations, voir «Mise au repos de la gestion des connexions serveur», à la page 223.

serveur

Adresse IP du serveur sous forme de nom symbolique ou en notation décimale à points.

Ou, si vous utilisez le partitionnement du serveur, entrez le nom unique du serveur logique. Pour plus d'informations, voir «Partitionnement du serveur : serveurs logiques configurés pour un serveur physique (adresse IP)», à la page 58.

now

N'utilisez l'option Mettre au repos "maintenant" que si le délai de maintien de routage est défini et que vous voulez que les nouvelles connexions soient envoyées à un autre serveur (différent du serveur mis au repos) avant que le délai de maintien de routage n'expire. Pour plus d'informations, voir «Mise au repos de la gestion des connexions serveur», à la page 223.

reach set

Définit l'intervalle, le niveau de consignation et la taille du journal pour le conseiller d'accessibilité.

refresh

Définit le nombre d'intervalles avant qu'il soit demandé à l'exécuteur de mettre à jour les informations relatives aux nouvelles connexions et aux connexions actives.

cycle de mise à jour

Nombre positif représentant le nombre d'intervalles. Valeur par défaut : 2

report

Affiche un rapport d'analyse sur les statistiques.

cluster

Adresse du cluster que vous souhaitez afficher dans le rapport. L'adresse peut

prendre la forme d'un nom symbolique ou d'une adresse IP. L'affichage par défaut est un rapport de gestionnaire portant sur tous les clusters.

Remarque : Chaque cluster supplémentaire doit être séparé du précédent par le signe plus (+).

restart

Relance tous les serveurs (qui ne sont pas arrêtés) en leur affectant des valeurs de pondération normalisées (la moitié de la pondération maximale).

message

Message à consigner dans le fichier journal du gestionnaire.

sensitivity

Définit la sensibilité minimale à partir de laquelle les pondérations sont mises à jour. Cette valeur définit le moment où le gestionnaire doit modifier sa pondération pour le serveur en fonction des informations externes.

weight

Nombre compris entre 1 et 100 à utiliser comme pourcentage de pondération. La valeur par défaut de 5 crée une sensibilité minimale de 5%.

smoothing

Définit un indice de lissage des variations des pondérations lors de l'équilibrage de charge. Plus l'indice de lissage est élevé, moins les pondérations des serveurs varient lorsque les conditions réseau sont modifiées. Plus cet index est faible, plus les pondérations des serveurs varient.

indice

Nombre positif en virgule flottante. Valeur par défaut : 1,5.

start

Lance le gestionnaire

fichier journal

Nom du fichier dans lequel les données de gestion sont consignées. Chaque enregistrement du journal est horodaté.

Le fichier par défaut se trouve dans le répertoire **logs**. Voir Annexe C, «Exemples de fichiers de configuration», à la page 481. Pour changer le répertoire dans lequel les fichiers journaux sont enregistrés, voir «Modification des chemins des fichiers journaux», à la page 267.

port_mesures

Port utilisé par Metric Server pour signaler les charges du système. Si vous indiquez un port de décompte, vous devez spécifier un nom de fichier journal. Le port de décompte par défaut est 10004.

status

Affiche l'état en cours de toutes les valeurs du gestionnaire qui peuvent être affectées globalement, ainsi que les valeurs par défaut associées.

stop

Arrête le gestionnaire.

unquiesce

Indique que le gestionnaire peut commencer à attribuer une pondération supérieure à 0 à un serveur préalablement mis au repos, sur chaque port pour lequel il est défini.

serveur

Adresse IP du serveur sous forme de nom symbolique ou en notation décimale à points.

version

Affiche la version en cours du gestionnaire.

Exemples

- Pour définir un délai de 5 secondes entre les mises à jour du gestionnaire, entrez :
`dscontrol manager interval 5`
- Pour affecter au niveau de consignation la valeur de 0, afin d'optimiser les performances, entrez :
`dscontrol manager loglevel 0`
- Pour fixer la taille du journal du gestionnaire à 1 000 000 octets, entrez :
`dscontrol manager logsize 1000000`
- Pour indiquer qu'aucune autre connexion ne doit être envoyée au serveur à l'adresse 130.40.52.153, entrez :
`dscontrol manager quiesce 130.40.52.153`
- Pour affecter la valeur 3 au nombre d'intervalles avant la mise à jour des pondérations, entrez :
`dscontrol manager refresh 3`
- Pour obtenir une analyse statistique du gestionnaire, entrez :
`dscontrol manager report`

Cette commande génère des résultats similaires à l'exemple suivant :

SERVEUR	ADRESSE IP	ETAT
mach14.dmz.com	10.6.21.14	ACTIF
mach15.dmz.com	10.6.21.15	ACTIF

LEGENDE ETAT GESTIONNAIRE

ACTV	Connexions actives
NEWC	Nouvelles connexions
SYS	Mesure du système
NOW	Pondération actuelle
NEW	Nouvelle pondération
WT	Pondération
CONN	Connexions

www.dmz.com 10.6.21.100 PORT: 21	POND. ACT NOUV	ACTV 49%	NOUVC 50%	PORT 1%	SYS 0%
mach14.dmz.com	10 10	0	0	-1	0
mach15.dmz.com	10 10	0	0	-1	0

www.dmz.com 10.6.21.100 PORT: 80	POND. ACT NOUV	ACTV 49%	NOUVC 50%	PORT 1%	SYS 0%
mach14.dmz.com	10 10	0	0	23	0
mach15.dmz.com	9 9	0	0	30	0

CONSEILLER	CLUSTER:PORT	DELAI
http	80	unlimited
ftp	21	unlimited

- Pour relancer tous les serveurs en leur affectant des pondérations normalisées et pour consigner un message dans le fichier journal du gestionnaire, entrez :
dscontrol manager restart Relance du gestionnaire pour mise à jour du code

Cette commande génère des résultats similaires à l'exemple suivant :

320-14:04:54

Relance du gestionnaire pour mettre à jour le code

- Pour affecter la valeur 10 à la sensibilité aux modifications de pondération, entrez :
dscontrol manager sensitivity 10
- Pour affecter la valeur 2 à l'indice de lissage, entrez :
dscontrol manager smoothing 2.0
- Pour démarrer le gestionnaire et indiquer le fichier journal nommé ndmgr.log (les chemins ne peuvent pas être définis), entrez :
dscontrol manager start ndmgr.log
- Pour afficher l'état en cours des valeurs associées au gestionnaire, entrez :
dscontrol manager status

Cette commande génère des résultats similaires à l'exemple suivant :

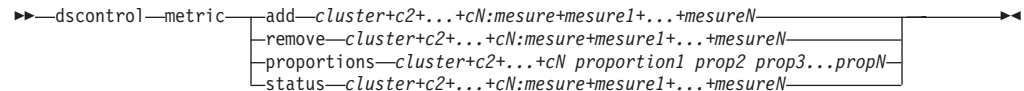
Etat du gestionnaire :

=====

```
Port de mesures..... 10004
Nom du fichier journal du gestionnaire..... manager.log
Niveau de consignation du gestionnaire..... 1
Taille maxi du journal du gestionnaire (octets) ... unlimited
Niveau de sensibilité..... 0.05
Indice de lissage..... 1.5
Intervalle de mise à jour (secondes)..... 2
Cycle de mise à jour des pondérations..... 2
Niveau du journal de contacts..... 1
Taille maximale du journal de contact (octets)... unlimited
Intervalle de mise à jour des contacts (secondes). 7
Nom du fichier journal du contrôleur de mesures... MetricMonitor.log
Niveau de consignation du contrôleur de mesures... 1
Taille maxi du journal du contr mesures (octets).. 1048576
```

- Pour arrêter le gestionnaire, entrez :
dscontrol manager stop
- Pour indiquer qu’aucune autre nouvelle connexion ne doit être envoyée à un serveur à l’adresse 130.40.52.153, entrez (Remarque : Ne mettez le serveur au repos “maintenant” que si le délai de maintien de routage est défini et que vous voulez que les nouvelles connexions soient envoyées à un autre serveur avant expiration du délai de maintien de routage.):
dscontrol manager quiesce 130.40.52.153 now
- Pour indiquer qu’aucune autre nouvelle connexion ne doit être envoyée à un serveur à l’adresse 130.40.52.153, entrez (Remarque : si le délai de maintien de routage est défini, les nouvelles connexions qui auront lieu par la suite à partir du client sont envoyées à ce serveur jusqu’à expiration du délai de maintien de routage.):
dscontrol manager quiesce 130.40.52.153
- Pour indiquer que le gestionnaire peut commencer à attribuer une pondération supérieure à 0 à un serveur à l’adresse 130.40.52.153 qui a préalablement été mis au repos, entrez :
dscontrol manager unquiesce 130.40.52.153
- Pour afficher le numéro de version en cours du gestionnaire, entrez :
dscontrol manager version

dscontrol metric — Configuration des mesures du système



add

Permet d'ajouter la mesure spécifiée.

cluster

Adresse de connexion des clients. Il peut s'agir du nom d'hôte de la machine ou de l'adresse IP. Chaque cluster supplémentaire doit être séparé du précédent par le signe plus (+).

measure

Nom de la mesure du système. Il doit s'agir du nom d'un fichier exécutable ou d'un fichier script du répertoire script du Metric Server.

remove

Supprime la mesure spécifiée.

proportions

Définit les proportions de toutes les mesures associées à cet objet.

status

Affiche les valeurs actuelles de cette mesure.

Exemples

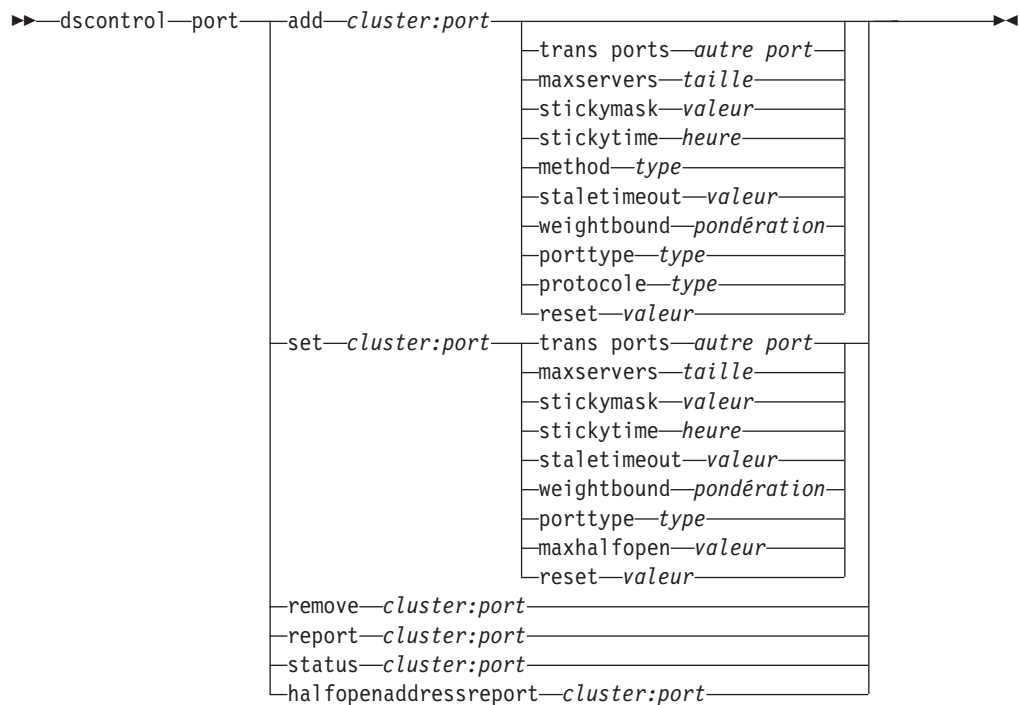
- Pour ajouter une mesure de système, entrez :
dscontrol metric add site1:metric1
- Pour définir les proportions d'un nom de site disposant de deux mesures de système, entrez :
dscontrol
metric proportions site1 0 100
- Pour afficher l'état en cours des valeurs associées à la mesure spécifiée, entrez :
dscontrol metric status site1:metric1

Cette commande génère des résultats similaires à l'exemple suivant :

Etat des mesures :

```
Cluster ..... 10.10.10.20
Nom de la mesure ..... metric1
Proportion de la mesure ..... 50
  Serveur ..... plm3
  Données de mesure ..... -1
```

dscontrol port — Configuration des ports



add

Ajoute un port à un cluster. Avant d'ajouter un serveur à un port, vous devez ajouter un port à un cluster. Si aucun port n'est destiné à un cluster, toutes les requêtes client sont traitées en local. Cette commande permet d'ajouter plusieurs ports à la fois.

cluster

Adresse du cluster sous forme de nom symbolique ou d'adresse IP. Vous pouvez utiliser le signe deux-points (:) comme caractère générique. Par exemple, la commande `dscontrol port add :80` permet d'ajouter le port 80 à tous les clusters.

Remarque : Chaque cluster supplémentaire doit être séparé du précédent par le signe plus (+).

port

Numéro du port. Le numéro de port 0 (zéro) est utilisé pour indiquer un port générique.

Remarque : Chaque port supplémentaire doit être séparé du précédent par le signe plus (+).

trans ports

Trans ports permet d'étendre la fonction d'affinité/maintien de routage à des ports multiples de façon à ce que les requêtes des clients reçue sur différents ports continuent à être envoyées au même serveur lors des requêtes suivantes. En ce qui concerne la valeur trans ports, indiquez le numéro *autre port* avec lequel vous souhaitez partager la fonction d'affinité trans ports. Pour utiliser cette fonction, les ports doivent :

- partager la même adresse de cluster
- partager les mêmes serveurs

- avoir le même (non nul) délai de maintien de routage
- avoir le même masque de maintien de routage

Pour supprimer la fonction trans ports, remplacez la valeur trans ports sur son propre numéro de port. Pour plus d'informations sur la fonction de l'affinité trans ports, voir «Affinité de ports croisés», à la page 222.

Remarque : L'affinité trans ports ne s'applique qu'aux méthodes d'acheminement MAC et NAT/NATP du composant Dispatcher.

autre port

Valeur de trans ports. La valeur par défaut de est la même que celle de son numéro de *port*.

maxservers

Nombre maximal de serveurs. La valeur par défaut du paramètre maxservers est 32. La valeur par défaut de maxservers est 32.

size

Valeur de maxservers.

Masque de maintien de routage

La fonction de masque d'adresse de l'affinité regroupe les requêtes des clients entrants, en fonction de leurs adresses de sous-réseaux communes. Lorsqu'une requête client établit une connexion avec un port pour la première fois, toutes les adresses ultérieures des clients possédant la même adresse de sous-réseau (indiquée par la partie masquée de l'adresse) sont acheminées vers ce même serveur. Pour activer le masque de maintien de routage, le maintien de routage (stickytime) du port doit être défini à zéro. «Masque d'adresse de l'affinité (masque de maintien de routage)», à la page 222.

Remarque : Le mot clé du masque de maintien de routage ne s'applique qu'au composant Dispatcher.

valeur

La valeur du masque de maintien de routage correspond au nombre de bits à poids fort, parmi les adresses IP 32 bits, que vous souhaitez masquer. Les valeurs possibles sont 8, 16, 24 et 32. La valeur par défaut est 32 et elle désactive la fonction du masque d'adresse d'affinité.

stickytime

Délai entre la fermeture d'une connexion et l'ouverture d'une nouvelle connexion au cours de laquelle un client sera renvoyé au même serveur utilisé lors de la première connexion. Passé le délai de maintien de routage, le client peut être envoyé à un serveur autre que le premier.

Composant Dispatcher :

- Méthode d'acheminement CBR de Dispatcher
 - Le délai de maintien de routage ne peut être défini (valeur différente de zéro) que sur un port SSL (et non sur un port HTTP) car sa définition active l'affinité d'ID SSL ID.
 - Lorsque vous définissez le délai de routage du port, le type d'affinité par défaut (none) doit être associé à la commande rule. L'affinité basée sur les règles (cookie passif, URI) ne peut pas être utilisée lorsqu'un délai de maintien de routage est défini pour le port.
- Méthodes d'acheminement MAC et NAT de Dispatcher

- Lorsque vous définissez le délai de routage du port (valeur différente de zéro), vous ne pouvez pas associer de type d'affinité à la règle. L'affinité basée sur les règles ne peut pas être utilisée lorsqu'un délai de maintien de routage est défini pour le port.
- La définition d'une valeur de délai de maintien de routage active l'affinité de l'adresse IP.

Composant CBR : lorsque vous définissez un délai de routage différent de zéro, le type d'affinité par défaut (none) doit être associé à la commande rule. L'affinité basée sur les règles (cookie passif, URI, cookie actif) ne peut pas être utilisée lorsqu'un délai de maintien de routage est défini pour le port.

time

Délai de maintien de routage du port (en secondes). la valeur Zéro signifie que le port n'est pas maintenu.

method

Il s'agit de la méthode d'acheminement. Les méthodes d'acheminement possibles sont les suivantes : MAC, NAT ou CBR (routage par contenu). Il se peut que vous n'ajoutiez *pas* de méthode d'acheminement NAT ou CBR si vous n'indiquez pas d'abord une adresse IP non nulle pour le paramètre clientgateway de la commande dscontrol executor. Voir «Réacheminement NAT/NAPT de Dispatcher (méthode d'acheminement nat)», à la page 53 et «Fonction CBR de Dispatcher (méthode d'acheminement cbr)», à la page 55 pour plus d'informations.

Remarques :

1. La méthode ne s'applique qu'au composant Dispatcher.
2. Si le serveur dorsal se trouve sur le même sous-réseau que l'adresse de retour et que vous utilisez la méthode d'acheminement CBR ou NAT, vous devez définir l'adresse de routeur en tant qu'adresse de serveur dorsal.
3. Si vous ajoutez une méthode d'acheminement MAC, pour devez attribuer au paramètre "protocol" la valeur HTTP ou SSL.

type

Il s'agit du type de la méthode d'acheminement. Les valeurs possibles sont les suivantes : mac, nat ou cbr. La méthode d'acheminement par défaut est MAC.

staletimeout

Nombre de secondes d'inactivité possible sur une connexion avant que cette dernière soit supprimée. Pour le composant Dispatcher, la valeur par défaut est 900 pour le port 21 (FTP) et 32 000 000 pour le port 23 (Telnet). Pour tous les autres ports Dispatcher et CBR, la valeur par défaut est 300. Le paramètre staletimeout peut également être défini au niveau de l'exécuteur ou du cluster. Pour plus d'informations, voir «Utilisation de la valeur du délai d'attente», à la page 268.

valeur

Valeur du paramètre **staletimeout** (en secondes).

weightbound

Définit la pondération maximale des serveurs sur ce port. Cela a une incidence sur le degré de différence possible entre le nombre des demandes transmises par l'exécuteur à chaque serveur. La valeur par défaut est 20.

weight

Nombre compris entre 1 et 100 représentant la limite de pondération maximale.

porttype

Type de port.

Remarque : Le paramètre Porttype ne s'applique qu'à Dispatcher.

type

La valeurs possibles sont **tcp**, **udp** et **both**. La valeur par défaut est (tcp/udp).

Protocole

Type de protocole. Ce paramètre doit être défini pour le composant Dispatcher si vous spécifiez la méthode "cbr" sur le port. Si vous sélectionnez le type de protocole **SSL**, vous devez également définir le délai de maintien de routage (valeur différente de zéro) pour activer l'affinité d'ID SSL. Si vous sélectionnez le type de protocole **HTTP**, vous pouvez définir l'affinité du serveur à l'aide de règles Contenu. Pour plus d'informations, voir «Fonction CBR de Dispatcher (méthode d'acheminement cbr)», à la page 55.

Remarque : Le protocole ne s'applique qu'à la méthode d'acheminement CBR de Dispatcher.

type

La valeurs possibles sont **HTTP** ou **SSL**.

maxhalfopen

Nombre maximal de connexions partielles. Ce paramètre permet de détecter les refus de service éventuelles qui génèrent un nombre élevé de connexions TCP partielles sur les serveurs.

Une valeur positive indique qu'une vérification est effectuée pour déterminer si le nombre de connexions partielles en cours dépasse la limite autorisée. Si tel est le cas, un script d'alerte est appelé. Pour plus d'informations, voir «Détection d'attaque de refus de service», à la page 239.

Remarque : Le paramètre maxhalfopen ne s'applique qu'à Dispatcher.

valeur

Valeur de maxhalfopen. La valeur par défaut est zéro (aucune vérification n'est effectuée).

reset

Reset permet d'indiquer si Load Balancer envoie des réinitialisations TCP aux serveurs arrêtés du port. Une réinitialisation A TCP provoque la fermeture immédiate de la connexion. Pour plus d'informations, voir «Envoie d'une réinitialisation TCP à un serveur arrêté (composant Dispatcher uniquement)», à la page 182.

Remarque : Reset ne s'applique qu'au composant Dispatcher. Le paramètre clientgateway de la commande dscontrol executor doit avoir pour valeur une adresse de routeur pour utiliser le mot clé reset.

valeur

Les valeurs acceptées pour reset sont yes et no. La valeur par défaut est no (pas de réinitialisation TCP des serveurs arrêtés). Lorsque reset a la valeur yes, des réinitialisations TCP sont envoyées aux serveurs arrêtés.

set

Définit les zones d'un port.

remove

Supprime ce port.

report

Génère un rapport sur ce port.

status

Affiche l'état des serveurs sur ce port. Pour visualiser l'état de tous les ports, n'indiquez pas de *port* dans cette commande. N'oubliez pas les deux-points.

nbSecondes

Durée en secondes avant la réinitialisation des connexions partielles.

halfopenaddressreport

Génère des entrées dans le journal (halfOpen.log) pour toutes les adresses client (jusqu'à environ 8000 paires d'adresses) qui ont accédé à des serveurs disposant de connexions partielles. De plus, les données statistiques sont affichées sur la ligne de commande, telles que le nombre total, moyen ou plus élevé de connexions partielles, et le temps moyen de connexion partielle (en secondes). Pour plus d'informations, voir «Détection d'attaque de refus de service», à la page 239.

Exemples

- Pour ajouter les ports 80 et 23 à l'adresse de cluster 130.40.52.153, entrez :
dscontrol port add 130.40.52.153:80+23
- Pour ajouter un port générique à l'adresse de cluster 130.40.52.153, entrez :
dscontrol port set 130.40.52.153:0
- Pour affecter la pondération maximale de 10 au port 80 à l'adresse de cluster 130.40.52.153, entrez :
dscontrol port set 130.40.52.153:80 weightbound 10
- Pour porter à 60 secondes la valeur de délai de maintien de routage des ports 80 et 23 de l'adresse de cluster 130.40.52.153, entrez :
dscontrol port set 130.40.52.153:80+23 stickytime 60
- Pour définir l'affinité trans ports du port 80 au port 23, à l'adresse de cluster 130.40.52.153, entrez :
dscontrol port set 130.40.52.153:80 crossport 23
- Pour supprimer le port 23 de l'adresse de cluster 130.40.52.153, entrez :
dscontrol port remove 130.40.52.153:23
- Pour obtenir l'état du port 80 à l'adresse de cluster 9.67.131.153, entrez :
dscontrol port status 9.67.131.153:80

Cette commande génère des résultats similaires à l'exemple suivant :

Etat du port :

```

Numéro du port ..... 80
Cluster ..... 9.67.131.153
Délai d'expiration ..... 300
Limite de pondération ..... 20
Nombre maximal de serveurs ..... 32
Délai de maintien de routage ..... 0
Type de port ..... tcp/udp
Affinité trans ports ..... 80
Bits du masque de rappel ..... 32
Nombre maximal de connexions partielles 0
Envoyer des réinitialisations TCP ..... no

```

- Pour obtenir le rapport du port 80 à l'adresse de cluster 9.62.130.157, entrez :
dscontrol port report 9.62.130.157:80

Cette commande génère des résultats similaires à l'exemple suivant :

Rapport du port :

Adresse de cluster 9.62.130.157
Numéro du port 80
Nombre de serveurs 5
Pondération maximale des serveurs 10
Nombre total de connexions actives 55
Nombre de connexions par seconde 12
Ko transférés par seconde 298
Nombre de connexions partielles 0
Réinitialisations TCP envoyées 0
Méthode d'acheminement acheminement MAC

- Pour obtenir le rapport d'adresses partielles pour le port 80 à l'adresse de cluster 9.67.127.121, entrez :

dscontrol port halfopenaddressreport 9.67.127.121:80

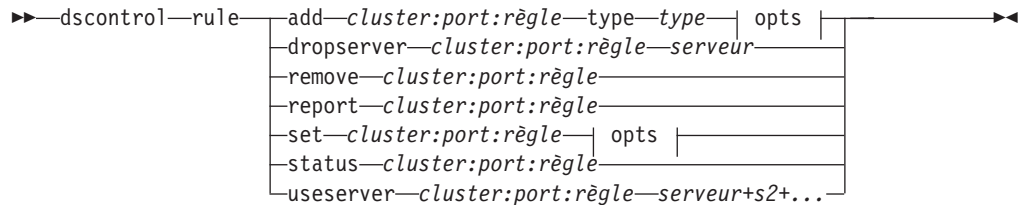
Cette commande génère des résultats similaires à l'exemple suivant :

Rapport

- Connexions partielles créé :

Rapport - Adresses avec connexions partielles pour cluster:port = 9.67.127.121:80
Rapport - Adresses avec connexions partielles pour cluster:port ... 0
Nombre total de connexions partielles consignées 0
Plus grand nombre de connexions partielles consignées 0
Nombre moyen de connexions partielles consignées 0
Temps moyen de connexion partielle (en secondes) consignées 0
Nombre total de connexions partielles reçues 0

dscontrol rule — Configuration des règles



opts :

beginrange	faible	endrange	élevée
priority	niveau		
motif	motif		
tos	valeur		
stickytime	heure		
affinity	type	affinité	
cookie	nom	valeur	
evaluate	niveau		
sharelevel	niveau		

add

Ajoute cette règle à un port.

cluster

Adresse du cluster sous forme de nom symbolique ou d'adresse IP. Vous pouvez utiliser le signe deux-points (:) comme caractère générique. Par exemple, la commande suivante `dscontrol rule add :80:RuleA type type`, permet d'ajouter RuleA au port 80 pour tous les clusters.

Remarque : Chaque cluster supplémentaire doit être séparé du précédent par le signe plus (+).

port

Numéro du port. Vous pouvez utiliser le signe deux-points (:) comme caractère générique. Par exemple, la commande suivante `dscontrol rule add clusterA::RuleA type type` permet d'ajouter RuleA à tous les ports pour ClusterA.

Remarque : Chaque port supplémentaire doit être séparé du précédent par le signe plus (+).

règle

Nom choisi pour la règle. Ce nom peut contenir tout caractère alphanumérique, des traits de soulignement, des traits d'union ou des points. Il peut comporter de 1 à 20 caractères, et ne doit contenir aucun espace.

Remarque : Chaque règle supplémentaire doit être séparée de la précédente par le signe plus (+).

type

Type de règle.

type

Les différents *types* de règles possibles sont les suivants :

ip La règle est définie en fonction de l'adresse IP du client.

time La règle est définie en fonction de l'heure.

connection

La règle est définie en fonction du nombre de connexions par seconde du port. Cette règle ne fonctionne que si le gestionnaire est en cours d'exécution.

active La règle est définie en fonction du nombre total de connexions actives sur le port. Cette règle ne fonctionne que si le gestionnaire est en cours d'exécution.

port La règle est définie en fonction du port client.

Remarque : Port ne s'applique qu'au composant Dispatcher.

service

Cette règle est fondée sur la zone d'octets type de service (TOS) de l'en-tête IP.

Remarque : la règle Service ne s'applique qu'au composant Dispatcher.

reservedbandwidth

Cette règle est fonction du nombre de kilo-octets par seconde de largeur de bande délivrés par un ensemble de serveurs. Pour plus d'informations, voir «Utilisation de règles basées sur la largeur de bande réservée et sur la largeur de bande partagée», à la page 216 et «Règle de largeur de bande réservée», à la page 216.

Remarque : Reservedbandwidth ne s'applique qu'au composant Dispatcher.

sharedbandwidth

Cette règle est fonction du nombre de kilo-octets par seconde de largeur de bande partagés au niveau de l'exécuteur ou du cluster. Pour plus d'informations, voir «Utilisation de règles basées sur la largeur de bande réservée et sur la largeur de bande partagée», à la page 216 et «Règle de largeur de bande partagée», à la page 216.

Remarque : Sharedbandwidth ne s'applique qu'au composant Dispatcher.

true Cette règle est toujours vraie. Considérez-la comme une instruction else en logique de programmation.

content

Cette règle décrit une expression régulière qui sera comparée aux URL demandées par les clients. Elle ne fonctionne que pour Dispatcher et CBR.

beginrange

Valeur de début de la fourchette utilisée pour déterminer si la règle est vraie.

faible

Dépend du type de règle. Le type de valeur et les valeurs par défaut sont précisés ci-après par type de règle :

ip Adresse du client sous forme de nom symbolique ou d'adresse IP. La valeur par défaut est 0.0.0.0.

time Entier. La valeur par défaut est 0 (minuit).

connection

Entier. La valeur par défaut est 0.

active

Entier. La valeur par défaut est 0.

port

Entier. La valeur par défaut est 0.

reservedbandwidth

Entier (kilo-octets par seconde). La valeur par défaut est 0.

endrange

Valeur de fin de la fourchette utilisée pour déterminer si la règle est vraie.

élevée

Dépend du type de règle. Le type de valeur et les valeurs par défaut sont précisés ci-après par type de règle :

ip

Adresse du client sous forme de nom symbolique ou d'adresse IP. La valeur par défaut est 255.255.255.254.

time

Entier. La valeur par défaut est 24 (minuit).

Remarque : Lors de la définition des intervalles de temps (début et fin d'une fourchette horaire), notez que chaque valeur doit être un entier représentant seulement l'heure. Les subdivisions de l'heure ne sont pas indiquées. Pour cette raison, pour indiquer une heure, par exemple entre 3:00 et 4:00 du matin, attribuez la valeur 3 à beginrange (début) et 3 à endrange (fin). Cela signifiera toutes les minutes comprises entre 3:00 et 3:59. Si vous indiquez 3 au paramètre beginrange et 4 au paramètre endrange, vous couvrirez la période de deux heures allant de 3:00 à 4:59.

connections

Entier. La valeur par défaut est 2 à la puissance 32 moins 1.

active

Entier. La valeur par défaut est 2 à la puissance 32 moins 1.

port

Entier. La valeur par défaut est 65535.

reservedbandwidth

Entier (kilo-octets par seconde). La valeur par défaut est 2 à la puissance 32 moins 1.

priority

Ordre dans lequel les règles sont consultées.

niveau

Entier. Si vous ne spécifiez pas la priorité de la première règle que vous ajoutez, Dispatcher lui affecte par défaut la valeur 1. Une règle ajoutée par la suite se verra affecter par défaut une priorité égale à la priorité la plus basse existante + 10. Supposons que vous avez une règle dont la priorité est 30. Vous ajoutez une nouvelle règle et définissez sa priorité à 25 (priorité *supérieure* à 30). Vous ajoutez ensuite une troisième règle, sans lui affecter de priorité. La priorité de la troisième règle sera de 40 (30 + 10).

pattern

Indique le motif à utiliser pour une règle type de contenu.

motif

Motif à utiliser. Pour plus d'informations sur les valeurs acceptées, voir Annexe B, «Syntaxe des règles de contenu (modèle)», à la page 477.

tos

Indique la valeur de "type de service" (TOS) utilisée par la règle type **service**.

Remarque : TOS ne s'applique qu'au composant Dispatcher.

valeur

Chaîne de 8 caractères à utiliser pour la valeur TOS. Les caractères valides sont : 0 (zéro binaire), 1 (un binaire), et x (peu importe). Par exemple : 0xx1010x. Pour plus d'informations, voir «Utilisation de règles basées sur le type de services (TOS)», à la page 214.

stickytime

Indique le délai de maintien de routage à utiliser pour une règle. Lorsque vous attribuez la valeur "activecookie" au paramètre affinity dans la commande rule, vous devez affecter une valeur différente de zéro au délai de maintien de routage pour activer ce type d'affinité. Le délai de maintien de routage de la commande rule ne s'applique pas aux types d'affinité "passivecookie" ou "uri".

Pour plus d'informations, voir «Affinité de cookie actif», à la page 225.

Remarque : La règle de délai de maintien de routage ne s'applique qu'au composant CBR.

time

Heure en secondes

Affinity

Indique le type d'affinité à utiliser pour une règle : cookie actif, cookie passif, URI ou aucune affinité.

Le type d'affinité "activecookie" permet l'équilibrage de charge du trafic Web et une affinité avec le même serveur en fonction des cookies générés par Load Balancer.

Le type d'affinité "passivecookie" permet l'équilibrage de la charge du trafic Web et une affinité avec le même serveur en fonction des cookies d'auto-identification générés par les serveurs. Vous devez utiliser le paramètre cookienam avec l'affinité de cookie passif.

Le type d'affinité "URI" permet l'équilibrage de la charge du trafic Web vers des serveurs Caching Proxy dans le but d'augmenter la mémoire cache.

Pour plus d'informations, voir «Affinité de cookie actif», à la page 225, «Affinité de cookie passif», à la page 226 et «Affinité d'URI», à la page 227.

Remarque : L'affinité s'applique aux règles configurées avec la méthode d'acheminement CBR de Dispatcher et au composant CBR.

type_affinité

Les valeurs possibles pour le type d'affinité sont les suivantes : none (valeur par défaut), activecookie, passivecookie ou uri.

cookienam

Nom arbitraire défini par l'administrateur qui agit comme identificateur pour Load Balancer. Il s'agit du nom que Load Balancer doit rechercher dans la demande d'en-tête HTTP client. Le nom de cookie et la valeur associée sert d'identificateur à Load Balancer, lui permettant d'envoyer les demandes suivantes d'un site Web au même serveur. Le nom de cookie n'est applicable qu'avec l'affinité de cookie passif.

Pour plus d'informations, voir «Affinité de cookie passif», à la page 226.

Remarque : Le nom de cookie s'applique aux règles configurées avec la méthode d'acheminement CBR de Dispatcher et au composant CBR.

valeur

Valeur du nom de cookie.

evaluate

Cette option est disponible pour le composant Dispatcher uniquement. Indique s'il faut évaluer la condition de règle sur tous les serveurs sur un port ou sur tous les serveurs de la règle. C'est option n'est valide que pour les règles qui fondent leurs décisions sur des caractéristiques des serveurs, telles que les règles de type connection, active et reservedbandwidth. Pour plus d'informations, voir «Option d'évaluation de serveur», à la page 220.

Pour la règle de type de connexion, vous pouvez également indiquer une option d'évaluation — upserversonrule. Grâce à cette option, les serveurs restants ne seront pas surchargés si l'un ou plusieurs d'entre eux s'arrêtent.

niveau

Les valeurs acceptées sont port, rule ou upserversonrule. La valeur par défaut est port. upserversonrule ne s'applique qu'à la règle de type de connexion.

sharelevel

Ce paramètre ne s'applique qu'à la règle relative à la largeur de bande partagée. Indique si la largeur de bande est partagé au niveau du cluster ou au niveau de l'exécuteur. Le partage de la largeur de bande au niveau du cluster permet à un ou des ports de partager une quantité maximale de largeur de bande sur plusieurs ports dans le même cluster. Le partage de la largeur de bande au niveau de l'exécuteur permet à un ou des clusters de la configuration Dispatcher de partager une quantité maximale de largeur de bande. Pour plus d'informations, voir «Règle de largeur de bande partagée», à la page 216.

niveau

Les valeurs acceptées sont executor ou cluster.

dropserver

Supprime un serveur d'un jeu de règles.

serveur

Adresse IP de la machine serveur TCP sous forme de nom symbolique ou d'adresse IP.

Ou, si vous utilisez le partitionnement du serveur, entrez le nom unique du serveur logique. Pour plus d'informations, voir «Partitionnement du serveur : serveurs logiques configurés pour un serveur physique (adresse IP)», à la page 58.

Remarque : Chaque serveur supplémentaire doit être séparé du précédent par le signe plus (+).

remove

Supprime une ou plusieurs règles séparées entre elles par des signes plus.

report

Affiche les valeur internes d'une ou plusieurs règles.

set

Définit les valeurs de cette règle.

status

Affiche les valeurs paramétrables d'une ou plusieurs règles.

useserver

Insère des serveurs dans un jeu de règles.

Exemples

- Pour ajouter une règle qui sera toujours vraie, ne spécifiez pas de valeur de début ni de fin. Entrez :

```
dscontrol rule add 9.37.67.100:80:trule type  
true priority 100
```
- Pour créer une règle qui interdit l'accès à une série d'adresses IP, en l'occurrence, celles qui commencent par "9:", entrez :

```
dscontrol rule add 9.37.131.153:80:ni type ip b 9.0.0.0 e 9.255.255.255
```
- Pour créer une règle qui limitera l'utilisation d'un serveur donné de 11:00 à 15:00, entrez :

```
dscontrol rule add cluster1:80:timerule type time beginrange 11 endrange 14  
dscontrol rule useserver cluster1:80:timerule server05
```
- Pour créer une règle fondée sur le contenu de la zone d'octets TOS, dans l'en-tête IP, entrez :

```
dscontrol rule add 9.67.131.153:80:tosrule type service tos 0xx1001x
```
- Pour créer une règle basée sur la largeur de bande réservée qui allouera un ensemble de serveurs (évalués dans la règle) pour délivrer des données jusqu'à 100 kilo-octets par seconde :

```
dscontrol rule add 9.67.131.153:80:rbwrule type  
reservedbandwidth  
beginrange 0 endrange 100 evaluate rule
```
- Pour créer une règle basée sur la largeur de bande partagée qui va rechercher la largeur de bande inutilisée au niveau du cluster, (Remarque : vous devez d'abord spécifier la quantité maximale de largeur de bande (en kilo-octets par seconde) pouvant être partagée au niveau du cluster à l'aide de la commande `dscontrol cluster`), entrez :

```
dscontrol cluster set 9.67.131.153  
sharedbandwidth 200  
  
dscontrol rule add 9.67.131.153:80:shbwrule type sharedbandwidth  
sharelevel cluster
```

dscontrol server — Configuration des serveurs



add

Permet d'ajouter ce serveur.

cluster

Adresse du cluster sous forme de nom symbolique ou d'adresse IP. Vous pouvez utiliser le signe deux-points (:) comme caractère générique. Par exemple, la commande `dscontrol server add :80:ServerA` permet d'ajouter ServerA au port 80 sur tous les clusters.

Remarque : Chaque cluster supplémentaire doit être séparé du précédent par le signe plus (+).

port

Numéro du port. Vous pouvez utiliser le signe deux-points (:) comme caractère générique. Par exemple, la commande `dscontrol server add ::ServerA`, permet d'ajouter ServerA à tous les clusters sur tous les ports.

Remarque : Chaque port supplémentaire doit être séparé du précédent par le signe plus (+).

serveur

Le **serveur** est l'adresse IP unique de la machine serveur TCP sous forme de nom symbolique ou d'adresse IP.

Ou, si vous utilisez un nom unique qui ne se résout pas en adresse IP, vous devez fournir le paramètre **address** du serveur dans la commande **dscontrol server add**. Pour plus d'informations, voir «Partitionnement du serveur : serveurs logiques configurés pour un serveur physique (adresse IP)», à la page 58.

Remarque : Chaque serveur supplémentaire doit être séparé du précédent par le signe plus (+).

address

Adresse IP unique du serveur TCP sous forme de nom d'hôte ou d'adresse IP. S'il n'est pas possible de résoudre le serveur, vous devez fournir l'adresse de la machine serveur physique. Pour plus d'informations, voir «Partitionnement du serveur : serveurs logiques configurés pour un serveur physique (adresse IP)», à la page 58.

address

Valeur de l'adresse du serveur.

collocated

L'option Co-implanté permet d'indiquer si Dispatcher est installé sur l'un des serveurs dont il équilibre les charges.

Remarque : Le paramètre *collocated* n'est valide que si vous utilisez les méthodes d'acheminement MAC, NAT ou CBR de Dispatcher. Site Selector et CBR peuvent être co-implantés sur toutes les plateformes mais ne requièrent pas ce mot clé. Pour plus d'informations, voir «Utilisation de serveurs implantés au même endroit», à la page 202.

valeur

Valeur de Co-implanté : oui ou non. Elle est fixée à non par défaut.

sticky

Permet à un serveur de remplacer sur son port les paramètres de maintien de routage. Lorsque la valeur par défaut est "oui," le serveur garde l'affinité normale définie sur le port. Lorsque cette valeur est "non," le client ne sera *pas* renvoyé à ce serveur lors de sa prochaine requête à ce port et ce, indépendamment des paramètres de maintien de routage du port. Ceci peut s'avérer utile quand vous utilisez des règles. Pour plus d'informations, voir «Substitution d'affinité de port», à la page 219.

valeur

Valeur du maintien de routage : oui ou non. Elle est fixée à non par défaut.

weight

Nombre compris entre 0 et 100 (mais qui ne doit pas dépasser la valeur de pondération du port spécifiée) représentant la pondération relative à ce serveur. L'affectation de la valeur zéro à la pondération empêche l'envoi de nouvelles demandes au serveur mais ne met pas fin aux connexions actuellement actives à ce serveur. La valeur par défaut correspond à la moitié de la valeur de pondération maximale du port indiqué. Si le gestionnaire est en cours d'exécution, ce paramètre est rapidement remplacé.

valeur

Valeur de la pondération du serveur.

fixedweight

L'option *fixedweight* vous permet d'indiquer si vous souhaitez, ou non, que le gestionnaire modifie la pondération du serveur. Si vous fixez sur oui la valeur *fixedweight*, le gestionnaire en activité ne sera pas autorisé à modifier la pondération du serveur. Pour plus d'informations, voir «Pondérations fixées par le gestionnaire», à la page 182.

valeur

Valeur de *fixedweight* : oui ou non. La valeur par défaut est non.

cookievalue

Cookievalue est une valeur arbitraire qui représente le côté serveur de la paire de valeur nom de cookie/cookie. La valeur de cookie, associée au nom de

cookie, sert d'identificateur permettant à Load Balancer d'envoyer les demandes client suivantes au même serveur. Pour plus d'informations, voir «Affinité de cookie passif», à la page 226.

Remarque : Cookievalue est valide pour Dispatcher (avec la méthode d'acheminement CBR) et pour CBR.

valeur

Il s'agit de n'importe quelle valeur arbitraire. Par défaut, il n'y a pas de valeur de cookie.

mapport

Mappe le numéro du port de destination de la demande client (pour Dispatcher) au numéro de port du serveur que Dispatcher utilise pour équilibrer la charge de la demande du client. Permet à Load Balancer de recevoir une demande de client sur un port et de la transmettre à un autre port de la machine serveur. Le paramètre mapport permet d'équilibrer la charge des demandes d'un client sur un serveur sur lequel peuvent s'exécuter plusieurs démons serveur.

Remarque : Mapport s'applique à Dispatcher (avec les méthodes d'acheminement nat ou cbr) ainsi qu'à CBR. Pour Dispatcher, voir «Réacheminement NAT/NAPT de Dispatcher (méthode d'acheminement nat)», à la page 53 et «Fonction CBR de Dispatcher (méthode d'acheminement cbr)», à la page 55. Pour CBR, voir «Equilibrage de charge client-proxy dans SSL et proxy-serveur dans HTTP», à la page 108.

Protocole

Les valeurs valides pour le protocole sont HTTP et HTTPS. La valeur par défaut est HTTP.

Remarque : Le protocole ne s'applique qu'au composant CBR.

valeur du port

Valeur du numéro de port de mappage. La valeur par défaut est le numéro de port de destination de la demande du client.

router

Si vous définissez un réseau étendu, il s'agit de l'adresse du routeur vers le serveur éloigné. La valeur par défaut est 0, correspondant à un serveur local. Notez que, lorsqu'une adresse de routeur est définie avec une valeur autre que zéro (ce qui désigne un serveur éloigné), elle ne peut pas être redéfinie par 0 pour rechanger le serveur en serveur local. Le serveur doit être supprimé, puis ajouté à nouveau sans adresse de routeur spécifiée. De même, un serveur local (adresse de routeur = 0) ne peut pas être changé en serveur éloigné en changeant l'adresse du routeur. Le serveur doit être supprimé, puis ajouté de nouveau. Pour plus d'informations, voir «Configuration du support de réseau étendu pour Dispatcher», à la page 228.

Remarque : Le routeur ne concerne que Dispatcher. Si vous utilisez les méthodes d'acheminement NAT ou CBR, indiquez l'adresse de routeur lors de l'ajout d'un serveur à la configuration.

adr

Adresse du routeur.

returnaddress

Adresse IP ou nom d'hôte unique. Il s'agit d'une adresse configurée sur la

machine Dispatcher que Dispatcher utilise comme adresse source lors de l'équilibrage de charge des demandes du client sur le serveur. Elle permet de garantir que le serveur renverra le paquet à la machine Dispatcher pour traiter le contenu de la demande, au lieu de l'envoyer directement au client. (Dispatcher transmettra ensuite le paquet IP au client.) Vous devez indiquer la valeur d'adresse de retour lors de l'ajout du serveur. L'adresse de retour ne peut pas être modifiée sauf si vous supprimez le serveur et l'ajoutez à nouveau. Elle ne peut pas être identique à l'adresse de cluster, de serveur ou NFA.

Remarque : Le paramètre `returnaddress` ne s'applique qu'à Dispatcher. Si vous utilisez les méthodes d'acheminement NAT ou CBR, indiquez l'adresse de retour lors de l'ajout d'un serveur à la configuration.

adr

Valeur de l'adresse de retour.

advisorrequest

Le conseiller HTTP ou HTTPS utilise la chaîne `advisor request` pour interroger l'état des serveurs. Elle n'est valide que pour les serveurs qui sont traités par le conseiller HTTP ou HTTPS. Vous devez démarrer le conseiller HTTP ou HTTPS pour activer cette valeur. Pour plus d'informations, voir «Configuration du conseiller HTTP ou HTTPS à l'aide de l'option de demande ou de réponse (URL)», à la page 190.

Remarque : La chaîne `advisorrequest` s'applique aux composants Dispatcher et CBR.

chaîne

Valeur de la chaîne utilisée par le conseiller HTTP ou HTTPS. La valeur par défaut est `HEAD / HTTP/1.0`.

Remarque : Si la chaîne comporte un espace —

- Lorsque vous lancez la commande à partir de l'invite du shell **dscontrol>>**, vous devez mettre la chaîne entre guillemets. Par exemple : **server set cluster:port:serveur advisorrequest "head / http/1.0"**
- Lorsque vous lancez la commande **dscontrol** à partir de l'invite du système d'exploitation, vous devez placer les caractères `"\"` et `\\"` respectivement avant et après le texte. Par exemple : **dscontrol server set cluster:port:serveur advisorrequest "\"head / http/1.0\""**

advisorresponse

Chaîne `advisor response` que le conseiller HTTP ou HTTPS recherche dans la réponse HTTP. Elle n'est valide que pour les serveurs qui sont traités par le conseiller HTTP ou HTTPS. Vous devez démarrer le conseiller HTTP ou HTTPS pour activer cette valeur. Pour plus d'informations, voir «Configuration du conseiller HTTP ou HTTPS à l'aide de l'option de demande ou de réponse (URL)», à la page 190.

Remarque : La chaîne `advisorresponse` s'applique aux composants Dispatcher et CBR.

chaîne

Valeur de la chaîne utilisée par le conseiller HTTP ou HTTPS. La valeur par défaut est null.

Remarque : Si la chaîne comporte un espace —

- Lorsque vous lancez la commande à partir de l'invite du shell **dscontrol>>**, vous devez mettre la chaîne entre guillemets.
- Lorsque vous lancez la commande **dscontrol** à partir de l'invite du système d'exploitation, vous devez placer les caractères "\" et \"\"\" respectivement avant et après le texte.

down

Marque ce serveur comme étant arrêté. Cette commande permet d'interrompre toutes les connexions actives à ce serveur et d'empêcher l'envoi d'autres connexions ou paquets à ce serveur.

Lorsqu'un serveur est arrêté à l'aide de la commande **server down**, si le délai de maintien de routage a une valeur différente de zéro pour ce serveur, les clients existants continuent à être servis par ce serveur jusqu'à expiration du délai. Le serveur n'est arrêté qu'après expiration du délai de maintien de routage.

remove

Permet de supprimer ce serveur.

report

Génère un rapport sur ce serveur. Le rapport contient pour chaque serveur les informations suivantes : nombre actuel de connexions par seconde, nombre de Ko transférés en une seconde, nombre total de connexions existantes, nombre de connexions actives, nombre de connexions à l'état FIN et nombre de connexions terminées.

set

Permet de définir des valeurs pour ce serveur.

status

Affiche l'état des serveurs.

up Marque ce serveur comme étant activé. Dispatcher envoi désormais de nouvelles connexions à ce serveur.

Exemples

- Pour ajouter le serveur 27.65.89.42 au port 80 sur un cluster à l'adresse 130.40.52.153, entrez :
`dscontrol server add 130.40.52.153:80:27.65.89.42`
- Pour fixer le serveur, à l'adresse 27.65.89.42, dans la position Maintien de routage (fonction de substitution d'affinité de port), entrez :
`dscontrol server set 130.40.52.153:80:27.65.89.42 sticky no`
- Pour marquer le serveur 27.65.89.42 comme étant arrêté, entrez :
`dscontrol server down 130.40.52.153:80:27.65.89.42`
- Pour supprimer le serveur 27.65.89.42 de tous les ports de tous les clusters, entrez :
`dscontrol server remove ::27.65.89.42`
- Pour fixer le serveur, à l'adresse 27.65.89.42, dans la position co-implanté (le serveur et Load Balancer sont situés sur la même machine), entrez :
`dscontrol server set 130.40.52.153:80:27.65.89.42 colocated yes`
- Pour affecter la valeur 10 à la pondération du serveur 27.65.89.42 au port 80 sur un cluster à l'adresse 130.40.52.153, entrez :
`dscontrol server set 130.40.52.153:80:27.65.89.42 weight 10`

- Pour marquer le serveur 27.65.89.42 comme étant activé, entrez :
dscontrol server up 130.40.52.153:80:27.65.89.42
- Pour ajouter un serveur éloigné :
dscontrol server add 130.40.52.153:80:130.60.70.1 router 130.140.150.0
- Pour permettre au conseiller HTTP d'interroger une demande d'URL HTTP HEAD / HTTP/1.0 pour le serveur 27.65.89.42 sur le port HTTP 80, entrez :
dscontrol server set 130.40.52.153:80:27.65.89.42
advisorrequest "\"HEAD / HTTP/1.0\""
- Pour afficher l'état du serveur 9.67.143.154 sur le port 80, entrez :
dscontrol server status 9.67.131.167:80:9.67.143.154

Cette commande génère des résultats similaires à l'exemple suivant :

Etat du serveur :

```
-----
Serveur ..... 9.67.143.154
Numéro du port ..... 80
Cluster ..... 9.67.131.167
Adresse de cluster ..... 9.67.131.167
Mis au repos ..... N
Serveur en fonction ..... Y
Pondération ..... 10
Pondération fixe ..... N
Rappel pour les règles ..... Y
Serveur éloigné ..... N
Adresse réseau du routeur ..... 0.0.0.0
Co-implanté ..... N
Demande du conseiller ..... HEAD / HTTP/1.0
Réponse du conseiller .....
Valeur du cookie ..... n/a
ID clone ..... n/a
```

dscontrol set — Configuration du journal du serveur



loglevel

Niveau auquel le serveur dsserver consigne ses activités.

niveau

La valeur par défaut de **loglevel** est 0. La fourchette va de 0 à 5. Les valeurs possibles sont les suivantes : 0 (Aucun), 1 (Minimal), 2 (De base), 3 (Modéré), 4 (Avancé), 5 (Prolixe).

logsize

Nombre maximal d’octets à consigner dans le fichier journal.

size

La taille de fichier journal par défaut est 1 Mo.

dscontrol status — Indique par affichage si le gestionnaire et les conseillers sont en cours d'exécution

►►—dscontrol—status—◀◀

Exemples

- Pour visualiser les éléments en cours d'exécution, entrez :
`dscontrol status`

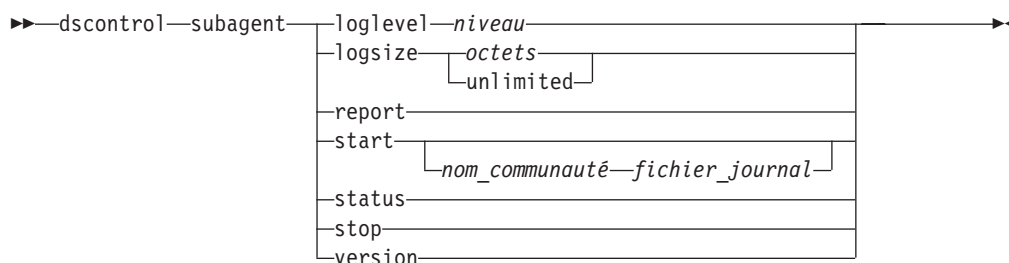
Cette commande génère des résultats similaires à l'exemple suivant :

L'exécuteur (executor) a été lancé.
Le gestionnaire (manager) a été lancé

CONSEILLER	CLUSTER:PORT	DELAI
reach	0	unlimited
http	80	unlimited
ftp	21	unlimited

dscontrol subagent — Configuration du sous-agent SNMP

Remarque : Les diagramme de syntaxe de la commande dscontrol subagent ne s'applique qu'au composant Dispatcher.



loglevel

Niveau auquel le sous-agent consigne ses activités dans un fichier.

niveau

Valeur du niveau (0 à 5). Plus la valeur est élevée, plus la quantité d'informations consignées dans le journal du gestionnaire est importante. Valeur par défaut : 1. Les valeurs possibles sont les suivantes : 0 (Aucun), 1 (Minimal), 2 (De base), 3 (Modéré), 4 (Avancé), 5 (Prolixe).

logsize

Taille maximale en octets à consigner dans le journal du sous-agent. La valeur par défaut est 1 Mo. Lorsque vous attribuez une taille maximale au fichier journal, ce dernier fonctionne en boucle. Lorsque le fichier atteint la taille indiquée, les entrées suivantes sont écrites à partir du haut du fichier et remplacent les entrées existantes. La taille du journal ne peut pas être moins élevée que la taille actuelle du journal. Les entrées de fichier sont horodatées, ce qui permet de déterminer l'ordre dans lequel elles ont été créées. Plus le niveau de consignation est élevé, plus la taille du journal doit être choisie attentivement car l'espace peut être saturé rapidement lors d'une consignation à des niveaux plus élevés.

octets

Taille maximale (en octets) du fichier journal du sous-agent. Vous pouvez indiquer un nombre positif supérieur à zéro, ou le mot **unlimited**. Il se peut que le fichier journal n'atteigne pas la taille maximale exacte avant l'écrasement, car la taille des entrées de journal elles-mêmes varie. La valeur par défaut est unlimited.

report

Affiche un rapport d'analyse sur les statistiques.

start

Lance le sous-agent.

nom_communauté

Nom de la valeur SNMP du nom de communauté que vous pouvez utiliser comme mot de passe de sécurité. La valeur par défaut est public.

Pour la plateforme **Windows**, le nom de communauté du système d'exploitation est utilisé.

fichier journal

Nom du fichier dans lequel les données du sous-agent SNMP sont consignées. Chaque enregistrement du journal est horodaté. La valeur par défaut est

subagent.log. Le fichier par défaut se trouve dans le répertoire **logs**. Voir Annexe C, «Exemples de fichiers de configuration», à la page 481. Pour changer le répertoire dans lequel les fichiers journaux sont enregistrés, voir «Modification des chemins des fichiers journaux», à la page 267.

status

Affiche l'état en cours de toutes les valeurs du sous-agent SNMP qui peuvent être affectées globalement, ainsi que les valeurs par défaut associées.

version

Affiche la version en cours du sous-agent.

Exemples

- Pour démarrer le sous-agent avec le nom de communauté **bigguy**, entrez la commande suivante :
`dscontrol subagent start bigguy bigguy.log`

Chapitre 28. Guide des commandes Site Selector

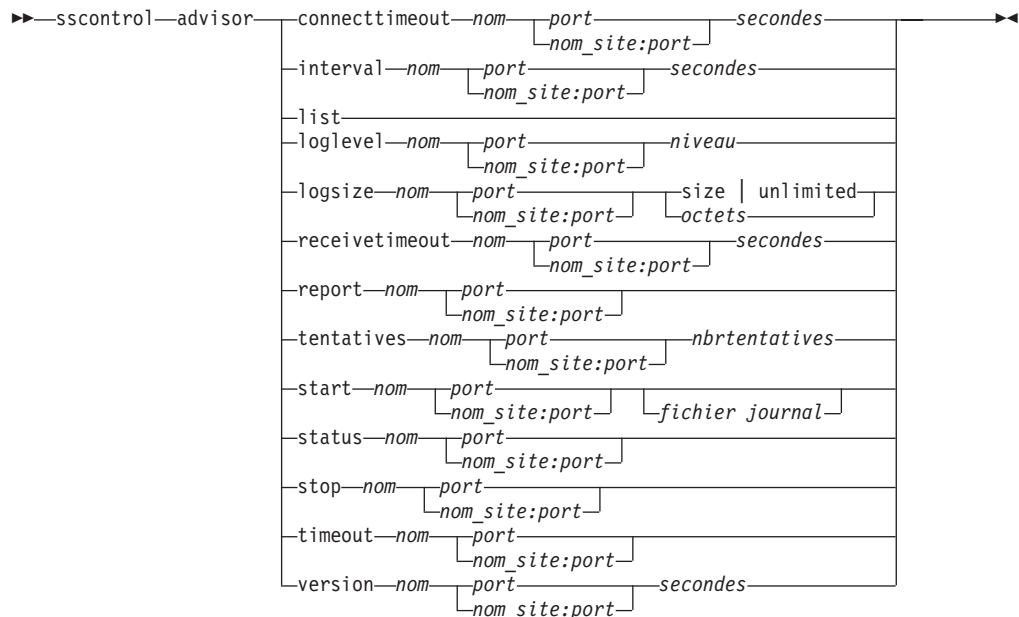
Le présent chapitre explique comment utiliser les commandes **sscontrol** Site Selector suivantes :

- «sscontrol advisor — Contrôle du conseiller», à la page 402
- «sscontrol file — Gestion des fichiers de configuration», à la page 407
- «sscontrol help — Affichage ou impression de l'aide relative à cette commande», à la page 409
- «sscontrol logstatus — Affichage des paramètres du journal du serveur», à la page 410
- «sscontrol manager — Contrôle du gestionnaire», à la page 411
- «sscontrol metric — Configuration des mesures du système», à la page 416
- «sscontrol nameserver — Contrôle de NameServer», à la page 417
- «sscontrol rule — Configuration des règles», à la page 418
- «sscontrol server — Configuration des serveurs», à la page 421
- «sscontrol set — Configuration du journal du serveur», à la page 423
- «sscontrol sitename — Configuration d'un nom de site», à la page 424
- «sscontrol status — Affiche si le gestionnaire et les conseillers sont en cours d'exécution», à la page 427

Vous pouvez entrer une version abrégée des paramètres de commandes **sscontrol**. Il suffit d'entrer les lettres spécifiques des paramètres. Par exemple, pour obtenir l'aide correspondant à la commande **file save**, entrez **sscontrol he f** à la place de **sscontrol help file**.

Remarque : Les valeurs des paramètres de commande doivent être saisies en anglais. Les seules exceptions s'appliquent aux noms d'hôte (utilisés dans les commandes **cluster** et **server**) et aux noms de fichiers (utilisés dans les commandes **file**).

sscontrol advisor — Contrôle du conseiller



connecttimeout

Permet de définir le délai d'attente à l'expiration duquel un conseiller signale qu'une connexion à un serveur a échoué. Pour plus d'informations, voir «Délai de connexion du conseiller et délai de réception pour les serveurs», à la page 188.

nom

Nom du conseiller. Les valeurs possibles sont les suivantes **http**, **https**, **ftp**, **sip**, **ssl**, **smtp**, **imap**, **pop3**, **ldap**, **nntp**, **telnet**, **connect**, **ping**, **WLM** et **WTE**. Les noms des conseillers personnalisés sont au format xxxx, ADV_xxxx étant le nom de la classe mettant en oeuvre le conseiller personnalisé.

port

Numéro du port contrôlé par le conseiller.

secondes

Il s'agit d'un entier positif représentant le délai d'attente à l'expiration duquel le conseiller signale qu'une connexion à un serveur a échoué. La valeur par défaut est trois fois la valeur spécifiée pour l'intervalle du conseiller.

interval

Définit la fréquence à laquelle le conseiller demande des informations aux serveurs.

secondes

Il s'agit d'un entier positif qui représente le nombre de secondes entre les demandes envoyées aux serveurs pour connaître leurs états en cours. Valeur par défaut : 7.

list

Affiche la liste des conseillers qui fournissent des informations au gestionnaire.

loglevel

Définit le niveau de consignation relatif à un journal de conseiller.

niveau

Valeur du niveau (0 à 5). La valeur par défaut est 1. Plus la valeur est élevée, plus la quantité d'informations consignée dans le journal du conseiller est importante. Les valeurs admises sont les suivantes :

- 0 correspond à Aucun
- 1 correspond à Minimal
- 2 correspond à De base
- 3 correspond à Modéré
- 4 correspond à Avancé
- 5 correspond à Prolixe

logsize

Définit la taille maximale d'un journal de conseiller. Lorsque vous affectez une taille maximale au fichier journal, celui-ci fonctionne en boucle, c'est-à-dire que lorsqu'il atteint la taille indiquée, les nouvelles entrées sont écrites à partir du haut du fichier et remplacent les entrées existantes. La valeur indiquée par logsize ne peut pas être inférieure à la taille actuelle du fichier journal. Les entrées du journal sont horodatées, ce qui permet de déterminer l'ordre dans lequel elles ont été enregistrées. Plus le niveau de consignation est élevé, plus la taille du journal doit être choisie avec soin car l'espace peut être saturé rapidement.

taille | unlimited

Taille maximale (en octets) du fichier journal du conseiller. Vous pouvez indiquer un nombre positif supérieur à zéro, ou **unlimited**. Il se peut que le fichier journal n'atteigne pas la taille maximale exacte avant le remplacement des entrées existantes, car la taille des entrées de journal elles-même varie. La valeur par défaut est 1 Mo.

receivetimeout

Permet de définir le délai à l'expiration duquel un conseiller signale que la réception d'un envoi provenant d'un serveur a échoué. Pour plus d'informations, voir «Délai de connexion du conseiller et délai de réception pour les serveurs», à la page 188.

secondes

Il s'agit d'un entier positif qui représente le délai en secondes à l'expiration duquel le conseiller signale que la réception d'un envoi provenant d'un serveur a échoué. La valeur par défaut est trois fois la valeur spécifiée pour l'intervalle du conseiller.

report

Affiche un rapport sur l'état du conseiller.

tentatives

Nombre de tentatives accordées à un conseiller avant de déclarer un serveur arrêté.

nbrtentatives

Entier supérieur ou égal à zéro. Il est préférable que le nombre de tentatives ne dépasse pas 3. Par défaut, le nombre de tentatives est égal à zéro.

start

Lance le conseiller. Il existe des conseillers pour chaque protocole. Les ports par défaut sont les suivants :

Nom du conseiller	Protocole	Port
Connect	Non disp.	Défini par l'utilisateur
db2	privé	50000
ftp	FTP	21
http	HTTP	80
https	SSL	443
imap	IMAP	143
ldap	LDAP	389
nntp	NNTP	119
PING	PING	Non disp.
pop3	POP3	110
sip	SIP	5060
smtp	SMTP	25
ssl	SSL	443
telnet	Telnet	23

nom

Indique le nom du conseiller.

nom_site:port

Le nom de site est facultatif dans les commandes du conseiller, mais la valeur de port est requise. Par défaut, le conseiller démarre sur tous les sites disponibles configurés. Si vous indiquez un nom de site, le conseiller démarre uniquement sur le site précisé. Les noms de site supplémentaires sont séparés par le signe plus (+).

fichier journal

Nom du fichier dans lequel les données de gestion sont consignées. Chaque enregistrement du journal est horodaté.

Le fichier par défaut est *nom_conseiller_port.log*, par exemple, **http_80.log**. Pour changer le répertoire dans lequel les fichiers journaux seront enregistrés, voir «Modification des chemins des fichiers journaux», à la page 267.

Vous ne pouvez lancer qu'un conseiller par nom de site.

status

Affiche l'état et les valeurs par défaut en cours pour toutes les valeurs globales d'un conseiller.

stop

Arrête le conseiller.

timeout

Définit le nombre de secondes pendant lesquelles le gestionnaire considère que les informations provenant du conseiller sont valides. Si le gestionnaire considère que les informations du conseiller sont antérieures à ce délai, il n'utilise pas ces informations pour déterminer les pondérations relatives aux serveurs sur le port contrôlé par le conseiller. Il est fait exception à ce délai lorsque le conseiller a informé le gestionnaire qu'un serveur spécifique est hors service. Le gestionnaire utilise ces informations relatives au serveur même après le dépassement du délai imparti pour les informations du conseiller.

secondes

Nombre positif représentant le nombre de secondes, ou **unlimited**. La valeur par défaut est unlimited.

version

Affiche la version en cours du conseiller.

Exemples

- Pour définir le délai d'attente (30 secondes) à l'expiration duquel un conseiller HTTP (pour le port 80) signale qu'une connexion à un serveur a échoué, entrez :

```
sscontrol advisor  
connecttimeout http 80 30
```

- Pour définir un intervalle de 6 secondes pour le conseiller FTP (pour le port 21), entrez :

```
sscontrol advisor interval ftp 21 6
```

- Pour afficher la liste des conseillers qui fournissent des informations au gestionnaire, entrez :

```
sscontrol advisor list
```

Cette commande génère des résultats similaires à l'exemple suivant :

ADVISOR	SITENAME:PORT	TIMEOUT
http	80	unlimited
ftp	21	unlimited

- Pour attribuer la valeur 0 au niveau de consignation du journal du conseiller http pour le site monSite afin d'optimiser les performances, entrez :
- Pour attribuer la valeur 5000 octets à la taille du journal du conseiller ftp pour le site monSite, entrez :

```
sscontrol advisor loglevel http monSite:80 0  
sscontrol advisor logsize ftp monSite:21 5000
```

- Pour définir le délai d'attente (60 secondes) à l'expiration duquel un conseiller HTTP (pour le port 80) signale que la réception d'un envoi provenant d'un serveur a échoué, entrez :

```
sscontrol advisor receivetimeout http 80 60
```

- Pour afficher un rapport sur l'état du conseiller ftp (pour le port 21), entrez :

```
sscontrol advisor report ftp 21
```

Cette commande génère des résultats similaires à l'exemple suivant :

Rapport du conseiller :

```
-----  
Nom du conseiller ..... http  
Numéro du port ..... 80  
  
Nom du site ..... monSite  
Adresse du serveur ..... 9.67.129.230  
Charge ..... 8
```

- Pour démarrer le conseiller à l'aide du fichier ftpadv.log, entrez :

```
sscontrol advisor start ftp 21 ftpadv.log
```

- Pour afficher l'état actuel des valeurs associées au conseiller http, entrez :

```
sscontrol advisor status http 80
```

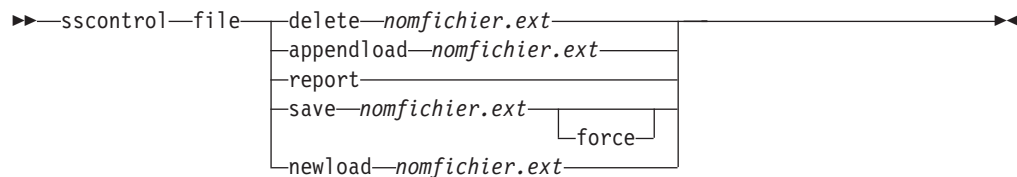
Cette commande génère des résultats similaires à l'exemple suivant :

Etat du conseiller :

Intervalle (secondes)..... 7
Délai d'expiration (secondes) Unlimited
Délai d'expiration de connexion (secondes)... 21
Délai d'expiration de réception (secondes)... 21
Nom du fichier journal du conseiller Http_80.log
Niveau de consignation 1
Taille maximale du journal (octets) Unlimited
Nombre de tentatives 0

- Pour arrêter le conseiller http sur le port 80, entrez :
sscontrol advisor stop http 80
- Pour définir un délai d'attente de 5 secondes pour les informations du conseiller, entrez :
sscontrol advisor timeout ftp 21 5
- Pour rechercher le numéro de version en cours du conseiller ssl, entrez :
sscontrol advisor version ssl 443

sscontrol file — Gestion des fichiers de configuration



delete

Supprime le fichier.

nomfichier.ext

Fichier de configuration.

Vous pouvez indiquer n'importe quelle extension de fichier (.ext) ou n'en indiquer aucune.

appendload

Ajoute un fichier de configuration à la configuration en cours et le charge dans Site Selector.

report

Génère un rapport sur les fichiers disponibles.

save

Sauvegarde la configuration en cours de Site Selector dans le fichier.

Remarque : Les fichiers sont sauvegardés dans les répertoires suivants et chargés à partir de ces mêmes répertoires :

- systèmes Linux et UNIX : **/opt/ibm/edge/lb/servers/configurations/ss**
- Systèmes Windows : **C:\Program Files\ibm\edge\lb\servers\configurations\composant**

force

Si vous voulez sauvegarder votre fichier dans un fichier existant du même nom, utilisez **force** pour supprimer le fichier existant avant de sauvegarder le nouveau fichier. Si vous n'utilisez pas l'option force, le fichier existant n'est pas remplacé.

newload

Charge un nouveau fichier de configuration dans Site Selector. Le nouveau fichier de configuration remplacera la configuration actuelle.

Exemples

- Pour supprimer un fichier, entrez :
`sscontrol file delete fichier3`

Le fichier (fichier3) est supprimé.
- Pour charger un nouveau fichier de configuration afin de remplacer la configuration actuelle, entrez :
`sscontrol file newload fichier1.sv`

Le fichier (fichier1.sv) a été chargé dans Dispatcher.
- Pour charger et ajouter un fichier de configuration à la configuration actuelle, entrez :

```
sscontrol file appendload fichier2.sv
```

Le fichier (fichier2.sv) a été chargé et ajouté à la configuration actuelle.

- Pour visualiser un rapport de vos fichiers (à savoir les fichiers que vous avez sauvegardés précédemment), entrez :

```
sscontrol file report
```

RAPPORT SUR LES FICHIERS :

fichier1.sauv

fichier2.sv

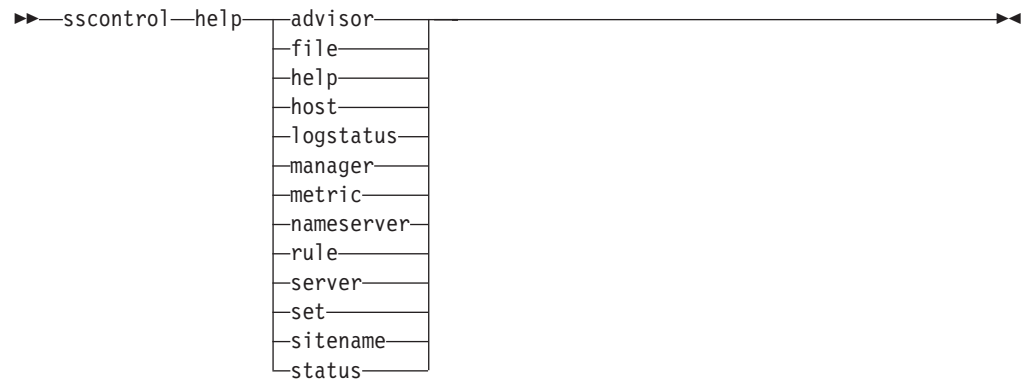
fichier3

- Pour sauvegarder votre configuration dans un fichier intitulé fichier3, entrez :

```
sscontrol file save fichier3
```

La configuration est sauvegardée dans fichier3.

sscontrol help — Affichage ou impression de l'aide relative à cette commande



Exemples

- Pour obtenir l'aide sur la commande sscontrol, entrez :
sscontrol help

Cette commande génère des résultats similaires à l'exemple suivant :

ARGUMENTS DE LA COMMANDE HELP :

Syntaxe : help <option>

Exemple: help name

help	- Affichage des informations d'aide
advisor	- Aide sur la commande advisor
file	- Aide sur la commande file
host	- Aide sur la commande host
manager	- Aide sur la commande manager
metric	- Aide sur la commande metric
sitename	- Aide sur la commande sitename
nameserver	- Aide sur la commande nameserver
server	- Aide sur la commande server
subagent	- Aide sur la commande subagent
set	- Aide sur la commande set
status	- Aide sur la commande status
logstatus	- Aide sur la commande logstatus

Les paramètres entre caractères < > sont des variables.

- L'aide affiche parfois des choix de variables en utilisant le caractère | pour séparer les options disponibles :
logsize <nombre d'octets | unlimited>
- Nombre maximal d'octets à consigner dans le journal

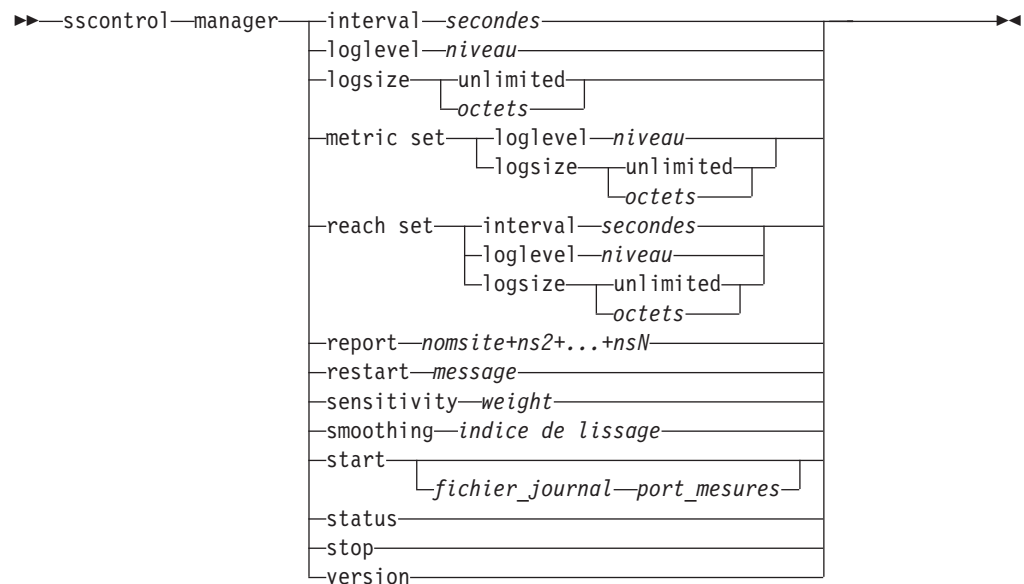
sscontrol logstatus — Affichage des paramètres du journal du serveur

►►—sscontrol—logstatus—◄◄

logstatus

Affiche les paramètres du journal du serveur (nom, niveau de consignation et taille du journal).

sscontrol manager — Contrôle du gestionnaire



interval

Définit la fréquence de mise à jour, par le gestionnaire, des pondérations des serveurs.

secondes

Nombre entier positif représentant la fréquence (en secondes) de mise à jour des pondérations par le gestionnaire. Valeur par défaut : 2

loglevel

Permet de définir le niveau de consignation relatif au journal du gestionnaire.

niveau

Valeur du niveau (0 à 5). Plus la valeur est élevée, plus la quantité d'informations consignées dans le journal du gestionnaire est importante. Valeur par défaut : 1. Les valeurs admises sont les suivantes :

- 0 correspond à Aucun
- 1 correspond à Minimal
- 2 correspond à De base
- 3 correspond à Modéré
- 4 correspond à Avancé
- 5 correspond à Prolixe

logsize

Définit la taille maximale du journal du gestionnaire. Lorsque vous attribuez une taille maximale au fichier journal, ce dernier fonctionne en boucle. Lorsque le fichier atteint la taille indiquée, les entrées suivantes sont écrites à partir du haut du fichier et remplacent les entrées existantes. La valeur indiquée par logsize ne peut pas être inférieure à la taille actuelle du fichier journal. Les entrées du journal sont horodatées, ce qui permet de déterminer l'ordre dans lequel elles ont été enregistrées. Plus le niveau de consignation est élevé, plus la taille du journal doit être choisie avec soin car l'espace peut être saturé rapidement.

octets

Taille maximale (en octets) du fichier journal du gestionnaire. Vous pouvez indiquer un nombre positif supérieur à zéro, ou **unlimited**. Il se peut que le fichier journal n'atteigne pas la taille maximale exacte avant le remplacement des entrées existantes, car la taille des entrées de journal elles-même varie. La valeur par défaut est 1 Mo.

metric set

Définit le **niveau de consignation** et la **taille** du journal du contrôleur de mesures. Les niveaux de consignation admis sont les suivants : 0 (Aucun), 1 (Minimal), 2 (De base), 3 (Modéré), 4 (Avancé), 5 (Prolixe). Le niveau par défaut est 1. La taille du journal définit le nombre maximum d'octets pouvant être consignés dans le journal du contrôleur de mesures. Vous pouvez indiquer un nombre positif supérieur à zéro ou la valeur "Unlimited". La valeur par défaut est 1.

reach set

Définit l'intervalle, le niveau de consignation et la taille du journal pour le conseiller d'accessibilité.

report

Affiche un rapport d'analyse sur les statistiques.

nom_site

Nom du site à afficher dans le rapport. Il s'agit d'un nom d'hôte ne pouvant être résolu qui sera demandé par le client. Le nom de site doit être un nom de domaine qualifié complet.

Remarque : Les noms de site supplémentaires sont séparés par le signe plus (+).

restart

Relance tous les serveurs (qui ne sont pas arrêtés) en leur affectant des valeurs de pondération normalisées (la moitié de la pondération maximale).

message

Message à consigner dans le fichier journal du gestionnaire.

sensitivity

Définit la sensibilité minimale à partir de laquelle les pondérations sont mises à jour. Cette valeur définit le moment où le gestionnaire doit modifier sa pondération pour le serveur en fonction des informations externes.

pondération

Nombre compris entre 0 et 100 indiquant le pourcentage de pondération. La valeur par défaut 5 crée une sensibilité minimale de 5%.

smoothing

Définit un indice de lissage des variations des pondérations lors de l'équilibrage de charge. Plus l'indice de lissage est élevé, moins les pondérations des serveurs varient lorsque les conditions réseau sont modifiées. Plus cet indice est faible, plus les pondérations des serveurs varient.

indice

Nombre positif en virgule flottante. Valeur par défaut : 1,5.

start

Lance le gestionnaire

fichier journal

Nom du fichier dans lequel les données de gestion sont consignées. Chaque enregistrement du journal est horodaté.

Le fichier par défaut se trouve dans le répertoire **logs**. Voir Annexe C, «Exemples de fichiers de configuration», à la page 481. Pour changer le répertoire dans lequel les fichiers journaux sont enregistrés, voir «Modification des chemins des fichiers journaux», à la page 267.

port_mesures

Port sur lequel Metric Server renvoie l'état des charges du système. Si vous indiquez un port de décompte, vous devez spécifier un nom de fichier journal. Le port de décompte par défaut est 10004.

status

Affiche l'état et les valeurs par défaut en cours pour toutes les valeurs globales du gestionnaire.

stop

Arrête le gestionnaire.

version

Affiche la version en cours du gestionnaire.

Exemples

- Pour définir un délai de 5 secondes entre les mises à jour du gestionnaire, entrez :
`sscontrol manager interval 5`
- Pour affecter au niveau de consignation la valeur de 0, afin d'optimiser les performances, entrez :
`sscontrol manager loglevel 0`
- Pour définir la taille du journal du gestionnaire à 1 000 000 octets, entrez :
`sscontrol manager logsize 1000000`
- Pour obtenir une analyse statistique du gestionnaire, entrez :
`sscontrol manager report`

Cette commande génère des résultats similaires à l'exemple suivant :

SERVEUR	ETAT
9.67.129.221	ACTIF
9.67.129.213	ACTIF
9.67.134.223	ACTIF

LEGENDE ETAT GESTIONNAIRE

CPU	Charge de la CPU
MEM	Charge de la mémoire
SYS	Mesure du système
NOW	Pondération actuelle
NEW	Nouvelle pondération
WT	Pondération

monSite	WEIGHT	CPU 49%	MEM 50%	PORT 1%	SYS 0%
	NOW NEW	WT LOAD	WT LOAD	WT LOAD	WT LOAD
9.37.56.180	10 10	-99 -1	-99 -1	-99 -1	0 0
TOTAUX :	10 10	-1	-1	-1	0

ADVISOR	SITENAME:PORT	TIMEOUT
http	80	unlimited

- Pour relancer tous les serveurs en leur affectant des pondérations normalisées et pour consigner un message dans le fichier journal du gestionnaire, entrez :
sscontrol manager restart Relance du gestionnaire pour mettre à jour

Cette commande génère des résultats similaires à l'exemple suivant :

320-14:04:54

Relance du gestionnaire pour mettre à jour le code

- Pour affecter la valeur 10 à la sensibilité aux modifications de pondération, entrez :
sscontrol manager sensitivity 10
- Pour affecter la valeur 2 à l'indice de lissage, entrez :
sscontrol manager smoothing 2.0
- Pour démarrer le gestionnaire et indiquer le fichier journal nommé ndmgr.log (les chemins ne peuvent pas être définis), entrez :
sscontrol manager start ndmgr.log
- Pour afficher l'état en cours des valeurs associées au gestionnaire, entrez :
sscontrol manager status

Cette commande génère des résultats similaires à l'exemple suivant :

Etat du gestionnaire :

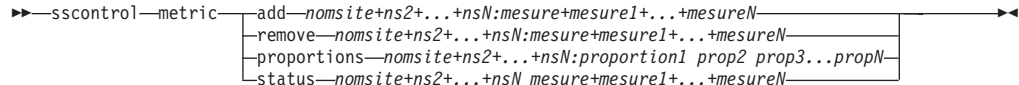
=====

```
Port de mesures..... 10004
Nom du fichier journal du gestionnaire ..... manager.log
Niveau de consignation du gestionnaire ..... 1
Taille maxi du journal du gestionnaire (octets) .. unlimited
Niveau de sensibilité..... 5
Indice de lissage ..... 1.5
Intervalle de mise à jour (secondes) ..... 2
```

Cycle de mise à jour des pondérations 2
Niveau du journal de contacts 1
Taille maximale du journal des contacts (octets) . unlimited
Intervalle mise à jour des contacts (secondes) ... 7

- Pour arrêter le gestionnaire, entrez :
sscontrol manager stop
- Pour afficher le numéro de version en cours du gestionnaire, entrez :
sscontrol manager version

sscontrol metric — Configuration des mesures du système



add

Permet d'ajouter la mesure spécifiée.

nom_site

Nom du site configuré. Les noms de site supplémentaires sont séparés par le signe plus (+).

measure

Nom de la mesure du système. Il doit s'agir du nom d'un fichier exécutable ou d'un fichier script du répertoire script du Metric Server.

remove

Supprime la mesure spécifiée.

proportions

Ce paramètre détermine la correspondance des mesures entre elles lorsqu'elles sont regroupées en une seule charge système pour un serveur.

status

Affiche les valeurs serveur actuelles de cette mesure.

Exemples

- Pour ajouter une mesure de système, entrez :
`sscontrol metric add site1:metric1`
- Pour définir les proportions d'un nom de site disposant de deux mesures de système, entrez :
`sscontrol metric proportions site1 0 100`
- Pour afficher l'état en cours des valeurs associées à la mesure spécifiée, entrez :
`sscontrol metric status site1:metric1`

Cette commande génère des résultats similaires à l'exemple suivant :

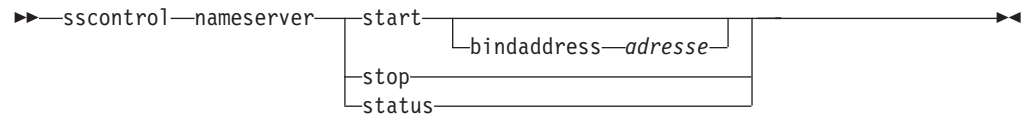
Etat des mesures :

```

Nom du site..... site1
Nom de la mesure..... metric1
1Metric proportion ..... 50
  Serveur ..... 9.37.56.100
  Données de mesure... -1

```

sscontrol nameserver — Contrôle de NameServer



start

Démarre le serveur de noms

bindaddress

Démarre le serveur de noms lié à l'adresse indiquée. Le serveur de noms ne répond qu'aux demandes destinées à cette adresse.

adresse

Adresse (IP ou symbolique) configurée sur le système Site Selector.

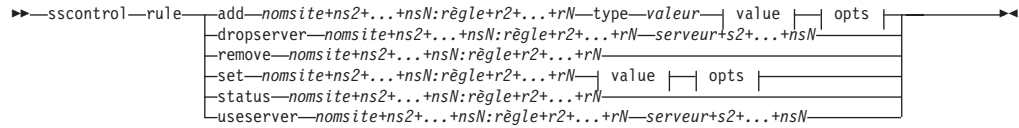
stop

Arrête le serveur de noms

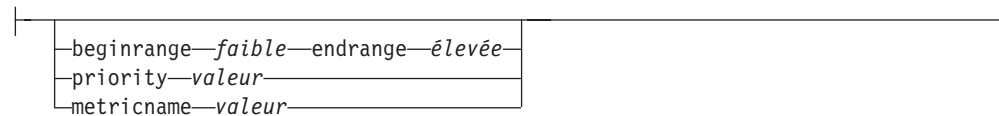
status

Affiche l'état du serveur de noms

sscontrol rule — Configuration des règles



opts :



add

Ajoute cette règle à un nom de site.

nom_site

Nom d'hôte ne pouvant être résolu qui sera demandé par le client. Le nom de site doit être un nom de domaine qualifié complet. Les noms de site supplémentaires sont séparés par le signe plus (+).

règle

Nom choisi pour la règle. Ce nom peut contenir tout caractère alphanumérique, des traits de soulignement, des traits d'union ou des points. Il peut comporter de 1 à 20 caractères, et ne doit contenir aucun espace.

Remarque : Chaque règle supplémentaire doit être séparée de la précédente par le signe plus (+).

type

Type de règle

type

Les différents *types* de règles possibles sont les suivants :

ip La règle est définie en fonction de l'adresse IP du client.

metricall

La règle est basée sur la valeur de mesure actuelle pour tous les serveurs de l'ensemble de serveurs.

metricavg

La règle est basée sur la moyenne des valeurs de mesure actuelles pour tous les serveurs de l'ensemble de serveurs.

time La règle est définie en fonction de l'heure.

true Cette règle est toujours vraie. Considérez-la comme une instruction else en logique de programmation.

beginrange

Valeur de début de la fourchette utilisée pour déterminer si la règle est vraie.

faible

Dépend du type de règle. Le type de valeur et les valeurs par défaut sont précisés ci-après par type de règle :

ip Adresse du client sous forme de nom symbolique ou d'adresse IP. La valeur par défaut est 0.0.0.0.

time Entier. La valeur par défaut est 0 (minuit).

metricall

Entier. Valeur par défaut : 100.

metricavg

Entier. Valeur par défaut : 100.

endrange

Valeur de fin de la fourchette utilisée pour déterminer si la règle est vraie.

élevée

Dépend du type de règle. Le type de valeur et les valeurs par défaut sont précisés ci-après par type de règle :

ip Adresse du client sous forme de nom symbolique ou d'adresse IP. La valeur par défaut est 255.255.255.254.

time Entier. La valeur par défaut est 24 (minuit).

Remarque : Lors de la définition des intervalles de temps (début et fin d'une fourchette horaire), notez que chaque valeur doit être un entier représentant seulement l'heure. Les subdivisions de l'heure ne sont pas indiquées. Pour cette raison, pour indiquer une heure, par exemple entre 3 h et 4 h du matin, attribuez la valeur 3 à beginrange (début) et 3 à endrange (fin). Cela signifiera toutes les minutes comprises entre 3 h et 3 h 59. Si vous indiquez la valeur 3 pour le paramètre beginrange et la valeur 4 pour le paramètre endrange, vous couvrirez la période de deux heures allant de 3 h à 4 h 59.

metricall

Entier. La valeur par défaut est 2 à la puissance 32 moins 1.

metricavg

Entier. La valeur par défaut est 2 à la puissance 32 moins 1.

priority

Ordre dans lequel les règles sont consultées.

niveau

Entier. Si vous ne spécifiez pas la priorité de la première règle que vous ajoutez, Site Selector lui affectera par défaut la valeur 1. Une règle ajoutée par la suite se verra affecter par défaut une priorité égale à la priorité la plus basse existante + 10. Supposons que vous ayez une règle dont la priorité est 30. Vous ajoutez une nouvelle règle et définissez sa priorité à 25 (priorité *supérieure* à 30). Vous ajoutez ensuite une troisième règle, sans lui affecter de priorité. La priorité de la troisième règle sera de 40 (30 + 10).

metricname

Nom de la mesure calculée pour une règle

dropserver

Supprime un serveur d'un jeu de règles.

serveur

Adresse IP de la machine serveur TCP sous forme de nom symbolique ou d'adresse IP.

Remarque : Les noms de site supplémentaires sont séparés par le signe plus (+).

remove

Supprime une ou plusieurs règles séparées entre elles par des signes plus.

set

Définit les valeurs de cette règle.

status

Affiche toutes les valeurs d'une ou de plusieurs règles.

useserver

Insère un serveur dans un jeu de règles.

Exemples

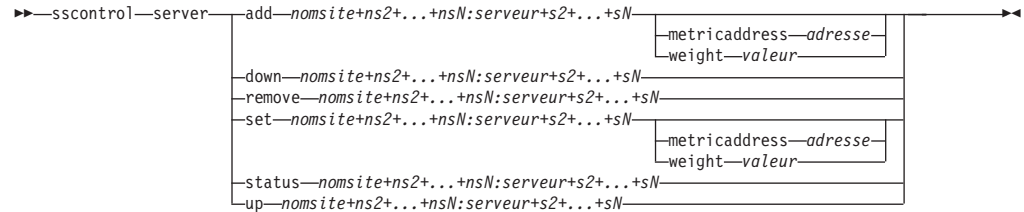
- Pour ajouter une règle qui sera toujours vraie, ne spécifiez pas de valeur de début ni de fin. Entrez :

```
sscontrol  
rule add sitename:rulename type true priority 100
```
- Pour créer une règle qui interdit l'accès à une plage d'adresses IP, en l'occurrence celles qui commencent par "9:", entrez :

```
sscontrol rule add sitename:rulename type ip b 9.0.0.0 e 9.255.255.255
```
- Pour créer une règle qui limitera l'utilisation d'un serveur donné de 11:00 à 15:00, entrez :

```
sscontrol rule add sitename:rulename type time beginrange 11 endrange 14  
sscontrol rule useserver sitename:rulename server05
```

sscontrol server — Configuration des serveurs



add

Permet d'ajouter ce serveur.

nom_site

Nom d'hôte ne pouvant être résolu demandé par le client. Le nom de site doit être un nom de domaine qualifié complet. Les noms de site supplémentaires sont séparés par le signe plus (+).

serveur

Adresse IP de la machine serveur TCP sous forme de nom symbolique ou d'adresse IP.

Remarque : Chaque serveur supplémentaire doit être séparé du précédent par le signe plus (+).

metricaddress

Adresse du serveur de mesures

adresse

Adresse du serveur sous forme de nom symbolique ou d'adresse IP.

weight

Nombre compris dans la fourchette 0–100 (mais qui ne doit pas dépasser la valeur de limite de pondération maximale pour le nom de site indiqué) représentant la pondération pour ce serveur. L'affectation de la valeur zéro à la pondération empêche l'envoi de nouvelles demandes au serveur. La valeur par défaut correspond à la moitié de la pondération maximale pour le nom de site indiqué. Si le gestionnaire est en cours d'exécution, ce paramètre est rapidement remplacé.

valeur

Valeur de pondération du serveur

down

Marque ce serveur comme étant arrêté. Empêche la transmission de toute autre demande à ce serveur.

remove

Permet de supprimer ce serveur.

set

Permet de définir des valeurs pour ce serveur.

status

Affiche l'état des serveurs.

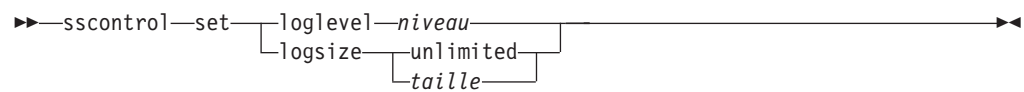
up

Marque ce serveur comme étant activé. Site Selector envoie désormais de nouvelles demandes à ce serveur.

Exemples

- Pour ajouter le serveur avec l'adresse 27.65.89.42 à un site dont le nom est site1, entrez :
`sscontrol server add site1:27.65.89.42`
- Pour marquer le serveur 27.65.89.42 comme étant arrêté, entrez :
`sscontrol server down site1:27.65.89.42`
- Pour supprimer le serveur avec l'adresse pour les sites, entrez :
`sscontrol server remove :27.65.89.42`
- Pour marquer le serveur 27.65.89.42 comme étant activé, entrez :
`sscontrol server up site1:27.65.89.42`

sscontrol set — Configuration du journal du serveur



loglevel

Niveau de consignation de ses activités par le serveur ssserver.

niveau

La valeur par défaut de **loglevel** est 0. Les valeurs possibles sont les suivantes :

- 0 correspond à Aucun
- 1 correspond à Minimal
- 2 correspond à De base
- 3 correspond à Modéré
- 4 correspond à Avancé
- 5 correspond à Prolixe

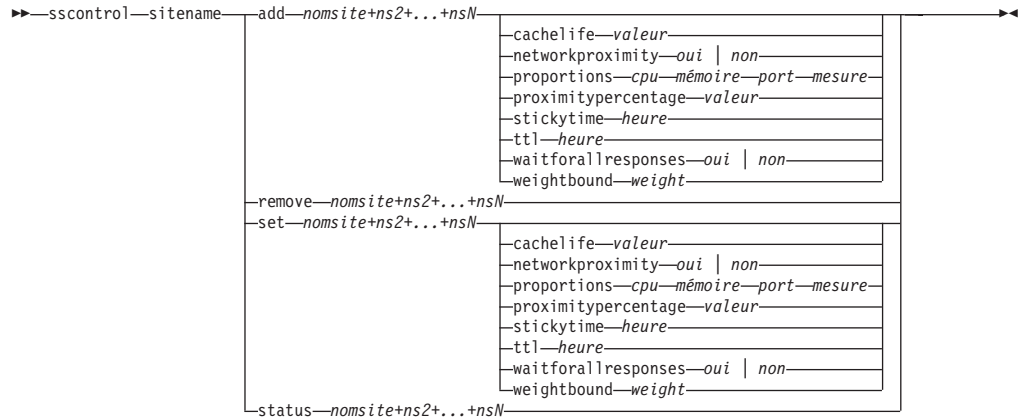
logsize

Nombre maximal d’octets à consigner dans le fichier journal.

taille

La taille de fichier journal par défaut est 1 Mo.

sscontrol sitename — Configuration d'un nom de site



add

Ajoute un nouveau nom de site

nom_site

Nom d'hôte ne pouvant être résolu, demandé par le client. Les noms de site supplémentaires sont séparés par le signe plus (+).

cachelife

Période pendant laquelle une réponse de proximité reste valide et enregistrée dans la mémoire cache. La valeur par défaut est 1800. Pour plus d'informations, voir «Utilisation de la fonction de proximité réseau (Network Proximity)», à la page 129.

valeur

Nombre positif représentant la période (en nombre de secondes) pendant laquelle une réponse de proximité reste valide et enregistrée dans la mémoire cache.

networkproximity

Détermine la proximité du réseau de chaque serveur par rapport au client demandeur. Tient compte de cette réponse de proximité pour la décision d'équilibrage de charge. Active ou désactive la fonction de proximité. Pour plus d'informations, voir «Utilisation de la fonction de proximité réseau (Network Proximity)», à la page 129.

valeur

Les options sont oui ou non. La valeur par défaut est non qui signifie que la fonction de proximité du réseau est désactivée.

proportions

Définit la proportion de la CPU, de la mémoire, du port (donnée provenant des conseillers) et des mesures système pour le serveur de mesures. Les proportions sont utilisées par le gestionnaire pour définir les pondérations du serveur. Chacune de ces valeurs est exprimée en pourcentage du total et, par conséquent, leur somme doit toujours être égale à 100.

cpu

Pourcentage de l'unité centrale utilisé sur chaque serveur avec équilibrage de charge (donnée provenant de l'agent Metric Server).

mémoire

Pourcentage de mémoire utilisé (donnée provenant de l'agent Metric Server) sur chaque serveur avec équilibrage de charge.

port Donnée provenant des conseillers en mode écoute sur ce port.

système

Donnée provenant de Metric Server.

proximitypercentage

Définit l'importance de la réponse de proximité, par rapport au bon fonctionnement du serveur (pondération du gestionnaire). Pour plus d'informations, voir «Utilisation de la fonction de proximité réseau (Network Proximity)», à la page 129.

valeur

La valeur par défaut est 50.

stickytime

Délai pendant lequel un client va recevoir l'ID serveur renvoyé pour la première demande. Le délai de maintien de routage par défaut est 0 (le nom de site n'est pas conservé).

heure

Nombre entier positif différent de zéro, représentant la période (en nombre de secondes) pendant laquelle le client reçoit l'ID serveur renvoyé pour la première demande.

ttl Définit la durée de vie. Indique la période pendant laquelle la réponse résolue reste en mémoire cache sur un autre serveur de noms. La valeur par défaut est 5.

valeur

Nombre positif représentant la période (en nombre de secondes) pendant laquelle la réponse résolue reste en mémoire cache sur le serveur de noms.

waitforallresponses

Permet de définir s'il est nécessaire d'attendre toutes les réponses de proximité des serveurs avant de répondre à la demande du client. Pour plus d'informations, voir «Utilisation de la fonction de proximité réseau (Network Proximity)», à la page 129.

valeur

Les valeurs possibles sont oui ou non. La valeur par défaut est oui.

weightbound

Nombre correspondant à la pondération maximale pouvant être définie pour les serveurs sur ce site. La valeur de limite de pondération définie pour le nom de site peut être remplacée pour des serveurs individuels à l'aide de la valeur de **pondération du serveur**. La valeur de limite de pondération par défaut pour le nom de site est égale à 20.

pondération

Valeur de la limite de pondération

set

Définit les propriétés du nom de site.

remove

Supprime le nom de site.

status

Affiche l'état en cours d'un nom de site spécifique.

Exemples

- Pour ajouter un nom de site, entrez :
`sscontrol sitename add 130.40.52.153`
- Pour activer la fonction de proximité du réseau, entrez :
`sscontrol sitename set monSite networkproximity yes`
- Pour définir une durée de vie de 1900000 secondes pour la mémoire cache, entrez :
`sscontrol sitename set monSite cachelife 1900000`
- Pour définir un pourcentage de proximité de 45, entrez :
`sscontrol sitename set monSite proximitypercentage 45`
- Pour empêcher un nom de site d'attendre toutes les réponses avant de répondre, entrez :
`sscontrol sitename set monSite waitforallresponses no`
- Pour définir une durée de vie de 7 secondes, entrez :
`sscontrol sitename set monSite ttl 7`
- Pour définir le niveau d'importance pour CpuLoad, MemLoad, Port et System Metric, entrez respectivement, :
`sscontrol sitename set monSite proportions 50 48 1 1`
- Pour supprimer un nom de site, entrez :
`sscontrol sitename remove 130.40.52.153`
- Pour afficher l'état du nom de site monSite, entrez :
`sscontrol sitename
status monSite`

Cette commande génère des résultats similaires à l'exemple suivant :

```
Etat
SiteName :
-----
SiteName ..... monSite
WeightBound ..... 20
TTL ..... 5
StickyTime ..... 0
Nombre de serveurs..... 1
Proportion accordée à CpuLoad ..... 49
Proportion accordée à MemLoad..... 50
Proportion accordée au port..... 1
Proportion accordée aux mesures du système... 0
Conseiller fonctionnant sur le port..... 80
Utilisation de la fonction de proximité..... N
```

sscontrol status — Affiche si le gestionnaire et les conseillers sont en cours d'exécution

►►—sscontrol—status—◀◀

Exemples

- Pour visualiser les programmes en cours d'exécution, entrez :
`sscontrol status`

Cette commande génère des résultats similaires à l'exemple suivant :

```
      NameServer
a été lancé.
      Le gestionnaire (manager) a été lancé
-----
| ADVISOR | SITENAME:PORT | TIMEOUT |
-----
|   http |           80 | unlimited |
-----
```

Chapitre 29. Guide des commandes Cisco CSS Controller

Le présent chapitre explique comment utiliser les commandes **ccocontrol** pour Cisco CSS Controller :

- «ccocontrol consultant — Configuration et contrôle d'un consultant», à la page 430
- «ccocontrol controller — Gestion du contrôleur», à la page 433
- «ccocontrol file — Gestion des fichiers de configuration», à la page 435
- «ccocontrol help — Affichage ou impression de l'aide relative à cette commande», à la page 437
- «ccocontrol highavailability — Contrôle de la haute disponibilité», à la page 438
- «ccocontrol metriccollector — Configuration du programme de collecte de mesures», à la page 441
- «ccocontrol ownercontent — Contrôle du nom de propriétaire et de la règle de contenu», à la page 443
- «ccocontrol service — Configuration d'un service», à la page 446

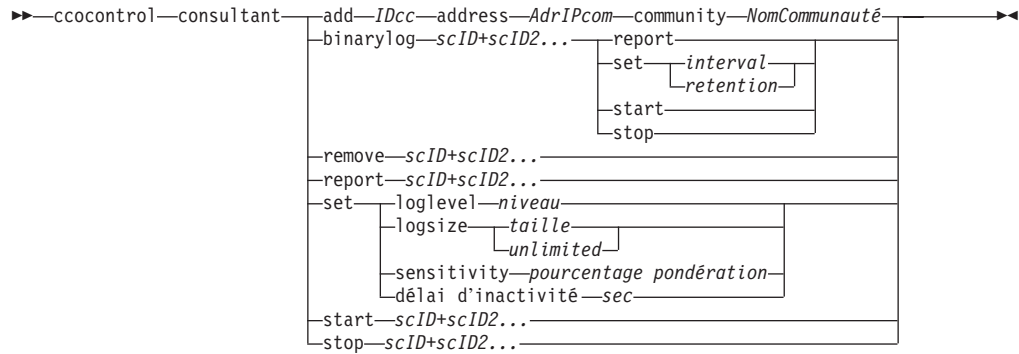
Vous pouvez utiliser une version abrégée des paramètres de la commande **ccocontrol** en entrant simplement la ou les quelques lettres d'identification des paramètres. Ainsi, pour obtenir de l'aide sur la commande **file save**, vous pouvez entrer **ccocontrol he f** au lieu de **ccocontrol help file**.

Pour afficher l'invite de la commande de **ccocontrol**, entrez : **ccocontrol**.

Pour fermer l'interface de ligne de commande, entrez **exit** or **quit**.

Remarque : Utilisez les lettres de l'anglais pour toutes les valeurs des paramètres des commandes. Les seules exceptions s'appliquent aux noms d'hôte (utilisés dans les commandes **server**) et aux noms de fichiers (utilisés dans les commandes **file**).

cococontrol consultant — Configuration et contrôle d'un consultant



add

Ajoute un consultant de commutateur.

IDcc (ID du consultant du commutateur)

Chaîne définie par l'utilisateur, qui désigne le consultant.

address

Adresse IP à laquelle Cisco CSS Switch fournit des pondérations.

AdrIPCom (Adresse IP du commutateur)

Adresse IP du commutateur.

community

Nom utilisé dans SNMP pour établir et définir des communications avec Cisco CSS Switch.

NomCommunauté

Nom de communauté en lecture/écriture de Cisco CSS Switch.

binarylog

Contrôle la consignation binaire d'un consultant.

report

Rapports concernant les caractéristique de la consignation binaire.

set

Fixe l'intervalle, en secondes, entre chaque enregistrement d'informations dans les journaux binaires. L'option de consignation binaire permet d'enregistre dans des fichiers journaux binaires des informations concernant tous les services de la configuration. Les informations ne sont écrites dans les journaux que si l'intervalle de consignation indiqué a expiré depuis l'écriture du dernier enregistrement dans le journal. L'intervalle par défaut de consignation binaire est de 60 secondes.

interval

Intervalle, en secondes, entre chaque entrée dans le journal binaire.

retention

Nombre d'heures pendant lesquelles les fichiers journaux binaires sont conservés.

start

Lance la consignation binaire.

stop

Arrête la consignation binaire.

remove

Supprime un consultant de commutateur.

report

Rapports concernant les caractéristique des consultants de commutateur.

set

Définit les caractéristique des consultants de commutateur.

loglevel

Définit le niveau auquel le consultant de commutateur consigne les activités.
La valeur par défaut est 1.

niveau

Niveau compris entre 0 et 5. La valeur par défaut est 1. Les valeurs admises sont les suivantes :

- 0 = Aucun
- 1 = Minimal
- 2 = De base
- 3 = Modéré
- 4 = Avancé
- 5 = Prolixe

logsize

Nombre maximal d'octets consignés dans le journal. La valeur par défaut est 1 048 576. Lorsque vous attribuez une taille maximale au fichier journal, ce dernier fonctionne en boucle. Lorsque le fichier atteint la taille indiquée, les entrées suivantes sont écrites à partir du haut du fichier et remplacent les entrées existantes. La valeur indiquée par logsize ne peut pas être inférieure à la taille actuelle du fichier journal. Les entrées du journal sont horodatées, ce qui permet de déterminer l'ordre dans lequel elles ont été enregistrées. Plus le niveau de consignation est élevé, plus la taille du journal doit être choisie avec soin car l'espace peut être saturé rapidement.

taille

Nombre maximal d'octets consignés dans le journal du consultant. Vous pouvez indiquer un nombre positif supérieur à zéro, ou le mot **unlimited**. Il est possible que le fichier journal n'atteigne pas la taille maximale exacte avant le remplacement des entrées existantes, car la taille des entrées de journal elles-même varie.

sensitivity

Quantité de modifications requises entre la nouvelle et l'ancienne pondérations pour que la pondération soit changée. La différence entre la nouvelle et l'ancienne pondérations doit être supérieure au pourcentage de sensibilité afin que la pondération puisse changer. Les valeurs admises vont de 0 à 100, la valeur par défaut étant 5.

pourcentage de pondération

Nombre compris entre 1 et 100 correspondant à la sensibilité.

sleeptime

Nombre de secondes entre chaque cycle de définition des pondérations. Valeur par défaut : 7.

sec

Entier correspondant au nombre de secondes d'inactivité. Les valeurs admises vont de 0 à 2 147 460.

start

Lance la collecte de mesures et la définition de pondérations.

stop

Arrête la collecte de mesures et la définition de pondérations.

Exemples

- Pour ajouter un consultant de commutateur dont l'identificateur de commutateur est cc1, l'adresse IP 9.37.50.17 et le nom de communauté comm1, entrez :

```
ccocontrol consultant add cc1 address 9.37.50.17 community comm2
```

- Pour démarrer la consignation binaire, entrez :

```
ccocontrol consultant binarylog cc1 start
```

- Pour afficher un rapport concernant les caractéristiques du consultant de commutateur cc1, entrez :

```
ccocontrol consultant report cc1
```

Cette commande génère des résultats similaires à l'exemple suivant :

```
Consultant cc1 connecté au commutateur à l'adresse 9.37.50.1:cn1
```

```
Consultant démarré
```

```
Délai d'inactivité = 7
```

```
Sensibilité = 5
```

```
Niveau consignation = 5
```

```
Taille du journal = 1 048 576
```

```
Contenu(s) de propriétaire :
```

```
contenu de propriétaire cp1
```

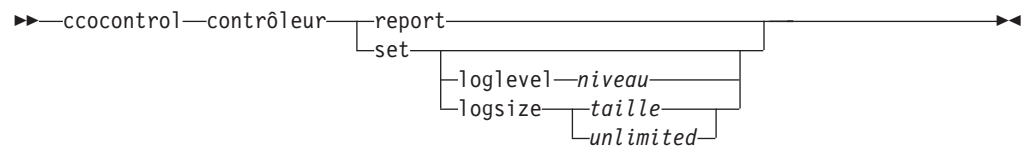
- Pour fixer à 10 secondes le délai d'inactivité entre les cycles de définition des pondérations pour l'ID de commutateur cc1, entrez :

```
ccocontrol consultant set cc1 sleeptime 10
```

- Pour lancer la collecte des mesures et la définition des pondérations pour l'ID du consultant de cc1, entrez :

```
ccocontrol consultant start cc1
```

cococontrol controller — Gestion du contrôleur



report

Affiche les caractéristiques du contrôleur. Ce rapport affiche également les informations relatives à la version.

set

Définit les caractéristiques du contrôleur.

loglevel

Définit le niveau auquel le contrôleur consigne les activités. La valeur par défaut est 1.

niveau

Niveau compris entre 0 et 5. La valeur par défaut est 1. Les valeurs admises sont les suivantes :

- 0 = Aucun
- 1 = Minimal
- 2 = De base
- 3 = Modéré
- 4 = Avancé
- 5 = Prolixe

logsize

Nombre maximal d'octets consignés dans le journal. La valeur par défaut est 1 048 576. Lorsque vous attribuez une taille maximale au fichier journal, ce dernier fonctionne en boucle. Lorsque le fichier atteint la taille indiquée, les entrées suivantes sont écrites à partir du haut du fichier et remplacent les entrées existantes. La valeur indiquée par logsize ne peut pas être inférieure à la taille actuelle du fichier journal. Les entrées du journal sont horodatées, ce qui permet de déterminer l'ordre dans lequel elles ont été enregistrées. Plus le niveau de consignation est élevé, plus la taille du journal doit être choisie avec soin car l'espace peut être saturé rapidement.

taille | unlimited

Nombre maximal d'octets consignés dans le journal du consultant. Vous pouvez indiquer un nombre positif supérieur à zéro, ou le mot **unlimited**. Il est possible que le fichier journal n'atteigne pas la taille maximale exacte avant le remplacement des entrées existantes, car la taille des entrées de journal elles-même varie.

Exemples

- Pour afficher un rapport concernant le contrôleur, entrez :

```
cococontrol controller report
```

Cette commande génère des résultats similaires à l'exemple suivant :

Rapport du contrôleur :

```
-----
Version . . . . . Version: 05.00.00.00 - 03/21/2002-09:49:57-EST
Niveau consignation . . . 1
Taille journal . . . . . 1048576
```

Fichier configuration . . config1.xml

Consultants :

Consultant consult1 -démarré

- Pour affecter au niveau de consignation la valeur zéro afin d'optimiser les performances, entrez :

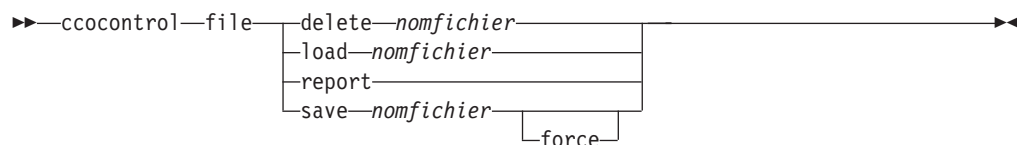
cococontrol set loglevel 0

- Pour fixer la taille du journal du contrôleur à 1 000 000 octets, entrez :

cococontrol controller

set logsize 1000000

cococontrol file — Gestion des fichiers de configuration



delete

Supprimer le fichier de configuration indiqué.

nomfichier

Fichier de configuration. L'extension doit être .xml. Si cette extension n'est pas indiquée, elle est ajoutée par défaut.

load

Charge la configuration enregistrée dans le fichier indiqué.

Remarque : Le chargement d'un fichier ajoute la configuration stockée dans ce fichier à la configuration en cours. Si vous voulez charger une *nouvelle* configuration, vous devez arrêter puis redémarrer le serveur avant de charger le fichier de la nouvelle configuration.

report

Liste des fichiers de configuration.

save

Sauvegarde la configuration en cours dans le fichier indiqué.

Remarque : Les fichiers sont sauvegardés dans les répertoires suivants et chargés à partir de ces mêmes répertoires :

- Systèmes AIX : `/opt/ibm/edge/lb/servers/configurations/cco`
- Systèmes Linux : `/opt/ibm/edge/lb/servers/configurations/cco`
- Systèmes Solaris : `/opt/ibm/edge/lb/servers/configurations/cco`
- Systèmes Windows :
Répertoire d'installation (par défaut) : `C:\Program Files\ibm\edge\lb\servers\configurations\cco`

force

Sauvegarde dans un fichier existant.

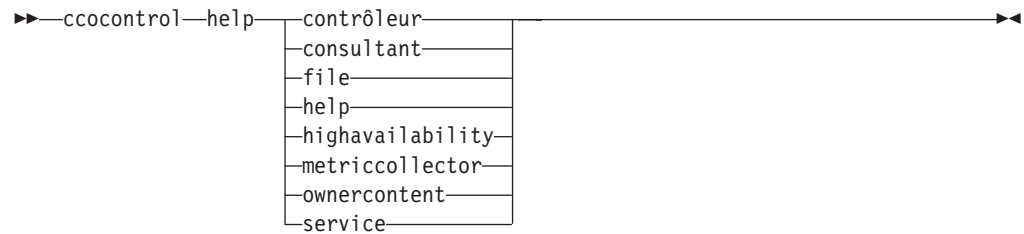
Exemples

- Pour supprimer un fichier nommé fichier1, entrez :
`cococontrol file delete fichier1`
- Pour ajouter la configuration du fichier à la configuration en cours, entrez :
`cococontrol file load config2`
- Pour afficher un rapport des fichiers précédemment sauvegardés, entrez :
`cococontrol file report`
Cette commande génère des résultats similaires à l'exemple suivant :
RAPPORT SUR LES FICHIERS :

fichier1.xml
fichier2.xml
fichier3.xml
- Pour sauvegarder votre configuration dans un fichier nommé fichier3 :

```
cococontrol file save config2
```

cococontrol help — Affichage ou impression de l'aide relative à cette commande



Exemples

- Pour obtenir de l'aide sur la commande cococontrol, entrez :

```
cococontrol help
```

Cette commande génère des résultats similaires à l'exemple suivant :

Les commandes suivantes sont

disponibles :

controller	- s'applique au contrôleur
consultant	- s'applique aux consultants du commutateur
file	- s'applique aux fichiers de configuration
help	- s'applique à l'aide
highavailability	- s'applique à la haute disponibilité
metriccollector	- s'applique aux programmes de collecte de mesures
ownerContent	- s'applique au contenu des propriétaires
service	- s'applique aux services

- Les symboles suivants sont utilisés dans la syntaxe de l'aide en ligne :

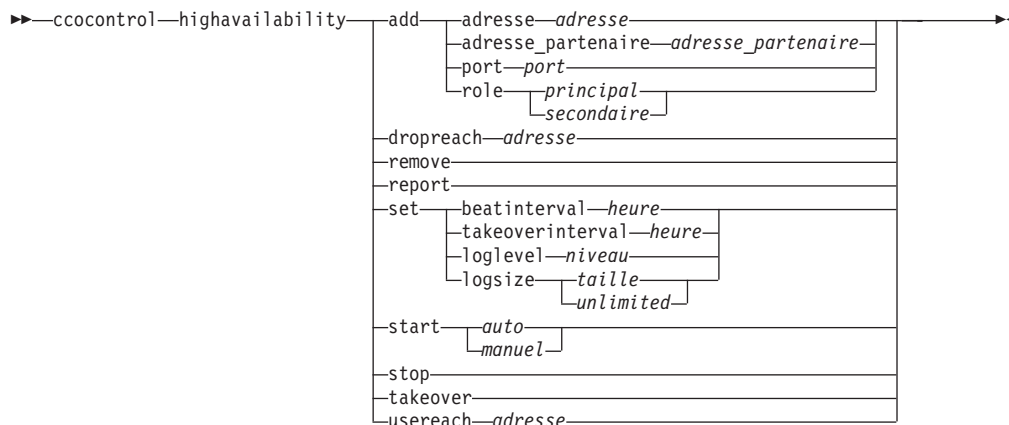
< > Paramètres entre accolades ou séquence de caractères.

[] Éléments facultatifs entre accolades.

| Barre verticale de séparation des choix possibles au sein des parenthèses et des accolades.

: Signe deux-points pour séparer les noms, par exemple,
Consultant1:ContenuPropriétaire1.

cococontrol highavailability — Contrôle de la haute disponibilité



add

Configure un noeud de haute disponibilité, un partenaire et des cibles à contacter.

address

Adresse pour la réception des signaux de présence.

adresse

Adresse IP du noeud à haute disponibilité.

partneraddress

Adresse pour l'envoi des signaux de présence. Il s'agit de l'adresse IP ou du nom d'hôte configuré sur le noeud partenaire. Cette adresse permet de communiquer avec la machine haute disponibilité partenaire.

adresse

Adresse IP du partenaire.

port

Port utilisé pour communiquer avec le partenaire. La valeur par défaut est 12345.

port

Numéro du port.

role

Rôle de haute disponibilité.

principal | secondaire

Rôles principal et secondaire de haute disponibilité.

dropreach

Supprime cette cible à contacter des critères de haute disponibilité.

adresse

Adresse IP de la cible à contacter.

remove

Supprime le noeud, le partenaire et la cible à contacter de la configuration de haute disponibilité. Vous devez arrêter la haute disponibilité avant de lancer cette commande.

report

Affiche les informations de haute disponibilité.

set

Définit les caractéristiques de la haute disponibilité.

beatinterval

Définit l'intervalle, en millisecondes, au bout duquel les signaux de présence sont envoyés au partenaire. La valeur par défaut est 500.

heure

Entier positif correspondant à l'écart, en millisecondes, entre chaque signal de présence.

takeoverinterval

Définit la durée, en millisecondes, qui doit s'écouler (pendant laquelle aucun signal de présence n'est reçu) avant qu'une reprise ne se produise. La valeur par défaut est 2000.

heure

Entier positif correspondant à l'écart, en millisecondes, entre chaque reprise.

loglevel

Définit le niveau auquel les activités sont consignées. La valeur par défaut est 1.

niveau

Niveau compris entre 0 et 5. La valeur par défaut est 1. Les valeurs admises sont les suivantes :

- 0 = Aucun
- 1 = Minimal
- 2 = De base
- 3 = Modéré
- 4 = Avancé
- 5 = Prolixe

logsize

Nombre maximal d'octets consignés dans le journal de haute disponibilité. La valeur par défaut est 1 048 576. Lorsque vous attribuez une taille maximale au fichier journal, ce dernier fonctionne en boucle. Lorsque le fichier atteint la taille indiquée, les entrées suivantes sont écrites à partir du haut du fichier et remplacent les entrées existantes. La valeur indiquée par logsize ne peut pas être inférieure à la taille actuelle du fichier journal. Les entrées du journal sont horodatées, ce qui permet de déterminer l'ordre dans lequel elles ont été enregistrées. Plus le niveau de consignation est élevé, plus la taille du journal doit être choisie avec soin car l'espace peut être saturé rapidement.

taille | unlimited

Nombre maximal d'octets consignés dans le journal de haute disponibilité. Vous pouvez indiquer un nombre positif supérieur à zéro, ou le mot **unlimited**. Il est possible que le fichier journal n'atteigne pas la taille maximale exacte avant le remplacement des entrées existantes, car la taille des entrées de journal elles-mêmes varie.

start

Active la haute disponibilité. Avant d'utiliser cette commande, vous devez configurer un noeud de haute disponibilité, un partenaire et une cible à contacter.

auto | manuelle

Détermine si la haute disponibilité démarre avec une stratégie de reprise automatique ou manuelle.

stop

Arrête la haute disponibilité.

takeover

Reprend le contrôle au noeud de haute disponibilité actif.

usereach

Adresse de la cible à contacter qui commencer à utiliser la haute disponibilité.
Ajoutez une cible à contacter à l'aide de ping de sorte que les partenaires de haute disponibilité puissent déterminer dans quelle mesure leurs cibles sont accessibles.

adresse

Adresse IP de la cible à contacter.

Exemples

- Pour ajouter le noeud de haute disponibilité d'adresse IP 9.37.50.17, avec le rôle principal sur le port 12345 et l'adresse de partenaire 9.37.50.14, entrez :

```
cococontrol highavailability add
address 9.37.50.17 role primary port 12345 partneraddress 9.37.50.14
```

- Pour ajouter l'adresse de cible à contacter 9.37.50.9, entrez :

```
cococontrol highavailability usereach
9.37.50.9
```

- Pour supprimer l'adresse de cible à contacter 9.37.50.9, entrez :

```
cococontrol highavailability dropreach
9.37.50.9
```

- Pour démarrer la haute disponibilité démarre avec une stratégie de reprise manuelle, entrez :

```
cococontrol
highavailability start manual
```

- Pour obtenir une analyse statistique de la haute disponibilité, entrez :

```
cococontrol highavailability report
```

Cette commande génère des résultats similaires à l'exemple suivant :

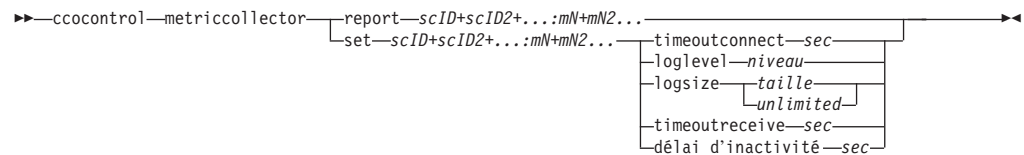
Etat de la haute disponibilité :

```
Noeud . . . . . principal
Adresse du noeud . . . . 9.37.50.17
Port . . . . . 12345
Adresse du partenaire . . 9.37.50.14
Stratégie de reprise . . . manual
Int. entre sig. prés. . . 500
Int. entre reprises . . . 2000
Etat . . . . . inactif
Sous-état . . . . . non synchronisé
```

Etat d'accessibilité : Noeud/Partenaire

Aucune cible configurée

cococontrol metriccollector — Configuration du programme de collecte de mesures



report

Affiche les caractéristiques du programme de collecte de mesures.

IDcc (ID du consultant du commutateur)

Chaîne définie par l'utilisateur, qui désigne le consultant.

Nm (nom de la mesure)

Nom qui identifie la mesure personnalisée ou fournie.

set

Définit les caractéristiques du programme de collecte de mesures.

timeoutconnect

Délai d'attente observé par le programme de collecte de mesures avant de signaler l'échec d'une connexion.

sec Entier positif correspondant au nombre de secondes à l'expiration desquelles le programme de collecte de mesures signale qu'une connexion à un service a échoué.

loglevel

Définit le niveau auquel le consultant indiqué consigne les activités. Valeur par défaut : 1.

niveau

Numéro du niveau. La valeur par défaut est 1. Plus le numéro est élevé, plus la quantité d'informations consignée dans le journal du consultant est importante. Les valeurs admises sont les suivantes :

- 0 = Aucun
- 1 = Minimal
- 2 = De base
- 3 = Modéré
- 4 = Avancé
- 5 = Prolixe

logsize

Nombre maximal d'octets consignés dans le journal. La valeur par défaut est 1 048 576. Lorsque vous attribuez une taille maximale au fichier journal, ce dernier fonctionne en boucle. Lorsque le fichier atteint la taille indiquée, les entrées suivantes sont écrites à partir du haut du fichier et remplacent les entrées existantes. La valeur indiquée par logsize ne peut pas être inférieure à la taille actuelle du fichier journal. Les entrées du journal sont horodatées, ce qui permet de déterminer l'ordre dans lequel elles ont été enregistrées. Plus le niveau de consignation est élevé, plus la taille du journal doit être choisie avec soin car l'espace peut être saturé rapidement.

taille | unlimited

Nombre maximal d'octets consignés dans le journal du consultant. Vous pouvez indiquer un nombre positif supérieur à zéro, ou le mot **unlimited**. Il

est possible que le fichier journal n'atteigne pas la taille maximale exacte avant le remplacement des entrées existantes, car la taille des entrées de journal elles-même varie.

timeoutreceive

Délai d'attente observé par un consultant avant de signaler l'échec d'une procédure de réception.

sec Il s'agit d'un entier positif qui représente le délai en secondes à l'expiration duquel le consultant signale que la réception d'un envoi provenant d'un service a échoué.

sleeptime

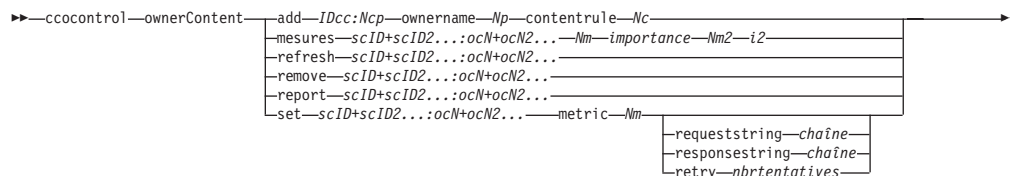
Nombre de secondes d'inactivité du programme de collecte de mesures entre chaque cycle de collecte de mesures.

Entier positif correspondant au nombre de secondes d'inactivité.

Exemples

- Pour afficher un rapport concernant les caractéristiques d'un programme de collecte de mesures, entrez :
`ccocontrol metriccollector report cc1:http`
Cette commande génère des résultats similaires à l'exemple suivant :
Collecte de mesures cc1:http
 measure(s) collectée(s).... http
 niveau consignation..... 5
 taille journal..... 1048576
 Nbr. sec. inactivité..... 7
 délai expiration conn..... 21
 délai expiration récep.... 21
- Pour définir un délai d'expiration de connexion de 15 secondes et une taille de journal illimitée pour le consultant de commutateur cc1 et la mesure http, entrez :
`ccocontrol metriccollector set cc1:http timeoutconnect 15 logsize unlimited`

cococontrol ownercontent — Contrôle du nom de propriétaire et de la règle de contenu



add

Ajoute un contenu de propriétaire au consultant désigné.

IDcc (ID du consultant du commutateur)

Chaîne, définie par l'utilisateur, qui représente le consultant.

Nomcp (nom du contenu de propriétaire)

Chaîne, définie par l'utilisateur, qui représente le nom de propriétaire et la règle de contenu définis sur le commutateur.

ownername

Nom configuré sur le commutateur, qui identifie la configuration du propriétaire.

Np (nom du propriétaire)

Chaîne de texte unique ne contenant aucun espace. Le nom de propriétaire indiqué ici doit être identique à celui spécifié sur le commutateur Cisco.

contentrule

Nom configuré sur le commutateur, qui identifie la configuration de la règle de contenu du propriétaire.

Nc (nom du contenu)

Chaîne de texte unique ne contenant aucun espace. Le nom de contenu indiqué ici doit être identique à celui spécifié sur le commutateur Cisco.

metrics

Désigne l'ensemble des mesures utilisées pour calculer les pondérations et l'importance de chaque mesure. L'importance est exprimée en pourcentage du total. Par conséquent, la somme des valeurs d'importance doit toujours être égale à 100. Les mesures peuvent être toute combinaison de mesures de données de connexion, mesures de conseiller d'application et mesures de serveur de mesures. Par défaut il s'agit du nombre de connexions actives (activeconn) et du débit de la connexion (connrate) avec une importance équivalente de 50/50.

Nm (nom de la mesure)

Nom qui identifie le programme chargé de collecter les mesures afin de déterminer la pondération du serveur.

Voici la liste des noms de mesure admis et des ports qui leur sont associés.

Nom du conseiller	Protocole	Port
connect	ICMP	12345
db2	privé	50000
dns	DNS	53
ftp	FTP	21

Nom du conseiller	Protocole	Port
http	HTTP	80
https	SSL	443
cachingproxy	HTTP (via Caching Proxy)	80
imap	IMAP	143
ldap	LDAP	389
nntp	NNTP	119
ping	PING	0
pop3	POP3	110
sip	SIP	5060
smtp	SMTP	25
ssl	SSL	443
telnet	Telnet	23
WLM	privé	10,007
activeconn	Non disp.	Non disp.
connrate	Non disp.	Non disp.
cpuload	Non disp.	Non disp.
memload	Non disp.	Non disp.

importance

Nombre compris entre 0 et 100 qui correspond à l'importance de la mesure dans le calcul des pondérations du serveur.

refresh

Actualise les services configurés avec la configuration issue de Cisco CSS Switch.

remove

Supprime un contenu de propriétaire.

report

Répertoire les caractéristiques des contenus de propriétaire.

set

Définit les caractéristiques des contenus de propriétaire.

mesure

Définit les caractéristique d'une mesure.

Nm

Nom de la mesure souhaitée.

requeststring

Définit une chaîne de demande pour la mesure désignée. Cette chaîne correspond à la demande envoyée par un programme de collecte de mesures pour rassembler des informations de mesure.

chaîne

Chaîne de demande envoyée au serveur par le programme de collecte de mesures.

responsestring

Définit une chaîne de réponse pour la mesure désignée. Le programme de

collecte de mesures utilise cette chaîne pour comparer les réponses qu'il reçoit des serveurs et déterminer ainsi la disponibilité du serveur.

chaîne

Chaîne de réponse à laquelle le programme de collecte de mesures compare les réponses reçues du serveur.

tentative

Nombre de tentatives accordées avant qu'un serveur ne soit déclaré arrêté.

nbrtentatives

Entier supérieur ou égal à zéro. Il est préférable que le nombre de tentatives ne dépasse pas 3. Par défaut, le nombre de tentatives est égal à zéro.

Exemples

- Pour ajouter le nom de contenu de propriétaire cp1 (de nom de propriétaire propriétaire1 et de nom de contenu contenu1) à l'ID de consultant de commutateur cc1, entrez :

```
ccocontrol ownerContent add cc1:cp1 ownername  
propriétaire1 contentruler contenu1
```
- Pour appliquer une proportion de 50 à chacune des mesures activeconn et http, entrez :

```
ccocontrol ownerContent metrics cc1:cp1 activeconn 50 http 50
```
- Pour afficher le rapport des caractéristiques des contenus de propriétaire, entrez :

```
ccocontrol ownerContent  
report cc1:cp1
```

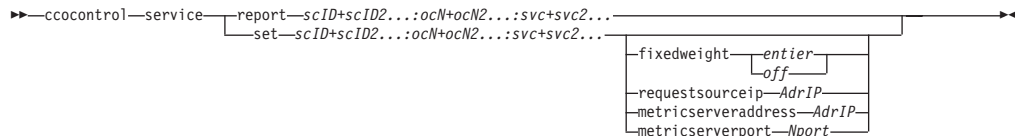
Cette commande génère des résultats similaires à l'exemple suivant :

```
ownerContent cc1:cp1  
  Limite de pondération = 10  
  La mesure activeconn a la proportion 25  
    Chaîne de réponse... non disp.  
    Chaîne de demande.... non disp.  
  La mesure http a la proportion 50  
    Chaîne de réponse... non disp.  
    Chaîne de demande.... non disp.  
  La mesure connrate a la proportion 25  
    Chaîne de réponse... non disp.  
    Chaîne de demande.... non disp.  
  Contient le service t3  
  Contient le service t2  
  Contient le service t1
```

- Pour définir une chaîne de demande http, entrez :

```
ccocontrol ownerContent set cc1:cp1 metric http  
requeststring getCookie
```

cococontrol service — Configuration d'un service



report

Affiche les caractéristiques des services.

IDcc (ID du consultant du commutateur)

Chaîne, définie par l'utilisateur, qui représente le consultant.

Nomcp (nom du contenu de propriétaire)

Chaîne, définie par l'utilisateur, qui représente le nom de propriétaire et la règle de contenu définis sur le commutateur.

svc (service)

Chaîne, définie par l'utilisateur sur le commutateur, qui représente le service.

set

Définit les caractéristiques des services.

fixedweight

Affecte une pondération fixe au service. La valeur par défaut est off.

entier | off

Entier positif, compris entre 0 et 10, correspondant à la pondération fixe du service ou le mot **off** si aucune pondération fixe ne doit être définie.

requestsourceip

Définit l'adresse à partir de laquelle contacter le service pour les demandes d'application.

AdrIP (adresse IP)

Adresse IP à partir de laquelle le service doit être contacté, sous forme d'adresse symbolique ou d'adresse IP.

metricserveraddress

Définit l'adresse à partir de laquelle contacter le service pour les demandes de serveur de mesures.

AdrIP (adresse IP)

Adresse IP du serveur de mesures, sous forme d'adresse symbolique ou d'adresse IP.

metricserverport

Définit le port à utiliser pour contacter le serveur de mesures.

Nport (numéro de port)

Numéro du port utilisé pour contacter le serveur de mesures.

Exemples

- Pour afficher un rapport concernant le service t1 du consultant cc1, entrez :

```
cococontrol service report
cc1:cp1:t1
```

Cette commande génère des résultats similaires à l'exemple suivant :

```
Service cc1:cp1:ta a la
pondération 10
La pondération fixe est off
```

IP source de la demande..... 9.27.24.156
Port d'application..... 80
Adresse du serveur de mesures... 1.0.0.1
Port du serveur de mesures..... 10004
La mesure activeconn a la valeur -99
La mesure http a la valeur -99
La mesure connrate a la valeur -99

- Pour définir une adresse de serveur de mesures pour le service t2, entrez :

```
cococontrol service set  
cc1:cp1:t2 metricserveraddress 9.37.50.17
```

Chapitre 30. Guide des commandes Nortel Alteon Controller

Le présent chapitre explique comment utiliser les commandes **nalcontrol** pour Nortel Alteon Controller :

- «nalcontrol consultant — Configuration et contrôle d'un consultant», à la page 450
- «nalcontrol controller — Gestion du contrôleur», à la page 453
- «nalcontrol file — Gestion des fichiers de configuration», à la page 455
- «nalcontrol help — Affichage ou impression de l'aide relative à cette commande», à la page 457
- «nalcontrol highavailability — Contrôle de la haute disponibilité», à la page 458
- «nalcontrol metriccollector — Configuration du programme de collecte de mesures», à la page 461
- «nalcontrol service — Configuration d'un service», à la page 465
- «nalcontrol server — Configuration d'un serveur», à la page 463

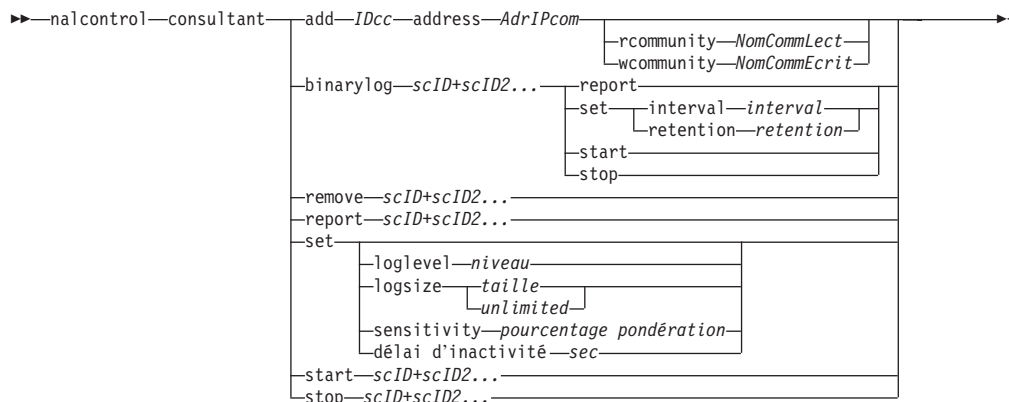
Vous pouvez utiliser une version abrégée des paramètres de la commande **nalcontrol** en entrant simplement la ou les quelques lettres d'identification des paramètres. Par exemple, pour obtenir l'aide correspondant à la commande **file save**, vous pouvez entrer **nalcontrol he f** au lieu de **nalcontrol help file**.

Pour afficher l'invite de la commande de **nalcontrol**, entrez : **nalcontrol**.

Pour fermer l'interface de ligne de commande, entrez **exit** or **quit**.

Remarque : Utilisez les lettres de l'anglais pour toutes les valeurs des paramètres des commandes. Les seules exceptions s'appliquent aux noms d'hôte (utilisés dans les commandes **server**) et aux noms de fichiers (utilisés dans les commandes **file**).

nalcontrol consultant — Configuration et contrôle d'un consultant



add

Ajoute un consultant de commutateur.

IDcc

Chaîne définie par l'utilisateur, qui désigne le consultant.

address

Adresse IP à laquelle Nortel Alteon Web Switch fournit des pondérations.

AdrIPcom

Adresse IP du commutateur.

rcommunity

Nom de communauté en lecture utilisé dans les communications SNMP get avec Nortel Alteon Web Switch. La valeur par défaut est public.

NomCommLect

Chaîne correspondant au nom de communauté en lecture, tel que configuré sur Nortel Alteon Web Switch. La valeur par défaut est public.

wcommunity

Nom de communauté en écriture utilisé dans les communications SNMP set.

writeCommName

Chaîne correspondant au nom de communauté en écriture, tel que configuré sur Nortel Alteon Web Switch. La valeur par défaut est privée.

binarylog

Contrôle la consignation binaire d'un consultant.

report

Rapports concernant les caractéristique de la consignation binaire.

set

Fixe l'intervalle, en secondes, entre chaque enregistrement d'informations dans les journaux binaires. L'option de consignation binaire permet d'enregistrer dans des fichiers journaux binaires des informations concernant tous les services de la configuration. Les informations ne sont écrites dans les journaux que si l'intervalle de consignation indiqué a expiré depuis l'écriture du dernier enregistrement dans le journal. L'intervalle par défaut de consignation binaire est de 60 secondes.

interval

Intervalle, en secondes, entre chaque entrée dans le journal binaire.

rétenction

Nombre d'heures pendant lesquelles les fichiers journaux binaires sont conservés.

start

Lance la consignation binaire.

stop

Arrête la consignation binaire.

remove

Supprime un consultant de commutateur.

report

Rapports concernant les caractéristique des consultants de commutateur.

set

Définit les caractéristique des consultants de commutateur.

loglevel

Définit le niveau auquel le consultant de commutateur consigne les activités. La valeur par défaut est 1.

niveau

Niveau compris entre 0 et 5. La valeur par défaut est 1. Les valeurs admises sont les suivantes :

- 0 = Aucun
- 1 = Minimal
- 2 = De base
- 3 = Modéré
- 4 = Avancé
- 5 = Prolixe

logsize

Nombre maximal d'octets consignés dans le journal. La valeur par défaut est 1 048 576. Lorsque vous attribuez une taille maximale au fichier journal, ce dernier fonctionne en boucle. Lorsque le fichier atteint la taille indiquée, les entrées suivantes sont écrites à partir du haut du fichier et remplacent les entrées existantes. La valeur indiquée par logsize ne peut pas être inférieure à la taille actuelle du fichier journal. Les entrées du journal sont horodatées, ce qui permet de déterminer l'ordre dans lequel elles ont été enregistrées. Plus le niveau de consignation est élevé, plus la taille du journal doit être choisie avec soin car l'espace peut être saturé rapidement.

taille

Nombre maximal d'octets consignés dans le journal du consultant. Vous pouvez indiquer un nombre positif supérieur à zéro, ou le mot **unlimited**. Il est possible que le fichier journal n'atteigne pas la taille maximale exacte avant le remplacement des entrées existantes, car la taille des entrées de journal elles-même varie.

sensitivity

Quantité de modifications requises entre la nouvelle et l'ancienne pondérations pour que la pondération soit changée. La différence entre la nouvelle et l'ancienne pondérations doit être supérieure au pourcentage de sensibilité afin que la pondération puisse changer. Les valeurs admises vont de 0 à 100, la valeur par défaut étant 5.

pourcentage de pondération

Nombre compris entre 1 et 100 correspondant à la sensibilité.

sleeptime

Nombre de secondes entre chaque cycle de définition des pondérations. Valeur par défaut : 7.

secondes

Entier correspondant au nombre de secondes d'inactivité. Les valeurs admises vont de 0 à 2 147 460.

start

Lance la collecte de mesures et la définition de pondérations.

stop

Arrête la collecte de mesures et la définition de pondérations.

Exemples

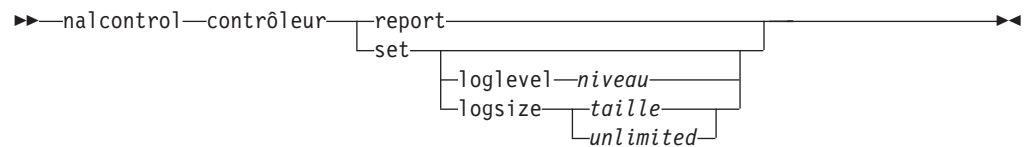
- Pour ajouter un consultant de commutateur dont l'identificateur de commutateur est cc1 et l'adresse IP 9.37.50.17, entrez :
nalcontrol consultant add cc1 address 9.37.50.17
- Pour démarrer la consignation binaire, entrez :
nalcontrol consultant binarylog cc1 start
- Pour afficher un rapport concernant les caractéristiques du consultant de commutateur cc1, entrez :
nalcontrol consultant report cc1

Cette commande génère des résultats similaires à l'exemple suivant :

```
ID
consultant : cc1 Adresse IP commutateur : 9.37.50.1
Communauté en lecture : publique
Communauté en écriture : privée
Consultant démarré
    Délai d'inactivité = 7
    Sensibilité       = 5
    Niveau consignation = 5
    Taille du journal  = 1 048 576
    Service(s) :
        Service svc1
```

- Pour fixer à 10 secondes le délai d'inactivité entre les cycles de définition des pondérations pour l'ID de commutateur cc1, entrez :
nalcontrol consultant set cc1 sleeptime 10
- Pour lancer la collecte des mesures et la définition des pondérations pour l'ID du consultant de cc1, entrez :
nalcontrol consultant start cc1

nalcontrol controller — Gestion du contrôleur



report

Affiche les caractéristiques du contrôleur. Ce rapport affiche également les informations relatives à la version.

set

Définit les caractéristiques du contrôleur.

loglevel

Définit le niveau auquel le contrôleur consigne les activités. La valeur par défaut est 1.

niveau

Niveau compris entre 0 et 5. La valeur par défaut est 1. Les valeurs admises sont les suivantes :

- 0 = Aucun
- 1 = Minimal
- 2 = De base
- 3 = Modéré
- 4 = Avancé
- 5 = Prolixe

logsize

Nombre maximal d'octets consignés dans le journal. La valeur par défaut est 1 048 576. Lorsque vous attribuez une taille maximale au fichier journal, ce dernier fonctionne en boucle. Lorsque le fichier atteint la taille indiquée, les entrées suivantes sont écrites à partir du haut du fichier et remplacent les entrées existantes. La valeur indiquée par logsize ne peut pas être inférieure à la taille actuelle du fichier journal. Les entrées du journal sont horodatées, ce qui permet de déterminer l'ordre dans lequel elles ont été enregistrées. Plus le niveau de consignation est élevé, plus la taille du journal doit être choisie avec soin car l'espace peut être saturé rapidement.

taille | unlimited

Nombre maximal d'octets consignés dans le journal du consultant. Vous pouvez indiquer un nombre positif supérieur à zéro, ou le mot **unlimited**. Il est possible que le fichier journal n'atteigne pas la taille maximale exacte avant le remplacement des entrées existantes, car la taille des entrées de journal elles-même varie.

Exemples

- Pour afficher un rapport concernant le contrôleur, entrez :

```
nalcontrol controller report
```

Cette commande génère des résultats similaires à l'exemple suivant :

Rapport du contrôleur :

```
-----
Version . . . . . Version: 05.00.00.00 - 03/21/2002-09:49:57-EST
Niveau consignation . . . 1
Taille journal . . . . . 1048576
```

Fichier configuration . . config1.xml

Consultants :

Consultant consult1 -démarré

- Pour affecter au niveau de consignation la valeur zéro afin d'optimiser les performances, entrez :

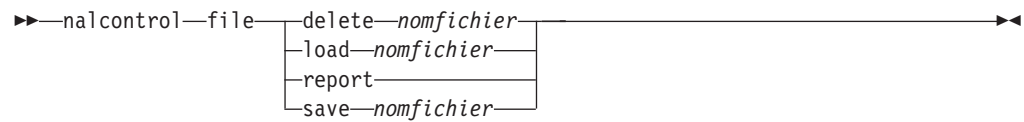
nalcontrol set loglevel 0

- Pour fixer la taille du journal du contrôleur à 1 000 000 octets, entrez :

nalcontrol controller set

logsize 1000000

nalcontrol file — Gestion des fichiers de configuration



delete

Supprimer le fichier de configuration indiqué.

nomfichier

Fichier de configuration. L'extension doit être .xml. Si cette extension n'est pas indiquée, elle est ajoutée par défaut.

load

Charge la configuration enregistrée dans le fichier indiqué.

Remarque : Le chargement d'un fichier ajoute la configuration stockée dans ce fichier à la configuration en cours. Si vous voulez charger une *nouvelle* configuration, vous devez arrêter puis redémarrer le serveur avant de charger le fichier de la nouvelle configuration.

report

Liste des fichiers de configuration.

save

Sauvegarde la configuration en cours dans le fichier indiqué.

Remarque : Les fichiers sont sauvegardés dans les répertoires suivants et chargés à partir de ces mêmes répertoires :

- Systèmes AIX : `/opt/ibm/edge/lb/servers/configurations/nal`
- Systèmes Linux : `/opt/ibm/edge/lb/servers/configurations/nal`
- Systèmes Solaris : `/opt/ibm/edge/lb/servers/configurations/nal`
- Systèmes Windows :

Chemin d'accès courant au répertoire d'installation —

`C:\Program Files\ibm\edge\lb\servers\configurations\nal`

Chemin d'accès au répertoire d'installation natif — `C:\Program Files\ibm\lb\servers\configurations\nal`

Exemples

- Pour supprimer un fichier nommé fichier1, entrez :
`nalcontrol file delete fichier1`
- Pour charger un nouveau fichier de configuration afin de remplacer la configuration actuelle, entrez :
`nalcontrol file load config2`
- Pour afficher un rapport des fichiers précédemment sauvegardés, entrez :
`nalcontrol file report`

Cette commande génère des résultats similaires à l'exemple suivant :

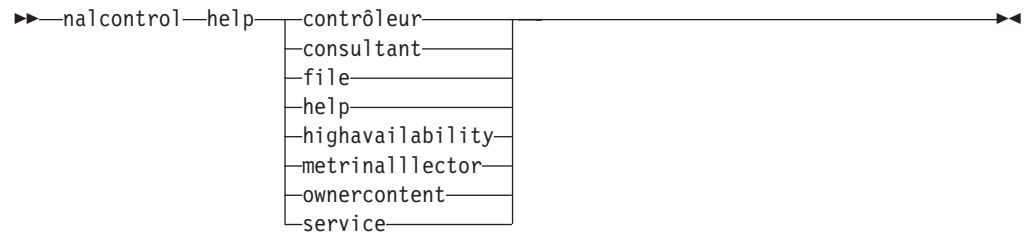
LISTE DES FICHIERS :

```
-----
fichier1.xml
fichier2.xml
fichier3.xml
```

- Pour sauvegarder votre configuration dans un fichier nommé config2, entrez :

```
nalcontrol file save config2
```

nalcontrol help — Affichage ou impression de l'aide relative à cette commande



Exemples

- Pour obtenir de l'aide sur la commande `nalcontrol`, entrez :

```
nalcontrol help
```

Cette commande génère des résultats similaires à l'exemple suivant :

Les commandes suivantes sont disponibles :

<code>controller</code>	- s'applique au contrôleur
<code>consultant</code>	- s'applique aux consultants du commutateur
<code>file</code>	- s'applique aux fichiers de configuration
<code>help</code>	- s'applique à l'aide
<code>highavailability</code>	- s'applique à la haute disponibilité
<code>metriccollector</code>	- s'applique aux programmes de collecte de mesures
<code>server</code>	- s'applique aux serveurs
<code>service</code>	- s'applique aux services

- Les symboles suivants sont utilisés dans la syntaxe de l'aide en ligne :

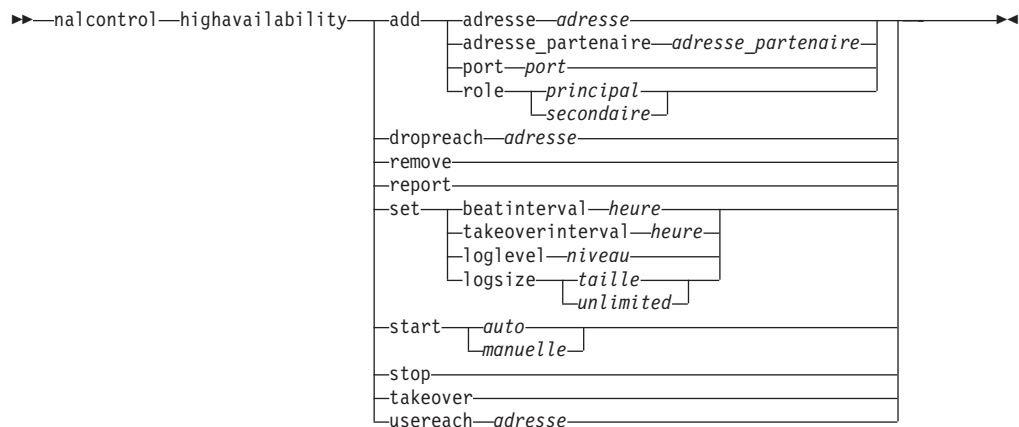
`< >` Paramètres entre accolades ou séquence de caractères.

`[]` Eléments facultatifs entre accolades.

`|` Barre verticale de séparation des choix possibles au sein des parenthèses et des accolades.

`:` Signe deux-points pour séparer les noms, par exemple, **consultant1:service1**.

nalcontrol highavailability — Contrôle de la haute disponibilité



add

Configure un noeud de haute disponibilité, un partenaire et des cibles à contacter.

address

Adresse pour la réception des signaux de présence.

adresse

Adresse IP du noeud à haute disponibilité.

partneraddress

Adresse pour l'envoi des signaux de présence. Il s'agit de l'adresse IP ou du nom d'hôte configuré sur le noeud partenaire. Cette adresse permet de communiquer avec la machine haute disponibilité partenaire.

adresse

Adresse IP du partenaire.

port

Port utilisé pour communiquer avec le partenaire. La valeur par défaut est 12345.

port

Numéro du port.

role

Rôle de haute disponibilité.

principal | secondaire

Rôles principal et secondaire de haute disponibilité.

dropreach

Supprime cette cible à contacter des critères de haute disponibilité.

adresse

Adresse IP de la cible à contacter.

remove

Supprime le noeud, le partenaire et la cible à contacter de la configuration de haute disponibilité. Vous devez arrêter la haute disponibilité avant de lancer cette commande.

report

Affiche les informations de haute disponibilité.

set

Définit les caractéristiques de la haute disponibilité.

beatinterval

Définit l'intervalle, en millisecondes, au bout duquel les signaux de présence sont envoyés au partenaire. La valeur par défaut est 500.

heure

Entier positif correspondant à l'écart, en millisecondes, entre chaque signal de présence.

takeoverinterval

Définit la durée, en millisecondes, qui doit s'écouler (pendant laquelle aucun signal de présence n'est reçu) avant qu'une reprise ne se produise. La valeur par défaut est 2000.

heure

Entier positif correspondant à l'écart, en millisecondes, entre chaque reprise.

loglevel

Définit le niveau auquel les activités sont consignées. La valeur par défaut est 1.

niveau

Niveau compris entre 0 et 5. La valeur par défaut est 1. Les valeurs admises sont les suivantes :

- 0 = Aucun
- 1 = Minimal
- 2 = De base
- 3 = Modéré
- 4 = Avancé
- 5 = Prolixe

logsize

Nombre maximal d'octets consignés dans le journal de haute disponibilité. La valeur par défaut est 1 048 576. Lorsque vous attribuez une taille maximale au fichier journal, ce dernier fonctionne en boucle. Lorsque le fichier atteint la taille indiquée, les entrées suivantes sont écrites à partir du haut du fichier et remplacent les entrées existantes. La valeur indiquée par logsize ne peut pas être inférieure à la taille actuelle du fichier journal. Les entrées du journal sont horodatées, ce qui permet de déterminer l'ordre dans lequel elles ont été enregistrées. Plus le niveau de consignation est élevé, plus la taille du journal doit être choisie avec soin car l'espace peut être saturé rapidement.

taille | unlimited

Nombre maximal d'octets consignés dans le journal de haute disponibilité. Vous pouvez indiquer un nombre positif supérieur à zéro, ou le mot **unlimited**. Il est possible que le fichier journal n'atteigne pas la taille maximale exacte avant le remplacement des entrées existantes, car la taille des entrées de journal elles-mêmes varie.

start

Active la haute disponibilité. Avant d'utiliser cette commande, vous devez configurer un noeud de haute disponibilité, un partenaire et une cible à contacter.

auto | manuelle

Détermine si la haute disponibilité démarre avec une stratégie de reprise automatique ou manuelle.

stop

Arrête la haute disponibilité.

takeover

Reprend le contrôle au noeud de haute disponibilité actif.

usereach

Adresse de la cible à contacter qui commencer à utiliser la haute disponibilité.
Ajoutez une cible à contacter à l'aide de ping de sorte que les partenaires de haute disponibilité puissent déterminer dans quelle mesure leurs cibles sont accessibles.

adresse

Adresse IP de la cible à contacter.

Exemples

- Pour ajouter le noeud de haute disponibilité d'adresse IP 9.37.50.17, avec le rôle principal sur le port 12345 et l'adresse de partenaire 9.37.50.14, entrez :

```
nalcontrol highavailability add
address 9.37.50.17 role principal port 12345 partneraddress 9.37.50.14
```

- Pour ajouter l'adresse de cible à contacter 9.37.50.9, entrez :

```
nalcontrol highavailability usereach
9.37.50.9
```

- Pour supprimer l'adresse de cible à contacter 9.37.50.9, entrez :

```
nalcontrol highavailability dropreach
9.37.50.9
```

- Pour démarrer la haute disponibilité démarre avec une stratégie de reprise manuelle, entrez :

```
nalcontrol
highavailability start manual
```

- Pour obtenir une analyse statistique de la haute disponibilité, entrez :

```
nalcontrol highavailability report
```

Cette commande génère des résultats similaires à l'exemple suivant :

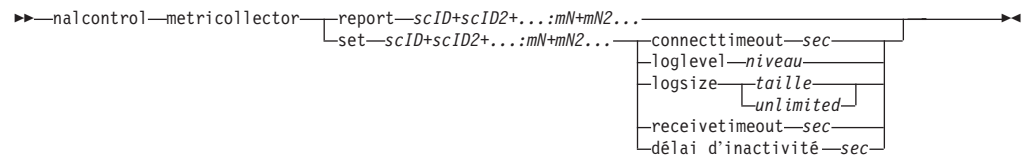
Etat de la haute disponibilité :

```
-----
Noeud . . . . . principal
Adresse du noeud . . . . 9.37.50.17
Port . . . . . 12345
Adresse du partenaire . . 9.37.50.14
Stratégie de reprise . . . manual
Int. entre sig. prés . . . 500
Int. entre reprises . . . 2000
Lancé . . . . . N
Etat . . . . . inactif
Sous-état . . . . . non synchronisé
```

Etat d'accessibilité : Noeud/Partenaire

```
-----
```

nalcontrol metriccollector — Configuration du programme de collecte de mesures



report

Affiche les caractéristiques du programme de collecte de mesures.

IDcc (ID du consultant du commutateur)

Chaîne définie par l'utilisateur, qui désigne le consultant.

Nm (nom de la mesure)

Nom qui identifie la mesure personnalisée ou fournie.

set

Définit les caractéristiques du programme de collecte de mesures.

connecttimeout

Délai d'attente observé par le programme de collecte de mesures avant de signaler l'échec d'une connexion.

sec Entier positif correspondant au nombre de secondes à l'expiration desquelles le programme de collecte de mesures signale qu'une connexion à un service a échoué.

loglevel

Définit le niveau auquel le consultant indiqué consigne les activités. Valeur par défaut : 1.

niveau

Numéro du niveau. La valeur par défaut est 1. Plus le numéro est élevé, plus la quantité d'informations consignée dans le journal du consultant est importante. Les valeurs admises sont les suivantes :

- 0 = Aucun
- 1 = Minimal
- 2 = De base
- 3 = Modéré
- 4 = Avancé
- 5 = Prolixe

logsize

Nombre maximal d'octets consignés dans le journal. La valeur par défaut est 1 048 576. Lorsque vous attribuez une taille maximale au fichier journal, ce dernier fonctionne en boucle. Lorsque le fichier atteint la taille indiquée, les entrées suivantes sont écrites à partir du haut du fichier et remplacent les entrées existantes. La valeur indiquée par logsize ne peut pas être inférieure à la taille actuelle du fichier journal. Les entrées du journal sont horodatées, ce qui permet de déterminer l'ordre dans lequel elles ont été enregistrées. Plus le niveau de consignation est élevé, plus la taille du journal doit être choisie avec soin car l'espace peut être saturé rapidement.

taille | unlimited

Nombre maximal d'octets consignés dans le journal du consultant. Vous pouvez indiquer un nombre positif supérieur à zéro, ou le mot **unlimited**. Il

est possible que le fichier journal n'atteigne pas la taille maximale exacte avant le remplacement des entrées existantes, car la taille des entrées de journal elles-même varie.

receivetimeout

Délai d'attente observé par un consultant avant de signaler l'échec d'une procédure de réception.

sec Il s'agit d'un entier positif qui représente le délai en secondes à l'expiration duquel le consultant signale que la réception d'un envoi provenant d'un service a échoué.

sleeptime

Nombre de secondes d'inactivité du programme de collecte de mesures entre chaque cycle de collecte de mesures.

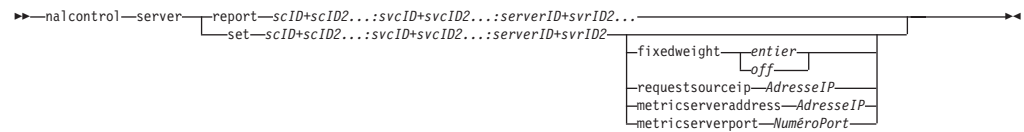
sec Entier positif correspondant au nombre de secondes d'inactivité.

Exemples

- Pour afficher un rapport concernant les caractéristiques d'un programme de collecte de mesures, entrez :
`nalcontrol metrinalllector report cc1:http`
Cette commande génère des résultats similaires à l'exemple suivant :

```
Metrinalllector cc1:http
measure(s) collectée(s).... http
niveau consignation..... 5
taille journal..... 1048576
Nbr. sec. inactivité..... 7
délai expiration conn..... 21
délai expiration récep.... 21
```
- Pour définir un délai d'expiration de connexion de 15 secondes et une taille de journal illimitée pour le consultant de commutateur cc1 et la mesure http, entrez :
`nalcontrol metrinalllector set cc1:http connecttimeout 15 logsize unlimited`

nalcontrol server — Configuration d'un serveur



report

Affiche les caractéristiques des serveurs.

IDcc

Chaîne, définie par l'utilisateur, qui représente le consultant.

IDsvc

Chaîne définie par l'utilisateur qui représente l'identificateur du service virtuel et le numéro du port virtuel sur le commutateur.

IDserveur

Entier qui représente le serveur sur le commutateur.

set

Définit les caractéristiques des serveurs.

fixedweight

Affecte une pondération fixe au serveur. La valeur par défaut est off. La valeur de pondération fixe maximale est 48.

entier | off

Entier positif correspondant à la pondération fixe du serveur ou le mot **off** si aucune pondération fixe ne doit être définie.

requestsourceip

Définit l'adresse à partir de laquelle contacter le serveur pour les demandes d'application.

AdresseIP

Adresse IP à partir de laquelle le serveur doit être contacté, sous forme d'adresse symbolique ou d'adresse IP.

metricserveraddress

Définit l'adresse à partir de laquelle contacter le serveur pour les demandes du serveur de mesures.

AdresseIP

Adresse IP du serveur de mesures, sous forme d'adresse symbolique ou d'adresse IP.

metricserverport

Définit le port à utiliser pour contacter le serveur de mesures.

NuméroPort

Numéro du port utilisé pour contacter le serveur de mesures.

Exemples

- Pour afficher un rapport concernant le serveur 1 du consultant cc1, entrez :
nalcontrol server report
cc1:svc1:1
Cette commande génère des résultats similaires à l'exemple suivant :

```

Le serveur ccl:svcl:1 a la
pondération -99
  La pondération fixe est off
  Ip source de la demande..... 9.27.24.156
  Port de l'application..... 99
  Adresse du serveur de mesures..... 9.99.99.98
  Port du serveur de mesures..... 10004
    La mesure activeconn a la valeur -99
    La mesure connrate a la valeur -99

```

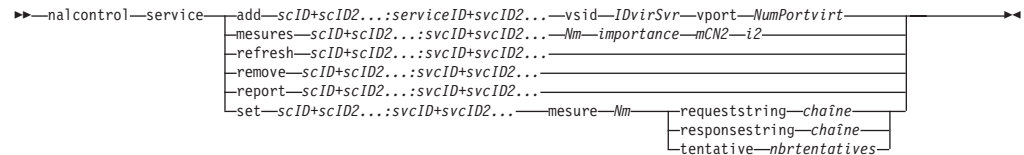
- Pour définir une adresse de serveur de mesures pour le service 2, entrez :

```

nalcontrol server set ccl:svcl:2
metricserveraddress 9.37.50.17

```

nalcontrol service — Configuration d'un service



add

Ajoute un service au consultant désigné.

IDcc (ID du consultant du commutateur)

Chaîne définie par l'utilisateur, qui désigne le consultant.

IDSvc (ID du service)

Chaîne définie par l'utilisateur qui identifie le service.

vsid

Mot clé d'identification du service virtuel.

IDSrvr (ID du serveur virtuel)

Numéro sur le commutateur qui représente le serveur virtuel.

vport

Mot clé d'identification du port virtuel.

NumPortvirt (numéro du port virtuel)

Numéro de port du service actuellement configuré sur le commutateur.

metrics

Désigne l'ensemble des mesures utilisées pour calculer les pondérations et l'importance de chaque mesure. L'importance est exprimée en pourcentage du total. Par conséquent, la somme des valeurs d'importance doit toujours être égale à 100. Les mesures peuvent être toute combinaison de mesures de données de connexion, mesures de conseiller d'application et mesures de serveur de mesures. Par défaut il s'agit du nombre de connexions actives (activeconn) et du débit de la connexion (connrate) avec une importance équivalente de 50/50.

Nm (nom de la mesure)

Nom qui identifie le programme chargé de collecter les mesures afin de déterminer la pondération du serveur.

Voici la liste des noms de mesure admis et des ports qui leur sont associés.

Nom du conseiller	Protocole	Port
connect	ICMP	12345
db2	privé	50000
dns	DNS	53
ftp	FTP	21
http	HTTP	80
https	SSL	443
cachingproxy	HTTP (via Caching Proxy)	80
imap	IMAP	143
ldap	LDAP	389
nntp	NNTP	119

Nom du conseiller	Protocole	Port
ping	PING	0
pop3	POP3	110
sip	SIP	5060
smtp	SMTP	25
ssl	SSL	443
telnet	Telnet	23
WLM	privé	10,007
activeconn	Non disp.	Non disp.
connrate	Non disp.	Non disp.
cpuload	Non disp.	Non disp.
memload	Non disp.	Non disp.

importance

Nombre compris entre 0 et 100 qui correspond à l'importance de la mesure dans le calcul des pondérations du serveur.

refresh

Actualise un service avec des informations provenant de Nortel Alteon Web Switch.

remove

Supprime un service.

report

Répertorie les caractéristiques d'un service.

set

Définit les caractéristiques d'un service.

mesure

Définit les caractéristique d'une mesure configurée.

Nm (nom de la mesure)

Nom de la mesure souhaitée.

requeststring

Définit une chaîne de demande pour la mesure désignée. Cette chaîne correspond à la demande envoyée par un programme de collecte de mesures pour rassembler des informations de mesure.

chaîne

Chaîne de demande envoyée au serveur par le programme de collecte de mesures.

responsestring

Définit une chaîne de réponse pour la mesure désignée. Le programme de collecte de mesures utilise cette chaîne pour comparer les réponses qu'il reçoit des serveurs et déterminer ainsi la disponibilité du serveur.

chaîne

Chaîne de réponse à laquelle le programme de collecte de mesures compare les réponses reçues du serveur.

tentative

Nombre de tentatives accordées avant qu'un serveur ne soit déclaré arrêté.

nbrtentatives

Entier supérieur ou égal à zéro. Il est préférable que le nombre de tentatives ne dépasse pas 3. Par défaut, le nombre de tentatives est égal à zéro.

Exemples

- Pour ajouter le service svc1 (d’ID de serveur virtuel 1 et de port virtuel 80) à l’ID de consultant de commutateur cc1, entrez :
nalcontrol service add cc1:svc1 vsid 1 vport 80
- Pour appliquer une proportion de 50 à chacune des mesures activeconn et http, entrez :
nalcontrol service metrics cc1:svc1 activeconn 50 http 50
- Pour afficher le rapport des caractéristiques des contenus de propriétaire, entrez :

```
nalcontrol service  
report cc1:svc1
```

Cette commande génère des résultats similaire à l’exemple suivant :

```
Service cc1:svc1  
  Limite de pondération = 48  
  La mesure activeconn a la proportion 50  
  La mesure connrate a la proportion 50  
  Contient le serveur 4  
  Contient le serveur 3  
  Contient le serveur 2  
  Contient le serveur 1
```

- Pour définir une chaîne de demande http, entrez :
nalcontrol service set cc1:svc1 metric http requeststring
getLastErrorCode

Annexe A. Interface graphique utilisateur : Instructions générales

Dans l'interface graphique de Load Balancer, la partie gauche du panneau affiche une arborescence comportant Load Balancer au niveau supérieur et Dispatcher, Content Based Routing (CBR), Site Selector, Cisco CSS Controller et Nortel Alteon Controller en tant que composants.

Si vous utilisez l'installation Load Balancer pour IPv4 et IPv6, seul le composant Dispatcher est disponible. Pour plus d'informations, voir Chapitre 8, «Déploiement de Dispatcher sur Load Balancer pour IPv4 et IPv6», à la page 81.

Les exemples d'affichage de l'interface graphique de Load Balancer illustrant les différents composants renvoient aux figures suivantes :

- La figure 41, à la page 470 pour Dispatcher
- La figure 42, à la page 471 pour CBR
- La figure 43, à la page 472 pour Site Selector
- La figure 44, à la page 473 pour Cisco CSS Controller
- La figure 45, à la page 474 pour Nortel Alteon Controller

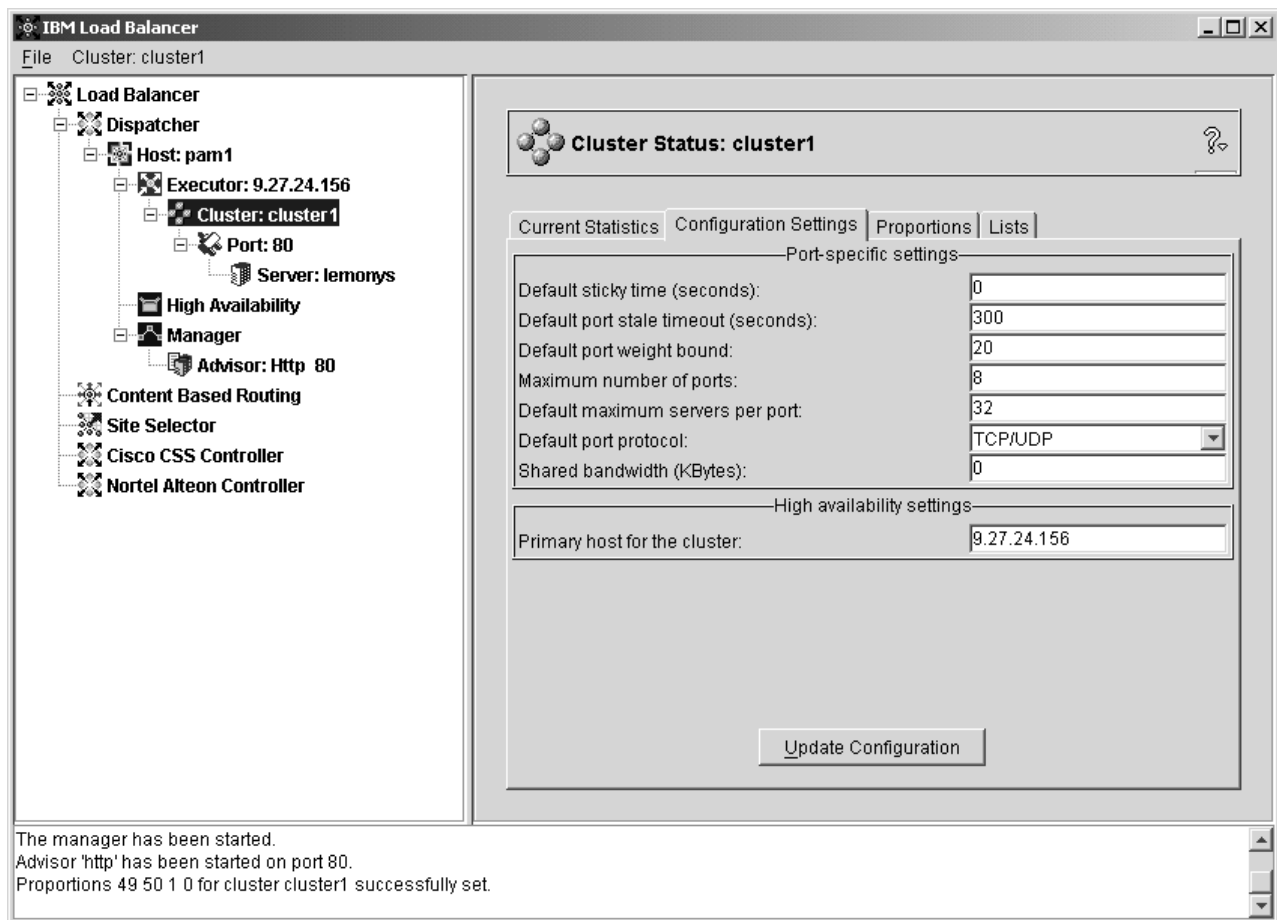


Figure 41. Interface graphique affichant l'arborescence du composant Dispatcher

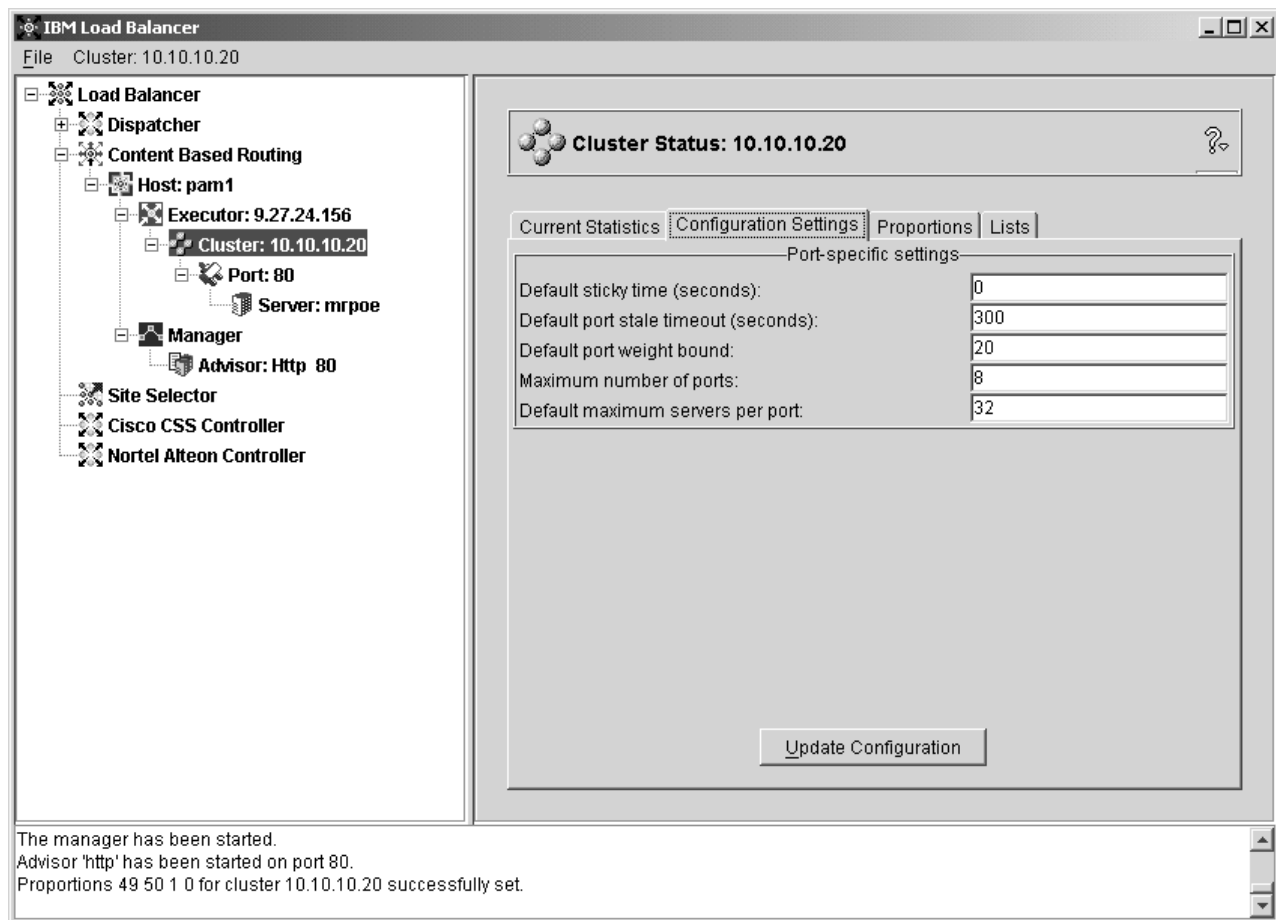


Figure 42. Interface graphique affichant l'arborescence du composant CBR

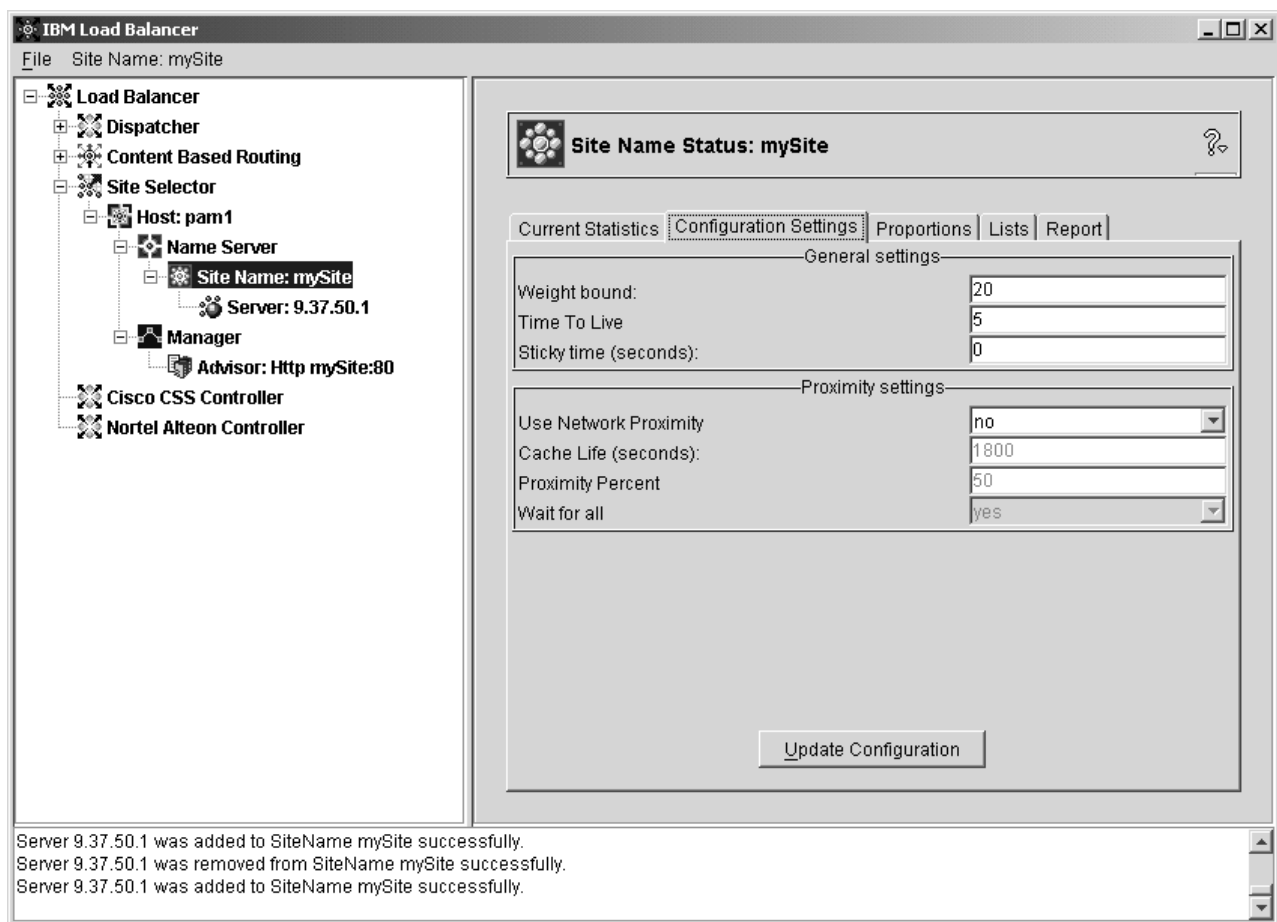


Figure 43. Interface graphique affichant l'arborescence du composant Site Selector

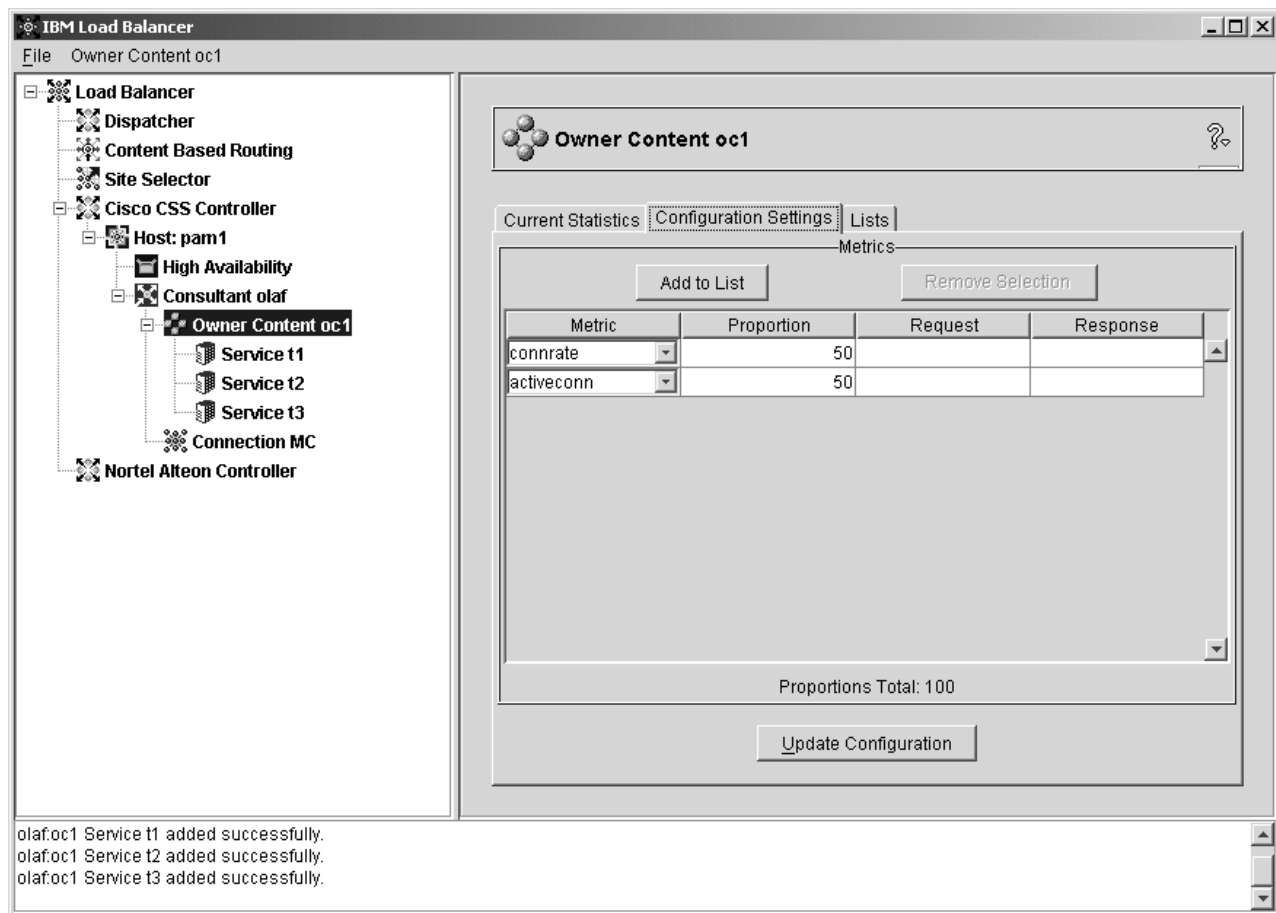


Figure 44. Interface graphique affichant l'arborescence du composant Cisco CSS Controller

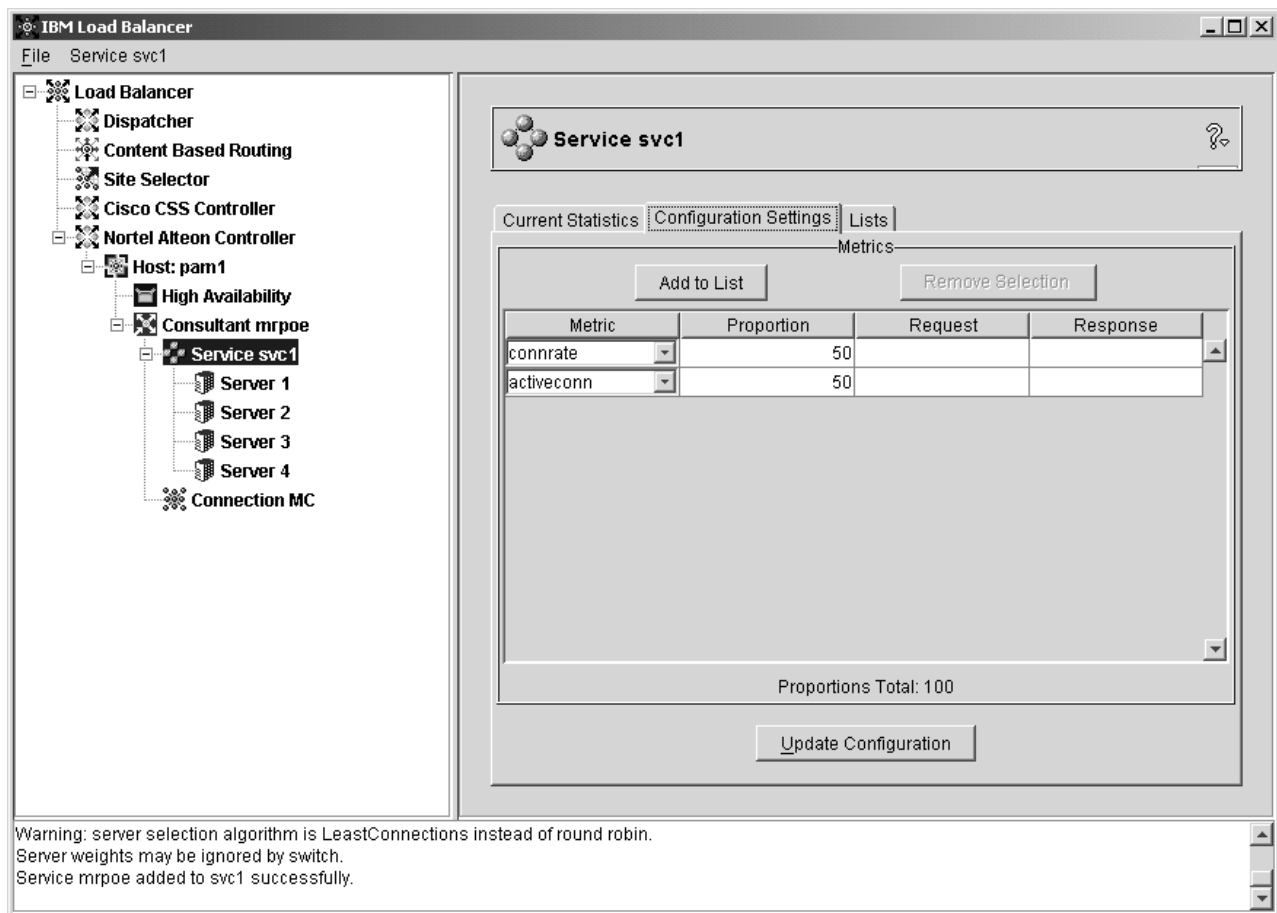


Figure 45. Interface graphique affichant l'arborescence du composant Nortel Alteon Controller

Tous les composants peuvent être configurés à partir de l'interface graphique. Pour sélectionner des éléments dans l'arborescence, cliquez à l'aide du bouton un de la souris (normalement le bouton gauche) et pour afficher les menus en incrustation, cliquez à l'aide du bouton deux (normalement le bouton droit). Les menus en incrustation des éléments de l'arborescence sont également disponibles à partir de la barre de menus située dans la partie supérieure de la fenêtre.

Cliquez sur les signes plus ou moins pour développer ou réduire les éléments de l'arborescence.

Pour exécuter une commande à partir de l'interface graphique : mettez le noeud Hôte en surbrillance dans l'arborescence de l'interface graphique, puis sélectionnez **Envoyer la commande...** dans le menu en incrustation Hôte. Dans la zone d'entrée de commande, entrez la commande à exécuter, par exemple **executor report**. Les résultats et l'historique des commandes exécutées lors de la session courante s'affichent dans la fenêtre ouverte.

La partie droite de la fenêtre contient deux listes d'indicateurs d'état relatifs à l'élément sélectionné.

- L'onglet **Statistiques actuelles** fournit des informations statistiques sur l'élément. Cet onglet n'apparaît pas pour tous les éléments de l'arborescence.
- Le bouton **Régénération des statistiques** permet d'afficher les dernières données statistiques. Si aucun bouton Régénération des statistiques n'apparaît, les statistiques sont rafraîchies dynamiquement et sont toujours actuelles.

- L'onglet **Paramétrages configuration** présente les paramètres de configuration qui peuvent être définis en utilisant les procédures décrites dans les chapitres de configuration pour chacun des composants. Cet onglet n'apparaît pas pour tous les éléments de l'arborescence.
- Le bouton **MAJ configuration** applique les dernières modifications à la configuration en cours.
- L'onglet **Proportions** présente les paramètres de proportion (ou de pondération) pouvant être définis à l'aide des informations du Chapitre 22, «Fonctions avancées de Dispatcher, CBR et Site Selector», à la page 201. Cet onglet n'apparaît pas pour tous les éléments de l'arborescence.
- L'onglet **Liste** présente des détails supplémentaires sur l'arborescence sélectionnée. Cet onglet n'apparaît pas pour tous les éléments de l'arborescence.
- Le bouton **Retrait** supprime les éléments mis en évidence dans les listes.
- L'onglet **Etat** présente les informations d'état du gestionnaire concernant l'élément. Cet onglet n'apparaît pas pour tous les éléments de l'arborescence.
- Le bouton **Régénération de l'état** permet d'afficher les dernières données d'état du gestionnaire.

Pour accéder à l'**aide**, cliquez sur le point d'interrogation (?) situé dans l'angle supérieur droit de la fenêtre Load Balancer.

- **Aide sur les zones** — décrit les valeurs par défaut de chaque zone.
- **Procédures** — affiche la liste des tâches pouvant être effectuées dans cet écran.
- **Centre de documentation** — permet d'accéder à des informations sur le produit, par exemple : présentation et mise en évidence des nouvelles fonctions, lien au site Web du produit, index des fichiers d'aide en ligne, glossaire des termes

Annexe B. Syntaxe des règles de contenu (modèle)

La présente annexe décrit comment utiliser la syntaxe de règle de contenu (modèle) pour le composant CBR et la méthode d'acheminement CBR de Dispatcher. Elle contient en outre des scénarios et des exemples de syntaxe.

Syntaxe de règle de contenu (modèle) :

Ne s'applique que si vous avez sélectionné le type de règle Contenu.

Entrez la syntaxe voulue en tenant compte des restrictions suivantes :

- Le motif ne doit pas contenir d'espace.
- Caractères tenant lieu de caractères spéciaux lorsqu'ils ne sont pas précédés d'une barre oblique inverse (\) :
 - * caractère générique (correspond à n'importe quel caractère)
 - (parenthèse gauche utilisée pour un regroupement logique
 -) parenthèse droite utilisée pour un regroupement logique
 - & ET logique
 - | OU logique
 - ! NON logique

Mots clés réservés

Les mots clés réservés sont toujours suivis d'un signe égal «=».

Méthode

Méthode HTTP de la demande, par exemple GET, POST, etc.

URI Chemin de la demande d'URL (respect des majuscules et minuscules)

Version

Version spécifique de la demande, HTTP/1.0 ou HTTP/1.1

Hôte valeur de l'hôte : en-tête(non respect des majuscules et minuscules)

Remarque : Facultatif dans les protocoles HTTP/1.0

<clé> Tout nom d'en-tête HTTP valide pouvant être recherché par HTTP. Par exemple, User-Agent, Connection, Referer, etc.

Un navigateur demandant la page `http://www.entreprise.com/path/webpage.htm` peut obtenir des résultats du type suivant :

```
Method=GET
URI=/path/webpage.htm
Version=HTTP/1.1
Host=www.company.com
Connection=Keep-Alive
Referer=http://www.company.com/path/parentwebpage.htm
```

Remarque : Le shell du système d'exploitation peut interpréter les caractères spéciaux tels que la perluète ("&") et les convertir en texte de remplacement avant leur évaluation avec **cbrcontrol**.

Par exemple, la commande suivante n'est valide qu'avec l'invite **cbrcontrol>>**.

```
rule add 10.1.203.4:80:cbr_prod_rule_ek type content
  pattern uri=/nipoek/*
```

Pour que cette commande fonctionne à partir de l'invite du système d'exploitation (ou dans le fichier de configuration) lorsque vous utilisez des caractères spéciaux, placez le motif entre guillemets (" ") comme suit :

```
cbrcontrol rule add 10.1.203.4:80:cbr_prod_rule_ek type content
  pattern "uri=/nipoek/*"
```

Si vous omettez les guillemets, le motif sera peut-être tronqué lors de la sauvegarde de la règle dans CBR. Les guillemets ne sont pas pris en charge avec l'invite cbrcontrol>>.

Ci-dessous, se trouve un ensemble de scénarios et des exemples d'utilisation des syntaxes de modèle

Scénario 1 :

La configuration d'un cluster implique un ensemble de serveurs Web pour le contenu HTML standard, un autre ensemble de serveurs Web avec WebSphere Application Server pour les demandes de servlet, un autre ensemble de serveurs Lotus Notes pour les fichiers NSF, etc. Pour effectuer la différence entre les pages demandées, l'accès aux données du client est requis. Il est également nécessaire de les envoyer aux serveurs appropriés. Les règles de correspondance de modèle de contenu permettent d'effectuer une séparation, nécessaire à l'accomplissement de ces tâches. Plusieurs règles sont configurées afin que la séparation des demandes nécessaire se produise automatiquement. Par exemple, les commandes suivantes provoquent les trois divisions indiquées :

```
>>rule add cluster1:80:servlets type content pattern uri=*/servlet/*priority 1
>>rule uses cluster1:80:servlets server1+server2

>>rule add cluster1:80:notes type content pattern uri=*.nsf* priority 2
>>rule uses cluster1:80:notes server3+server4

>>rule add cluster1:80:regular type true priority 3
>>rule uses cluster1:80:regular server5+server6
```

Si une demande de fichier NSF est reçue par Load Balancer, la règle servlets est vérifiée mais ne correspond pas. La demande est ensuite vérifiée par la règle notes qui trouve une correspondance. L'équilibrage de charge du client est effectué entre le serveur 3 et le serveur 4.

Scénario 2

Le contrôle par le site Web principal de plusieurs groupes internes constitue un autre scénario classique. Par exemple, l'ensemble de serveurs et le contenu de www.company.com/software sont différents de ceux de www.company.com/hardware. Etant donné que les demandes dépendent toutes du cluster www.company.com racine, les règles de contenu sont requises pour trouver les différences d'URI et pour effectuer l'équilibrage de charge. La règle du scénario peut être du type suivant :

```
>>rule add cluster1:80:div1 type content pattern uri=/software/* priority 1
>>rule uses cluster1:80:div1 server1+server2

>>rule add cluster1:80:div2 type content pattern uri=/hardware/* priority 2
>>rule uses cluster1:80:div2 server3+server4
```

Scénario 3

Pour certaines associations, l'ordre de recherche des règles est important. Par exemple, dans le scénario 2, les clients ont été séparés en fonction d'un répertoire dans leur chemin de demande. Cependant, le répertoire cible peut apparaître à plusieurs niveaux du chemin et la signification est différente en fonction de l'emplacement. Par exemple, la cible de `www.company.com/pcs/fixed/software` est différente de celle de `www.company.com/mainframe/fixed/software`. Les règles doivent être définies pour prendre en compte cette possibilité et ne doivent pas inclure trop de scénarios en même temps. Par exemple, la recherche générique du test «`uri=*/software/*` » est trop large. D'autres règles peuvent être structurées de la manière suivante :

Vous pouvez associer plusieurs recherches afin de restreindre la recherche :

```
>>rule add cluster1:80:pcs type content pattern (uri=/pcs/*)&(uri=*/software/*)
>>rule uses cluster 1:80:pcs server1
```

Dans le cas où il n'est pas possible d'utiliser des combinaisons, l'ordre est important :

```
>>rule add cluster1:80:pc1 type content pattern uri=/pcs/*
>>rule uses cluster1:80:pc1 server2
```

La deuxième règle permet de détecter l'apparition de «`pcs`» dans des répertoires ultérieurs du chemin et non le premier.

```
>>rule add cluster1:80:pc2 type content pattern uri=*/pcs/*
>>rule uses cluster1:80:pc2 server3
```

Dans la plupart des cas, vous pouvez compléter les règles par une règle par défaut **toujours vrai** afin de détecter tous les éléments non pris en compte par les autres règles. Un serveur peut également renvoyer un message du type «Ce site est actuellement indisponible, faites une nouvelle tentative ultérieurement» dans les scénarios pour lesquels aucun autre serveur ne peut renvoyer de réponse pour ce client.

```
>>rule add cluster1:80:sorry type true priority 100
>>rule uses cluster1:80:sorry server5
```

Annexe C. Exemples de fichiers de configuration

La présente annexe contient des exemples de fichiers de configuration pour le composant Dispatcher de Load Balancer.

IMPORTANT : Si vous utilisez l'installation Load Balancer pour IPv4 et IPv6, n'oubliez pas de remplacer le signe deux-points (:) par le signe at (@), comme délimiteur des commandes dscontrol, dans ces exemples de fichiers de configuration.

Exemples de fichiers de configuration Load Balancer

Les exemples de fichiers se trouvent dans le répertoire ...ibm/edge/lb/servers/samples/.

Dispatcher Fichier de configuration — systèmes AIX, Linux et Solaris

```
#!/bin/bash
#
# configuration.sample - Exemple de fichier de configuration pour
# composant Dispatcher
#
#
# Ce script doit être lancé par le superutilisateur.
#
# iam=`whoami`

# if [ "$iam" != "root" ]if [ "$iam" != "root" ]
# then
# echo "Vous devez vous connecter en tant que superutilisateur
# pour exécuter ce script"
# exit 2
# fi

#
# Démarrez d'abord le serveur
#
# dsserver start
# sleep 5

#
# Démarrez ensuite l'exécuteur
#
# dscontrol executor start

#
# Il est possible d'arrêter le répartiteur à tout moment à l'aide
# des commandes "dscontrol executor stop" et "dsserver stop"
# pour arrêter respectivement l'exécuteur et le serveur avant
# d'arrêter le logiciel Dispatcher.
#
# L'étape suivante dans la configuration du répartiteur est de définir
# l'adresse NFA (adresse de non-réacheminement) et les adresses de clusters.
#
# L'adresse NFA permet d'accéder à distance au répartiteur
# afin d'effectuer des opérations d'administration et de configuration.
# Cette adresse est obligatoire car le répartiteur doit acheminer
# des paquets vers les adresses de clusters.
#
```

```

# L'adresse cluster correspond au nom d'hôte (ou à l'adresse IP)
# auquel les clients éloignés se connecteront.
#
# Vous pouvez indifféremment utiliser les noms d'hôte et les adresses IP
# dans ce fichier.
#

# NFA=nomhôte.domaine.nom
# CLUSTER=www.votresociété.com

# echo "Chargement de l'adresse de non réacheminement"
# dscontrol executor set nfa $NFA

#
# L'étape suivante dans la configuration du répartiteur consiste à créer
# un cluster. Le répartiteur acheminera les requêtes envoyées
# à l'adresse de cluster vers les serveurs associés
# à ce cluster. Vous pouvez configurer plusieurs adresses de
# clusters et leur associer plusieurs serveurs à l'aide du répartiteur.

# Utilisez une configuration similaire pour CLUSTER2, CLUSTER3, etc.
#

# echo "Chargement de la première adresse de cluster"
# dscontrol cluster add $CLUSTER

#
# L'étape suivante consiste à définir les ports utilisés par ce cluster.
# Toute requête reçue par le répartiteur sur un port défini
# sera réacheminée vers le port correspondant de l'un
# des serveurs.
#

# echo "Création de ports pour le cluster : $CLUSTER"

# dscontrol port add $CLUSTER:20+21+80

#
# La dernière étape consiste à associer chaque serveur
# aux ports de ce cluster.
# Vous pouvez utiliser indifféremment le nom
# d'hôte ou l'adresse IP des serveurs.
#

# SERVER1=server1name.domain.name
# SERVER1=nomserveur1.domaine.nom
# SERVER2=nomserveur2.domaine.nom
# SERVER3=nomserveur3.domaine.nom

# echo "Ajout des serveurs"
# dscontrol server add $CLUSTER:20+21+80:
# $SERVER1+$SERVER2+$SERVER3

#
# Nous allons maintenant lancer les composants d'équilibrage de charge
# du répartiteur. Le premier composant s'appelle le gestionnaire.
# Les autres composants d'équilibrage de charge sont les
# conseillers. Si le gestionnaire et les conseillers ne fonctionnent pas,
# le répartiteur envoie des requêtes au format de permutation circulaire
# (round-robin). Une fois le gestionnaire lancé, les décisions de pondération
# basées sur le nombre de connexions nouvelles et actives sont utilisées, et
# les requêtes entrantes sont envoyées au meilleur serveur. Les conseillers
# fournissent au gestionnaire des informations supplémentaires sur la capacité
# du serveur à répondre aux requêtes, et à détecter si le serveur est actif
# ou non. Si un conseiller détecte qu'un serveur est arrêté, cela sera
# consigné (à condition que les proportions du gestionnaire soient définies
# pour inclure les entrées de conseiller) et aucune autre requête ne sera

```



```

# acheminée vers le serveur.

# La dernière étape de configuration des composants d'équilibrage de charge
# est la définition des proportions du gestionnaire. Ce dernier met à
# jour la pondération de chaque serveur en fonction de quatre règles :
# 1. Nombre de connexions actives sur chaque serveur.
# 2. Nombre de nouvelles connexions sur chaque serveur.
# 3. Informations fournies par les conseillers.
# 4. Informations fournies par le conseiller au niveau système.
# La somme de ces proportions doit être égale à 100. Par exemple,
# si l'on définit les proportions du gestionnaire de la façon suivante :
# dscontrol manager proportions 48 48 0 0
# 48 % des informations proviendront des connexions nouvelles, 48%,
# des connexions actives, 4%, des conseillers et les entrées
# système ne seront pas prises en compte. #
#
# REMARQUE : par défaut, les proportions du gestionnaire sont définies à 50 50 0 0.
#

# echo "Démarrage du gestionnaire (manager)..."
# dscontrol manager start

# echo "Démarrage du conseiller (advisor) FTP sur le port 21..."
# dscontrol advisor start ftp 21
# echo "Démarrage du conseiller (advisor) HTTP sur le port 80..."
# dscontrol advisor start http 80
# echo "Démarrage du conseiller (advisor) Telnet sur le port 23..."
# dscontrol advisor start telnet 23
# echo "Démarrage du conseiller (advisor) SMTP sur le port 25..."
# dscontrol advisor start smtp 25
# echo "Démarrage du conseiller (advisor) POP3 sur le port 110..."
# dscontrol advisor start pop3 110
# echo "Démarrage du conseiller (advisor) NNTP sur le port 119..."
# dscontrol advisor start nntp 119
# echo "Démarrage du conseiller (advisor) SSL sur le port 443..."
# dscontrol advisor start ssl 443
#

# echo "Définition des proportions du gestionnaire..."
# dscontrol manager proportions 58 40 2 0

#
# L'étape finale dans la configuration du répartiteur est d'affecter
# un alias à la Carte d'interface réseau (NIC).
#
# REMARQUE : N'utilisez pas cette commande dans un environnement
# haute disponibilité. Les scripts go* configureront les cartes NIC et
# le bouclage si nécessaire.
# dscontrol executor configure $CLUSTER

# Si votre adresse de cluster se trouve sur une autre carte NIC
# ou sur un sous-réseau autre que ceux de la NFA, utilisez le format
# suivant comme commande de configuration de cluster.
# dscontrol executor configure $CLUSTER tr0 0xfffff800
# tr0 étant votre carte NIC (tr1, la seconde carte réseau en anneau à jeton,
# en0 la première carte ethernet) et 0xfffff800 étant un masque
# de sous-réseau valide pour votre site.
#

#
# Les commandes suivantes permettent de définir les valeurs par défaut.
# Utilisez ces commandes pour modifier les valeurs par défaut.
# dscontrol manager loglevel 1
# dscontrol manager logsize 1048576
# dscontrol manager sensitivity 5
# dscontrol manager interval 2
# dscontrol manager refresh 2

```

```
#
# dscontrol advisor interval ftp 21 5
# dscontrol advisor loglevel ftp 21 1
# dscontrol advisor logsize ftp 21 1048576
# dscontrol advisor timeout ftp 21 unlimited
# dscontrol advisor interval telnet 23 5
# dscontrol advisor loglevel telnet 23 1
# dscontrol advisor logsize telnet 23 1048576
# dscontrol advisor timeout telnet 23 unlimited
# dscontrol advisor interval smtp 25 5
# dscontrol advisor loglevel smtp 25 1
# dscontrol advisor logsize smtp 25 1048576
# dscontrol advisor timeout smtp 25 unlimited
# dscontrol advisor interval http 80 5
# dscontrol advisor loglevel http 80 1
# dscontrol advisor logsize http 80 1048576
# dscontrol advisor timeout http 80 unlimited
# dscontrol advisor interval pop3 110 5
# dscontrol advisor loglevel pop3 110 1
# dscontrol advisor logsize pop3 110 1048576
# dscontrol advisor timeout pop3 110 unlimited
# dscontrol advisor interval nntp 119 5
# dscontrol advisor loglevel nntp 119 1
# dscontrol advisor logsize nntp 119 1048576
# dscontrol advisor timeout nntp 119 unlimited
# dscontrol advisor interval ssl 443 5
# dscontrol advisor loglevel ssl 443 1
# dscontrol advisor logsize ssl 443 1048576
# dscontrol advisor timeout ssl 443 unlimited
#
```

Dispatcher Fichier de configuration — systèmes Windows

Voici un exemple de fichier de configuration de Load Balancer intitulé **configuration.cmd.sample** à utiliser sous Windows.

```
@echo off
rem configuration.cmd.sample - Exemple de fichier de configuration pour
rem le composant Dispatcher.
rem

rem Démarrez dsserver à partir du panneau Services

rem

rem
rem Démarrez ensuite l'exécuteur
rem
rem call dscontrol executor start

rem

rem L'étape suivante dans la configuration du répartiteur est de définir
rem l'adresse NFA (adresse de non-réacheminement) et les
rem adresses de clusters.
rem

rem L'adresse NFA permet d'accéder à distance au répartiteur
rem afin d'effectuer des opérations d'administration de configuration.
rem Cette adresse est obligatoire car le répartiteur doit réacheminer
rem des paquets vers les adresses de clusters.

rem
rem L'adresse CLUSTER est le nom d'hôte (ou l'adresse IP)
rem à laquelle les clients éloignés se connecteront.
rem
```

```

rem Vous pouvez indifféremment utiliser les noms d'hôte et les adresses IP
rem dans ce fichier.
rem NFA=[adresse de non-réacheminement]
rem CLUSTER=[nom du cluster]
rem

rem set NFA=nom_hôte.domaine.nom
rem set CLUSTER=www.votresociété.com

rem echo "Chargement de l'adresse de non réacheminement"
rem call dscontrol executor set nfa %NFA%

rem
rem Les valeurs par défaut sont affectées aux commandes suivantes.
rem Utilisez ces commandes pour modifier les valeurs par défaut.

rem call dscontrol executor set fintimeout 30
rem
rem L'étape suivante dans la configuration du répartiteur consiste à
rem créer
rem un cluster. Le répartiteur acheminera les requêtes envoyées
rem à l'adresse de cluster vers les serveurs associés
rem à ce cluster. Vous pouvez configurer plusieurs adresses de
rem clusters à l'aide du répartiteur.
rem Utilisez une configuration similaire pour CLUSTER2, CLUSTER3, etc.
rem

rem echo "Chargement de la première adresse de cluster"
rem call dscontrol cluster add %CLUSTER%

rem
rem L'étape suivante consiste à définir les ports utilisés par ce cluster.
rem Toute requête reçue par le répartiteur sur un port défini
rem sera réacheminée vers le port correspondant
rem de l'un des serveurs.
rem

rem echo "Création des ports de CLUSTER : %CLUSTER%"
rem call dscontrol port add %CLUSTER%:20+21+80

rem
rem La dernière étape consiste à associer chaque serveur aux
rem ports définis pour le cluster. Vous pouvez utiliser indifféremment
rem le nom d'hôte ou l'adresse IP des machines serveurs.
rem

rem set SERVER1=nomserveur1.domaine.nom
rem set SERVER2=nomserveur2.domaine.nom
rem set SERVER3=nomserveur3.domaine.nom

rem echo "Ajout des serveurs"
rem call dscontrol server add %CLUSTER%:20+21+80:
rem %SERVER1%+%SERVER2%+%SERVER3%

rem
rem Nous allons maintenant lancer les composants d'équilibrage de charge
rem du répartiteur. Le premier composant s'appelle le gestionnaire.
rem Les autres composants d'équilibrage de charge sont les
rem conseillers. Si le gestionnaire et les conseillers ne sont pas
rem actifs, le répartiteur envoie des requêtes au format de permutation circulaire
rem (round-robin). Une fois le gestionnaire lancé, les décisions de
rem pondération basées sur le nombre de connexions nouvelles et actives
rem sont utilisées et les requêtes entrantes sont envoyées au meilleur
rem serveur. Les conseillers permettent au gestionnaire de disposer d'informations
rem supplémentaires sur la capacité du serveur à répondre aux requêtes
rem et à détecter si un serveur est actif ou non. Si un conseiller
rem détecte qu'un serveur est arrêté, cela sera consigné (à condition que

```

```

rem les proportions du gestionnaire aient été définies pour inclure les
rem entrées du conseiller) et aucune requête ne sera acheminée vers le serveur.
rem La dernière étape de configuration des composants d'équilibrage de charge
rem est la définition des proportions du gestionnaire. Ce dernier met à
rem jour la pondération manager de chaque serveur sur
la base de quatre règles :

```

```

rem 1. Nombre de connexions actives sur chaque serveur.
rem 2. Nombre de nouvelles connexions pour chaque serveur.
rem 3. Informations fournies par les conseillers.
rem 4. Informations fournies par le conseiller au niveau système.
rem
rem La somme de ces proportions doit être égale à 100. Par exemple,
rem si l'on définit les proportions avec
rem dscontrol cluster set <cluster> proportions 48 48 4 0
rem 48 % des informations proviendront des connexions nouvelles,
rem des connexions actives, 4 % des conseillers et les entrées
rem système ne seront pas prises en compte.
rem
rem REMARQUE : par défaut, les propriétés du gestionnaires sont
définies comme suit :
rem 50 50 0 0

```

```

rem echo "Démarrage du gestionnaire (manager)..."
rem call dscontrol manager start

```

```

rem echo "Démarrage du conseiller (advisor) FTP sur le port 21..."
rem call dscontrol advisor start ftp 21
rem echo "Démarrage du conseiller (advisor) HTTP sur le port 80..."
rem call dscontrol advisor start http 80
rem echo "Démarrage du conseiller Telnet sur le port 23 ..."
rem call dscontrol advisor start telnet 23
rem echo "Démarrage du conseiller SMTP sur le port 25 ..."
rem call dscontrol advisor start smtp 25
rem echo "Démarrage du conseiller POP3 sur le port 110 ..."
rem call dscontrol advisor start pop3 110
rem echo "Démarrage du conseiller NNTP sur le port 119 ..."
rem call dscontrol advisor start nntp 119
rem echo "Démarrage du conseiller SSL sur le port 443 ..."
rem call dscontrol advisor start ssl 443
rem

```

```

rem echo "Définition des proportions du cluster..."
rem call dscontrol cluster set %CLUSTER% proportions 58 40 2 0

```

```

rem
rem L'étape finale de configuration du répartiteur est
rem l'affectation d'un alias à la carte d'interface réseau (NIC).
rem
rem
rem REMARQUE : N'utilisez pas cette commande dans un environnement à
rem haute disponibilité. Les scripts go* configureront les cartes NIC
rem et l'unité de bouclage si nécessaire.
rem
rem dscontrol executor configure %CLUSTER%

```

```

rem Si votre adresse de cluster se trouve sur une autre carte NIC
rem ou sur un sous-réseau autre que l'adresse NFA, utilisez la
rem commande de configuration de cluster suivante.
rem dscontrol executor configure %CLUSTER% tr0 0xfffff800
rem tr0 étant votre carte NIC (tr1 la seconde carte réseau en anneau
rem à jeton, en0 la première carte Ethernet) et 0xfffff800,
rem un masque de sous-réseau valide de votre site.
rem

```

```

rem
rem Les valeurs par défaut sont affectées aux commandes suivantes.

```

rem Utilisez ces commandes pour modifier les valeurs par défaut.

Conseiller type

Vous trouverez ci-dessous un exemple de fichier de conseiller intitulé **ADV_type**.

```
/**
 * ADV_type : Conseiller HTTP de Load Balancer
 *
 *
 * Cette classe définit un exemple de conseiller personnalisé pour Load Balancer.
 * Comme tous les conseillers, le conseiller personnalisé étend la fonction de
 * la base du conseiller, appelée ADV_Base. En fait, c'est cette base qui
 * effectue la plupart des fonctions du conseiller, telles que la communication
 * des charges à Load Balancer pour l'algorithme de pondération de Load Balancer.
 * La base du conseiller effectue également les opérations de connexion
 * et de fermeture de la connexion et fournit des méthodes d'envoi et de
 * réception qui seront utilisées par le conseiller. Le conseiller n'est
 * lui-même utilisé que pour l'envoi de données vers le port du serveur
 * conseillé et pour la réception de données sur ce dernier. Les méthodes
 * TCP de la base du conseiller sont programmées pour calculer la charge.
 * Un indicateur du constructeur de ADV_base remplace, si vous le souhaitez,
 * la charge existante par la nouvelle charge renvoyée par le conseiller.
 *
 * Remarque : En fonction d'une valeur définie dans le constructeur, la base
 * du conseiller fournit la charge à l'algorithme de pondération à
 * un intervalle donné.
 * Si le véritable conseiller n'a pas terminé ses opérations afin de
 * renvoyer une charge valide, la base du conseiller utilise
 * la charge précédente.
 *
 *
 * ATTRIBUTION DE NOM
 */
```

```

* La convention d'attribution de nom est la suivante :
*
* - Le fichier doit se trouver dans le répertoire Load Balancer suivant :
*
*    lb/servers/lib/CustomAdvisors/ (lb\servers\lib\CustomAdvisors sous Windows)
*
* - Le nom du conseiller doit être précédé de "ADV". Il peut
*   cependant n'être lancé qu'à partir du nom. Par exemple, le
*   conseiller peut être lancé avec "modèle".
*
* - Le nom du conseiller doit être en minuscules.
*
* En respectant ces règles, le chemin et le nom du conseiller
* donné en exemple sont les suivants :
*
*    <répertoire de base>/lib/CustomAdvisors/ADV_sample.class
*
* Les conseillers, tout comme les autres éléments de Load Balancer,
* doivent être compilés avec la version Java recommandée. Pour
* garantir l'accès aux classes Load Balancer, vérifiez que le fichier
* lbmlb.jar (situé dans le sous-répertoire lib du répertoire
* de base) figure dans la classe d'environnement CLASSPATH du système.
*
* Méthodes fournies par ADV_Base :
*
* - ADV_Base (Constructeur) :
*
*   - Paramètres
*     - String sName = Nom du conseiller
*     - String sVersion = Version du conseiller
*     - int iDefaultPort = Numéro de port par défaut utilisé par le conseiller
*     - int iInterval = Intervalle que doivent utiliser les serveurs
*     - String sDefaultName = Non utilisé. indiquer "".
*     - boolean replace = True - remplacement de la valeur de la charge
*                               par la base du conseiller
*                               False - ajout à la valeur de la charge calculée
*                               par la base du conseiller
*   - Return
*     - Les constructeurs n'ont pas de valeurs de retour.
*
* la base de conseiller étant basée sur une arborescence,
* le conseiller a de nombreuses autres méthodes
* d'utilisation à sa disposition. Ces méthodes peuvent
* être référencées en utilisant le paramètre CALLER dans
* getLoad().
*
* Ces méthodes sont les suivantes :
*
* - send - Envoie un paquet de données concernant la connexion
*         socket établie sur le port spécifié du serveur.
*   - Paramètres
*     - String sDataString - Les données à envoyer se présentent sous
*       forme de chaîne
*   - Return
*     - int RC - Indique si les données ont été correctement envoyées ;
*       un zéro indique que les données ont été envoyées ; un
*       entier négatif indique une erreur.
*
* - receive - Reçoit des informations de la connexion socket.
*   - Paramètres
*     - StringBuffer sbDataBuffer - Données reçues pendant l'appel
*   - Return
*     - int RC - Indique si les données ont été correctement
*       reçues ; un zéro indique que les données ont été
*       envoyées ; un entier négatif indique une
*       erreur.

```

```

*
* Si la fonction fournie par la base du conseiller n'est pas
* suffisante, vous pouvez créer la fonction appropriée dans le
* conseiller et les méthodes fournies par la base du conseiller
* seront alors ignorées.
*
* Il est essentiel de savoir si la charge renvoyée doit être
* appliquée à la charge générée dans la base du conseiller ou
* ou la remplacer ; les deux situations sont possibles.
*
* Cet exemple concerne principalement le conseiller HTTP de Load Balancer.
* Il fonctionne très simplement :une demande d'envoi (demande http
* principale) est émise. Dès que la réponse est reçue, la méthode getLoad
* se termine, indiquant à la base du conseiller d'arrêter de chronométrer.
* La méthode est alors terminée. Les informations renvoyées ne sont pas
* analysées, la charge est fonction du temps nécessaire pour effectuer
* les opérations d'envoi et de réception.
*/

```

```

package CustomAdvisors;
import com.ibm.internet.nd.advisors.*;

public class ADV_sample extends ADV_Base implements ADV_MethodInterface
{
    Chaîne COPYRIGHT = "(C) Copyright IBM Corporation 1997, Tous
droits réservés.\n";

    static final String  ADV_NAME          = "Type";
    static final int     ADV_DEF_ADV_ON_PORT = 80;
    static final int     ADV_DEF_INTERVAL  = 7;

    // Remarque : La plupart des protocoles de serveur
    // requièrent un retour chariot ("\r") et un passage à
    // la ligne ("\n") à la fin des messages. Si tel est le
    // cas, incluez-les dans la chaîne ci-dessous.
    static final String  ADV_SEND_REQUEST  =
"HEAD / HTTP/1.0\r\nAccept: */*\r\nUser-Agent: " +
"IBM_Load_Balancer_HTTP_Advisor\r\n\r\n";

    /**
     * Constructeur.
     *
     * Paramètres : Aucun, mais le constructeur de ADV_Base
     * comporte plusieurs paramètres qui doivent lui être transmis.
     */
    public ADV_type()
    {
        super( ADV_NAME,
              "2.0.0.0-03.27.98",
              ADV_DEF_ADV_ON_PORT,
              ADV_DEF_INTERVAL,
              "", // not used
              false);
        super.setAdvisor( this );
    }

    /**
     * ADV_AdvisorInitialize
     *
     * Toute initialisation spécifique au conseiller qui doit être mise
     * en oeuvre après le démarrage de la base du conseiller.
     * Cette méthode n'est appelée qu'une seule fois et n'est généralement
     * pas utilisée.
     */
    public void ADV_AdvisorInitialize()

```

```

{
    return;
}

/**
 * getLoad()
 *
 * Cette méthode est appelée par la base du conseiller pour terminer
 * l'opération du conseiller, basée sur des détails propres au protocole.
 * Dans cet exemple de conseiller, une seule opération d'envoi
 * et de réception est nécessaire ; si une logique plus
 * complexe est nécessaire, il est possible d'émettre des envois
 * et réceptions multiples. Par exemple, une réponse peut être
 * reçue et sa syntaxe peut être analysée. Sur la base
 * des informations obtenues, un autre ordre d'envoi et de
 * réception peut alors être émis.
 *
 * Paramètres :
 *
 * - iConnectTime - la charge actuelle car elle fait référence au temps
 *                  nécessaire à l'établissement d'une connexion
 *                  avec le serveur sur le port spécifié.
 *
 * - caller - Une référence à la classe de la base du conseiller
 *             dans laquelle les méthodes fournies par Load
 *             Balancer doivent répondre à de simples demandes
 *             TCP, principalement des demandes d'envoi et de
 *             réception.
 *
 * Résultats :
 *
 * - La charge - Valeur exprimée en millisecondes, pouvant être
 *   ajoutée à la charge existante, ou la remplacer, suivant la
 *   valeur de l'indicateur "replace" du constructeur.
 *
 * Plus la charge est importante, plus le serveur met de temps
 * à répondre et donc plus la charge de Load Balancer diminue.
 *
 * Si la valeur est négative, il s'agit d'une erreur. Une erreur
 * du conseiller indique que le serveur auquel le conseiller tente
 * d'accéder n'est pas accessible et qu'une défaillance a été
 * identifiée. Load Balancer ne tentera pas d'équilibrer un serveur
 * défaillant. Load Balancer recommencera à équilibrer la charge
 * du serveur à la réception d'une valeur positive.
 */
public int getLoad(int iConnectTime, ADV_Thread caller)
{
    int iRc;
    int iLoad = ADV_HOST_INACCESSIBLE; // -1

    // Send tcp request
    iRc = caller.send(ADV_SEND_REQUEST);
    if (iRc >= 0)
    {
        // Réception
        StringBuffer sbReceiveData = new StringBuffer("");
        iRc = caller.receive(sbReceiveData);

        /**
         * En mode conseiller normal (l'indicateur "replace" a la valeur false),
         * la valeur renvoyée est 0 ou 1 selon que le serveur est actif ou inactif.
         * En cas de bonne réception, une charge de zéro est renvoyée pour
         * indiquer que la charge élaborée dans la base du conseiller
         * n'est pas utilisée.

```



```

*
* Sinon (l'indicateur "replace" a la valeur true), renvoie la valeur de
* charge souhaitée.
*/

if (iRc >= 0)
{
    iLoad = 0;
}
}
return iLoad;
}
} // End - ADV_type

```

Annexe D. Exemple de configuration de haute disponibilité à deux niveaux utilisant Dispatcher, CBR et Caching Proxy

Cette annexe décrit comment définir une configuration de haute disponibilité à deux niveaux regroupant les capacités de deux composants Load Balancer (Dispatcher et CBR) et Caching Proxy.

Configuration de la machine serveur

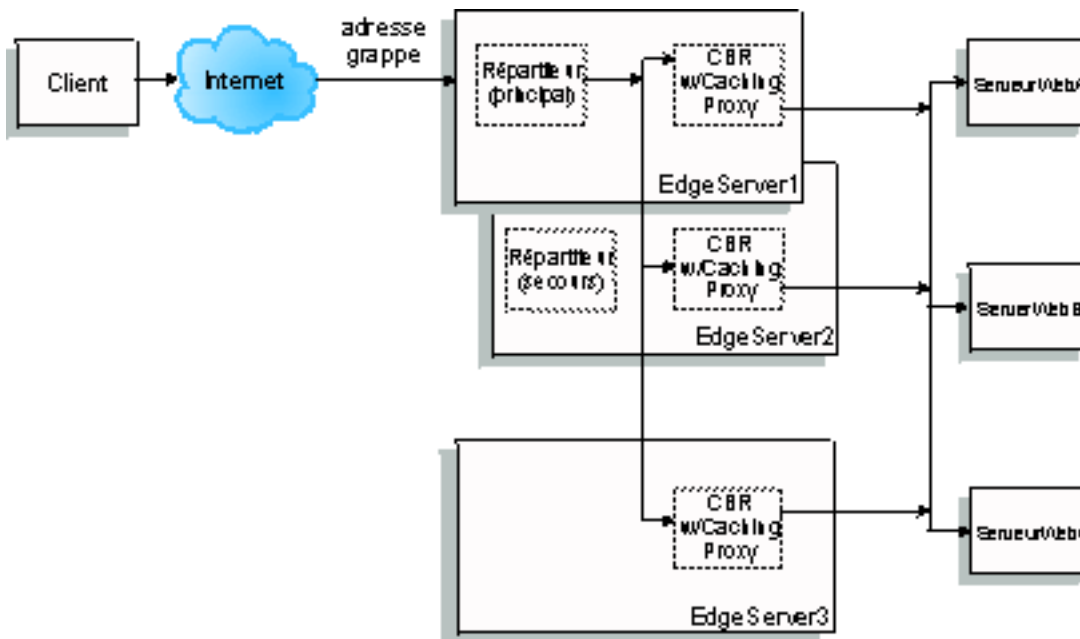


Figure 46. Exemple de configuration de haute disponibilité à deux niveaux utilisant Dispatcher, CBR et Caching Proxy

Le serveur suivant est configuré pour la figure 46 :

- EdgeServer1 : machine Dispatcher principale (haute disponibilité) co-implantée avec CBR et Caching Proxy équilibrant la charge des serveurs Web.
- EdgeServer2 : machine Dispatcher de secours (haute disponibilité) co-implantée avec CBR et Caching Proxy.
- EdgeServer3 : CBR et machine Caching Proxy.
- ServeurWebA, ServeurWebB, ServeurWebC : serveurs Web d'arrière-plan.

La figure 46 montre une représentation de base de l'équilibrage de la charge de plusieurs serveurs (EdgeServer1, EdgeServer2, EdgeServer3) sur plusieurs serveurs Web d'arrière-plan. Le composant CBR utilise Caching Proxy pour acheminer les demandes vers les serveurs Web d'arrière-plan en fonction de l'URL. Le composant Dispatcher permet d'équilibrer la charge des composants CBR sur les serveurs EdgeServer. La fonction haute disponibilité de Dispatcher permet d'assurer l'acheminement des demandes vers les serveurs dorsaux même en cas de défaillance de la machine haute disponibilité principale (EdgeServer1).

Instructions de configuration de base :

- Configurez Caching Proxy de la même façon sur tous les serveurs EdgeServer. Pour améliorer l'accessibilité globale aux pages Web sur les serveurs dorsaux, configurez Caching Proxy pour permettre le stockage en mémoire cache. Les serveurs EdgeServer peuvent ainsi y placer les pages Web les plus souvent demandées. Pour plus d'informations sur la configuration de Caching Proxy, reportez-vous au *Guide d'administration de Caching Proxy*.
- Définissez l'adresse du cluster et les ports de la même façon dans les composants CBR et Dispatcher de Load Balancer.
- Configurez le composant CBR de la même façon sur tous les serveurs EdgeServer. Utilisez les serveurs Web A, B et C sur les ports que vous souhaitez définir pour le cluster. Pour plus d'informations sur la configuration de CBR, voir Chapitre 11, «Configuration de CBR (Content Based Routing)», à la page 109.
- Configurez le composant Dispatcher de la même façon sur les serveurs EdgeServer1 et EdgeServer2. Définissez tous les serveurs EdgeServer comme serveurs des ports que vous souhaitez définir pour le cluster dont la charge est équilibrée par Dispatcher. Pour plus d'informations sur la configuration de Dispatcher, voir Chapitre 7, «Configuration de Dispatcher», à la page 63.
- Configurez EdgeServer1 comme serveur de haute disponibilité principal et EdgeServer2 comme serveur de haute disponibilité de secours. Pour plus d'informations, voir «Haute disponibilité», à la page 204.

Remarque :

1. Pour éviter l'affichage des adresses des serveurs dorsaux dans l'URL d'un client, définissez l'instruction ReversePass pour chaque adresse de serveur dorsal du fichier de configuration de Caching Proxy.
2. Pour assurer un stockage efficace en mémoire cache des pages Web, associez la valeur "ON" à l'instruction "Caching" et augmentez la valeur affectée à l'instruction "CacheMemory" en fonction de la taille requise dans le fichier de configuration de Caching Proxy.
3. Lignes de l'exemple auxquelles font référence les remarques 1 et 2 ci-dessus :


```
Caching          ON
CacheMemory      128000 K
ReversePass /* http://websrvA.company.com/* http://www.company.com/*
```
4. N'oubliez pas d'affecter un alias à l'adresse de cluster de la carte d'interface réseau de EdgeServer1 et à l'adresse de cluster de de l'unité de bouclage des autres serveurs EdgeServer.
5. Si les serveurs EdgeServers fonctionnent sous Linux, il peut s'avérer nécessaire d'installer un correctif pour le noyau Linux ou d'utiliser une autre méthode que l'affectation d'alias à l'unité de bouclage. Pour plus d'informations, voir «Solutions alternatives pour l'affectation d'alias à l'unité de bouclage sous Linux lors de l'utilisation de la méthode d'acheminement MAC de Load Balancer», à la page 78.
6. Avec CBR, l'affinité de port (délai de maintien de routage) ne doit pas être utilisée conjointement aux règles Contenu, sinon, ces dernières ne sont pas appliquées lors du traitement des demandes envoyées aux serveurs Web d'arrière-plan.

Fichiers de configuration exemple :

Les fichiers de configuration exemple suivants sont identiques à ceux créés lors de la définition de la configuration de Edge Components comme illustré à la figure 46, à la page 493. Ils correspondent aux fichiers des composants Dispatcher et CBR de Load Balancer. Dans ces fichiers, une seule carte de réseau Ethernet est utilisée pour chaque machine EdgeServer et toutes les adresses sont représentées dans un sous-réseau privé. Les fichiers de configuration exemple utilisent les adresses IP suivantes pour les machines spécifiées :

- EdgeServer1 (EdgeServer haute disponibilité principal) : 192.168.1.10
- EdgeServer2 (EdgeServer haute disponibilité de secours) : 192.168.1.20
- EdgeServer3 (EdgeServer de mise en cache des pages Web) : 192.168.1.30
- Adresse du cluster du site Web : 192.168.1.11
- ServeurWebA à C (serveurs Web d'arrière-plan) : 192.168.1.71, 192.168.1.72 et 192.168.1.73

Fichier de configuration exemple du composant Dispatcher d'un serveur EdgeServer haute disponibilité principal :

```
dscontrol executor start

dscontrol cluster add 192.168.1.11 primaryhost 192.168.1.10

dscontrol port add 192.168.1.11:80

dscontrol server add 192.168.1.11:80:edgeserver1 address 192.168.1.10

dscontrol server add 192.168.1.11:80:edgeserver2 address 192.168.1.20

dscontrol server add 192.168.1.11:80:edgeserver3 address 192.168.1.30

dscontrol manager start manager.log 10004

dscontrol highavailability heartbeat add 192.168.1.10 192.168.1.20
dscontrol highavailability backup add primary auto 4567
```

Fichier de configuration exemple du composant CBR des serveurs EdgeServer :

```
cbrcontrol set loglevel 1
cbrcontrol executor start

cbrcontrol cluster add 192.168.1.11

cbrcontrol port add 192.168.1.11:80

cbrcontrol server add 192.168.1.11:80:webserverA address 192.168.1.71

cbrcontrol server add 192.168.1.11:80:webserverB address 192.168.1.72

cbrcontrol server add 192.168.1.11:80:webserverC address 192.168.1.73

cbrcontrol rule add 192.168.1.11:80:webA_rule type content
  pattern (URI=*WSA*)|(URI=*wsA*) priority 21
cbrcontrol rule useserver 192.168.1.11:80:webA_rule webserverA

cbrcontrol rule add 192.168.1.11:80:webB_rule type content
  pattern (URI=/WS_B*) priority 22
cbrcontrol rule useserver 192.168.1.11:80:webB_rule webserverB

cbrcontrol rule add 192.168.1.11:80:webC_rule type content
  pattern URI=*webC* priority 23
cbrcontrol rule useserver 192.168.1.21:80:webC_rule webserverC
```

Annexe E. Remarques

Le présent document peut contenir des informations ou des références concernant certains produits, logiciels ou services IBM non annoncés dans ce pays. Pour plus de détails, référez-vous aux documents d'annonce disponibles dans votre pays, ou adressez-vous à votre partenaire commercial IBM. Toute référence à un produit, logiciel ou service IBM n'implique pas que seul ce produit, logiciel ou service puisse être utilisé. Tout autre élément fonctionnellement équivalent peut être utilisé, s'il n'enfreint aucun droit d'IBM. Il est de la responsabilité de l'utilisateur d'évaluer et de vérifier lui-même les installations et applications réalisées avec des produits, logiciels ou services non expressément référencés par IBM.

IBM peut détenir des brevets ou des demandes de brevet sur certains sujets décrits dans le présent document. La remise de ce document ne vous donne aucun droit de licence sur ces brevets ou demandes de brevet. Si vous désirez recevoir des informations concernant l'acquisition de licences, veuillez en faire la demande par écrit à l'adresse suivante :

IBM Director of
Licensing
IBM Corporation
500 Columbus Avenue
Thornwood, NY 10594
U.S.A.

Les informations sur les licences concernant les produits utilisant un jeu de caractères double octet peuvent être obtenues par écrit à l'adresse suivante :

IBM World Trade Asia Corporation
Licensing
2-31 Roppongi 3-chome, Minato-ku
Tokyo 106, Japan

Le paragraphe suivant ne s'applique ni au Royaume-Uni, ni dans aucun pays dans lequel il serait contraire aux lois locales.

LE PRESENT DOCUMENT EST LIVRE EN L'ETAT. IBM DECLINE TOUTE RESPONSABILITE, EXPLICITE OU IMPLICITE, RELATIVE AUX INFORMATIONS QUI Y SONT CONTENUES, Y COMPRIS EN CE QUI CONCERNE LES GARANTIES DE VALEUR MARCHANDE OU D'ADAPTATION A VOS BESOINS. Certaines juridictions n'autorisent pas l'exclusion des garanties implicites, auquel cas l'exclusion ci-dessus ne vous sera pas applicable.

Le présent document peut contenir des inexactitudes ou des coquilles. Il est mis à jour périodiquement. Chaque nouvelle édition inclut les mises à jour. IBM peut modifier sans préavis les produits et logiciels décrits dans ce document.

Les références à des sites Web non IBM sont fournies à titre d'information uniquement et n'impliquent en aucun cas une adhésion aux données qu'ils contiennent. Les éléments figurant sur ces sites Web ne font pas partie des éléments du présent produit IBM et l'utilisation de ces sites relève de votre seule responsabilité.

IBM pourra utiliser ou diffuser, de toute manière qu'elle jugera appropriée et sans aucune obligation de sa part, tout ou partie des informations qui lui seront fournies.

Les licenciés souhaitant obtenir des informations permettant : (i) l'échange des données entre des logiciels créés de façon indépendante et d'autres logiciels (dont celui-ci), et (ii) l'utilisation mutuelle des données ainsi échangées, doivent adresser leur demande à :

IBM Corporation
Attn.: G7IA./503.
P.O. Box 12195
3039 Cornwallis Rd.
Research Triangle Park, N.C. 27709-2195
U.S.A.

Ces informations peuvent être soumises à des conditions particulières, prévoyant notamment le paiement d'une redevance.

Le logiciel sous licence décrit dans le présent document et tous les éléments sous licence disponibles pour celui-ci sont fournis par IBM dans les conditions internationales d'utilisation des logiciels IBM ou autre contrat de clientèle IBM.

Les données de performance indiquées dans ce document ont été déterminées dans un environnement contrôlé. Par conséquent, les résultats peuvent varier de manière significative selon l'environnement d'exploitation utilisé. Certaines mesures évaluées sur des systèmes en cours de développement ne sont pas garanties sur tous les systèmes disponibles. En outre, elles peuvent résulter d'extrapolations. Les résultats peuvent donc varier. Il incombe aux utilisateurs de ce document de vérifier si ces données sont applicables à leur environnement d'exploitation.

Les informations concernant des produits non IBM ont été obtenues auprès des fournisseurs de ces produits, par l'intermédiaire d'annonces publiques ou via d'autres sources disponibles. IBM n'a pas testé ces produits et ne peut confirmer l'exactitude de leurs performances ni leur compatibilité. Elle ne peut recevoir aucune réclamation concernant des produits non IBM. Toute question concernant les performances de produits non IBM doit être adressée aux fournisseurs de ces produits.

Toute instruction relative aux intentions d'IBM pour ses opérations à venir est susceptible d'être modifiée ou annulée sans préavis, et doit être considérée uniquement comme un objectif.

Le présent document peut contenir des exemples de données et de rapports utilisés couramment dans l'environnement professionnel. Ces exemples mentionnent des noms fictifs de personnes, de sociétés, de marques ou de produits à des fins illustratives ou explicatives uniquement. Toute ressemblance avec des noms de personnes, de sociétés ou des données réelles serait purement fortuite.

Si vous visualisez ces informations en ligne, il se peut que les photographies et illustrations en couleur n'apparaissent pas à l'écran.

Marques

Les termes qui suivent sont des marques d'IBM Corporation aux Etats-Unis et/ou dans certains autres pays.

AFS
AIX
DFS
IBM
iSeries
NetView
OS/2
Redbooks
RS/6000
SecureWay
ViaVoice
WebSphere
zSeries

Java ainsi que toutes les marques incluant Java sont des marques de Sun Microsystems, Inc. aux Etats-Unis et/ou dans certains autres pays.

Microsoft, Windows, Windows NT et le logo Windows sont des marques de Microsoft Corporation aux Etats-Unis et/ou dans certains autres pays.

Intel, Intel Inside (logos), MMX et Pentium sont des marques d'Intel Corporation aux Etats-Unis et/ou dans certains autres pays.

UNIX est une marque enregistrée de The Open Group aux Etats-Unis et /ou dans certains autres pays.

Linux est une marque de Linus Torvalds aux Etats-Unis et/ou dans certains autres pays.

Les autres noms de sociétés, de produits et de services peuvent appartenir à des tiers.

Glossaire

Nortel Alteon Controller : Composant d'IBM Load Balancer. Nortel Alteon Controller utilise la technologie Load Balancer pour fournir en temps réel des informations sur l'équilibrage de charge à Nortel Alteon Web Switch.

Nortel Alteon Web Switch : Systèmes Nortel Alteon ACE Director Series Switch et Nortel Alteon 180 Series Switch de la gamme Alteon Web Switching, utilisés pour l'acheminement de paquets et le routage de contenus.

Dispatcher : Composant de Load Balancer qui permet l'équilibrage efficace du trafic TCP ou UDP entre des groupes de serveurs reliés. La machine Dispatcher correspond au serveur qui exécute le code Dispatcher.

A

ACK : Bit de contrôle n'occupant aucun espace séquentiel, qui indique que la zone d'accusé de réception de ce segment précise le prochain numéro de séquence à recevoir par l'émetteur de ce segment, d'où accusant réception de tous les numéros de séquence précédents.

adresse : Code unique affecté à chaque unité ou poste de travail connecté à un réseau. Une adresse IPv4 standard est une zone d'adresse 32 bits contenant deux parties. La première correspond à l'adresse réseau et la seconde au numéro d'hôte. Une adresse IPv6 est une zone d'adresse de 128 bits prenant en charge un nombre d'adresses bien plus important qu'IPv4. En outre, elle prend en charge des fonctions supplémentaires telles que l'adressage multidiffusion et anycast.

adresse à contacter : Dans le cadre des fonctions de haute disponibilité de Dispatcher, adresse du poste cible auquel le conseiller doit envoyer des pings pour vérifier si ce poste répond.

adresse cible : Adresse du poste partenaire de haute disponibilité auquel des signaux de présence et des réponses sont envoyés.

adresse de cluster : Sous Dispatcher, adresse à laquelle les clients se connectent.

adresse de la mesure : Adresse à laquelle se connecte le serveur de mesures.

adresse de non-réacheminement (NFA) : Adresse IP principale du poste Load Balancer, utilisée pour l'administration et la configuration.

Adresse de retour : Adresse IP ou nom d'hôte unique. Ils sont configurés sur la machine Dispatcher et utilisés par Dispatcher comme adresse source lors de l'équilibrage de charge des demandes client sur le serveur.

adresse du serveur : Code unique attribué à chaque ordinateur fournissant des services partagés à d'autres ordinateurs d'un réseau ; par exemple, un serveur de fichiers, un serveur d'impression ou un serveur de messagerie. Il peut s'agir de l'adresse IP ou du nom d'hôte.

adresse IP : Adresse Internet Protocol. Adresse unique, indiquant l'emplacement réel de chaque unité ou poste de travail d'un réseau. Elle est également désignée par adresse Internet.

adresse MAC : De l'anglais Media Access Control. Adresse matérielle d'un support réseau partagé.

adresse source : Dans le cadre des fonctions de haute disponibilité de Dispatcher, adresse du poste partenaire de haute disponibilité envoyant des signaux de présence.

affinité trans ports : L'affinité trans ports se définit comme l'extension à plusieurs ports de la fonction affinité (maintien de routage). Voir également délai de maintien de routage.

agent : (1) En gestion de systèmes, utilisateur qui, pour une interaction particulière, a assumé un rôle d'agent. (2) Entité représentant un ou plusieurs objets gérés (a) en émettant des notifications relatives aux objets et (b) en traitant les demandes d'opérations de gestion émises par les gestionnaires pour modifier ou consulter les objets.

alias : Autre nom attribué à un serveur. Grâce à l'alias, le serveur est indépendant du nom de sa machine hôte. L'alias doit être défini dans le serveur de noms de domaine.

alias de bouclage : Seconde adresse IP associée à l'interface de bouclage. Cette adresse présente l'avantage de ne pas apparaître sur une interface réelle.

API : Interface de programme d'application (Application programming interface). Interface (conventions d'appellation) par laquelle un programme d'application accède au système d'exploitation et autres services. L'API est définie au niveau du code source et apporte un niveau d'abstraction entre l'application et le noyau (ou autres utilitaires privilégiés) afin de garantir la portabilité du code.

assistant : Dans une application, dialogue utilisant des instructions étape par étape pour guider l'utilisateur dans une tâche spécifique.

C

Caching Proxy : Serveur proxy de mise en mémoire cache pouvant contribuer à accélérer les temps de réponse grâce à des schémas de mise en mémoire cache extrêmement efficaces. Un filtrage PICS souple aide les administrateurs réseau à contrôler l'accès aux informations du Web à un emplacement central.

CBR : Abréviation de Content Based Routing. Composant de Load Balancer. CBR utilise Caching Proxy pour équilibrer la charge des demandes entrantes, en fonction du contenu de la page Web qui utilise des types de règle précis, sur des serveurs HTTP ou HTTPS.

cbrcontrol : Fournit l'interface du composant Content Based Router de Load Balancer.

cbrserver : Dans Content Based Router, traite les demandes adressées à l'exécuteur, au gestionnaire et aux conseillers à partir de la ligne de commande.

ccocontrol : Dans Cisco CSS Controller, fournit l'interface de Cisco CSS Switch.

ccoserver : Dans Cisco CSS Controller, traite les demandes adressées aux consultants à partir de la ligne de commande.

CGI : Abréviation de Common Gateway Interface. Norme permettant l'échange d'informations entre un serveur Web et un programme externe. Ce dernier peut être écrit dans n'importe quel langage pris en charge par le système d'exploitation et permet d'effectuer des tâches qui ne sont généralement pas exécutées par le serveur, par exemple, le traitement des formulaires.

Cisco CSS Controller : Composant d'IBM Load Balancer. Cisco CSS Controller utilise la technologie Load Balancer pour fournir en temps réel des informations sur l'équilibrage de charge à Cisco Content Services Switch.

Cisco CSS Switch : Il s'agit de l'un des commutateurs CSS série 11000 de Cisco, utilisé pour l'acheminement des paquets et le routage du contenu.

client : Système ou processus informatique demandant un service à un autre système ou processus informatique. Par exemple, un poste de travail ou un ordinateur personnel demandant des documents HTML à un serveur Web Lotus Domino Go est un client de ce serveur.

cluster : Sous Dispatcher, groupe de serveurs TCP ou UDP utilisés pour le même but et identifiés par un seul nom d'hôte. Voir aussi cellule.

cohabitation : Lorsque Load Balancer est installé sur la machine dont il assure l'équilibrage de charge.

Co-implantation d'adresses multiples : Dans une configuration de co-implantation d'adresses multiples, le client peut indiquer une adresse de serveur co-implanté différente de celle de non-acheminement (NFA). Voir aussi Co-implantation.

conseiller : Fonction de Load Balancer. Les conseillers collectent et analysent les informations renvoyées par chaque serveur, puis en informent le gestionnaire.

conseiller de contact (reach) : Sous Dispatcher, conseiller qui lance des pings vers une cible déterminée, puis indique si cette cible répond ou non.

consignation binaire : Permet de stocker les informations du serveur dans des fichiers binaires et de traiter ces fichiers afin d'analyser les informations du serveur recueillies.

consultant : Collecte les mesures des serveurs dont la charge est équilibrée et envoie des informations de pondération de serveur au commutateur chargé de l'équilibrage de la charge.

contenu propriétaire : Nom de propriétaire et règle de contenu pour un propriétaire, tous deux définis sur Cisco CSS Switch.

contrôleur : Collection d'un ou plusieurs consultants.

Conversion d'adresses réseau : NAT ou Network Address Translator, Virtual LAN. Périphérique en cours de développement permettant d'étendre les adresses Internet déjà utilisées. Il permet l'utilisation d'adresses dupliquées dans une entreprise et d'une adresse unique à l'extérieur.

D

délai de maintien de routage : Délai entre la fermeture d'une connexion et l'ouverture d'une nouvelle connexion au cours de laquelle un client sera renvoyé au même serveur utilisé lors de la première connexion. Passé le délai de maintien de routage, le client peut être envoyé à un serveur autre que le premier.

délai d'expiration : Intervalle de temps alloué à une opération.

démon : Contrôleur de disque et d'exécution. Programme non impliqué explicitement, qui attend certaines conditions pour agir. L'idée est que le logiciel n'a pas besoin de savoir que le démon est actif (bien que le programme effectue souvent une opération uniquement parce qu'il appellera implicitement un démon).

dscontrol : Fournit l'interface du composant Dispatcher de Load Balancer.

dsserver : Dans Dispatcher, traite les demandes adressées à l'exécuteur, au gestionnaire et aux conseillers à partir de la ligne de commande.

E

état down (hors service) : Interruption de toutes les connexions actives à un serveur et arrêt de l'envoi de nouvelles connexions ou de nouveaux paquets à ce serveur.

état FIN : Etat d'une transaction terminée. Lorsqu'une transaction est à l'état FIN, le programme récupérateur de Load Balancer peut libérer la mémoire réservée à la connexion.

état up (en service) : Permettre à un serveur de recevoir de nouvelles connexions.

Ethernet : Type standard de réseau local (lan). Il permet à plusieurs systèmes d'accéder à tout moment au support de transmission sans coordination préalable, évite les conflits à l'aide des fonctions de détection et de report de porteuse et résout les conflits éventuels en utilisant les options de détection et de transmission des collisions. Les protocoles de logiciels utilisés par les systèmes Ethernet varient, mais incluent TCP/IP.

Evolutivité : Capacité d'adaptation d'un système à une intensité d'utilisation, de volume ou de demande supérieure ou inférieure. Par exemple, un système évolutif peut s'adapter efficacement pour gérer de petits ou grands réseaux exécutant des tâches dont la complexité est variable.

exécuteur : Fonctions de Load Balancer. L'exécuteur achemine des demandes aux serveurs TCP ou UDP et contrôle également le nombre des connexions nouvelles, actives et terminées, puis regroupe les connexions terminées ou réinitialisées afin de récupérer de l'espace mémoire. L'exécuteur fournit les connexions nouvelles et actives au gestionnaire.

F

FIN : Bit de contrôle occupant un seul numéro de séquence, qui indique que l'émetteur n'enverra aucun autre contrôle ou donnée occupant un espace séquentiel.

FQDN : Abréviation de Fully Qualified Domain Name. Nom complet d'un système, comprenant son nom d'hôte local et son nom de domaine avec un domaine de niveau supérieur (tld). Par exemple, "venera" est un nom d'hôte et

"venera.isi.edu" est un nom FQDN. Le nom FQDN doit permettre de déterminer l'adresse Internet unique de n'importe quel hôte présent sur Internet. Ce processus, appelé "résolution de nom", utilise le système DNS (Domain Name System).

FTP (File Transfer Protocol) : Protocole d'application permettant le transfert de fichiers à destination et en provenance d'ordinateurs en réseau. Le protocole FTP requiert un ID utilisateur et, le cas échéant, un mot de passe permettant l'accès aux fichiers d'un système hôte éloigné.

G

gestionnaire : Fonctions de Load Balancer. Le gestionnaire définit des pondérations basées sur les compteurs internes de l'exécuteur et sur les informations renvoyées par les conseillers. L'exécuteur utilise ensuite les pondérations pour effectuer un équilibrage de charge.

Gestionnaire de paquets Red Hat (RPM) : Gestionnaire de paquets Red Hat (Red Hat Package Manager).

GRE : Abréviation de Generic Routing Encapsulation. Protocole permettant de transmettre un protocole réseau arbitraire A à un autre protocole arbitraire B en encapsulant les paquets du protocole A dans des paquets GRE, qui sont à leur tour intégrés à des paquets du protocole B.

H

haute disponibilité : Fonction de Load Balancer permettant à un Load Balancer de prendre le contrôle d'un autre si ce dernier est défaillant pour une raison quelconque.

haute disponibilité réciproque : La haute disponibilité réciproque permet à deux machines Dispatcher de jouer, l'une pour l'autre, le rôle de machine principale et de secours. Voir aussi machine de secours (backup), haute disponibilité, machine principale.

hôte : Ordinateur connecté à un réseau, fournissant un point d'accès à ce réseau. Il peut s'agir d'un client et/ou d'un serveur.

HTTP (Hypertext Transfer Protocol) : Protocole permettant de transférer et d'afficher des documents hypertexte.

I

ICMP : Abréviation d'Internet Control Message Protocol. Protocole qui permet la génération de messages d'erreur, de test et d'information relatifs aux conditions de transmission entre un serveur hôte et une passerelle vers Internet.

IMAP : Abréviation d'Internet Message Access Protocol. Protocole permettant au client d'accéder à des messages électroniques placés dans un serveur et de les manipuler. Il permet la manipulation à distance de dossiers de messages (boîtes aux lettres), suivant un procédé équivalent à celle des boîtes aux lettres locales.

interface de bouclage : Interface qui ignore les fonctions de communication non nécessaires lorsque les informations sont adressées à une entité du même système.

Internet : Ensemble mondial de réseaux interconnectés utilisant la série de protocoles Internet et permettant un accès public.

intranet : Réseau privé sécurisé intégrant des normes et applications Internet (par exemple, des navigateurs Web) à l'infrastructure de gestion de réseau informatique existante d'une entreprise.

IP : Abréviation de Internet Protocol. Protocole sans connexion, acheminant des données via un réseau ou des réseaux interconnectés. Le protocole IP sert d'intermédiaire entre les couches de protocole supérieures et la couche physique.

IPSEC : Abréviation de Internet Protocol Security. Norme de développement pour la sécurité au niveau de la couche réseau ou de traitement des paquets transmis lors des communications réseau.

L

langage HTML (Hypertext Markup Language) : Langage permettant de créer des documents hypertexte. Ces derniers comportent des liens avec d'autres documents contenant des informations complémentaires sur le terme ou le sujet mis en évidence. HTML contrôle, par exemple, le format du texte et de la position des zones d'entrée de formulaire, ainsi que les liens navigables.

largeur de bande : Différence entre la fréquence la plus élevée et la fréquence la plus basse d'une voie de transmission ; quantité de données pouvant être envoyée par un circuit de communication particulier par seconde.

M

machine de secours (backup) : Dans le cadre des fonctions de haute disponibilité de Dispatcher, partenaire de la machine principale (primary). Elle contrôle l'état de la machine principale (primary) et la remplace si nécessaire. Voir aussi haute disponibilité, machine principale

machine principale (primary) : Dans le cadre des fonctions de haute disponibilité de Dispatcher, poste qui démarre en tant que poste d'acheminement actif des paquets. Son partenaire, la machine de secours (backup), contrôle l'état de la machine principale et la remplace si nécessaire. Voir aussi machine de secours (backup), haute disponibilité.

masque de réseau : Pour IPv4, masque de 32 bits permettant d'identifier les bits d'adresse de sous-réseau de la partie hôte d'une adresse IP.

masque de sous-réseau : Pour IPv4, masque de 32 bits permettant d'identifier les bits d'adresse de sous-réseau de la partie hôte d'une adresse IP.

mesure : Processus ou commande qui renvoie une valeur numérique pouvant être utilisée lors de l'équilibrage de charge sur le réseau, par exemple, le nombre d'utilisateurs connectés.

mettre au repos : Mettre fin à un processus en permettant l'arrêt normal des opérations.

MIB : (1) Abréviation de Management Information Base. Ensemble d'objets accessibles via un protocole de gestion de réseau. (2) Définition des informations de gestion indiquant les informations accessibles à partir d'un hôte ou d'une passerelle, ainsi que les opérations autorisées.

N

nalcontrol : Fournit l'interface du composant Nortel Alteon Controller de Load Balancer.

nalserver : Dans Nortel Alteon Controller, traite les demandes adressées aux consultants à partir de la ligne de commande.

Network Address Port Translation : NAT, également appelé mappage de port. Il permet la configuration de plusieurs démons de serveurs dans un serveur physique unique qui écoutent sur différents numéros de ports.

NIC : Abréviation de Network Interface Card. Carte à circuit installé sur un ordinateur pour établir des connexions physiques à un réseau.

NNTP : Abréviation de Network News Transfer Protocol. Protocole TCP/IP permettant le transfert d'éléments d'informations.

noeud géré : Dans le cadre des communications Internet, poste de travail, serveur ou routeur contenant un agent de gestion de réseau. Dans le cadre du protocole IP (Internet Protocol), le noeud géré contient généralement un agent SNMP.

nom de site : Un nom de site est un nom d'hôte impossible à résoudre, qui sera demandé par le client. Par exemple, un site Web a 3 serveurs (1.2.3.4, 1.2.3.5 et 1.2.3.6) configurés pour le nom de site *www.dnsload.com*. Lorsqu'un client demande ce nom de site, l'une des trois adresses IP du serveur est renvoyée en tant que résolution. Le nom de site doit être un nom de domaine complet, comme *dnsload.com*. Un nom de site incomplet tel que *dnsload* n'est pas un nom de site valide.

nom d'hôte : Nom symbolique attribué à un hôte. Les noms d'hôte sont résolus à des adresses IP via un serveur de noms de domaine.

notation décimale : Représentation syntaxique d'un entier 32 bits constitué de quatre nombres de 8 bits, écrits dans la base 10 et séparés par des points. Elle sert à représenter des adresses IPv4.

P

paquet : Unité de données acheminées entre une source et une cible sur Internet ou tout autre réseau de commutation de paquets.

pare-feu : Ordinateur qui relie un réseau privé tel qu'une entreprise à un réseau public tel qu'Internet. Il contient des programmes limitant l'accès entre deux réseaux. Voir aussi *passerelle de proxy*.

passerelle : Unité fonctionnelle permettant d'interconnecter deux ordinateurs réseau ayant des architectures différentes.

PICS : Abréviation de Platform for Internet Content Selection. Les clients prenant en charge PICS permettent aux utilisateurs de déterminer les "rating services" (services de contrôle d'accès) qu'ils souhaitent utiliser et, pour chacun de ces rating services, s'il est acceptable ou non.

ping : Commande qui envoie des paquets de demande d'écho ICMP (Internet Control Message Protocol) à un hôte, à une passerelle ou à un routeur, en espérant recevoir une réponse.

POP3 : Abréviation de Post Office Protocol 3. Protocole permettant l'échange de courrier réseau et l'accès aux boîtes aux lettres.

port : Nombre qui identifie une unité de communication abstraite. Les serveurs Web utilisent par défaut le port 80.

poste de gestion de réseau : Dans le cadre du protocole SNMP (Simple Network Management Protocol), poste exécutant des protocoles d'application de gestion qui surveillent et contrôlent les éléments de réseau.

poste serveur : Serveur relié par Dispatcher à d'autres serveurs dans un seul serveur virtuel. Dispatcher équilibre le trafic entre les postes serveur. Synonyme de serveur groupé.

poste serveur TCP : Serveur relié par Load Balancer à d'autres serveurs dans un seul serveur virtuel. Load Balancer équilibre le trafic TCP entre les postes serveur TCP. Synonyme de serveur groupé.

priorité : Dans le cadre de l'équilibrage de charge basé sur des règles, niveau d'importance accordé à une règle donnée. Dispatcher évalue les règles du premier niveau de priorité jusqu'au dernier.

programme de collecte de mesures : Programme résidant sur le consultant et chargé de collecter une ou des mesures.

Protocole : Ensemble de règles régissant le fonctionnement des éléments d'un système de communication si des communications sont établies. Les protocoles peuvent déterminer les détails de niveau inférieur des interfaces entre des machines, par exemple, l'ordre de transmission des bits d'un octet ; ils peuvent également déterminer les échanges de haut niveau entre les programmes d'application, par exemple, le transfert de fichier.

protocole HTTPS (Hypertext Transfer Protocol, Secure) : Protocole permettant de transférer et d'afficher des documents hypertexte à l'aide de SSL.

proximité réseau : Proximité de deux entités connectées en réseau (par exemple, un client et un serveur), déterminée par Site Selector par la mesure de la durée de la procédure aller-retour.

Q

QoS (Quality of Service) : Propriétés de performances d'un service réseau, notamment le débit, le délai de transmission et la priorité. Certains protocoles permettent aux paquets ou aux flots de données d'inclure des spécifications QoS.

R

règle : Dans le cadre de l'équilibrage de charge basé sur des règles, mécanisme de regroupement de serveurs permettant de choisir un serveur en fonction des informations autres que l'adresse de destination ou le port.

réseau : Système de communication pour les données logicielles et matérielles. Les réseaux sont généralement classés en fonction de leur étendue géographique, réseau local (LAN), réseau métropolitain (MAN) ou réseau étendu (WAN), et des protocoles utilisés.

Réseau local (LAN) : Réseau local (Local Area Network). Réseau informatique d'unités connectés entre elles, à des fins de communication, dans une zone de taille limitée. Ce réseau local peut être connecté à un réseau plus étendu.

réseau privé : Réseau distinct sur lequel Dispatcher communique avec des serveurs groupés pour des raisons de performances.

RMI : Abréviation de Remote Method Invocation. Partie de la bibliothèque de langages de programmation Java qui permet à un programme Java d'être exécuté sur une machine et d'accéder aux objets et aux méthodes d'un autre programme Java exécuté sur une autre machine.

routeur : Périphérique chargé d'acheminer des paquets entre des réseaux. La décision d'acheminement dépend des informations de la couche réseau et des tables de routage, souvent élaborées par des logiciels de routage.

S

script CGI : Programme CGI écrit dans un langage de script tel que Perl ou REXX qui utilise l'interface CGI (Common Gateway Interface) pour effectuer des tâches qui ne sont généralement pas exécutées par le serveur, par exemple, le traitement des formulaires.

serveur : Ordinateur fournissant des services partagés à d'autres ordinateurs d'un réseau ; par exemple, un serveur de fichiers, un serveur d'impression ou un serveur de messagerie.

serveur de noms de domaines : DNS. Service de requête global de données réparties et répliquées, généralement utilisé sur Internet pour la conversion des noms d'hôte en adresses Internet. Type de nom d'hôte utilisé sur Internet, bien que le terme approprié soit "nom de domaine complet". DNS peut être configuré pour utiliser une série de serveurs de noms, en fonction des domaines figurant dans le nom recherché, jusqu'à ce qu'une valeur soit trouvée.

serveur groupé : Serveur relié par Dispatcher à d'autres serveurs dans un seul serveur virtuel. Load Balancer équilibre le trafic TCP ou UDP entre ces serveurs groupés.

service : (1) Fonction fournie par un ou plusieurs noeuds, par exemple, HTTP, FTP, Telnet. (2) Pour Nortel Alteon Controller, un service est la fonction ou les informations demandées par un utilisateur final à partir d'un site. Il est identifié par une adresse IP virtuelle et un numéro de port virtuel à la demande d'un utilisateur final. Sur le commutateur, il est défini par un identifiant de serveur virtuel, correspondant à un entier, un numéro de port virtuel ou un nom de service. (3) Pour Cisco CSS Consultant, un service est un emplacement cible où réside un segment de contenu physique. Par exemple, un serveur local ou éloigné et un port.

shell : Logiciel acceptant et traitant des lignes de commande à partir du poste de travail d'un utilisateur. Le Bash Korn constitue un des nombreux shells UNIX disponibles.

signal de présence (heartbeat) : Paquet simple transmis entre deux postes Load Balancer en mode haute disponibilité, utilisé par le poste Load Balancer de secours pour surveiller l'état du poste Load Balancer actif.

Site Selector : Composant Load Balancer pour l'équilibrage de charge utilisant le système DNS. Site Selector équilibre la charge sur des serveurs faisant partie d'un réseau étendu (WAN) en utilisant des mesures et des pondérations recueillies à l'aide du composant Metric Server qui s'exécute sur ces serveurs.

SMTP : Abréviation de Simple Mail Transfer Protocol. Dans la série des protocoles, protocole d'application permettant le transfert de messages entre les utilisateurs de l'environnement Internet. Le protocole SMTP indique les séquences d'échange et le format des messages. Il considère que TCP (Transmission Control Protocol) est le protocole sous-jacent.

SNMP : Abréviation de Simple Network Management Protocol. Protocole Internet standard défini par les spécifications STD 15, RFC 1157 et conçu pour gérer les noeuds d'un réseau IP. SNMP ne se limite pas à TCP/IP. Il peut être utilisé pour gérer et surveiller différents types d'équipements, notamment les ordinateurs, les routeurs, les concentrateurs et les chargeurs automatiques de disques.

SPARC : Architecture de processeur modulable.

sscontrol : Fournit l'interface du composant Site Selector de Load Balancer.

SSL : Abréviation de Secure Sockets Layer. Plan de sécurité courant développé par Netscape Communications Corp. conjointement avec RSA Data Security Inc. SSL permet l'authentification du serveur par le client, ainsi que le chiffrement de l'ensemble des données et demandes. L'URL d'un serveur sécurisé protégé par SSL commence par https (et non HTTP).

ssserver : Dans Site Selector, gère les demandes adressées au nom de site, au gestionnaire et aux conseillers à partir de la ligne de commande.

stratégie : Dans le cadre des fonctions de haute disponibilité de Dispatcher, mot clé permettant d'indiquer le mode de récupération d'espace à la suite de la défaillance de la machine active.

superutilisateur : Droits non restreints permettant d'accéder à une partie du système d'exploitation AIX, Red Hat Linux ou Solaris, et de modifier celle-ci, et généralement associés à l'utilisateur qui gère le système.

SYN : Bit de contrôle du segment entrant, occupant un seul numéro de séquence, utilisé au début d'une connexion, afin d'indiquer la position de début de la numérotation des séquences.

Système Metric Server : Précédemment appelé Server Monitor Agent (SMA). Metric Server fournit au gestionnaire Load Balancer des mesures spécifiques au système.

T

TCP : Abréviation de Transmission Control Protocol. Protocole de communication utilisé sur le réseau Internet. TCP permet un échange d'informations hôte à hôte fiable. Il utilise IP comme protocole sous-jacent.

TCP/IP : Abréviation de Transmission Control Protocol/Internet Protocol. Une série de protocoles permettant la communication entre les réseaux, quelles que soient les techniques de communication utilisées dans chaque réseau.

Telnet : Protocole d'émulation de terminal ; protocole d'application TCP/IP destiné au service de connexion éloigné. Telnet permet à un utilisateur d'un site d'accéder à un hôte éloigné comme si le poste de l'utilisateur était connecté directement à cet hôte éloigné.

TTL : Durée (en nombre de secondes) pendant laquelle un client peut enregistrer en mémoire cache la réponse de résolution de nom.

type de règle : Dans le cadre de l'équilibrage de charge basé sur des règles, indicateur des informations devant être évaluées pour déterminer si une règle est vraie.

Type de service (TOS) : Type de service (Type of service). Zone d'octets de 1, dans l'en-tête IP du paquet SYN.

U

UDP : Abréviation de User Datagram Protocol. Dans la série des protocoles Internet, protocole fournissant un service de datagramme sans connexion et non fiable. Il permet à un programme d'application d'une machine ou d'un processus d'envoyer un datagramme à un programme d'application d'une autre machine ou d'un autre processus. UDP utilise le protocole IP (Internet Protocol) pour transmettre des datagrammes.

URI : Universal Resource Identifier. Adresse codée de toute ressource disponible sur le Web (par exemple, un document HTML, une image, un clip video, un programme).

URL : Uniform Resource Locator. Méthode standard permettant d'indiquer l'emplacement d'un objet, généralement une page Web sur Internet. Les adresses URL possèdent la forme des adresses utilisées sur Internet. Elles sont utilisées dans des documents HTML afin de spécifier la cible d'un lien hypertexte qui correspond en général à un autre document HTML (pouvant être stocké sur une autre machine).

V

valeur de début : Dans le cadre de l'équilibrage de charge basé sur des règles, valeur inférieure indiquée pour une règle. La valeur par défaut dépend du type de règle.

valeur de fin : Dans le cadre de l'équilibrage de charge basé sur des règles, valeur supérieure indiquée pour une règle. Cette valeur dépend du type de règle.

valeur par défaut : Valeur, attribut ou option utilisé si aucune valeur n'est indiquée de façon explicite.

voie d'acheminement : Chemin du trafic de réseau entre l'origine et la destination.

VPN : Abréviation de Virtual Private Network (VPN). Réseau composé d'un ou plusieurs tunnels IP sécurisés reliant deux réseaux ou plus.

W

WAN : Wide Area Network. Réseau fournissant des services de communication dans une zone plus étendue que celle prise en charge par un réseau local ou métropolitain et qui peut être utilisé pour fournir des fonctions de transmission publiques.

WAP : Abréviation de Wireless Application Protocol. Norme internationale ouverte pour les applications utilisant des communications sans fil, par exemple l'accès Internet à partir d'un téléphone mobile.

WAS : WebSphere Application Server.

Web : Réseau de serveurs HTTP contenant des programmes et des fichiers dont la plupart correspond à des documents hypertexte qui contiennent des liens avec d'autres documents des serveurs HTTP. Désigné également par World Wide Web.

WLM : Abréviation de Workload Manager. Conseiller fourni avec Dispatcher. Le conseiller WLM est conçu pour travailler uniquement avec les serveurs sur gros systèmes OS/390 exécutant le composant MVS Workload Manager (WLM).

Index

A

accessibilité xvii
add
 Cisco CSS Controller 430
 Nortel Alteon Controller 450
administration à distance 35, 38, 39, 40
 administration basée sur le Web 261, 263
 RMI 261, 262
administration à distance (basée sur le Web)
 refresh 265
administration basée sur le Web 261, 263
 refresh 265
adresse de non-acheminement
 définition 69, 361
affichage
 compteurs internes 360
 état
 serveurs d'un port 382
 un cluster ou tous les clusters 356
liste
 conseillers fournissant actuellement des mesures 350, 405
numéro de version
 conseiller 351, 405, 406
 gestionnaire 376, 413, 415
paramètres globaux et leurs valeurs par défaut
 pour le gestionnaire 375, 413, 414
 pour un conseiller 351, 404, 405
rapport statistique 374, 412, 413
rapport sur l'état d'un conseiller 350, 403, 405
affinité (maintien de routage)
 affinité de ports croisés 222, 223
 affinité trans ports 378
 cookie actif 224, 225, 387
 cookie passif 224, 226, 387
 délai de maintien de routage 55, 221, 222
 fonctionnement 221
 ID SSL (acheminement CBR) 55
 maintien de routage (substitution d'affinité de port) 219, 391
 masque d'adresse de l'affinité 222
 masque de maintien de routage 222, 223
 mettre au repos maintenant 224, 372, 376
 option de règle 224
 stickymask 379
 stickytime 379, 387
 substitution d'affinité de port 219
 URI 224, 227, 387
affinité d'URI 224, 227, 387
affinité de cookie actif 224, 225, 387
affinité de cookie passif 224, 226, 387
affinité de ports croisés 222
affinité trans ports 378

AIX
 configuration requise 31
 installation 32
ajout
 cluster 356
 port à un cluster 70, 382
 serveur à un port 71, 394, 422
alertes
 contrôleurs 257
 Dispatcher, CBR, Site Selector 184
alias
 carte d'interface réseau 69, 116
 unité de bouclage 72
arrêt
 Cisco CSS Controller 278
 conseiller 350, 404, 406
 exécuteur 361
 gestionnaire 376, 413, 415
 Nortel Alteon Controller 278
assistant, configuration
 CBR 113
 Dispatcher 66
 site Selector 133

B

binlog
 cbrcontrol 352
 dscontrol 352
journal binaire, pour les statistiques des serveurs 352

C

Caching Proxy 107
 configuration pour utiliser CBR 114
carte d'interface réseau
 alias 69
 ethernet (pour Solaris) 67
 mappage (pour Windows) 70
carte d'interface réseau Ethernet
 ibmlb.conf
 configuration Solaris 67
CBR
 alias pour la carte d'interface réseau 116
 avec Caching Proxy
 configuration 118
 connexions SSL 107
 conseiller ssl2http 108
 mot clé mappage 108
 présentation générale 106
 caractères nationaux Latin-1 endommagés (sous Windows) 327
 comportement inattendu de l'interface graphique lors de l'utilisation de cartes vidéo Matrox AGP 326
configuration
 Configuration du poste CBR 113
 présentation des tâches 109

CBR (suite)

déconnexion de l'hôte en cours d'administration Web 327
démarrage et arrêt 277
échec de cbrcontrol 325
échec de la commande cbrcontrol sous Solaris 326
échec de lbadmin 325
erreur de syntaxe ou de configuration 326
exemple de démarrage rapide 99
fonctions appropriées 23
ifconfig, commande 116
incident d'exécution 325
incident lors de la résolution de l'adresse IP en nom d'hôte (Windows) 328
les conseillers et les cibles à contacter marquent tous les serveurs comme étant arrêtés (Windows) 327
mémoire insuffisante pour Java/erreur d'unité d'exécution (HP-UX) 327
non-équilibre des requêtes 326
paramètres de l'équilibrage de charge 180
 tentative du serveur du conseiller 188
planification 105
tableau de résolution des incidents 292
utilisation du composant Dispatcher 55
CBR (Content Based Routing) 5
configuration
 Configuration du poste CBR 113
 présentation des tâches 109
paramètres de l'équilibrage de charge 180
planification 105
tableau de résolution des incidents 292
utilisation 276
utilisation du composant Dispatcher 55
cbrserver
 démarrage 100
ccocontrol, commande consultant 430, 433
 file 435
 help 437
 hôte 443
 invite 429
 mesure 441
 serveur 446
ccoserver
 démarrage 140
 échec du démarrage 301, 331
Cisco CSS Controller
 alertes 257
 co-implantation 243
 commandes 429

- Cisco CSS Controller *(suite)*
 - comportement inattendu de l'interface graphique lors de l'utilisation de cartes vidéo Matrox AGP 332
 - configuration
 - configuration de la machine CSS 152
 - exemple 16
 - présentation des tâches 149
 - configuration matérielle et logicielle requises 143
 - conseiller Workload manager 255
 - conseillers 248
 - consignation binaire pour les statistiques des serveurs 255
 - déconnexion de l'hôte en cours d'administration Web 333
 - démarrage 278
 - démarrage et arrêt 278
 - échec de ccocontrol 331
 - échec de lbadmim 331
 - échec du démarrage 331
 - erreur de connexion de consultant 332
 - exemple de démarrage rapide 139
 - fonctions appropriées 27
 - haute disponibilité 243
 - impossible de créer un registre sur le port 13099 331
 - la commande de régénération n'a pas actualisé la configuration du consultant 332
 - mémoire insuffisante pour Java/erreur d'unité d'exécution (HP-UX) 333
 - Metric Server 253
 - paramètres de l'équilibrage de charge 246
 - planification 143
 - pondérations non actualisées sur le commutateur 332
 - report
 - contrôleur 433
 - tableau de résolution des incidents 295
 - utilisation 278
- clé privée
 - pour authentification distante 262
- clé publique
 - pour authentification distante 262
- clés
 - lbkeys 196, 253, 262
- cluster
 - affichage
 - état de ce cluster 356
 - ajout 356
 - cbrcontrol 353
 - configuration de l'adresse 69
 - définition 69, 356
 - définition du niveau d'importance des informations 72
 - dscontrol 353
 - générique 69
 - proportions 353
 - suppression 356, 425
- cluster générique 69, 356
 - avec Caching Proxy pour le proxy transparent 238
- cluster générique *(suite)*
 - pour combiner les configurations serveurs 236
 - pour équilibrer la charge des pare-feux 237
- co-implantation
 - Cisco CSS Controller 243
 - Nortel Alteon Controller 243
 - remarque sur IPv6 86
- co-implantation d'adresses multiples 71
- co-implanté avec NAT 204
- co-implanter, Load Balancer et client 241
- co-implanter, Load Balancer et le serveur 66, 71, 202, 231, 391, 394
- collecte d'informations 281
- collocated (mot clé) 203, 394
- commande cbrcontrol
 - advisor 347
 - binlog 352
 - cluster 353
 - executor 357
 - file 362
 - help 364
 - host 369
 - logstatus 370
 - manager 371
 - metric 377
 - port 378
 - rule 384
 - serveur 390
 - set 396
 - status 397
- commande ndcontrol
 - highavailability 438
- commande netstat 76
- commande route 76, 77
- commandes
 - cbrcontrol
 - advisor 347
 - binlog 352
 - cluster 353
 - executor 357
 - file 362
 - help 364
 - host 369
 - logstatus 370
 - manager 371
 - metric 377
 - port 378
 - rule 384
 - serveur 390
 - set 396
 - status 397
 - cococontrol
 - consultant 430, 433
 - file 435
 - help 437
 - hôte 443
 - invite 429
 - mesure 441
 - serveurs, configuration 446
- Cisco CSS Controller 429
- dscontrol
 - advisor 347
 - binlog 352
 - cluster 353
 - contrôle du conseiller 71
- commandes *(suite)*
 - dscontrol *(suite)*
 - contrôle du gestionnaire 71
 - définition d'un port 70
 - définition d'un serveur 71
 - définition de l'adresse de non-acheminement 69, 361
 - executor 357
 - file 362
 - haute disponibilité, contrôle 365, 458
 - help 364
 - host 369
 - invite 346
 - logstatus 370
 - manager 371
 - metric 377
 - port 378
 - rule 384
 - serveur 390
 - set 396
 - sous-agent, configuration
 - SNMP 398
 - status 397
- ifconfig 70, 231
 - affectation d'alias à l'unité de bouclage 73
- nalcontrol
 - consultant 450, 453
 - file 455
 - help 457
 - hôte 465
 - invite 449
 - mesure 461
 - serveurs, configuration 463
- ndcontrol
 - haute disponibilité, contrôle 438
- netstat
 - contrôle des adresses IP et des alias 76
- Nortel Alteon Controller 449
- Site Selector 401
- sscontrol
 - advisor 402
 - file 407
 - help 409
 - logstatus 410
 - manager 411
 - mesure 416
 - nameserver 417
 - rule 418
 - server 421
 - set 423
 - sitename 424
 - status 427
- voie d'acheminement
 - suppression d'une voie supplémentaire 76, 77
- composant Cisco CSS Controller
 - caractères nationaux Latin-1 endommagés (sous Windows) 333
- composant Dispatcher
 - adresse de routeur non spécifiée ou non valide pour la méthode port 315
 - adresse IP non résolue correctement sur la connexion éloignée 309

- composant Dispatcher (*suite*)
 - affichage d'un écran bleu lors du démarrage de l'exécuteur 306
 - affichage incorrect de l'interface graphique 306
 - alias renvoyé au lieu de l'adresse locale 310
 - arrêt des processus Load Balancer (Solaris) 316
 - caractères nationaux Latin-1 endommagés (sous Windows) 312
 - charge des demandes non équilibrée 302
 - comportement imprévu lors du chargement d'un fichier de configuration volumineux 308
 - comportement inattendu avec "rmmod ibmlb" 310
 - comportement inattendu de l'interface graphique lors de l'utilisation de cartes vidéo Matrox AGP 310
 - configuration
 - configuration d'un réseau privé 235
 - configuration de la machine Load Balancer 66
 - présentation des tâches 63
 - conflit d'adresses IP lors de l'utilisation de la fonction de haute disponibilité 316
 - connexion à une machine éloignée 304
 - déconnexion de l'hôte en cours d'administration Web 311
 - délai lors du chargement d'une configuration Load Balancer 316
 - démarrage 268
 - démarrage incorrect de l'interface graphique 305
 - disparition des fenêtres d'aide 306
 - dysfonctionnement de MS IIS et de SSL 304
 - dysfonctionnement des conseillers 304
 - échec de dscontrol 304
 - échec de lbadmim 304
 - erreur lorsque Caching Proxy est installé 305
 - évitement du trafic retour avec Load Balancer par la fonction Path MTU Discovery 306
 - fonction CBR (content-based routing) 55
 - fonction haute disponibilité de Load Balancer inopérante en mode réseau étendu 307
 - fonction haute disponibilité inopérante 303
 - haute disponibilité, conseils de configuration 318
 - impossible d'ajouter un signal de présence 303
 - impossible d'ouvrir la fenêtre d'aide 305
 - incident d'exécution 302
- composant Dispatcher (*suite*)
 - incident lors de la résolution de l'adresse IP en nom d'hôte (Windows) 313
 - lbadmim se déconnecte du serveur après mise à jour de la configuration 309
 - les conseillers et les cibles à contacter marquent tous les serveurs comme étant arrêtés (Windows) 313
 - les conseillers ne fonctionnent pas dans une configuration en haute disponibilité après une panne réseau (Windows) 314
 - les demandes client échouent lors de la tentative de renvoi de réponses de grande page 317
 - machines principale et de secours activées en mode haute disponibilité 317
 - mémoire insuffisante pour Java/erreur d'unité d'exécution (HP-UX) 312
 - message d'avertissement Java lors de l'installation de correctifs de service 324
 - mise à niveau de Java fournie avec l'installation 325
 - n'utilisez pas la commande IP address add lorsque vous attribuez un alias sur une unité de bouclage (Linux) 315
 - NAT / NAPT 53
 - non réponse du serveur 302
 - paramètres de l'équilibrage de charge 180
 - délai de rapport du conseiller 187
 - délai du serveur du conseiller 188
 - indice de lissage 184
 - Intervalle conseiller 187
 - intervalles gestionnaire 183
 - pondérations 181
 - proportion de l'importance accordée aux données d'état 180
 - seuil de sensibilité 183
 - tentative du serveur du conseiller 182, 188
 - pas d'enregistrement des charges des serveurs 311
 - planification 51
 - polices de caractères coréennes non souhaitées sous AIX et Linux 309
 - réacheminement MAC 53
 - réinitialisation d'un serveur arrêté 182
 - réinitialisation des serveurs arrêtés 381
 - routes supplémentaires (Windows) 303
 - serveur Web établissant une liaison à 0.0.0.0 311
 - sous Linux, fuite de mémoire lors de l'utilisation du gestionnaire et des conseillers 322
 - sous Linux, limitations d'utilisation de serveurs zSeries ou S/390 320
- composant Dispatcher (*suite*)
 - sous Linux, paquets acheminés par Dispatcher mais non reçus par le serveur dorsal 322
 - sous Linux, possibilité d'échec de synchronisation de HA Dispatcher 318
 - sous Solaris, impossible d'ajouter des serveurs IPv6 à la configuration 324
 - sous Windows, incident lors de la reprise de la haute disponibilité 323
 - sous Windows, le message d'erreur "le serveur ne répond pas" apparaît 318
 - support IPv6 81
 - tableau de résolution des incidents 286
 - temps de réponse important 310
 - transmission d'un cadre impossible 306
 - utilisation 268
 - composants du produit 51
 - configuration
 - CBR (Content Based Routing) 109
 - cbrwizard 113
 - Cisco CSS Controller 149
 - composant Dispatcher 63
 - définition de consultant de commutateur 175
 - dswizard 66
 - fichiers exemple 481
 - haute disponibilité 153, 175, 176
 - lancement du consultant 153, 175
 - mesures 153, 175
 - méthodes
 - assistant (CBR) 113
 - assistant (Dispatcher) 66
 - assistant (Site Selector) 133
 - GUI (CBR) 112
 - interface graphique (Cisco CSS Controller) 151
 - interface graphique (Dispatcher) 64
 - interface graphique (Nortel Alteon Controller) 173
 - interface graphique (Site Selector) 132
 - ligne de commande (CBR) 110
 - ligne de commande (Cisco CSS Controller) 149
 - ligne de commande (Dispatcher) 63
 - ligne de commande (Nortel Alteon Controller) 171
 - ligne de commande (Site Selector) 131
 - scripts (CBR) 111
 - scripts (Cisco CSS Controller) 150
 - scripts (Dispatcher) 64
 - scripts (Nortel Alteon Controller) 172
 - scripts (Site Selector) 132
 - Nortel Alteon Controller 171
 - service 175
 - Site Selector 131

- configuration (*suite*)
 - sswizard 133
 - tâches avancées 179, 201
 - test 154, 176
 - vérification 77
- configuration logicielle requise
 - Cisco CSS Controller 143
 - Nortel Alteon Controller 161
- configuration matérielle requise
 - Cisco CSS Controller 143
 - Nortel Alteon Controller 161
- configuration requise
 - AIX 31
 - HP-UX 35
 - Linux 36
 - solaris 38
 - Windows : 40
- connecttimeout
 - Site Selector 402
- connexions, définition du pourcentage
 - d'importance 181, 356
- connexions SSL
 - configuration de ibmproxy 108
 - conseiller HTTPS 188
 - conseiller SSL 189
 - incident d'activation 304
 - pour CBR 107, 108
- conseiller Caching Proxy 189
- conseiller DB2 190
- conseiller ftp 347, 402
- conseiller http 347, 402
- conseiller personnalisé 192, 250
 - exemple 487
- conseiller ssl2http 108, 189
- conseiller WAS 190, 193
- conseiller Workload manager (WLM) 198, 255
- conseillers
 - cbrcontrol 347
 - composant CBR
 - conseiller ssl2http 189
 - composant Dispatcher 185
 - arrêt 350
 - conseiller Caching Proxy 189
 - conseiller self 190, 191
 - délai d'expiration de réception du serveur 348, 350
 - délai de connexion du serveur 188, 350
 - délai de rapport 187, 349
 - délai de réception du serveur 188
 - démarrage 71, 350
 - démarrage/arrêt 186
 - dépassement du délai de connexion du serveur 347
 - détection d'erreur rapide 188
 - intervalle 187, 350
 - liste 188, 350
 - noms 347
 - personnaliser 192
 - port 354
 - rapport sur l'état de 350
 - report 351
 - tentative du serveur 182, 188, 348
 - version de 351
- contrôleurs 248
 - délai d'inactivité 249

- conseillers (*suite*)
 - contrôleurs (*suite*)
 - délai de connexion du serveur 249
 - délai de réception du serveur 249
 - détection d'erreur rapide 249
 - personnaliser 250
 - tentative du serveur 249
 - demande/réponse de conseiller
 - HTTP 190
 - dscontrol 347
 - exemple de fichier de configuration 487
 - exemple personnalisé 487
 - liste 349
 - option d'URL, conseiller HTTP 190
 - remarque sur IPv6 85
 - restriction sous Solaris 185
 - site Selector
 - délai de connexion du serveur 188
 - délai de réception du serveur 188
 - détection d'erreur rapide 188
 - tentative du serveur 188
- Site Selector
 - arrêt 404, 406
 - délai de connexion du serveur 402, 405
 - délai de rapport 404, 406
 - délai de réception du serveur 403, 405
 - démarrage 403, 405
 - intervalle 402
 - intervalle 405
 - list 402
 - liste 403, 405
 - loglevel 402
 - noms 402
 - port 347, 402
 - rapport sur l'état de 403, 405
 - tentatives du serveur 403
 - version de 405, 406
- sscontrol 402, 409
- conseillers, composant Load Balancer
 - démarrage 71
- consignation binaire pour les statistiques
 - des serveurs 240, 266, 268
 - contrôleurs 255
- consultant
 - cococontrol 430, 433
 - Cisco CSS Controller
 - add 430
 - binarylog 430
 - report 430
 - démarrage 153, 175
 - nalcontrol 450, 453
 - Nortel Alteon Controller
 - add 450
 - binarylog 450
 - report 450
- consultant de commutateur
 - définition 175
- contrôleur
 - Cisco CSS Controller
 - loglevel 431, 433
 - logsize 431, 433
 - report 433

- contrôleur (*suite*)
 - Cisco CSS Controller (*suite*)
 - set 433
 - Nortel Alteon Controller
 - loglevel 451, 453
 - logsize 451, 453
 - report 453
 - set 453
 - pondération fixée 247
- contrôleurs
 - conseiller personnalisé 250
 - paramètres de l'équilibrage de charge
 - délai d'attente du serveur du conseiller 249
 - délai d'inactivité 247
 - délai d'inactivité du conseiller 249
 - importance accordée aux informations de mesure 246
 - pondérations 247
 - seuil de sensibilité 248
 - tentative du serveur du conseiller 249
- conversion d'adresse réseau (NAT) 53

D

- default.cfg 68, 115, 134
- définition
 - adresse de cluster 70
 - adresse de non-acheminement 66, 69, 361
 - cluster 356
 - fréquence d'interrogation de l'exécuteur par le gestionnaire 183, 374
 - indice de lissage 184, 375, 412, 414
 - intervalle de temps
 - interrogation des serveurs par le conseiller 350, 405
 - mise à jour de l'exécuteur par le gestionnaire 183, 374, 411, 413
 - niveau de consignation
 - pour le conseiller 265, 350, 405
 - pour le gestionnaire 411
 - nom du fichier journal 404
 - pour le gestionnaire 412
 - pondération maximale
 - pour les serveurs d'un port spécifique 181, 382
 - pondération pour un serveur 374, 376, 394, 421
 - port à un cluster 70, 382
 - pourcentage d'importance de l'équilibrage de charge 356
 - sensibilité aux mises à jour de pondération 183, 375, 412, 414
 - serveur à un port 71, 394, 422
 - taille maximale du journal
 - pour le conseiller 265, 350, 403, 405
 - pour le gestionnaire 374, 411, 413
- délai d'attente 268, 355, 358, 380
- démarrage
 - CBR 100
 - Cisco CSS Controller 140, 278
 - conseiller 71, 350, 403, 405

- démarrage (*suite*)
 - Dispatcher 47
 - exécuteur 68, 361
 - gestionnaire 71, 375, 412, 414
 - Nortel Alteon Controller 158, 278
 - serveur 68, 69
 - site Selector 124
 - Site Selector 277
 - système Metric Server 279
- démarrage et arrêt
 - CBR 277
 - Dispatcher 268
- désinstallation
 - sous AIX 32
 - sous HP-UX 36
 - sous Linux 37
 - sous Solaris 39
 - sous Windows 40
- détection d'attaque de refus de service 239
 - halfopenaddressreport 382
 - maxhalfopen 381
- diagnostic des incidents
 - adresse de routeur non spécifiée ou non valide pour la méthode port 315
 - adresse IP non résolue correctement sur la connexion éloignée 309
 - affichage d'un écran bleu lors du démarrage de l'exécuteur Load Balancer 306
 - affichage incorrect de l'interface graphique 306
 - alias renvoyé au lieu de l'adresse locale 310
 - arrêt des processus Load Balancer (Solaris) 316
 - caractères nationaux Latin-1 endommagés (sous Windows) 312, 327, 330, 333, 335
 - CBR ne fonctionne pas 325
 - ce message d'erreur s'affiche lorsque l'on tente de visualiser l'aide en ligne 305
 - charges non indiquées par Metric Server 336
 - comportement imprévu lors du chargement d'un fichier de configuration volumineux 308
 - comportement inattendu avec "rmmmod ibmlb" 310
 - comportement inattendu de l'interface graphique lors de l'utilisation de cartes vidéo Matrox AGP 310, 326, 329, 332, 334
 - configuration du serveur de mesures dans une configuration de second niveau 337
 - conflit d'adresses IP lors de l'utilisation de la fonction de haute disponibilité 316
 - déconnexion de l'hôte en cours d'administration Web 311, 327, 330, 333, 335
 - délai lors du chargement d'une configuration Load Balancer 316
 - diagnostic des incidents (*suite*)
 - démarrage incorrect de l'interface graphique 305
 - disparition des panneaux d'aide 306
 - dysfonctionnement de Dispatcher, de Microsoft IIS et de SSL 304
 - échec de la commande cbrcontrol ou lbadm 325
 - échec de la commande cbrcontrol sous Solaris 326
 - échec de la commande ccocontrol ou lbadm 331
 - échec de la commande dscontrol ou lbadm 304
 - échec de la commande nalcontrol ou lbadm 333
 - échec de la commande sscontrol ou lbadm 328
 - échec du démarrage de ssserver sous Windows 329
 - échec du lancement de ccoserver 331
 - échec du lancement de nalserver 333
 - équilibre de charge de Site Selector incorrect 329
 - erreur de connexion de consultant 332, 335
 - erreur de syntaxe ou de configuration 326
 - erreur lors de l'exécution de Dispatcher lorsque Caching Proxy est installé 305
 - évitement du trafic retour avec Load Balancer par la fonction Path MTU Discovery 306
 - fonction haute disponibilité de Dispatcher inopérante 303
 - fonction haute disponibilité de Load Balancer inopérante en mode réseau étendu 307
 - haute disponibilité, conseils de configuration 318
 - impossible d'ajouter un signal de présence 303
 - impossible de créer un registre sur le port 13099 331
 - impossible de créer un registre sur le port 14099 334
 - incident lors de la résolution de l'adresse IP en nom d'hôte (Windows) 313, 328
 - incidents courants et solutions 302, 304, 325, 328, 331, 333, 336
 - IOException Metric Server sous Windows 336
 - journal de Metric Server indique qu'une signature est nécessaire pour accéder à l'agent 337
 - la commande de régénération n'a pas actualisé la configuration du consultant 332, 335
 - la valeur de mesure renvoie -1 après le démarrage de Metric Server 340
 - lbadm se déconnecte du serveur après mise à jour de la configuration 309
- diagnostic des incidents (*suite*)
 - les conseillers et les cibles à contacter marquent tous les serveurs comme étant arrêtés (Windows) 313, 327, 330
 - les conseillers ne fonctionnent pas 304
 - les conseillers ne fonctionnent pas dans une configuration en haute disponibilité après une panne réseau (Windows) 314
 - les demandes client échouent lors de la tentative de renvoi de réponses de grande page 317
 - les requêtes de Dispatcher ne sont pas acheminées 302
 - Load Balancer ne peut pas traiter et transmettre de cadre 306
 - machines principale et de secours activées en mode haute disponibilité 317
 - mémoire insuffisante pour Java/erreur d'unité d'exécution (HP-UX) 312, 327, 330, 333, 336
 - message d'avertissement Java lors de l'installation de correctifs de service 324
 - mise à niveau de Java fournie avec l'installation 325
 - n'utilisez pas la commande IP address add lorsque vous attribuez un alias sur une unité de bouclage (Linux) 315
 - non-équilibre des requêtes 326
 - non exécution de Dispatcher 302
 - non exécution de Site Selector 328
 - non réponse du composant Dispatcher et du serveur 302
 - numéros de port utilisés par CBR 299
 - numéros de port utilisés par Cisco CSS Controller 301
 - numéros de port utilisés par Dispatcher 299
 - numéros de port utilisés par Nortel Alteon Controller 301
 - numéros de port utilisés par Site Selector 300
 - pas d'enregistrement des charges des serveurs 311
 - permutation circulaire non effectuée par Site Selector (Solaris) 328
 - polices de caractères coréennes non souhaitées sous AIX et Linux 309
 - pondérations non actualisées sur le commutateur 332, 335
 - route supplémentaire 303
 - serveur Web établissant une liaison à 0.0.0.0 311
 - sous AIX, résultat de la commande ps -vg altéré 337
 - sous Linux, fuite de mémoire lors de l'utilisation du gestionnaire et des conseillers 322
 - sous Linux, limitations d'utilisation de serveurs zSeries ou S/390 320

- diagnostic des incidents *(suite)*
 - sous Linux, paquets acheminés par Dispatcher mais non reçus par le serveur dorsal 322
 - sous Linux, possibilité d'échec de synchronisation de HA Dispatcher 318
 - sous Solaris, impossible d'ajouter des serveurs IPv6 à la configuration 324
 - sous Solaris, les scripts génèrent des messages de console non souhaités 339
 - sous Windows, incident lors de la reprise de la haute disponibilité 323
 - sous Windows, le message d'erreur "le serveur ne répond pas" apparaît 318
 - sur les systèmes Linux, extraction impossible de valeurs de Metric Server 339
 - temps de réponse important 310
- diagrammes de syntaxe
 - exemples 343
 - lecture 343
 - paramètres 343
 - ponctuation 343
 - symboles 343
- Dispatcher
 - configuration
 - configuration des serveurs dorsaux 72
 - fonctions appropriées 19
- DPID2 271
- dscontrol, commande
 - advisor 347
 - binlog 352
 - cluster 353
 - conseiller 71
 - exécuteur 69
 - executor 357
 - file 362
 - gestionnaire 71
 - help 364
 - highavailability 365, 458
 - host 369
 - invite 346
 - logstatus 370
 - manager 371
 - metric 377
 - port 70, 378
 - réduction des paramètres de commande 346
 - rule 384
 - serveur 71, 390
 - set 396
 - status 397
 - subagent 398
- dsserver
 - démarrage 47

E

- équilibre basé sur des règles 212
 - adresse IP du client 213, 385, 389, 418, 420

- équilibre basé sur des règles *(suite)*
 - Connexions actives d'un port 215, 385
 - contenu de la requête 55, 219
 - Contenu de la requête 385
 - heure 385, 389, 418, 420
 - Heure 214
 - largeur de bande partagée 216, 385, 389
 - largeur de bande réservée 216, 385, 389
 - mesure de tous les serveurs 217
 - metricall 418
 - metricavg 418
 - moyenne des mesures 218
 - nombre de connexions par seconde 215, 385
 - option d'évaluation 220
 - option d'évaluation de serveur 220
 - port du client 214, 385
 - sélection de règles, par composant 212
 - toujours vraie 385, 389, 418, 420
 - toujours vraies 218
 - type de service (TOS) 214, 385, 389
- état, affichage
 - serveurs d'un port spécifique 382
- exécuteur
 - arrêt 361
 - démarrage 361
- executor
 - cbrcontrol 357
 - dscontrol 357
- exemple de démarrage rapide 45
 - CBR 99
 - Cisco CSS Controller 139
 - Nortel Alteon Controller 157
 - site Selector 123
- exemples
 - démarrage 45
 - Démarrage
 - CBR 99
 - Cisco CSS Controller 139
 - Nortel Alteon Controller 157
 - Site Selector 123
 - gestion de serveurs locaux 10, 11, 13, 15, 16
- exemples de fichiers de configuration 481
 - composant Dispatcher (Windows) 484
 - conseiller 487
- exploitation de Load Balancer 261

F

- fichier de mappage d'adresses
 - exemple 236
- fichiers de configuration exemple
 - composant Dispatcher (AIX) 481
- file
 - cbrcontrol 111, 362
 - cococontrol 435
 - dscontrol 64, 362
 - nalcontrol 455
 - sscontrol 132, 407
- Firewall (restriction) 40

G

- gestion de Load Balancer 261
- gestionnaire
 - arrêt 376, 413, 415
 - démarrage 71, 375, 412, 414
 - niveau d'importance des informations 180
 - pondération fixée 182
 - version de 376, 413, 415
- goActive 210
- goIdle 210
- goInOp 210
- goStandby 210
- GRE (Generic Routing Encapsulation)
 - linux 234
 - OS/390 234
 - support de réseau étendu 234
- GUI
 - CBR 112
 - Cisco CSS Controller 151
 - Dispatcher 65
 - instructions générales 469
 - Nortel Alteon Controller 173
 - résolution 306
 - site Selector 132

H

- haute disponibilité 5, 6, 60, 204
 - Cisco CSS Controller 243
 - configuration 153, 175, 176, 205
 - dscontrol 365
 - hôte principal 355
 - linux pour S/390 211
 - NAT, acheminement 210
 - ndcontrol 438
 - Nortel Alteon Controller 243
 - primaryhost 356
 - réciroque 61, 207, 355, 356, 367
 - remarque sur IPv6 85
- scripts 209
 - goActive 210
 - goIdle 210
 - goInOp 210
 - goStandby 210
 - highavailChange 210
- haute disponibilité réciroque 61, 205, 207
 - de relais 209
 - primaryhost 355, 356
 - scripts 209
- help
 - cbrcontrol 364
 - cococontrol 437
 - dscontrol 364
 - nalcontrol 457
- high availability
 - dscontrol 458
- highavailChange 210
- host
 - cbrcontrol 369
 - dscontrol 369
- hôte
 - cococontrol 443
 - nalcontrol 465
- hôte principal 207

HP-UX
 arp publish, commande 70
 configuration requise 35
 installation 35

I

IBM Firewall (restriction) 40
 ibmlb.conf
 configuration Solaris 67
 ibmproxy 108, 114
 ifconfig, commande 70, 73, 116, 231
 indice de lissage, définition 184, 375, 412, 414
 informations, collecte 281
 installation
 Load Balancer 31
 sous AIX 32
 sous HP-UX 35
 sous Linux 37
 sous Solaris 38
 sous Windows 40
 Interface graphique
 CBR 112
 Cisco CSS Controller 151
 Dispatcher 64
 instructions générales 469
 Nortel Alteon Controller 173
 site Selector 132
 intervalle, définition de la fréquence
 interrogation de l'exécuteur par le
 gestionnaire 183, 374
 interrogation des serveurs par le
 conseiller 350, 405
 mise à jour des pondérations par le
 gestionnaire pour l'exécuteur 183, 374, 411, 413

J

journal
 binaire, pour les statistiques des
 serveurs 240
 fichier, définition du nom
 pour le conseiller 404
 pour le gestionnaire 412
 niveau, définition
 pour le conseiller 265, 350, 405
 pour le consultant 267
 pour le gestionnaire 265, 411
 pour le serveur 265, 267
 pour le sous-agent 265
 taille, définition
 pour le conseiller 265, 350, 403, 405
 pour le consultant 267
 pour le gestionnaire 265, 374, 411, 413
 pour le serveur 265, 267
 pour le sous-agent 265, 267
 utilisation des journaux Cisco CSS
 Controller 278, 279
 utilisation des journaux de CBR 277
 utilisation des journaux Load
 Balancer 265

journal (*suite*)
 utilisation des journaux Metric
 Server 279
 utilisation des journaux Site
 Selector 278

L

lbkeys 197, 254, 262
 lbwebaccess 264
 lien explicite 235
 ligne de commande
 envoyer la commande (interface
 graphique) 474
 exemple de configuration
 CBR 100
 Cisco CSS Controller 140
 Dispatcher 47
 Nortel Alteon Controller 158
 site Selector 124
 linux
 configuration requise 36
 haute disponibilité sous S/390 211
 installation 37
 load Balancer
 support IPv6 81
 Load Balancer
 avantages 4
 configuration
 CBR 109
 Cisco CSS Controller 149
 composant Dispatcher 66, 113, 134
 Nortel Alteon Controller 171
 Site Selector 131
 e xemple de démarrage rapide
 Nortel Alteon Controller 157
 exemple de démarrage rapide 45
 CBR 99
 Cisco CSS Controller 139
 site Selector 123
 fonctions 3, 9
 identification des incidents 281
 installation 31
 présentation générale 3, 9
 remarques relatives à la
 planification 51, 127
 tâches de configuration
 avancées 179, 201
 utilisation et gestion 261, 277, 278
 Utilisation et gestion 261
 Load Balancer pour IPv4 et IPv6 81
 activation des paquets IPv6 88
 adresse lien-local 84
 autoconf6, AIX 88
 co-implantation 86
 commandes dscontrol 92
 conseillers, utilisation 85
 création d'un alias pour l'unité de
 bouclage 88
 différences entre les syntaxes des
 commandes 92
 dsconfig 88
 fonctions non prises en charge 84
 haute disponibilité 85
 ifconfig 88
 ip adr 88

Load Balancer pour IPv4 et IPv6 (*suite*)
 Metric Server 87
 modprobe, Linux 88
 prise en charge des plateformes 82
 remarques sur la configuration 84
 logstatus
 cbrcontrol 370
 dscontrol 370
 sscontrol 410

M

maintien de routage (affinité)
 affinité de ports croisés 222, 223
 affinité trans ports 378
 cookie actif 224, 225, 387
 cookie passif 224, 226, 387
 délai de maintien de routage 55, 221, 222
 fonctionnement 221
 maintien de routage (substitution
 d'affinité de port) 219, 391
 masque d'adresse de l'affinité 222
 masque de maintien de routage 222, 223
 mettre au repos maintenant 224, 372, 376
 stickymask 379
 stickytime 379, 387
 substitution d'affinité de port 219
 URI 224, 387
 manager
 cbrcontrol 371
 dscontrol 371
 sscontrol 411
 marquage d'un serveur à l'état
 défaillant 394, 421, 422
 en service 395, 421, 422
 marques 499
 masque d'adresse de l'affinité 222
 Masque d'adresse de l'affinité 379
 mesure
 ccocontrol 441
 nalcontrol 461
 sscontrol 416
 mesures
 configuration 153, 175
 mesures du système
 configuration 377, 416, 441, 461
 définition du pourcentage
 d'importance 181, 246, 353, 354
 méthode d'acheminement
 cbr 55, 57
 mac 53, 54
 mac, nat ou cbr 56, 380
 nat 57
 NAT 53
 méthode d'acheminement cbr 55, 57
 délai de maintien de routage 55
 méthode d'acheminement mac 53
 méthode d'acheminement nat 57
 méthode d'acheminement NAT
 haute disponibilité, scripts 210
 Méthode d'acheminement NAT 53
 metric
 cbrcontrol 377
 dscontrol 377

- Metric Server
 - démarrage et arrêt 279
 - présentation générale 196, 253
 - remarque sur IPv6 87
- migration 31
- mise à l'état de repos d'un serveur 223, 372, 374, 376

N

- nalcontrol, commande
 - consultant 450, 453
 - file 455
 - help 457
 - hôte 465
 - invite 449
 - mesure 461
 - serveur 463
- nalserver
 - démarrage 158
 - échec du démarrage 333
- nameserver
 - sscontrol 417
- NAPT (network address port translation) 53
- NAT, co-implantation de serveur avec 204
- Nortel Alteon Consultant
 - fonctions appropriées 28
- Nortel Alteon Controller
 - alertes 257
 - caractères nationaux Latin-1 endommagés (sous Windows) 335
 - co-implantation 243
 - commandes 449
 - comportement inattendu de l'interface graphique lors de l'utilisation de cartes vidéo Matrox AGP 334
 - configuration
 - configuration de la machine Nortel Alteon Controller 174
 - présentation des tâches 171
 - configuration matérielle et logicielle requises 161
 - conseiller Workload manager 255
 - conseillers 248
 - consignation binaire pour les statistiques des serveurs 255
 - déconnexion de l'hôte en cours d'administration Web 335
 - démarrage et arrêt 278
 - échec de lbadmin 333
 - échec de nalcontrol 333
 - échec du démarrage 333
 - erreur de connexion de consultant 335
 - exemple de démarrage rapide 157
 - haute disponibilité 243
 - impossible de créer un registre sur le port 14099 334
 - la commande de régénération n'a pas actualisé la configuration du consultant 335
 - mémoire insuffisante pour Java/erreur d'unité d'exécution (HP-UX) 336
 - Metric Server 253

- Nortel Alteon Controller (*suite*)
 - paramètres de l'équilibrage de charge 246
 - planification 161
 - pondérations non actualisées sur le commutateur 335
 - report
 - contrôleur 453
 - tableau de résolution des incidents 296
 - utilisation 278
- nouvelles connexions, définition du pourcentage d'importance 181, 354
- nouvelles fonctions, version 6.1
 - configuration de client co-implanté 7
 - conseiller SIP 7
 - prise en charge de HP-UX Itanium 64 bits 7
 - prise en charge de l'espace utilisateur 6
 - prise en charge de Linux zSeries 64 bits 7
 - prise en charge du navigateur Firefox 7
 - prise en charge hors noyau 6

O

- option de menu Contrôler 269
- options de proximité 129
- OS/390
 - support GRE 234

P

- paramètres, affichage de toutes les valeurs globales
 - pour le gestionnaire 375, 413, 414
 - pour un conseiller 351, 404, 405
- paramètres de l'équilibrage de charge (optimisation) 180, 246
- planification
 - CBR 105
 - Cisco CSS Controller 143
 - composant Dispatcher 51
 - Nortel Alteon Controller 161
 - Site Selector 127
- planification de l'installation 3, 9, 51, 127
- pondération
 - contrôleurs 247
 - définition
 - limite pour tous les serveurs d'un port 181, 382
 - pour un serveur 394, 421
 - mode de définition par le gestionnaire 182
- pondération maximale, définition pour les serveurs d'un port spécifique 181, 382
- port
 - cbrcontrol 378
 - dscontrol 378
 - port générique 70, 382
 - conseiller ping 190

- port générique (*suite*)
 - pour acheminer le trafic destiné à un port non configuré 238
 - pour le traitement du trafic FTP 238
- ports
 - affichage
 - état des serveurs sur ce port 382
 - ajout 382
 - définition de la pondération maximale 181, 382
 - définition pour un cluster 70, 382
 - générique 70
 - pour les conseillers 347, 402
 - suppression 382
- pourcentage d'importance pour l'équilibrage de charge, définition 181, 356
- présentation générale
 - configuration de CBR 109
 - configuration de Cisco CSS Controller 149
 - configuration de Nortel Alteon Controller 171
 - configuration de Site Selector 131
 - configuration du composant Dispatcher 63
- primaryhost 356
- prise en charge de l'espace utilisateur 82
- prise en charge hors noyau 82
- proximité réseau 129

R

- rapport d'analyse des statistiques, affichage 374, 412, 413
- redémarrage de tous les serveurs avec des pondérations normalisées 375, 412, 414
- références de commandes
 - comment lire 343
- régénération à distance de la configuration 265
- règle de contenu 55, 219
- remarques 497
- report
 - Cisco CSS Controller 433
 - Nortel Alteon Controller 453
- réseau privé, utilisation avec Dispatcher 235
- résolution, interface graphique 306
- résolution des incidents 281
 - adresse de routeur non spécifiée ou non valide pour la méthode port 315
 - adresse IP non résolue correctement sur la connexion éloignée 309
- affichage d'un écran bleu lors du démarrage de l'exécuteur Load Balancer 306
- affichage incorrect de l'interface graphique 306
- alias renvoyé au lieu de l'adresse locale 310
- arrêt des processus Load Balancer (Solaris) 316

- résolution des incidents (*suite*)
 - caractères nationaux Latin-1 endommagés (sous Windows) 312, 327, 330, 333, 335
 - CBR ne fonctionne pas 325
 - ce message d'erreur s'affiche lorsque l'on tente de visualiser l'aide en ligne 305
 - charges non indiquées par Metric Server 336
 - comportement imprévu lors du chargement d'un fichier de configuration volumineux 308
 - comportement inattendu avec "rmmmod ibmlb" 310
 - comportement inattendu de l'interface graphique lors de l'utilisation de cartes vidéo Matrox AGP 310, 326, 329, 332, 334
 - configuration du serveur de mesures dans une configuration de second niveau 337
 - conflit d'adresses IP lors de l'utilisation de la fonction de haute disponibilité 316
 - déconnexion de l'hôte en cours d'administration Web 311, 327, 330, 333, 335
 - délai lors du chargement d'une configuration Load Balancer 316
 - démarrage incorrect de l'interface graphique 305
 - disparition des panneaux d'aide 306
 - dysfonctionnement de Dispatcher, de Microsoft IIS et de SSL 304
 - échec de la commande cbrcontrol ou lbadm 325
 - échec de la commande cbrcontrol sous Solaris 326
 - échec de la commande ccocontrol ou lbadm 331
 - échec de la commande dscontrol ou lbadm 304
 - échec de la commande nalcontrol ou lbadm 333
 - échec de la commande sscontrol ou lbadm 328
 - échec du démarrage de sserver sous Windows 329
 - échec du lancement de ccoserver 331
 - échec du lancement de nalservice 333
 - équilibre de charge de Site Selector incorrect 329
 - erreur de connexion de consultant 332, 335
 - erreur de syntaxe ou de configuration 326
 - erreur lors de l'exécution de Dispatcher lorsque Caching Proxy est installé 305
 - évitement du trafic retour avec Load Balancer par la fonction Path MTU Discovery 306
 - fonction haute disponibilité de Dispatcher inopérante 303

- résolution des incidents (*suite*)
 - fonction haute disponibilité de Load Balancer inopérante en mode réseau étendu 307
 - haute disponibilité, conseils de configuration 318
 - impossible d'ajouter un signal de présence 303
 - impossible de créer un registre sur le port 13099 331
 - impossible de créer un registre sur le port 14099 334
 - incident lors de la résolution de l'adresse IP en nom d'hôte (Windows) 313, 328
 - incidents courants et solutions 302, 304, 325, 328, 331, 333, 336
 - IOException Metric Server sous Windows 336
 - journal de Metric Server indique qu'une signature est nécessaire pour accéder à l'agent 337
 - la commande de régénération n'a pas actualisé la configuration du consultant 332, 335
 - la valeur de mesure renvoie -1 après le démarrage de Metric Server 340
 - lbadm se déconnecte du serveur après mise à jour de la configuration 309
 - les conseillers et les cibles à contacter marquent tous les serveurs comme étant arrêtés (Windows) 313, 327, 330
 - les conseillers ne fonctionnent pas 304
 - les conseillers ne fonctionnent pas dans une configuration en haute disponibilité après une panne réseau (Windows) 314
 - les demandes client échouent lors de la tentative de renvoi de réponses de grande page 317
 - les requêtes de Dispatcher ne sont pas acheminées 302
 - Load Balancer ne peut pas traiter et transmettre de cadre 306
 - machines principale et de secours activées en mode haute disponibilité 317
 - mémoire insuffisante pour Java/erreur d'unité d'exécution (HP-UX) 312, 327, 330, 333, 336
 - message d'avertissement Java lors de l'installation de correctifs de service 324
 - mise à niveau de Java fournie avec l'installation 325
 - n'utilisez pas la commande IP address add lorsque vous attribuez un alias sur une unité de bouclage (Linux) 315
 - non-équilibre des requêtes 326
 - non exécution de Dispatcher 302
 - non exécution de Site Selector 328
 - non réponse du composant Dispatcher et du serveur 302

- résolution des incidents (*suite*)
 - numéros de port utilisés par CBR 299
 - numéros de port utilisés par Cisco CSS Controller 301
 - numéros de port utilisés par Dispatcher 299
 - numéros de port utilisés par Nortel Alteon Controller 301
 - numéros de port utilisés par Site Selector 300
 - pas d'enregistrement des charges des serveurs 311
 - permutation circulaire non effectuée par Site Selector (Solaris) 328
 - polices de caractères coréennes non souhaitées sous AIX et Linux 309
 - pondérations non actualisées sur le commutateur 332, 335
 - route supplémentaire 303
 - serveur Web établissant une liaison à 0.0.0.0 311
 - sous AIX, résultat de la commande ps -vg altéré 337
 - sous Linux, fuite de mémoire lors de l'utilisation du gestionnaire et des conseillers 322
 - sous Linux, limitations d'utilisation de serveurs zSeries ou S/390 320
 - sous Linux, paquets acheminés par Dispatcher mais non reçus par le serveur dorsal 322
 - sous Linux, possibilité d'échec de synchronisation de HA Dispatcher 318
 - sous Solaris, impossible d'ajouter des serveurs IPv6 à la configuration 324
 - sous Solaris, les scripts génèrent des messages de console non souhaités 339
 - sous Windows, incident lors de la reprise de la haute disponibilité 323
 - sous Windows, le message d'erreur "le serveur ne répond pas" apparaît 318
 - sur les systèmes Linux, extraction impossible de valeurs de Metric Server 339
 - temps de réponse important 310
- RMI (Remote Method Invocation) 35, 38, 39, 40, 261, 262
- route supplémentaire 76, 77
- routes, supplémentaires 76
- routes, suppression des routes supplémentaires 77
- rule
 - cbrcontrol 384
 - dscontrol 384
 - sscontrol 418

S

- sauvegarde, haute-disponibilité 60, 365, 438, 458
- configuration 205

- scripts 209
 - ccoserverdown 257
 - exit utilisateur 184, 257
 - goActive 210
 - goIdle 210
 - goInOp 210
 - goStandby 210
 - highavailChange 210
- scripts d'exit utilisateur 184, 257
 - ccoallserversdown 257
 - ccoserverdown 257
 - ccoserverup 257
 - détection de refus de service 239
 - managerAlert 184
 - managerClear 184
 - nalallserversdown 257
 - naloserverup 257
 - nalserverdown 257
 - serverDown 184
 - serverUp 184
- secure Sockets Layer 71
- sensibilité aux mises à jour de
 - pondération, définition 183, 375, 412, 414
- server
 - sscontrol 421
- serveur
 - adresse 391
 - advisorrequest 393
 - advisorresponse 393
 - ajout 394, 422
 - cbrcontrol 390
 - ccocontrol 446
 - co-implanté avec NAT 204
 - collocated 391, 394
 - cookievalue 391
 - définition de la pondération 394, 421
 - définition sur un port 71, 394, 422
 - dscontrol 390
 - fixedweight 391
 - logique 58
 - mapport 108, 392
 - marqué défaillant 394, 421, 422
 - marqué en service 395, 421, 422
 - mise à l'état de repos 223, 372, 374, 376
 - nalcontrol 463
 - partitionnement 58
 - physiques 58
 - pondération 391
 - protocole 392
 - réactivation 376
 - redémarrage global avec des
 - pondérations normalisées 375, 412, 414
 - réinitialisation d'un serveur
 - arrêté 182
 - returnaddress 392
 - routage non maintenu (substitution d'affinité de port) 391, 394
 - routeur 392
 - suppression 394, 421, 422
- serveur marqué défaillant 394, 421, 422
- serveur marqué en service 395, 421, 422
- serveurs de liaison 70, 71, 185, 231
- service
 - configuration 175
- set
 - cbrcontrol 396
 - dscontrol 396
 - sscontrol 423
- seuil de sensibilité 248
- simple Network Management Protocol (Protocole simplifié de gestion de réseau) 269
- site Selector
 - caractères nationaux Latin-1
 - endommagés (sous Windows) 330
 - comportement inattendu de l'interface graphique lors de l'utilisation de cartes vidéo Matrox AGP 329
 - déconnexion de l'hôte en cours d'administration Web 330
 - équilibre de la charge des machines Dispatcher HA 211
 - exemple de configuration 15
 - exemple de démarrage rapide 123
 - fonctions appropriées 25
 - les conseillers et les cibles à contacter marquent tous les serveurs comme étant arrêtés (Windows) 330
 - mémoire insuffisante pour Java/erreur d'unité d'exécution (HP-UX) 330
 - paramètres de l'équilibrage de charge
 - délai du serveur du conseiller 188
 - tentative du serveur du conseiller 188
- Site Selector
 - commandes 401
 - configuration
 - configuration de la machine 134
 - présentation des tâches 131
 - démarrage et arrêt 277
 - échec de lbadm 328
 - échec de sscontrol 328
 - échec du démarrage de sserver sous Windows 329
 - équilibre de charge incorrect avec des chemins en double 329
 - incident d'exécution 328
 - paramètres de l'équilibrage de charge 180
 - planification 127
 - présentation générale 14
 - tableau de résolution des incidents 293
 - trafic à permutation circulaire à partir des clients Solaris non exécuté 328
 - utilisation 277
- sitename
 - sscontrol 424
- sNMP 269
- SNMP 265
- solaris
 - arp publish, commande 70
 - configuration de la machine Dispatcher 67
 - configuration requise 38
 - installation 38
- sous-agents 265, 269
 - dscontrol 398
- spécifique d'un cluster
 - proportions 424
- sscontrol, commande
 - advisor 402
 - file 407
 - help 409
 - logstatus 410
 - manager 411
 - mesure 416
 - nameserver 417
 - rule 418
 - server 421
 - set 423
 - sitename 424
 - status 427
- sSL 71
- sserver
 - démarrage 124
- status
 - cbrcontrol 397
 - dscontrol 397
- substitution d'affinité de port
 - serveur 219, 391, 394
- soutien de réseau étendu 228
 - exemple de configuration 232
 - linux 234
 - utilisation d'une machine Dispatcher éloignée 229
 - utilisation de conseillers éloignés 230
 - utilisation de GRE 234
- soutien IPv6 81
 - activation des paquets IPv6 88
 - adresse lien-local 84
 - autoconf6, AIX 88
 - co-implantation 86
 - commandes dscontrol 92
 - conseillers, utilisation 85
 - création d'un alias pour la carte NIC 88
 - différences entre les syntaxes des commandes 92
 - dsconfig 88
 - fonctions non prises en charge 84
 - haute disponibilité 85
 - ifconfig 88
 - ip adr 88
 - Metric Server 87
 - modprobe, Linux 88
 - prise en charge des plateformes 82
 - remarques sur la configuration 84
- suppression
 - cluster 356, 425
 - port d'un cluster 382
 - serveur d'un port 394, 421, 422
 - voie d'acheminement supplémentaire 77
- supprimer
 - cluster 356, 425
 - port d'un cluster 382
 - serveur d'un port 394, 421, 422
 - voie d'acheminement supplémentaire 77
- système Metric Server
 - charges non indiquées par Metric Server 336
 - configuration du serveur de mesures dans une configuration de second niveau 337

- système Metric Server (*suite*)
 - IOException Metric Server sous Windows 336
 - journal de Metric Server indique qu'une signature est nécessaire pour accéder à l'agent 337
 - la valeur de mesure renvoie -1 après le démarrage de Metric Server 340
 - sous AIX, résultat de la commande ps -vg altéré 337
 - sous Solaris, les scripts génèrent des messages de console non souhaités 339
 - sur les systèmes Linux, extraction impossible de valeurs de Metric Server 339
 - tableau de résolution des incidents 297
 - utilisation 279

T

- tableaux de résolution des incidents
 - CBR 292
 - Cisco CSS Controller 295
 - composant Dispatcher 286
 - Nortel Alteon Controller 296
 - Site Selector 293
 - système Metric Server 297
- test
 - configuration 154, 176

U

- unité de bouclage
 - alias 72
 - solutions alternatives pour l'affectation d'alias sous Linux 78

V

- vérification
 - voie d'acheminement supplémentaire 76
- version, affichage
 - conseiller 351, 405, 406
 - gestionnaire 376, 413, 415

W

- WAS (WebSphere Application Server)
 - conseiller WAS 190, 193
- Windows :
 - commande de configuration de l'exécuteur 70
 - configuration de la machine Dispatcher 68
 - configuration requise 40
 - installation 40



GC11-2541-00



Spine information:



WebSphere Application Server

Load Balancer - Guide d'administration

Version 6.1

GC11-2541-00