

IBM DB2 Universal Database



Extension Net Search

Guide d'administration et d'utilisation

Version 8.2

IBM DB2 Universal Database



Extension Net Search

Guide d'administration et d'utilisation

Version 8.2

Important

Avant de consulter les informations suivantes et d'utiliser le produit associé, veuillez lire les informations générales contenues à l'Annexe O, «Remarques», à la page 259.

Deuxième édition - mai 2004

Réf. US : SH12-6740-04

LE PRESENT DOCUMENT EST LIVRE "EN L'ETAT". IBM DECLINE TOUTE RESPONSABILITE, EXPRESSE OU IMPLICITE, RELATIVE AUX INFORMATIONS QUI Y SONT CONTENUES, Y COMPRIS EN CE QUI CONCERNE LES GARANTIES DE QUALITE 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.

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*

Remarque : Les captures d'écrans ne sont pas disponibles en français à la date d'impression.

© Copyright IBM France 2004. Tous droits réservés.

© Copyright International Business Machines Corporation 1995, 2004. All rights reserved.

Avis aux lecteurs canadiens

Le présent document a été traduit en France. Voici les principales différences et particularités dont vous devez tenir compte.

Illustrations

Les illustrations sont fournies à titre d'exemple. Certaines peuvent contenir des données propres à la France.

Terminologie

La terminologie des titres IBM peut différer d'un pays à l'autre. Reportez-vous au tableau ci-dessous, au besoin.

IBM France	IBM Canada
ingénieur commercial	représentant
agence commerciale	succursale
ingénieur technico-commercial	informaticien
inspecteur	technicien du matériel

Claviers

Les lettres sont disposées différemment : le clavier français est de type AZERTY, et le clavier français-canadien de type QWERTY.








OS/2 et Windows - Paramètres canadiens

Au Canada, on utilise :

- les pages de codes 850 (multilingue) et 863 (français-canadien),
- le code pays 002,
- le code clavier CF.

Nomenclature

Les touches présentées dans le tableau d'équivalence suivant sont libellées différemment selon qu'il s'agit du clavier de la France, du clavier du Canada ou du clavier des États-Unis. Reportez-vous à ce tableau pour faire correspondre les touches françaises figurant dans le présent document aux touches de votre clavier.

France	Canada	Etats-Unis
 (Pos1)		Home
Fin	Fin	End
 (PgAr)		PgUp
 (PgAv)		PgDn
Inser	Inser	Ins
Suppr	Suppr	Del
Echap	Echap	Esc
Attn	Intrp	Break
Impr écran	ImpEc	PrtSc
Verr num	Num	Num Lock
Arrêt défil	Défil	Scroll Lock
 (Verr maj)	FixMaj	Caps Lock
AltGr	AltCar	Alt (à droite)

Brevets

Il est possible qu'IBM détienne des brevets ou qu'elle ait déposé des demandes de brevets portant sur certains sujets abordés dans ce document. Le fait qu'IBM vous fournisse le présent document ne signifie pas qu'elle vous accorde un permis d'utilisation de ces brevets. Vous pouvez envoyer, par écrit, vos demandes de renseignements relatives aux permis d'utilisation au directeur général des relations commerciales d'IBM, 3600 Steeles Avenue East, Markham, Ontario, L3R 9Z7.

Assistance téléphonique

Si vous avez besoin d'assistance ou si vous voulez commander du matériel, des logiciels et des publications IBM, contactez IBM direct au 1 800 465-1234.

Table des matières

Avis aux lecteurs canadiens.	iii
-------------------------------------	------------

Préface.	ix
A qui s'adresse ce manuel.	ix
Comment utiliser ce manuel.	ix
Comment lire les diagrammes de syntaxe.	ix
Informations connexes.	x
Comment nous faire part de vos commentaires.	xi
Comment prendre contact avec IBM.	xi
Informations produit.	xii

Partie 1. Guide de l'utilisateur 1

Chapitre 1. Présentation et concepts . . 3

Concepts clés	3
Utilisation d'une fonction de recherche scalaire SQL	5
Utilisation d'une recherche de procédure mémorisée	6
Utilisation d'une fonction de valeur table SQL pour la recherche.	7
Concepts supplémentaires	8
Fonction de transformation de colonne	8
Services de l'instance	9
Données stockées en externe	9
Vues d'administration	9
Prise en charge des bases de données partitionnées	9
Index sur alias de type nickname dans une base de données fédérée	9
Fonctions clés	10
DB2 Extension Net Search dans l'environnement client/serveur DB2.	11

Chapitre 2. Installation 13

Configuration système requise	13
Présentation de l'installation pour un serveur DB2 partitionné (UNIX)	14
Installation sur UNIX	14
Etape 1 pour UNIX : Installation des composants du produit.	14
Etape 2 pour UNIX : Mise à jour de l'instance DB2.	15
Installation sous Windows	15
Noms de répertoires et noms de fichiers.	15
Installation des bibliothèques Outside-In	15
Vérification de l'installation	16
Vérification de l'installation sous Windows.	16
Vérification de l'installation sous UNIX	16

Chapitre 3. Scénarios utilisateur. . . . 19

Exemple simple avec la fonction de recherche scalaire SQL	19
---	----

Exemple simple avec utilisation de la mémoire cache et recherche de procédure mémorisée	21
Exemple simple avec fonction de valeur table SQL	22

Chapitre 4. Planification 23

Emplacements des répertoires et stockage des index	23
Formats de document et pages de codes prises en charge	23
Logiciel de filtrage Outside-In	24
Rôles utilisateur	25
Utilisation de la ligne de commande ou de l'interface du Centre de contrôle DB2 pour l'indexation	26

Chapitre 5. Services de l'instance de l'Extension Net Search 27

Démarrage et arrêt de DB2 Extension Net Search.	27
Services de verrouillage	27
Utilisation des services de verrouillage	27
Affichage d'une image instantanée d'un verrou	29
Services de mise à jour	30

Chapitre 6. Création et gestion d'un index de texte. 31

Introduction aux commandes db2text.	31
Activation d'une base de données	32
Désactivation d'une base de données	33
Création d'un index de texte	34
Création d'un index de texte sur des types de données binaires.	36
Création d'un index de texte sur un type de données non pris en charge	37
Création d'un index de texte pour les types de données DATALINK.	37
Installation du fichier jar Data Links	38
Création d'un index de texte sur un alias de type nickname à l'aide de la réplication avec mise à jour incrémentielle de l'index à l'aide de DB2 Replication	38
Création d'un index de texte pouvant être utilisé par la recherche de procédure mémorisée	39
Index de texte sur des vues	43
Remarques sur les performances	45
Gestion des index de texte	45
Mise à jour et réorganisation d'un index de texte	46
Modification d'un index de texte	47
Effacement des événements d'index	48
Suppression d'un index de texte	48
Affichage de l'état de l'index de texte.	49
Sauvegarde et restauration d'index	49
Suppression des fichiers du répertoire /tmp	50

Chapitre 7. Utilisation du Centre de contrôle DB2 51

Démarrage et arrêt des services de l'instance DB2	
Extension Net Search	52
Activation d'une base de données	52
Administration de l'index de texte.	53
Création d'un index de texte	55
Gestion d'un index de texte	65

Chapitre 8. Recherche 71

Recherche de texte à l'aide des fonctions de recherche scalaires SQL	72
Réalisation d'une requête.	72
Recherche et renvoi du nombre d'occurrences trouvées	72
Recherche et renvoi du score d'un document texte trouvé	73
Spécification des arguments de recherche SQL.	73
Recherche de termes dans n'importe quelle séquence	73
Recherche à l'aide des opérateurs booléens AND et OR	74
Recherche floue	74
Recherche de parties d'un terme (caractères génériques)	75
Recherche de termes contenant déjà un caractère générique	75
Recherche de termes dans une séquence fixe	75
Recherche de termes dans la même phrase ou le même paragraphe	76
Recherche de termes dans des sections de documents structurés	76
Recherche à l'aide de l'opérateur booléen NOT	76
Recherche de thésaurus	76
Recherche d'attributs numériques	77
Recherche de texte libre	77
Exemples supplémentaires de syntaxe de recherche	78
Recherche de texte à l'aide d'une recherche de procédure mémorisée	78
Recherche de texte à l'aide de la fonction de valeur table SQL	79
Utilisation de la fonction de mise en évidence (HIGHLIGHT)	80
Recherche sur plusieurs colonnes	82
Remarques sur les performances	82

Chapitre 9. Utilisation de documents structurés 83

Chapitre 10. Utilisation d'un thésaurus pour étendre les termes recherchés . . 85

Structure d'un thésaurus	85
Relations de thésaurus prédéfinies.	86
Définition de vos propres relations	87
Création et compilation d'un thésaurus	87
Création d'un fichier de définitions du thésaurus	87
Compilation d'un fichier de définitions en dictionnaire thésaurus	88

Partie 2. Références 89

Chapitre 11. Commandes d'administration pour le propriétaire d'instance	91
CONTROL	92
START	94
STOP	95

Chapitre 12. Commandes d'administration pour l'administrateur de base de données	97
ENABLE DATABASE	98
DISABLE DATABASE	100
DB2EXTDL (utilitaire)	102
DB2EXTHL (utilitaire)	103

Chapitre 13. Commandes d'administration pour le propriétaire de table de texte 105	
ACTIVATE CACHE	106
ALTER INDEX	108
CLEAR EVENTS	112
CREATE INDEX	114
DEACTIVATE CACHE	128
DROP INDEX	130
DB2EXTTH (Utilitaire)	132
UPDATE INDEX	134
HELP	138
COPYRIGHT	139

Chapitre 14. Syntaxe des arguments de recherche	141
Argument de recherche	142

Chapitre 15. Fonction de recherche scalaire SQL et fonction de valeur	
table SQL	149
Résumé des fonctions de recherche	149
CONTAINS	150
NUMBEROFMATCHES	151
SCORE	152
DB2EXT.TEXTSEARCH	153
DB2EXT.HIGHLIGHT	157

Chapitre 16. Fonction de recherche de procédure mémorisée.	161
DB2EXT.TEXTSEARCH (pour une recherche de procédure mémorisée)	162

Chapitre 17. Prise en charge de documents structurés.	165
Comment un modèle de document décrit des documents structurés.	165
Exemple de modèle de document	165
Modèles de document	166
Zones de texte et attributs de document	167
Modèles de document par défaut.	168

Définition d'un modèle de document pour les documents en texte clair structurés	169
Paramètres des éléments	170
Que se passe-t-il lorsqu'un document GPP est indexé ?	170
Définition d'un modèle de document pour les documents HTML	171
Paramètres des éléments	172
Définition d'un modèle de document pour les documents XML	173
Paramètres des éléments	174
Que se passe-t-il lorsqu'un document XML est indexé ?	176
Définition d'un modèle de document pour les documents filtrés Outside-In	177
Paramètres des éléments	178
Que se passe-t-il lorsqu'un document Outside-In est indexé ?	178

Chapitre 18. Prise en charge du thésaurus 179

Partie 3. Annexes 181

Annexe A. Migration 183

Migration de l'Extension Net Search Version 8.1 vers l'Extension Net Search Version 8.1.x et 8.2	184
Migration de l'Extension Net Search Version 7.2 vers l'Extension Net Search Version 8.1.x	184
Migration de Text Information Extender Version 7.2 vers l'Extension Net Search Version 8.1.x	184

Annexe B. Utilisation de grandes quantités de mémoire. 187

AIX (32 bits et 64 bits)	187
Windows (32 bits)	188
L'environnement d'exploitation Solaris (32 bits et 64 bits)	188
Linux	189
HP-UX (32 bits et 64 bits)	189

Annexe C. Catalogues d'informations DB2 Extension Net Search. 191

Vues pour informations de niveau de base de données	191
Table db2ext.proxyinformation	193
Vues pour les informations de niveau d'index	193
Vue db2ext.textindexes	194
Vue db2ext.indexconfiguration.	196
Vue db2ext.textindexformats	196
Vues de table pour un index de texte	197
Vue d'événements	197
Tables de journalisation, vues et alias de type nickname.	198

Annexe D. CCSID pris en charge . . . 201

CCSID.	201
----------------	-----

Annexe E. Langues prises en charge 207

Annexe F. Messages DB2 Extension

Net Search 209

Messages d'information et d'avertissement	209
Messages d'erreur	209

Annexe G. Référence du modèle de document 229

Définition de type de document pour des modèles de document	229
Sémantique des expressions du pointeur (XPath)	230
Limites des zones de texte et des attributs de document	232
Valeurs de l'attribut de marque Outside-In	232

Annexe H. Moteur de recherche de texte 235

Tokenisation.	235
Mots	235
Phrases	235
Paragraphes	235
Mots vides	235
Langues prenant en charge les mots vides.	236
Configuration	236
Normalisation d'un caractère	237
Utilisation de caractères spéciaux dans un même mot	237

Annexe I. Codes anomalie du moteur de recherche de texte. 239

Annexe J. Identification et résolution des incidents 247

Traçage des erreurs	247
Suppression des objets DB2 sans utiliser les commandes de l'Extension Net Search appropriées.	247
Suppression d'une table	247
Suppression d'une base de donnée	247

Annexe K. Messages Data Link. . . . 249

Annexe L. CCSID pris en charge par le thésaurus 251

CCSID.	251
----------------	-----

Annexe M. Messages renvoyés par les outils du thésaurus. 253

Annexe N. Erreurs système Windows 255

Erreurs système	255
---------------------------	-----

Annexe O. Remarques 259

Marques	261
-------------------	-----

Glossaire 263

Index	267
------------------------	------------

Préface

Le présent manuel décrit comment utiliser l'Extension IBM DB2 Universal Database Net Search afin de préparer et de gérer une base de données DB2 en vue de l'extraction de données texte. Il décrit également comment utiliser les fonctions SQL fournies pour accéder à ces types de données et les manipuler. En intégrant les fonctions de DB2 Extension Net Search dans les instructions SQL de votre programme, vous pouvez créer des programmes d'extraction de texte puissants et polyvalents.

Dans ce manuel, les références à "DB2" font allusion à DB2 UDB.

A qui s'adresse ce manuel

Ce manuel est destiné aux administrateurs de base de données DB2 qui sont familiarisés avec les concepts d'administration, les outils et les techniques DB2.

Il s'adresse également aux programmeurs d'applications DB2 qui connaissent le langage SQL et un ou plusieurs langages de programmation pouvant être utilisés pour les programmes d'application DB2.

Comment utiliser ce manuel

Le présent manuel est structuré de la manière suivante :

«Partie 1. Guide de l'utilisateur»

Cette partie présente DB2 Extension Net Search, décrit son installation et sa configuration et traite de remarques sur la planification. Elle décrit également comment préparer et gérer une base de données DB2 de sorte à pouvoir rechercher du texte.

Lisez cette partie si vous êtes un nouvel utilisateur de DB2 Extension Net Search et souhaitez apprendre à utiliser ses fonctions pour rechercher du texte.

«Partie 2. Références»

Cette partie présente des informations de référence pour les fonctions et les commandes de DB2 Extension Net Search.

Lisez cette partie si vous êtes déjà familiarisé avec les concepts et les tâches de DB2 Extension Net Search, mais souhaitez obtenir des informations sur une fonction ou une commande spécifique de DB2 Extension Net Search.

«Partie 3. Annexes»

Cette partie fournit des informations de référence supplémentaires pour DB2 Extension Net Search. Elle contient des informations sur la migration, l'utilisation de la mémoire, les vues, les modèles de document, les messages et les codes.

Lisez cette partie si vous avez besoin d'informations de référence spécifiques sur DB2 Extension Net Search.

Comment lire les diagrammes de syntaxe

Tout au long de ce manuel, la syntaxe est décrite à l'aide de la structure définie comme suit :

Préface

- Lisez les diagrammes de syntaxe de la gauche vers la droite et de haut en bas, en suivant le chemin de la ligne.

Le symbole ►— indique le début d’une instruction.

Le symbole —► indique que la syntaxe de l’instruction continue sur la ligne suivante.

Le symbole ►— indique qu’il s’agit de la suite de l’instruction de la ligne précédente.

Le symbole —◄ indique la fin d’une instruction.

- Les éléments obligatoires sont indiqués sur la ligne horizontale (le chemin principal).

►—élément obligatoire—◄

- Les éléments facultatifs sont indiqués sous le chemin principal.

►—
|élément facultatif|—◄

- S’il est possible de choisir deux ou plusieurs éléments, ils apparaissent empilés. Si vous *devez* choisir l’un des éléments indiqués, un élément de la pile figure sur le chemin principal.

►—
|choix1 obligatoire
|choix2 obligatoire|—◄

S’il est possible de ne choisir aucun des éléments indiqués, la pile entière figure sous le chemin principal.

►—
|choix1 facultatif
|choix2 facultatif|—◄

Une flèche de répétition au-dessus d’une pile indique que vous pouvez choisir plusieurs options parmi les éléments empilés.

►—
|choix1 facultatif
|choix2 facultatif|—◄

- Les mots clés sont indiqués en lettres majuscules. Leur orthographe doit être totalement respectée. Les variables sont indiquées en lettres minuscules (par exemple, srcpath). Elles représentent des noms ou des valeurs fournis par l’utilisateur dans la syntaxe.
- Si des marques de ponctuation, des parenthèses, des opérateurs arithmétiques ou d’autres symboles de ce genre sont indiqués, vous devez les saisir comme élément de la syntaxe.

Informations connexes

IBM DB2 Universal Database Version 8

- *IBM DB2 Universal Database version 8 - Mise en route* pour DB2 Servers (GC11-1879), pour DB2 Clients (GC11-1876), pour DB2 Connect Personal Edition (GC11-1878), pour DB2 Personal Edition (GC11-1881) et IBM Data Links Manager (GC11-1858-00). Ces manuels expliquent comment planifier, installer, configurer et faire migrer DB2 Universal Database sur la plateforme appropriée.
- *IBM DB2 Universal Database Administration Guide version 8* Planning (SC09-4822), Performance (SC09-4821) et Implementation (SC09-4820). Ces manuels décrivent comment concevoir et implémenter une base de données DB2.
- *IBM DB2 Universal Database Call Level Interface Guide and Reference, Volume 1 Version 8* (SC09-4849). Ce manuel explique comment développer des applications qui accèdent aux bases de données DB2 à l'aide de DB2 Call Level Interface, une interface SQL pouvant être appelée et qui est compatible avec la spécification ODBC de Microsoft ODBC.
- *IBM DB2 Universal Database Call Level Interface Guide and Reference, Volume 2 Version 8* (SC09-4850). Ce manuel explique comment développer des applications qui accèdent aux bases de données DB2 à l'aide de DB2 Call Level Interface, une interface SQL pouvant être appelée et qui est compatible avec la spécification ODBC de Microsoft ODBC.
- *IBM DB2 Universal Database Command Reference Version 8* (SC09-4828). Ce manuel explique comment utiliser l'interpréteur de commandes DB2 et fournit des informations de référence relatives aux commandes DB2.
- *IBM DB2 Universal Database Replication Guide and Reference Version 8* (SC27-1121). Ce manuel explique comment planifier, configurer et gérer un environnement de réplication des données DB2.

IBM DB2 Universal Database Enterprise Server Edition Version 8

- *IBM DB2 Connect - DB2 Connect Enterprise Edition Version 8 - Mise en route* (GC11-1877). Ce manuel explique comment planifier, installer et configurer DB2 Universal Database Enterprise - Extended Edition sur la plateforme appropriée.

Comment nous faire part de vos commentaires

Vos commentaires et suggestions peuvent nous permettre d'améliorer la qualité de la documentation. N'hésitez pas à nous faire part de vos commentaires sur ce manuel ainsi que sur que sur toute autre publication concernant les Extensions DB2. Pour nous faire parvenir vos commentaires, vous pouvez utiliser la méthode suivante :

- Envoyez vos commentaires par courrier électronique à l'adresse suivante : swsddid@de.ibm.com. Veillez à mentionner le titre du manuel, le numéro de la partie du manuel, la version du produit et le cas échéant, l'emplacement spécifique du texte sur lequel vous souhaitez faire des commentaires (par exemple, un numéro de page ou de tableau).

Si vous faites parvenir des informations à IBM, IBM pourra utiliser ou diffuser, de toute manière qu'elle jugera appropriée et sans aucune obligation de sa part, tout ou partie de ces informations.

Comment prendre contact avec IBM

Si vous rencontrez un problème technique, consultez et appliquez les actions suggérées dans le manuel *Troubleshooting Guide* avant de contacter le Service clients DB2. Ce manuel vous conseille sur les informations à rassembler pour aider le Service clients DB2 à mieux vous répondre.

Pour obtenir des informations ou commander des produits DB2, prenez contact avec votre partenaire commercial IBM ou avec un distributeur agréé de logiciels IBM.

Aux Etats-Unis, composez l'un des numéros suivants :

- 1-800-237-5511 pour obtenir le Service clients,
- 1-888-426-4343 pour connaître les options de service disponibles.

Informations produit

Aux Etats-Unis, composez l'un des numéros suivants :

- 1-800-IBM-CALL (1-800-426-2255) pour commander des produits ou obtenir des informations générales.
- Pour commander des manuels, composez le 1-800-879-2755.

<http://www.ibm.com/software/data/db2/>

Les pages Web consacrées à DB2 fournissent des informations récentes sur les nouveautés, les descriptions de produit, le calendrier des formations, etc.

<http://www.ibm.com/software/data/support/>

Les pages Web de support DB2 fournissent un accès aux questions fréquemment posées, aux correctifs, aux manuels et aux informations techniques les plus récentes concernant DB2.

Remarque : Il peut arriver que ces informations ne soient disponibles qu'en langue anglaise.

<http://www.ibm.com/software/data/db2/extendors/>

Les pages Web Extensions DB2 fournissent des informations sur toutes les Extensions DB2 couramment disponibles. Ceci inclut DB2 Extension XML, DB2 Extension Spatiale et DB2 Extension AIV.

<http://www.ibm.com/software/data/db2/extendors/support/>

Les pages Web de support Extensions DB2 fournissent un accès aux questions fréquemment posées, aux conseils et astuces, aux correctifs et à la documentation.

<http://www.ibm.com/software/data/db2/extendors/netsearch/index.html>

La page de DB2 Extension Net Search fournit des conseils relatifs aux performances.

www.elink.ibm.com/public/applications/publications/cgibin/pbi.cgi

Le Centre de publications fournit des informations pour commander ou télécharger les publications.

<http://www.ibm.com/certify/index.html>

Le programme de certification professionnelle (Professional Certification Program) du site Web IBM fournit les informations de test de certification d'un grand nombre de produits IBM, y compris DB2.

Sur Compuserve : GO IBMDB2

Entrez cette commande pour accéder aux forums IBM DB2 Family. Tous les produits DB2 sont supportés via ces forums.

Pour savoir comment contacter IBM en dehors des Etats-Unis, reportez-vous à l'annexe A du manuel *IBM Software Support Handbook*. Pour accéder à ce document, consultez la page Web suivante :

<http://techsupport.services.ibm.com/guides/contacts.html>

Remarque : Dans certains pays, les distributeurs agréés IBM peuvent contacter leur centre d'assistance au lieu de prendre contact avec le centre de support IBM.

Préface

Partie 1. Guide de l'utilisateur

Chapitre 1. Présentation et concepts

DB2 Extension Net Search version 8.1 fait partie de la famille des Extensions DB2*.

Elle remplace DB2 Text Information Extender version 7.2 et l'Extension Net Search version 7.2 et offre aux utilisateurs et aux programmeurs d'applications une méthode rapide, polyvalente et intelligente de recherche dans des documents texte stockés dans DB2, dans d'autres bases de données et dans des systèmes de fichiers à l'aide de requêtes SQL.

Concepts clés

Pour maîtriser totalement les fonctionnalités de DB2 Extension Net Search, il est nécessaire connaître certains termes clés qui figurent en **gras** dans cette section, ainsi que les diverses options disponibles. Il est également nécessaire d'avoir des connaissances de base sur les concepts et les termes propres à DB2.

En principe, DB2 Extension Net Search effectue la recherche dans les **documents texte** contenus dans la colonne d'une table de base de données.

Les documents texte doivent être identifiables de manière unique. Pour ce faire, l'Extension Net Search utilise la **clé primaire** de la table.

Les documents peuvent être de formats divers, tels que HTML ou XML.

Plutôt que de rechercher de manière séquentielle dans les documents texte, ce qui nécessiterait énormément de temps, l'Extension Net Search crée un **index de texte** permettant la recherche dans les documents.

Un index de texte est constitué des **termes** importants, extraits des documents de texte.

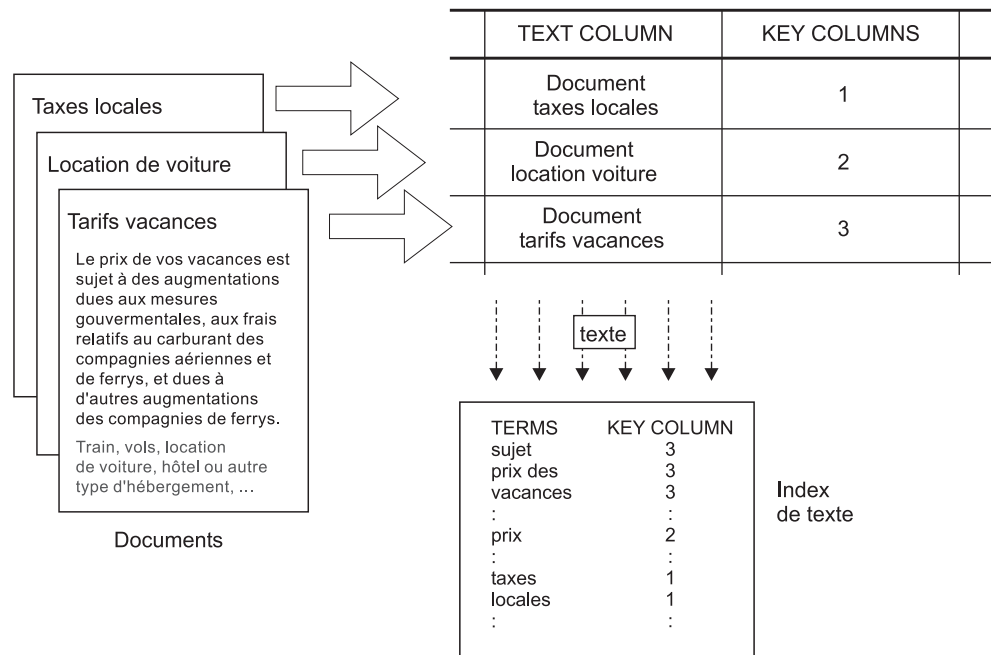


Figure 1. Création d'un index de texte

La **création d'un index de texte** est le processus qui consiste à définir et à déclarer les propriétés de l'index, comme, par exemple, son emplacement. Une fois créé, l'index de texte ne contient pas de données. La **mise à jour de l'index** est le processus qui consiste à ajouter des données dans l'index de texte. La première mise à jour de l'index ajoute tous les documents texte de la colonne de texte à l'index. Cette opération est appelée **mise à jour initiale**.

L'utilisation d'un index de texte pour la recherche engendre des problèmes de synchronisation entre la table et l'index de texte qui doivent être pris en considération, du fait que toutes les modifications de suivi apportées à la table (additions, suppressions et mises à jour de documents texte) doivent être reflétées dans l'index de texte.

La synchronisation dans l'Extension Net Search est basée sur des **déclencheurs** qui stockent automatiquement les informations relatives aux documents nouveaux, modifiés et supprimés dans une **table de journalisation**. Chaque index de texte dispose d'une table de journalisation. L'opération consistant à appliquer le contenu de la table de journalisation à l'index de texte correspondant est appelée **mise à jour incrémentielle**.

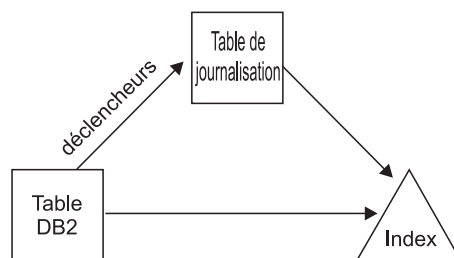


Figure 2. Processus de mise à jour incrémentielle

L'index de texte peut être mis à jour à l'aide d'une option **manuelle** ou **automatique**. L'option automatique utilise une programmation de la mise à jour pour définir les jours et les heures.

Notez qu'**aucune** de ces options ne synchronise l'index de texte dans le cadre d'une transaction de mise jour, de suppression et d'insertion de documents texte. L'indexation de texte asynchrone de l'Extension Net Search améliore les performances et les accès concurrents. La mise à jour est appliquée dans une transaction distincte pour copier une toute petite partie de l'index. Ce dernier est seulement verrouillé à l'accès en lecture pendant une très courte durée lorsque la copie est mise à la place de l'original. Ceci est transparent pour les opérations de recherche. Pour plus d'informations, reportez-vous au Chapitre 5, «Services de l'instance de l'Extension Net Search», à la page 27.

Un index de texte présente certaines propriétés, telles que des propriétés de mise à jour automatique et d'emplacement du fichier d'index. Si nécessaire, vous pouvez modifier certaines propriétés. Cette opération est également connue sous le nom de **modification** de l'index.

Une telle propriété existe si la clause ORDER BY doit pré-trier l'index de texte sur les colonnes de la table. Dans ce cas, la mise à jour initiale indexera le document texte dans l'ordre spécifié et renverra les résultats de la recherche dans cet ordre.

Par exemple, la spécification de résumés de livres pré-triés en fonction du prix du livre. Lors de la recherche des livres les moins chers relatifs aux systèmes de base de données relationnelle, vous pouvez limiter la recherche de texte afin qu'elle ne revoie que les deux premiers livres qui seront les meilleurs marchés. Cependant, sans index pré-trié, vous devez rechercher tous les livres et les associer aux livres les moins chers, ce qui est une opération beaucoup plus coûteuse.

L'Extension Net Search autorise plusieurs index pré-triés par colonne de texte. Par exemple, un index pour le pré-tri des livres en fonction de leur date de publication et un second pour le pré-tri des livres en fonction de leur prix.

Généralement, la première mise à jour qui suit la création d'un index de texte est une mise à jour initiale et les suivantes sont incrémentielles. Cependant, lors de l'utilisation d'index pré-triés, vous devez conserver l'ordre en cas de mise à jour. L'option **Recreate Index on Update** s'en charge, reconstruisant totalement l'index chaque fois qu'une mise à jour est effectuée.

Une fois l'index de texte mis à jour, la recherche peut s'effectuer à l'aide de l'une des options suivantes :

- une fonction de recherche scalaire SQL,
- une recherche de procédure mémorisée,
- une fonction de valeur table SQL.

Comme les options de recherche présentent des caractéristiques fonctionnelles différentes, elles sont expliquées dans les sections suivantes.

Utilisation d'une fonction de recherche scalaire SQL

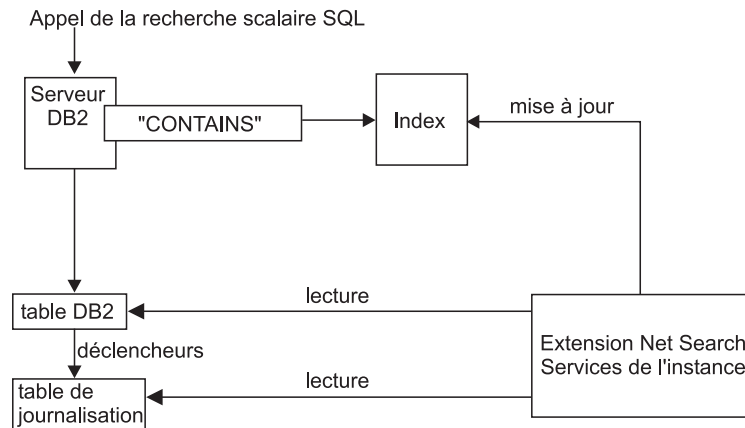


Figure 3. Utilisation d'une fonction de recherche scalaire SQL

L'Extension Net Search propose trois fonctions de recherche de texte scalaires (CONTAINS, NUMBEROFMATCHES et SCORE) qui sont totalement intégrées dans SQL. Vous pouvez utiliser les fonctions de recherche aux mêmes endroits que les expressions SQL standard dans les requêtes SQL. L'exemple suivant montre des requêtes typiques :

```
SELECT * FROM books WHERE CONTAINS (abstract, "relational databases") = 1
AND PRICE <10
```

```
SELECT ISBN, SCORE (abstract, "relational databases") as SCORE
from BOOKS
where NUMBEROFMATCHES (abstract, "relational databases")
>5 AND PRICE <10
order by SCORE
```

Les fonctions scalaires SQL renvoient un indicateur spécifiant avec quelle exactitude les documents texte correspondaient à une condition de recherche de texte donnée. Ensuite, la clause SELECT de la requête SQL détermine les informations renvoyées à l'utilisateur final.

Utilisez les fonctions de recherche scalaires comme méthode de recherche par défaut. Ces fonctions de recherche doivent être adaptées dans la majorité des cas, en particulier lorsque l'expression de recherche de texte est associée à d'autres conditions.

Notez que l'optimiseur DB2 est conscient du nombre de documents texte devant correspondre au prédicat CONTAINS et de la différence de coût entre chaque alternative de plan d'accès. L'optimiseur choisira le plan d'accès le moins coûteux.

Utilisation d'une recherche de procédure mémorisée

La recherche de procédure mémorisée fonctionne différemment de la recherche scalaire SQL. Lors de la création de l'index de texte, vous devez spécifier les colonnes de la table ou de la vue qui sont renvoyées à l'utilisateur final. Ces données sont stockées dans une **mémoire cache** dans la mémoire principale. Ceci permet à la recherche de procédure mémorisée de renvoyer les résultats de la recherche de manière extrêmement rapide. La mémoire cache doit être **activée** pour pouvoir être utilisée. Il existe également une commande de **désactivation** correspondante.

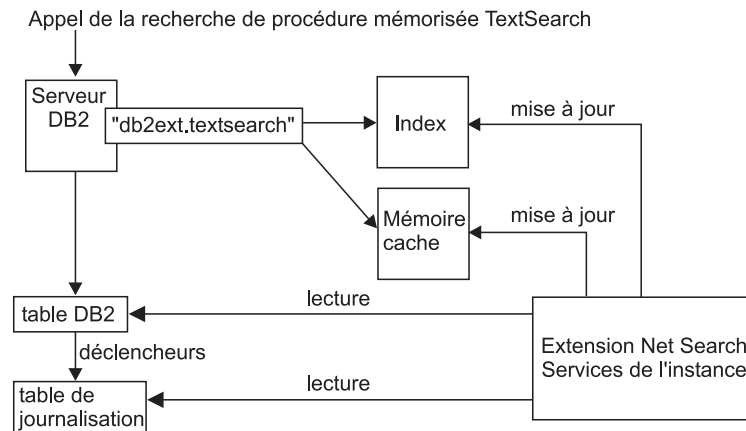


Figure 4. Utilisation d'une recherche de procédure mémorisée

La commande `ACTIVATE` charge les données dans une mémoire cache temporaire (totalement créée lors de l'activation) ou dans une mémoire cache persistante qui est gérée sur le disque.

L'utilisation de la procédure mémorisée pour la recherche nécessite des calculs de mémoire, tels que la quantité de mémoire requise et la quantité de mémoire disponible devant être prévue pour les mises à jour de l'index.

La procédure mémorisée peut fonctionner sur des index de texte créés sur des vues. Cependant, du fait que des déclencheurs ne peuvent pas être créés sur des vues, les modifications ne sont pas automatiquement reconnues. Vous devez ajouter manuellement les informations modifiées à la table de journalisation ou utiliser l'option `RECREATE`.

Utilisez la recherche de procédure mémorisée pour les applications fortement performantes/évolutives qui sont intéressées par les requêtes de recherche de texte uniquement. Par exemple, des requêtes ne nécessitant pas de joindre les résultats de la recherche de texte à ceux d'autres conditions SQL complexes.

Les principales différences fonctionnelles avec les fonctions de recherche scalaires SQL sont les suivantes :

- La recherche de procédure mémorisée ne peut pas être utilisée dans des requêtes SQL arbitraires, mais constitue une requête sur une table de mémoire cache prédéfinie.
- La recherche de procédure mémorisée peut exploiter les index sur des vues.
- La recherche de procédure mémorisée peut exploiter plusieurs index de texte pré-triés sur une colonne.

Notez que pour cette option, une grande quantité de mémoire principale doit être disponible. Pour plus d'informations, reportez-vous à l'Annexe B, «Utilisation de grandes quantités de mémoire», à la page 187.

Utilisation d'une fonction de valeur table SQL pour la recherche

Concepts clés

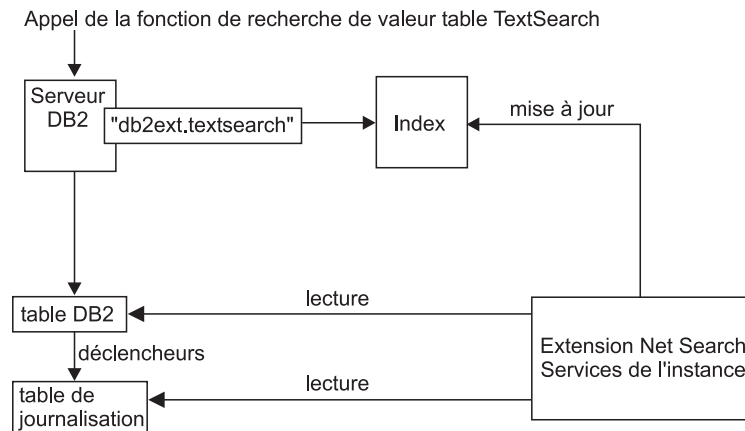


Figure 5. Utilisation d'une fonction de valeur table SQL pour la recherche

La fonction de valeur table SQL est un compromis entre les fonctions de recherche scalaires SQL et la recherche de procédure mémorisée. La fonction de valeur table SQL permet d'utiliser une fonction `db2ext.highlight` pour obtenir des informations sur les raisons pour lesquelles un document a été qualifié de résultat de la recherche.

Les principales différences fonctionnelles avec la recherche de procédure mémorisée sont les suivantes :

- Aucune mémoire cache n'est nécessaire (et aucune mémoire cache n'est exploitée).
- La fonction de valeur table peut être utilisée dans des instructions SQL arbitraires.
- Une grande quantité de mémoire principale n'est pas nécessaire.

La principale différence fonctionnelle avec les fonctions de recherche scalaires SQL est la suivante :

- La fonction de valeur table SQL peut exploiter les index sur des vues.

Utilisez la fonction de valeur table SQL dans les cas où vous utiliseriez normalement une fonction scalaire SQL, alors que vous souhaitez exploiter des index de texte sur des vues.

Concepts supplémentaires

Outre les concepts clés de l'Extension Net Search, vous devez également comprendre certains concepts supplémentaires ; ceux-ci sont expliqués dans cette rubrique.

Pour plus d'informations sur le développement des applications basées sur l'Extension Net Search, reportez-vous au Chapitre 4, «Planification», à la page 23.

Fonction de transformation de colonne

Vous pouvez utiliser votre propre fonction pour convertir un format ou un type de données non pris en charge en un format ou un type de données pris en charge. En spécifiant une fonction définie par l'utilisateur (fonction UDF), vous pouvez obtenir le document texte original comme entrée. La sortie de la fonction UDF doit être un format pris en charge pouvant être traité pendant l'indexation.

Vous pouvez également utiliser cette fonction pour indexer des documents stockés dans des magasins de données externes non pris en charge. Dans ce cas, la colonne DB2 contient des références de documents et la fonction renvoie le contenu du document portant la référence concernée.

Reportez-vous à la section «Création d'un index de texte sur un type de données non pris en charge» à la page 37.

Services de l'instance

Les services de l'instance de l'Extension Net Search prennent soin des services de verrouillage spécifiques à l'index et des services de mise à jour de l'index de texte (automatique et manuelle).

Pour plus d'informations, reportez-vous au Chapitre 5, «Services de l'instance de l'Extension Net Search», à la page 27.

Données stockées en externe

Dans la majorité des cas, les données sur lesquelles vous créez un index de texte sont stockées dans des colonnes natives de tables DB2, telles que CLOBs ou VARCHARs.

Cependant, les documents texte stockés en externe, dans des fichiers ou d'autres bases de données, par exemple, sont également pris en charge. La fonction DB2 Data Links est disponible pour les documents stockés dans des fichiers. Pour les documents stockés dans d'autres bases de données, utilisez des tables d'alias de type nickname DB2 pour créer un index de texte.

Reportez-vous à la section «Informations connexes» à la page x.

Vous pouvez également utiliser la fonction de transformation de colonne pour les données stockées dans des magasins de données externes non pris en charge. Reportez-vous à la section «Fonction de transformation de colonne» à la page 8.

Vues d'administration

Plusieurs vues sont disponibles dans DB2 Extension Net Search. Elles fournissent des informations sur les index de texte et leurs propriétés.

Pour plus d'informations, reportez-vous à l'Annexe C, «Catalogues d'informations DB2 Extension Net Search», à la page 191.

Prise en charge des bases de données partitionnées

Les fonctions de recherche de DB2 Extension Net Search utilisent la prise en charge de bases de données partitionnées comme suit :

- La fonction scalaire SQL exploite les index créés sur une table partitionnée.
- La recherche de procédure mémorisée et la fonction de valeur table SQL exploitent uniquement des tables sur un noeud dans un environnement partitionné.

Index sur alias de type nickname dans une base de données fédérée

Vous pouvez également créer un index de texte sur des alias de type nickname dans une base de données fédérée renvoyant à des tables dans une base de données éloignée. Dans ce cas, le rôle de la table de journalisation (pour les mises

à jour incrémentielles d'index) est différent de son rôle pour un index sur une table normale. Contrairement aux tables normales, les déclencheurs DB2 ne peuvent pas être créés sur des alias de type nickname, par conséquent, les informations de modification ne peuvent pas être insérées dans une table de journalisation utilisant des déclencheurs. Ainsi, les mises à jour incrémentielles peuvent créer un index sur un alias de type nickname de deux manières différentes :

- La table de journalisation est créée en local dans la base de données fédérée et l'application doit veiller à ce que la table de journalisation contienne des informations de modification correctes sur l'alias de type nickname. Pour les vues DB2, le cas est similaire à la mise à jour incrémentielle d'index. Il s'agit de l'option par défaut.
- DB2 Replication a été défini pour que les modifications à la table référencées par l'alias de type nickname soient capturées dans une "Table de modification des données" (table CD) pour des bases de données DB2 éloignées ou dans une "Table de modification cohérente des données" (table CCD) pour des bases de données relationnelles non DB2. DB2 Extension Net Search peut alors utiliser la table CD ou la table CCD au lieu de créer une table de journalisation pour un index sur un alias de type nickname. Dans ce cas, vous devez spécifier les caractéristiques de la table de capture dans la commande DB2TEXT CREATE INDEX.

Fonctions clés

DB2 Extension Net Search version 8.1 présente les fonctions clés suivantes :

- Indexation
 - Indexation rapide de très larges volumes de données
 - Mise à jour dynamique des index
 - Stockage des colonnes de table dans la mémoire principale au moment de l'indexation pour éviter les opérations de lecture physiques coûteuses au moment de la recherche
 - Choix de la ligne de commande ou de l'interface via le centre de contrôle DB2 pour l'indexation
 - Différents formats de texte, par exemple HTML et XML
 - Prise en charge de logiciels de filtrage tiers "Outside In"
 - Prise en charge de table des alias de type nickname
 - Prise en charge de DB2 Data Link Manager
 - Prise en charge des index de texte pré-triés
 - Prise en charge des bases de données partitionnées.
- Recherche
 - Opérations booléennes.
 - Recherche de proximité des mots dans la même phrase ou le même paragraphe.
 - Recherches de "correspondance partielle" de mots ayant une orthographe similaire à celle du terme recherché.
 - Recherches approximatives de mots entiers ou de caractères uniques, à l'aide de caractères génériques masquant le début, le milieu ou la fin.
 - Recherches de texte libre. Pour des documents contenant un texte spécifique, les critères de recherche sont exprimés en langage naturel.
 - Une fonction de mise en évidence pour montrer pourquoi un document particulier a été qualifié de résultat de la recherche.
 - Prise en charge du thésaurus.

- Recherche limitée aux sections dans les documents.
- Prise en charge d'attributs numériques.
- Recherche ultra rapide dans un grand nombre de documents texte comportant de nombreux utilisateurs concurrents.
- Résultats de la recherche
 - Il est possible de spécifier le mode de tri des résultats de la recherche au moment de l'indexation.
 - Vous pouvez spécifier des sous-ensembles de résultats de la recherche lorsque de grands volumes de données sont recherchés et que de longues listes de résultats sont attendues.
 - Vous pouvez définir une limite sur les termes recherchés à l'aide d'un nombre d'occurrences élevé.
 - Les fonctions SQL intégrées, associées à l'Optimiseur DB2 sélectionnent automatiquement le meilleur plan en fonction des résultats de la recherche attendus.

DB2 Extension Net Search dans l'environnement client/serveur DB2

La fonctionnalité de recherche de DB2 Extension Net Search est intégrée dans le langage SQL et exécutée sur le serveur. De ce fait, il n'est pas nécessaire d'installer l'Extension Net Search sur le client pour émettre des requêtes de recherche de texte.

DB2 Extension Net Search prend en charge les appels d'administration envoyés au serveur du côté client. Vous pouvez installer DB2 Extension Net Search du côté client et du côté serveur ou utiliser le Centre de contrôle DB2 pour gérer DB2 Extension Net Search du côté client.

Chapitre 2. Installation

Ce chapitre explique comment installer DB2 Extension Net Search sur des systèmes UNIX et Windows.

Si vous souhaitez migrer de l'Extension Net Search Version 8.1 à la Version 8.2, reportez-vous à l'Annexe A, «Migration», à la page 183.

Après l'installation, exécutez le script de vérification de l'installation DB2 Extension Net Search.

Configuration système requise

Les versions de logiciels suivantes sont requises pour exécuter DB2 Extension Net Search :

- DB2 Version 8.1 Fix Pack 4
- Java Runtime Environment (JRE). La version de JRE dépend de la version de DB2.

DB2 Extension Net Search est disponible sous les systèmes d'exploitation suivants :

Environnement d'exploitation Solaris :

- Solaris 7
- Solaris 8
- Solaris 9

AIX :

- AIX Version 4.3.3. L'ensemble des fichiers suivants est également requis : xlc.aix43.rte 6.0.0.5.
- AIX Version 5.1.0/5.2.0. L'ensemble des fichiers suivants est également requis pour la prise en charge du code à 64 bits : xlc.aix50.rte 6.0.0.5 ou supérieur.

Notez que vous pouvez télécharger des ensembles de fichiers AIX à partir du site Web suivant : <http://techsupport.services.ibm.com/server/fixed>

Windows :

- Windows NT Version 4
- Windows 2000
- Windows XP

HP :

- HP-UX 11i

Linux :

- Linux sur Intel x86 (32-bits) prend en charge DB2 Extension Net Search.
- Linux sur les machines zSeries supporte DB2 Extension Net Search.

Remarque

L'état de la validation des nouveaux noyaux et distributions Linux est fréquemment mis à jour. Pour obtenir les toutes dernières informations relatives aux niveaux du logiciel Linux pris en charge, reportez-vous au site : <http://www.ibm.com/software/data/db2/linux/validate>

Pour AIX, Solaris et HP-UX, DB2 Extension Net Search est disponible sous forme d'application 32 bits ou 64 bits. Sous Linux, Extension Net Search est disponible sous forme d'application 32-bits sur Intel x86 et sous forme d'application 31-bits sur zSeries.

Pour tous les systèmes d'exploitation DB2 Extension Net Search, la capacité mémoire minimale requise est de 30 Mo. L'espace disque minimal requis pour une installation DB2 Extension Net Search standard est de 50 Mo.

Les configurations logicielle et matérielle minimales requises pour DB2 Extension Net Search sont identiques à celles de DB2 Universal Database Version 8.1. Pour connaître ces conditions, ainsi que les modules de correction spécifiques au système d'exploitation, reportez-vous à la section correspondante dans le manuel *IBM Universal Database - Mise en route version 8*.

Toute configuration matérielle supplémentaire requise dépend de la taille et du type de l'index de texte sélectionné. Pour la documentation IBM, reportez-vous à la section «Informations connexes» à la page x.

Présentation de l'installation pour un serveur DB2 partitionné (UNIX)

Assurez-vous que l'installation et la configuration de DB2 est correcte sur tous les noeuds. Après avoir installé DB2, vous devez installer DB2 Extension Net Search sur chaque ordinateur.

Remarque

Un ID utilisateur isolé, différent de l'ID propriétaire d'instance ne fonctionne pas avec des bases de données partitionnées.

Installation sur UNIX

Pour procéder à une installation sur UNIX, effectuez les étapes suivantes :

1. Installation du produit.
2. Mise à jour l'instance DB2.

Etape 1 pour UNIX : Installation des composants du produit

Pour une installation sur UNIX, procédez comme suit :

1. Connectez-vous en tant qu'utilisateur root sur la machine cible.
2. Accédez au répertoire correspondant à votre plateforme :
 - `cd /<cdrom>` où `<cdrom>` correspond au chemin d'accès à votre lecteur de CD-ROM.
 - `cd /<platform>`

3. Appelez `./nsetup.sh` et suivez les instructions affichées sur l'écran.
Remarque : ignorez tous les messages 'Exited with'.

Etape 2 pour UNIX : Mise à jour de l'instance DB2

Pour mettre à jour l'instance DB2, procédez comme suit :

1. Vérifiez que vous êtes actif en tant qu'utilisateur root.
2. Selon la plateforme, utilisez l'une des commandes suivantes :
 - Pour AIX : `cd /usr/opt/db2_08_01/instance`
 - Pour Solaris, Linux, HP-UX : `cd /opt/IBM/db2/V8.1/instance`
3. Exécutez `db2iupdt` à l'aide de la commande `./db2iupdt <db2instance>`, où `<db2instance>` correspond à l'ID utilisateur d'une instance DB2 existante que vous souhaitez utiliser avec l'Extension Net Search.
4. Déconnectez-vous.

Remarque

Extension Net Search crée automatiquement de nouvelles instances DB2 pendant `db2icrt`.

Installation sous Windows

Pour procéder à une installation sous Windows, vous devez vous connecter avec un ID utilisateur doté des droits d'administrateur. Pour ce faire, effectuez les étapes suivantes :

1. Utilisez `<cdrom>:\windows\install\setup.exe` pour transférer les fichiers du module vers la machine cible. Notez que pour chaque service DB2, vous devez entrer un ID utilisateur et le mot de passe correspondant pour créer le service DB2 Extension Net Search adéquat.
2. Réinitialisez le système après le transfert des données.
3. Appelez la commande `db2text start` pour démarrer les services de l'instance DB2 Extension Net Search.

Remarque

Toutes les instances DB2 créent un service Windows. Assurez-vous que les services de l'instance DB2 s'exécutent sous un compte utilisateur, et non sous le compte système.

DB2 Extension Net Search ne prend pas en charge Microsoft Cluster Server.

Noms de répertoires et noms de fichiers

Vous devez spécifier les noms de répertoires et de fichiers en caractères SBCS pour toutes les commandes de l'Extension Net Search. La longueur maximale des noms de chemin d'accès (y compris le nom de fichier) ne doit pas dépasser 256 octets.

Installation des bibliothèques Outside-In

Pour utiliser l'Extension Net Search avec le logiciel Outside-In de Stellent, vous devez configurer les bibliothèques pour chaque plateforme :

Installation des bibliothèques Outside-In

- Sous Windows, assurez-vous que le répertoire où sont situés les bibliothèques est ajouté à la variable d'environnement PATH.
- Sous UNIX, ajoutez toutes les bibliothèques Outside-In dans le répertoire /opt/IBM/db2/V8.1/lib.

Vérification de l'installation

La vérification de l'installation de l'Extension Net Search est possible sur les plateformes Windows et UNIX.

Remarque

Pour les bases de données réparties, utilisez l'exemple de vérification suivant :

```
nom_base de données répartie nsesample [numéro_noeud]  
[nom du fichier_espace_table]
```

Vérification de l'installation sous Windows

Effectuez la procédure ci-dessous pour vérifier que l'Extension Net Search est correctement installée.

- Suivez ces étapes pour appeler le script d'administration nsesample.bat pour configurer les index de texte :
 1. Appelez db2cmd pour ouvrir une fenêtre de commande DB2.
 2. Passez au répertoire <sqllib>\samples\db2ext.
 3. A partir de la fenêtre de commande DB2, appelez nsesample.bat <votrebd> où <votrebd> est le nom d'une base de données. Notez que cette commande crée la base de données si elle n'existe pas encore.
 4. Vérifiez le fichier de sortie généré nsesample.log dans le répertoire en cours.
- Ensuite, appelez les exemples de requêtes suivants devant s'exécuter dans la fenêtre de commande DB2 :
 1. Connectez-vous à votre base de données à l'aide de la commande db2 connect to <votrebd>.
 2. Exécutez les exemples de requêtes à l'aide de la commande db2 -tvf search.
 3. Vérifiez les résultats des requêtes contenues dans le script. Notez que chaque requête doit renvoyer une ou plusieurs occurrences.

Si le fichier nsesample.log ne contient pas d'erreur et que toutes les requêtes aboutissent, l'Extension Net Search a été correctement installée.

Vérification de l'installation sous UNIX

Effectuez la procédure ci-dessous pour vérifier que l'Extension Net Search est correctement installée.

- Suivez ces étapes pour appeler le script d'administration nsesample pour configurer les index de texte :
 1. Accédez à <instance_owner_home>/sqllib/samples/db2ext.
 2. Appelez ./nsesample <yourdb>. Notez que cette commande crée la base de données si elle n'existe pas encore.
 3. Vérifiez le fichier de sortie généré nsesample.log dans votre répertoire principal.
- Ensuite, appelez des exemples de requêtes à exécuter dans la même fenêtre de commande DB2 :

1. Connectez-vous à votre base de données à l'aide de la commande `db2 connect to <votrebd>`.
2. Exécutez les exemples de requêtes à l'aide de la commande `db2 -tvf search`.
3. Vérifiez les résultats des requêtes contenues dans le script. Notez que chaque requête doit renvoyer une ou plusieurs occurrences.

Si le fichier `nsesample.log` ne contient pas d'erreur et que toutes les requêtes aboutissent, l'Extension Net Search a été correctement installée.

Remarque

Pour plus d'informations sur la migration, reportez-vous à l'Annexe A, «Migration», à la page 183.

Chapitre 3. Scénarios utilisateur

Ce chapitre permet d'acquérir des connaissances sur l'Extension Net Search en suivant les exemples pas à pas :

Exemple de recherche scalaire SQL

Cet exemple de ligne de commande explique les fonctions d'indexation et de recherche disponibles.

Exemple de procédure mémorisée

Cet exemple de ligne de commande utilise la commande d'index de l'exemple ci-dessus. Cependant avec l'ajout d'une mémoire cache, cet exemple explique les différentes fonctions d'indexation et de recherche disponibles.

Exemple de fonction de valeur table SQL

Pour plus d'informations sur l'utilisation de l'Extension Net Search, reportez-vous aux chapitres suivants :

- Chapitre 4, «Planification», à la page 23
- Chapitre 6, «Création et gestion d'un index de texte», à la page 31
- Chapitre 8, «Recherche», à la page 71

Remarque

Avant d'utiliser les exemples, assurez-vous que l'Extension Net Search est installée correctement à l'aide de la procédure de vérification de l'installation.

Exemple simple avec la fonction de recherche scalaire SQL

Effectuez les étapes suivantes dans l'exemple DB2 Extension Net Search :

1. Création d'une base de données
2. Activation d'une base de données pour la recherche de texte
3. Création d'une table
4. Création d'un index en texte intégral
5. Chargement des exemples de données
6. Synchronisation de l'index de texte
7. Recherche avec l'index de texte

Vous pouvez émettre des modèles de commandes sur la ligne de commande du système d'exploitation en utilisant une base de données existante. Dans les exemples suivants, le nom de la base de données est `sample`.

Création d'une base de données

Pour créer une base de données dans DB2, utilisez la commande suivante :

```
db2 "create database sample"
```

Activation d'une base de données pour la recherche de texte

Vous pouvez émettre des commandes DB2 Extension Net Search, comme vous le faites pour les commandes DB2 sur la ligne de commande du système d'exploitation. Par exemple, la commande suivante permet de démarrer les services de l'instance de l'Extension Net Search :

```
db2text "START"
```

La commande ci-dessous permet de préparer la base de données à utiliser avec DB2 Extension Net Search :

```
db2text "ENABLE DATABASE FOR TEXT CONNECT TO sample"
```

Cette étape ne doit être effectuée qu'une seule fois pour chaque base de données.

Création d'une table

```
db2 "CREATE TABLE books (isbn VARCHAR(18) not null PRIMARY KEY,  
    author VARCHAR(30), story LONG VARCHAR, year INTEGER)"
```

Cette commande DB2 permet de créer une table appelée books. Elle comporte des colonnes pour l'auteur (author), l'histoire (story), le numéro ISBN (isbn number) et l'année (year) de publication du livre.

Création d'un index en texte intégral

```
db2text "CREATE INDEX db2ext.myTextIndex FOR TEXT ON books (story)  
    CONNECT TO sample"
```

Cette commande crée un index en texte intégral pour la colonne story. Le nom de l'index de texte est db2ext.myTextIndex.

Chargement des exemples de données

```
db2 "INSERT INTO books VALUES ('0-13-086755-1','John', 'Un homme  
    descendait la rue en courant.',2001)"  
db2 "INSERT INTO books VALUES ('0-13-086755-2','Mike', 'Le chat chasse  
    les souris.', 2000)"  
db2 "INSERT INTO books VALUES ('0-13-086755-3','Peter', 'Des hommes se  
    tenaient debout près de la table.',1999)"
```

Ces commandes chargent le numéro ISBN (isbn), l'auteur (author), l'histoire (story) et l'année de publication (publishing year) de ces livres dans la table.

Synchronisation de l'index de texte

Pour mettre à jour l'index de texte avec des données de la table sample, utilisez la commande suivante :

```
db2text "UPDATE INDEX db2ext.myTextIndex FOR TEXT CONNECT TO sample"
```

Recherche avec l'index de texte

Pour rechercher dans l'index de texte, utilisez la fonction de recherche scalaire CONTAINS suivante :

```
db2 "SELECT author, story FROM books WHERE CONTAINS  
    (story, '\"chat\"') = 1 AND YEAR >= 2000"
```

Remarque

Selon le type de shell du système d'exploitation que vous utilisez, il se peut que vous deviez avoir recours à un autre caractère d'échappement devant les guillemets encadrant l'expression à rechercher dans le texte. L'exemple ci-dessus utilise "\"" comme caractère d'échappement.

Cette requête recherche tous les livres contenant le terme chat, parus en 2000 ou après 2000. La requête renvoie la table résultat suivante :

```
AUTHOR    STORY  
Mike      Le chat chasse les souris.
```

D'autres fonctions prises en charge incluent SCORE et NUMBEROFMATCHES. SCORE renvoie un indicateur sur la précision avec laquelle l'argument de recherche décrit un document trouvé. NUMBEROFMATCHES renvoie le nombre d'occurrences des termes de la requête trouvés dans un document résultant.

Exemple simple avec utilisation de la mémoire cache et recherche de procédure mémorisée

Effectuez les étapes suivantes dans l'exemple de recherche de procédure mémorisée DB2 Extension Net Search :

1. Création d'un index de texte avec option de mémoire cache.
2. Synchronisation de l'index et activation de la mémoire cache.
3. Recherche avec la procédure mémorisée TEXTSEARCH.

Remarque

L'exemple de procédure mémorisée présume que les étapes de l'exemple précédent sont terminées et que la base de données est toujours activée.

Création d'un index de texte avec option de mémoire cache

Comme la base de données est déjà activée, utilisez la commande suivante pour créer un index en texte intégral :

```
db2text "CREATE INDEX db2ext.mySTPTextIndex FOR TEXT ON books (story)
        CACHE TABLE (author, story) MAXIMUM CACHE SIZE 1
        CONNECT TO sample"
```

Dans cet exemple, l'index en texte intégral est destiné à la colonne story et il spécifie la table d'une mémoire cache sur les colonnes author et story. Le nom de l'index de texte est mySTPTextIndex.

Synchronisation de l'index et activation de la mémoire cache

Pour mettre à jour l'index en fonction des données insérées dans la table, utilisez la commande suivante :

```
db2text "UPDATE INDEX db2ext.mySTPTextIndex FOR TEXT CONNECT TO sample"
```

Pour activer la mémoire cache, utilisez la commande suivante :

```
db2text "ACTIVATE CACHE FOR INDEX db2ext.mySTPTextIndex FOR TEXT
        CONNECT TO sample"
```

Recherche avec la procédure mémorisée TEXTSEARCH

Vous pouvez uniquement utiliser la procédure mémorisée de DB2 Extension Net Search dans certains cas. Pour plus de détails, reportez-vous à la section «Utilisation d'une recherche de procédure mémorisée» à la page 6.

```
db2 "call db2ext.textSearch
    ('\'chat\'','DB2EXT','MYSTPTTEXTINDEX',0,2,0,0,?,?)"
```

Cette requête recherche tous les livres relatifs à un chat, mais ne renvoie que les deux premiers résultats. Dans ce cas, la requête renvoie la table résultat suivante :

Value of output parameters

```
-----
Parameter Name : SEARCHTERMCOUNTS
Parameter Value : 1
Parameter Name : TOTALNUMBEROFRESULTS
Parameter Value : 1
```

```
AUTHOR    STORY  
Mike     Le chat chasse les souris.
```

```
Return Status = 0
```

Pour d'autres modèles de syntaxe de recherche, consultez le fichier suivant : `sqllib\sample\db2ext\search`

Pour plus de détails sur les autres paramètres utilisés dans la requête, reportez-vous à la section «Recherche de texte à l'aide d'une recherche de procédure mémorisée» à la page 78.

Exemple simple avec fonction de valeur table SQL

La fonction de valeur table SQL peut être utilisée sur les index de texte créés dans les exemples précédents.

La requête de la fonction de valeur table SQL correspond à la requête CONTAINS précédemment utilisée. Pour plus d'informations, reportez-vous à la section "Synchronisation de l'index de texte" à la page 20.

```
db2 "SELECT author, story FROM books b, table (db2ext.textsearch  
      ('\"chat\"','DB2EXT','MYTEXTINDEX', 0, 2, CAST  
      (NULL AS VARCHAR(18)))) T where T.primKey = b.isbn
```

Pour plus de détails, reportez-vous à la section «Utilisation d'une fonction de valeur table SQL pour la recherche» à la page 7.

Remarque

CAST (NULL AS VARCHAR(18)) appelle la fonction de valeur table correspondant à la clé primaire pour la table books.

Pour chaque type de clé primaire, il existe une fonction de valeur table. Ceci identifie la fonction de valeur table correcte pour DB2.

Chapitre 4. Planification

Afin d'utiliser DB2 Extension Net Search de la manière la plus efficace possible, il est essentiel qu'une planification soit établie avant le développement. Celle-ci doit impliquer plusieurs groupes, dont les administrateurs de base de données, les concepteurs système et d'interface, les architectes système et les développeurs.

Les sections suivantes vous guident vers les points devant être pris en compte.

Pour plus d'informations sur le développement d'applications basées sur DB2 Extension Net Search, reportez-vous aux chapitres suivants :

- Chapitre 5, «Services de l'instance de l'Extension Net Search», à la page 27
- Chapitre 6, «Création et gestion d'un index de texte», à la page 31
- Chapitre 8, «Recherche», à la page 71

Emplacements des répertoires et stockage des index

L'espace disque requis pour un index dépend de la taille et du type de données que vous voulez indexer. Comme référence, pour l'indexation de documents mono-octets, vous devez réserver un espace disque d'environ 0,7 fois la taille des documents à indexer. Pour les documents à deux octets, réservez un espace disque de taille égale à celle des documents à indexer.

La quantité d'espace requis pour les fichiers temporaires dans le répertoire de travail est de 1 à 4 fois la quantité d'espace requis pour le fichier d'index final dans le répertoire d'index.

Si vous avez plusieurs index de grande taille, vous devez les stocker sur des périphériques de disque distincts, en particulier si vous disposez d'un accès concurrent aux index pendant leur mise à jour ou pendant la recherche.

Vous devez également spécifier le répertoire dans lequel l'index de texte sera stocké. Assurez-vous que l'espace disque est suffisant et que le propriétaire d'instance DB2 dispose des droits d'accès en écriture sur le répertoire.

Formats de document et pages de codes prises en charge

DB2 Extension Net Search doit connaître le format (ou type) des documents texte à rechercher. Ces informations sont nécessaires pour l'indexation des documents texte.

DB2 Extension Net Search prend en charge les formats de documents suivants :

TEXT	Texte en clair (par exemple, ASCII à plat)
HTML	Hypertext Markup Language
XML	Extended Markup Language
GPP	General Purpose Format (texte à plat comportant des balises définies par l'utilisateur)

Outside-In (INSO)

Logiciel de filtrage permettant d'extraire le contenu texte des PDF

et d'autres outils de formatage de texte courants, par exemple Microsoft Word. Pour plus d'informations, reportez-vous au «Logiciel de filtrage Outside-In».

Pour les formats de document HTML, XML, GPP et les formats de filtre Outside-In, la recherche peut être limitée à des parties spécifiques d'un document. Le Chapitre 9, «Utilisation de documents structurés», à la page 83, explique comment définir et utiliser des modèles de document.

Aux endroits où les filtres Outside-In ne peuvent pas être utilisés pour les formats de documents non pris en charge, vous pouvez écrire une fonction définie par l'utilisateur (UDF). Cette fonction UDF doit être spécifiée au moment de la création de l'index et elle convertit les données du format non accepté en un format pris en charge.

Pour plus de détails, reportez-vous à la section «CREATE INDEX» à la page 114.

Vous pouvez indexer des documents s'ils sont dans l'un des ID de jeu de caractères codés (CCSID) pris en charge. Ils sont également connus sous le nom de pages de codes. Pour en obtenir la liste, reportez-vous à l'Annexe D, «CCSID pris en charge», à la page 201.

Pour vérifier la page de codes de la base de données, utilisez la commande DB2 suivante :

```
db2 GET DB CFG for <nombd>
```

Dans un but de cohérence, DB2 convertit normalement la page de codes d'un document dans celle de la base de données. Cependant, lorsque vous stockez les données dans une base de données DB2, dans une colonne dont le type de données est binaire, BLOB, FOR BIT DATA ou une valeur datalink par exemple, DB2 ne convertit pas les données et les documents gardent leur CCSID d'origine.

Notez que le fait d'avoir deux pages de codes différentes peut engendrer des incidents lors de la création d'un index de texte ou lors de la recherche. Pour plus d'informations, reportez-vous à la section «Création d'un index de texte sur des types de données binaires» à la page 36.

Logiciel de filtrage Outside-In

DB2 Extension Net Search prend en charge le logiciel de filtrage de documents tiers. Vous pouvez utiliser le logiciel appelé Outside-In par Stellent pour extraire le contenu texte des fichiers PDF ou des documents écrits dans le format propriétaire des outils de formatage de texte courant sans avoir recours à des applications natives. Ces formats sont, par exemple, Microsoft Word et Lotus Word Pro.

Téléchargez des bibliothèques Outside-In en tant que modules d'extension pendant la mise à jour de l'index (UPDATE INDEX). Les bibliothèques ne font pas partie de l'Extension Net Search et doivent être installées séparément. Vous devez vous assurer que l'Extension Net Search peut trouver les bibliothèques Outside-In. Pour plus d'informations sur la configuration et l'utilisation des bibliothèques, reportez-vous à la section «Installation des bibliothèques Outside-In» à la page 15.

Le logiciel Outside-In peut non seulement générer du contenu texte mais aussi des informations structurelles, des zones, par exemple. L'Extension Net Search peut également personnaliser la partie des informations du document généré par

Outside-In qui doit être stockée dans l'index. Pour ce faire, vous devez appliquer un type spécifique de modèle de document, le modèle de document Outside-In.

Pour plus d'informations, reportez-vous à la section «Définition d'un modèle de document pour les documents filtrés Outside-In» à la page 177.

Le site Web de Stellent se trouve à l'adresse suivante : <http://www.stellent.com>.

Pour visualiser la liste des formats de filtrage, utilisez l'url suivant :

http://www.stellent.com/intradoc-cgi/nph-idc.cgi.exe/p31019225.pdf?IdcService=GET_FILE&noSaveAs=1&Rendition=Web&RevisionSelectionMethod=LatestReleased&allowInterrupt=1&dDocName=p31019225

Remarque

Le logiciel de filtrage Outside-In est disponible uniquement sur les instances 32 bits. Le support 64 bits n'est pas disponible.

Rôles utilisateur

Les différents rôles et autorisations qui suivent sont ceux des utilisateurs de l'Extension Net Search :

Propriétaire d'instance DB2

L'utilisateur propriétaire d'instance DB2 peut démarrer et arrêter les services de l'instance pour DB2 Extension Net Search et contrôler les services de verrouillage. En outre, l'utilisateur de l'instance DB2 devient DBADM (administrateur de base de données) pour chaque base de données activée. Ceci active un point central de contrôle pour toutes les modifications de la base de données conduites par DB2 Extension Net Search.

Autorisations DB2 requises

DBADM est accordé sur ENABLE DATABASE.

Autorisations du système de fichiers requises

Droits d'accès en lecture et en écriture pour tous les répertoires d'index de texte et accès en lecture aux fichiers modèles.

Commandes pour le propriétaire d'instance

DB2TEXT START, DB2TEXT STOP et DB2TEXT CONTROL.

Les commandes sont uniquement autorisées sur le serveur. Dans un environnement DB2 réparti, vous pouvez utiliser n'importe quel serveur. Chaque commande vérifie si l'utilisateur exécutant la commande est le propriétaire d'instance DB2. Notez que l'utilisation d'un ID utilisateur isolé et distinct sur les systèmes UNIX n'a pas d'effet sur le traitement de l'Extension Net Search en termes d'autorisations ou de performances.

Administrateurs de base de données

Les administrateurs de base de données peuvent activer et désactiver des bases de données à utiliser avec DB2 Extension Net Search.

Autorisations DB2 requises

DBADM (SYSADM pour ENABLE DATABASE).

Commandes pour l'administrateur de base de données

DB2TEXT ENABLE DATABASE et DB2TEXT DISABLE DATABASE.

Propriétaires de table de texte

Les propriétaires de table de texte peuvent créer, supprimer et modifier des index. Notez qu'ils doivent être en mesure de contrôler l'emplacement des index et des mises à jour des index en texte intégral.

Privilèges et autorisations DB2 requises

Propriétaire de table de texte.

Commandes pour le propriétaire de table de texte

DB2TEXT CREATE INDEX, DB2TEXT DROP INDEX, DB2TEXT ALTER INDEX, DB2TEXT ACTIVATE CACHE, DB2TEXT DEACTIVATE CACHE, DB2TEXT UPDATE INDEX, DB2TEXT CLEAR EVENTS et DB2EXTTH.

Notez que l'implémentation des commandes est en partie réalisée sous l'ID utilisateur du propriétaire d'instance DB2. De ce fait, octroyez au propriétaire d'instance les droits d'accès nécessaires au système de fichiers avant de créer ou de modifier les index de texte.

Utilisation de la ligne de commande ou de l'interface du Centre de contrôle DB2 pour l'indexation

Pour l'indexation, vous pouvez utiliser l'option de ligne de commande ou l'interface du Centre de contrôle DB2.

Chapitre 5. Services de l'instance de l'Extension Net Search

Les services de l'instance DB2 Extension Net Search sont constitués des :

- Services de verrouillage
- Services de mise à jour

Le présent chapitre explique comment démarrer et arrêter les services de l'instance DB2 Extension Net Search. Il présente également en détails les services de verrouillage et les services de mise à jour.

Démarrage et arrêt de DB2 Extension Net Search

Pour pouvoir créer un index de texte et rechercher dans les documents, vous devez préalablement démarrer les services de l'instance DB2 Extension Net Search.

Pour démarrer les services de l'instance, connectez-vous avec l'ID utilisateur du propriétaire d'instance DB2 (systèmes UNIX uniquement) et entrez la commande suivante :

```
db2text start
```

Pour arrêter les services de l'instance, entrez la commande suivante :

```
db2text stop
```

Notez qu'il doit y avoir un service d'instance de l'Extension Net Search par instance DB2. Le service de verrouillage gère les verrous pour plusieurs bases de données.

Services de verrouillage

Lorsque vous démarrez DB2 Extension Net Search, les services de verrouillage démarrent automatiquement. Les services de verrouillage sont requis pour synchroniser les accès concurrents aux index de texte dans l'Extension Net Search.

Les services de verrouillage veillent à ce que deux processus ne tentent pas de modifier un index de texte simultanément ou qu'aucun processus ne lit les données d'un index de texte alors qu'un autre processus est en train de les modifier. De ce fait, la plupart des processus demandent un verrouillage de l'index de texte avant de démarrer et le libèrent à nouveau une fois le traitement terminé.

Notez que les services de verrouillage pour les index de texte de l'Extension Net Search ne doivent pas être confondus avec les verrous DB2 qui contrôlent l'accès aux tables DB2.

Utilisation des services de verrouillage

Dans l'Extension Net Search, il existe différents types de verrous qui contrôlent l'accès concurrent à un index. Les verrous sont différents selon que l'index de texte est seulement lu, comme dans le cas d'une demande de recherche ou si les modifications doivent être calculées puis écrites dans des fichiers.

Pendant la commande `db2text start`, les services de verrouillage démarrent automatiquement. Les types de verrous suivants existent sur un index de texte :

Utilisation des services de verrouillage

Verrou S

Pour un accès partagé en lecture uniquement. Par exemple, des demandes de recherche.

Verrou U

Pour un accès en lecture et en écriture pendant le calcul des modifications sur un index (mise à jour) avec un accès concurrent en lecture. Par exemple, des demandes de recherche.

Verrou X

Pour un accès exclusif en lecture/écriture d'une courte durée pendant laquelle les modifications sont réellement écrites dans l'index.

Verrou IX

Pour un accès en lecture/écriture devant être exclusif, empêchant tout nouveau verrou de type S alors que le processus de mise à jour attend un verrou de type X.

Il existe un service de verrouillage de l'Extension Net Search par instance DB2. Le service de verrouillage gère les verrous pour plusieurs bases de données.

Le fichier de configuration des services de verrouillage est db2ext1m.cfg. Il est stocké dans <instance_owner_home>/sql1ib/db2ext pour des systèmes UNIX et dans <sql1ib>\<DB2INSTANCE>\db2ext pour Windows.

Les modifications apportées au fichier de configuration ne prennent effet que lorsque les services de l'instance de l'Extension Net Search sont démarrées lors de l'exécution de la commande db2text start. Pour plus d'informations, reportez-vous à la section «CONTROL» à la page 92. L'utilisateur peut définir les valeurs suivantes :

- le nombre maximal de bases de données,
- le nombre maximal d'index par base de données,
- le nombre maximal de verrous autorisés (utilisateurs concurrents) par index,
- les temps d'attente et le nombre de tentatives pour obtenir un verrou.

Les valeurs par défaut du fichier de configuration sont les suivantes :

```
<default
    maxDbs          = " 8"
    maxIdxPerDb     = " 50"
    maxLocksPerIdx  = "100"

    sWait = " 50"
    uWait = " 500"
    xWait = " 500"

    sAttempt = "50"
    uAttempt = "10"
    xAttempt = "60"

    latchTimeout = "80"

/>
```

La syntaxe est <default attribute=value.../> et la signification des attributs est la suivante :

maxDbs

Nombre de bases de données que les services de verrouillage peuvent gérer (nombre entier supérieur à 1).

maxIdxPerDb

Nombre d'index par base de données pouvant être verrouillés (nombre entier supérieur à 1). Cette valeur est la même pour toutes les bases de données.

maxLocksPerIdx

Nombre de verrous pouvant exister simultanément dans un index (nombre entier supérieur à 1). Cette valeur est la même pour toutes les index.

sWait/sAttempt

Lors de la demande d'un verrou de type S, sAttempt correspond au nombre de tentatives réalisées si le verrou n'est pas accordé immédiatement. sWait correspond au temps d'attente (intervalle) entre chaque tentative (nombre entier supérieur à 1). Ces paramètres s'appliquent également aux verrous de type IX.

uWait/uAttempt

Lors de la demande d'un verrou de type U, uAttempt correspond au nombre de tentatives réalisées si le verrou n'est pas accordé immédiatement. uWait correspond au temps d'attente (intervalle) entre chaque tentative (nombre entier supérieur à 1).

xWait/xAttempt

Lors de la demande d'un verrou de type X, xAttempt correspond au nombre de tentatives réalisées si le verrou n'est pas accordé immédiatement. xWait correspond au temps d'attente (intervalle) entre chaque tentative (nombre entier supérieur à 1).

latchTimeout

Temps d'attente supplémentaire à l'intervalle des services de verrouillage. Pour déterminer le temps d'attente total pour un verrou, utilisez le mode de calcul suivant :

$$\text{waiting time} = \# \text{ attempts} * (\# \text{ waits} + (2 * \# \text{ latchTimeout}))$$

Le temps d'attente est calculé en millisecondes. Notez qu'à chaque tentative, la valeur latchTimeout est doublée lorsqu'elle est ajoutée au temps d'attente global.

Affichage d'une image instantanée d'un verrou

Il est possible d'afficher une image instantanée d'un verrou en ayant recours à l'une des commandes suivantes :

- Pour un seul index de texte :
`db2text CONTROL LIST ALL LOCKS FOR DATABASE mabdd INDEX monindex`
- Pour tous les index de texte verrouillés d'une base de données :
`db2text CONTROL LIST ALL LOCKS FOR DATABASE mabdd`Notez que seuls les index réellement verrouillés figurent sur la liste.

La première fois qu'un index de texte est verrouillé, de la mémoire est réservée pour la base de données et l'index de texte dans les services de verrouillage. Si d'autres index de texte sont verrouillés, de la mémoire est également allouée à ces index dans les services de verrouillage. Cette quantité de mémoire n'est libérée que lorsque l'index de texte est supprimé ou que la base de données est désactivée ou encore chaque fois que les services de l'Extension Net Search sont redémarrés. Ceci signifie qu'un index de texte ou une base de données consomme de la mémoire dans les services de verrouillage, même si aucun verrou n'est actuellement défini.

Affichage d'une image instantanée d'un verrou

La commande "db2text CONTROL CLEAR ALL LOCKS" force la libération de tous les verrous établis sur une base de données ou un index. Pour plus de détails sur l'utilisation de cette commande, reportez-vous à la section «CONTROL» à la page 92. Notez que cette commande ne libère pas la mémoire allouée à la base de données ou aux index. Pour ce faire, vous devez supprimer l'index ou désactiver la base de données ou encore redémarrer les services de l'Extension Net Search.

Services de mise à jour

Les services de mise à jour démarrent lors de la commande db2text start. Ces services mettent automatiquement à jour l'index de texte à des moments spécifiques. Notez que l'index de texte n'est pas immédiatement synchronisé avec la table utilisateur.

Pendant la création de l'index, vous pouvez spécifier la périodicité de la vérification des services de mise à jour si une mise à jour de l'index est requise en faisant usage de la commande suivante :

```
db2text create index DB2EXT.TITLE for text on DB2EXT.TEXTTAB (TITLE)
      UPDATE FREQUENCY D(1,3) H(0,12) M(0) update minimum 5
```

Dans cet exemple, les services de mise à jour sont activés tous les lundi et mercredi à midi (12:00) et à minuit (0:00) et ils vérifient si une tâche doit être accomplie sur l'index db2ext.title. Notez qu'au moins cinq modifications sont nécessaires avant que la mise à jour automatique de l'index ne commence à se synchroniser avec la base de données.

Pour plus de détails sur les paramètres, reportez-vous à la section «CREATE INDEX» à la page 114.

Dans un environnement de base de données partitionné, les services de mise à jour démarrent seulement sur un noeud.

Remarque

L'utilisation de UPDATE FREQUENCY toutes les minutes engendre une grosse charge de travail sur votre machine. Pour éviter cela, utilisez uniquement UPDATE FREQUENCY pour un nombre d'index très limité.

Chapitre 6. Création et gestion d'un index de texte

Le présent chapitre fournit des informations sur la création et la gestion d'un index de texte et aborde les thèmes suivants :

- Introduction aux commandes db2text.
- Activation d'une base de données pour la recherche de texte.
- Création d'un index de texte pour différents types de données.
- Création d'un index de texte sur un alias de type nickname avec une mise à jour incrémentielle de l'index avec DB2 Replication
- Création d'un index de texte pouvant être utilisé par une procédure mémorisée.
- Index de texte sur des vues
- Gestion d'un index

Des informations permettent également d'éviter les problèmes de pages de codes qui peuvent se présenter et des remarques sur les performances que vous devrez peut être prendre en considération.

Avant de créer un index de texte, assurez-vous d'avoir pris en considération les conditions requises énoncées au Chapitre 4, «Planification», à la page 23. D'autres conditions préalables à l'indexation incluent le démarrage des services de l'instance DB2 Extension Net Search à l'aide de la commande db2text start.

Pour obtenir des exemples de création d'un index de texte et rendre la recherche possible sur un texte, reportez-vous au Chapitre 3, «Scénarios utilisateur», à la page 19.

Remarque

Vous pouvez également créer et gérer un index de texte à l'aide du Centre de contrôle DB2. Reportez-vous au Chapitre 7, «Utilisation du Centre de contrôle DB2», à la page 51.

Introduction aux commandes db2text

L'exemple suivant montre une commande DB2 Extension Net Search :

```
db2text ENABLE DATABASE FOR TEXT
```

Remarque

Les commandes db2text, telles que db2text ENABLE DATABASE FOR TEXT et db2text CREATE INDEX, sont également appelées commandes.

Démarrage du processeur de ligne de commande DB2 Extension Net Search

Conseil

Pour chaque commande de création et de gestion d'index, vous pouvez spécifier la base de données, l'utilisateur et le mot de passe.

```
db2text ... connect TO <base_de_données> USER <ID_utilisateur> USING  
<mot_de_passe>
```

Notez que si vous ignorez les options de connexion dans la commande db2text, la variable d'environnement DB2DBDFT spécifie la base de données.

Pour afficher une liste de commandes, entrez la commande suivante :

```
db2text ?
```

Pour afficher la syntaxe d'une commande individuelle, entrez la commande suivante :

```
db2text ? commande
```

Par exemple, pour afficher la syntaxe de la commande CREATE INDEX, utilisez la commande suivante :

```
db2text ? CREATE INDEX
```

db2text renvoie 0 si le traitement de la commande a abouti et 1 si la commande n'a pas été traitée. Notez qu'en cas d'erreur de document et que l'index poursuit ses mises à jour, la commande db2text renvoie 0.

Remarque

Le shell système interprète les caractères spéciaux, tels que ?, (,), *, ! et ". De ce fait, si la commande contient ces caractères, utilisez des guillemets ou un caractère d'échappement.

L'exemple suivant illustre une commande UNIX utilisant des caractères spéciaux :

```
db2 "SELECT * FROM sample WHERE CONTAINS (DESCRIPTION, '\"enable\"') = 1"
```

Activation d'une base de données

Résumé

Quand	Une fois pour chaque base de données contenant des colonnes de texte dans lesquelles la recherche doit s'effectuer.
Commande	ENABLE DATABASE FOR TEXT
Autorisation	SYSADM

Cette commande prépare la base de données connectée pour une utilisation par DB2 Extension Net Search.

Cette commande déclare également les fonctions et les procédures de recherche DB2 Extension Net Search décrites dans le Chapitre 15, «Fonction de recherche scalaire SQL et fonction de valeur table SQL», à la page 149.

Lorsque vous activez une base de données, vous créez les tables et les vues suivantes :

db2ext.dbdefaults

Affiche les valeurs par défaut de la base de données pour les caractéristiques de traitement, de texte et d'index.

db2ext.textindexformats

Affiche la liste des formats pris en charge et des fichiers modèles utilisés.

db2ext.indexconfiguration

Affiche les paramètres de configuration de l'index.

db2ext.textindexes

Vue de catalogue qui effectue le suivi de tous les index de texte.

db2ext.proxyinformation

Affiche les informations sur le proxy pour accéder aux fichiers à l'aide d'un serveur proxy.

Pour plus d'informations sur toutes les vues, reportez-vous à l'Annexe C, «Catalogues d'informations DB2 Extension Net Search», à la page 191.

Lorsqu'une base de données est activée, elle le reste jusqu'à ce que vous la désactiviez.

Remarque

Un ID utilisateur isolé, différent de l'ID propriétaire d'instance ne fonctionne pas avec des bases de données partitionnées.

Désactivation d'une base de données

Résumé

Quand

Lorsque vous n'avez plus l'intention d'effectuer des recherches de texte sur cette base de données.

Commande

DISABLE DATABASE FOR TEXT

Autorisation

DBADM sur la base de données

Lorsque DB2 Extension Net Search prépare la base de données pour une utilisation, certains changements administratifs sont effectués. La présente section décrit les fonctions qui permettent d'inverser le processus.

Pour désactiver la base de données connectée, utilisez la commande suivante :

db2text DISABLE DATABASE FOR TEXT

Désactivation d'une base de données

Lorsque vous désactivez une base de données, vous supprimez également les objets suivants :

- les vues de catalogue DB2 Extension Net Search créées lors de l'activation du base de données,
- la déclaration des fonctions SQL de DB2 Extension Net Search (UDF).

Pour désactiver la base de données et supprimer tous les index de texte, utilisez la commande suivante :

```
db2text DISABLE DATABASE for text force
```

Remarque

La désactivation d'une base de données échouera si des index de texte sont définis dans la base de données. Il est recommandé de supprimer ces index l'un après l'autre, puis de s'assurer qu'aucun incident n'apparaît. Si vous utilisez la commande `DISABLE DATABASE for text force`, celle-ci garantit uniquement le fait que les tables de catalogue Extension Net Search sont supprimées de la base de données.

Cependant, si certains index n'ont pas pu être complètement supprimés, il restera toujours des ressources qui nécessiteront un nettoyage manuel. Ceci inclut :

- Les fichiers dans l'index, le répertoire de travail et le répertoire cache
- Les entrées de planificateur dans `ctedem.dat`
- Les endroits où un index a été créé à l'aide de l'option de capture de réplication : les entrées `IBMSNAP_SIGNAL`, `IBMSNAP_PRUNE_SET` et `IBMSNAP_PRUNCNTL` dans les tables de la base de données éloignée doivent être supprimées manuellement. Ces entrées peuvent être facilement identifiées avec la commande `APPLY_QUAL="NSE" || <nom instance> and TARGET_SERVER= <nom bdd>`.

Dans l'exemple suivant, l'instance est DB2 et la base de données est SAMPLE.

```
DELETE FROM <ccSchema>.IBMSNAP_SIGNAL
WHERE SIGNAL_INPUT_IN IN
      (SELECT MAP_ID FROM <ccSchema>.IBMSNAP_PRUNCNTL
       WHERE APPLY_QUAL= 'NSEDB2' AND TARGET_SERVER= 'SAMPLE');
```

```
DELETE FROM <ccSchema>.IBMSNAP_PRUNCNTL
WHERE APPLY_QUAL= 'NSEDB2' AND TARGET_SERVER= 'SAMPLE';
```

```
DELETE FROM <ccSchema>.IBMSNAP_PRUNE_SET
WHERE APPLY_QUAL= 'NSEDB2' AND TARGET_SERVER= 'SAMPLE';
```

Création d'un index de texte

Résumé

Quand	Une fois pour chaque colonne contenant du texte dans lequel la recherche doit s'effectuer.
Commande	CREATE INDEX ... FOR TEXT ... (Voir les exemples ci-dessous)
Autorisation	CONTROL sur la table

Il est possible de créer un index de texte sur des types de données pris en charge, bien que les conditions requises soient différentes pour les types de données suivants :

- types de données binaires,
- types de données non pris en charge,
- types de données Datalink.

Différentes conditions préalables doivent être également respectées pour la création d'un index de texte pour une recherche de procédure mémorisée.

Lorsque vous procédez à la création d'un index de texte, les objets suivants sont également créés :

Table de journalisation

Cette table conserve une trace de toutes les lignes modifiées dans la table utilisateur. Notez que si vous sélectionnez l'option **Recreate index on Update** ou utilisez des tables de capture de réplication, la table de journalisation n'est pas créée.

Table d'événements

Cette table recueille les informations relatives aux incidents pendant la mise à jour des index de texte.

Déclencheurs sur la table utilisateur

Ils ajoutent des informations dans la table de journalisation chaque fois qu'un document est ajouté, supprimé ou modifié dans la colonne. Ces informations sont nécessaires pour la synchronisation de l'index lors de la prochaine indexation.

Notez que vous ne pouvez créer des déclencheurs que si vous créez une table de journalisation et que l'index de texte est créé sur une table de base et non sur des vues ou des tables d'alias de type nickname.

Afin d'optimiser les performances et l'espace disque, utilisez la commande CREATE INDEX pour spécifier un autre espace table pour les tables.

Remarque

L'utilisation de la commande DB2 LOAD pour importer les documents peut entraîner des incidents, du fait que les déclencheurs ne se mettent pas en application et que l'indexation incrémentielle des documents chargés est impossible.

De ce fait, il est préférable d'utiliser la commande DB2 IMPORT car elle active les déclencheurs.

L'exemple suivant crée un index de texte sur la colonne de texte HTMLFILE dans la table htmltab.

```
db2text create index DB2EXT.HTMLIDX for text on DB2EXT.HTMLTAB  
(HTMLFILE) format HTML
```

Une clé primaire est nécessaire sur cette table.

Les valeurs par défaut pour la création de l'index émanent de la vue db2ext.dbdefaults.

Création d'un index de texte

Si des erreurs se produisent pendant l'indexation, des **événements de mise à jour d'index** sont ajoutés à la table d'événements. Par exemple, lorsqu'un document en file d'attente pour indexation est introuvable. Pour plus d'informations, reportez-vous à la section «Vue d'événements» à la page 197.

Pour inverser les modifications apportées par la commande CREATE INDEX, utilisez la commande DROP INDEX. Pour plus d'informations, reportez-vous à la section «Suppression d'un index de texte» à la page 48.

Pour synchroniser l'index de texte avec la base de données, utilisez la commande suivante :

```
db2text update index DB2EXT.HTMLIDX for text
```

Notez qu'il n'est possible de trouver des documents qu'après synchronisation.

Résumé de la recherche

En fonction des options sélectionnées au cours de la création de l'index, plusieurs méthodes de recherche sont possibles :

- Les fonctions de recherche scalaires s'appliquent à tous les index de texte, excepté ceux créés sur des vues.
- La fonction de recherche de procédure mémorisée ne s'applique qu'aux index de texte, créés avec une mémoire cache.
- La fonction de valeur table SQL s'applique à tous les index de texte, y compris ceux créés sur des vues.

Création d'un index de texte sur des types de données binaires

Lorsque vous stockez des données dans une colonne ayant un type de données binaire, BLOB, FOR BIT DATA, ou une valeur data link par exemple, DB2 ne convertit pas les données. Ceci signifie que les documents conservent leurs pages de codes d'origine (CCSID), ce qui peut entraîner des problèmes lors de la création d'un index de texte, du fait que vous pouvez avoir deux pages de codes différentes. Aussi, vous devez déterminer si vous utilisez la page de codes de la base de données ou celle spécifiée dans la commande CREATE INDEX.

Pour éviter cet incident, spécifiez la page de codes lors de la création de l'index de texte :

```
db2text CREATE INDEX db2ext.comment FOR TEXT ON db2ext.texttab (comment)
        CCSID 1252
```

Si la page de codes n'est pas indiquée, vérifiez quel CCSID a été utilisé pour créer l'index, en appelant :

```
db2 SELECT ccsid FROM db2ext.textindexes WHERE INDSHEMA = 'COMMENT'
        and INDNAME = 'DB2EXT'
```

Notez qu'il n'y a pas de prise en charge pour des documents avec des pages de code différentes dans un index de texte. Pour plus d'informations sur le mode de conversion des paramètres de page de codes des documents par DB2, reportez-vous au manuel *DB2 Universal Database Administration Guide*. Reportez-vous à la section «Informations connexes» à la page x.

Notez qu'un tel problème ne se pose pas avec la création d'index sur des types de données caractères.

Création d'un index de texte sur un type de données non pris en charge

Pour créer un index, les colonnes de texte doivent être de l'un des types de données suivants :

- CHAR
- VARCHAR
- LONG VARCHAR,
- CLOB
- GRAPHIC
- VARGRAPHIC
- LONG VARGRAPHIC
- DBCLOB
- BLOB
- DATALINK

Si les documents se trouvent dans une colonne d'un autre type, UDT (type défini par l'utilisateur) par exemple, vous devez fournir une fonction qui prend le type utilisateur comme entrée et qui fournit l'un des types mentionnés ci-dessus comme type de sortie.

Spécifiez le nom de cette fonction de transformation. Pour plus d'informations, reportez-vous à la section «CREATE INDEX» à la page 114.

Exemple : vous avez l'intention de stocker un texte compressé dans une table.

1. Créez un type défini par l'utilisateur (UDT) pour le texte dans une session SQL interactive :

```
db2 "CREATE DISTINCT TYPE COMPRESSED_TEXT AS CLOB(1M)"
```

2. Créez une table et insérez le texte dedans :

```
db2 "CREATE TABLE UDTTABLE (author VARCHAR(50) not null,  
                             text COMPRESSED_TEXT, primary key (author))"  
db2 "INSERT ..."
```

3. Créez une fonction UDF, par exemple uncompress. Elle reçoit une valeur de type COMPRESSED_TEXT et renvoie le texte non compressé correspondant sous forme de valeur CLOB(10M) par exemple.
4. Créez l'index de texte de la manière suivante pour spécifier la fonction UDF uncompress :

```
db2text "CREATE INDEX UDTINDEX for text ON UDTTABLE  
        (uncompress(text))  
        ..."
```

Création d'un index de texte pour les types de données DATALINK

DB2 Extension Net Search prend en charge le type de données DATALINK.

1. Si vous utilisez des proxy, ajoutez une ligne à la table db2ext.proxyinformation, pour le nom d'hôte, une valeur de délai d'expiration (en secondes) et un port 'proxy' ou 'socks'.

```
db2 INSERT into db2ext.proxyinformation values  
    ('hostname', '80' 'proxy', 10)
```

Notez qu'une seule ligne est autorisée dans cette table. Un déclencheur s'en assure.

2. Pour plus de détails sur la configuration de l'environnement Java sur différentes plateformes, dans le Centre d'aide et d'information DB2 par exemple, reportez-vous à la documentation de DB2 Universal Database version 8. Par principe, vous devez régler la configuration du gestionnaire de bases de données.

La fonction UDF Data Link renvoie une valeur blob (100 Ko). Pour modifier la taille renvoyée, utilisez DB2EXTDL et mettez à jour le paramètre de configuration du gestionnaire de bases de données, `java_heap_sz`.

Installation du fichier jar Data Links

Avec l'Extension Net Search, vous pouvez indexer des données stockées dans des fichiers référencés à l'aide de la fonction DB2 Data Links. Pour ce faire, vous devez installer le fichier jar Data Links `ctedludf.jar`. Dans le processeur de ligne de commande DB2, exécutez la commande suivante :

- Pour UNIX :

```
call sqlj.install_jar  
('file:/<instance_owner_home>/sqllib/java/ctedludf.jar','ctedludf.jar')
```

- Pour Windows :

```
call sqlj.install_jar  
('file:D:\sqllib\java\ctedludf.jar','ctedludf.jar')
```

Vous devez également mettre à jour la taille de la mémoire dynamique Java à l'aide de la commande suivante :

```
db2 update dbm cfg using JAVA_HEAP_SZ 2048
```

Utilisez la commande suivante pour désenregistrer le fichier jar de toutes les plateformes :

```
call sqlj.remove_jar('ctedludf.jar')
```

Pour obtenir la liste des messages d'erreur, reportez-vous à l'Annexe K, «Messages Data Link», à la page 249.

Création d'un index de texte sur un alias de type nickname à l'aide de la réplication avec mise à jour incrémentielle de l'index à l'aide de DB2 Replication

Avant de procéder à la création d'un index de texte sur un alias de type nickname à l'aide d'une table de capture de réplication, vous devez effectuer les étapes suivantes :

Remarque

Les étapes ci-dessous fournissent uniquement une présentation du processus et ne constituent pas un exemple.

1. Configuration de la base de données fédérée DB2 avec toutes les définitions de serveur et d'encapsuleur.
2. Configuration des tables de contrôle de réplication et des programmes de capture sur le serveur éloigné. C'est à cet endroit que réside la table source pour l'alias de type nickname. Reportez-vous au Chapitre 2, "Setting up for Replication" du manuel *DB2 Replication Guide and Reference, Version 8*. Si DB2 ne

Création d'un index de texte sur un alias de type nickname à l'aide de la réplication

crée pas automatiquement les alias de type nickname, vous devez en créer dans la base de données DB2 fédérée à l'aide d'un nom de schéma pour les tables suivantes :

- IBMSNAP_SIGNAL
- IBMSNAP_PRUNE_SET
- IBMSNAP_PRUNCNTL
- IBMSNAP_REGISTER
- IBMSNAP_REG_SYNC (sources non DB2 éloignées uniquement)

Après cette étape, les alias de type nickname destinés aux tables de contrôle de réplication sont disponibles sous un "schéma de contrôle de capture" dans la base de données fédérée DB2. Le nom du schéma est important pour la commande DB2TEXT CREATE INDEX.

3. Enregistrez la table en tant que source de réplication. Pour plus de détails, reportez-vous au Chapitre 3, "Registering tables and views as replication sources" du manuel *DB2 Replication Guide and Reference, Version 8*. Pour consulter les restrictions concernant l'enregistrement de l'alias de type nickname sur lequel l'index doit être créé, reportez-vous à la page 120 pour de plus amples informations.
4. Si DB2 ne crée pas automatiquement un alias de type nickname lors de l'étape d'enregistrement, créez un alias de type nickname pour la table de capture de réplication dans la base de données fédérée. La table de capture de réplication peut être une table CD (capture des données) ou une table CCD (capture des modifications de données). Cet alias de type nickname est un paramètre pour la commande DB2TEXT CREATE INDEX.

Notez que les noms de colonne IBMSNAP_OPERATION, IBMSNAP_COMMITSEQ, IBMSNAP_INTENTSEQ et ceux des colonnes de la clé primaire ne doivent pas être changés.

5. Si vous utilisez une source de réplication DB2, assurez-vous que le programme de capture est actif. Il est fortement recommandé de ne pas utiliser un démarrage à froid pour le programme de capture. Si tel est le cas, toutes les lignes de la table IBMSNAP_SIGNAL pour APPLY_QUAL LIKE 'NSE%' doivent être réinsérées. Dans l'instruction SQL suivante, vous pouvez voir cette opération est effectuée :

```
INSERT INTO <schéma contrôle capture>.IBMSNAP_SIGNAL
SELECT CURRENT_TIMESTAMP, 'CMD', 'CAPSTART', MAP_ID, 'P'
FROM <schéma contrôle capture>.IBMSNAP_PRUNCNTL
WHERE APPLY_QUAL LIKE 'NSE%';
```

6. Vous pouvez utiliser l'exemple suivant pour créer un index de texte sur un alias de type nickname à l'aide de la réplication :

```
DB2TEXT
CREATE INDEX <nom_index> FOR TEXT ON <alias de type nickname>
(<colonne texte>)
REPLICATION CAPTURE TABLE <alias de type nickname capture>
CONTROL TABLE SCHEMA <schéma contrôle capture>
```

Création d'un index de texte pouvant être utilisé par la recherche de procédure mémorisée

Pour pouvoir utiliser la recherche de procédure mémorisée, vous devez spécifier les options de la mémoire cache pendant la commande CREATE INDEX. Ceci permet d'obtenir de hautes performances, en déplaçant toutes les données spécifiées dans la mémoire principale.

Création d'un index de texte placé en mémoire cache et pouvant être utilisé par la recherche de procédure mémorisée

Cependant, avant la première mise à jour de l'index pour la recherche, vérifiez que votre table contient des documents pour éviter la mise à jour d'un index sur une table vide. Ceci permet d'obtenir de meilleures performances d'indexation et une solide estimation de la configuration requise pour la mémoire cache.

La recherche de procédure mémorisée vous permet de renvoyer rapidement des données prédéfinies, associées à un document. Utilisez l'option 'cache table' pour définir cela dans la commande CREATE INDEX. La commande ACTIVATE CACHE déplace alors les données spécifiées dans une mémoire cache.

Remarque

Les fonctions de recherche scalaires SQL peuvent également utiliser cet index de texte, s'il n'a pas été créé sur une vue.

Dans un environnement DB2 réparti, vous devez spécifier un espace table pour les tables d'administration sur un seul noeud pour la procédure mémorisée et émettre un appel sur ce noeud de manière explicite.

Pour être sûr que vous vous connectez au bon noeud, utilisez la variable d'environnement DB2NODE.

Lors de la création d'un index de texte pour la recherche de procédure mémorisée, vous devez déterminer et calculer les paramètres suivants :

- Le type de mémoire cache.
- Le mode de mise à jour de l'index.
- La quantité maximale de mémoire que l'Extension Net Search peut utiliser, MAXIMUM CACHE SIZE.
- La quantité de mémoire disponible nécessaire pour les mises à jour ultérieures de document, PCTFREE. Notez que ceci ne concerne que les mises à jour incrémentielles.

Les types de mémoire cache suivants sont disponibles :

Mémoire cache temporaire

Elle est reconstituée à chaque commande DB2TEXT ACTIVATE CACHE et requiert le chargement des données d'une table DB2 dans la mémoire. Cette opération est plus longue que l'activation d'une mémoire cache persistante, en particulier pour les index de grande taille. Cependant, elle peut fournir des performances de recherche légèrement meilleures.

Mémoire cache persistante

Elle est gérée sur le disque et peut être rapidement mappée vers la mémoire au moyen du système d'exploitation à chaque commande DB2TEXT ACTIVATE CACHE. Dans des scénarios de mise à jour incrémentielle de l'index, elle doit restée activée pour permettre la synchronisation entre l'index et la mémoire cache. Si ce n'est pas le cas, la commande DB2TEXT ACTIVATE CACHE suivante recrée totalement la mémoire cache.

Les méthodes de mise à jour de l'index de texte suivantes sont disponibles :

Sans l'option Recreate index on update

Evitez de supprimer ou de ré-insérer un document dans la table car l'emplacement d'un document supprimé ne peut pas être réutilisé dans la mémoire cache. En conséquence de quoi, la modification des colonnes de clé doit être évitée sur un index activé.

Création d'un index de texte placé en mémoire cache et pouvant être utilisé par la recherche de procédure mémorisée

Cette opération est également connue sous le nom de mise à jour incrémentielle.

Avec l'option **Recreate index on update**

Cette opération recrée l'index à chaque mise à jour. Utilisez autant que possible des types de données variables dans les expressions de la colonne de la mémoire cache. Ceci économisera de l'espace en mémoire cache. Utilisez les expressions de transtypage correspondantes dans la clause `CACHE TABLE`.

Utilisez cette option si vous pensez insérer plus de 50% de vos documents après l'activation initiale de l'index.

L'Extension Net Search fournit deux fonctions SQL vous permettant de déterminer les paramètres de la mémoire `CREATE INDEX` : `MAXIMUM CACHE SIZE` et `PCTFREE`.

Pour les mises à jour incrémentielles et avec recréation

La fonction UDF suivante renvoie la valeur recommandée de la taille maximale de la mémoire cache `MAXIMUM CACHE SIZE` en Mo :

```
DB2EXT.MAXIMUM_CACHE_SIZE(maximumNumberDocs INTEGER,  
    averageRowLength INTEGER, numberOfCacheColumns INTEGER)
```

La commande suivante renvoie le paramètre de longueur moyenne d'une ligne de votre table :

```
SELECT AVG(LENGTH(cache column_1) + ... + LENGTH(cache column_n))
```

Notez que la moyenne peut changer considérablement si d'autres documents sont insérés dans la table. Le nombre de colonnes en mémoire cache fait référence au nombre d'expressions de colonne utilisées dans la clause `CACHE TABLE` de la commande `DB2TEXT CREATE INDEX`.

Pour plus d'informations, reportez-vous à l'Annexe B, «Utilisation de grandes quantités de mémoire», à la page 187.

Pour les mises à jour incrémentielles uniquement

La fonction UDF suivante renvoie la valeur `PCTFREE` recommandée, en fonction des nombres réel et maximal de documents.

```
DB2EXT.PCTFREE(actualNumberDocs INTEGER, maximumNumberDocs INTEGER)
```

Le nombre réel de documents correspond au nombre de lignes de la table au moment de la première commande `ACTIVATE CACHE` qui crée la mémoire cache.

Le nombre maximal de documents est une estimation du nombre maximal de documents contenus dans la table avant que la commande `DB2TEXT ACTIVATE` suivante (dans le cas d'une mémoire cache temporaire) ou `DB2TEXT ACTIVATE CACHE RECREATE` (dans le cas d'une mémoire cache persistante) ne soit exécutée.

Si l'index est recréé à chaque mise à jour, définissez la valeur de `PCTFREE` à 0.

Exemples

Présumez que votre table comporte 10 000 lignes et que vous ne pensez pas qu'elle en dépassera 20 000. Utilisez l'appel suivant pour calculer la valeur de PCTFREE requise :

```
db2 "values DB2EXT.PCTFREE(10000,20000) "
```

Présumez que la taille maximale de ligne est 20 000 et que vous disposez de deux colonnes dans la mémoire cache avec une taille moyenne de 76. Utilisez l'appel suivant pour renvoyer la taille :

```
db2 " values DB2EXT.MAXIMUM_CACHE_SIZE(20000,76,2) "
```

Une fois les paramètres adéquats déterminés, vous pouvez créer votre index et la table de mémoire cache à l'aide de l'appel suivant :

```
db2text CREATE INDEX db2ext.comment FOR TEXT ON db2ext.texttab (comment)
        CACHE TABLE (docid) PCTFREE 10 MAXIMUM CACHE SIZE 5
```

Dans cet exemple, la colonne docid est constituée en plus de l'index, utilisant la mémoire principale pour un renvoi rapide de la table de résultats. 10% de la mémoire cache sont réservés aux futurs documents et la mémoire cache est limitée à 5 Mo maximum.

Mise à jour de l'index de texte

Pour rechercher dans cet index, vous devez le mettre à jour, puis l'activer. Cette opération copie l'expression 'cache table' spécifiée de la base de données dans la mémoire.

Si pendant les opérations ACTIVATE ou UPDATE, les valeurs de MAXIMUM CACHE SIZE sont dépassées, il est recommandé de procéder comme suit :

- Générez à nouveau la mémoire cache à l'aide de la séquence de commandes DB2TEXT suivante : DEACTIVATE CACHE, ALTER INDEX, MAXIMUM CACHE SIZE et ACTIVATE CACHE RECREATE.

Si vous attendez de fréquentes mises à jour des documents, envisagez l'utilisation de types de données de taille fixe pour les expressions de la colonne de mémoire cache dans la clause CACHE TABLE. L'exemple suivant montre comment utiliser la même mémoire cache pendant les opérations de mise à jour.

```
CACHE TABLE(cast(C1 as char(20)), cast(substr(C2,1,10) as char(10))....
```

Dans ce cas, assurez-vous que seules des types de données invariables sont utilisées.

Activation et désactivation de la mémoire cache pour un index de texte

Avant l'activation de la mémoire cache, effectuez toutes les mises à jour incrémentielles en attente afin d'éviter un mauvais calcul de la valeur de PCTFREE.

Pour activer l'index de texte, utilisez la commande suivante :

```
db2text ACTIVATE CACHE FOR INDEX db2ext.comment FOR TEXT
```

Cette commande extrait les données de la table de la mémoire cache spécifiée hors de la base de données et les stocke en mémoire. La durée requise dépend de la taille de la table.

Remarque

Si vous appelez `update index` lorsqu'un index est activé, les tables de la mémoire cache seront également mises à jour. Les documents supprimés pouvant occuper des emplacements en mémoire cache, veillez à définir une valeur suffisante pour `PCTFREE`.

L'appel `ACTIVATE CACHE` doit être réitéré chaque fois que vous arrêtez votre système. Si vous utilisez la mémoire cache persistante, la nouvelle activation sera plus rapide.

Remarque

Si une mise à jour a lieu lorsque la mémoire cache persistante n'est pas activée, cette dernière est supprimée et recrée lors de l'appel d'activation.

Pour économiser les ressources, vous devez également désactiver les index qui ne sont pas couramment requis.

Pour vérifier la quantité de mémoire restante, utilisez l'appel suivant :

```
db2text control show cache status for database cte index db2ext.comment
```

Ceci indique si l'index a été activé et la quantité d'espace restant dans la mémoire cache spécifiée.

Index de texte sur des vues

Vous pouvez créer des index de texte sur des vues en utilisant des procédures mémorisées ou des fonctions de valeur table, cependant, vous ne pouvez pas inclure de fonction scalaire comme `CONTAINS`, par exemple.

Un autre inconvénient majeur réside dans le fait que vous ne pouvez pas créer de déclencheurs sur des vues, aussi les modifications des tables de base sous-jacentes ne sont pas reconnues.

Ainsi, avec les mises à jour incrémentielles d'index, l'utilisateur doit savoir quel document a été ajouté, mis à jour ou supprimé afin de synchroniser l'index de texte avec la base de données. Pour ce faire, vous devez ajouter toutes les modifications à la table de journalisation. Ce processus est illustré dans le modèle suivant :

1. Pour créer la table de base, utilisez la commande suivante :

```
db2 "create table sample (key INTEGER not null PRIMARY KEY, name  
      VARCHAR(50) not null, comment VARCHAR(90))"
```

2. Pour ajouter des entrées, utilisez les commandes suivantes :

```
db2 "insert into sample values(1,'Claus','travaille dans le bureau 301')"  
db2 "insert into sample values(2,'Manja','se trouve dans le même bureau  
      que Juergen')"  
db2 "insert into sample values(2,'Juergen','est le plus loin de  
      Raiko')"  
db2 "insert into sample values(3,'Raiko','est assis dans le bureau  
      à côté de Claus ')"
```

3. Pour créer la vue, utilisez la commande suivante :

```
db2 "create view sampleview as select key, comment from sample"
```

4. Utilisez les commandes suivantes pour créer, mettre à jour et activer l'index de texte :

```
db2text "create index indexview for text on hde.sampleview(comment)
        cache table (comment) maximum cache size 1 key columns
        for index on view (key)"
db2text "update index indexview for text"
db2text "activate cache for index indexview for text"
```

Remarque

Vous devez spécifier la table de la mémoire cache pour pouvoir créer un index de texte sur une vue. Pour créer la table de journalisation correcte, vous devez spécifier les colonnes clés pour l'index sur une vue. Si vous créez un index de cette façon, vous pouvez également l'utiliser avec la fonction valeur table.

Lorsque vous utilisez la recherche de procédure mémorisée dans un environnement DB2 réparti, vous devez spécifier un espace table pour les tables d'administration sur un seul noeud ou appeler sur ce noeud de manière explicite. Pour être sûr que vous vous connectez au bon noeud, utilisez la variable d'environnement DB2NODE.

5. Pour mettre à jour la table, utilisez les commandes suivantes :

```
db2 "insert into sample values(4,'Bernhard','travaille au même étage
        que Manja, mais pas que Claus')"
db2 "insert into sample values(5,'Guenter','partage le même bureau
        que Raiko')"
```

6. Puis, mettez à jour la table de journalisation. Pour obtenir le nom de la table de journalisation, utilisez la commande suivante :

```
db2 "select INDSHEMA,INDNAME,LOGVIEWSCHEMA,LOGVIEWNAME
        from db2ext.textindexes"
```

La table de journalisation se présente comme suit :

sqltype	sqllen	sqlname.data	sqlname.length
496 INTEGER	4	OPERATION	9
392 TIMESTAMP	26	TIME	4
497 INTEGER	4	PK01	4

Pour ajouter les entrées à la table de journalisation, utilisez les commandes suivantes :

```
db2 "insert into sample values(0,CURRENT TIMESTAMP,4)"
db2 "insert into sample values(0,CURRENT TIMESTAMP,5)"
```

La première valeur décrit l'opération (0 = insertion, 1 = mise à jour, 2 = suppression). La seconde doit toujours être CURRENT TIMESTAMP (horodatage en cours) et la dernière, la clé qui a été insérée.

7. Utilisez la commande suivante pour mettre de nouveau à jour l'index :

```
db2text "update index indexview for text"
```

Vous pouvez désormais utiliser la procédure mémorisée pour rechercher dans les nouvelles valeurs.

Remarques sur les performances

Afin d'améliorer les performances au cours de l'indexation, tenez compte des éléments suivants :

- Préférez l'utilisation du type de données VARCHAR pour stocker les documents texte plutôt que LONG VARCHAR ou CLOB.
- Utilisez des disques durs différents pour stocker l'index de texte et les fichiers de la base de données.
- Utilisez des petites colonnes de clé primaire, par exemple TIMESTAMP et INTEGER au lieu de types VARCHAR.
- Assurez-vous que votre système dispose d'une quantité de mémoire réelle disponible suffisante pour toutes ces données. Si la quantité de mémoire est insuffisante, le système d'exploitation utilise à la place l'espace de pagination. Ceci réduit les performances de la recherche.

Pour plus d'informations sur la configuration de la mémoire requise pour les différentes plateformes, reportez-vous à l'Annexe B, «Utilisation de grandes quantités de mémoire», à la page 187.

- Le paramètre du nombre de validations pour la mise à jour, utilisé lors d'une mise à jour automatique ou manuelle de l'index, ralentit les performances d'indexation pendant une indexation incrémentielle. Notez que ce paramètre n'est pas utilisé pendant le processus de mise à jour initiale.

Si le nombre de validations COMMITCOUNT n'est pas défini, alors le paramètre NUMBER_DOCS de db2ext.textindexes n'est pas mis à jour. Par conséquent, pour visualiser le nombre de documents pendant le processus de mise à jour, utilisez la commande CONTROL LIST. Pour plus d'informations, reportez-vous à la section «CONTROL» à la page 92.

- Les performances peuvent diminuer pendant la mise à jour des index si des messages d'erreur ou d'avertissement s'affichent dans la table du journal des événements.

Remarque

Pour obtenir les conseils les plus récents en matière de performances, consultez le site Web DB2 Extension Net Search à l'adresse :
www.ibm.com/software/data/db2/extendernetsearch/index.html

Gestion des index de texte

Ce chapitre explique comment gérer les index de texte et obtenir des informations utiles les concernant. Les tâches de maintenance sont les suivantes :

1. Mise à jour et réorganisation d'un index de texte
2. Modification d'un index de texte
3. Suppression des événements de mise à jour d'un index
4. Suppression d'un index de texte
5. Affichage de l'état des index

Vous pouvez exécuter ces tâches à tout moment et dans n'importe quel ordre. Le présent chapitre inclut également des informations sur la sauvegarde et la restauration des index et des bases de données activées.

Pour connaître les commandes permettant d'afficher des informations sur les index de texte, les noms de répertoire et les mises à jour, reportez-vous à l'Annexe C, «Catalogues d'informations DB2 Extension Net Search», à la page 191.

Vous pouvez également gérer un index de texte à l'aide du Centre de contrôle DB2. Reportez-vous à la section «Gestion d'un index de texte» à la page 65.

Mise à jour et réorganisation d'un index de texte

Une fois l'index de texte créé et mis à jour pour la première fois, vous devez le maintenir à jour. Par exemple, lorsque vous ajoutez un document texte à une base de données ou que vous modifiez un document existant dans une base de données, vous devez indexer le document pour maintenir le contenu de l'index synchronisé avec celui de la base de données. De même, lorsque vous supprimez un document texte d'une base de données, ses termes doivent être retirés de l'index.

Si l'index de texte a été créé sans l'option `RECREATE INDEX ON UPDATE`, des déclencheurs stockent automatiquement les informations sur les documents nouveaux, modifiés ou supprimés dans une table de journalisation interne. Aussi, la prochaine fois qu'une mise à jour de l'index a lieu, les documents référencés dans la table de journalisation sont indexés. Pour plus d'informations sur un index de texte sur des vues, reportez-vous à la section «Index de texte sur des vues» à la page 43.

Si vous spécifiez l'option `RECREATE` dans la commande `CREATE INDEX`, l'index est totalement reconstitué à chaque mise à jour. Cette option ne crée pas de table de journalisation, ni de déclencheur.

Généralement, la mise à jour d'un index s'effectue selon certains intervalles. Il est possible de modifier la fréquence de mise à jour d'un index existant à l'aide de la commande `ALTER INDEX`.

La fréquence de la mise à jour de l'index est spécifiée en termes de quand la mise à jour doit avoir lieu et du nombre minimal de modifications de texte devant se trouver en file d'attente. Si le nombre de modifications figurant dans la table de journalisation est insuffisant au jour et à l'heure donnés, l'index n'est pas mis à jour.

Vous devez planifier soigneusement une indexation périodique. L'indexation de documents texte est une tâche qui demande du temps et des ressources. La durée dépend de plusieurs facteurs. Ceci inclut la taille des documents, le nombre de documents texte ayant été ajoutés ou modifiés depuis la précédente mise à jour de l'index et la puissance du processeur.

Remarque

Dans une table DB2, des récupérations amont et des interblocages peuvent se présenter dans les cas présents :

- hautes fréquences de mise à jour,
- haute fréquence de modifications,
- longues transactions.

La commande `UPDATE INDEX` vous permet de mettre à jour un index immédiatement sur demande.

Résumé

Quand

Lorsqu'un index doit être immédiatement mis à jour sans attendre qu'une indexation périodique ait lieu.

Commande

UPDATE INDEX

Autorisation

CONTROL sur la table

La commande suivante met à jour l'index :

```
db2text UPDATE INDEX comment FOR TEXT
```

Cette commande est utile lorsque vous avez ajouté plusieurs documents texte à une base de données et souhaitez les rechercher immédiatement.

Pour déterminer si une réorganisation manuelle est nécessaire, interrogez la vue `db2ext.textindexes` à l'aide de la commande suivante :

```
db2 "select reorg_suggested from db2ext.textindexes where INDNAME = 'comment'"
```

Si vous spécifiez `MANUAL REORGANIZATION` et mettez souvent à jour une colonne, le processus de mise à jour ralentit. Pour effectuer une réorganisation manuelle, utilisez la commande suivante :

```
db2text UPDATE INDEX comment FOR TEXT reorganize
```

Cependant, si vous spécifiez `AUTOMATIC REORGANIZE` pendant `CREATE INDEX`, l'index sera automatiquement réorganisé si nécessaire.

Modification d'un index de texte

Résumé

Quand

Lorsque la fréquence de mise à jour ou les répertoires de travail et de l'index doivent être modifiés.

Commande

ALTER INDEX

Autorisation

CONTROL sur la table

Utilisez cette commande pour modifier le répertoire de travail de l'index, la fréquence de mise à jour d'un index ou les caractéristiques de la mémoire cache, principalement `MAXIMUM CACHE SIZE` ou `PCTFREE`. Si vous ne spécifiez pas de fréquence de mise à jour, les paramètres courants restent inchangés. Si une mise à jour de l'index ou une recherche est en cours d'exécution, un message d'erreur s'affiche. Ceci indique que l'index est actuellement verrouillé et qu'aucune modification ne peut être effectuée.

L'exemple suivant modifie la fréquence de mise à jour de l'index.

Modification d'un index de texte

```
db2text ALTER INDEX comment FOR TEXT
        UPDATE FREQUENCY d(1,2,3,4,5) h(12,15) m(00) UPDATE MINIMUM 100
```

Dans cet exemple, l'index doit être mis à jour à 12:00 ou 15:00, du lundi au vendredi, si un minimum de 100 documents texte se trouve dans la file d'attente.

Utilisez la commande suivante pour arrêter la mise à jour périodique d'un index :

```
db2text ALTER INDEX comment FOR TEXT
        UPDATE FREQUENCY NONE
```

Si l'index est copié d'un répertoire à un autre, il est verrouillé pendant ce processus. Cependant, après la copie, l'index est déverrouillé et peut à nouveau être utilisé.

Effacement des événements d'index

Résumé

Quand

Lorsque vous n'avez plus besoin des messages dans la table d'événements d'un index.

Commande

```
CLEAR EVENTS FOR INDEX
```

Autorisation

CONTROL sur la table

Des informations sur les événements d'indexation, telles que les heures de début et de fin de mise à jour, le nombre de documents indexés ou les erreurs de document qui se sont produites pendant la mise à jour sont stockées dans la table des événements de l'index. Ceci peut vous aider à déterminer la cause du problème. Lorsque vous n'avez plus besoin de ces messages, vous pouvez les supprimer.

L'exemple suivant supprime les messages de l'index de texte spécifié :

```
db2text CLEAR EVENTS FOR INDEX comment FOR TEXT
```

Suppression d'un index de texte

Résumé

Quand

Lorsque vous n'avez plus l'intention d'effectuer des recherches de texte dans une colonne de texte.

Commande

```
DROP INDEX FOR TEXT
```

Autorisation

CONTROL sur la table

Exemple :

```
db2text DROP INDEX comment FOR TEXT
```


Lors de la suppression d'un index de texte, vous supprimez également les tables et les vues suivantes :

- vue et table de journalisation,
- vue et table des événements,
- déclencheurs de la table de journalisation (si présents).

Remarque

Supprimez toujours les index de la table avant de la supprimer. Si vous supprimez la table en premier, les index existent toujours.

Affichage de l'état de l'index de texte

Pour obtenir des informations sur les index de texte courants dans la base de données, utilisez les vues. Par exemple, si vous souhaitez connaître les valeurs par défaut de la base de données courante, utilisez la commande suivante :

```
db2 "select * from db2ext.dbdefaults"
```

Pour plus d'informations sur les index couramment disponibles, leurs tables correspondantes et le nombre de documents indexés, utilisez la commande suivante :

```
db2 "select indschema, indname, tabschema, tabname, number_docs  
from db2ext.textindexes"
```

Utilisez la commande suivante pour obtenir des informations sur les formats d'un index spécifique :

```
db2 "select format, modelname from db2ext.textindexformats where  
indschema = 'DB2EXT' and indname = 'TITLE'"
```

Pour plus d'informations, reportez-vous à l'Annexe C, «Catalogues d'informations DB2 Extension Net Search», à la page 191.

Si le nombre de validations COMMITCOUNT n'est pas défini, alors le paramètre NUMBER_DOCS de db2ext.textindexes n'est pas mis à jour. Pour visualiser le nombre de documents mis à jour pendant le processus de mise à jour, utilisez la commande suivante :

```
db2text CONTROL LIST ALL LOCKS FOR DATABASE sample INDEX db2ext.title
```

Pour plus d'informations, reportez-vous à la section «CONTROL» à la page 92.

Sauvegarde et restauration d'index

Suivez la procédure ci-dessous pour sauvegarder les bases de données activées et les index de texte créés par DB2 Extension Net Search :

1. Pour savoir quels index DB2 Extension Net Search a créé et où il sont stockés, appelez une instruction SELECT sur la vue db2ext.textindexes :

```
db2 "select indschema, indname, indexdirectory from db2ext.textindexes"
```
2. Assurez-vous qu'aucune mise à jour de l'index n'est en cours d'exécution, puis arrêtez les services DB2 Extension Net Search à l'aide de la commande suivante :

```
db2text stop
```
3. Une fois la base de données sauvegardée, sauvegardez les répertoires et les sous-répertoires de l'index.

Sauvegarde et restauration d'index

4. Redémarrez les services DB2 Extension Net Search à l'aide de la commande suivante :

```
db2text start
```

Suivez la procédure ci-dessous pour restaurer les bases de données activées et les index de texte créés par DB2 Extension Net Search :

1. Arrêtez DB2 Extension Net Search à l'aide de la commande suivante :

```
db2text stop
```

2. Restaurez les copies de sauvegarde des répertoires de l'index dans le même chemin qu'auparavant.

3. Redémarrez DB2 Extension Net Search à l'aide de la commande :

```
db2text start
```

Suppression des fichiers du répertoire /tmp

Lorsque les services de l'Extension Net Search sont en cours d'exécution, vous ne pouvez supprimer aucun des fichiers suivants dans le répertoire /tmp :

- Fichiers sémaphore et mémoire partagée :

```
fichiers CTE*  
<propriétaire_instance>TEXT.0000.LATCH  
<propriétaire_instance>TEXT.0000  
<propriétaire_instance>CACHE.0000  
<propriétaire_instance>SCHEDULER.LATCH  
<propriétaire_instance>DEMON.SEM  
<propriétaire_instance>DEMON:MEM
```

- Pendant la création d'index, si la mémoire cache est temporaire :

```
<nom_base de données>.IX123456  
<nom_base de données>.IX123456.data0
```

Chapitre 7. Utilisation du Centre de contrôle DB2

Utilisez le Centre de contrôle DB2 pour gérer les fonctions d'administration DB2 Extension Net Search, les instances DB2, les bases de données et les objets de base de données, tels que des tables, des vues et des groupes utilisateur.

Vous pouvez invoquer les commandes sur différents objets du Centre de contrôle DB2, par exemple :

- les objets instance,
- les objets de base de données,
- Les objets index

Les principaux éléments du Centre de contrôle DB2 sont la barre de menus, la barre d'outils, l'arborescence d'objets et le panneau de contenu.

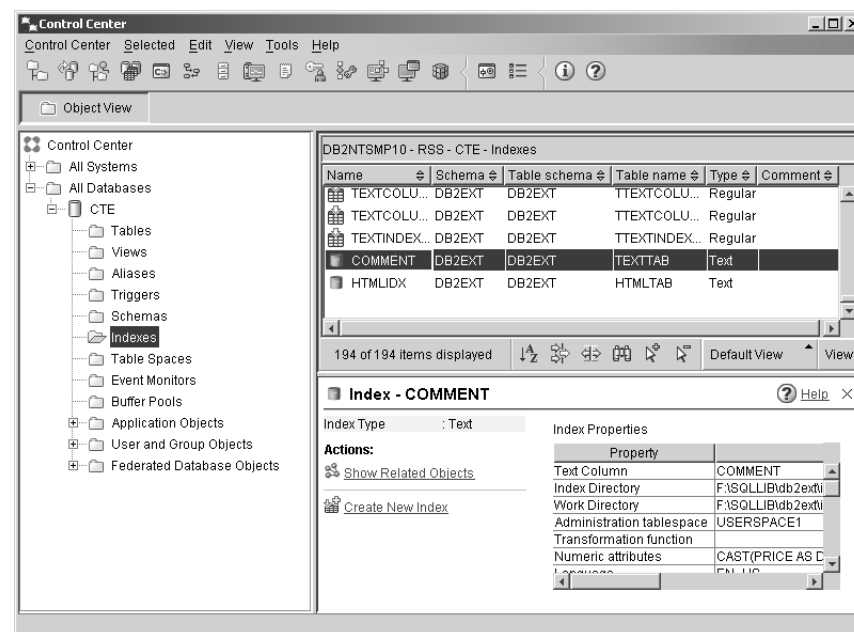


Figure 6. Centre de contrôle DB2

Vous pouvez également utiliser la ligne de commande. Pour plus d'informations, reportez-vous aux chapitres suivants :

- Chapitre 6, «Création et gestion d'un index de texte», à la page 31
- Chapitre 5, «Services de l'instance de l'Extension Net Search», à la page 27

Remarque

Pour pouvoir utiliser les exemples et le Centre de contrôle DB2 pour l'Extension Net Search, une licence valide doit être installée pour l'extension sur le serveur de bases de données.

Seules les fonctions d'indexation et d'administration de DB2 Extension Net Search sont abordées dans ce chapitre. Pour plus d'informations sur l'utilisation du Centre de contrôle DB2, reportez-vous à la section «Informations connexes» à la page x.

Démarrage et arrêt des services de l'instance DB2 Extension Net Search

A partir de l'arborescence d'objets, cliquez sur un système pour afficher les instances disponibles. Mettez en évidence l'instance et à l'aide du bouton droit de la souris, cliquez pour afficher le menu en incrustation d'objets instance. Mettez en évidence l'**Extension Net Search** et sélectionnez l'une des commandes suivantes dans le menu en incrustation :

Démarrage des services de l'instance DB2 Extension Net Search

Cette opération démarre les services de l'instance s'ils ne le sont pas déjà.

Arrêt des services de l'instance DB2 Extension Net Search

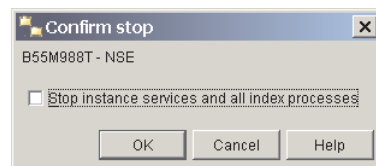


Figure 7. Boîte de dialogue d'arrêt des services de l'Extension Net Search

Une boîte de dialogue s'affiche. Cochez la case pour arrêter les services de l'instance et les processus de l'index. Dans la syntaxe de commande, ceci est connu comme l'option **FORCE**. Cliquez sur le bouton **OK**.

Etat de l'instance

Une boîte de dialogue s'affiche indiquant l'état de l'instance.

Pour plus d'informations, reportez-vous au Chapitre 11, «Commandes d'administration pour le propriétaire d'instance», à la page 91.

Activation d'une base de données

Dans l'arborescence d'objets, cliquez sur l'objet d'instance pour afficher les bases de données disponibles. Mettez en évidence la base de données et à l'aide du bouton droit de la souris, cliquez pour afficher le menu en incrustation. Mettez en évidence l'**Extension Net Search** et sélectionnez l'une des commandes suivantes dans le menu développé :

Activation de la base de données pour le texte

Une boîte de dialogue s'affiche si la base de données n'est pas activée. Cliquez sur le bouton **OK** pour activer la base de données. Si elle est activée, un message apparaît.

Administration de la base de données à partir du Centre de contrôle DB2

Désactivation de la base de données pour le texte

Une boîte de dialogue s'affiche si la base de données n'est pas désactivée. Cliquez sur le bouton **OK** pour désactiver la base de données. Si elle l'est déjà, un message apparaît.

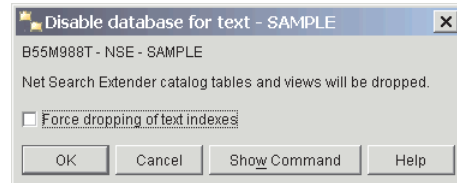


Figure 8. Boîte de dialogue Désactivation de la base de données pour le texte

Cochez la case si vous souhaitez désactiver la base de données et supprimer tous les index de texte.

Modification de la taille de renvoi de la fonction Datalink

Une boîte de dialogue s'affiche indiquant la taille de renvoi courante de la fonction Datalink. Entrez la nouvelle valeur Datalink en kilo-octets (Ko) et cliquez sur le bouton **OK**.

Pour plus d'informations, reportez-vous au Chapitre 12, «Commandes d'administration pour l'administrateur de base de données», à la page 97.

Notez que dans toutes les boîtes de dialogue, le bouton **Afficher la commande** affiche l'alternative de ligne de commande.

Administration de l'index de texte

Dans l'arborescence d'objets, sous l'objet de base de données, vous pouvez voir l'objet index. Cliquez sur l'objet index pour afficher les index dans le panneau de contenu. Vous pouvez identifier les objets index de texte par le type 'texte' dans la colonne Type.

Administration de l'index de texte

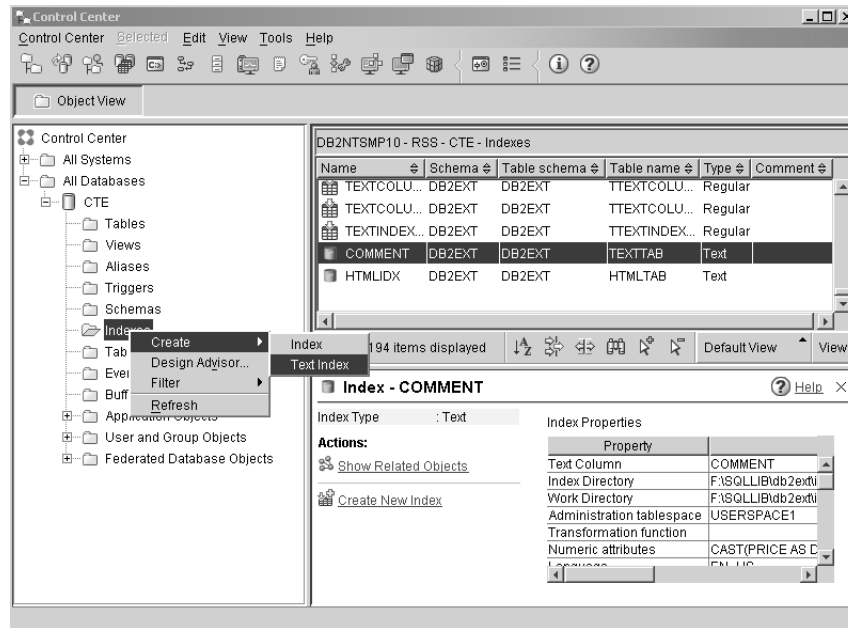


Figure 9. Centre de contrôle DB2

A l'aide du bouton droit de la souris, cliquez sur l'objet index et sélectionnez l'une des commandes suivantes dans le menu en incrustation :

Créer Une boîte de dialogue s'affiche. Sélectionnez **Index de Texte** dans la boîte de dialogue pour lancer un assistant permettant de créer un index de texte. Pour plus de détails, reportez-vous à la section «Création d'un index de texte» à la page 55.

Filtrer Une boîte de dialogue s'affiche dans laquelle vous pouvez sélectionner les objets index à afficher dans la vue du panneau de contrôle.

Régénérer

Les informations de l'arborescence d'objets et du panneau de contrôle sont régénérées.

Pour gérer les index de texte, reportez-vous à la section «Gestion d'un index de texte» à la page 65.

Remarque

Pour accéder aux commandes d'objet d'instance, index de texte et de base de données sans avoir recours au bouton droit de la souris, cliquez sur la commande de menu **Sélection** et mettez en évidence l'**Extension Net Search** pour accéder aux commandes adéquates.

Avant de créer un index de texte, assurez-vous d'avoir pris en considération les conditions requises énoncées au Chapitre 4, «Planification», à la page 23.

D'autres conditions requises pour l'indexation incluent les éléments suivants :

- Démarrage des services de l'instance DB2 Extension Net Search
- Activation de la base de données

Création d'un index de texte

Sélectionnez la commande **Créer**, mettez en évidence **Index de Texte** dans le menu développé, l'assistant Création d'index de texte apparaît. Cet assistant permet de spécifier les options de configuration pour l'index de texte dans un certain nombre de panneaux.

Pour passer d'un panneau à l'autre, entrez toutes les informations obligatoires et cliquez sur le bouton **Suivant** jusqu'à ce que le bouton **Fin** soit activé. Cliquez sur le bouton **Fin** pour créer l'index de texte.

Pour créer un index de texte sur des vues, utilisez la commande CREATE INDEX décrite à la section «CREATE INDEX» à la page 114.

Ecran Nom

Cet écran permet de spécifier le schéma et le nom de l'index de texte. Vous pouvez également spécifier un répertoire de travail et d'index pour les fichiers d'index de texte. Créez les tables administratives pour l'index sur l'espace table d'administration.

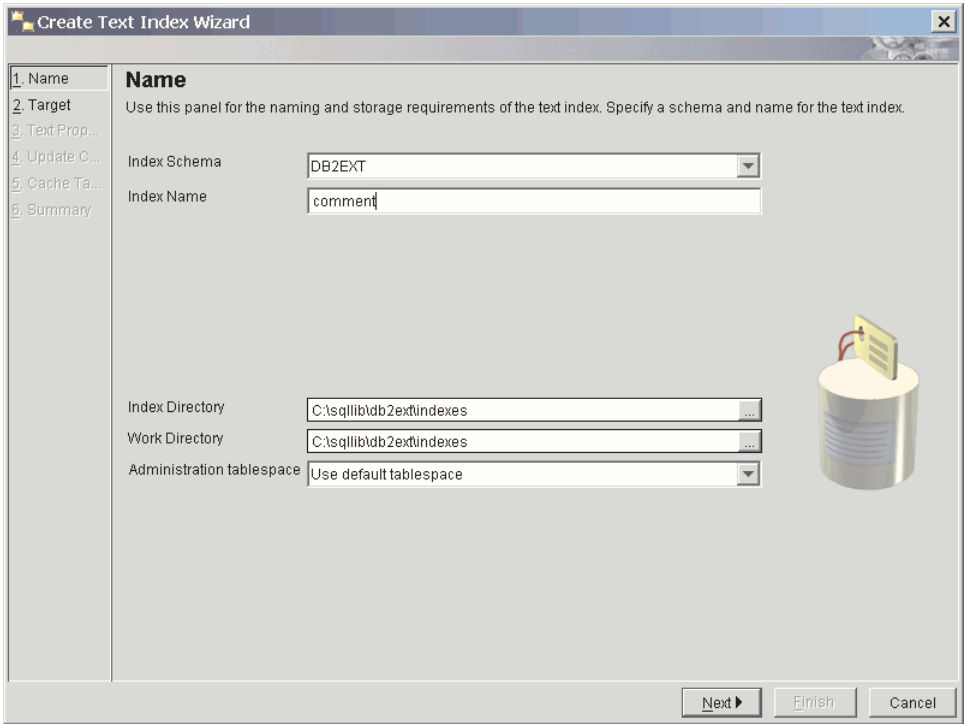


Figure 10. Assistant Création d'index texte : écran Nom

Ci-dessous figure la description des zones de cet écran.

Tableau 1. Zones texte de l'écran Nom

Nom de la zone	Obligatoire/ Facultatif	Valeur par défaut	Description
Schéma d'index	Obligatoire	ID utilisateur	Sélectionnez un nom de schéma de l'index de texte. Il s'agit du nom de schéma DB2 pour les tables d'administration spécifiques à l'index.

Création d'un index de texte

Tableau 1. Zones texte de l'écran Nom (suite)

Nom d'index	Obligatoire	N/A	Entrez un nom d'index DB2 valide pour l'index de texte. Avec le schéma d'index, il identifie de manière unique un index en texte intégral dans la base de données.
Répertoire d'index	Facultatif	Voir le chemin d'accès	Spécifiez le chemin d'accès au répertoire dans lequel l'index de texte sera stocké. Le répertoire doit exister avec des droits d'accès en lecture, écriture et exécution pour l'ID utilisateur du propriétaire d'instance DB2.
Répertoire de travail	Facultatif	Voir le chemin d'accès	Spécifiez le répertoire de travail dans lequel les fichiers temporaires seront stockés pendant les opérations de recherche et d'administration. Le répertoire doit exister avec des droits d'accès en lecture, écriture et exécution pour l'ID utilisateur du propriétaire d'instance DB2.
Espace table d'administration	Facultatif	Utilisez l'espace table par défaut	Sélectionnez un nom d'espace table pour les tables d'administration de l'index de texte. Vous devez définir l'espace de table sur le même groupe de noeuds que l'espace table pour la table utilisateur.

Ecran Cible

Cet écran permet de spécifier le schéma et le nom de la table ou de la table d'alias de type nickname, ainsi que le nom de la colonne de texte contenant les données à indexer. Il est possible d'utiliser une fonction de transformation pour modifier le contenu de la colonne de texte. En plus de la colonne de texte, vous pouvez également spécifier des attributs numériques si vous souhaitez indexer le contenu de l'expression d'une colonne de table.

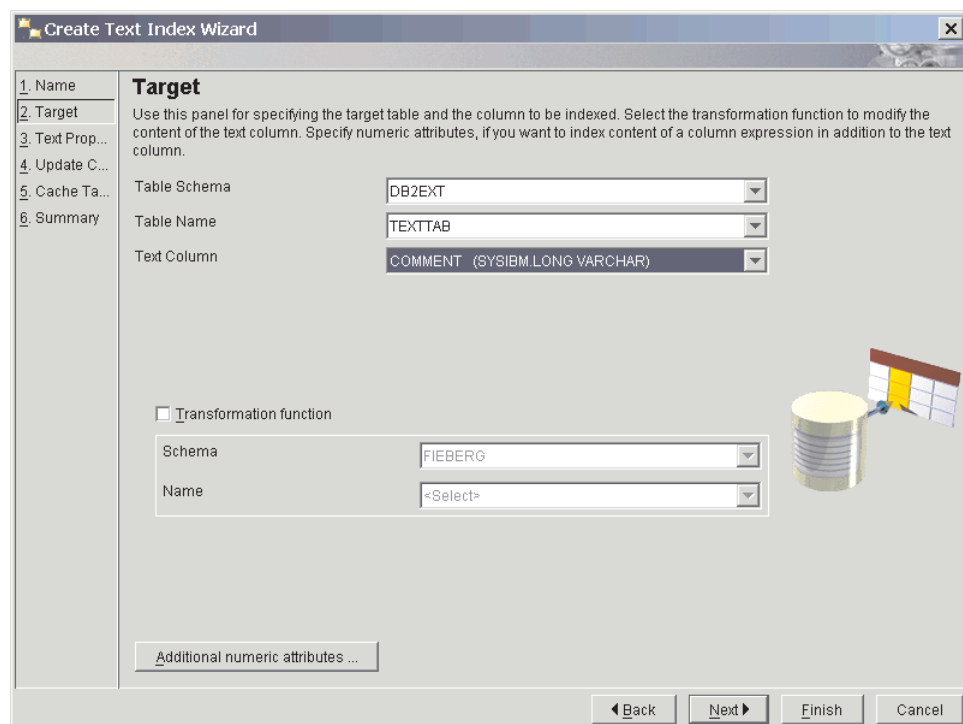


Figure 11. Assistant Création d'index texte : écran Cible

Ci-dessous figure la description des zones de cet écran.

Tableau 2. Zones texte de l'écran Cible

Nom de la zone	Obligatoire/Facultatif	Valeur par défaut	Description
Schéma de la table (1)	Obligatoire	ID utilisateur	Sélectionnez le schéma de la table ou de la table d'alias de type nickname sur laquelle vous créez un index de texte.
Nom de la table (2)	Obligatoire	N/A	Sélectionnez le nom de la table ou de la table d'alias de type nickname sur laquelle vous créez un index. La table doit comporter une clé primaire.
Colonne texte (3)	Obligatoire	N/A	Sélectionnez le nom de la colonne utilisée pour la création de l'index de texte. La colonne doit être ou être transformée dans l'un des types suivants : CHAR (pour données binaires), VARCHAR (pour données binaires), LONG VARCHAR (pour données binaires), CLOB, DBCLOB, BLOB, GRAPHIC, VARGRAPHIC, LONG VARGRAPHIC et DATALINK.
Fonction de transformation	Facultatif	Désactivée	Cochez cette case pour utiliser une fonction de transformation.
Fonction de transformation: Schéma	Obligatoire (si la fonction est sélectionnée)	ID utilisateur	Sélectionnez le schéma de la fonction UDF utilisée pour accéder aux documents texte.
Fonction de transformation: Nom	Comme ci-dessus.	N/A	Sélectionnez le nom de la fonction UDF utilisée pour accéder aux documents texte.

Notez que vous ne pouvez spécifier que le schéma de table (1), le nom de table (2) et la colonne de texte (3), dans cet ordre.

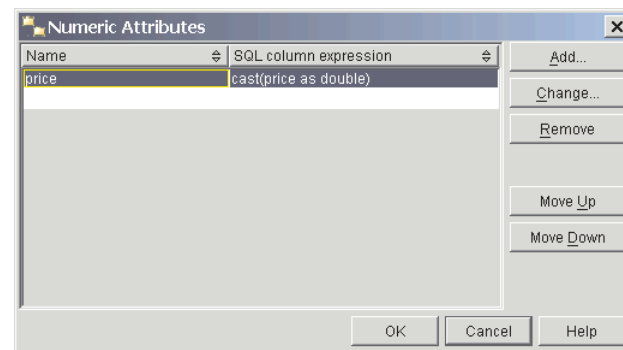


Figure 12. Boîte de dialogue Attributs numériques

Pour afficher ou ajouter des attributs, cliquez sur le bouton **Attributs numériques**. Une fenêtre s'ouvre. Pour ajouter des attributs numériques à l'index, cliquez sur le bouton **Ajout**. Une seconde fenêtre s'ouvre. Spécifiez l'expression et le nom de la colonne SQL pour l'attribut.

Vous pouvez également sélectionner un attribut et appuyer sur les boutons appropriés pour modifier, déplacer ou supprimer une entrée.

Explication des attributs numériques

Les attributs numériques permettent d'indexer des expressions de colonne en plus de la colonne de texte. Par exemple, si vous souhaitez indexer la date de la colonne de type `TIMESTAMP` en plus de la colonne de texte, indiquez un attribut numérique `"cast(julian_day(date) as double)"` et un nom pour l'attribut.

Spécifiez un attribut numérique si vous souhaitez utiliser une expression numérique dans une requête de recherche. Si vous effectuez des recherches à l'aide de requêtes SQL, vous pouvez utiliser une recherche combinée, au lieu d'attributs numériques, par exemple : `WHERE numattrib = 123 AND contains('...')`.

Ecran Propriétés de type texte

Cet écran permet de spécifier la langue et le format des documents texte. Si les documents n'ont pas le même CCSID que la base de données et que la colonne de texte est de type binaire, spécifiez le CCSID. Notez que celui de la base de données est initialement sélectionné. Si vos documents sont de format structuré GPP, HTML, Outside-In ou XML, vous pouvez spécifier un modèle de document.

Remarque

Dans la boîte à liste de format, le format de filtrage Outside-In est aussi connu sous le nom d'INSO.

Create Text Index Wizard

1. Name
2. Target
3. Text Prop...
4. Update C...
5. Cache Ta...
6. Summary

Text Properties
Use this panel for specifying the properties of the document you want to index. If your documents are a certain structured format, you can specify a document model.

Language: English / U.S. (EN_US)
CCSID: 1252
Format: TEXT

☒ Default Document Model
☐ User Document Model

Model name:
Model file:
Model CCSID: 1252

☐ Treat numbers as words
☒ Index stopwords

Back Next Finish Cancel

Figure 13. Assistant Création d'index texte : écran Propriétés de type texte

Ci-dessous figure la description des zones de cet écran.

Tableau 3. Zones texte de l'écran Propriétés de type texte

Nom de la zone	Obligatoire/ Facultatif	Valeur par défaut	Description
Langue	Facultatif	EN_US	Sélectionnez une langue afin de définir les délimiteurs de fin de phrase et de fin de paragraphe lors de l'indexation des documents.
CCSID	Facultatif	CCSID de la base de données	Sélectionnez le CCSID pour l'indexation des documents texte.
Format	Facultatif	TEXT	Sélectionnez le format du document texte : HTML, XML, TEXT, INSO ou GPP.
Modèle de document par défaut	Facultatif	Activé	Utilisez le modèle de document par défaut.
Modèle de document utilisateur	Facultatif	Désactivé	Utilisez le modèle de votre document.
Nom du modèle	Obligatoire (si l'option Modèle de document utilisateur est sélectionnée)	N/A	Entrez le nom du modèle de document. Pour les formats HTML, XML, Outside-In et GPP, vous pouvez spécifier un modèle de document. Notez que le nom ne figure que dans le fichier du modèle.
Fichier du modèle	Comme ci-dessus.	N/A	Spécifiez le fichier du modèle de document. Le fichier doit être lisible par les propriétaires d'instance DB2.
CCSID du modèle	Comme ci-dessus.	CCSID de la base de données	Sélectionnez le CCSID pour interpréter le contenu du fichier du modèle de document.
Considérer les nombres comme des mots	Facultatif	Désactivé	Cochez cette case pour que les séquences de chiffres soient interprétées comme des mots distincts, même si elles sont accolées à des caractères.
Mots vides de l'index	Facultatif	Activé	Cochez cette case pour activer le traitement des mots vides spécifiques à la langue. Le fichier <langue>.tsw dans le répertoire sqllib/db2ext/resources contient la liste des mots vides.

Ecran Mise à jour des caractéristiques

Cet écran permet de spécifier si l'index est mis à jour de manière incrémentielle ou s'il est totalement recréé. Vous pouvez définir les paramètres de mise à jour de telle sorte que l'index se mette à jour automatiquement au moment spécifié.

Création d'un index de texte

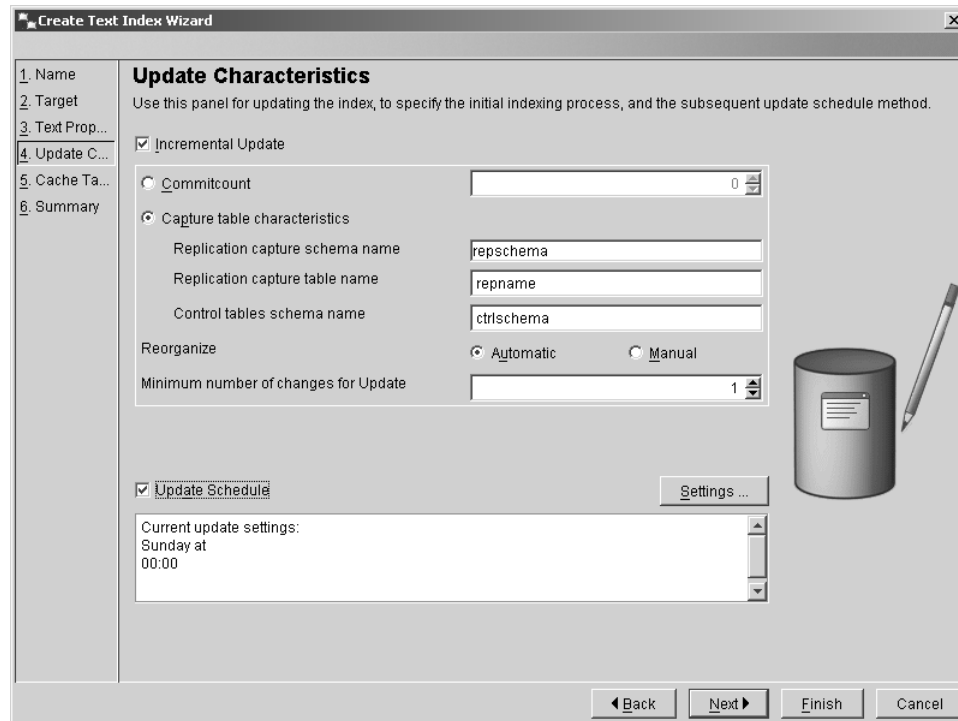


Figure 14. Assistant Création d'index texte : écran Mise à jour des caractéristiques

Ci-dessous figure la description des zones de cet écran.

Tableau 4. Zones texte de l'écran Mise à jour des caractéristiques

Nom de la zone	Obligatoire/Facultatif	Valeur par défaut	Description
Mise à jour incrémentielle	Facultatif	Activé	Cochez cette case pour que l'index soit mis à jour de manière incrémentielle. Si vous ne la cochez pas, l'index sera recréé chaque fois qu'une opération de mise à jour est effectuée.
Nombre de validations	Facultatif	0	Nombre de modifications traitées au cours d'une mise à jour dans une transaction. Le nombre de validations a des effets sur les performances. Pour plus d'informations, reportez-vous à la section «Remarques sur les performances» à la page 45.
Caractéristiques de la table de capture	Facultatif	N/A	Cette option permet d'utiliser une table de capture de réplication pour capturer les modifications sur la table source. La table de capture de réplication doit être une table CD (capture des données) ou une table CCD (capture des modifications de données).
Nom du schéma de capture de réplication	Facultatif	ID utilisateur	Nom du schéma de la table de capture de réplication. Notez que la table doit d'abord avoir été créée à l'aide de la réplication DB2.
Nom de la table de capture de réplication	Obligatoire, si l'option Caractéristiques de la table de capture est activée	N/A	Nom de la table de capture de réplication. Notez que la table doit d'abord avoir été créée à l'aide de la réplication DB2.

Tableau 4. Zones texte de l'écran Mise à jour des caractéristiques (suite)

Nom du schéma de la table de contrôle	Obligatoire, si l'option Caractéristiques de la table de capture est activée	N/A	Nom du schéma de la table de contrôle. Notez que les tables doivent d'abord avoir été créées à l'aide de la réplication DB2.
Réorganisation Automatique ou Manuelle	Facultatif ou Obligatoire	Activé/ Désactivé	Ces boutons d'option permettent d'effectuer automatiquement ou manuellement la réorganisation de l'index.
Nombre minimal de modifications dans la mise à jour	Facultatif	1	Spécifiez le nombre minimal de modifications des documents texte avant que l'index ne soit mis à jour de manière incrémentielle au moment spécifié.
Programmation de la mise à jour	Facultatif	Désactivée	Cochez cette case pour ajouter des paramètres de mise à jour automatique.

Pour ajouter des paramètres de mise à jour de l'index, cliquez sur le bouton **Paramètres**. Notez que ce bouton est uniquement activé si vous sélectionnez **Programmation de la mise à jour**. Dans la boîte de dialogue, sélectionnez les jours, les heures et les minutes pour la mise à jour. Notez que si vous sélectionnez plusieurs jours, la mise à jour a lieu à la même heure, tous les jours sélectionnés.

Ecran Table de mémoire cache

Cet écran permet de spécifier une table mise en cache en plus de l'index. Vous pouvez spécifier la mise en mémoire cache des colonnes de résultats et vous pouvez effectuer la recherche sur la mémoire cache à l'aide d'une procédure mémorisée. D'autres paramètres de mémoire cache peuvent également être spécifiés, par exemple le type, la taille maximale et l'ordre dans lequel le contenu de la table utilisateur est extrait au cours de l'indexation initiale.

Création d'un index de texte

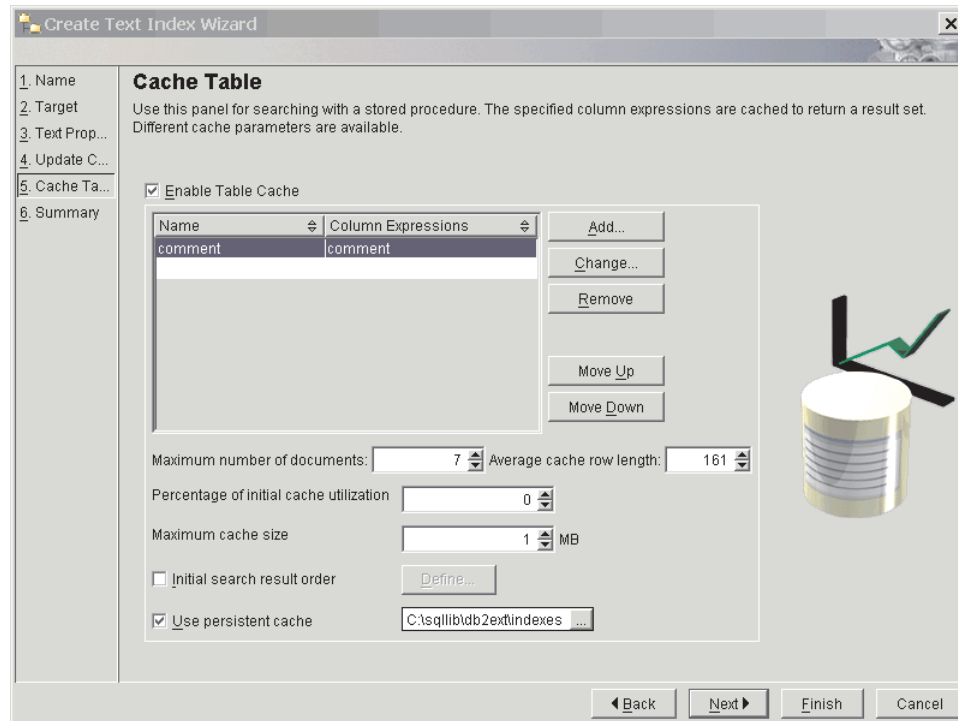


Figure 15. Assistant Création d'index texte : écran Table de mémoire cache

Ci-dessous figure la description des zones de cet écran.

Tableau 5. Zones texte de l'écran Table de mémoire cache

Nom de la zone	Obligatoire/ Facultatif	Valeur par défaut	Description
Activer la mémoire cache de la table	Facultatif	Désactivé	Cochez cette case pour activer la génération d'une table mise en cache.
Table de colonne de résultats	Obligatoire (si la case Activer la mémoire cache de la table est cochée)	N/A	Affiche la liste des expressions de la colonne SQL spécifiant les colonnes de résultats de la recherche.
Nombre maximal de documents	Obligatoire	Nombre de lignes d'une table	Voir la section suivante : Détermination de l'utilisation et de la taille de la mémoire cache.
Longueur moyenne de ligne de cache	Obligatoire	N/A	Voir la section suivante : Détermination de l'utilisation et de la taille de la mémoire cache.
Pourcentage d'utilisation initiale de la mémoire cache	Facultatif	50%	Sélectionnez le pourcentage de mémoire cache, gardé disponible pour les documents supplémentaires.
Taille de la mémoire cache maximale	Facultatif	N/A	Spécifiez la taille maximale de la table mise en cache générée pendant l'activation de l'index. Si la valeur est trop faible, l'activation échouera.
Ordre initial d'affichage des résultats de la recherche	Facultatif	Désactivé	Cochez cette case pour définir l'ordre d'affichage des résultats de la recherche. Les documents sont renvoyés dans le même ordre d'indexation que celui de la table placée en mémoire cache. Cet ordre ne peut pas être garanti après une mise à jour incrémentielle.

Tableau 5. Zones texte de l'écran Table de mémoire cache (suite)

Utiliser la mémoire cache persistante	Facultatif	Activé	Cette option permet une exécution rapide de l'activation, après une désactivation ou un réamorçage du système. Notez que vous devez spécifier le chemin d'accès au répertoire pour la mémoire cache persistante. N'activez pas cette option si la mémoire cache doit être temporaire.
---------------------------------------	------------	--------	---

Détermination de l'utilisation et de la taille de la mémoire cache

La valeur Pourcentage d'utilisation initiale de la mémoire cache indique le pourcentage de mémoire cache libre pour des documents supplémentaires. La valeur Taille de la mémoire cache maximale indique la taille maximale de la table mise en cache devant être générée pendant l'activation du cache. Ces options dépendent des facteurs suivants :

- Le nombre réel de documents présents dans la table.
- Le nombre attendu de mises à jour.
- La taille moyenne des expressions SQL à mettre en cache.

Vous pouvez entrer les valeurs recommandées pour le Pourcentage d'utilisation initiale de la mémoire cache et la Taille de la mémoire cache maximale. Vous pouvez également les laisser être calculées chaque fois que vous entrez des valeurs dans les zones Nombre maximal de documents ou Longueur moyenne de ligne de cache.

La valeur Nombre maximal de documents est initialement définie au nombre de lignes de la table. Modifiez cette valeur en fonction du nombre de documents et du nombre de modifications attendues. Incluez toutes les mises à jour, ajouts et suppressions de documents.

Lorsque vous ajoutez une expression SQL à la liste de table de mémoire cache, la valeur Longueur moyenne de ligne de cache est calculée en fonction de la longueur du résultat. Comme le nombre de lignes de votre table est pris en considération, le calcul peut prendre beaucoup de temps. Si vous savez qu'en moyenne, cette valeur est plus petite, modifiez-la.

Par exemple, si votre table comporte 10 entrées et que la somme des expressions de colonne est 100, cela signifie que ces valeurs sont initialement définies. Si vous vous attendez à ce que le nombre maximal de documents (y compris ceux qui ont été supprimés) soit égal à 10 000, entrez ce nombre. Si vous savez que les expressions de colonne en moyenne sont inférieures à la valeur calculée, par exemple VARCHAR(100) et une taille de texte plein de 10, utilisez ce nombre comme taille de ligne moyenne.

Pour définir l'ordre initial d'affichage des résultats de la recherche, cliquez sur le bouton **Définition**. Notez que ce bouton n'est disponible que si vous avez coché la case Ordre initial d'affichage des résultats de la recherche. Une boîte de dialogue affiche toutes les expressions de la colonne SQL spécifiée. Pour ajouter un ordre d'affichage, cliquez sur le bouton **Ajouter**, puis dans la boîte de dialogue, spécifiez l'ordre d'affichage des résultats SQL.

Pour modifier, déplacer ou supprimer une entrée, sélectionnez l'expression et cliquez sur les boutons appropriés.

Création d'un index de texte

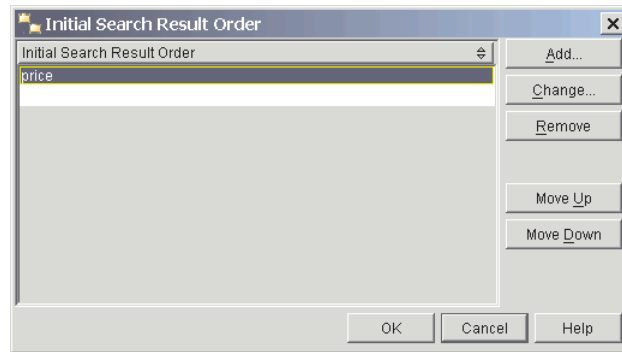


Figure 16. Boîte de dialogue Ordre initial d'affichage des résultats de la recherche

Pour ajouter des expressions de la colonne SQL, cliquez sur le bouton **Ajouter**, en regard de la table Colonne de résultats. Dans la boîte de dialogue, spécifiez l'expression et le nom de la colonne de résultats.

Pour modifier ou déplacer une entrée, cliquez sur l'expression de colonne qui active les boutons appropriés.

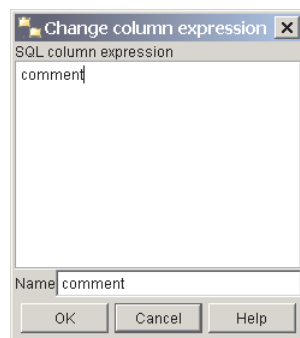


Figure 17. Boîte de dialogue Modification d'une expression de colonne

Ecran Récapitulatif

Cet écran présente un aperçu des paramètres précédemment sélectionnés.

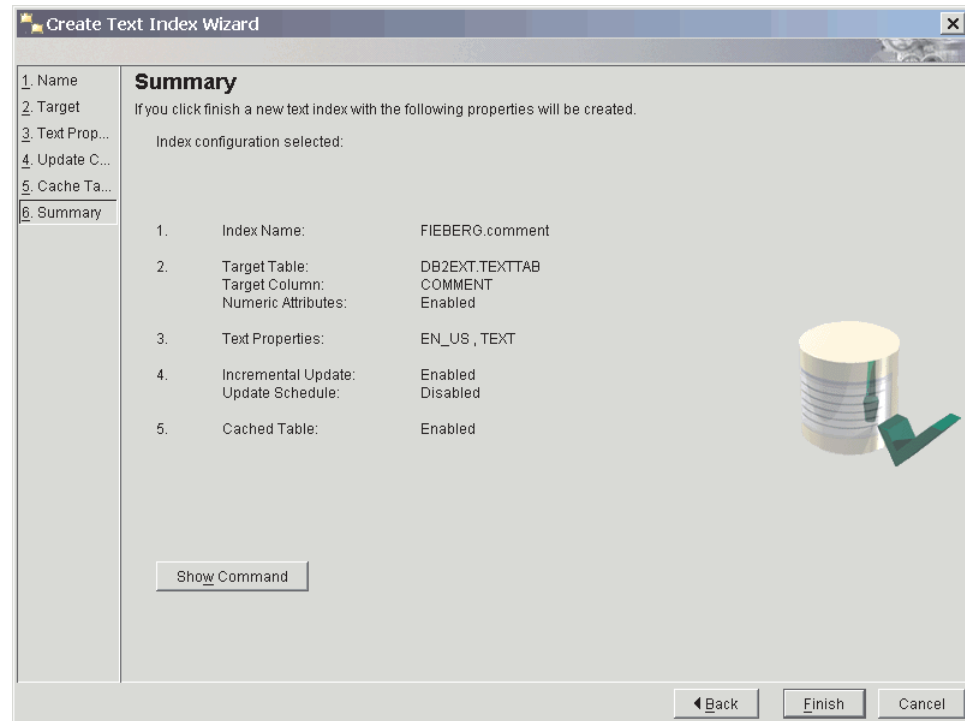


Figure 18. Assistant Création d'index texte : écran Récapitulatif

Cliquez sur le bouton **Afficher la commande** pour afficher les commandes exécutées lorsque vous cliquez sur le bouton **Fin**. Cette action aboutit à la création de l'index de texte.

Gestion d'un index de texte

Pour gérer les index de texte, sélectionnez l'index de texte dans le panneau de contenu, puis sélectionnez l'option de menu **Sélection**. Vous pouvez sélectionner l'une des commandes suivantes à partir du menu :

1. La commande ALTER pour modifier un index de texte.
2. La commande DROP pour supprimer un index de texte.
3. La commande UPDATE pour mettre à jour un index de texte.
4. La commande SHOW INDEX EVENTS pour afficher les événements d'index.
5. La commande ACTIVATE INDEX MEMORY pour activer une mémoire cache d'index.
6. La commande DEACTIVATE INDEX MEMORY pour désactiver une mémoire cache d'index.
7. La commande SHOW STATUS pour afficher l'état des index.

Notez que les commandes d'activation et de désactivation ne s'affichent que si vous avez créé l'index avec l'option de mémoire cache.

Modification d'un index de texte

Sélectionnez la commande **Alter** (Modifier) et une boîte de dialogue affiche une série d'écrans. Ils présentent les paramètres de l'index de texte. Notez que vous **ne** pouvez **pas** modifier tous les paramètres.

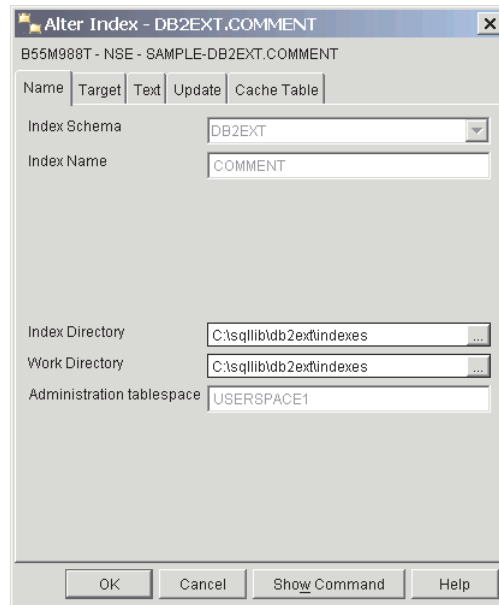


Figure 19. Boîte de dialogue Modification de l'index : onglet Nom

L'onglet **Nom** présente le nom et les configurations de stockage de l'index. Il est possible de modifier les répertoires d'index et de travail.

L'onglet **Cible** présente les paramètres des attributs numériques et cibles de l'index. Il n'est pas possible de changer ces configurations.

L'onglet **Texte** présente les configurations des documents texte. Il n'est pas possible de modifier ces paramètres.

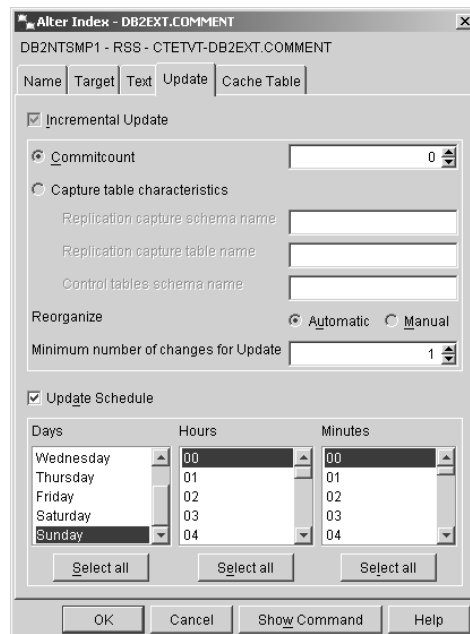


Figure 20. Boîte de dialogue Modification de l'index : onglet Mise à jour

L'onglet **Mise à jour** présente les caractéristiques de mise à jour de la configuration de l'index. La planification des mises à jour peut être modifiée. Si

L'index a été créé pour des mises à jour incrémentielles, vous pouvez également modifier le nombre minimal des modifications. Si l'index a été créé avec l'option Nombre de validations, vous pouvez également modifier la valeur de ce nombre.

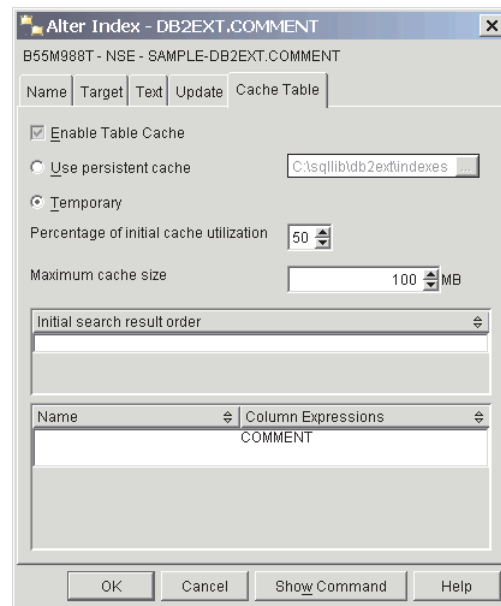


Figure 21. Boîte de dialogue Modification de l'index : onglet Table de mémoire cache

L'onglet **Table de mémoire cache** présente les paramètres de l'option de mémoire cache. Si la mémoire cache des résultats est déjà activée, vous pouvez modifier le répertoire persistant ou rendre temporaire la mémoire cache de l'index. Vous pouvez également modifier les valeurs Taille de la mémoire cache maximale et Pourcentage d'utilisation initiale de la mémoire cache.

Suppression d'un index de texte

Sélectionnez la commande **Drop** (Supprimer) et une boîte de dialogue affiche les index de texte disponibles.

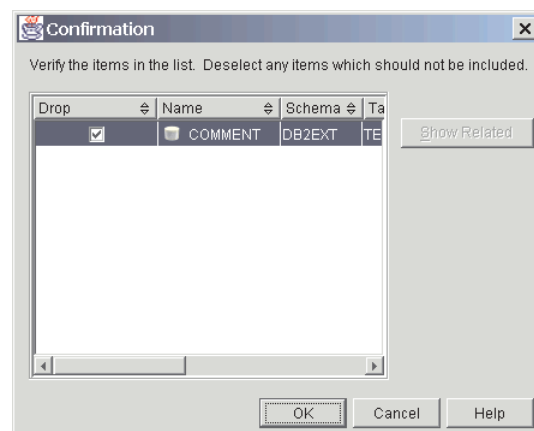


Figure 22. Boîte de dialogue Suppression d'index

Sélectionnez l'index et cliquez sur le bouton **OK**.

Mise à jour d'un index de texte

Sélectionnez la commande **Update** (Mettre à jour) et une boîte de dialogue affiche un certain nombre d'options de mise à jour.

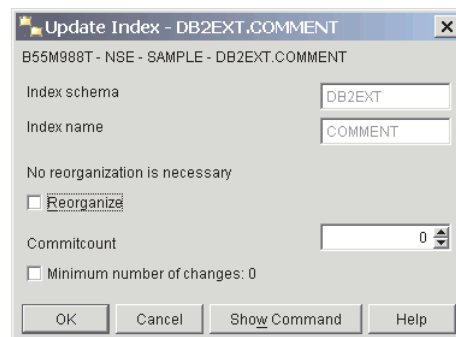


Figure 23. Boîte de dialogue Mise à jour de l'index

Il est possible de spécifier un nombre de validations pour l'opération de mise à jour. Si vous souhaitez mettre à jour le minimum spécifié au cours de la création ou de la modification de l'index, cochez la case Nombre minimal de modifications. Pour réorganiser l'index, cochez la case correspondante. Notez que si une réorganisation est recommandée, la case est cochée.

Affichage des événements d'index

Sélectionnez la commande **Affichage des événements d'index** et le contenu de la table des événements apparaît dans la boîte de dialogue.

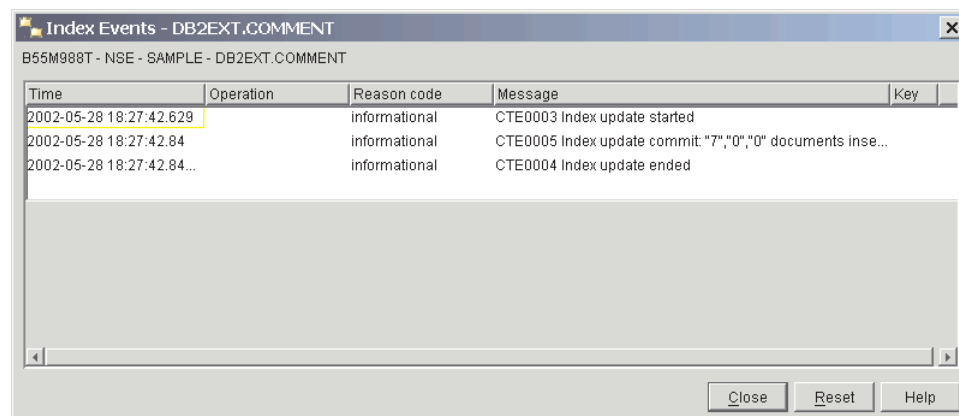


Figure 24. Boîte de dialogue Événements d'index

Pour effacer les événements de l'index, cliquez sur le bouton **Réinitialiser**.

Activation de la mémoire cache d'un index de texte

Sélectionnez la commande **Activate index memory** et une boîte de dialogue s'affiche.

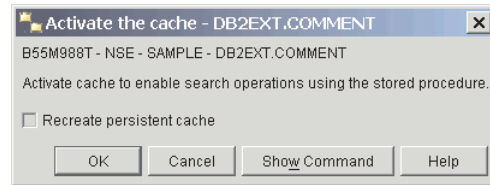


Figure 25. Boîte de dialogue Activation de la mémoire cache

Pour activer la mémoire cache, cliquez sur le bouton **OK**. Si vous souhaitez recréer la mémoire cache, cochez la case correspondante.

Désactivation de la mémoire cache d'un index de texte

Sélectionnez la commande **Deactivate index memory** et une boîte de dialogue s'affiche.



Figure 26. boîte de dialogue Désactivation de la mémoire cache

Pour libérer la mémoire cache, cliquez sur le bouton **OK**.

Affichage de l'état des index

Sélectionnez la commande **Affichage de l'état** et une boîte de dialogue affiche l'état de l'index de texte.

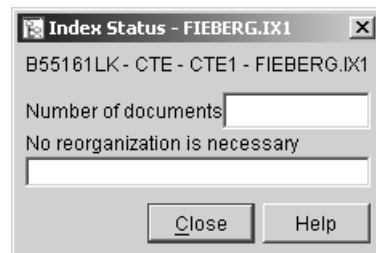


Figure 27. Boîte de dialogue Etat des index

Elle comporte des informations sur le nombre de documents indexés, l'indicateur de réorganisation recommandée et des informations supplémentaires sur l'index.

Chapitre 8. Recherche

DB2 Extension Net Search fournit les méthodes suivantes pour rechercher dans du texte :

Fonctions de recherche scalaires SQL

Elles permettent d'inclure des sous-requêtes de recherche de texte dans les requêtes SQL. L'Extension Net Search fournit ces fonctions en plus de celles normalement disponibles dans SQL.

Fonction de recherche de procédure mémorisée

Elle permet de renvoyer les tables de résultats prédéfinies et placées dans la mémoire cache.

Fonction de valeur table SQL

Vous pouvez utiliser cette recherche de la même manière que la procédure mémorisée.

Pour les fonctions de recherche scalaires SQL, le chapitre décrit les domaines suivants :

- Recherche de texte, à l'aide des fonctions CONTAINS, NUMBEROFMATCHES et SCORE.

Pour la description de la syntaxe, reportez-vous au Chapitre 15, «Fonction de recherche scalaire SQL et fonction de valeur table SQL», à la page 149.

- Spécification des arguments de la recherche à l'aide d'exemples avec la fonction CONTAINS.

Pour la description de la syntaxe, reportez-vous au Chapitre 14, «Syntaxe des arguments de recherche», à la page 141.

Pour la fonction de recherche de procédure mémorisée, le chapitre décrit les domaines suivants :

- Recherche de texte à l'aide de la recherche de procédure mémorisée.
- Pour la spécification des arguments de recherche, reportez-vous au Chapitre 14, «Syntaxe des arguments de recherche», à la page 141 pour la description des paramètres.

Pour la fonction de valeur table SQL, le chapitre décrit les domaines suivants :

- Recherche de texte à l'aide de la fonction de valeur table SQL et de la fonction HIGHLIGHT.
- Pour la description de la syntaxe, reportez-vous au Chapitre 15, «Fonction de recherche scalaire SQL et fonction de valeur table SQL», à la page 149.
- Pour la spécification des arguments de recherche, reportez-vous au Chapitre 14, «Syntaxe des arguments de recherche», à la page 141 pour la description des paramètres.

Il existe également des remarques sur les performances de recherche dont vous devez tenir compte.

Avant de procéder à la recherche, assurez-vous que toutes les étapes d'indexation, décrites au Chapitre 6, «Création et gestion d'un index de texte», à la page 31, impliquant les différents types de données sont pris en considération.

Remarque

Le shell système interprète les caractères spéciaux, tels que ?, (,), *, ! et ". De ce fait, si la commande contient ces caractères, vous devez utiliser des guillemets ou un caractère d'échappement.

L'exemple suivant montre une commande UNIX qui utilise des caractères spéciaux :

```
db2 "SELECT * from sample WHERE CONTAINS (DESCRIPTION, '\"enable\"') = 1"
```

Recherche de texte à l'aide des fonctions de recherche scalaires SQL

Cette section décrit, à l'aide d'exemples, comment utiliser les fonctions de recherche scalaires SQL de manières suivantes :

- Utilisation de la fonction CONTAINS pour effectuer une requête.
- Utilisation de la fonction NUMBEROFMATCHES pour déterminer le nombre d'occurrences trouvées dans un document texte.
- Utilisation de la fonction SCORE pour connaître la pertinence d'un document texte trouvé.

Pour la description de la syntaxe, reportez-vous au Chapitre 15, «Fonction de recherche scalaire SQL et fonction de valeur table SQL», à la page 149.

Réalisation d'une requête

Cet exemple montre comment la fonction CONTAINS recherche du texte dans la colonne comment de la table texttab. La valeur 1 est renvoyée si le texte répond à l'argument de recherche, sinon la valeur 0 est renvoyée.

```
SELECT AUTHOR,TITLE
      FROM DB2EXT.TEXTTAB
      WHERE CONTAINS(COMMENT, '"livre"') = 1
```

Dans cet exemple, la recherche porte sur le terme livre dans la colonne COMMENT.

Remarque

Pour augmenter les performances, il peut être avantageux d'ajouter des critères de recherche restrictifs, par exemple :

```
SELECT AUTHOR,TITLE
      FROM db2ext.texttab
      WHERE CONTAINS(COMMENT, '"livre"') = 1 AND PRICE < 20
```

Recherche et renvoi du nombre d'occurrences trouvées

Utilisez la fonction NUMBEROFMATCHES pour déterminer le nombre d'occurrences de la valeur recherchée, trouvé dans chaque document texte.

```
SELECT AUTHOR,TITLE,
      NUMBEROFMATCHES(COMMENT, '"livre"')
      FROM DB2EXT.TEXTTAB
```

NUMBEROFMATCHES renvoie un nombre entier.

Recherche et renvoi du score d'un document texte trouvé

SCORE est une valeur absolue qui indique avec quelle pertinence le document correspond à la valeur recherchée par rapport aux autres documents trouvés. La valeur indique le nombre d'occurrences trouvées dans le document en rapport avec la taille du document. Dans l'exemple suivant, vous pouvez obtenir le score d'un document trouvé à l'aide de la fonction SCORE :

```
WITH TEMPTABLE(docid,score)
  AS (SELECT docid,
            SCORE(COMMENT,'livre')
      FROM DB2EXT.TEXTTAB)
SELECT *
  FROM TEMPTABLE
 WHERE score > 0
 ORDER BY score ASC
```

SCORE renvoie une valeur DOUBLE comprise entre 0 et 1.

Remarque

Les fonctions de recherche CONTAINS, SCORE et NUMBEROFMATCHES ne peuvent pas être utilisées pour les index créés sur des vues.

Dans un environnement DB2 réparti, les valeurs SCORE sont différentes :

- Dans un environnement non réparti, tous les documents se trouvent dans une seule table. La valeur SCORE est basée sur une seule table et les relations des documents avec tous les autres documents de la table.
- Dans un environnement DB2 réparti, tous les documents se trouvent sur des noeuds différents. Pendant l'indexation, seuls les documents locaux sont utilisés pour construire des index de texte qui se trouvent en local sur tous les noeuds. Dans ce cas, la valeur SCORE est basée sur les relations des documents avec tous les documents sur un seul des noeuds.

Spécification des arguments de recherche SQL

Les fonctions CONTAINS, NUMBEROFMATCHES et SCORE utilisent toutes des arguments de recherche. Cette section utilise la fonction CONTAINS pour montrer différents exemples d'arguments de recherche dans les fonctions DB2 Extension Net Search.

Pour la description de la syntaxe, reportez-vous au Chapitre 14, «Syntaxe des arguments de recherche», à la page 141.

Recherche de termes dans n'importe quelle séquence

L'argument de recherche peut être composé d'un ou de plusieurs termes. Un moyen pour combiner plusieurs termes recherchés est de les assembler par des virgules, comme suit :

```
SELECT AUTHOR,TITLE
  FROM DB2EXT.TEXTTAB
 WHERE CONTAINS(COMMENT,
                '("enfant", "dinosaur")') = 1
```

Cette forme d'argument de recherche le texte contenant l'un des termes recherchés. Dans les termes logiques, un opérateur OR assemble les termes recherchés.

Recherche à l'aide des opérateurs booléens AND et OR

Vous pouvez associer les termes recherchés à d'autres termes recherchés à l'aide des opérateurs booléens "&" (AND) et "|" (OR) :

```
SELECT AUTHOR, TITLE
      FROM DB2EXT.TEXTTAB
      WHERE CONTAINS(COMMENT,
                     '"auteur" | "pulitzer"')
= 1
```

Vous pouvez également associer plusieurs termes à l'aide d'opérateurs booléens :

```
SELECT AUTHOR, TITLE
      FROM DB2EXT.TEXTTAB
      WHERE CONTAINS(COMMENT,
                     '"auteur" | "pulitzer" & "livre"') = 1
```

Si vous utilisez plusieurs opérateurs booléens, ils sont traités de la gauche vers la droite. Cependant, l'opérateur logique AND (&) édite un lien plus fort que l'opérateur logique OR (|). Vous pouvez voir cette évaluation dans l'exemple suivant qui ne comporte pas de parenthèse :

```
"livre" & "pulitzer" | année & "auteur"
```

Dans ce cas, l'Extension Net Search évalue les opérateurs booléens de la manière suivante :

```
("livre" & "pulitzer") | (année & "auteur")
```

Aussi, pour évaluer correctement les opérateurs booléens, vous devez ajouter des parenthèses :

```
"livre" & ("pulitzer" | année) & "auteur"
```

Vous pouvez également associer les opérateurs booléens aux termes recherchés qui sont liés par des virgules :

```
("auteur", "pulitzer") & "livre"
```

Dans ce cas, cependant, la virgule est interprétée comme un opérateur booléen OR :

```
("auteur" | "pulitzer") & "livre"
```

Pour plus d'informations, reportez-vous à la section «Recherche à l'aide de l'opérateur booléen NOT» à la page 76.

Recherche floue

Dans une recherche "floue", la recherche s'effectue sur les mots orthographiés de la même manière que le terme recherché.

```
SELECT AUTHOR, TITLE
      FROM DB2EXT.TEXTTAB
      WHERE CONTAINS(COMMENT,
                     'fuzzy form of 80 "pullitzer"') =1
```

Dans cet exemple, la recherche peut trouver une occurrence du mot pulitzer mal orthographié.

Le niveau de correspondance, dans l'exemple "80", spécifie le degré de précision. Utilisez une recherche floue si les mots peuvent être mal orthographiés dans le document. C'est souvent le cas si le document a été créé à l'aide d'un appareil de reconnaissance optique de caractères (ROC) ou une entrée phonétique. Utilisez une

valeur comprise entre 1 et 100, 100 correspondant à une occurrence exacte et en-dessous de 80 la correspondance est inexacte.

Remarque

Si la recherche floue ne présente pas un degré de précision convenable, recherchez les parties d'un terme à l'aide du caractère de masquage.

Recherche de parties d'un terme (caractères génériques)

Les caractères de masquage, connus sous le nom de caractères "génériques" permettent une recherche plus flexible. Ceci est obtenu en augmentant le nombre de documents texte trouvés par une recherche.

DB2 Extension Net Search utilise deux caractères génériques : le signe pour cent (%) et le trait de soulignement (_).

- Le signe % représente **n'importe quel nombre de caractères arbitraires**. L'exemple suivant montre le signe % utilisé comme caractère générique dans le milieu d'un terme recherché :

```
SELECT AUTHOR, TITLE
      FROM DB2EXT.TEXTTAB
     WHERE CONTAINS(COMMENT, '"sus%ce"') = 1
```

Ce terme recherché trouve les documents texte contenant le mot "suspense".

- _ représente **un caractère** dans un terme recherché. L'exemple suivant trouve également les documents texte contenant le mot "suspense".

```
SELECT AUTHOR, TITLE
      FROM DB2EXT.TEXTTAB
     WHERE CONTAINS(COMMENT, '"su_pence"') = 1
```

Remarque

Utilisez les caractères génériques avec parcimonie car ils augmentent considérablement la taille de la liste des résultats, diminuant ainsi les performances et renvoyant des résultats de recherche inattendus.

Recherche de termes contenant déjà un caractère générique

Si vous recherchez un terme contenant le signe "%" ou le caractère "_", vous devez le faire précéder d'un caractère, dit d'*échappement*. Vous pouvez identifier le caractère d'échappement à l'aide du mot clé ESCAPE.

Dans l'exemple suivant, le caractère d'échappement est "!" :

```
SELECT AUTHOR, TITLE
      FROM DB2EXT.TEXTTAB
     WHERE CONTAINS(COMMENT,
                    '"100!%" ESCAPE "!"') = 1
```

Recherche de termes dans une séquence fixe

Si vous recherchez "clé primaire", vous ne trouverez les deux termes que s'ils sont adjacents et se présentent dans la séquence indiquée :

```
SELECT AUTHOR, TITLE
      FROM DB2EXT.TEXTTAB
     WHERE CONTAINS(COMMENT, '"clé primaire"') = 1
```

Recherche de termes dans la même phrase ou le même paragraphe

L'exemple suivant montre un argument de recherche qui trouve les documents texte dans lesquels les termes recherchés se trouvent dans la même phrase :

```
SELECT AUTHOR, TITLE
      FROM DB2EXT.TEXTTAB
     WHERE CONTAINS(COMMENT,
                    '"web" IN SAME SENTENCE AS "disque"') = 1
```

Il est également possible de rechercher plus de deux mots se présentant ensemble. Dans l'exemple suivant, la recherche s'effectue sur deux mots figurant dans le même paragraphe :

```
SELECT AUTHOR, TITLE
      FROM DB2EXT.TEXTTAB
     WHERE CONTAINS(COMMENT,
                    '"ordinateur" IN SAME PARAGRAPH AS "web"') = 1
```

Recherche de termes dans des sections de documents structurés

L'exemple suivant montre un argument de recherche qui trouve les documents texte dans lesquels le terme recherché IBM figure dans le sous-section H2 de documents structurés :

```
SELECT CATEGORY, DATE
      FROM DB2EXT.HTMLTAB
     WHERE CONTAINS(HTMLFILE,
                    'SECTIONS ("H2") "IBM"') = 1
```

Notez que les noms de section font la distinction entre minuscules et majuscules. Assurez-vous que le nom de section dans le fichier modèle et dans la requête sont identiques. Pour plus de détails, reportez-vous au Chapitre 9, «Utilisation de documents structurés», à la page 83.

Recherche à l'aide de l'opérateur booléen NOT

L'opérateur booléen NOT peut être utilisé pour exclure de la recherche des documents texte spécifiques :

```
SELECT AUTHOR, TITLE
      FROM DB2EXT.TEXTTAB
     WHERE CONTAINS(COMMENT,
                    '("auteur", "pulitzer") & NOT "livre"') = 1
```

Dans cet exemple, tous les documents texte contenant le terme "livre" sont exclus de la recherche sur "auteur" ou "pulitzer".

Recherche de thésaurus

La recherche de thésaurus est une fonction d'extension de recherche de terme puissante dans DB2 Extension Net Search. Les termes supplémentaires que vous recherchez sont tirés d'un thésaurus constitué par vous-même, aussi vous disposez d'un contrôle direct sur les termes. Ainsi, par exemple, une recherche sur "base de données" peut trouver des termes, tels que "référentiel" et "DB2".

Utilisez ce type de recherche pour des domaines d'intérêt spécifiques dans lesquels vous effectuez des recherches fréquentes et qui produisent des résultats considérablement plus efficaces.

Pour plus d'informations et une description de la création d'un thésaurus, reportez-vous au Chapitre 10, «Utilisation d'un thésaurus pour étendre les termes recherchés», à la page 85. L'exemple suivant montre la syntaxe permettant d'utiliser l'extension du thésaurus.

Cet exemple prend le terme "produit" et l'étend, ajoutant toutes les relations au terme trouvées dans le thésaurus "nsesamplethes". Ici, "marketing" est ajouté à la recherche.

```
SELECT CATEGORY, DATE
      FROM DB2EXT.HTMLTAB
      WHERE CONTAINS(HTMLFILE,
                     'THESAURUS "nsesamplethes"
                     EXPAND RELATED
                     TERM OF "produit"') = 1
```

L'exemple suivant prend le terme "produit". La recherche est étendue à tous les *synonymes* du terme recherché.

```
SELECT CATEGORY, DATE
      FROM DB2EXT.HTMLTAB
      WHERE CONTAINS(HTMLFILE,
                     'THESAURUS "nsesamplethes"
                     EXPAND SYNONYM
                     TERM OF "produit"') = 1
```

Recherche d'attributs numériques

Il est possible d'effectuer une recherche sur les attributs numériques stockés dans un document structuré, à l'aide de la syntaxe suivante :

```
SELECT AUTHOR, TITLE
      FROM DB2EXT.TEXTTAB
      WHERE CONTAINS(COMMENT,
                     'ATTRIBUTE "PRICE" between 9 and 20') = 1
```

Notez que les noms d'attribut font la distinction entre majuscules et minuscules. Assurez-vous que le nom de l'attribut contenu dans le fichier modèle est identique à celui de la requête. Pour plus de détails, reportez-vous au Chapitre 9, «Utilisation de documents structurés», à la page 83.

Recherche de texte libre

La "recherche de texte libre" consiste à exprimer le terme recherché sous forme de texte de format libre. Une expression ou une phrase décrit en langue normale le sujet à rechercher. La séquence de mots dans une requête en texte libre n'est pas importante. Cependant, dans le cas d'un ensemble de termes de requête, au moins l'un des termes doit figurer dans les documents sur lesquels la recherche doit être effectuée.

Notez que les caractères ou les mots génériques ne sont pas pris en charge pour les chaînes de recherche dans un argument en texte libre.

Par exemple :

```
SELECT AUTHOR, TITLE, SCORE(COMMENT,
                              'IS ABOUT EN_US "quelque chose en rapport avec les dinosaures"')
      FROM DB2EXT.TEXTTAB
      WHERE CONTAINS(COMMENT,
                     'IS ABOUT EN_US "quelque chose en rapport avec les dinosaures"') = 1
```

En associant la requête avec la fonction de recherche SCORE, la recherche renverra les documents contenant uniquement le mot "quelque chose".

Exemples supplémentaires de syntaxe de recherche

Pour vous familiariser avec les exemples supplémentaires de syntaxe de recherche, utilisez le fichier de commandes CLP, appelé `search`. Il contient des exemples de fonctions de recherche DB2 Extension Net Search qui s'exécutent sur la table échantillon.

Pour exécuter l'exemple, utilisez la syntaxe suivante :

```
db2 -tvf search
```

Si la table et les index n'ont pas été créés, exécutez l'une des commandes suivantes :

- Sous UNIX : `nsesample` dans le répertoire `<instance_owner_home>/<sqllib>/samples/db2ext`.
- Sous Windows : `nsesample (.bat)` dans le répertoire `<sqllib>/samples/db2ext`.

Recherche de texte à l'aide d'une recherche de procédure mémorisée

Utilisez l'interface de recherche de procédure mémorisée si votre application nécessite un sous-ensemble du résultat de la recherche de texte, mais de manière hautement performante. N'utilisez pas la procédure mémorisée si tous les résultats sont requis ou si un grand nombre de documents doit être indexé. En effet, comme des parties de la table utilisateur sont copiées dans la mémoire, une grande quantité de mémoire doit être disponible.

Vous pouvez utiliser la procédure mémorisée pour demander en premier les résultats de 0 à 20, puis de 21 à 40, etc. de manière semblable à la navigation du curseur. En associant cette capacité de curseur à l'utilisation d'une mémoire cache (calculée pendant l'indexation), la recherche est extrêmement rapide, en particulier si aucune jointure n'est nécessaire.

Si vous avez l'intention d'utiliser la procédure mémorisée, veillez à tenir compte des options suivantes :

- Dans un environnement DB2 réparti, vous devez spécifier un espace table sur un seul noeud pour la procédure mémorisée et appeler sur ce noeud de manière explicite.
- Les options des résultats de la recherche dans la mémoire cache ont été spécifiées pendant la commande `CREATE INDEX`.
- Les conditions requises pour la mémoire partagée présentes et futures, impliquant éventuellement des mises à jour incrémentielles, ont été totalement prises en compte. Pour plus d'informations, reportez-vous à la section «Création d'un index de texte pouvant être utilisé par la recherche de procédure mémorisée» à la page 39.
- La mémoire cache de l'index a été activée à l'aide de la commande `db2text activate`.

L'exemple suivant illustre une recherche de procédure mémorisée :

```
db2 "call db2ext.textSearch('livre','DB2EXT','COMMENT',0,2,1,1,?,?)"
```

Le premier paramètre est le terme recherché. La syntaxe est exactement la même que celle présente dans les fonctions SQL. Spécifiez ensuite le nom et le schéma de l'index. Si vous n'avez pas masqué le nom, il est converti en majuscules. Les deux nombres suivants indiquent la probabilité d'obtention du résultat en segments. Les

deux valeurs en nombre entier suivantes spécifient si le score et les informations sur les occurrences sont requis. Les deux dernières valeurs sont des valeurs de sortie.

Pour plus de détails sur les paramètres, reportez-vous au Chapitre 16, «Fonction de recherche de procédure mémorisée», à la page 161.

Remarque

Si vous avez besoin d'ensembles de résultats plus importants, vous devez utiliser un espace table utilisateur. Si aucun espace table n'est disponible, créez-en un. L'exemple suivant crée un espace table sur une plateforme UNIX :

```
db2 "create user temporary tablespace tempts managed by system
    using ('/work/tempts.ts')"
```

Dans un environnement DB2 réparti, vous devez spécifier un espace table pour les tables d'administration sur un seul noeud pour la procédure mémorisée et appeler sur ce noeud de manière explicite.

Recherche de texte à l'aide de la fonction de valeur table SQL

Utilisez la fonction de valeur table SQL si vous implémentez une interface, mais n'avez pas besoin que tous les résultats soient renvoyés ou si vous ne disposez pas de toute la mémoire réelle requise pour utiliser l'interface de procédure mémorisée.

Deux fonctions de valeur table SQL sont disponibles, toutes deux appelées `db2ext.textsearch`. L'une est dotée de paramètres supplémentaires à utiliser avec la fonction `db2ext.highlight`. Pour plus d'informations, reportez-vous à la section «Utilisation de la fonction de mise en évidence (HIGHLIGHT)» à la page 80.

Avec la fonction de valeur table SQL, vous disposez de la même interface de curseur que la procédure mémorisée pour obtenir que des parties du résultat. Cependant, vous devez effectuer la jointure des résultats avec la table utilisateur. L'exemple suivant en est l'illustration :

```
db2 "select docid , author, score from TABLE(db2ext.textsearch('\livre\' ',
    'DB2EXT','COMMENT',3,2,cast(NULL as integer))) as t, db2ext.texttab u
    where u.docid = t.primkey"
```

Les valeurs suivantes sont celles qui peuvent être renvoyées de la fonction de valeur table SQL :

```
--> primKey
<type de clé primaire simple>
la clé primaire

--> score          DOUBLE
la valeur du score dans le document trouvé

--> NbResults      INTEGER
le nombre total de résultats trouvés (même valeur pour toutes les lignes)

--> numberOfMatches INTEGER
le nombre d'occurrences dans le document
```

Remarque

Notez qu'une seule colonne de clé primaire est autorisée. Pour plus de détails sur les paramètres, reportez-vous à la section «DB2EXT.TEXTSEARCH» à la page 153.

Utilisation de la fonction de mise en évidence (HIGHLIGHT)

Deux fonctions de valeur table SQL sont disponibles, toutes deux appelées `db2ext.textsearch`. Pour utiliser la fonction `db2ext.highlight`, vous devez utiliser la fonction `db2ext.textsearch` avec les paramètres supplémentaires `numberOfHits` et `hitInformation`.

Dans cet exemple, appelez la fonction `db2ext.highlight` pour afficher le document complet sans mettre en évidence les occurrences trouvées par la fonction `db2ext.textsearch`.

```
select  p.docid,
        db2ext.highlight(p.comment, t.hitinformation, ' WINDOW_NUMBER = 0 ')
        as highlight
from    DB2EXT.TEXTTAB p,
        table (db2ext.textsearch('"bestseller" | "peacekeeping" | "soldiers"
        | "attention"', 'DB2EXT', 'COMMENT', 0, 20,
        cast(NULL as INTEGER), 10)) t
where   p.docid = t.primkey and p.docid = 2
```

L'argument de recherche renvoie le résultat suivant :

DOCID HIGHLIGHT

```
2      A New York Times bestseller about peacekeeping soldiers called
      "Keepers" who devise a shocking scheme to get the worlds
      attention after their tour of duty ends.
```

1 record(s) selected.

Remarque

Dans tous les exemples `db2ext.highlight`, la fonction de table `db2ext.textsearch` recherche l'un des termes suivants : "bestseller", "peacekeeping", "soldiers", ou "attention".

Dans cet exemple, appelez la fonction `db2ext.highlight` pour afficher le document complet et mettre en évidence toutes les occurrences trouvées par la fonction `db2ext.textsearch`.

```
select  p.docid,
        db2ext.highlight(p.comment, t.hitinformation, ' WINDOW_NUMBER = 0,
        TAGS = ("<bf>","</bf>" ) ') as highlight
from    DB2EXT.TEXTTAB p,
        table (db2ext.textsearch('"bestseller" | "peacekeeping" | "soldiers"
        | "attention"', 'DB2EXT', 'COMMENT', 0, 20,
        cast(NULL as INTEGER), 10)) t
where   p.docid = t.primkey and p.docid = 2
```

L'argument de recherche renvoie le résultat suivant :

DOCID HIGHLIGHT

```
2      A New York Times <bf>bestseller</bf> about <bf>peacekeeping</bf>
```


Exemples supplémentaires de syntaxe de recherche

`<bf>soldiers</bf>` called "Keepers" who devise a shocking scheme to get the worlds `<bf>attention</bf>` after their tour of duty ends.

1 record(s) selected.

Dans cet exemple, appelez la fonction `db2ext.highlight` pour afficher au maximum 10 parties (fenêtres) du document. La taille de toutes les fenêtres est de 24, ce qui correspond approximativement à 12 octets de données de chaque côté de l'occurrence. De plus, les occurrences trouvées par la fonction de table `db2.textsearch` sont mises en évidence.

```
select p.docid,  
       db2ext.highlight(p.comment, t.hitinformation, ' WINDOW_NUMBER = 10,  
       WINDOW_SIZE = 24, TAGS = ("<bf>", "</bf>" ) ') as highlight  
from DB2EXT.TEXTTAB p,  
     table (db2ext.textsearch('"bestseller" | "peacekeeping" | "soldiers"  
       | "attention"', 'DB2EXT', 'COMMENT', 0, 20,  
       cast(NULL as INTEGER), 10)) t  
where p.docid = t.primkey and p.docid = 2
```

L'argument de recherche renvoie le résultat suivant :

DOCID HIGHLIGHT

```
2      York Times <bf>bestseller</bf> about <bf>peacekeeping</bf> ...  
      <bf>peacekeeping</bf> <bf>soldiers</bf> called "Keepers" ... the  
      worlds <bf>attention</bf> after their
```

1 record(s) selected.

La première occurrence trouvée est `<bf>bestseller</bf>` et cette occurrence détermine la première fenêtre. La seconde occurrence, `<bf>peacekeeping</bf>` ne se trouve qu'à 8 octets de la première occurrence et totalement incluse dans la première fenêtre. La troisième occurrence, `<bf>soldiers</bf>` est en dehors de la première fenêtre et détermine une nouvelle fenêtre. Comme la seconde occurrence, `<bf>peacekeeping</bf>` n'est qu'à 2 octets sur la gauche de l'occurrence `<bf>soldiers</bf>`, elle est également incluse dans la seconde fenêtre et mise en évidence. La quatrième occurrence `<bf>attention</bf>` se trouve en dehors de la seconde fenêtre et en détermine donc une nouvelle. Comme aucune occurrence précédente, ni supplémentaire n'est contenue dans la taille de cette fenêtre, seules les données entourant l'occurrence sont contenues dans la fenêtre.

De plus, comme aucun paramètre `WINDOW_SEPARATOR` n'est spécifié, le séparateur de fenêtre par défaut, " ... " est adopté pour séparer les trois fenêtres du document.

Remarque

Pour garantir des performances élevées lors de l'utilisation de la fonction `db2ext.highlight`, l'utilisateur doit limiter les résultats de la recherche dans la fonction de valeur table `db2ext.textsearch`.

Pour plus de détails sur les paramètres, reportez-vous à la section «DB2EXT.HIGHLIGHT» à la page 157.

Recherche sur plusieurs colonnes

Au cas où vous devez créer un index de texte sur plusieurs colonnes, le moyen le plus facile est d'utiliser la fonction scalaire SQL et d'associer les recherches sur cette colonne. En voici un exemple :

```
SELECT AUTHOR,TITLE
      FROM DB2EXT.TEXTTAB
     WHERE CONTAINS(COMMENT,
                    '"livre"')=1 and CONTAINS(AUTHOR,'"Mike"')=1
```

La situation se complique pour une fonction de valeur table, comme vous devez utiliser une union pour des raisons de performances. Une autre possibilité avec la fonction de valeur table consiste à utiliser une vue et à associer vos colonnes de table dans une colonne de vue pour créer un index de texte. Ainsi, vous évitez d'avoir deux appels séparés de recherche de texte.

La combinaison de colonnes de texte peut améliorer les performances. Cependant, cela dépend fortement de vos conditions de recherche personnelles.

Remarques sur les performances

Afin d'améliorer les performances au cours de la recherche, tenez compte des éléments suivants :

- Lors de la recherche dans SQL :
 - Si vous constatez une diminution des performances, utilisez l'instruction explain pour vérifier le plan de traitement de l'Optimiseur DB2.
 - La recherche paramétrique peut accélérer la recherche, spécialement si vous utilisez d'autres prédicats de recherche pour réduire la taille des résultats.
 - Utilisez le mot clé RESULT LIMIT si vous n'avez pas besoin de tous les résultats.
- Lors de la recherche avec la procédure mémorisée :
 - L'expression de la table de mémoire cache spécifiée étant copiée de la base de données dans la mémoire, assurez-vous que votre poste de travail dispose d'une quantité de mémoire disponible suffisante pour ces données. Si la quantité de mémoire est insuffisante, l'espace de pagination est utilisé, ce qui diminue les performances de la recherche.
- Si vous utilisez la fonction NUMBEROFMATCHES ou SCORE sans la fonction CONTAINS, les performances de requêtes peuvent diminuer. De plus, pour éviter un double traitement, assurez-vous que la chaîne de caractères figurant dans la fonction CONTAINS corresponde exactement à la chaîne de caractères utilisée dans la fonction NUMBEROFMATCHES ou SCORE.

Remarque

Pour obtenir les conseils les plus récents en matière de performances, consultez le site Web DB2 Extension Net Search à l'adresse :
www.ibm.com/software/data/db2/extendernetsearch/index.html

Chapitre 9. Utilisation de documents structurés

DB2 Extension Net Search permet d'indexer des zones texte ou numériques et d'effectuer des recherches dessus, telles que le titre, l'auteur ou la description dans un document structuré. Les documents peuvent être au format XML, Outside-In ou HTML, ou contenir des balises définies par l'utilisateur (GGP).

Utilisez des balises de marquage et leurs noms de zone dans un *modèle de document* pour définir quelles zones des documents sont indexées, et de ce fait disponibles pour la recherche. Vous pouvez utiliser le nom de la zone (également connu comme nom de section) dans les requêtes sur cette zone.

Pour pouvoir rechercher dans ces zones, spécifiez un format et un modèle de fichier (FORMAT AND MODEL FILE) lorsque vous créez l'index de texte contenant les documents. Pour plus d'informations, reportez-vous à la section «CREATE INDEX» à la page 114.

Pour plus d'informations sur la création et la définition de modèles de document, reportez-vous au Chapitre 17, «Prise en charge de documents structurés», à la page 165.

Pour la syntaxe du modèle de document, reportez-vous à l'Annexe G, «Référence du modèle de document», à la page 229.

Chapitre 10. Utilisation d'un thésaurus pour étendre les termes recherchés

Il est possible d'élargir une requête en recherchant non seulement un terme spécifique, mais aussi les termes qui lui sont relatifs. Ce processus peut être automatisé à l'aide des fonctions de l'Extension Net Search pour rechercher et extraire d'un thésaurus les termes relatifs au terme recherché. Un thésaurus est un dictionnaire contrôlé de termes sémantiquement associés qui couvrent généralement un domaine spécifique.

DB2 Extension Net Search permet d'étendre un terme recherché en ajoutant des termes supplémentaires tirés d'un thésaurus que vous avez précédemment créé. Pour savoir comment utiliser l'extension d'un thésaurus dans une requête, reportez-vous au Chapitre 14, «Syntaxe des arguments de recherche», à la page 141.

La création d'un thésaurus en vue d'une utilisation dans une application de recherche nécessite qu'un fichier de définitions du thésaurus soit compilé dans un format interne, le dictionnaire thésaurus.

Le présent chapitre aborde les domaines suivants :

- **Structure d'un thésaurus**

Un thésaurus est structuré comme un réseau de noeuds, reliés par des relations. Cette section décrit les relations prédéfinies du Extension Net Search et comment définir vos propres relations.

- **Création et compilation d'un thésaurus**

La description ci-dessous est celle de la syntaxe d'un fichier de définitions du thésaurus et des outils utilisés pour le compiler en dictionnaire thésaurus.

Structure d'un thésaurus

Un thésaurus est structuré comme un réseau de noeuds, reliés par des relations. Le Moteur de recherche de texte recherche un terme dans un thésaurus en commençant par le terme en question, puis en suivant un chemin dessiné par les relations au terme et en délivrant les termes trouvés à mesure du processus.

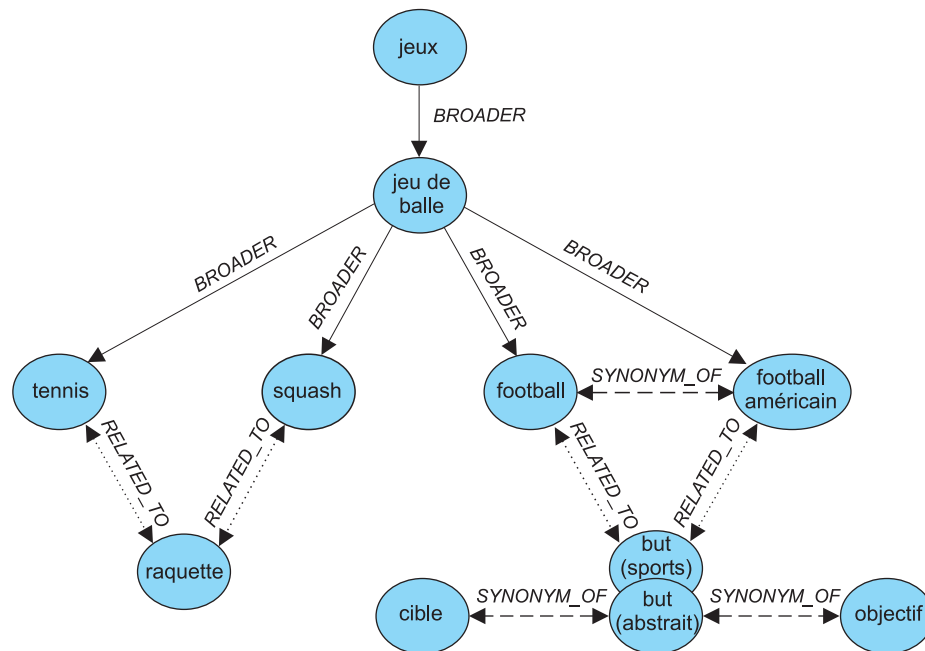


Figure 28. Exemple de structure d'un thésaurus

Les entrées du thésaurus sont connectées par des relations. Les noms de relation, tel que **BROADER**, permettent de limiter l'extension à certaines lignes désignées dans la hiérarchie des relations. Certaines relations sont bidirectionnelles, d'autres unidirectionnelles. **BROADER**, par exemple, est le nom d'une relation unidirectionnelle.

Relations de thésaurus prédéfinies

Les relations sont celles prédéfinies dans l'Extension Net Search:

- **Relations associatives**

Une relation associative est une relation bidirectionnelle entre deux termes qui n'expriment pas le même concept, mais qui se rapportent l'un à l'autre.

Relation associative prédéfinie : **RELATED_TO**

Exemples :

tennis **RELATED_TO** raquette
 football américain **RELATED_TO** but (sports)

- **Relations synonymes**

Une relation synonyme est une relation bidirectionnelle entre deux termes ayant la même signification ou une signification similaire et qui peuvent être utilisés l'un pour l'autre. Cette relation peut être utilisée, par exemple, entre un terme et son abréviation.

Relation synonyme prédéfinie : **SYNONYM_OF**

Exemples :

point **SYNONYM_OF** tache
 GB **SYNONYM_OF** Grande-Bretagne

La figure 28 montre deux termes **but** dans le même thésaurus. Le premier est spécifié par le commentaire (sports), l'autre par le commentaire (abstrait). Même si les termes ont la même orthographe, les relations synonymes peuvent se connecter à différents groupes de mots. Ceci peut être modélisé à l'aide de différentes relations lors de la définition du thésaurus.

- **Relations hiérarchiques**

Une relation hiérarchique est une relation unidirectionnelle entre deux termes, dont l'un a une signification plus large (plus générale) que l'autre. Selon sa direction, la relation peut être utilisée pour rechercher des termes plus spécialisés ou plus généraux.

Relations hiérarchiques prédéfinies :

- NARROWER pour modéliser des relations de réduction

Les relations NARROWER sont destinées à modéliser une séquence de termes plus spécialisés. Plus vous suivez une telle relation, plus les termes trouvés sont spécifiques. Par exemple, si vous recherchez le terme maison avec une relation NARROWER, le résultat peut être gratte-ciel palais église chapelle cathédrale etc. dans une liste de termes de plus en plus spécialisés.

- BROWDER pour modéliser des relations d'élargissement

Les relations BROWDER sont destinées à modéliser une séquence de termes de plus en plus généraux. Plus vous suivez une telle relation, moins les termes trouvés sont spécifiques. Par exemple, si vous recherchez le terme maison avec une relation BROWDER, le résultat peut être objet de construction d'édifice etc. dans une liste de termes de plus en plus généraux.

Définition de vos propres relations

L'Extension Net Search permet de définir vos propres relations de thésaurus RELATED_TO, NARROWER et BROWDER. Chaque nom de relation devant être unique, vous devez qualifier de tels noms de relations en ajoutant un numéro unique, tel que : RELATED_TO(42).

Vous pouvez utiliser le même numéro de relation pour définir une relation d'un autre type, tel que NARROWER(42). Le numéro 0 est utilisé pour faire référence aux relations prédéfinies de l'Extension Net Search.

Création et compilation d'un thésaurus

Utilisez les étapes ci-dessous pour créer un thésaurus pouvant être utilisé par les fonctions de l'Extension Net Search :

1. Création d'un fichier de définitions du thésaurus.
2. Compilation du fichier de définitions en dictionnaire thésaurus.

Création d'un fichier de définitions du thésaurus

Pour créer votre propre thésaurus, vous devez d'abord définir son contenu dans un fichier de définitions à l'aide d'un éditeur de texte.

Restrictions : La longueur du nom de fichier, y compris son extension, ne doit pas dépasser 256 caractères. Vous pouvez avoir plusieurs thésaurus dans le même répertoire, mais il est préférable de disposer d'un répertoire par thésaurus.

Un modèle de fichier de définitions de thésaurus anglais nsesamplethes.def est fourni. Sous Windows, le répertoire du thésaurus est :

<sql1lib>\db2ext\thes

Sous UNIX, le répertoire du thésaurus est :

<instance_owner_home>/sql1lib/db2ext/thes

Les premiers groupes de définitions de ce fichier sont les suivants :

```
:WORDS
  accounting
  .RELATED_TO account checking
  .RELATED_TO sale management
  .SYNONYM_OF account
  .SYNONYM_OF accountant

:WORDS
  acoustics
  .RELATED_TO signal processing

:WORDS
  aeronautical equipment
  .SYNONYM_OF turbocharger
  .SYNONYM_OF undercarriage

:WORDS
  advertising
  .RELATED_TO sale promotion
  .SYNONYM_OF advertisement
:
:
:
```

Figure 29. Extrait du modèle de fichier de définitions du thésaurus

Pour la syntaxe de chaque groupe de définitions, reportez-vous au Chapitre 18, «Prise en charge du thésaurus», à la page 179.

Chaque membre doit être écrit sur une seule ligne. Chaque terme associé doit être précédé du nom de la relation. Si les termes du membre sont liés les uns aux autres, spécifiez une relation de membre.

La longueur des termes du membre et des termes associés est limitée à 64 caractères. Les caractères mono-octet et les caractères à deux octets de la même lettre sont considérés comme identiques. Aucune distinction n'est faite entre les majuscules et les minuscules. Un terme peut contenir un espace et un point mono-octet "." ou deux points ":" peuvent être utilisés.

Les relations définies par l'utilisateur sont toutes basées sur le type *associatif*. Elles sont identifiées par des numéros uniques compris entre 1 et 128.

Compilation d'un fichier de définitions en dictionnaire thésaurus

Pour compiler un fichier de définitions du thésaurus, exécutez la commande `db2extth`. Pour la syntaxe de la commande, reportez-vous à la section «DB2EXTTH (Utilitaire)» à la page 132.

Pour utiliser un dictionnaire thésaurus dans un environnement partitionné, assurez-vous que tous les noeuds physiques peuvent accéder aux fichiers créés.

Conseil

Reportez-vous à l'Annexe M, «Messages renvoyés par les outils du thésaurus», à la page 253.

Partie 2. Références

Chapitre 11. Commandes d'administration pour le propriétaire d'instance

Ce chapitre décrit la syntaxe des commandes d'administration destinées au propriétaire d'instance. L'administration du propriétaire d'instance consiste à vérifier l'état des services de verrouillage et de mise à jour de DB2 Extension Net Search, ainsi que le démarrage et l'arrêt de ces services.

Pour plus d'informations, reportez-vous au Chapitre 5, «Services de l'instance de l'Extension Net Search», à la page 27.

Les commandes sont une variation de la commande DB2TEXT et permettent l'administration des services DB2 Extension Net Search spécifiques à une instance DB2.

Commande	Objectif	Page
CONTROL	Dresse la liste des verrous d'index en texte intégral et les supprime. Dresse également la liste des états de la mémoire cache.	92
START	Démarre les services de l'instance DB2 Extension Net Search.	94
STOP	Arrête les services de l'instance DB2 Extension Net Search.	95

CONTROL

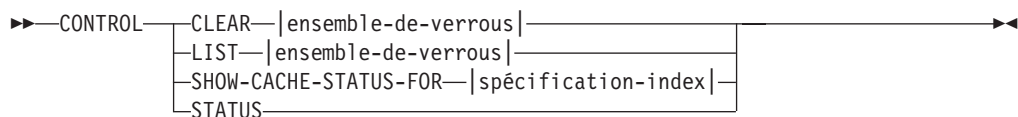
Cette commande permet de dresser la liste des verrous d'index en texte intégral, gérés par les services de l'instance de l'Extension Net Search et même d'en supprimer certains. Si les services de verrouillage et de mise à jour sont en cours d'exécution, vous pouvez afficher leur état, ainsi que des informations sur la mémoire cache activée.

Dans un environnement DB2 réparti, ceci ne s'applique qu'à la partition en cours. L'utilisateur est responsable d'appeler la commande DB2, db2_all pour les partitions souhaitées.

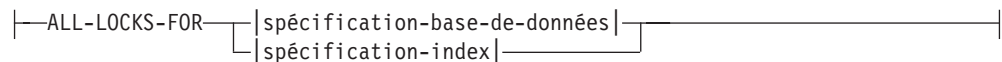
Autorisation

Vous devez exécuter cette commande en tant que propriétaire d'instance DB2 sur le serveur.

Syntaxe de commande



ensemble-de-verrous :



spécification-index :



Spécification-base-de-données :



Paramètres de commande

CLEAR

Utilisez CLEAR pour forcer l'effacement d'un jeu de verrous. Utilisez cette commande avec précaution, après avoir vérifié la cause du problème de verrouillage.

LIST

Utilisez LIST pour obtenir des informations sur les verrous courants, tenus sur un index ou une base de données spécifique. Dans le cas d'un verrou de mise à jour, il est possible d'obtenir des informations sur les documents qui ont été traités.

Vous pouvez également utiliser LIST pour afficher le nombre de documents mis à jour lors du processus de mise à jour. Notez que ceci est uniquement applicable pendant la période où le verrou détient l'index.

Lors de l'utilisation d'une table de capture de réplication, il n'y a pas de mises à jour. Les insertions proviennent soit d'une insertion, soit d'une mise à jour de la table source sur laquelle l'index a été créé.

ensemble-de-verrous

Fonctionne avec des verrous, uniquement dans la base de données ou l'index spécifié.

SHOW CACHE STATUS FOR

Affiche l'état d'activation d'une table mise en cache de l'index spécifié. Il peut s'agir de : "Not Activated" ou "Currently Activated". Si la mémoire cache est activée, elle affiche des détails sur l'utilisation de la mémoire cache. Par exemple, la taille maximale de la mémoire cache (en Mo), le nombre maximal de documents à insérer et l'espace restant dans la table de la mémoire cache (en Ko).

STATUS

En utilisant le mot clé STATUS, la commande indique si le verrouillage et la mise à jour des services de l'instance de l'Extension Net Search sont actifs.

DATABASE nom de base de données

Nom de la base de données sur le serveur utilisé.

INDEX schéma-index.nom-index

Schéma et nom de l'index de texte actuellement utilisé. Il est spécifié dans la commande CREATE INDEX.

Utilisation

Si un message d'erreur de commande d'administration indique qu'il y a un problème de verrouillage, assurez-vous qu'aucune tâche en conflit n'est en cours d'exécution. Par exemple, tenter une commande ALTER pendant qu'une commande UPDATE est en cours d'exécution. Libérez ensuite tous les verrous pour l'index.

Utilisez la commande SHOW CACHE STATUS pour la mise à jour incrémentielle d'un index afin de vérifier que la taille de la mémoire spécifiée est suffisamment large pour contenir toutes les informations de mise à jour lors de la prochaine mise à jour ou afin de vérifier si une activation a été effectuée.

START

Cette commande démarre un démon qui contrôle le verrouillage des index en texte intégral et leur mise à jour automatique sur le serveur DB2.

Remarque

Comme la commande n'active pas de table temporaire mise en cache pour les index, des commandes ACTIVATE CACHE individuelles sont nécessaires pour la recherche avec une procédure mémorisée.

Autorisation

Vous devez exécuter cette commande en tant que propriétaire d'instance DB2 sur un serveur ou l'un des serveurs dans un environnement DB2 réparti.

Syntaxe de commande

►►—START—◄◄

Paramètres de commande

Aucun.

Utilisation

Sous Windows, la commande démarre un service db2ext-<NomInstance>. Vous pouvez également démarrer cette commande à l'aide des méthodes Windows usuelles.

Pour le verrouillage des index en texte intégral, vous pouvez modifier un fichier de configuration afin de répondre aux conditions requises. Pour plus d'informations, reportez-vous à la section «Services de verrouillage» à la page 27.

STOP

Cette commande arrête les services de verrouillage et de mise à jour de l'Extension Net Search.

Autorisation

Vous devez exécuter cette commande en tant que propriétaire d'instance DB2 sur un serveur ou l'un des serveurs dans un environnement DB2 réparti.

Syntaxe de commande

```

>> STOP [FORCE]
  
```

Paramètres de commande

FORCE

Arrête les services, même si les processus détiennent des verrous ou si la table mise en cache est activée pour les index. Si vous ne spécifiez pas FORCE, la commande échoue.

Utilisation

L'arrêt des services de l'instance de l'Extension Net Search ne permet pas d'autre utilisation des commandes spécifiques à l'Extension Net Search. Lors du redémarrage des services, vous devez de nouveau activer la mémoire cache temporaire si vous avez précédemment utilisé une mémoire cache activée avec votre index.

Notez que la cache activée ou l'exécution des commandes de l'Extension Net Search n'arrête pas les services.

Pour le verrouillage des index en texte intégral, vous pouvez modifier un fichier de configuration afin de répondre aux conditions requises. Pour plus d'informations, reportez-vous à la section «Services de verrouillage» à la page 27.

Commande STOP

Chapitre 12. Commandes d'administration pour l'administrateur de base de données

Ce chapitre décrit la syntaxe des commandes d'administration destinées à l'administrateur de base de données. L'administration des bases de données consiste à configurer les bases de données pour être utilisées par DB2 Extension Net Search, puis à désactiver cette préparation.

Le Chapitre 6, «Création et gestion d'un index de texte», à la page 31 décrit comment utiliser ces commandes.

Seules les commandes ENABLE DATABASE et DISABLE DATABASE sont une variation de la commande DB2TEXT, cependant toutes ces commandes permettent une administration au niveau des bases de données.

Commande	Objectif	Page
ENABLE DATABASE	Active la base de données courante pour créer des index en texte intégral.	98
DISABLE DATABASE	Redéfinit le travail de préparation effectué par DB2 Extension Net Search pour une base de données.	100
DB2EXTDL (utilitaire)	Fonction UDF par défaut permettant d'extraire le contenu de la colonne de texte Data Link.	102
DB2EXTHL (utilitaire)	La fonction UDF prend un document de 100 ko et renvoie un objet CLOB de 200 ko.	103

Conseil

Si aucune connexion à une base de données n'a été spécifiée dans la commande db2text, l'exécutable db2text entraîne une connexion implicite à la base de données indiquée dans la variable d'environnement DB2DBDFT.

ENABLE DATABASE

Cette commande permet à une base de données de créer et d'exploiter des index en texte intégral sur des colonnes de texte.

Autorisation

Vous devez exécuter cette commande en tant qu'administrateur de base de données pour activer la base de données. Pour cela, vous devez disposer des droits SYSADM pour pouvoir octroyer les droits DBADM au propriétaire d'instance DB2.

Syntaxe de commande

►►—ENABLE-DATABASE-FOR-TEXT—|options-connexion|—►

options-connexion :

|—CONNECT-TO—nom-de-base-de-données—|
|—USER—id-utilisateur—USING—mot-de-passe—|

Paramètres de commande

CONNECT TO nom-de-base-de-données

Nom de la base de données qui constitue une cible pour cette commande. Ce paramètre peut être omis, si DB2DBDFT est défini et que l'utilisateur exécute la commande sous un ID utilisateur disposant des droits DB2 nécessaires.

USER id-utilisateur USING mot-de-passe

Utilisez un mot de passe et un ID utilisateur pour vous connecter à la base de données.

Utilisation

Cette commande prépare la base de données connectée à une utilisation par DB2 Extension Net Search. Cette étape est obligatoire pour pouvoir créer un index DB2 Extension Net Search sur des tables/colonnes dans la base de données.

Les valeurs par défaut de la base de données, établies après l'exécution de la commande, peuvent être affichées à l'aide de la vue de catalogue DB2EXT.DBDEFAULTS.

Modifications de la base de données

Cette commande accorde les droits DBADM au propriétaire d'instance DB2 associé à l'instance DB2 de la base de données activée.

La commande ENABLE DATABASE crée divers objets de base de données dans le schéma DB2EXT, tels que des catalogues DB2 Extension Net Search, des fonctions UDF et des procédures mémorisées. Une fois la commande exécutée, les vues de catalogue suivantes sont disponibles :

db2ext.dbdefaults
db2ext.textindexes
db2ext.textindexformats
db2ext.indexconfiguration
db2ext.proxyinformation

Commande ENABLE DATABASE

Notez que les vues DB2 Text Information Extender sont également disponibles pour des raisons de compatibilité avec les versions antérieures. Pour plus d'informations, reportez-vous à l'Annexe C, «Catalogues d'informations DB2 Extension Net Search», à la page 191.

Notez également que les tables énoncées ci-dessus se trouvent dans l'espace table par défaut de la base de données, connu comme IBMDEFAULTGROUP. Ceci est réparti sur tous les noeuds définis dans `db2nodes.cfg`.

Modifications du système de fichiers

Aucune.

DISABLE DATABASE

Cette commande annule les modifications DB2 Extension Net Search apportées à une base de données.

Autorisation

Vous devez exécuter cette commande en tant qu'administrateur de base de données pour désactiver la base de données. Pour cela, vous devez disposer des droits DBADM.

Syntaxe de commande

```
➤➤—DISABLE-DATABASE-FOR-TEXT—[FORCE]—[options-connexion]—➤➤
```

options-connexion :

```
|—[CONNECT-TO—nom-de-base-de-données—[USER—id-utilisateur—USING—mot-de-passe]—]|
```

Paramètres de commande

CONNECT TO **nom-de-base-de-données**

Nom de la base de données qui constitue une cible pour cette commande. Ce paramètre peut être omis, si DB2DBDFT est défini et que l'utilisateur exécute la commande sous un ID utilisateur disposant des droits DB2 nécessaires.

USER **id-utilisateur** USING **mot-de-passe**

Utilisez un mot de passe et un ID utilisateur pour vous connecter à la base de données.

FORCE

Force la suppression de tous les index DB2 Extension Net Search contenus dans la base de données. Pour plus de détails, reportez-vous à la section «DROP INDEX» à la page 130.

Utilisation

Cette commande redéfinit la base de données connectée, de sorte qu'elle ne puisse plus être utilisée par d'autres commandes DB2 Extension Net Search. Si des index en texte intégral existent dans la base de données, cette commande échoue à moins que l'option FORCE ne soit utilisée.

Cette commande ne retire pas les droits DBADM au propriétaire d'instance DB2.

Remarque

La désactivation d'une base de données échouera si des index de texte sont définis dans la base de données. Il est recommandé de supprimer ces index l'un après l'autre, puis de s'assurer qu'aucun incident n'apparaît. Si vous utilisez la commande DISABLE DATABASE for text force, celle-ci garantit uniquement le fait que les tables de catalogue Extension Net Search sont supprimées de la base de données.

Cependant, si certains index n'ont pas pu être complètement supprimés, il restera toujours des ressources qui nécessiteront un nettoyage manuel. Ceci inclut :

- Les fichiers dans l'index, le répertoire de travail et le répertoire cache
- Les entrées de planificateur dans ctedem.dat
- Les endroits où un index a été créé à l'aide de l'option de capture de réplication : les entrées IBMSNAP_SIGNAL, IBMSNAP_PRUNE_SET et IBMSNAP_PRUNCNTL dans les tables de la base de données éloignée doivent être supprimées manuellement. Ces entrées peuvent être facilement identifiées avec la commande APPLY_QUAL="NSE" || <nom instance> and TARGET_SERVER= <nom bdd>.

Dans l'exemple suivant, l'instance est DB2 et la base de données est SAMPLE.

```
DELETE FROM <ccSchema>.IBMSNAP_SIGNAL
WHERE SIGNAL_INPUT_IN IN
      (SELECT MAP_ID FROM <ccSchema>.IBMSNAP_PRUNCNTL
       WHERE APPLY_QUAL= 'NSEDDB2' AND TARGET_SERVER= 'SAMPLE');

DELETE FROM <ccSchema>.IBMSNAP_PRUNCNTL
WHERE APPLY_QUAL= 'NSEDDB2' AND TARGET_SERVER= 'SAMPLE';

DELETE FROM <ccSchema>.IBMSNAP_PRUNE_SET
WHERE APPLY_QUAL= 'NSEDDB2' AND TARGET_SERVER= 'SAMPLE';
```

Modifications de la base de données

Les modifications suivantes, effectuées dans la base de données pour activer DB2 Extension Net Search sont supprimées :

- Les vues de catalogues DB2 Extension Net Search de la base de données.
- Tous les objets de la base de données créés par DB2 Extension Net Search.

Modifications du système de fichiers et de la mémoire partagée

Si vous utilisez l'option FORCE, les fichiers d'index sont supprimés.

Si vous utilisez l'option FORCE, la mémoire cache est supprimée pour toute cache d'index activée. Pour plus d'informations, reportez-vous à la section «DROP INDEX» à la page 130.

DB2EXTDL (utilitaire)

Par défaut, la fonction UDF qui extrait le contenu d'une colonne de texte Data Link renvoie une valeur BLOB de 100 Ko. Selon la taille du plus grand document contenu dans la base de données référencée par la fonction Data Link, vous pouvez augmenter ou diminuer cette valeur.

Autorisation

Vous devez exécuter cette commande en tant qu'administrateur de base de données pour activer la base de données. Pour cela, vous devez disposer des droits SYSADM pour pouvoir octroyer les droits DBADM au propriétaire d'instance DB2.

Syntaxe de commande

►►—db2extdl—*nouvelle-taille-résultats*—————►◄

Paramètres de commande

nouvelle-taille-résultats

Nouvelle taille des résultats de la fonction UDF pour extraire le contenu Data Link en Ko. Il s'agit d'un entier positif inférieur à 2097152.

DB2EXTHL (utilitaire)

Par défaut, la fonction UDF de mise en évidence prend comme entrée un document de taille maximale égale à 100 octets et renvoie un objet CLOB de 200 octets. Selon la taille du document le plus grand dans la base de données, vous pouvez augmenter la taille de la valeur d'entrée à une taille maximale de 1 Go.

Autorisation

Vous devez exécuter cette commande en tant qu'administrateur de base de données pour activer la base de données. Pour cela, vous devez disposer des droits SYSADM pour pouvoir octroyer les droits DBADM au propriétaire d'instance DB2.

Syntaxe de commande

►►—db2exthl—*nouvelle-taille-de-l'entrée-mise-en-évidence* —————►◄

Paramètres de commande

nouvelle-taille-de-l'entrée-mise-en-évidence

La nouvelle taille du résultat de la fonction UDF de mise en évidence en Ko. Il s'agit d'un entier positif de <1048576.

Commande DB2EXTHL (utilitaire)

Chapitre 13. Commandes d'administration pour le propriétaire de table de texte

Ce chapitre décrit la syntaxe des commandes d'administration destinées au propriétaire de table de texte.

Le Chapitre 6, «Création et gestion d'un index de texte», à la page 31 décrit le mode d'utilisation de ces commandes.

Les commandes sont une variation de la commande DB2TEXT. Elles permettent au propriétaire d'une table de créer et de manipuler des index en texte intégral sur des colonnes de la table.

Commande	Objectif	Page
ACTIVATE CACHE	Active la mémoire cache de telle sorte que les opérations de recherche utilisant la procédure mémorisée soient possibles.	106
ALTER INDEX	Modifie les caractéristiques d'un index.	108
CLEAR EVENTS	Supprime les événements d'index d'une table d'événements d'index utilisée pendant la mise à jour de l'index.	112
CREATE INDEX	Crée un index en texte intégral.	114
DEACTIVATE CACHE	Désactive la mémoire cache de telle sorte que les opérations de recherche à l'aide de la procédure mémorisée ne sont plus possibles.	128
DB2EXTTH (Utilitaire)	Compile le fichier de définitions du thésaurus.	132
DROP INDEX	Supprime un index en texte intégral pour une colonne de texte.	130
ENABLE DATABASE	Active la base de données courante afin de créer des index en texte intégral.	98
UPDATE INDEX	Démarre le processus d'indexation en fonction du contenu courant des colonnes de texte.	134
HELP	Affiche la liste des options de commande DB2TEXT.	138
COPYRIGHT	Affiche les informations produit et de copyright de l'Extension Net Search.	139

Conseil

Si aucune connexion à une base de données n'a été spécifiée dans la commande db2text, l'exécutable db2text entraîne une connexion implicite à la base de données indiquée dans la variable d'environnement DB2DBDFT.

ACTIVATE CACHE

Cette commande active la table mise en cache à partir de la table utilisateur DB2 ou de la mémoire cache persistante. Une fois fait, les opérations de recherche ayant recours à la procédure mémorisée sont possibles. Pour plus d'informations, reportez-vous au Chapitre 16, «Fonction de recherche de procédure mémorisée», à la page 161.

Cette commande est disponible uniquement si l'index a été créé avec une option **CACHE TABLE**. Pour plus d'informations, reportez-vous à la section «**CREATE INDEX**» à la page 114.

Autorisation

Selon les vues de catalogue DB2, l'ID utilisateur dans cette commande doit disposer des droits **CONTROL** sur la table pour laquelle l'index en texte intégral a été créé.

Syntaxe de commande

```

>> ACTIVATE CACHE FOR INDEX [schéma-index"." ] nom-index FOR-TEXT
|
| RECREATE | | options-connexion |
|

```

options-connexion :

```

|
| CONNECT-TO nom-de-base-de-données |
| USER id-utilisateur USING mot-de-passe |
|

```

Paramètres de commande

schéma-index

Schéma de l'index de texte, tel qu'il est spécifié dans la commande **CREATE INDEX**. Si aucun schéma n'est spécifié, l'ID utilisateur de la connexion DB2 est utilisé.

nom-index

Nom de l'index de texte, tel qu'il est spécifié dans la commande **CREATE INDEX**.

RECREATE

S'applique uniquement aux index utilisant une mémoire cache persistante. La mémoire cache existante est supprimée. Si une mise à jour sans activation a été réalisée, la mémoire cache persistante est automatiquement reconstruite à partir de la base de données.

CONNECT TO nom-de-base-de-données

Nom de la base de données qui constitue la cible de cette commande. Ce paramètre peut être omis, si DB2DBDFT est défini et que l'utilisateur exécute la commande sur le serveur. Notez que l'ID utilisateur doit disposer des droits DB2 requis.

USER id-utilisateur USING mot-de-passe

Utilisez un mot de passe et un ID utilisateur pour vous connecter à la

base de données. Sans spécification de votre part, la connexion est tentée sous l'ID utilisateur courant et sans mot de passe.

Utilisation

Il n'est pas possible d'émettre la commande si l'une des commandes suivantes est en cours d'exécution sur l'index :

- UPDATE INDEX
- ALTER INDEX
- DROP INDEX
- CLEAR EVENTS
- DEACTIVATE CACHE

Remarque

L'activation d'une table mise en cache peut nécessiter qu'elle soit totalement recréée, même si une mémoire cache persistante a été utilisée. Ceci se produit si une opération de mise à jour a eu lieu alors que la mémoire cache persistante était désactivée.

La quantité de mémoire utilisée pour générer la mémoire cache est calculée dynamiquement à partir du nombre de documents courants et de la taille des colonnes résultantes. Utilisez la valeur PCTFREE pour augmenter la quantité de mémoire minimale calculée d'un facteur de $100/(100-PCTFREE)$. La valeur PCTFREE est spécifiée dans la commande CREATE ou ALTER INDEX.

De ce fait, PCTFREE décrit le pourcentage de mémoire cache allouée, réservée pour les opérations d'insertion alors que la mémoire cache est activée. Notez que pour chaque commande ACTIVATE CACHE, la taille de la mémoire réelle est ré-évaluée.

Modifications du système de fichiers

Des fichiers pour l'implémentation de la mémoire cache persistante sont créés.

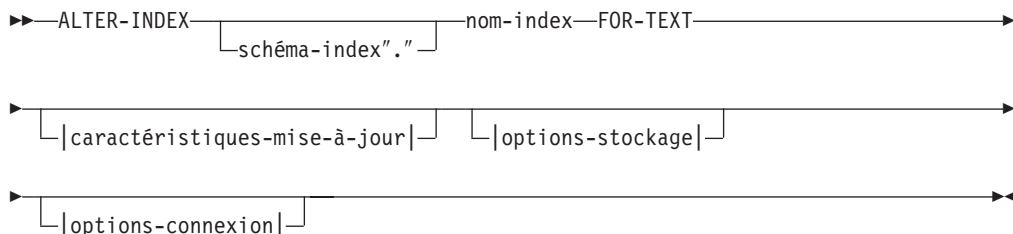
ALTER INDEX

Cette commande les caractéristiques d'un index en texte intégral, par exemple, les options de mise à jour et celles de stockage.

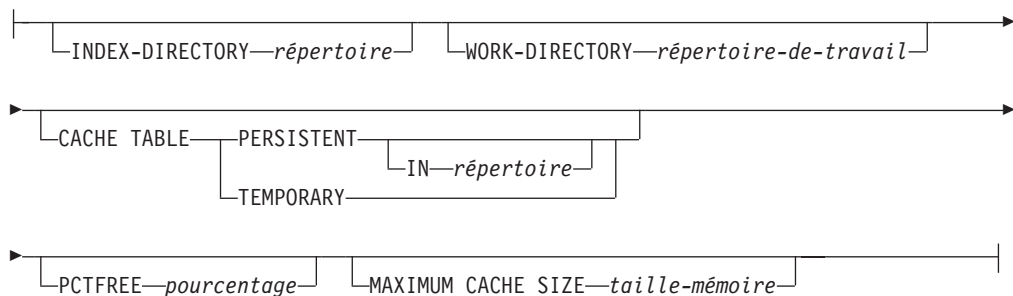
Autorisation

Selon les vues de catalogue DB2, l'ID utilisateur de cette commande doit disposer des droits CONTROL sur la table pour laquelle l'index en texte intégral a été créé.

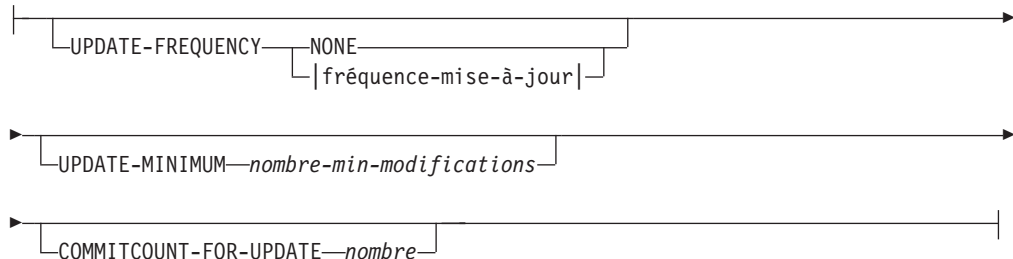
Syntaxe de commande



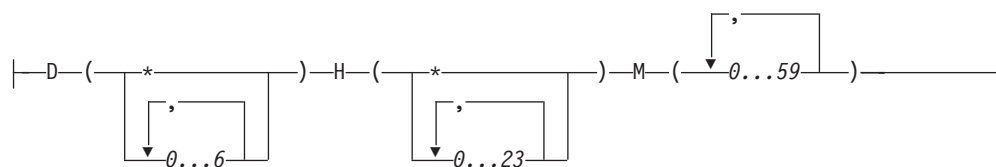
options-stockage :



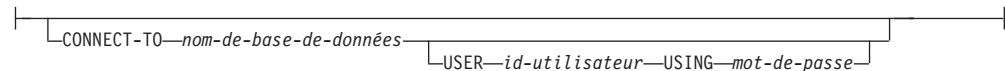
caractéristiques-mise-à-jour :



fréquence-mise-à-jour :



options-connexion :



Paramètres de commande

schéma-index

Schéma de l'index de texte, tel qu'il est spécifié dans la commande CREATE INDEX. Si aucun schéma n'est spécifié, l'ID utilisateur de la connexion DB2 est utilisé.

nom-index

Nom de l'index de texte, tel qu'il est spécifié dans la commande CREATE INDEX.

INDEX DIRECTORY répertoire

Chemin d'accès au répertoire dans lequel l'index de texte est stocké. Comme le répertoire est destiné à contenir les données de l'index, assurez-vous que le répertoire est accessible en lecture/écriture et exécutez les droits pour l'ID utilisateur du propriétaire d'instance DB2.

Notez que dans un environnement DB2 réparti, ce répertoire doit exister sur chaque noeud. Un sous-répertoire, NODE<nr>, est créé sous le répertoire pour distinguer les index sur des noeuds logiques d'un serveur. Tous les fichiers d'index du répertoire d'index précédent sont supprimés.

WORK DIRECTORY répertoire-de-travail

Stocke les fichiers temporaires pendant les opérations de recherche et d'administration. Vous pouvez modifier le répertoire de travail distinct indépendamment d'un nouveau répertoire d'index.

Si le répertoire n'existe pas, il est créé pour l'ID utilisateur du propriétaire d'instance DB2. S'il existe, assurez-vous que le propriétaire d'instance est doté des droits en lecture/écriture sur le répertoire sur les plateformes UNIX.

Notez que dans un environnement DB2 réparti, ce répertoire doit exister sur chaque noeud. Un sous-répertoire, NODE<nr>, est créé sous le répertoire pour distinguer les index sur des noeuds logiques d'un serveur. Tous les fichiers d'index temporaires du répertoire d'index précédent sont supprimés.

CACHE TABLE PERSISTENT IN répertoire

Indique qu'après une désactivation ou une relance du système, la table mise en cache dans CREATE INDEX est persistante. Quel que soit le cas, ceci permet une exécution rapide de ACTIVATE CACHE. La mémoire cache persistante est stockée dans le répertoire spécifié.

La mémoire cache persistante précédemment créée est déplacée vers un autre emplacement. Cet emplacement requiert toujours un index désactivé.

CACHE TABLE TEMPORARY

Indique que la table de résultats mise en cache est à présent temporaire et que toute cache persistante précédemment existante a été supprimée. Notez que cette modification requiert un index désactivé.

MAXIMUM CACHE SIZE taille-mémoire

Indique la nouvelle taille maximale de la table mise en cache devant être

Commande ALTER INDEX

générée pendant la commande ACTIVATE CACHE. Un entier positif spécifie le paramètre de la taille de la mémoire en mégaoctets (Mo).

Si le nombre entier est trop petit, la commande ACTIVATE CACHE échoue. La taille de la mémoire cache réelle est calculée pendant la commande ACTIVATE CACHE. Cette modification requiert un index désactivé.

PCTFREE pourcentage

Spécifie le pourcentage de mémoire cache, gardé libre pour les documents supplémentaires. Ce pourcentage doit être un nombre entier inférieur à 100 et supérieur ou égal à 0. Notez que la mémoire cache persistante précédente est supprimée et que cette modification requiert un index désactivé. Reportez-vous à la section «ACTIVATE CACHE» à la page 106.

UPDATE FREQUENCY

À l'aide des paramètres suivants, la fréquence de mise à jour de l'index détermine quand la mise à jour a lieu :

- **D.** Jour(s) de la semaine où l'index est mis à jour : * (tous les jours) ou 0..6 (0=Dimanche)
- **H.** Heure(s) de mise à jour de l'index : * (toutes les heures) ou 0..23
- **M.** Minute(s) de l'heure de mise à jour de l'index : 0..59
- **NONE.** Aucune autre mise à jour de l'index n'a lieu. Cette valeur est destinée à une colonne de texte dans laquelle aucun autre changement n'est apporté.

Si vous ne spécifiez pas le mot clé UPDATE FREQUENCY, les paramètres de fréquence restent inchangés.

UPDATE MINIMUM nombre-min-modifications

Nombre minimal de modifications autorisées sur les documents texte avant que l'index ne soit mis à jour de manière incrémentielle. Si vous ne spécifiez pas le mot clé UPDATE MINIMUM, le paramètre reste inchangé.

Notez qu'il n'est possible de modifier la valeur de UPDATE MINIMUM que si l'index n'a pas été créé à l'aide de l'option RECREATE ON UPDATE.

COMMITCOUNT FOR UPDATE nombre

Pour le traitement de la mise à jour, vous pouvez spécifier un nombre de validations. Pour plus d'informations, reportez-vous à la section «UPDATE INDEX» à la page 134. Ceci s'applique à la commande UPDATE et à la spécification UPDATE FREQUENCY qui planifie le traitement de la mise à jour.

Notez qu'il n'est possible de modifier la valeur de COMMITCOUNT que si l'index n'a pas été créé à l'aide de l'option RECREATE ON UPDATE.

Notez également que vous ne pouvez pas modifier la valeur de COMMITCOUNT, si vous avez créé l'index avec la clause REPLICATION.

CONNECT TO nom-de-base-de-données

Nom de la base de données qui constitue la cible de cette commande. Ce paramètre peut être omis, si DB2DBDFT est défini et que l'utilisateur exécute la commande sur le serveur. Notez que l'ID utilisateur doit disposer des droits DB2 requis.

USER id-utilisateur USING mot-de-passe

Utilisez un mot de passe et un ID utilisateur pour vous connecter à la base de données. Sans spécification de votre part, la connexion est tentée sous l'ID utilisateur courant et sans mot de passe.

Utilisation

Il n'est pas possible d'émettre la commande si l'une des commandes suivantes est en cours d'exécution sur l'index :

- ALTER INDEX
- CLEAR EVENTS
- ACTIVATE CACHE
- DROP INDEX
- UPDATE INDEX
- DEACTIVATE CACHE

Si vous créez l'index avec une option de mémoire cache, vous ne pouvez pas utiliser la commande ALTER INDEX pour le répertoire d'index lorsque l'index est activé. Vous devez d'abord désactiver la mémoire cache.

Dans un environnement DB2 réparti, un index de texte avec des options de la mémoire cache est uniquement autorisé sur un espace table à un seul noeud.

Modifications de la base de données

Modifiez les vues de catalogue DB2 Extension Net Search.

Modifications du système de fichiers

- Création des sous-répertoires NODE<nr> dans l'index et les répertoires de travail
- Déplacement de fichiers d'index
- Création de répertoires cache persistants
- Déplacement de fichiers cache persistants

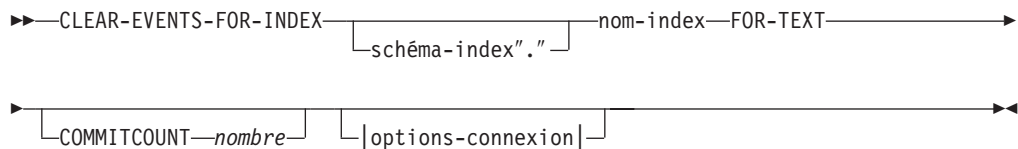
CLEAR EVENTS

Cette commande supprime les événements d'indexation d'une vue d'événements de l'index. Utilisez la vue d'événements à des fins d'administration. Le nom de la vue d'événements figure dans la colonne EVENTVIEWNAME de la vue DB2EXT.TEXTINDEXES.

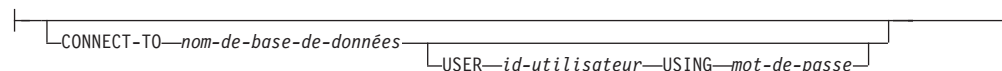
Autorisation

Selon les vues de catalogue DB2, l'ID utilisateur de cette commande doit disposer des droits CONTROL sur la table pour laquelle l'index en texte intégral a été créé.

Syntaxe de commande



options-connexion :



Paramètres de commande

schéma-index

Schéma de l'index de texte, tel qu'il est spécifié dans la commande CREATE INDEX. Si aucun schéma n'est spécifié, l'ID utilisateur de la connexion DB2 est utilisé.

nom-index

Nom de l'index de texte, tel qu'il est spécifié dans la commande CREATE INDEX.

COMMITCOUNT nombre

Un entier supérieur ou égal à 0 affiche le nombre de lignes supprimées dans une transaction par DB2.

CONNECT TO nom-de-base-de-données

Nom de la base de données qui constitue la cible de cette commande. Ce paramètre peut être omis, si DB2DBDFT est défini et que l'utilisateur exécute la commande sur le serveur. Notez que l'ID utilisateur doit disposer des droits DB2 requis.

USER id-utilisateur USING mot-de-passe

Utilisez un mot de passe et un ID utilisateur pour vous connecter à la base de données. Sans spécification de votre part, la connexion est tentée sous l'ID utilisateur courant et sans mot de passe.

Utilisation

Si vous planifiez des mises à jour régulières à l'aide de l'option UPDATE FREQUENCY dans les commandes CREATE ou ALTER INDEX, vérifiez régulièrement

la table d'événements. Utilisez la commande `CLEAR EVENTS` pour vider les tables d'événements, une fois que vous avez vérifié la cause de l'événement et supprimé la source de l'erreur.

Essayez de maintenir une cohérence entre le contenu des colonnes de texte de la table et l'index, en particulier lors de la ré-indexation des documents.

Il n'est pas possible d'émettre la commande si l'une des commandes suivantes est en cours d'exécution sur l'index :

- `UPDATE INDEX`
- `ALTER INDEX`
- `ACTIVATE CACHE`
- `DEACTIVATE CACHE`
- `DROP INDEX`

CREATE INDEX

Cette commande crée un index en texte intégral sur une colonne de texte en vue d'une utilisation dans des requêtes en texte intégral de DB2 Extension Net Search.

Dans un environnement DB2 réparti, un index en texte intégral est créé sur chaque partition de l'espace table sur lequel la table utilisateur est définie. Les modifications ultérieures de la répartition de l'espace table ne sont pas autorisées et mènent à un comportement inattendu dans les commandes d'administration et lors du processus de la recherche.

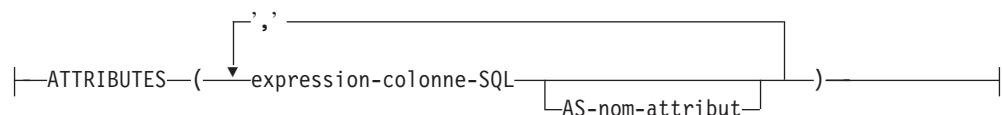
Autorisation

Selon les vues de catalogue DB2, l'ID utilisateur de cette commande doit disposer des droits CONTROL sur la table sur laquelle l'index en texte intégral a été créé.

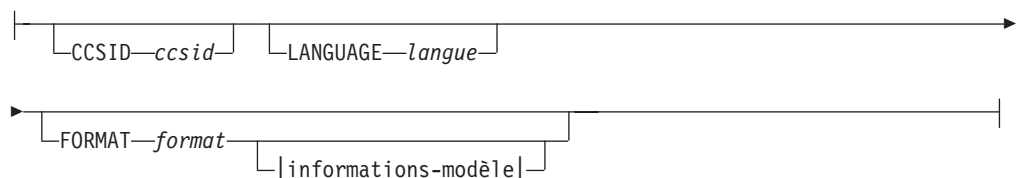
Syntaxe de commande



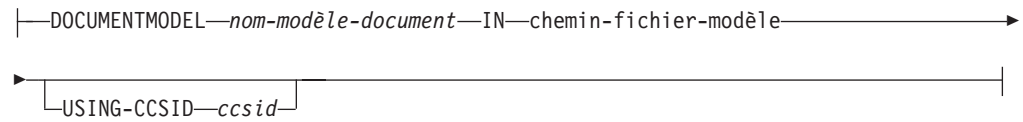
liste-attributs :



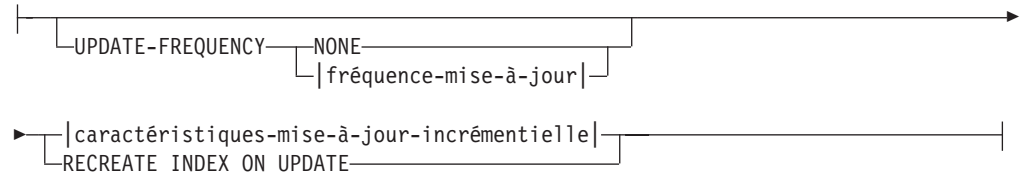
informations-par-défaut-texte :



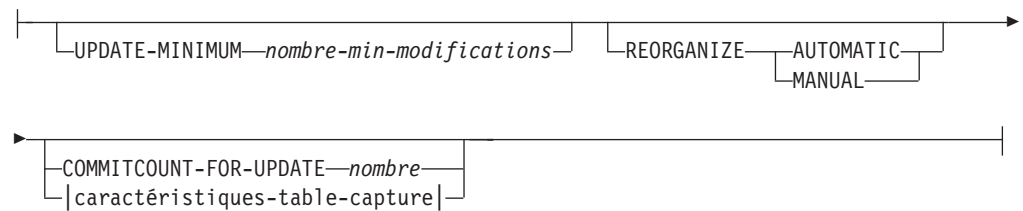
informations-modèle :



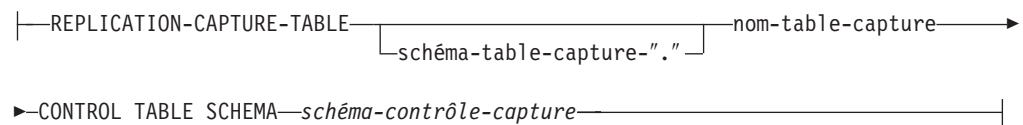
caractéristiques-mise-à-jour :



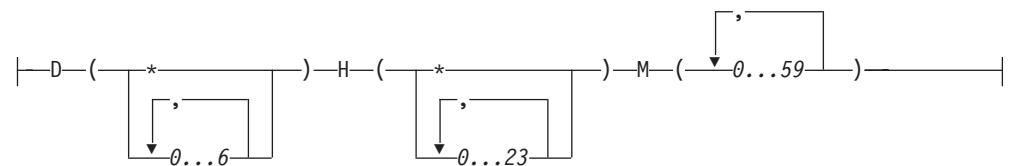
caractéristiques-mise-à-jour-incrémentielle :



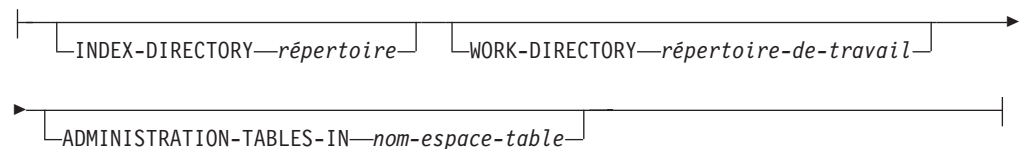
caractéristiques-table-capture :



fréquence-mise-à-jour :

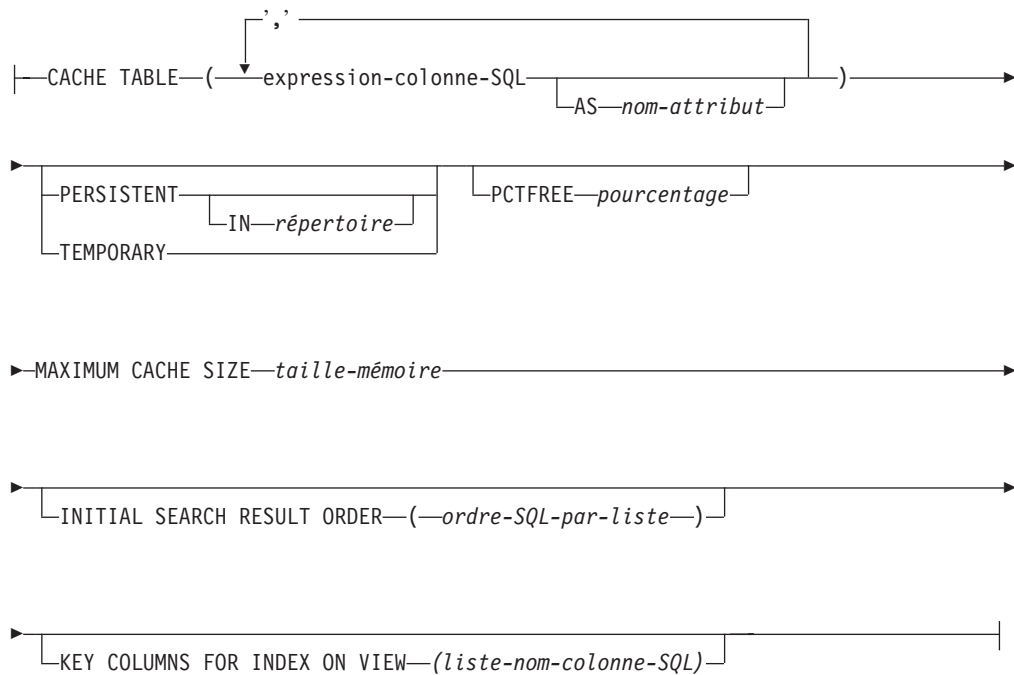


options-stockage :



Commande CREATE INDEX

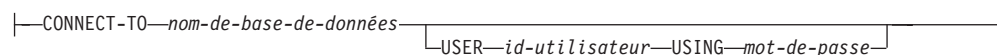
options-résultats-recherche-cache :



options-configuration-index :



options-connexion :



Paramètres de commande

schéma-index

Schéma de l'index de texte. A utiliser comme nom de schéma DB2 pour les tables d'administration spécifiques à l'index. Si aucun schéma n'est spécifié, l'ID utilisateur de la connexion DB2 est utilisé. Notez que le schéma de l'index doit être un nom de schéma DB2 valide.

nom-index

Nom de l'index. Avec le schéma d'index, il identifie de manière unique un index en texte intégral dans une base de données. Il sert également de nom pour la table d'événements d'index.

Pour plus de détails, reportez-vous à l'Annexe C, «Catalogues d'informations DB2 Extension Net Search», à la page 191. Notez que le nom de l'index doit être un nom d'index DB2 valide.

schéma-table

Schéma de table, d'alias de type nickname ou de vue, pour lequel l'index est créé. Si aucun schéma n'est spécifié, l'ID utilisateur de la connexion DB2 est utilisé.

nom-table

Nom de la table de texte, de l'alias de type nickname ou de la vue dans la base de données connectée qui contient la colonne pour laquelle l'index en texte intégral est créé.

Notez que si le nom de table ne fait pas référence à une table de base DB2, les restrictions suivantes sont appliquées :

- Une vue ne permet qu'une recherche de procédure mémorisée ou de fonction de valeur table. De ce fait, vous devez spécifier les colonnes clés pour l'index ou les vues à l'aide de la clause KEY COLUMNS FOR INDEX ON VIEW.
- Pour les mises à jour d'index incrémentielles sur des alias de type nickname sans table de capture, une table de journalisation est créée. Elle doit être remplie manuellement, si des modifications sont apportées aux données de la vue ou de la table d'alias de type nickname. Avec les tables de base, cette opération est effectuée automatiquement. L'utilisateur **ne doit donc pas** toucher à la table de journalisation. Pour la présentation de la table de journalisation, reportez-vous à l'Annexe C, «Catalogues d'informations DB2 Extension Net Search», à la page 191.
- Les prédicats DB2 CONTAINS, SCORE et NUMBEROFMATCHES ne sont autorisés que pour les index sur des tables de base ou des alias de type nickname, mais pas sur des vues.
- Les index sur des vues ne sont autorisés que si vous spécifiez les options des résultats de la recherche dans la mémoire cache dans la commande.

nom-colonne-texte

Nom de la colonne contenant le texte utilisé pour la création d'un index en texte intégral. Le type de la colonne doit être l'un des suivants :

- CHAR (FOR BIT DATA)
- VARCHAR (FOR BIT DATA)
- LONG VARCHAR (FOR BIT DATA)
- CLOB
- DBCLOB
- BLOB
- GRAPHIC
- VARGRAPHIC
- LONG VARGRAPHIC
- DATALINK

Si le type de la colonne n'est pas l'un de ceux-là, spécifiez une fonction de transformation à l'aide de **schéma-fonction.nom-fonction** pour convertir le type de colonne.

Notez que, si vous utilisez une colonne Data Link, le contenu référencé **doit** être extrait pour l'indexation. Cette opération a lieu via le protocole qui fait partie de la valeur Data Link, par exemple Http. Si vous utilisez des protocoles différents de "file" ou "unc", assurez-vous qu'ils sont pris en charge par des serveurs faisant partie des valeurs Data Link. Comme des serveurs proxy peuvent être nécessaires pour extraire le contenu du fichier,

Commande CREATE INDEX

l'administrateur de base de données peut les spécifier dans la table DB2EXT.PROXYINFORMATION avant la création de l'index.

Notez que plusieurs index sont autorisés sur les mêmes colonnes, mais **seulement** dans les conditions suivantes :

L'index est créé sur une vue

De ce fait, vous ne pouvez pas utiliser l'index dans les arguments de recherche CONTAINS, SCORE ou NUMBEROFMATCHES.

L'index est créé sur une table

Si tous les index sont synchronisés, ils ont des propriétés identiques sur la même colonne dans les détails de la commande CREATE INDEX suivante :

- Schéma et nom de fonction
- ATTRIBUTES
- CCSID
- LANGUAGE
- FORMAT
- DOCUMENTMODEL
- INDEX CONFIGURATION

Ainsi, peu importe l'index choisi par les arguments CONTAINS, SCORE ou NUMBEROFMATCHES.

schéma-fonction.nom-fonction

Schéma et nom d'une fonction UDF utilisée pour accéder aux documents texte qui se trouvent dans une colonne d'un type pris en charge. La fonction effectue une conversion du type de colonne, à l'aide du paramètre d'entrée d'un type de colonne arbitraire. Elle renvoie la valeur de l'un des types pris en charge par l'Extension Net Search.

ATTRIBUTES (expression-colonne-SQL AS nom-attribut, ...)

Garantit que le contenu d'une expression de colonne est indexé en plus de la colonne de texte. Ce contenu peut être également recherché par la clause ATTRIBUTE dans une instruction de recherche. Les expressions de la colonne SQL doivent être définies à l'aide des noms de colonne non qualifiés de la table sur laquelle l'index est créé. Le seul type de données autorisé est DOUBLE. Les opérateurs de transtypage peuvent être utilisés dans les expressions de colonnes, mais le transtypage implicite de DB2 est **impossible**. Les noms d'attribut doivent respecter les règles des noms d'attribut de modèles de document et doivent être différents des noms d'attribut dans le fichier de définition de modèle d'index.

Déterminez les noms d'attribut pour les expressions en respectant les règles suivantes :

- S'il est explicitement désigné par la clause SQL AS dans l'expression de colonne, utilisez le nom spécifié. Par exemple : ATTRIBUTES (C1+C2 AS monNom)
- Si une colonne de la table spécifiée est utilisée sans la clause AS, le nom de la colonne est utilisé. Par exemple : CACHE TABLE (C1)
- Si une expression est utilisée sans la clause AS et ne fait pas référence à une colonne nommée, la commande CREATE INDEX génère une erreur.

Par exemple : ATTRIBUTES (CAST(JULIAN_DAY(date) AS DOUBLE) as day, (price1+price2)/2 as avg_price)

Notez que les attributs sans guillemet sont mappés en majuscules et doivent être spécifiés de cette façon pendant la recherche.

CCSID ccsid

L'ID de jeu de caractères codés (CCSID) est utilisé lors de l'indexation des documents texte. La valeur par défaut est tirée de la vue DB2EXT.DBDEFAULTS dans laquelle DEFAULTNAME='CCSID'.

LANGUAGE langue

Pour la liste, reportez-vous à l'Annexe E, «Langues prises en charge», à la page 207. La valeur par défaut est tirée de la vue DB2EXT.DBDEFAULTS dans laquelle DEFAULTNAME='LANGUAGE'.

FORMAT format

Format des documents texte dans la colonne, par exemple, HTML. Ces informations sont nécessaires pour l'indexation des documents. Pour obtenir la liste des formats de document pris en charge pour les documents structurés, reportez-vous à la section «Formats de document et pages de codes prises en charge» à la page 23.

Pour les formats de documents structurés, vous pouvez spécifier des informations dans un fichier de modèle de document. Si aucun modèle de document n'est spécifié, le texte du document est indexé à l'aide du modèle de document par défaut. Reportez-vous à la section «Modèles de document» à la page 166.

Si le mot clé du format n'est pas spécifié, la valeur par défaut est tirée de la vue DB2EXT.DBDEFAULTS dans laquelle DEFAULTNAME='FORMAT'.

DOCUMENTMODEL nom-modèle-document IN chemin-fichier-modèle

Le chemin-fichier-modèle spécifie l'emplacement d'un fichier modèle. Il contient une définition de modèle pour le format dans la clause FORMAT. Il doit être lisible par le propriétaire d'instance DB2. Un modèle de document permet d'indexer des sections spécifiques d'un document et d'y effectuer des recherches. Il est possible de définir des balises de marquage et des noms de section dans un modèle de document. Ce dernier est lié à un format de document, prenant en charge les structures HTML, XML ou GPP. Il n'est possible de spécifier qu'un seul modèle de document dans un fichier modèle.

Comme les modèles de documents n'ont pas besoin d'être référencés dans les conditions de recherche, utilisez à la place tous les noms de section dans le fichier modèle. Pour plus de détails sur les modèles de documents, reportez-vous au Chapitre 9, «Utilisation de documents structurés», à la page 83. Notez que comme le modèle de document est en lecture seule pendant la commande CREATE INDEX, toutes les modifications ultérieures ne sont pas reconnues pour cet index.

Notez que dans un environnement DB2 réparti, le chemin-fichier-modèle doit être accessible sur chaque noeud à l'aide d'un système de fichiers partagés.

USING CCSID ccsid

Spécifiez un CCSID pour interpréter le contenu du fichier modèle. La valeur par défaut est tirée de la vue DB2EXT.DBDEFAULTS dans laquelle DEFAULTNAME='MODELCCSID'.

UPDATE FREQUENCY

La fréquence de mise à jour de l'index détermine quand la mise à jour a lieu. Si le nombre de modifications de la table utilisateur est inférieur à la valeur spécifiée par l'option UPDATE MINIMUM, l'index n'est pas mis à

Commande CREATE INDEX

jour. Si la valeur de UPDATE FREQUENCY n'est pas spécifiée, la valeur par défaut NONE est utilisée, de sorte qu'aucune autre mise à jour de l'index n'est effectuée. Ceci est utile lorsqu'aucune autre modification ne doit avoir lieu sur une colonne de texte.

- **D.** Jour(s) de la semaine où l'index est mis à jour : * (tous les jours) ou 0..6 (0=Dimanche)
- **H.** Heure(s) de mise à jour de l'index : * (toutes les heures) ou 0..23
- **M.** Minute(s) de l'heure de mise à jour de l'index : 0..59
- **NONE.** Aucune autre mise à jour de l'index n'est effectuée. La mise à jour doit être démarrée manuellement.

La valeur par défaut est tirée de la vue DB2EXT.DBDEFAULTS dans laquelle DEFAULTNAME='UPDATEFREQUENCY'.

UPDATE MINIMUM nombre-min-modifications

Nombre minimal de modifications autorisées sur les documents texte avant que l'index ne soit mis à jour automatiquement par la commande UPDATE FREQUENCY. Les valeurs d'entiers positifs sont autorisées. La valeur par défaut est tirée de la vue DB2EXT.DBDEFAULTS dans laquelle DEFAULTNAME='UPDATEMINIMUM'.

Notez que cette valeur est ignorée dans une commande DB2TEXT UPDATE. Cette option ne peut pas être utilisée avec l'option RECREATE INDEX ON UPDATE, comme le nombre de modifications n'est pas disponible sans table de journalisation, ni déclencheurs de mise à jour incrémentielle.

Pour des bases de données réparties, l'option UPDATE MINIMUM est vérifiée sur tous les noeuds.

REORGANIZE AUTOMATIC/MANUAL

Les mises à jour effectuées à l'aide de UPDATE FREQUENCY ne reconnaissent l'index que si l'option REORGANIZE AUTOMATIC est spécifiée. Cette étape s'effectue automatiquement en fonction de la valeur de select REORGSGGESTED from DB2EXT.TEXTINDEXES après la mise à jour.

REORGANIZE MANUAL ne peut être effectué qu'avec une commande UPDATE manuelle, à l'aide de l'option REORGANIZE.

Si la clause REORGANIZE est omise, la valeur par défaut est tirée de la vue DB2EXT.DBDEFAULTS dans laquelle DEFAULTNAME='AUTOMATICREORG'.

Pour plus d'informations sur l'option REORGANIZE, reportez-vous à la section «UPDATE INDEX» à la page 134.

REPLICATION CAPTURE TABLE schéma-table-capture. CONTROL TABLE SCHEMA schéma-contrôle-capture

Pour le traitement de mise à jour incrémentielle, la table de capture de réplication spécifiée est utilisée à la place d'une table de journalisation, normalement créée pour l'index. Ainsi, nom-schéma, nom-table et le nom de la table de capture font référence à des objets de la base de données DB2 (fédérée) locale.

Le schéma-contrôle-capture correspond au nom de schéma des tables de contrôle de réplication, par exemple IBMSNAP_PRUNE_SET sur le système DB2 local. Les tables de contrôle de réplication doivent être disponibles en tant qu'alias de type nickname sur le système DB2 local après la configuration de la réplication.

Des alias de type nickname doivent être disponibles au minimum pour les tables de contrôle de capture suivantes :

- IBMSNAP_SIGNAL
- IBMSNAP_PRUNE_SET
- IBMSNAP_PRUNCNTL
- IBMSNAP_REGISTER
- IBMSNAP_REG_SYNC (sources éloignées non DB2 uniquement)

Comme le Centre de réplication DB2 ne garantit pas automatiquement la création d'alias de type nickname locaux pour une table de capture éloignée et des tables de contrôle de capture, cela peut être fait manuellement. Cette tâche est similaire à la création d'un alias de type nickname pour la table sur laquelle l'index de texte est créé.

Les noms des colonnes de clé primaire dans l'alias de type nickname de table utilisateur et l'alias de type nickname de table de capture doivent correspondre. De plus, les noms des colonnes IBMSNAP_OPERATION, IBMSNAP_COMMITSEQ et IBMSNAP_INTENTSEQ ne doivent pas être modifiées dans l'alias de type nickname de la table de capture.

Après la création de l'index, les noms de colonne DB2EXT.TEXTINDEXES(LOGVIEWNAME) et DB2EXT.TEXTINDEXES(LOGVIEWSCHEMA) font référence au nom local de la table de capture de réplication.

Comme l'Extension Net Search ne requiert pas toutes les fonctionnalités du Centre de réplication DB2, la table CD ou CCD doit satisfaire aux règles suivantes :

- Utilisation de l'enregistrement de capture de modification, et non d'une copie en mode régénération intégrale.
- Aucune définition de sous-ensemble horizontal des modifications de capture n'est autorisée. Par des déclencheurs, par exemple. Reportez-vous au Chapitre 6, "Subsetting data in your replication environment" du manuel *DB2 Replication Guide and Reference, Version 8*.
- L'enregistrement des modifications pour un sous-ensemble de colonnes n'est autorisé que si les colonnes de la clé primaire, la colonne de texte et toutes les colonnes impliquées dans les expressions d'attribut et de table de mémoire cache de la commande DB2TEXT CREATE INDEX sont incluses.
- Les colonnes de la clé primaire doivent être incluses dans la table de capture. Notez que l'image-après est suffisante.
- Les tables de capture ne doivent pas être condensées. Pour chaque clé primaire, il doit y avoir une entrée avec les données les plus récentes. Cependant, DB2 Extension Net Search exige qu'un historique complet soit disponible.
- La table doit utiliser l'option D/I. Ceci permet aux mises à jour de clés primaires sur la table source d'être transformées en paire d'insertions/suppressions.

Les autres conditions requises incluent les éléments suivants :

- Le type et la version du serveur de la table source sur laquelle l'index est créé, doivent être :
 - DB2/AIX V8.1 ou suivante
 - DB2/NT V8.1 ou suivante

Commande CREATE INDEX

- DB2/HP V8.1 ou suivante
- DB2/LINUX V8.1 ou suivante
- DB2/SUN V8.1 ou suivante
- DB2 z/OS V7.2 ou suivante
- DB2 OS/400 V5.2 ou suivante
- Informix IDS 9.33
- ORACLE 9i
- SYBASE ASE 12.5
- Microsoft SQL Server 2000
- Les encapsuleurs pris en charge sont les suivants :
 - DB2: DRDA*
 - Informix: Informix
 - ORACLE : NET8, (SQLNET)
 - SYBASE : CTLIB
 - MSSQLSERVER : MSSQLODBC3

Remarques et restrictions

Assurez-vous d'insérer le nom correct de la table source dans la table d'enregistrement. Selon le type de Système de gestion de base de données, il faut utiliser le nom de table éloigné ou l'alias de type nickname local :

- DB2 : nom de table éloigné (le nom de table sur le serveur éloigné)
- Non DB2 : alias de type nickname (alias de type nickname correspondant dans la base de données DB2 fédérée)

Un mappage utilisateur doit exister pour que l'utilisateur local puisse accéder à la source de données éloignée via des alias de type nickname et l'utilisateur éloigné doit disposer des droits de contrôle sur les tables.

Si l'ID utilisateur du propriétaire d'instance DB2 est différent de l'ID utilisateur local, un mappage utilisateur supplémentaire pour l'ID utilisateur du propriétaire d'instance DB2 est nécessaire.

Le nom de la table de base spécifié ne doit pas être une vue sur un alias de type nickname. Ceci est dû au fait qu'une vue peut être sur plusieurs alias de type nickname et que plusieurs tables CD et CCD peuvent être également impliquées. Du fait qu'une seule table CD ou CCD peut être spécifiée dans la clause de capture de réplication, une vue sur des alias de type nickname ne peut pas être prise en charge. En outre, des alias de type nickname sur une vue éloignée ne peuvent pas être pris en charge car la clé primaire est manquante.

La table CD ou CCD doit être un alias de type nickname et ne peut pas être une vue ou un alias.

Pour plus d'informations sur le manuel *DB2 Replication Guide and Reference Version 8*, reportez-vous à la section «Informations connexes» à la page x.

COMMITCOUNT FOR UPDATE nombre

Pour le traitement de mise à jour **incrémentielle**, un nombre de validations

peut être spécifié. Pour plus d'informations, reportez-vous à la section «UPDATE INDEX» à la page 134. Si ce n'est pas le cas, une valeur par défaut est tirée de la vue DB2EXT.DBDEFAULTS, dans laquelle DEFAULTNAME='COMMITCOUNT'.

La valeur COMMITCOUNT FOR UPDATE destinée à l'index figure dans DB2EXT.TEXTINDEXES.COMMITCOUNT. Elle peut être modifiée pour chaque index, à l'aide de la commande ALTER INDEX. Elle s'applique également au traitement de la mise à jour planifiée en fonction de la spécification de la commande UPDATE FREQUENCY. Une valeur égale à 0 signifie que la mise à jour est terminée dans une transaction, avec des valeurs supérieures à 0 spécifiant le nombre de documents à traiter dans une transaction.

L'utilisation du nombre de validations a des effets sur les performances. Pour plus d'informations, reportez-vous à la section «Remarques sur les performances» à la page 45.

RECREATE INDEX ON UPDATE

Ceci ne permet pas les mises à jour incrémentielles d'index, mais recrée l'index lorsqu'une mise à jour est effectuée (mise à jour planifiée ou par commande). Pour plus d'informations, reportez-vous aux Notes d'utilisation de la section «UPDATE INDEX» à la page 134.

Remarque

Aucun déclencheur n'est créé sur la table utilisateur et aucune table de journalisation n'est créée.

INDEX DIRECTORY répertoire

Chemin d'accès au répertoire dans lequel l'index de texte doit être stocké. Comme le répertoire est destiné à contenir les données de l'index, assurez-vous que le répertoire est accessible en lecture/écriture et exécutez les droits pour l'ID utilisateur du propriétaire d'instance DB2.

La valeur par défaut est tirée de la vue DB2EXT.DBDEFAULTS, dans laquelle DEFAULTNAME=INDEXDIRECTORY'. Un sous-répertoire, NODE<nr>, est créé sous le répertoire pour distinguer les index sur des noeuds logiques d'un serveur.

Notez que dans un environnement DB2 réparti, ce répertoire doit exister sur chaque noeud physique.

WORK DIRECTORY répertoire

Un répertoire de travail distinct peut être spécifié en option. Il sera utilisé pour stocker les fichiers temporaires pendant la recherche sur l'index et les opérations d'administration. Le répertoire doit exister avec des droits d'accès en lecture, écriture et exécution pour l'ID utilisateur du propriétaire d'instance DB2.

La valeur par défaut est tirée de la vue DB2EXT.DBDEFAULTS, dans laquelle DEFAULTNAME='WORKDIRECTORY'. Un sous-répertoire, NODE<nr>, est créé sous le répertoire pour distinguer les index sur des noeuds logiques d'un serveur.

Notez que dans un environnement DB2 réparti, ce répertoire doit exister sur chaque noeud physique.

ADMINISTRATION TABLES IN nom-espace-table

Nom de l'espace table normal pour les tables d'administration créées pour

Commande CREATE INDEX

l'index. L'espace table doit exister. S'il n'est pas spécifié, l'espace table de la table utilisateur est choisi, si l'index est créé sur une table de base.

Dans le cas d'un alias de type nickname ou d'une vue, un espace table par défaut est choisi par DB2.

Lors de la création d'index de texte sur des vues, des alias de type nickname ou d'index de texte pour une recherche de procédure mémorisée dans un environnement DB2 réparti, l'espace table ne doit posséder qu'un seul noeud.

CACHE TABLE (liste-expression-colonne-SQL)

Une table mise en cache est générée en plus de l'index, constituée des expressions de la colonne spécifiée. Cette mémoire cache est utilisée pour renvoyer le jeu de résultats via une recherche de procédure mémorisée sans jointure des résultats de la recherche en texte intégral avec une table DB2. Notez qu'une recherche DB2 normale à l'aide de l'index en texte intégral et contenant la fonction CONTAINS est toujours possible.

Définissez les expressions de la colonne SQL à l'aide de noms de colonne non qualifiés de la table sur laquelle l'index est créé. Les types d'expression de la colonne SQL autorisés sont tous les types distincts intégrés et définis par l'utilisateur. Les noms de colonne dans le jeu de résultats sont déterminés en respectant les règles suivantes :

- S'il est explicitement désigné par la clause SQL AS dans l'expression de la colonne, le nom spécifié est utilisé. Par exemple : `CACHE TABLE (C1+C2 AS monNom)`
- Si une colonne de la table spécifiée est utilisée sans la clause AS, le nom de la colonne est utilisé. Par exemple : `CACHE TABLE(C1)`
- Si une expression est utilisée sans la clause AS et ne fait pas référence à une colonne nommée, la commande CREATE INDEX génère une erreur.
- Aucun nom de colonne en double n'est permis.

Les types de données CLOB ne sont pas pris en charge comme types de données de mémoire cache. Vous devez les transtyper VARCHARS.

Remarque

Notez que si les noms de colonne du jeu de résultats ne sont pas disjoints, la commande CREATE INDEX renvoie une erreur. Notez également que la table mise en cache n'est pas implicitement activée après création. Par exemple, la recherche par procédure mémorisée n'est pas possible tant que la commande DB2TEXT ACTIVATE CACHE n'a pas été effectuée.

Cette option peut être utilisée dans un environnement DB2 réparti, uniquement si la table utilisateur est stockée dans un espace table ne possédant qu'un seul noeud.

PERSISTENT IN répertoire

Indique que la mémoire cache est également créée de manière persistante et peut être activée peu de temps après une désactivation ou une relance du système. La mémoire cache persistante est stockée dans le répertoire spécifié.

Notez que si le répertoire n'est pas spécifié, la valeur par défaut est tirée de la vue `db2ext.dbdefaults`, dans laquelle `DEFAULTNAME='CACHEDIRECTORY'`.

TEMPORARY

Spécifie que la mémoire cache n'est pas conservée de manière persistante. Si aucune des valeurs PERSISTENT ou TEMPORARY n'est spécifiée, la valeur par défaut est tirée de la vue DB2EXT.DBDEFAULTS, dans laquelle DEFAULTNAME='USEPERSISTENTCACHE'.

MAXIMUM CACHE SIZE **taille-mémoire**

Indique la taille maximale de la table mise en cache devant être générée pendant la commande DB2TEXT ACTIVATE CACHE. La taille doit être spécifiée en mégaoctets (Mo) sous forme d'un entier positif. Il n'existe aucune valeur par défaut pour ce paramètre. Si le nombre entier est trop petit, la commande ACTIVATE CACHE échouera. La taille de la mémoire cache réelle est calculée pendant la commande ACTIVATE CACHE.

Les valeurs limites de la taille maximale de la mémoire cache pour les différentes plateformes sont les suivantes :

- Windows : 1024 Mo (1 Go = 1073741824 octets)
- AIX : 1536 Mo (1,5 Go = 1610612736 octets)
- Solaris, Linux, HP-UX : 2048 Mo (2 Go = 2147483647 octets)

Pour plus d'informations, reportez-vous à la section Annexe B, «Utilisation de grandes quantités de mémoire», à la page 187.

PCTFREE **pourcentage**

Spécifie le pourcentage de mémoire cache à garder libre pour les documents supplémentaires. Le pourcentage doit être un entier inférieur à 100 et supérieur ou égal à 0. S'il n'est pas spécifié, la valeur par défaut est tirée de la vue db2ext.dbdefaults, dans laquelle DEFAULTNAME='PCTFREE'.

Pour plus de détails, voir «ACTIVATE CACHE» à la page 106.

INITIAL SEARCH RESULT ORDER (**ordre-SQL-par-liste**)

Spécifie l'ordre utilisé pour extraire le contenu de la table utilisateur pendant l'indexation initiale. En utilisant cette option et en passant le rang dynamique des résultats de la recherche en texte intégral, les documents sont renvoyés dans leur ordre d'indexation, tels qu'ils sont stockés dans la table de résultats mise en cache.

Pour plus d'informations, reportez-vous au Chapitre 16, «Fonction de recherche de procédure mémorisée», à la page 161.

Vous ne pouvez travailler avec des index pré-enregistrés que si vous utilisez l'interface de recherche de procédure mémorisée.

Remarque

L'ordre de l'index **ne peut pas** être garanti pour les documents nouveaux ou modifiés après une mise à jour incrémentielle. Par exemple : INITIAL RESULT ORDER(length(column1) asc, column2+column3 desc)

KEY COLUMNS FOR INDEX ON VIEW (**liste-nom-colonne-SQL**)

Si des index sur des vues sont créés, la clause KEY COLUMNS FOR INDEX ON VIEW doit être spécifiée, sinon elle NE DOIT PAS l'être. La liste de noms de colonne spécifie les colonnes qui identifient DE MANIERE UNIQUE une ligne dans la vue.

Commande CREATE INDEX

Comme ce caractère unique ne peut pas être vérifié par DB2 comme dans le cas de clés primaires, il incombe à l'utilisateur d'assurer le caractère unique équivalent. Les colonnes spécifiées constituent une partie de la table de journalisation pour l'index.

INDEX CONFIGURATION (valeur-option), ...

Il s'agit des valeurs de configuration de l'index. Les valeurs par défaut sont soulignées.

Option	Valeurs	Description
TreatNumbersAsWords	<u>0</u> ou 1	Interprète les séquences de chiffres comme des mots distincts, même si elles sont adjacentes à des caractères. Par exemple, la valeur par défaut 0 signifie que tea42at5 est considéré comme un mot.
IndexStopWords	<u>0</u> ou 1	Considère ou ignore les mots vides lors de l'indexation. Actuellement, la liste de mots vides se trouve dans un fichier UCS-2 <langue>.tsw dans le répertoire <instance>/sql11b/db2ext/resources. Les modifications sur ce fichier n'ont aucun effet après la création de l'index. Notez également que <langue> correspond à la valeur LANGUAGE de la commande CREATE INDEX.
UpdateDelay	secondes	Indique la durée en secondes pour une mise à jour incrémentielle sans tables de capture. Seules les entrées antérieures à cette durée seront prises dans la table de journalisation. Ainsi vous éviterez les mises à jour perdues, par exemple, des modifications de document qui ne sont pas reflétées dans l'index dans les scénarios de transaction où les transactions de l'utilisateur interfèrent avec les commandes de mise à jour. Par conséquent, le paramètre UpdateDelay doit être défini à une durée maximale d'une transaction d'écriture utilisateur sur la table sur laquelle l'index a été créé.

CONNECT TO nom-de-base-de-données

Nom de la base de données qui constitue la cible de cette commande. Ce paramètre peut être omis, si DB2DBDFT est défini et que l'utilisateur exécute la commande sur le serveur. Notez que l'ID utilisateur doit disposer des droits DB2 requis.

USER id-utilisateur USING mot-de-passe

Utilisez un mot de passe et un ID utilisateur pour vous connecter à la base de données. Sans spécification de votre part, la connexion est tentée sous l'ID utilisateur courant et sans mot de passe.

Modifications de la base de données

- Modifiez les vues de catalogue DB2 Extension Net Search.
- Créez une table de journalisation d'index dans l'espace table spécifié. Ceci s'applique uniquement si l'option RECREATE INDEX et la table de capture ne sont pas spécifiées,
- Créez une table d'événements d'index dans l'espace table spécifié.
- Différées jusqu'à la première mise à jour : création des déclencheurs sur la table texte utilisateur (uniquement si RECREATE INDEX n'est pas spécifié et qu'aucune table de capture n'est utilisée).
- Si une table de capture de réplication est utilisée, la modification suivante est effectuée sur les tables de contrôle de capture :
 - une insertion dans les tables IBMSNAP_PRUNCTNL et IBMSNAP_PRUNE_SET

Les entrées de ces tables sont identifiées de manière unique par les colonnes :

- APPLY_QUAL = 'NSE' || <instance DB2 exécutant NSE>
- SET_NAME = <identificateur index interne>
- TARGET_SERVER = <nom de base de données DB2 cible de l'opération DB2TEXT>

Reportez-vous à la page 34 pour plus d'informations sur les colonnes.

Modifications de la mémoire partagée

Différées jusqu'à l'exécution de ACTIVATE : si la clause CACHE TABLE est utilisée, une mémoire cache est générée pour la table de résultats dans la *mémoire partagée*.

Modifications du système de fichiers

- Des sous-répertoires NODE<nr> sont créés sous les répertoires d'index, de travail et de mémoire cache.
- Le répertoire <nom index interne> est créé sous <répertoire-index>/NODE<nr>, où répertoire-index fait référence au paramètre correspondant de cette commande et NODE<nr> correspond au numéro du noeud dans un environnement DB2 réparti.

Utilisation

La création d'un index en texte intégral requiert une clé primaire sur la table utilisateur. Dans DB2 Extension Net Search Version 8.1, une clé primaire DB2 comportant plusieurs colonnes peut être utilisée sans restriction de type. Cependant, pour utiliser la recherche de valeur table, aucune clé primaire composée n'est autorisée.

Le nombre de colonnes de clé primaire est limité à 14, la longueur totale de toutes les colonnes de clé primaire est limitée à $1024 - 14 = 1010$ octets.

- La taille totale des expressions SQL pour ATTRIBUTES, CACHE TABLE et INITIAL SEARCH RESULT ORDER ne doit pas dépasser 24 Ko.
- Les mises à jour de l'index initial sont toujours effectuées comme une transaction logique. Dans ce cas, il n'y a pas de nombre de validation.

Remarque

Après la création de l'index, la longueur des colonnes de clé primaire ou de vue ne doit pas être modifiée par les commandes ALTER TABLE.

La synchronisation entre la table utilisateur, l'index en texte intégral et la table de résultats mise en cache est réalisée pendant la commande UPDATE INDEX. Pour plus d'informations, reportez-vous à la section «UPDATE INDEX» à la page 134.

DEACTIVATE CACHE

Cette commande libère une table mise en cache. Une mémoire cache persistante est conservée pour être réutilisée lors de la prochaine commande ACTIVATE. Jusqu'à la prochaine activation, les opérations de recherche via la procédure mémorisée ne sont plus possibles sur la mémoire cache désactivée.

Autorisation

Selon les vues de catalogue DB2, l'ID utilisateur de cette commande doit disposer des droits CONTROL sur la table pour laquelle l'index en texte intégral a été créé.

Syntaxe de commande

```

▶▶ DEACTIVATE-CACHE-FOR-INDEX [schéma-index"." ] nom-index FOR-TEXT
|
| options-connexion |
|
▶▶
  
```

options-connexion :

```

|
| CONNECT-TO nom-de-base-de-données
| USER id-utilisateur USING mot-de-passe
|
|
  
```

Paramètres de commande

schéma-index

Schéma de l'index de texte, tel qu'il est spécifié dans la commande CREATE INDEX. Si aucun schéma n'est spécifié, l'ID utilisateur de la connexion DB2 est utilisé comme nom de schéma.

nom-index

Nom de l'index de texte, tel qu'il est spécifié dans la commande CREATE INDEX.

CONNECT TO nom-de-base-de-données

Nom de la base de données qui constitue la cible de cette commande. Ce paramètre peut être omis, si DB2DBDFT est défini et que l'utilisateur exécute la commande sur le serveur. Notez que l'ID utilisateur doit disposer des droits DB2 requis.

USER id-utilisateur USING mot-de-passe

Utilisez un mot de passe et un ID utilisateur pour vous connecter à la base de données. Sans spécification de votre part, la connexion est tentée sous l'ID utilisateur courant et sans mot de passe.

Utilisation

Notez que cette commande ne peut pas être émise si l'une des commandes suivantes est en cours d'exécution sur l'index :

- ACTIVATE CACHE
- DEACTIVATE CACHE
- UPDATE INDEX
- ALTER INDEX

- DROP INDEX
- CLEAR EVENTS

Remarque

Après désactivation d'une mémoire cache persistante, la mémoire cache est rendue inaccessible à la recherche par procédure mémorisée. Cependant, elle peut être utilisée pour une commande rapide ACTIVATE, à moins qu'une mise à jour n'ait été effectuée entre temps.

Dans ce cas, la mémoire cache persistante est automatiquement recrée à l'aide de la commande ACTIVATE CACHE.

DROP INDEX

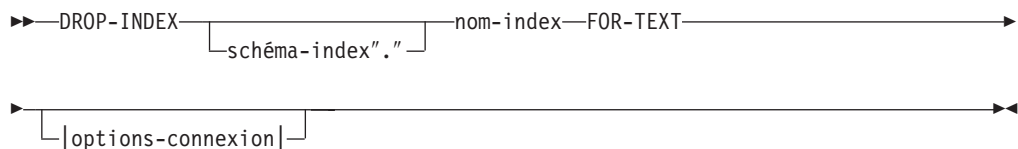
Cette commande supprime un index en texte intégral d'une colonne de texte. Si la mémoire cache pour l'index est activée, elle est supprimée à l'aide de cette commande.

Autorisation

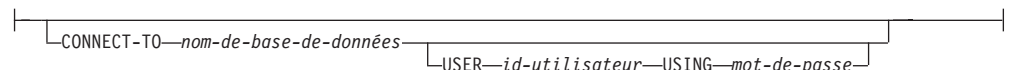
Selon les vues de catalogue DB2, l'ID utilisateur de cette commande doit disposer des droits CONTROL sur la table pour laquelle l'index en texte intégral a été créé. L'utilisateur peut aussi être l'administrateur de la base de données (DBADM).

L'administrateur de la base de données (DBADM) peut également supprimer l'index du fait qu'il doit pouvoir désactiver la base de données à l'aide de l'option FORCE.

Syntaxe de commande



options-connexion :



Paramètres de commande

schéma-index

Schéma de l'index de texte, tel qu'il est spécifié dans la commande CREATE INDEX. Si aucun schéma n'est spécifié, l'ID utilisateur de la connexion DB2 est utilisé comme nom de schéma.

nom-index

Nom de l'index, tel qu'il est spécifié dans la commande CREATE INDEX. Avec le schéma d'index, il identifie de manière unique l'index en texte intégral dans une base de données.

CONNECT TO nom-de-base-de-données

Nom de la base de données qui constitue la cible de cette commande. Ce paramètre peut être omis, si DB2DBDFT est défini et que l'utilisateur exécute la commande sur le serveur. Notez que l'ID utilisateur doit disposer des droits DB2 requis.

USER id-utilisateur USING mot-de-passe

Utilisez un mot de passe et un ID utilisateur pour vous connecter à la base de données. Sans spécification de votre part, la connexion est tentée sous l'ID utilisateur courant et sans mot de passe.

Utilisation

L'index est supprimé, quel que soit l'état d'activation de sa table mise en cache. Pour plus d'informations, reportez-vous à la section «ACTIVATE CACHE» à la page 106.

Notez que la commande ne peut pas être émise si l'une des commandes suivantes est en cours d'exécution sur l'index :

- UPDATE INDEX
- CLEAR EVENTS
- ALTER INDEX
- ACTIVATE CACHE
- DEACTIVATE CACHE
- DROP INDEX

Remarque

Les index doivent être supprimés manuellement, avant ou après que la table utilisateur ne soit supprimée de DB2. Sinon, les résultats ne sont pas effacés correctement.

Modifications de la base de données

- Modifiez les vues de catalogue DB2 Extension Net Search.
- Supprimez l'index DB2.
- Supprimez les tables d'événements/de journalisation de l'index.
- Supprimez les déclencheurs de la table texte utilisateur.

Lorsque vous utilisez des tables de capture de réplication, les entrées dans les tables IBMSNAP_PRUNE_SET et IBMSNAP_PRUNCTRL sont supprimées.

Modifications de la mémoire partagée

La table mise en cache est supprimée.

Modifications du système de fichiers

- Le répertoire <nom index interne> est supprimé de l'index et les répertoires de travail de l'index supprimé.
- Suppression de la mémoire cache persistante pour l'index.

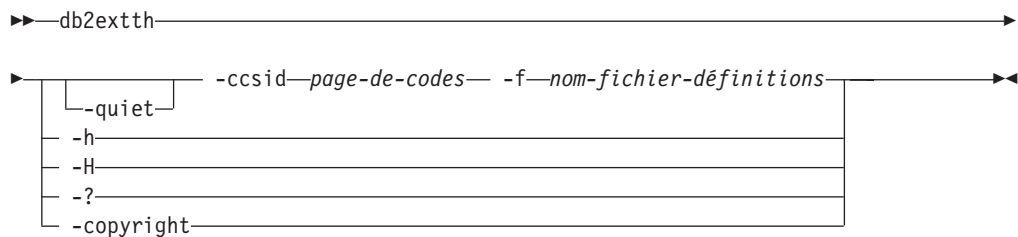
DB2EXTTH (Utilitaire)

Cet utilitaire indépendant compile un fichier de définitions de thésaurus. Après l'exécution du compilateur de thésaurus, les fonctions liées au thésaurus de la syntaxe de l'argument de recherche peuvent être utilisées.

Autorisation

Aucune. Cette commande ne se limite pas nécessairement qu'au propriétaire de la table, mais n'a de sens que dans le contexte d'une requête.

Syntaxe de commande



Paramètres de commande

-f nom-fichier-définitions

Nom du fichier contenant les définitions du thésaurus. Le nom de fichier doit contenir le chemin d'accès absolu ou relatif au fichier. La longueur du nom de fichier est limitée à 8+3 caractères, l'extension étant facultative.

Le dictionnaire thésaurus est généré dans le même répertoire que le fichier de définitions et porte le même nom. La seule différence est que le dictionnaire a les extensions suivantes : wdf, wdv, grf, grv, MEY, ROS, NEY, SOS, et lkn, où n est un chiffre. Notez que si des fichiers de thésaurus existants portent le même nom, ils s'écrasent.

-ccsid page-de-codes

Page de codes dans laquelle le fichier de définitions du thésaurus est écrit. Pour la liste des pages de codes prises en charge pour un thésaurus, reportez-vous à l'annexe Annexe L, «CCSID pris en charge par le thésaurus», à la page 251.

-quiet Les informations de sortie ne sont pas affichées.

-copyright

Renvoie le numéro de compilation interne du produit. Utilisez ce numéro si vous devez signaler des problèmes.

-h, -H ou -?

Affiche des informations d'aide.

Utilisation

Cette commande permet de compiler un fichier de définition du thésaurus dans un format de définition de thésaurus binaire.

Remarque

Le format est le même que dans DB2 Text Information Extender Version 7.2. Seules les modifications suivantes ont été apportées :

- Les nouvelles relations BROADER et NARROWER sont équivalentes aux relations HIGHER_THAN et LOWER_THAN utilisées précédemment dans Text Information Extender. Pour faire référence à ces anciennes relations de recherche, les nouvelles relations doivent être utilisées.
- Notez également que les fichiers de dictionnaire thésaurus doivent être stockés dans <selon-OS>/sql11ib/db2ext/thes pour être utilisables pendant la recherche, à moins que le thésaurus ne soit totalement qualifié dans la requête.

Pour plus d'informations, reportez-vous au Chapitre 10, «Utilisation d'un thésaurus pour étendre les termes recherchés», à la page 85.

UPDATE INDEX

Cette commande démarre immédiatement le processus d'indexation en mettant à jour l'index pour qu'il reflète le contenu courant des colonnes de texte avec lesquelles il est associé.

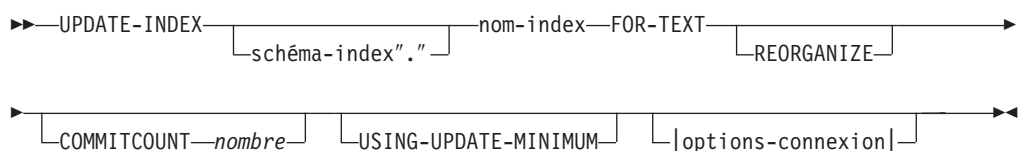
Pendant l'exécution de la mise à jour, la recherche avec le prédicat CONTAINS est possible. Pour un index avec une table de résultats mise en cache et activée, la recherche par procédure mémorisée est également possible pendant la mise à jour. Cependant, les colonnes de la table mise en cache peuvent afficher de nouvelles valeurs, même si le texte modifié n'est pas encore validé dans l'index en texte intégral.

L'utilisation de l'option RECREATE INDEX ON UPDATE dans la commande CREATE INDEX permet de vider l'index avant de le recréer. Jusqu'à la fin de la mise à jour, des résultats vides sont renvoyés.

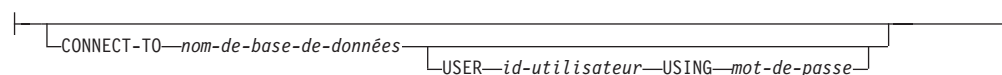
Autorisation

Selon les vues de catalogue DB2, l'ID utilisateur de cette commande doit disposer des droits CONTROL sur la table pour laquelle l'index en texte intégral a été créé.

Syntaxe de commande



options-connexion :



Paramètres de commande

schéma-index

Schéma de l'index de texte. Il est spécifié dans la commande CREATE INDEX. Si aucun schéma n'est spécifié, l'ID utilisateur de la connexion DB2 est utilisé.

nom-index

Nom de l'index de texte. Il est spécifié dans la commande CREATE INDEX.

REORGANIZE

Si une colonne de texte est fréquemment mise à jour, les mises à jour ultérieures de l'index peuvent devenir inefficaces. Pour rendre le processus de mise à jour à nouveau efficace, réorganisez l'index. La vue DB2EXT.TEXTINDEXES permet de déterminer si un index doit être réorganisé.

L'option REORGANIZE AUTOMATIC de la commande CREATE INDEX permet d'éviter la vérification et la réorganisation manuelles de l'index.

Remarque

Le processus de réorganisation a lieu après une mise à jour standard.

USING UPDATE MINIMUM

Utilise les paramètres UPDATE MINIMUM de la commande CREATE INDEX et lance une mise à jour incrémentielle uniquement si le nombre de modifications spécifié a été atteint. Par défaut, la mise à jour est lancée sans condition.

Pour des bases de données réparties, l'option UPDATE MINIMUM est vérifiée sur tous les noeuds.

Pour plus d'informations, reportez-vous à la section «CREATE INDEX» à la page 114.

COMMITCOUNT nombre

Un entier supérieur ou égal à 0 affiche le nombre de documents traités dans une transaction par le moteur de recherche et par DB2 pour les mises à jour incrémentielles de l'index.

Cependant, pour les mises à jour initiales, telles que la première mise à jour après la commande CREATE INDEX ou toute mise à jour avec l'option RECREATE INDEX ON UPDATE, une seule transaction logique ignore COMMITCOUNT. Ceci peut être changé grâce à la commande ALTER INDEX.

CONNECT TO nom-de-base-de-données

Nom de la base de données qui constitue la cible de cette commande. Ce paramètre peut être omis, si DB2DBDFT est défini et que l'utilisateur exécute la commande sur le serveur. Notez que l'ID utilisateur doit disposer des droits DB2 requis.

USER id-utilisateur USING mot-de-passe

Utilisez un mot de passe et un ID utilisateur pour vous connecter à la base de données. Sans spécification de votre part, la connexion est tentée sous l'ID utilisateur courant et sans mot de passe.

Utilisation

Cette commande s'exécute de manière synchrone. Elle démarre le traitement de la mise à jour sur tous les noeuds logiques/physiques DB2 requis dans un environnement DB2 réparti. La durée dépend du nombre de fichiers à indexer et du nombre de documents déjà indexés. L'état de la mise à jour peut être affiché dans une vue créée pour chaque index. Le nom de cette vue peut être extrait de DB2EXT.TEXTINDEXES dans la colonne EVENTVIEWNAME. Pour plus d'informations, reportez-vous à l'Annexe C, «Catalogues d'informations DB2 Extension Net Search», à la page 191.

Deux options sont possibles pour afficher le nombre de documents validés, ayant été traités. Pour déterminer si une mise à jour est toujours en cours d'exécution et combien de documents ont été validés dans l'index, utilisez la vue DB2EXT.TEXTINDEXES (NUMBER_DOCS). Pour obtenir des informations sur le démarrage, la validation des modifications et la fin du traitement de mise à jour, utilisez la vue des événements associée à l'index.

Pour afficher le nombre de documents non validés, ayant été traités, utilisez la commande CONTROL LIST ALL LOCKS FOR INDEX.

Remarque

Les vues affichent uniquement des informations relatives au noeud connecté.

Pour les mises à jour incrémentielles sur une table de base avec des noeuds physiques, l'heure doit être synchronisée sur chaque noeud. Si les heures ne sont pas synchronisées, les mises à jour peuvent être perdues ou ne pas prendre effet.

Il n'est pas possible d'émettre la commande si l'une des commandes suivantes est en cours d'exécution sur l'index :

- CLEAR EVENTS
- ALTER INDEX
- DROP INDEX
- ACTIVATE CACHE
- DEACTIVATE CACHE
- UPDATE INDEX

Après la mise à jour d'un index avec une table de résultats mise en cache persistante et désactivée, la mémoire cache persistante est supprimée de telle sorte que la commande suivante ACTIVATE CACHE la recrée en fonction du contenu de la base de données.

Si l'utilisateur interrompt cette commande, tous les processus impliqués dans la fonction de mise à jour s'arrêtent. Si un nombre de validations a été utilisé dans une mise à jour incrémentielle, certaines mises à jour peuvent être visibles dans l'index, alors que d'autres peuvent nécessiter une nouvelle commande de mise à jour.

Pour mettre fin à la mise à jour automatique d'un index, recherchez le processus du propriétaire d'instance DB2 exécutant la commande de mise à jour de l'index sur la partition utilisée pour les services de mise à jour. Arrêtez ce processus et le traitement de la mise à jour sur toutes les partitions.

Remarque

Comme la commande fonctionne en deux phases séparées pour la création de l'index sur toutes les partitions et les mises à jour initiales de l'index, émettez une commande `db2text drop index` pour s'assurer que l'index n'est pas en partie disponible. Si cette commande n'est pas émise, la mise à jour suivante qui peut être déclenchée par une commande de mise à jour ou l'option `UPDATE FREQUENCY`, effectuera une complète ré-indexation pour garantir un état cohérent.

Modifications de la base de données

- Insertions dans la table des événements.
- Suppression de la table de journalisation de l'index.

Lorsque vous utilisez des tables de capture de réplication, les modifications suivantes sont apportées à la base de données.

- Un signal est ajouté à la table `IBMSNAP_SIGNAL` avant de démarrer la mise à jour initiale

Commande UPDATE INDEX

- La valeur Synchpoint de la commande IBMSNAP_PRUNE_SET est modifiée après la mise à jour incrémentielle

HELP

Cette commande permet d'afficher la liste des commandes DB2TEXT disponibles ou la syntaxe d'une commande DB2TEXT individuelle.

Autorisation

Aucune requise.

Syntaxe de commande



Paramètres de commande

HELP ou ?

Affiche l'aide sur la commande spécifiée ou le code anomalie.

commande

Les premiers mots clés qui identifient une commande DB2TEXT :

- ENABLE
- DISABLE
- CREATE
- DROP
- ALTER
- UPDATE
- CLEAR
- START
- STOP
- CONTROL
- ACTIVATE
- DEACTIVATE

code-anomalie

Code anomalie d'une commande DB2 Extension Net Search.

Utilisation

Si d'autres mots clés que le premier sont spécifiés, le reste est ignoré et la syntaxe de la commande identifiée est affichée.

Si aucun paramètre 'commande' n'est spécifié après '?' ou 'HELP' (ou s'il n'y a aucun paramètre du tout), DB2TEXT dresse la liste de tous les paramètres de commande DB2TEXT disponibles.

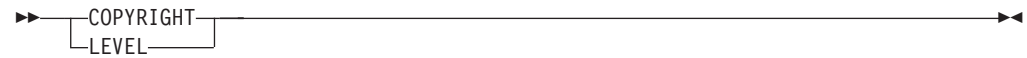
COPYRIGHT

Fournit les informations produit et de copyright de l'Extension Net Search.

Authorisation

Aucune requise.

Syntaxe de commande



Paramètres de commande

COPYRIGHT / LEVEL

Fournit des informations de copyright, le numéro de version et de compilation du produit.

Commande COPYRIGHT

Chapitre 14. Syntaxe des arguments de recherche

Un argument de recherche est la condition que vous spécifiez lorsque vous recherchez des termes dans des documents texte. Il est constitué de paramètres de recherche et d'un ou de plusieurs termes recherchés.

Des exemples d'arguments de recherche figurent à la section «Spécification des arguments de recherche SQL» à la page 73 et dans un fichier appelé search. Reportez-vous à la section «Exemples supplémentaires de syntaxe de recherche» à la page 78.

Les fonctions de recherche scalaires SQL utilisant des arguments de recherche sont les suivantes :

CONTAINS

Cette fonction utilise un argument de recherche pour rechercher du texte dans un document texte spécifique. Elle renvoie la valeur INTEGER 1 si le document contient le texte recherché ou toute relation spécifiée dans l'argument de recherche. Sinon, la valeur renvoyée est 0.

NUMBEROFMATCHES

Cette fonction utilise un argument de recherche pour effectuer une recherche dans les documents texte et renvoie une valeur INTEGER indiquant le nombre d'occurrences par document.

SCORE

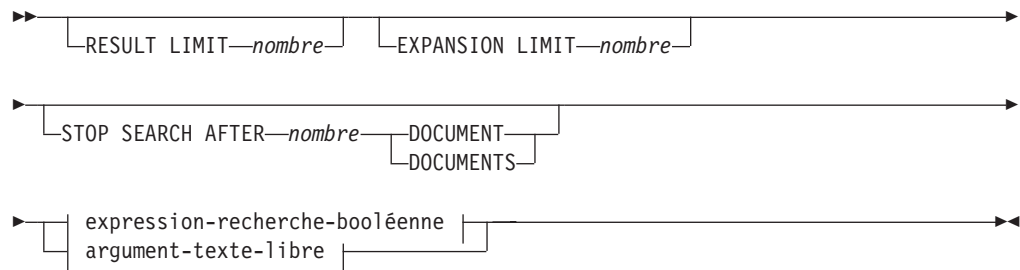
Cette fonction utilise un argument de recherche pour effectuer une recherche dans les documents texte. Elle renvoie une valeur pour chaque document trouvé, indiquant la précision avec laquelle le document trouvé est décrit par l'argument de recherche.

Remarque

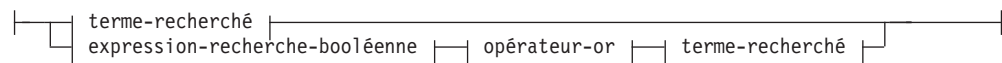
La même syntaxe est utilisée dans les arguments de la recherche de procédure mémorisée et de la fonction de valeur table SQL.

Argument de recherche

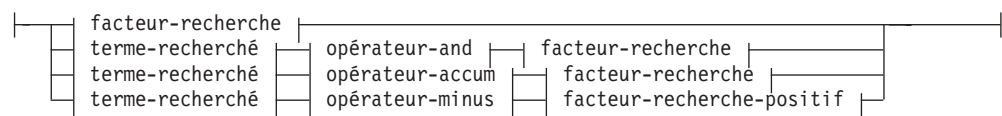
Syntaxe des arguments de recherche



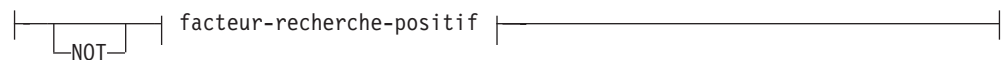
Expression-recherche-booléenne :



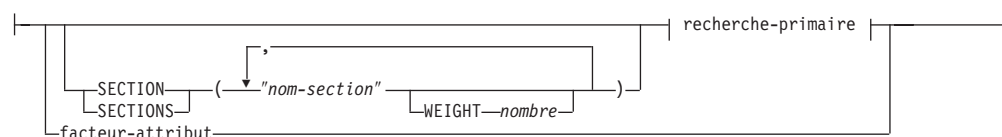
terme-recherché :



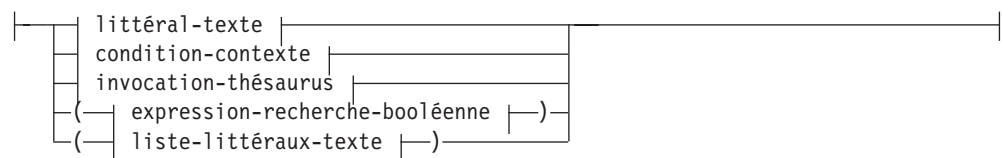
Facteur-recherche :



Facteur-recherche-positif :



Recherche-primaire :



Opérateur-and :



Opérateur-or :



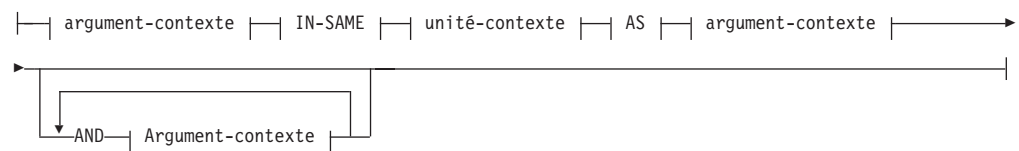
Opérateur-accum :



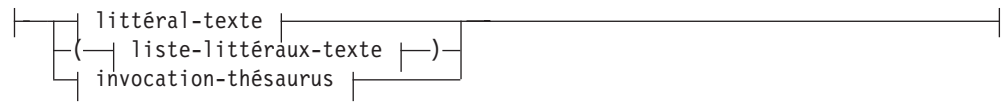
Opérateur-minus :



Condition-contexte :



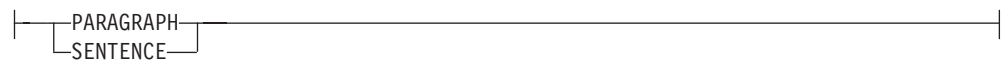
Argument-contexte :



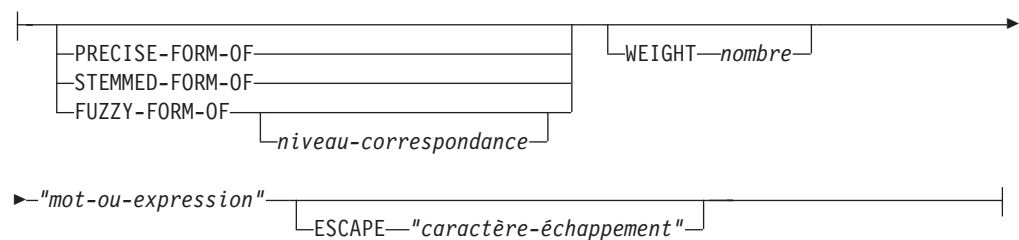
Liste-littéraux-texte :



Unité-contexte :

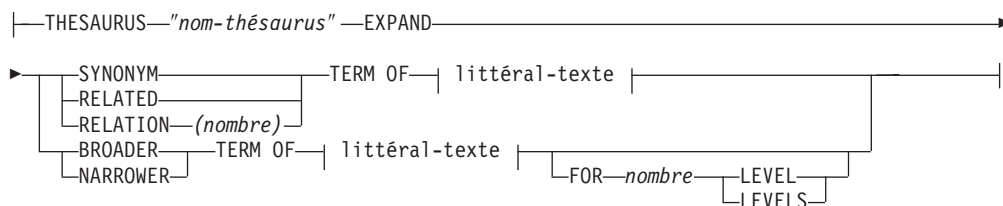


Littéral-texte :

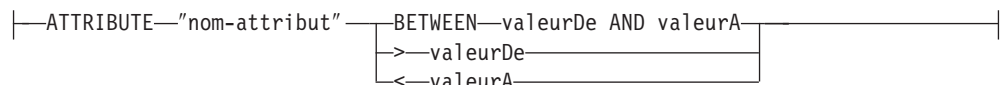


Syntaxe des arguments de recherche

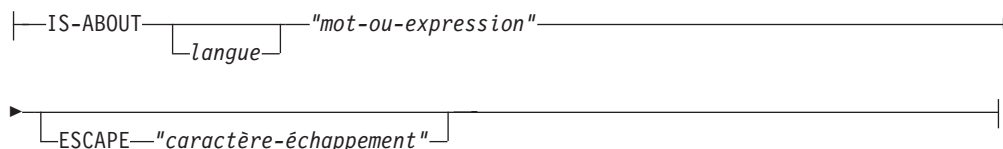
invocation-thésaurus :



Facteur-attribut :



argument-texte-libre :



Exemples

Des exemples figurent à la section «Spécification des arguments de recherche SQL» à la page 73.

Paramètres de recherche

RESULT LIMIT nombre

Mot clé spécifiant le nombre maximal de résultats à renvoyer par la recherche en texte intégral.

L'argument RESULT LIMIT doit être utilisé avec la fonction SCORE pour garantir l'évaluation des résultats renvoyés et le traitement des meilleurs résultats uniquement.

EXPANSION LIMIT nombre

Mot clé spécifiant le nombre maximal de fois qu'un terme peut être développé pour la recherche. Par exemple, pour déterminer combien de fois vous pouvez développer le terme recherché 'a*'.

STOP SEARCH AFTER nombre DOCUMENTS(S)

Mot clé spécifiant le seuil de recherche. La recherche est arrêtée lorsque le nombre de documents est atteint pendant la recherche et qu'un résultat intermédiaire est renvoyé. Une valeur plus faible augmente les performances de la recherche, mais aboutit à des résultats moins nombreux avec un rang potentiellement élevé.

Notez qu'il n'y a pas de valeur par défaut et que la valeur de *nombre* doit être un entier positif.

expression-recherche-booléenne

Les termes recherchés et les facteurs de recherche peuvent être associés à l'aide d'opérateurs booléens NOT, AND, OR, ACCUM et MINUS en

fonction des diagrammes de syntaxe. Les opérateurs ont priorité les uns par rapport aux autres, dans l'ordre suivant (le premier ayant la plus forte priorité) : NOT > MINUS = ACCUM = AND > OR comme le montre l'exemple suivant :

```
"Pilote" MINUS "passager" & "véhicule" |
"transport" & "public"
```

est évalué comme :

```
((("Pilote" MINUS "passager") & ("véhicule"))) | ("transport" &
"public")
```

L'opérateur correspond à la valeur true, si l'un des arguments booléens correspond à cette même valeur (qui est comparable à l'opérateur OR). La valeur de rang est calculée en accumulant les valeurs de rang des deux opérandes. L'opérateur ACCUM a la même liaison (priorité) que AND. L'opérateur MINUS correspond à la valeur true, si l'opérande de gauche correspond à la même valeur. La valeur de rang est calculée en prenant la valeur de rang de l'opérande de gauche et en soustrayant une pénalité, si l'opérande de droite correspond à la valeur true.

recherche-primaire

Une recherche primaire constituée d'un appel de thésaurus est évaluée à la valeur true, si l'un des littéraux texte étendus se trouve dans la section spécifiée du document. Une recherche primaire constituée d'une liste de littéraux texte est évaluée à la valeur true, si l'un des littéraux texte se trouve dans la section spécifiée du document.

SECTION(S) *nom-section*

Mot clé spécifiant une ou plusieurs sections dans un document structuré sur la ou lesquelles la recherche doit se limiter. Le nom de la section doit être spécifié dans un fichier modèle au moment de la création de l'index. Reportez-vous à la section «CREATE INDEX» à la page 114.

Les noms de section font la distinction majuscules-minuscules.

Assurez-vous que le nom de la section dans le fichier modèle est identique à celui de la requête.

Ce modèle décrit la structure des documents qui contiennent les sections identifiables, de telle sorte que des recherches peuvent être effectuées individuellement dans le contenu de ces sections. Les noms de section ne peuvent pas contenir de caractères génériques. Le *facteur-recherche-positive* utilisant la clause SECTION correspond à la valeur true, si la recherche primaire se trouve dans l'une des sections.

argument-contexte IN SAME unité-contexte AS argument-contexte AND argument-contexte ...

Cette condition permet de rechercher une combinaison de littéraux texte se trouvant dans le même paragraphe ou la même phrase. Les arguments de contexte sont toujours équivalents aux listes de littéraux texte et l'extension du thésaurus peut être utilisée pour étendre un littéral texte en une telle liste.

La condition est évaluée à true, si une unité-contexte se trouve dans le document (paragraphe, respectivement phrase) qui contient au moins l'un des littéraux texte de chaque argument-contexte développé, comme le montre l'exemple suivant :

```
("a","b") IN SAME PARAGRAPH AS ("c","d")
AND THESAURUS "t1" EXPAND SYNONYM TERM OF "e".
```

Si l'on présume que e1, e2 sont synonymes de e, les paragraphes suivants correspondent :

```
".. a c e .." ,  ".. a c e1.." ,  "a c e2..",  
".. a d e .." ,  ".. a d e1.." ,  "a d e2..",  
".. b c e .." ,  ".. b c e1.." ,  "b c e2..",  
".. b d e .." ,  ".. b d e1.." ,  "b d e2..".
```

PRECISE FORM OF

Mot clé entraînant le mot (ou chaque mot de l'expression) suivant PRECISE FORM OF à être recherché exactement tel qu'il est saisi. Cette forme de recherche fait la distinction majuscules-minuscules. Autrement dit, le respect de la casse est important. Par exemple, si vous recherchez le terme souris, vous ne trouverez pas "Souris".

STEMMED FORM OF

Mot clé entraînant le mot (ou chaque mot de l'expression) suivant STEMMED FORM OF à être réduit à sa racine avant d'être recherché. Cette forme de recherche ne fait pas de distinction entre les majuscules et les minuscules. Par exemple, si vous recherchez le terme souris, vous trouverez également "Souris".

La façon dont les mots sont réduits à leur racine dépend de la langue. Actuellement, seule la langue anglaise est prise en charge et le mot doit respecter les terminaisons standard.

FUZZY FORM OF

Mot clé destiné à la réalisation d'une recherche "floue" qui est une recherche des termes ayant une orthographe similaire au terme recherché. Ceci est particulièrement utile lors de la recherche dans des documents qui ont été créés par un programme de reconnaissance optique des caractères (OCR). De tels documents incluent souvent des mots mal orthographiés. Par exemple, le mot économie peut être reconnu comme écononie par un programme OCR. Notez que les trois premiers caractères doivent correspondre et que la recherche floue ne peut pas être utilisée si un mot de l'atome de recherche contient un caractère générique.

niveau de correspondance

Entier compris entre 1 et 100, spécifiant le degré de similitude, où 100 est plus proche que 1. 100 spécifie une "correspondance exacte" et 60 est déjà considéré comme une valeur très "floue". Plus le niveau de correspondance est flou, plus la durée de la recherche est longue, du fait que davantage de documents sont qualifiés pour la recherche. Le niveau de correspondance par défaut est 70.

WEIGHT nombre

Associe un littéral texte à une valeur WEIGHT pour modifier le score par défaut. Les valeurs WEIGHT autorisées sont des entiers compris entre 0 (valeur du score la plus faible) et 100 (la plus élevée). La valeur par défaut est 100.

mot-ou-expression

Mot ou expression à rechercher. Les caractères pouvant être utilisés dans un mot dépendent de la langue. Le fait que les mots doivent être séparés par des caractères séparateurs dépend également de la langue. En anglais et dans la plupart des autres langues, chaque mot d'une expression doit être séparé par un espace.

Pour rechercher une chaîne de caractères contenant des guillemets, saisissez les guillemets deux fois. Par exemple, pour rechercher le texte caractère "générique", tapez :

"caractère ""générique""

Notez que dans l'exemple, il n'est possible de rechercher qu'un seul ensemble de guillemets. Il n'est pas possible de rechercher deux guillemets dans une séquence. La longueur de chaque mot ou expression est également limitée à 128 octets.

Caractères génériques

Un mot peut contenir les caractères génériques suivants :

_ (trait de soulignement)

Représente n'importe quel caractère simple.

% (pour cent)

Représente n'importe quel nombre de caractères arbitraires. Si un mot est constitué d'un seul caractère %, il représente alors un mot facultatif de n'importe quelle longueur. Un mot ne peut pas être composé exclusivement de caractères génériques, sauf lorsqu'un seul % est utilisé pour représenter un mot facultatif. Si vous utilisez un caractère générique, vous ne pouvez pas utiliser le THESAURUS. Les caractères génériques ne peuvent pas suivre un caractère non alphanumérique.

ESCAPE caractère-échappement

Caractère identifiant le caractère suivant comme étant à rechercher et non à utiliser comme caractère générique. Par exemple, si un caractère d'échappement est \$, alors \$%, \$_ et \$\$ représentent respectivement %, _ et \$. Tous les caractères % et _ non précédés de \$ représentent des caractères génériques.

Lors de la recherche, vous ne pouvez utiliser que des caractères mono-octets. Aucun caractère à deux octets n'est autorisé.

THESAURUS *nom-thésaurus*

Mot clé utilisé pour spécifier le nom du thésaurus à utiliser pour étendre un littéral texte. Le nom du thésaurus est celui du fichier (sans son extension) d'un thésaurus compilé à l'aide du compilateur de thésaurus. Il doit se trouver dans <selon-OS>/sql11b/db2ext/thes. Le chemin d'accès peut être également spécifié, précédant le nom de fichier.

EXPAND *relation*

Spécifie la relation devant être utilisée pour étendre le littéral texte à l'aide du thésaurus. Le thésaurus a prédéfini les relations décrites dans la commande DB2EXTTH. Elles sont référées à l'aide des mots clés suivants :

- SYNONYM, relation symétrique exprimant l'équivalence.
- RELATED, relation symétrique exprimant l'association.
- BROADER, relation hiérarchique dirigée, pouvant être suivie de niveaux de profondeur spécifiés.
- NARROWER, relation hiérarchique dirigée, pouvant être suivie de niveaux de profondeur spécifiés.

Pour les relations définies par l'utilisateur, utilisez RELATION(nombre) qui correspond à la définition de relation dans DB2TEXTTH.

TERM OF *littéral-texte*

Littéral texte auquel d'autres termes de recherche doivent être ajoutés du thésaurus.

nombre LEVELS

Mot clé utilisé pour spécifier le nombre de niveaux (la profondeur) des termes du thésaurus, utilisés pour étendre la recherche pour une relation donnée. Si ce mot clé n'est pas spécifié, un nombre égal à 1 est présumé. La valeur de profondeur doit être un entier positif.

ATTRIBUTE *Nom-attribut*

Recherches de documents ayant des attributs correspondant à la condition spécifiée. Le nom-attribut fait référence au nom d'une expression d'attribut dans la commande CREATE INDEX ou à une définition d'attribut dans le fichier modèle de document.

Le facteur-attribut est autorisé pour les attributs de type DOUBLE uniquement. La précision de la valeur est garantie jusqu'à 15 chiffres. Les nombres de 16 chiffres et plus sont arrondis. L'utilisation de caractères génériques n'est pas autorisée dans nom-attribut, valeurDe et valeurA. Pour plus d'informations, reportez-vous à ce qui suit :

BETWEEN *valeurDe AND valeurA*

Un facteur d'attribut BETWEEN correspond à la valeur true, si la valeur de l'attribut est supérieure à (différente de) valeurDe et inférieure à (différente de) valeurA.

>valeurDe

Un facteur d'attribut ">" correspond à la valeur true si la valeur de l'attribut est supérieure (différente de) à valeurDe.

<valeurA

Un facteur d'attribut "<" correspond à la valeur true, si la valeur de l'attribut est inférieure (différente de) valeurA.

Si le nom de l'attribut dans la commande CREATE INDEX est spécifié avec des guillemets ou est défini dans un fichier modèle, le nom de l'attribut spécifié doit correspondre exactement. En revanche, si aucun guillemet n'est spécifié dans la commande CREATE INDEX, le nom de l'attribut doit être en majuscules.

IS ABOUT *langue mot-ou-expression*

Option permettant de spécifier un argument de recherche en texte libre. Elle doit être utilisée pour obtenir un type différent d'algorithme de score, du fait qu'elle vérifie le positionnement des termes dans les documents. Plus les termes utilisés dans le mot-ou-expression sont proches, plus le nombre de mots inclus dans le document sont nombreux et plus le score renvoyé est élevé.

Les valeurs autorisées pour la langue sont décrites à l'Annexe E, «Langues prises en charge», à la page 207 et ne s'appliquent qu'à la langue Thai. Si elle n'est pas spécifiée, la langue en_US est utilisée par défaut. La langue est utilisée uniquement pour la tokenisation du mot ou de l'expression.

Notez que l'argument IS ABOUT est utile uniquement si les valeurs des scores sont demandées et que les résultats de la recherche sont triés par valeur de score.

Chapitre 15. Fonction de recherche scalaire SQL et fonction de valeur table SQL

DB2 Extension Net Search fournit des fonctions de recherche scalaires SQL et une fonction de valeur table SQL pour la recherche dans des documents texte stockés dans DB2 Universal Database.

Ce chapitre décrit les fonctions de recherche SQL suivantes.

Résumé des fonctions de recherche

Fonction de recherche	Objectif	Page
CONTAINS	Recherche du texte dans un document spécifique.	150
NUMBEROFMATCHES	Recherche et renvoie le nombre d'occurrences trouvées.	151
SCORE	Recherche et renvoie la valeur du score d'un document texte trouvé.	152
DB2EXT.TEXTSEARCH	La fonction de valeur table SQL renvoie une table de clés primaires trouvées, le nombre d'occurrences et/ou les valeurs des scores.	153
DB2EXT.HIGHLIGHT	Pour obtenir des informations sur les raisons pour lesquelles un document a été qualifié de résultat de la recherche,	157

Pour obtenir des exemples sur l'utilisation de la fonction de recherche scalaire SQL et de la fonction de valeur table SQL, reportez-vous au Chapitre 8, «Recherche», à la page 71.

CONTAINS

La fonction scalaire CONTAINS recherche du texte dans un document texte indexé par l'Extension Net Search. Elle renvoie un entier de valeur 1 si le document contient le texte ou toute relation spécifiée dans l'argument de recherche. Sinon, la valeur renvoyée est 0.

Syntaxe de la fonction

►►—CONTAINS—(—*nom-colonne*—,—*argument-recherche*—)————►►

Paramètres de fonction

nom-colonne

Nom d'une colonne. Un index de texte doit être associé à la colonne. Il est possible de créer des index de texte à l'aide de la commande d'administration DB2TEXT CREATE INDEX.

argument-recherche

Chaîne de type VARCHAR contenant les termes à rechercher. Reportez-vous au Chapitre 14, «Syntaxe des arguments de recherche», à la page 141.

Remarque

Vous ne pouvez pas utiliser la requête CONTAINS sur un index de texte créé sur une vue.

NUMBEROFMATCHES

La fonction scalaire NUMBEROFMATCHES effectue une recherche dans des documents texte et renvoie un entier indiquant le nombre d'occurrences trouvées par document.

Syntaxe de la fonction

►►—NUMBEROFMATCHES—(—*nom-colonne*—,—*argument-recherche*—)—►►

Paramètres de fonction

nom-colonne

Nom d'une colonne. Un index de texte doit être associé à la colonne. Il est possible de créer des index de texte à l'aide de la commande d'administration DB2TEXT CREATE INDEX.

argument-recherche

Chaîne de type VARCHAR contenant les termes à rechercher. Reportez-vous au Chapitre 14, «Syntaxe des arguments de recherche», à la page 141.

Remarque

Vous ne pouvez pas utiliser la requête NUMBEROFMATCHES sur un index de texte créé sur une vue.

SCORE

La fonction scalaire SCORE effectue une recherche dans des documents texte et renvoie la valeur du score pour chaque document trouvé, indiquant la précision avec laquelle le document trouvé est décrit par l'argument de recherche.

SCORE renvoie une valeur DOUBLE. Plus le terme recherché dans le document est fréquent, plus le score du document augmente.

Syntaxe de la fonction

►►—SCORE—(—*nom-colonne*—,—*argument-recherche*—)—►►

Paramètres de fonction

nom-colonne

Nom d'une colonne. Un index de texte doit être associé à la colonne. Il est possible de créer des index de texte à l'aide de la commande d'administration DB2TEXT CREATE INDEX.

argument-recherche

Chaîne de type VARCHAR contenant les termes à rechercher.
Reportez-vous au Chapitre 14, «Syntaxe des arguments de recherche», à la page 141.

Remarque

Vous ne pouvez pas utiliser la requête SCORE sur un index de texte créé sur une vue.

DB2EXT.TEXTSEARCH

En plus des fonctions de recherche de procédure mémorisée et de recherche scalaire SQL, l'Extension Net Search fournit deux fonctions de valeur table SQL qui ressemblent beaucoup à la procédure mémorisée.

Les deux fonctions de valeur table s'appellent `db2ext.textsearch`. Le seul point qui les différencie est que l'une prend en charge la fonction `HIGHLIGHT` et comporte deux paramètres supplémentaires `numberOfHits` et `hitInformation`.

Notez que vous ne pouvez pas utiliser la fonction de valeur table sur des tables avec un clé primaire composée.

Pour plus d'informations sur l'utilisation de la fonction `HIGHLIGHT`, reportez-vous à la section «`DB2EXT.HIGHLIGHT`» à la page 157.

Remarque

La fonction valeur table peut être utilisée dans un environnement DB2 réparti, uniquement si la table utilisateur est stockée dans un espace table ne possédant qu'un seul noeud. Vous devez également vous assurer que vous vous connectez au bon noeud en utilisant la variable d'environnement `DB2NODE`.

Syntaxe de la fonction

1. `db2ext.textsearch` sans prise en charge de `HIGHLIGHT`

```
db2ext.textSearch
(
  query           VARCHAR(4096),
  indexSchema     VARCHAR(128),
  indexName       VARCHAR(128),
  resultFirstRow  INTEGER,
  resultNumberRows INTEGER,
  primKeyBinding  <supported types>, // same type as primary key
)

return table
(
  primKey          <supported types>, // same type as primary key
  numberOfMatches  INTEGER,
  score            DOUBLE,
  totalNbResults   INTEGER
)
```

2. `db2ext.textsearch` avec prise en charge de `HIGHLIGHT`

```
db2ext.textSearch
(
  query           VARCHAR(4096),
  indexSchema     VARCHAR(128),
  indexName       VARCHAR(128),
  resultFirstRow  INTEGER,
  resultNumberRows INTEGER,
  primKeyBinding  <supported types>, // same type as primary key
  numberOfHits    INTEGER
)
```

```
return table
(
  primaryKey          <supported types>, // same type as primary key
  numberOfMatches    INTEGER,
  score              DOUBLE,
  totalNbResults     INTEGER
  hitInformation     BLOB(20K)
)
```

Paramètres de fonction

Les paramètres suivants sont les paramètres d'entrée :

query Pour plus d'informations, reportez-vous au Chapitre 14, «Syntaxe des arguments de recherche», à la page 141.

indexSchema, indexName

Identifie l'index dans lequel rechercher. Pour plus d'informations, reportez-vous à la section «CREATE INDEX» à la page 114.

resultFirstRow

La liste des résultats de la requête est renvoyée en plusieurs parties. Ce paramètre indique quelle est la ligne de la liste des résultats de la requête devant être saisie en premier dans la table de résultats de la fonction valeur table. La valeur doit être supérieure ou égale à 0.

Notez que le chiffre 0 identifie la première ligne dans la liste des résultats de la requête.

resultNumberRows

Ce paramètre indique le nombre de lignes de la liste des résultats de la requête entrées dans la table de résultats de la fonction valeur table et où 0 signifie que tous les résultats doivent être renvoyés.

Notez que ceci est différent du paramètre correspondant au nombre maximal de résultats qui détermine la taille maximale de la liste des résultats de la requête.

primaryKeyBinding

Le type de ce paramètre détermine le type du paramètre de sortie primaryKey. Si l'index de texte a été créé pour une table de base avec une clé primaire de type <type1>, primaryKeyBinding doit dans ce cas être également de type <type1>.

En outre, le paramètre détermine la portée de la recherche de texte. Si primaryKeyBinding a pour valeur NULL ("CAST(NULL as <type1>)", la portée de la recherche sera tous les documents stockés dans l'index. Il est également possible de limiter la recherche aux documents auxquels primaryKeyBinding est lié.

Par exemple, si primaryKeyBinding a pour valeur CAST(5 as BIGINT), la recherche est limitée à un seul document avec la valeur de clé primaire BIGINT égale à "5".

Notez que seules les clés primaires de colonnes uniques de types suivants sont prises en charge : SMALLINT, INTEGER, BIGINT, REAL, DOUBLE, VARCHAR FOR BIT DATA, DATE, TIME et TIMESTAMP.

numberOfhits

Cette option spécifie le nombre maximal d'informations sur les occurrences renvoyées par la fonction db2ext.textsearch. Si la valeur 0 est spécifiée,

les informations pour 1100 occurrences au maximum sont fournies. Ce processus peut prendre beaucoup de temps.

Notez que ce paramètre est nécessaire uniquement pour la construction d'informations HIGHLIGHT requises par la fonction `db2ext.highlight`.

Paramètres de fonction

Les valeurs renvoyées énoncées ci-dessous sont stockées dans une table temporaire qui doit être jointe à la table normale, pour le cas où d'autres résultats seraient demandés. Notez que les valeurs `NUMBEROFMATCHES`, `SCORE`, `TOTALNUMBEROFRESULTS` et `HITINFORMATION` ne sont calculées que si elles sont demandées dans l'instruction `select`.

primKey

Clé primaire du document trouvé.

numberofmatches

`NUMBEROFMATCHES` est un entier (`INTEGER`) indiquant le nombre d'occurrences trouvées pour chaque document.

score `SCORE` renvoie une valeur `DOUBLE`. Plus le terme recherché dans le document est fréquent, plus le score du document augmente.

totalNumberOfResults

La liste des résultats de la requête indique le nombre de résultats trouvés. Notez que chaque ligne a la même valeur.

Notez également que si vous utilisez `STOP SEARCH AFTER` ou `RESULT LIMIT` avec la syntaxe `SCORE` dans une requête, ce nombre n'est plus fiable.

hitInformation

Les informations sur les occurrences renvoyées par `db2ext.textsearch` sont nécessaires pour le traitement de la mise en évidence. Ce paramètre de sortie peut contenir couramment les informations relatives à approximativement 1100 occurrences. Si le nombre d'occurrences dépasse ce seuil, les informations relatives aux autres occurrences sont ignorées.

Notez que cette valeur est renvoyée uniquement si vous spécifiez la valeur de `numberOfHits`.

Utilisation

Avec la fonction de valeur table SQL, vous pouvez effectuer des recherches dans des vues, de la même façon qu'avec la recherche de procédure mémorisée. Cependant, aucune mémoire partagée n'est nécessaire, ainsi l'index n'a pas besoin d'être activé.

Cette fonction est principalement destinée aux utilisateurs ayant utilisé une requête SQL dans la recherche de procédure mémorisée. Cependant, uniquement une clé primaire de colonne seule est prise en charge sur des tables de base.

L'exemple suivant montre comment travailler dans une table de clé primaire à plusieurs colonnes :

```
select s.id from
db2ext.sample s, table (db2ext.textSearch(
    'characteristics',
    'DB2EXT',
    'COMMANDS',
```

Fonction de valeur table SQL

```
1,  
20,  
cast(NULL as INTEGER))) t  
where s.id = t.primkey
```

Dans cet exemple, vous devez d'abord créer une vue sur cette table avec une seule clé unique, puis créer l'index sur la vue.

Pour obtenir un exemple d'utilisation de la fonction de valeur table avec la fonction `db2ext.highlight`, reportez-vous à la page 158.

DB2EXT.HIGHLIGHT

La fonction `db2ext.highlight` permet d'obtenir des informations sur les raisons pour lesquelles un document a été qualifié de résultat de la recherche. Plus spécifiquement, elle peut être utilisée pour :

- obtenir les occurrences,
- obtenir les occurrences et le texte qui les entoure,
- obtenir le document avec les balises de mise en évidence définies par l'utilisateur et entourant les occurrences.

Notez que la fonction `db2ext.highlight` ne peut être utilisée qu'avec la fonction de valeur table `db2ext.textsearch`. La fonction de valeur table recherche dans l'index, fournissant les résultats à utiliser pour la fonction `HIGHLIGHT`.

Pour plus d'informations sur l'utilisation de la fonction `db2ext.textsearch`, reportez-vous à la section «DB2EXT.TEXTSEARCH» à la page 153.

Syntaxe de la fonction

```

➤—db2ext.highlight———→
➤—(—document-content—,—hit-information—,—hit-processing-information—)—→

```

Paramètres de fonction

Les paramètres suivants sont les paramètres d'entrée :

document content CLOB(100K)

Seuls les documents UTF8 de format TEXT ou XML sont pris en charge. Pour augmenter cette valeur, reportez-vous à la section «DB2EXTHL (utilitaire)» à la page 103.

hit information BLOB(20K)

Chaîne contenant les informations sur les occurrences. Elle est renvoyée par la fonction `db2ext.textsearch`, si le paramètre `numberOfHits` est spécifié.

hit processing information VARCHAR(1024)

Ce paramètre est une liste de paires de valeurs d'options, séparées par une virgule et chaque caractère de la chaîne étant placé entre guillemets. Il précise comment la mise en évidence doit être traitée pour le document spécifié. Si aucune des options n'est indiquée, le document original est renvoyé.

TAGS = ("STRING", "STRING")

Cette option permet à l'utilisateur de spécifier les balises à insérer avant et après une occurrence dans le document. Si cette option est omise, aucune balise n'est ajoutée ni avant, ni après une occurrence dans le document.

WINDOW_NUMBER = INTEGER

Cette option spécifie le nombre de parties (ou fenêtres) du document devant être renvoyées par la fonction de mise en évidence (`HIGHLIGHT`). Chaque fenêtre comporte une ou plusieurs occurrences et la première occurrence de chaque fenêtre détermine la partie du document renvoyé à l'utilisateur. Celles-ci peuvent être ou non entourées par du texte.

Fonction de mise en évidence SQL

Si cette option est omise, la valeur 0 est prise par défaut et le document entier contenant les balises de départ et de fin (si spécifié) est renvoyé. Dans ce cas, l'option WINDOW_SIZE est ignorée.

WINDOW_SIZE = INTEGER

Cette option spécifie la taille de fenêtre recommandée (en octets). La taille réelle peut varier, selon le nombre d'occurrences, la longueur des occurrences et les tailles de la balise de départ et de la balise de fin. Si l'option est omise, la valeur par défaut est 0. Ceci signifie que seules les occurrences sans texte autour seront renvoyées.

WINDOW_SEPARATOR = "STRING"

Cette option spécifie la balise utilisée pour séparer une fenêtre de la fenêtre suivante. Si l'option est omise, "..." est la valeur utilisée par défaut.

FORMAT = "STRING"

Cette option spécifie le format du document. Les valeurs valides sont XML ou TEXT. Si cette option est omise, TEXT est la valeur prise par défaut. Assurez-vous que la valeur du format est la même que celle spécifiée lors de l'indexation.

MODEL_NAME = "STRING"

Cette option spécifie le nom du modèle lié au document XML spécifié. Notez que si le FORMAT est TEXT, cette option a pour résultat une condition d'erreur.

SECTIONS = ("section-name1", ..., "section-nameN")

Pour les documents XML, la mise en évidence peut être limitée aux sections concernées. Par exemple, elles peuvent être définies dans le fichier modèle. Pour spécifier ces sections, séparez les noms de section par des virgules. Si cette option est omise, la mise en évidence est effectuée sur la totalité du document XML. Notez que si le FORMAT est TEXT, cette option est ignorée.

Paramètres de fonction

Les paramètres suivants sont les paramètres de renvoi :

CLOB(200K)

La fonction de mise en évidence HIGHLIGHT renvoie une valeur CLOB contenant les parties du document qu'elle a modifiées.

Utilisation

L'exemple suivant montre comment utiliser la fonction HIGHLIGHT :

```
select p.id,
       p.title,
       db2ext.highlight(p.content,
                        t.hitinformation,
                        'TAGS = (<bf>', '</bf>'),
                        WINDOW_NUMBER = 5,
                        WINDOW_SIZE = 200,
                        WINDOW_SEPARATOR = "...",
                        FORMAT = "XML",
                        SECTIONS = ("section1-name", "section2-name"))
FROM patent p, table (db2ext.textsearch(
    "relational database systems",
    'DB2EXT',
    'TI_FOR_CONTENT',
```

```
0,
20,
CAST(NULL as BIGINT),
15)) t
```

```
WHERE p.id = t.primkey
```

L'utilisation de documents de plus de 100 octets entraînera l'arrêt de la requête SQL et produira une erreur SQL (SQL 1476N et erreur sql -433). Pour éviter cela, utilisez la commande `db2exthl` pour augmenter la taille du contenu du document. Pour plus d'informations, reportez-vous à la section «DB2EXTHL (utilitaire)» à la page 103.

Remarque

Les caractères spéciaux comme "newline" sera renvoyé tel quel.

Restrictions

- Seuls les documents texte à plat ou XML sont pris en charge.
- Seules les bases de données UTF8 sont prises en charge. Pour des documents binaires ou DataLink, vous devez vous assurer que les documents sont en UTF8.
- Les documents Thaï ne sont pas pris en charge.
- Si le format de document utilisé pendant l'indexation ne correspond pas à celui utilisé au moment de la requête, la fonction HIGHLIGHT renverra des résultats imprévisibles.
- Seules les occurrences trouvées dans les parties de texte d'un document sont mises en évidence.
- La fonction HIGHLIGHT ne peut être utilisée qu'avec la fonction `db2ext.textsearch`.
- Les valeurs de chaîne ne peuvent pas contenir de guillemet (").

Chapitre 16. Fonction de recherche de procédure mémorisée

L'Extension Net Search fournit une recherche de procédure mémorisée pour renvoyer des tables de résultats prédéfinis. La table de résultats est spécifiée dans la section de la table de mémoire cache pendant la création de l'index. Utilisez la recherche de procédure mémorisée lorsque vous devez renvoyer un petit nombre de résultats dans un ordre spécifique.

Par exemple, une application Internet dans laquelle les 20 premières lignes sont renvoyées, mais où le reste des résultats peut être également renvoyé par paquets de 20 lignes.

Remarque

La fonction de procédure mémorisée peut être utilisée dans un environnement DB2 réparti, uniquement si la table utilisateur est stockée dans un espace table ne possédant qu'un seul noeud.

Vous devez également vous assurer que vous vous connectez au bon noeud en utilisant la variable d'environnement DB2NODE.

DB2EXT.TEXTSEARCH (pour une recherche de procédure mémorisée)**Syntaxe de la fonction**

```
db2ext.TextSearch(
    IN      query          VARCHAR(4096),
    IN      indexSchema    VARCHAR(128),
    IN      indexName      VARCHAR(128),
    IN      resultFirstRow INTEGER,
    IN      resultNumberRows INTEGER,
    IN      scoringFlag    INTEGER,
    IN      searchTermCountsFlag INTEGER,
    OUT     searchTermCounts VARCHAR(4096),
    OUT     totalNumberOfResults INTEGER )
```

Paramètres de fonction

Les paramètres suivants sont les paramètres d'entrée :

Query Pour plus d'informations, reportez-vous au Chapitre 14, «Syntaxe des arguments de recherche», à la page 141.

indexSchema, indexName

Identifient l'index dans lequel effectuer la recherche. Reportez-vous à la section «CREATE INDEX» à la page 114.

resultFirstrow

La liste des résultats de la requête est renvoyée en plusieurs parties. Le paramètre indique la ligne de la liste des résultats de la requête devant être placée en premier dans l'ensemble des résultats de la procédure mémorisée. La première ligne de la liste des résultats de la requête est identifiée par le chiffre 0.

resultNumberRows

Ce paramètre indique le nombre de lignes de la liste des résultats de la requête à placer dans l'ensemble des résultats de la procédure mémorisée.

Il ne doit pas être confondu avec l'expression "result limit" dans la requête qui détermine la taille maximale de la liste des résultats de la requête.

La valeur doit être supérieure ou égale à 0. 0 signifie que tous les résultats sont renvoyés.

Remarque

Si un ensemble de résultats plus grand est demandé, veillez à ce qu'un espace table utilisateur temporaire soit disponible. Si aucun espace table n'est disponible, créez-en un. L'exemple suivant crée un espace table sur une plateforme UNIX :

```
db2 "create user temporary tablespace tempts managed by system
    using ('/work/tempts.ts')"
```

scoringFlag

0 signifie qu'il n'y a pas de score et 1 qu'il y en a. Si l'évaluation par score est demandée, une colonne supplémentaire avec les valeurs de score sera renvoyée avec en premier lieu, la valeur la plus élevée.

searchTermCountsFlag

Ceci contrôle le traitement du nombre de termes recherchés. Si cet indicateur est égal à 0, le nombre n'est pas calculé.

Paramètres de fonction

Les paramètres suivants sont les paramètres de sortie :

searchTermCounts

Nombre d'occurrences de chaque requête de terme recherché dans l'index. Ces nombres sont renvoyés sous forme de liste, séparés par des espaces, dans l'ordre des termes recherchés dans la requête.

Pour plus d'informations, reportez-vous à **searchTermCountsFlag**.

totalNumberOfResults

Nombre total de résultats trouvés dans la liste des résultats de la requête.

Notez également que si vous utilisez STOP SEARCH AFTER ou RESULT LIMIT avec la syntaxe scoringFlag dans une requête, ce nombre n'est plus fiable.

Utilisation

Les colonnes de l'ensemble de résultats renvoyés par la procédure mémorisée sont données par l'option CACHE TABLE de la commande DB2TEXT CREATE INDEX. Si scoringFlag=1, alors une colonne de type DOUBLE est ajoutée. Celle-ci contient la valeur SCORE.

Les options suivantes permettent d'augmenter les performances d'une seconde requête avec la même chaîne que la première requête. Notez que ceci doit avoir lieu dans une autre fenêtre sans la valeur totalNumberOfResults requise :

- Si vous n'avez pas besoin de l'évaluation par score, ajoutez la syntaxe suivante : STOP SEARCH AFTER x DOCUMENTS, où x est la valeur de resultFirstRow + resultNumberRows.
- Si vous avez besoin de l'évaluation par score, ajoutez la syntaxe suivante : STOP SEARCH AFTER y DOCUMENTS, où y est égal à la valeur de totalNumberOfResults dans la première requête.

Pour être sûr que vous vous connectez au bon noeud pour la recherche, il peut être nécessaire de définir la variable d'environnement DB2NODE.

Pour UNIX, utilisez la commande suivante :

```
export DB2NODE=<no>
```

Notez qu'il est important que tous les noeuds physiques ait un temps synchronisé.

Pour Windows, utilisez :

```
set DB2NODE= <no>
```

Remarque

Un ID utilisateur isolé, différent de l'ID propriétaire d'instance ne fonctionne pas avec des bases de données partitionnées.

Chapitre 17. Prise en charge de documents structurés

Les documents structurés sont constitués de modèles de document et définitions de fichier document.

Comment un modèle de document décrit des documents structurés

Les documents au format HTML ou XML sont des exemples de documents structurés ; ils contiennent des balises qui identifient des zones de texte ou des attributs de document. Les zones de texte peuvent contenir des informations telles que le titre, l'auteur ou une description du document. Les attributs de documents peuvent contenir un nombre.

Vous trouverez ci-dessous un extrait d'un document en texte clair structuré. Il contient des zones et des attributs délimités par des balises de type HTML.

```
[head]Gestion des documents structurés
[/head]
```

```
[abstract]Ce document décrit le concept des documents structurés
et l'utilisation des modèles de document pour ...
[/abstract]
:
:
```

Lorsque l'Extension Net Search indexe des documents structurés, il doit reconnaître leur structure pour pouvoir indexer la zone de texte et les attributs, et les enregistrer sous un nom unique qui identifie la zone ou l'attribut qui les contient. Cela permet à l'Extension Net Search de faire une recherche sélective dans une zone de texte ou de rechercher des documents ayant un attribut particulier.

Pour que l'Extension Net Search puisse reconnaître la structure d'un format de document spécifique, vous devez lui en donner une définition dans un *modèle de document*.

Indiquez le nom du modèle de document en tant qu'argument lorsque vous appelez la commande CREATE INDEX pour indexer les documents.

Avant d'indexer des documents à l'aide d'un modèle de document, vous devez tout d'abord définir un modèle de document puis l'ajouter à l'index.

Exemple de modèle de document

Vous devez définir un modèle de document pour chaque format de document que vous souhaitez indexer. Voici un modèle de document simple pour les documents en texte clair structurés. Notez que, dans l'exemple, GPP signifie General Purpose Parser.

```
<?xml version="1.0"?>
<GPPModel>                                - Début du modèle de document GPP

  <GPPFieldDefinition                       - Début de définition de la zone
    name="Head"                            - Nom que vous donnez à la zone
    start="[head]"                         - Chaîne de limite au début de la zone
    end="[/head]"                          - Chaîne de limite à la fin de la zone
    exclude="YES" />
```

Exemple de modèle de document

```
<GPPFieldDefinition      - Début de définition de la zone
name="Abstract"
start="[abstract]"
end="[/abstract]"
exclude="NO" />
:
:
</GPPModel>
```

Les modèles de document sont écrits en langage XML, qui utilise des balises comme défini dans l'Annexe G, «Référence du modèle de document», à la page 229. Un modèle de document est composé de définitions de zones de texte et d'attributs. Cet exemple illustre uniquement les définitions de zones de texte spécifiées dans les éléments GPPFieldDefinition. De la même façon, vous pouvez utiliser GPPAttributeDefinition pour définir des attributs de document.

Dans l'exemple, la première ligne, `<?xml version="1.0"?>`, indique que le modèle de document est écrit à l'aide de balises XML. Chaque définition de zone de texte spécifie les chaînes de limite pour identifier le début et la fin de la définition de la zone dans le document source. Ainsi, chaque fois qu'un document contient la séquence de caractères `[head]` suivi d'une partie de texte puis de la séquence de caractères `[/head]`, le texte compris entre les chaînes de limite est considéré comme étant le contenu de la zone de texte identifiée par le nom `head`.

Vous affectez un nom de zone à chaque définition de zone. Ce nom constitue le moyen par lequel une requête peut limiter la recherche au contenu d'une zone de texte. Il peut être fixe ou provenir d'une règle définie à partir du contenu de l'unité structurelle. Dans ce cas, ce pourrait être, par exemple, le nom de la balise pour une entité XML, ou le nom d'un attribut XML.

Modèles de document

Un modèle de document contrôle principalement quelles sont les parties de la structure d'un document à indexer et comment elles le sont. Son objectif est :

- D'identifier les zones de texte devant être distinguées dans le document source.
- De déterminer le type d'une telle zone de texte.
- D'Affecter un nom de zone à la zone de texte.

Lorsque le modèle de document identifie le texte comme appartenant à une zone de texte, il est considéré comme faisant partie du contenu textuel du document et les termes sont extraits et stockés dans l'index.

Les éléments d'un modèle de document varient en fonction de l'analyseur syntaxique utilisé pour ce format de document :

- Pour le format HTML, un modèle de document utilise les noms de balise HTML pour définir quelles balises doivent être indexées et comment traiter les informations des balises meta.
- Pour le format XML, il n'existe pas d'ensemble de balises prédéfini, aussi un modèle de document doit d'abord définir quelles sont les balises présentant un intérêt. Les éléments XML de même nom peuvent être également distingués en fonction des autres éléments dans lesquels ils sont imbriqués.
- Pour le format GPP (general purpose parser), le modèle de document agit encore plus en profondeur avec l'analyseur syntaxique car il doit déterminer les limites des zones de texte. Dans ce cas, la définition de zone doit spécifier des chaînes pour détecter les limites des zones.

- Pour les formats Outside-In, un modèle de document utilise des balises similaires aux noms de balise HTML pour définir les balises à indexer et comment traiter les informations des balises meta. Notez que le format de filtrage Outside-In est aussi connu sous le nom d'INSO.

Pour plus d'informations, reportez-vous à la section appropriée "Définition d'un modèle de document".

Pour plus d'informations sur la syntaxe du modèle de document sous forme de DTD (Document Type Definition) et les informations des zones de texte, reportez-vous à l'Annexe G, «Référence du modèle de document», à la page 229.

Zones de texte et attributs de document

Un modèle de document vous permet d'identifier les parties ou sections de document qui soit appartiennent à une zone de texte spécifique, soit soit un attribut de document, soit les deux.

Zones de texte

Le texte d'un document est totalement indexé qu'il fasse ou non partie d'une zone de texte. Lorsque le modèle de document identifie le texte comme appartenant à une zone de texte, il est considéré comme faisant partie du contenu textuel du document et est indexé normalement. Autrement dit, les termes significatifs sont extraits et stockés dans l'index. Ceci signifie que les recherches de texte illimitées incluent une recherche dans ce texte.

Ainsi, vous pouvez faire une recherche sélective d'un texte dans une zone particulière. Vous pouvez, par exemple, rechercher des documents contenant le mot structure dans la zone de texte Abstract.

Une zone de texte peut apparaître plusieurs fois dans un même document. Vous pouvez, par exemple, définir une zone de texte contenant toutes les légendes des figures. Une zone de texte peut également chevaucher une autre zone de texte.

Si vous ne souhaitez pas indexer le contenu de certaines zones de texte, vous pouvez spécifier une définition de zone contenant `exclude="YES"`. Vous trouverez une liste des limites pour les zones de texte et les attributs à la section «Limites des zones de texte et des attributs de document» à la page 232.

Attributs de document

Les attributs de document contiennent des informations formatées, brèves de type nombre. Vous pouvez utiliser des fourchettes de valeur pour rechercher des documents contenant ces attributs.

Les attributs ne sont pas stockés avec le texte indexé, mais dans un index des éléments distinct. Ainsi, pour rechercher un document par le contenu d'un attribut, vous devez faire une recherche spécifique sur l'attribut.

Attributs de nombre : L'Extension Net Search fournit un analyseur qui reconnaît les nombres en virgule flottante. Vous trouverez ci-après quelques exemples de formats corrects et incorrects pour les valeurs INTEGER.

Tableau 6. Comportement des modèles de document par défaut pour les formats de document pris en charge

Format correct	Format incorrect
1000 1 000 1.000 - où le point est un caractère décimal	1,000
100 000 100 000.00123	1 000 000 - deux espaces entre 1 et 0

Notez que les espaces ne sont pas autorisés dans les fractions décimales d'un nombre. Exemple : 1 000.000 100 est traité comme deux nombres, 1000.000 et 100.

Les séparateurs de langue et les formats monétaires spécifiques ne sont pas pris en charge.

Modèles de document par défaut

Pour les documents HTML, XML et pour les documents filtrés Outside-In, l'Extension Net Search fournit des modèles de document par défaut qui sont utilisés si vous ne définissez pas de modèle de document. Pour les documents en texte en clair et structurés, vous devez fournir et spécifier un modèle de document.

Si vous utilisez l'un des modèles de document par défaut :

- Toutes les zones sont indexées et aucune information spéciale, telle qu'une information de balise META n'est extraite.
 - Pour les formats HTML et INSO, chaque zone reçoit le nom de la balise correspondante.
 - Pour le format XML, tous les noeuds XML d'un document XML sont mappés pour chevaucher des zones qui sont identifiées par les chemins d'accès complets à la balise des noeuds correspondants. Par exemple, le chemin d'accès /play/role/name.
- Aucun attribut numérique est indexé (du fait qu'aucun attribut numérique est défini dans le modèle de document par défaut).

Tableau 7. Comportement des modèles de document par défaut pour les formats de document pris en charge

Type de document	Comportement du modèle de document par défaut
HTML	Les zones suivantes sont acceptées comme zone de texte : <a> <address> <au> <author> <h1> <h2> <h3> <h4> <h5> <h6> <title>. Le nom de la zone est le nom de la balise, par exemple "address".
XML	Accepte toutes les balises comme zone de texte. Le nom de la zone est le chemin d'accès complet à la balise, par exemple "/play/title".
Texte en clair structuré (GPP)	Aucun modèle de document par défaut.
Outside-In (INSO)	Accepte comme zones de texte les propriétés du document illustrées à la section «Paramètres des éléments» à la page 178, telles qu'elles sont renvoyées par les filtres Outside-In. Le nom de la zone est celui de la propriété du document utilisé par Outside-In, par exemple : "SCCCA_TITLE".

Pour chaque type de document, un modèle de document est défini. Comme les modèles sont tous différents, un exemple et des explications sont fournis pour chacun d'eux.

Remarque

Bien que les modèles de document par défaut traitent correctement les documents, il est préférable de définir vos propres modèles de document pour obtenir de meilleures performances d'indexation et de recherche.

Avec le modèle de document par défaut, le texte d'un document est totalement indexé qu'il fasse ou non partie d'une zone de texte. Ceci signifie que les recherches de texte illimitées incluent une recherche dans ce texte.

Définition d'un modèle de document pour les documents en texte clair structurés

L'exemple suivant est celui d'un document en texte clair structuré GPP :

```
[head]Gestion des documents structurés
[/head]
[year]2002
[/year]
[abstract]Ce document décrit le concept des documents structurés
et l'utilisation des modèles de document pour ...
[/abstract]
```

Exemple de modèle de document GPP :

```
<?xml version="1.0"?>
<GPPModel>

  <GPPFieldDefinition
    name="Head"
    start="[head]"
    end="[/head]"
    exclude="YES" />

  <GPPFieldDefinition
    name="Abstract"
    start="[abstract]"
    end="[/abstract]"
    exclude="NO" />

  <GPPAttributeDefinition
    name="year"
    start="[year]"
    end="[/year]"
    type="NUMBER" />

</GPPModel>
```

- Début d'une zone de texte

- Fin d'une zone de texte

- Début d'un attribut de document

- Fin d'un attribut de document

La première ligne, `<?xml version="1.0"?>`, indique que le modèle de document est écrit à l'aide de balises XML. Notez que ce modèle n'est pas écrit pour les documents au format XML.

Chaque zone est définie dans une balise `GPPFieldDefinition` ou `GPPAttributeDefinition` qui contient les paramètres des éléments.

Toutes les définitions doivent être contenues dans la balise `<GPPModel>`.

Paramètres des éléments

Il s'agit des paramètres des éléments du modèle de document :

- name** Vous affectez un nom à la zone de texte ou à l'attribut de document pour chaque définition. Les noms permettent de limiter une requête de recherche au contenu d'une zone de texte spécifique ou d'un attribut de document. A l'aide des exemples ci-dessus, vous pouvez rechercher des documents contenant le mot `structure` dans la zone de texte appelée `Abstract`.
- start** Chaîne de limite en page de codes UTF-8 qui marque le début de la zone de texte ou de l'attribut de document. La spécification de ces chaînes ne suit aucune règle. Il peut s'agir de n'importe quelle chaîne arbitraire UTF-8. Exemples : `start="introduction:"`, `start="note!"`, `start="$$. "`.
- Les caractères non imprimables et les caractères XML spéciaux "<" et "&" doivent être spécifiés à l'aide du caractère d'échappement XML standard ("`<`;" pour "<" et "`&`;" pour "&").
- end** Facultatif. Chaîne de limite en page de codes UTF-8 qui marque la fin de la zone de texte ou de l'attribut de document. Si vous ne spécifiez pas de balise de fin, la prochaine balise de début trouvée est considérée comme la fin de la zone. Si aucune balise de début suivante n'est trouvée, la zone s'étend jusqu'à la fin du document et aucune autre zone n'est identifiée.
- type** Le type d'attribut de document doit toujours être "NUMBER". Ce paramètre ne s'applique pas aux définitions de zone.
- exclude** YES ou NO. Paramètre qui détermine si le texte d'une définition de zone doit être exclu et par conséquent, non indexé. Ce paramètre ne s'applique pas aux définitions d'attribut.
- Dans l'exemple, la définition de zone "head" serait exclue, mais la définition "abstract" serait incluse.

Restrictions :

- Deux définitions de zone ou définitions d'attribut ne peuvent pas avoir la même balise de début. Cependant, une définition de zone et une définition d'attribut peuvent avoir les mêmes balises de début et de fin.
- Une balise de début ne doit pas être le préfixe d'une autre. Par exemple, vous ne pouvez pas avoir une balise de début "author" et une balise de début "authority".
- Les balises de début et de fin ne doivent pas être des chaînes vides.

Pour plus d'informations sur les DTD (définitions de type de document), reportez-vous à la section «Définition de type de document pour des modèles de document» à la page 229.

Pour connaître les limites supplémentaires, reportez-vous à la section «Limites des zones de texte et des attributs de document» à la page 232.

Que se passe-t-il lorsqu'un document GPP est indexé ?

L'analyseur syntaxique GPP analyse le document à la recherche des chaînes de limite de début. Lorsqu'il trouve une chaîne de début, il analyse la zone qui suit jusqu'à ce qu'il trouve la chaîne de limite de fin correspondante.

Le contenu de la zone est alors indexé en fonction du terme de la définition, autrement dit, comme zone de texte ou attribut de document. Si la zone de texte et l'attribut de document ont les mêmes chaînes de limite de début et de fin, le contenu de la zone est indexé comme zone de texte et attribut de document.

Aucune imbrication de zones n'est permise. Si une nouvelle chaîne de limite de début est trouvée dans une zone avant qu'une chaîne de limite de fin ne soit atteinte, la nouvelle chaîne de limite de début est interprétée comme du texte normal.

Si aucune chaîne de limite de fin correspondante n'est trouvée, la zone est présumée s'étendre jusqu'à la fin du document et un code anomalie est signalé.

Si aucune chaîne de limite de fin n'est spécifiée dans le modèle de document, la nouvelle chaîne de limite de début signale la fin de la zone précédente.

Définition d'un modèle de document pour les documents HTML

L'analyseur syntaxique HTML convertit le texte en page de codes UTF-8. Il effectue une reconnaissance des balises HTML et les classifie en classes de balise :

- Informations référencées à ignorer, telles que les informations sur les polices.
- Balises qui fournissent des informations positionnelles, par exemple <p>; pour un nouveau paragraphe.
- Balises qui fournissent des informations structurelles, telle que <Title>

Il reconnaît toutes les références d'entité de caractère définies en HTML 4, par exemple "ä" (ä) et les résout en points de code correspondants de la page de codes UTF-8.

Il reconnaît les balises META et analyse le texte de ces balises.

Exemple de document HTML :

```
<HTML>
<HEAD>
<META NAME="year" CONTENT="2002">
<TITLE> The Firm </TITLE>
</HEAD>
<BODY>
<H1>Synopsis</H1>;
```

```
<H1>Prologue</H1>;:
:
</BODY>
```

Exemple de modèle de document HTML :

```
<?xml version="1.0"?>
<HTMLModel>
```

```
<HTMLFieldDefinition
name="subtitle"
tag="title"
exclude="YES" />
```

```
<HTMLFieldDefinition
name="header1"
tag="h1"
exclude="YES" />
```

- Début d'une zone de texte

- Fin de la zone de texte

Modèle de document

```
<HTMLAttributeDefinition      - Début de l'attribut de
name="year"                  document
tag="meta"
meta-qualifier="year"
type="NUMBER" />           - Fin de l'attribut de
                                document

</HTMLModel>
```

La première ligne, `<?xml version="1.0"?>`, indique que le modèle de document est écrit à l'aide de balises XML. Notez que ce modèle n'est pas écrit pour les documents au format XML.

Chaque zone est définie dans une balise `HTMLFieldDefinition` ou `HTMLAttributeDefinition` qui contient les paramètres des éléments.

Toutes les définitions de zone de texte doivent être contenues dans la balise `<HTMLModel>`.

Paramètres des éléments

Il s'agit des paramètres des éléments du modèle de document :

name	Vous affectez un nom à la zone de texte ou à l'attribut de document pour chaque définition. Les noms permettent de limiter une requête de recherche au contenu d'une zone de texte spécifique ou d'un attribut de document. A l'aide des exemples ci-dessus, vous pouvez rechercher des documents contenant le mot <code>firm</code> dans la zone de texte appelée <code>subtitle</code> .
tag	<p>Identifie un élément dont les balises de début et de fin (impliquée) marquent la zone de texte ou l'attribut de document. Le texte à l'intérieur d'un élément de ce nom constitue le contenu de la zone définie.</p> <p>La casse de la balise n'a pas d'importance.</p> <p>Dans les exemples ci-dessus, le texte suivant une balise <code>H1</code> est indexé comme faisant partie de la zone <code>"header1"</code>. Dans ce cas, <code>"synopsis"</code> et <code>"prologue"</code> seront indexés.</p>
meta-qualifier	<p>Cette balise doit être utilisée avec l'élément tag. En spécifiant <code>tag="meta"</code>, la valeur du contenu qui correspond au qualificatif <code>meta</code> est extraite.</p> <p>Dans l'exemple de document HTML, la balise <code>meta</code> dispose des éléments suivants :</p> <pre><META NAME="Author" CONTENT="J. Grisham"></pre> <p>Dans l'exemple de modèle de document, <code>meta-qualifier="author"</code>. De ce fait, le contenu <code>"J. Grisham"</code> est indexé comme valeur de l'attribut de chaîne <code>"author"</code>.</p>
type	Le type d'attribut de document doit être <code>"NUMBER"</code> . Ce paramètre ne s'applique pas aux définitions de zone.
exclude	YES ou NO. Paramètre qui détermine si le texte d'une définition de zone doit être exclu et par conséquent, non indexé. Ce paramètre ne s'applique pas aux définitions d'attribut.

Dans l'exemple, la définition de zone "header1" serait exclue, mais la définition "subtitle" serait incluse.

Tout autre texte d'un document est indexé, mais pas comme partie de zone.

Pour plus d'informations sur les DTD (définitions de type de document), reportez-vous à la section «Définition de type de document pour des modèles de document» à la page 229.

Pour connaître les limites, reportez-vous à la section «Limites des zones de texte et des attributs de document» à la page 232.

Définition d'un modèle de document pour les documents XML

Pour les documents XML, un modèle de document vous permet de définir comment un noeud d'un document XML est mappé vers une zone, un attribut de document ou les deux.

Exemple de document XML :

```
<?xml version="1.0"?>
<purchaseOrder orderDate="2001-01-20">           [4]
  <shipAddress countryCode="US">                  [1]
    <name>Alice Smith</name>                       [2]
    <street>123 Maple Street</street>
    <city>Mill Hill</city>
    <state>CA</state>
    <zip>90999</zip>
  </shipAddress>
  <item partNo="123" quantity="1">
    <name>S&B Lawnmower Type ABC-x</name>          [3]
    <price>239.90</price>
    <shipDate>2001-01-25</shipDate>
  </item>
  <item partNo="987" quantity="1">                 [3]
    <name>Multifunction Rake ZYX</name>
    <price>69.90</price>
    <shipDate>2001-01-24</shipDate>
  </item>
</purchaseOrder>
```

Exemple de modèle de document XML :

```
<?xml version="1.0"?>
<XMLModel>

<XMLFieldDefinition                               [1]
name="addresses"
locator="/purchaseOrder/shipAddress"

<XMLFieldDefinition                               [2]
name="customerName"
locator="//shipAddress/name"

<XMLAttributeDefinition                           [3]
name="partNumber"
type="NUMBER"
locator="/purchaseOrder//item/@partNo" />

<XMLFieldDefinition                               [4]
name="none"
```

Modèle de document

```
locator="/purchaseOrder/@orderDate"
exclude="yes"/>

</XMLModel>
```

La première ligne, `<?xml version="1.0"?>`, indique que le modèle est écrit en langage XML. Chaque zone est définie dans une balise `XMLFieldDefinition` ou `XMLAttributeDefinition` qui contient les paramètres des éléments.

Notez que toutes les définitions de zone de texte doivent être contenues dans la balise `<XMLModel>`. Pour connaître les limites, reportez-vous à la section «Limites des zones de texte et des attributs de document» à la page 232.

Paramètres des éléments

Il s'agit des paramètres des éléments du modèle de document :

name Vous affectez un nom à la zone de texte ou à l'attribut de document pour chaque définition. Ces noms permettent de limiter une requête de recherche au contenu d'une zone de texte spécifique ou d'un attribut de document.

Vous pouvez utiliser l'une des variables suivantes dans un nom. La variable est remplacée par une chaîne générée à partir de l'élément correspondant dans le document source.

Variable	Valeur
\$(NAME)	Nom qualifié réel (QName) de l'élément XML qui correspondait à XPath.
\$(LOCALNAME)	Nom local réel (sans préfixe) de l'élément XML qui correspondait à XPath.
\$(PATH)	Chemin d'accès absolu réel sous forme de séquence de barres obliques et de balises de l'élément XML qui correspondait à XPath.

type Le type d'attribut de document doit être "NUMBER". Ce paramètre ne s'applique pas aux définitions de zone.

locator

Expressions dans le langage XPath qui sélectionne les parties des documents source à utiliser comme zones de recherche.

Lors de l'écriture d'un fichier de modèle de document XML, les noms qualifiés, appelés QNames, dans un pointeur doivent être identiques aux balises du document XML, faute de quoi les zones ne seront pas reconnues et les requêtes sur les zones ne renverront pas de résultat.

Vous trouverez ci-après quelques exemples de pointeurs. Pour plus d'informations, reportez-vous à la syntaxe dans la section «Sémantique des expressions du pointeur (XPath)» à la page 230.

purchaseOrder	salesOrder
Tous les éléments purchaseOrder et salesOrder.	
shipAddress	Tous les éléments shipAddress.
*	Tous les éléments (abréviation de <code>child::*</code> – pour plus d'informations, voir la syntaxe).
name/item	Tous les éléments item ayant un parent name.

purchaseOrder//itemTous les éléments `item` ayant un ancêtre `purchaseOrder`.

/

Le noeud racine.

comment()

Tous les noeuds de commentaire.

processing-instruction()

Toutes les instructions de traitement.

attribute::* (ou @*)

Tous les noeuds d'attribut.

Un littéral est une chaîne placée entre guillemets simples ou doubles. Pour la définition exacte des "terminal tokens", reportez-vous aux recommandations XML.

Les pointeurs XPath sont similaires aux schémas XSLT (XML Stylesheet Language Transformation). Ils comprennent exactement le sous-ensemble de schémas XSLT ne contenant pas de prédicat, ni de fonction 'id' et 'key', ni de tests de noeud 'text()' et 'node()'.

ignore YES ou NO. Ce paramètre permet de constituer des exceptions au pointeur.

Vous pouvez spécifier un pointeur général, tel que *, pour correspondre aux noeuds à indexer. Mais vous pouvez également spécifier que certains noeuds correspondant à un pointeur plus spécifique ne doivent pas être indexés.

Pour formuler cela, vous devez inclure une définition de zone avec le pointeur le plus spécifique pour les noeuds à ignorer pendant l'indexation. Vous donnez alors à ce pointeur une priorité plus élevée que celle du pointeur général et spécifiez `ignore="yes"`. Ceci indique à l'indexeur qu'il ne doit pas générer d'informations de zone pour les noeuds correspondants.

Notez que lorsqu'un tel noeud ignoré est imbriqué dans un noeud générant une zone, le contenu du noeud ignoré est indexé car il appartient également au contenu du noeud générant la zone.

priority

Nombre réel compris entre -1 et +1 et qui spécifie la priorité à donner à une définition trouvée par un pointeur spécifique.

Si vous ne spécifiez pas de priorité, les priorités par défaut sont utilisées :

- Plusieurs alternatives séparées par | sont traitées comme un ensemble de définitions, un pour chaque alternative.
- Les pointeurs qui correspondent par un seul nom, autrement dit, les pointeurs de l'une des formes suivantes ont la priorité par défaut 0 :
 - `ChildOrAttributeAxisSpecifier QName`
 - `ChildOrAttributeAxisSpecifier processing-instruction(Literal)`
- Les pointeurs de la forme `ChildOrAttributeAxisSpecifier NCName:*` ont la priorité par défaut -0,25.
- D'autres pointeurs de la forme `ChildOrAttributeAxisSpecifier NodeTest` ont la priorité par défaut -0,5.
- Tous les autres pointeurs ont la priorité par défaut 0,5.

Notez que plus le pointeur est spécifique, plus la valeur de priorité par défaut est élevée. Par exemple, le pointeur non spécifique * donne une

faible priorité à la définition trouvée, alors qu'un nom est un pointeur plus spécifique et donne une priorité plus élevée.

Notez également que si à un noeud correspond plusieurs pointeurs, vous pouvez déterminer quelles définitions sont choisies en leur affectant des priorités. La définition ayant la priorité la plus élevée est choisie. Si deux définitions ont la même priorité, la plus récente est choisie.

Cette résolution de conflit est la même que celle utilisée en langage XSLT (XML Stylesheet Language Transformation).

exclude

YES ou NO. Paramètre qui détermine si le texte d'une définition de zone doit être exclu et par conséquent, non indexé. Ce paramètre ne s'applique pas aux définitions d'attribut.

Dans l'exemple, la définition de zone "customerName" serait exclue, mais la définition "addresses" serait incluse.

Restrictions :

- L'imbrication de zones est permise si, par exemple, l'emplacement Xpath d'une spécification sélectionne un noeud figurant dans un élément XML sélectionné par une autre définition d'attribut. Toutefois, du fait de la sémantique de l'expression du pointeur, il n'est pas possible de mapper le même noeud XML vers plusieurs zones.
- L'Extension Net Search ne tente pas de détecter la page de codes d'un document XML. Le CCSID spécifié pendant la commande CREATE INDEX ou s'il n'est pas spécifié, la page de codes DB2.

Que se passe-t-il lorsqu'un document XML est indexé ?

Le tableau ci-dessous indique ce qui est placé dans l'index.

Tableau 8. Entrées de l'index de texte

Informations sur les limites de zone	Texte indexé	
début de la zone "addresses" début de la zone "customerName"	Alice Smith	[1] et [2]
fin de la zone "customerName"	123 Maple Street Mill Hill CA 90999	[2]
fin de la zone "addresses"	123 1 S&B Lawnmower Type ABC-x 239.90 2001-01-25 987Z 1 Multifunction Rake ZYX 69.90 2001-01-24	
Nom d'attribut	valeur(s) d'attribut	
Numéro de composant	123, 987	

Notez que les zones imbriquées sont faisables, comme le montre l'exemple. La zone addresses sélectionne, dans le document XML, un noeud qui domine le noeud sélectionné par la zone customerName. Par conséquent, le contenu de ce noeud imbriqué appartient en toute logique aux deux zones. Bien que les zones de

texte puissent se chevaucher, le texte contenu dans ces zones n'est indexé qu'une seule fois. Dans cet exemple, lors de la recherche avec une limite de zone, Alice Smith est trouvé dans `addresses` ainsi que dans `customerName`.

Le contenu des zones est déterminé par les règles suivantes :

- Pour une zone dont le pointeur correspond à un commentaire, une instruction de traitement ou un attribut XML, le contenu de la zone est le texte réel du commentaire, de l'instruction de traitement ou de la valeur de l'attribut.
- Pour une zone correspondant à un élément XML ou au noeud racine, le contenu de la zone est constitué d'un texte de n'importe quel élément imbriqué, excepté pour les éléments correspondant aux zones qui ont pour spécification `exclude="YES"`.

Le document doit contenir un XML syntaxiquement correct, mais il n'est pas nécessaire qu'un DTD soit spécifié dans le document XML. Aucune validation DTD, ni de résolution d'entité externe n'est effectuée, l'Extension Net Search compare seulement le document XML au modèle de document. Les entités internes sont substituées comme requis par XML.

Pour plus d'informations sur les DTD (définitions de type de document), reportez-vous à la section «Définition de type de document pour des modèles de document» à la page 229.

Pour connaître les limites, reportez-vous à la section «Limites des zones de texte et des attributs de document» à la page 232.

Définition d'un modèle de document pour les documents filtrés Outside-In

Les modèles de document pour le format Outside-In sont très similaires aux modèles de document HTML, du fait qu'ils permettent de mapper les éléments structurels, identifiés par un ensemble de balises, vers des zones de texte et des attributs de document NSE. Présumons que vous disposez d'un ensemble de documents Microsoft Word et souhaitez indexer les propriétés du document "title", "subject" et "keyword" en tant que zones et les propriétés de document "author" et "category" en tant qu'attributs de document. L'exemple suivant pour un modèle de document Outside-In effectue ce mappage :

```
<?xml version="1.0"?>
<INSOModel>

<INSOFieldDefinition
name="title"
tag="SCCCA_TITLE"/>

<INSOFieldDefinition
name="title"
tag="SCCCA_SUBJECT"/>

<INSOFieldDefinition
name="title"
tag="SCCCA_KEYWORDS"/>

<INSOAttributeDefinition
name="author"
tag="SCCCA_AUTHOR"
type="STRING"/>

<INSOAttributeDefinition
```

```
name="category"  
tag="SCCA_CATEGORY"  
type="STRING"/>  
  
</INSOModel>
```

Paramètres des éléments

Il s'agit des paramètres des éléments du modèle de document :

- name** Nom que vous affectez à la zone de texte ou à l'attribut de document. Vous affectez un nom de zone à chaque définition de zone et un nom d'attribut à chaque définition d'attribut. Ces noms constituent le moyen par lequel une requête peut limiter la recherche au contenu d'une certaine zone de texte et peut rechercher des documents ayant un certain attribut.
- tag** Identifie une balise dont les éléments de début et de fin ou de fin impliquée marquent la zone de texte ou l'attribut de document. Le texte à l'intérieur d'un élément de ce nom constitue le contenu de la zone ou de l'attribut défini. La casse de la balise n'a pas d'importance. Les valeurs possibles sont décrites ci-dessous.
- type** Le type de l'attribut de document peut être "NUMBER", "DATE" ou "STRING". Ce paramètre ne s'applique pas aux définitions de zone.
- exclude** YES ou NO. Paramètre qui détermine si le texte d'une définition de zone doit être exclu et par conséquent, non indexé. Ce paramètre ne s'applique pas aux définitions d'attribut.

Les modèles de document Outside-In sont constitués de définitions de zone et/ou d'attribut qui définissent un nom (name) ou une balise (tag). Pour les définitions d'attribut, un type est également requis, alors que les définitions de zone comportent un indicateur facultatif "exclude". Comme avec les modèles HTML, l'attribut 'name' d'une telle définition définit le nom de la zone ou de l'attribut de l'Extension Net Search vers lequel la partie du document doit être mappée. Il peut s'agir d'une chaîne de texte UTF-8 arbitraire. Pour plus d'informations, reportez-vous à la spécification Outside In Content Access Specification, Version 7.5.

Pour obtenir la liste des valeurs possibles pour l'attribut de balise en rapport aux balises Outside-In de début, fin et propriété de document, reportez-vous à la section «Valeurs de l'attribut de marque Outside-In» à la page 232.

Que se passe-t-il lorsqu'un document Outside-In est indexé ?

Par défaut, tout le texte est indexé comme ne faisant pas partie d'une zone. Chaque fois qu'une balise de début apparaissant dans le flux du texte correspond à un élément de définition dans le modèle de document actif, le texte placé entre la balise de début et la balise de fin correspondante est traité en fonction du terme de cette définition. Par exemple, en tant que zone indexée, zone exclue, attribut ou les deux.

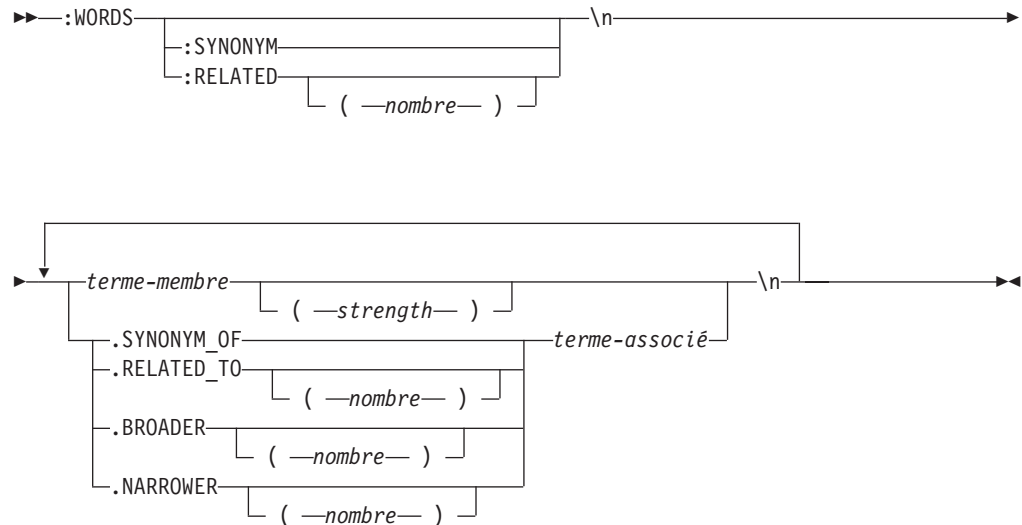
Si aucune définition correspondante n'existe, la balise de début et la balise de fin correspondante sont ignorées.

Comme les filtres Outside-In reconnaissent automatiquement le format et la page de codes du document, la spécification du CCSID est sans effet. Si les filtres Outside-In ne parviennent pas à déterminer le format et la page de codes corrects, le document est traité comme un fichier ASCII.

Chapitre 18. Prise en charge du thésaurus

La syntaxe ci-dessous est celle de chaque groupe de définitions.

Syntaxe d'une définition de thésaurus



Notez que \n ne fait pas partie de la syntaxe, mais représente la fin d'une ligne dans le fichier de définitions du thésaurus.

Vous pouvez insérer les lignes de commentaire dans un fichier de définitions de thésaurus comme suit :

mon texte de commentaire

:WORDS

Mot clé qui commence un groupe de mots associés.

:SYNONYM,

:RELATED [(nombre)],

Nom de relation.

Les noms de relation sont constitués d'un type de relation et d'un nombre. Si ce dernier est omis, la valeur 0 est utilisée, ce qui correspond au nom d'une relation fournie par le système. Le nom de relation fourni par le système est toujours :SYNONYM.

Les noms de relation commençant par deux points, tel que :SYNONYM, précèdent une liste de mots, liés les uns aux autres par la même relation. Par exemple :

```
:WORDS
:SYNONYM
  air steward
  cabin staff member
  flight attendant
```

terme-membre

Terme à inclure dans le dictionnaire thésaurus.

Prise en charge du thésaurus

- La longueur maximale est 64 octets (42 octets pour la page de codes UTF-8).
- Les caractères mono-octet et les caractères à deux octets de la même lettre sont considérés comme identiques.
- Aucune distinction n'est faite entre majuscules et minuscules.
- Un terme peut contenir un espace.
- Le caractère mono-octet, point "." ou deux points ":" ne peut pas être utilisé.

Ce paramètre peut être utile si vous ne souhaitez pas qu'une recherche de thésaurus inclut des mots ayant une faible relation avec le terme recherché. La valeur du paramètre Strength est comprise entre 1 et 100. La valeur par défaut est 100.

.SYNONYM_OF,
.RELATED_TO [(nombre)], .BROADER [(nombre)],
.NARROWER [(nombre)]

Nom de relation. Les noms de relation sont constitués d'un type de relation et d'un nombre. Si ce dernier est omis, la valeur 0 est utilisée, ce qui correspond au nom d'une relation fournie par le système. Le nom de relation fourni par le système est toujours **.SYNONYM**.

Les noms de relation qui comment par un point, tel que **.SYNONYM_OF**, définissent la relation entre un mot et un autre. Par exemple :

```
:WORDS
  air steward
  .SYNONYM_OF cabin staff member
  .SYNONYM_OF flight attendant
```

Le *nombre* facultatif identifie une relation définie par l'utilisateur. Il doit s'agir d'un nombre unique tiré du fichier complet de définitions du thésaurus (actuellement 1 à 128). Par exemple : **RELATED_TO(42)**.

Si vous souhaitez utiliser des noms symboliques pour les relations de thésaurus dans votre application au lieu du nom et du nombre, votre application doit traiter le mappage nom-vers-nombre. Par exemple, si vous définissez la relation *opposite_of* comme **RELATED_TO(1)**, votre application doit mapper ce nom vers le nom de relation interne **RELATED_TO(1)**.

terme-associé

Chaque terme associé doit être précédé par le nom de relation. Le terme associé est lié à chaque terme du membre en fonction de la relation spécifiée. Si tous les termes du membre sont liés les uns aux autres, ceci peut être spécifié à l'aide d'une relation de membre.

- La longueur maximale est 64 octets (42 octets pour la page de codes UTF-8).
- Les caractères mono-octet et les caractères à deux octets de la même lettre sont considérés comme identiques.
- Aucune distinction n'est faite entre majuscules et minuscules.
- Un terme peut contenir un espace.
- Le caractère mono-octet, point "." ou deux points ":" ne peut pas être utilisé.

Exemple de terme associé :

```
:WORDS:SYNONYM
  reject
  decline
  RELATED_TO(1) accept
```

Partie 3. Annexes

Annexe A. Migration

De nombreuses modifications ont été apportées à DB2 Extension Net Search Version 8.1.x, il inclut à présent l'interface de recherche et les fonctions de DB2 Text Information Extender Version 7.2. Par conséquent, il y a trois options de migration :

- Migration de Text Information Extender Version 7.2 vers l'Extension Net Search Version 8.1.x.

Si vous migrez de DB2 Version 7.2 et de Text Information Extender Version 7.2, vous devez commencer par migrer de l'instance DB2 et des bases de données d'instance. Consultez la documentation DB2 pour connaître la syntaxe et utiliser correctement les commandes DB2, telles que `db2imigr` que vous devez utiliser pour migrer l'instance DB2 de la version 7 à la version 8, et `db2 migrate db <nom de bd>` que vous devez utiliser pour migrer les bases de données d'instance.

- Migration de l'Extension Net Search Version 7.2 vers l'Extension Net Search Version 8.1.x
- Migration de l'Extension Net Search Version 8.1 vers l'Extension Net Search Version 8.1.x et 8.2

La procédure de migration de l'Extension DB2 Net Search est appelée `db2extmdb`, elle renvoie des messages en anglais. Si une erreur se produit au cours de la migration, corrigez-la puis relancez `db2extmdb`.

`db2extmdb` récupère toutes les informations sur l'administration `db2ext` pertinentes pour la migration dans une nouvelle table appelée `DB2EXT.TMIGRATION`. Dans cette table, chaque index de texte est représenté sous forme d'une seule ligne. Cette table d'informations relatives à la migration est conservée jusqu'à ce que la migration de la base de données soit effectuée. Vous ne pouvez pas supprimer cette table.

Lors de la migration des index de texte, le traitement diffère selon l'état de la table de journal et la disponibilité des fichiers index :

- Si la table de journal est vide et que l'on peut accéder aux fichiers index, c'est que la migration de l'index a été effectuée. Il s'agit de l'option la plus rapide.
- Si la table de journal n'est pas vide ou si les fichiers index sont introuvables, l'index sera recréé, un état cohérent avec la base de données ne pouvant être assuré.

Notez que cette migration peut prendre beaucoup de temps. Pendant l'exécution de la migration, aucune modification ne doit être apportée aux tables utilisateur.

Remarque

Pour obtenir les informations les plus récentes relatives à la migration, consultez le fichier `release.txt` sur le CD-ROM ainsi que le site Web DB2 Extension Net Search.

Migration de l'Extension Net Search Version 8.1 vers l'Extension Net Search Version 8.1.x et 8.2

La migration de l'Extension Net Search Version 8.1 vers l'Extension Net Search Version 8.1.x ne nécessite aucune étape préalable. Vous pouvez donc supprimer votre ancienne installation et installer la dernière version de Extension Net Search.

Migrez votre instance DB2 de la version 8.1 à la version 8.1.x ou 8.2 en utilisant db2iupdt pour UNIX et Windows. Migrez ensuite toutes les bases de données activées de cette instance à l'aide du nouvel outil db2extmdb. Le demandeur doit être le propriétaire d'instance et la syntaxe est la suivante :

db2extmdb <nom de la base de données>

Pendant l'exécution de la migration, aucune modification ne doit être apportée aux tables utilisateur avec les index de texte.

Remarques et recommandations

Les étapes de la migration sont journalisées dans le fichier suivant :

<os-dependent>/sql1ib/db2ext/db2extm <nom-bdd>.log

Avant d'appeler le programme db2extmdb, l'utilisateur doit effectuer une sauvegarde de tous les répertoires d'index et de la base de données.

Migration de l'Extension Net Search Version 7.2 vers l'Extension Net Search Version 8.1.x

Si vous utilisez Extension Net Search Version 7.2 et que vous n'avez pas besoin de toutes les nouvelles options, continuez simplement à utiliser les anciennes interfaces en installant les interfaces de compatibilité de l'Extension Net Search Version 7.2 disponibles sur le CD-ROM ou par téléchargement. Reportez-vous au fichier release.txt pour de plus amples informations. Pour utiliser la nouvelle fonction Extension Net Search, vous devez changer vos scripts d'administration de sorte à ce qu'ils correspondent à la syntaxe d'administration modifiée ainsi que changer votre syntaxe de recherche dans les nouvelles interfaces de recherche.

Notez qu'il n'existe pas de migration automatique de Extension Net Search Version 7.2 à Extension Net Search Version 8.1.x.

Notez également que ces interfaces de compatibilité sont déconseillées et qu'elles ne seront pas disponibles dans les versions à venir.

Migration de Text Information Extender Version 7.2 vers l'Extension Net Search Version 8.1.x

Migrez votre instance DB2 de la version 7.2 à la version 8.1.x en utilisant, dans un premier temps, l'outil DB2 db2imgr puis db2iupdt, pour UNIX et Windows. Par exemple :

```
db2imgr -u <idutilisateur isolé> <nom instance>
db2iupdt -u <idutilisateur isolé> <nom instance>
```

Migrez ensuite toutes les bases de données de cette instance à l'aide de la commande DB2 migrate et du nouvel outil db2extmdb. Par exemple :


```
db2 migrate database <nom de la base de données>
db2extmdb <nom de la base de données>
```

Le demandeur doit être le propriétaire de l'instance.

Le programme récupère toutes les informations sur l'administration db2ext pertinentes pour la migration dans une nouvelle table, appelée DB2EXT.TMIGRATION. Dans cette table, chaque index de texte est représenté sous forme d'une seule ligne. La table d'informations sur la migration continuera d'exister jusqu'à ce que la migration de la base de données aboutisse, et l'utilisateur ne devra pas la supprimer.

Si une erreur se produit, corrigez-la puis relancez db2extmdb.

Lors de la migration de l'index de texte, un autre traitement a lieu. Cela dépend de l'état de la 'table de journalisation'. Si la table de journalisation est vide, l'index sera migré ; il s'agit du cas le plus rapide. Si la table de journalisation n'est pas vide, un état cohérent avec la base de données ne pourra pas être assuré et l'index doit être recréé. Ce processus peut prendre beaucoup de temps.

Pendant l'exécution de la migration, aucune modification ne doit être apportée aux tables utilisateur avec les index de texte.

Recommandation

Avant d'appeler le programme db2extmdb, l'utilisateur doit effectuer une sauvegarde de tous les répertoires d'index et de la base de données, ainsi que vérifier que tous les fichiers de modèles de documents utilisés pour la création des index de texte version 7.2 existent toujours et qu'ils sont accessibles à la lecture.

Une fois la migration de l'instance DB2 effectuée, le répertoire précédent sqllib est renommé sqllib_v71. Si les index de texte sont stockés dans le répertoire d'index par défaut, déplacez le répertoire sqllib_v71/db2ext/indexes vers le nouveau répertoire sqllib, sqllib/db2ext/indexes.

Remarques et recommandations

Les étapes de la migration sont journalisées dans le fichier suivant :

```
<os-dependent>/sqllib/db2ext/db2extm <nom-bdd>.log
```

Pour HP-UX ou Linux, cette étape n'est pas nécessaire comme Text Information Extender n'est pas disponible sur ces plateformes.

Veuillez également noter que DB2 Text Information Extender Version 7.2 n'est plus disponible.

Annexe B. Utilisation de grandes quantités de mémoire

L'utilisation de la mémoire cache pour une recherche de procédure mémorisée nécessite une grande quantité de mémoire et différentes configurations de mémoire requises pour les plateformes suivantes :

- AIX
- Windows
- Environnement d'exploitation Solaris
- Linux
- HP-UX

AIX (32 bits et 64 bits)

Configuration des limites du système :

- Vérifiez les limites du système en utilisant la commande `ulimit -a`
- S'il y a des valeurs différentes de "aucune limite", procédez comme suit :
 - Connectez-vous en tant qu'utilisateur root.
 - Sauvegardez le fichier `/etc/security/limits`, puis éditez le fichier pour augmenter le plafond ferme.
 - Réglez toutes les valeurs sur "aucune limite" (valeur -1) pour le propriétaire d'instance DB2 utilisé.

Configuration des limites de mémoire partagée :

- Sur AIX, vous n'avez pas besoin de configurer les limites de la mémoire partagée.

Configuration de l'espace de pagination :

- Vous obtiendrez la taille de la mémoire RAM du système en utilisant la commande `lsattr -E -l sys0`
- Vous obtiendrez la taille de l'espace de pagination en utilisant la commande `lsp -a`.
- Réglez la taille de l'espace de pagination sur au moins 1,5 - 2 fois la taille de la mémoire RAM de votre système ou utilisez le paramètre `MAXIMUM CACHE SIZE` que vous fournissez dans la commande `CREATE INDEX`. Utilisez l'utilitaire `SMIT` pour sélectionner un nombre plus élevé.

Notez que la limite de la taille de la mémoire cache maximale sur AIX est de 1536 Mo (1,5 Go = 1610612736 octets).

Sur AIX (32-bits), si vous souhaitez rechercher des tables volumineuses de la mémoire cache en utilisant l'interface de recherche de procédure mémorisée, il vous faut remplacer le paramètre `MAXDATA` par l'exécutable `db2fmp`. Pour ce faire, connectez-vous en tant que propriétaire de l'instance et exécutez les commandes suivantes :

```
db2stop
cd ~/sqllib/adm
cp db2fmp db2fmp.org /usr/bin/echo '\0040\0\0\0' | dd of=db2fmp bs=4 count=1
seek=19 conv=notrunc
```

Si vous n'avez toujours pas assez de mémoire, réduisez encore la quantité de segments de mémoire comme suit :

Utilisation de grandes quantités de mémoire

```
cp db2fmp db2fmp.org /usr/bin/echo '\0000 \0 \0 \0' | dd of=db2fmp bs=4 count=1  
seek=19 conv=notrunc
```

Le paramètre maxdata sera ramené à 1 (256 Mo).

Si vous souhaitez revenir aux paramètres d'origine, exécutez les commandes suivantes :

```
db2stop  
cd ~/sqllib/adm  
cp db2fmp.org db2fmp
```

Notez que cette modification ramène la taille de segment maximale de 1,2 à 0,5 Go. Cela peut affecter d'autres applications utilisant l'instance DB2.

Windows (32 bits)

Réglage de la taille du fichier de pagination :

- Réglez la taille du fichier de pagination de la mémoire virtuelle Windows sur au moins 1,5 - 2 fois la taille de la mémoire RAM de votre système ou utilisez le paramètre MAXIMUM CACHE SIZE que vous fournissez dans la commande CREATE INDEX. Sélectionnez un nombre plus élevé. Reportez-vous à la documentation Windows pour plus d'informations relatives au changement de taille du fichier de pagination.

Notez que la limite de la taille de la mémoire cache maximale sur Windows est de 1000 Mo (1 Go = 1073741824 octets).

L'environnement d'exploitation Solaris (32 bits et 64 bits)

Configuration des limites du système :

- Configurez les limites du système en utilisant la commande : `ulimit -a`
- Exécutez les étapes suivantes :
 - Connectez-vous en tant qu'utilisateur root.
 - Sauvegardez le fichier `/etc/system`, puis éditez le fichier pour augmenter le plafond ferme.
 - Ajoutez ou vérifiez que les lignes suivantes sont réglées sur les valeurs affichées au moins :

```
rlim_fd_cur -> Par défaut 64, recommandé >= 1024  
rlim_fd_cur_max -> Par défaut 1024, recommandé >= 4096
```

Configuration des limites de mémoire partagée :

- Vérifiez les paramètres en cours en utilisant la commande `sysdef -i`
- Editez le fichier `/etc/system` pour définir la limite de la taille de la mémoire en utilisant : `set shmsys:shminfo_shmmax=0xffffffff`
Vous devrez éventuellement augmenter les valeurs des paramètres suivants :

```
set shmsys:shminfo_shmmni=512  
set shmsys:shminfo_shmseg=128,
```

redémarrez ensuite le système.

Configuration de l'espace de pagination :

- Vous obtiendrez la taille de la mémoire RAM du système en utilisant la commande `/usr/sbin/prtconf`
- Vous obtiendrez la taille de l'espace de pagination en utilisant la commande `swap -l`.

- Réglez la taille de l'espace de pagination sur au moins 1,5 - 2 fois la taille de la mémoire RAM de votre système ou utilisez le paramètre MAXIMUM CACHE SIZE que vous fournissez dans la commande CREATE INDEX. Sélectionnez un nombre plus élevé.

Reportez-vous à la documentation Solaris pour davantage d'informations sur comment ajouter de l'espace de pagination.

Notez que la limite de la taille de la mémoire cache maximale sur Solaris est de 2000 Mo (2 Go = 2147483647 octets).

Linux

Recherchez les paramètres du noyau recommandés sous Linux dans le manuel de mise en route DB2.

L'état de la validation des nouveaux noyaux et distributions Linux est fréquemment mis à jour. Pour obtenir les toutes dernières informations relatives aux niveaux du logiciel Linux pris en charge, reportez-vous au site : <http://www.ibm.com/software/data/db2/linux/validate>

Pour voir les limites de vos ressources partagées courantes, utilisez `ipcs -l`. Pour vérifier les limites du système, utilisez la commande `ulimit -a`.

HP-UX (32 bits et 64 bits)

Recherchez les paramètres du noyau recommandés sous HP-UX dans le manuel de mise en route DB2.

Annexe C. Catalogues d'informations DB2 Extension Net Search

DB2 Extension Net Search stocke d'importantes informations relatives aux valeurs par défaut, aux configurations, aux index de texte et aux formats dans les tables du catalogue. Pour visualiser ces informations, vous pouvez interroger des vues sur les tables.

Les vues et tables suivantes reflètent la configuration actuelle de votre système :

- Vues pour informations de niveau de base de données :
 - db2ext.dbdefaults
 - table db2ext.proxyinformation
- Vues pour informations de niveau d'index :
 - db2ext.textindexes
 - db2ext.textindexformats
 - db2ext.indexconfiguration

Notez que pour des raisons de compatibilité, les vues DB2 Text Information Extender sont toujours disponibles : db2ext.textcolumns, db2ext.models et db2ext.formats.

- Vues de table pour un index de texte :
 - Vue d'événements
 - Vue de table de journalisation

Vues pour informations de niveau de base de données

La vue db2ext.dbdefaults affiche toutes les valeurs par défaut pour la base de données.

Les valeurs par défaut sur le niveau de base de données ne peuvent pas être modifiées et sont disponibles comme paires attribut-valeur dans cette vue :
db2ext.dbdefaults

```
db2 select DEFAULTNAME, DEFAULTVALUE from DB2EXT.DBDEFAULTS
```

Tableau 9. vue db2ext.dbdefaults

Attribut	Valeur par défaut	Remarques
CCSID	ID de jeu de caractères codés (CCSID) de base de données	ID de jeu de caractères codés (CCSID) par défaut pour documents. Ceci est applicable si aucun ID de jeu de caractères codés n'est spécifié dans la commande CREATE INDEX.
FORMAT	TEXT	Format par défaut du document. Ceci est applicable si aucun format n'est spécifié dans la commande CREATE INDEX.

Vues pour informations de niveau de base de données

Tableau 9. vue db2ext.dbdefaults (suite)

Attribut	Valeur par défaut	Remarques
INDEXDIRECTORY	Consultez le nom du chemin d'accès sous Remarques	Répertoire pour fichiers index en texte intégral. Ceci est applicable si aucun répertoire d'index n'est spécifié dans la commande CREATE INDEX. Le nom du chemin d'accès est : \$DB2EXT_INSTOWNERHOMEDIR/sqllib/db2ext/indexes
LANGUAGE	FR_FR	La langue du document.
MODELCCSID	ID de jeu de caractères codés (CCSID) de base de données	ID de jeu de caractères codés (CCSID) de fichiers de modèles de documents.
UPDATECOMMITCOUNT	0	Nombre de modifications traitées dans une transaction lors d'une mise à jour.
CLEARCOMMITCOUNT	0	Nombre de modifications traitées dans une transaction lors d'une commande CLEAR INDEX.
UPDATEFREQUENCY	NONE	Quand vérifier les mises à jour dans les nouveaux index.
UPDATEMINIMUM	1	Nombre minimal de modifications avant l'exécution de la mise à jour.
WORKDIRECTORY	Consultez le nom du chemin d'accès sous Remarques	Répertoire pour les fichiers temporaires d'index. Le nom du chemin d'accès est : <os_dependent>/sqllib/db2ext/ indexes
CACHEDIRECTORY	Consultez le nom du chemin d'accès sous Remarques	Répertoire par défaut pour l'option PERSISTENT CACHE de la commande CREATE INDEX. Le nom du chemin d'accès est : <os_dependent>/sqllib/db2ext/memory
PCTFREE	50	Pourcentage de mémoire cache libre restante pour de nouvelles insertions.
USERPERSISTENTCACHE	1	Utilisez la mémoire cache persistante.
AUTOMATICREORG	1	Option REORGANIZE dans la commande CREATE INDEX. Cela implique une réorganisation automatique.
TREATNUMBERSASWORDS	0	N'interprétez pas les séquences comme des mots séparés, même s'ils sont des caractères adjacents. Par exemple, la valeur par défaut 0 signifie que tea42at5 est considéré comme un seul mot.
INDEXSTOPWORDS	1	Ignorez les mots vides lors de l'indexation.
VERSION		NSE V8.1.4 Numéro de version en cours de L'Extension Net Search.

Tableau 9. vue db2ext.dbdefaults (suite)

Attribut	Valeur par défaut	Remarques
UPDATEDELAY	0	Seules les entrées du journal sont utilisées pour les mises à jour incrémentielles à l'aide d'un horodatage. Si l'entrée du journal est antérieure à l'horodatage en cours, il s'agit d'un Updatedelay. Ce paramètre doit seulement être utilisé pour de longues transactions en cours sur la table utilisateur lors d'une commande de mise à jour pour éviter les mises à jour perdues. Notez qu'avec les entrées du journal et les mises à jour incrémentielles, aucune table de capture n'est utilisée.

Table db2ext.proxyinformation

Si vous utilisez des liaisons de données et que vous voulez accéder à des fichiers à l'aide d'un serveur proxy, vous devez spécifier des informations sur le proxy dans la table db2ext.proxyinformation.

Tableau 10. vue db2ext.proxyinformation

Attribut	Type	Remarques
PROXYHOST	VARCHAR(254)	Nom d'hôte sur le serveur proxy.
PROXYPORT	VARCHAR(6)	Port utilisé pour le serveur proxy.
PROXYTYPE	VARCHAR(10)	Type de serveur proxy (PROXY ou SOCKS).
PROXYTIMEOUT	INTEGER	Délai d'attente en secondes.

Vous pouvez insérer une ligne au maximum. Si vous avez un serveur proxy sur le port 123, vous pouvez insérer une ligne avec l'instruction SQL suivante :

```
db2 insert into db2ext.proxyinformation values ('proxy1', '123', 'PROXY', 10)
```

Vues pour les informations de niveau d'index

Vous pouvez demander des informations à un niveau d'index en utilisant les vues DB2 Extension Net Search :

- db2ext.textindexes
- db2ext.textindexformats
- db2ext.indexconfiguration
- <schéma du nom de la vue d'événements d'index>.<nom de la vue d'événements d'index>

Pour des raisons de compatibilité amont, les vues DB2 Text Information Extender db2ext.textcolumns, db2ext.formats et db2ext.models sont toujours prises en charge mais déconseillées.

Notez que dans la vue db2ext.textcolumns, les colonnes OPERATION, OPERATIONBEGIN et OPERATIONEND ne sont plus prises en charge.

Vue db2ext.textindexes

Chaque base de données activée pour DB2 Extension Net Search comporte une vue db2ext.textindexes. Elle contient des informations sur les paramètres, les statistiques et les valeurs par défaut pour les index de texte créés dans cette base de données.

Lorsque vous créez un index de texte, de nouvelles entrées sont créées dans db2ext.textindexes. Lorsque vous supprimez les index de texte, ces entrées sont supprimées.

Vous pouvez interroger la vue pour obtenir des informations sur les index. Voici un exemple utilisant le schéma d'index :

```
db2 "select COLNAME from DB2EXT.TEXTINDEXES where INDSHEMA='myschema'
and INDNAME='myindex'"
```

Notez cependant que vous ne pouvez pas modifier la vue avec des commandes de manipulation de données SQL normales, ni créer explicitement ni supprimer la vue de catalogue. Le tableau suivant présente les autres éléments de la vue.

Notez également que les paramètres de réplication ne sont pas inclus dans cette vue.

Tableau 11. vue db2ext.textindexes

Attribut	Type	Remarques
INDSCHEMA	VARCHAR(128)	Nom de schéma de l'index de texte.
INDNAME	VARCHAR(128)	Nom de l'index de texte.
TABSCHEMA	VARCHAR(128)	Nom de table du schéma pour les tables de base, les alias de type nickname et les vues.
TABNAME	VARCHAR(128)	Nom d'alias sur lequel l'index a été créé.
COLNAME	VARCHAR(128)	Colonne sur laquelle l'index a été créé.
CCSID	INTEGER	ID de jeu de caractères codés de document pour cet index.
LANGUAGE	VARCHAR(5)	Langue du document pour cet index.
FUNCTIONSCHEMA	VARCHAR(128)	Schéma de la fonction de mappage de colonne.
FUNCTIONNAME	VARCHAR(18)	Nom de la fonction de mappage de la colonne.
INDEXDIRECTORY	VARCHAR(256)	Répertoire pour fichiers index en texte intégral.
WORKDIRECTORY	VARCHAR(256)	Répertoire pour les fichiers temporaires d'index.
CACHEDIRECTORY	VARCHAR(256)	Répertoire pour la mémoire cache persistante (si persistentcache=1).
UPDATEFREQUENCY	VARCHAR(300)	Critère de déclenchement pour l'application de mises à jour automatiques à cet index.
UPDATEMINIMUM	INTEGER	Nombre minimal de documents devant être modifié avant l'exécution d'une mise à jour.

Vues pour les informations de niveau d'index

Tableau 11. vue db2ext.textindexes (suite)

Attribut	Type	Remarques
EVENTVIEWSHEMA	VARCHAR(128)	Schéma de la table d'événement créée pour cet index.
EVENTVIEWNAME	VARCHAR(128)	Nom de la vue d'événements créée pour cet index.
LOGVIEWSHEMA	VARCHAR(128)	Schéma de la vue journal créée pour un index.
LOGVIEWNAME	VARCHAR(128)	Nom de la vue journal créée pour un index (important pour la mise à jour incrémentielle sur les vues).
COMMITCOUNT	INTEGER	Valeurs par défaut pour la mise à jour du nombre de validations.
NUMBER_DOCS	INTEGER	Nombre total de documents en cours dans l'index. Notez que lors d'une mise à jour de l'index, la valeur est uniquement mise à jour si le nombre de validations est défini.
REORG_SUGGESTED	INTEGER	Indique si les performances peuvent être optimisées en exécutant UPDATE INDEX REORGANIZE. Ce paramètre est seulement vrai (1) si au moins l'un des noeuds dispose d'une réorganisation d'index suggérée.
REORGAUTOMATIC	INTEGER	1, si l'index est automatiquement réorganisé pendant l'opération de mise à jour.
RECREATEONUPDATE	INTEGER	1, si l'index est automatiquement réorganisé pendant l'opération de mise à jour.
CREATIONTIME	TIMESTAMP	Heure de la création d'index.
UPDATETIME	TIMESTAMP	Heure de la dernière mise à jour. Si UPDATE TIME est égal à CREATION TIME alors aucune mise à jour n'a été effectuée.
PERSISTENTCACHE	INTEGER	1, si la mémoire cache persistante est utilisée.
MAXIMUMCACHESIZE	INTEGER	Taille maximale de la mémoire cache.
PCTFREE	INTEGER	Pourcentage de mémoire cache libre restante pour de nouvelles insertions.
CACHETABLE	VARCHAR(32000)	Liste des expressions de la colonne pour CACHE TABLE.
RESULTORDER	VARCHAR(32000)	Par ordre SQL pour INITIAL RESULT ORDER.
ATTRIBUTES	VARCHAR(32000)	Liste des expressions de la colonne pour ATTRIBUTES.
VIEWKEYCOLUMNS	VARCHAR(32000)	Colonnes clés pour index sur la vue.

Vue db2ext.indexconfiguration

Les paramètres de configuration d'index sont disponibles dans la vue db2ext.indexconfiguration. La vue est disponible par le biais de fonctions de requêtes SQL normales. Voici un exemple utilisant le nom d'index :

```
db2 "select VALUE from DB2EXT.INDEXCONFIGURATION where INDSHEMA='myschema'
and INDNAME='myindex' and PARAMETER = 'INDEXSTOPWORDS'"
```

Les tableaux suivants présentent les autres éléments de la vue.

Tableau 12. vue db2ext.indexconfiguration

Attribut	Type	Remarques
INDSCHEMA	VARCHAR(128)	Nom de schéma de l'index.
INDNAME	VARCHAR(128)	Nom de l'index.
PARAMETER	VARCHAR(30)	Type de paramètre.
VALUE	VARCHAR(512)	Valeur du paramètre.

Pour les attributs PARAMETER et VALUE, plusieurs valeurs sont disponibles.

Tableau 13. vue db2ext.indexconfiguration

Attribut et valeurs	Attribut et valeurs
PARAMETER	VALUE
- TREATNUMBERSASWORDS	- 0 ou 1
- INDEXSTOPWORDS	- 0 ou 1
- UPDATEDELAY	- secondes >= 0

Pour plus d'informations, consultez l'option CONFIGURATION de la commande CREATE INDEX.

Vue db2ext.textindexformats

Des informations sur le format et le modèle pour les index sont disponibles dans la vue db2ext.textindexformats. Voici un exemple utilisant le nom d'index :

```
db2 "select FORMAT from DB2EXT.TEXTINDEXFORMATS where INDSHEMA='myschema'
and INDNAME='myindex'"
```

Le tableau suivant présente les autres éléments de la vue.

Tableau 14. vue db2ext.textindexformats

Attribut	Type	Remarques
INDSCHEMA	VARCHAR(128)	Nom de schéma pour l'index (utilisé comme préfixe pour nom-table et nom-schéma dans la table de journalisation).
INDNAME	VARCHAR(128)	Nom d'index spécifié dans la commande CREATE INDEX.
FORMAT	VARCHAR(30)	Le modèle est lié à ce format.
MODELNAME	VARCHAR(30)	Nom d'un modèle de document.
MODELFILE	VARCHAR(256)	Fichier contenant la définition du modèle.

Tableau 14. vue db2ext.textindexformats (suite)

Attribut	Type	Remarques
MODELCCSID	INTEGER	ID de jeu de caractères codés de MODELFILE.
DEFAULT	INTEGER	Actuellement 1, comme de nombreux formats dans un index ne sont pas pris en charge.

Vues de table pour un index de texte

Vous pouvez demander des informations à un niveau d'index en utilisant ces vues DB2 Extension Net Search :

- Vue d'événements
- Vue de table de journalisation

Vue d'événements

Cette vue vous permet d'avoir des informations sur l'état d'indexation, les événements d'erreurs et lorsque des problèmes surviennent lors de l'indexation, par exemple, lorsqu'un document est introuvable. Ces événements de mise à jour d'index sont alors écrits dans la table des événements de l'index.

Le schéma et le nom sont stockés dans la vue db2ext.textindexes. Pour obtenir le nom d'une vue d'événements, utilisez l'exemple suivant :

```
db2 "select EVENTVIEWSHEMA, EVENTVIEWNAME from DB2EXT.TEXTINDEXES
      where INDSHEMA = 'myschema' and INDNAME = 'myindex'
```

La vue d'événements d'un index comporte les colonnes suivantes.

Tableau 15. La vue d'événements

Attribut	Type	Remarques
OPERATION	INTEGER	L'opération sur la table utilisateur doit être reflétée sur l'index en texte intégral (insertion = 0/ mise à jour = 1/ suppression = 2). Lorsque vous utilisez une table de capture de réplication, les opérations de mise à jour sont divisées entre opérations de suppression et d'insertion. Dans ce cas, une insertion dans la table des événements provient soit d'une insertion, soit d'une mise à jour de la table source sur laquelle l'index a été créé.
TIME	TIMESTAMP	Horodatage de la création de l'entrée d'événement.
REASON	INTEGER	Code anomalie. Pour une liste des codes anomalie, reportez-vous à l'Annexe I, «Codes anomalie du moteur de recherche de texte», à la page 239.
SEVERITY	INTEGER	Gravité de l'entrée de table. Par exemple, 1 indique des éléments donnés à titre d'information, 4 représente un avertissement et 8 une erreur d'entrée de table.
MESSAGE	VARCHAR(1024)	Informations supplémentaires.
KEY1, ... KEY14	Dépend de la table utilisateur	De la première colonne de la clé primaire de la table utilisateur à la dernière colonne de la clé primaire (un maximum de 14).

Vues pour les informations de niveau d'index

Tableau 15. La vue d'événements (suite)

Attribut	Type	Remarques
PARTITION	INTEGER	Numéro de partition de base de données où survient l'erreur. Dans un environnement non réparti, il s'agit de 0.

Les événements peuvent être supprimés par la commande DB2TEXT CLEAR EVENTS, reportez-vous à la rubrique «CLEAR EVENTS» à la page 112 pour plus d'informations.

Remarque

Des événements informationnels, tels que le démarrage, la validation et la fin de la mise à jour sont également disponibles dans cette vue.

Dans ce cas, Key1, ... Key14 et OPERATION ont tous une valeur NULL.

Dans le cas des index sur les vues, les colonnes PK01, ..., PK14 font référence aux colonnes spécifiées dans la clause KEY COLUMNS de la commande CREATE INDEX.

Tables de journalisation, vues et alias de type nickname

La table de journalisation est destinée à stocker les opérations de modification sur la vue ou la table utilisateur qui nécessite alors une synchronisation avec l'index externe en texte intégral.

Pour les index créés sur des tables normales ou des tables d'alias de type nickname, les déclencheurs créés sur la table utilisateur permettent d'insérer les informations sur les modifications dans la table de journalisation. Cependant, si les tables de capture de réplication sont utilisées, aucune table de journalisation n'est créée et la table de capture de réplication est alors utilisée.

Pour les tables de journalisation, la commande de mise à jour lit les entrées et les supprime après l'aboutissement de la synchronisation.

Cependant, dans le cas d'index sur les vues, les déclencheurs ne peuvent pas remplir la table de journalisation. Comme vous pouvez mettre à jour la vue, l'utilisateur est responsable de cette tâche.

Tableau 16. Vue de la table de journalisation

Attribut	Type	Remarques
OPERATION	INTEGER	Type de modification sur la table utilisateur nécessitant une synchronisation d'index : (0 = insertion, 1 = mise à jour, 2 = suppression).
TIME	TIMESTAMP	Horodatage pour la création d'une ligne dans cette table.
PK01 ... PKnm	Même que la table utilisateur	En cas d'erreurs, colonne dans laquelle l'erreur est survenue. Il s'agit d'une copie des colonnes de clé primaire de la table utilisateur ou des colonnes de clé équivalentes dans le cas d'un index sur une vue.

Vues pour les informations de niveau d'index

L'utilisateur qui crée la table peut sélectionner, mettre à jour, insérer et supprimer cette vue.

Si vous spécifiez une table de capture de réplication dans la commande de création d'index, aucune table de journalisation n'est créée et la table de capture de réplication est alors utilisée. La table de capture de réplication doit comporter les colonnes suivantes :

Tableau 17. La table de capture de réplication

Attribut	Type	Remarques
IBMSNAP_OPERATION	INTEGER	Type de modification sur des tables CD ou CCD nécessitant une synchronisation d'index : (I = insertion, U= mise à jour, D= suppression). Lorsque vous utilisez une table de capture de réplication, les opérations de mise à jour sont divisées entre opérations de suppression et d'insertion. Dans ce cas, une insertion dans la table des événements provient soit d'une insertion, soit d'une mise à jour de la table source sur laquelle l'index a été créé.
IBMSNAP_COMMITSEQ	CHAR	Mappe vers la colonne correspondante de la table CD ou CCD.
IBMSNAP_INTENTSEQ	CHAR	Mappe vers la colonne correspondante de la table CD ou CCD.
PK01 ... PKnm	Même que la table utilisateur	En cas d'erreurs, colonne dans laquelle l'erreur est survenue. Il s'agit de colonnes de clé primaire de la table utilisateur.

L'utilisateur qui définit la table peut réaliser des sélections, des mises à jour, des insertions et des suppression avec l'option Octroi.

Annexe D. CCSID pris en charge

Les CCSID suivants sont pris en charge sur DB2 Extension Net Search.

CCSID

37	Etats-Unis/Canada - CECP
273	Allemagne F.R./Autriche - CECP
274	Ancienne page de codes de la Belgique
277	Danemark, Norvège - CECP
278	Finlande, Suède - CECP
280	Italie - CECP
284	Espagne/Amérique latine - CECP
285	Royaume-Uni - CECP
290	Japonais (Katakana) étendu
297	France - CECP
301	Japon DB PC
367	ASCII
420	Arabe Bilingue
423	Grèce - 183
424	Israël (Hébreu)
437	Anglais US
500	International #5
737	MS DOS Grec
806	Hindi
813	Grec
819	Latin-1
833	Coréen étendu
836	Chinois simplifié étendu
838	Thaïlandais avec marques de tonalité basse & Caractères anciens
848	PC, Cyrillique, Ukrainien avec l'Euro
850	Latin-1
852	Latin-2
855	Bulgare
857	Turc
858	Ordinateur personnel - Multilingue avec l'Euro
860	Portugais

862	Hébreu
863	Canadien
864	Arabe
866	Russe
867	Israël - Ordinateur personnel
869	Grec
870	Latin 2 - EBCDIC Multilingue
871	Islande - CECP
872	Cyrillique - PC avec l'Euro
874	Thaïlandais
875	Grèce
891	Corée - Ordinateur personnel
895	Japon Latin 7 bits
901	PC Baltique Multi avec l'Euro
902	Estonie 8 bits avec l'Euro
912	Latin-2
915	Russe
916	Hébreu
920	Turc
921	Letton, Estonien
922	Estonien
923	Latin 9
924	Latin 9 EBCDIC
927	Taïwan PC
930	Japon EBCDIC
932	Japonais, combiné SBCS/DBCS
933	Coréen
935	Chinois (simplifié)
937	Chinois (traditionnel)
938	Taïwan PC
939	Japonais
941	Japon OPEN
942	Japonais, combiné SBCS/DBCS
943	Japonais, combiné SBCS/DBCS
948	Chinois (traditionnel), combiné SBCS/DBCS
949	Coréen
950	Chinois (traditionnel), combiné SBCS/DBCS

954	Japonais
964	Chinois (traditionnel), combiné SBCS/DBCS
970	Coréen
1025	Cyrillique, Multilingue
1026	Latin #5 - Turquie
1027	Japonais (Latin) étendu
1040	Coréen étendu - Ordinateur personnel
1041	Japonais étendu - Ordinateur personnel
1043	Chinois traditionnel étendu - PC
1046	Arabe
1047	Latin 1/Open Systems
1051	H-P Emulation, Roman 8
1088	Coréen révisé - Ordinateur personnel
1089	Arabe
1112	Baltique - Multilingue, EBCDIC
1115	République populaire de Chine (RPC)-PC
1122	Estonie, EBCDIC
1123	Cyrillique, Ukraine
1124	Ukrainien
1125	Ukrainien
1131	Vietnamien
1137	Devanagari EBCDIC
1140	Etats-Unis, Canada, etc. ECECP
1141	Autriche, Allemagne ECECP
1142	Danemark, Norvège ECECP
1143	Finlande, Suède ECECP
1144	Italie ECECP
1145	Espagne, Amérique latine (Espagnol) ECECP
1146	Royaume-Uni ECECP
1147	France ECECP
1148	International ECECP
1149	Islande ECECP
1153	EBCDIC Latin 2 Multilingue avec l'Euro
1154	EBCDIC Cyrillique, Multilingue avec l'Euro
1155	EBCDIC Turquie avec l'Euro
1156	EBCDIC Baltique Multi avec l'Euro
1157	EBCDIC Estonie avec l'Euro

1158	EBCDIC Cyrillique, Ukraine avec l'Euro
1160	Thaïlandais avec marques de tonalité basse & Caractères anciens
1161	Thaïlandais avec marques de tonalité basse & Caractères anciens - PC
1162	MS Windows thaïlandais
1163	Vietnamien ISO-8 avec l'Euro
1164	Vietnamien EBCDIC avec l'Euro
1200	UCS2
1208	UTF8
1250	Latin-2, Biélorusse
1251	Russe
1252	Latin-1
1253	Tchèque
1254	Turc
1255	Hébreu
1256	Arabe
1258	Vietnamien
1275	Apple, Latin 1
1280	Apple Grec
1281	Apple Turc
1282	Apple Europe Centrale
1283	Apple Cyrillique
1351	Japon OPEN
1363	Coréen
1364	Coréen
1381	Chinois (simplifié), combiné SBCS/DBCS
1383	Chinois (simplifié), combiné SBCS/DBCS
1386	Chinois (simplifié), combiné SBCS/DBCS
1388	Chinois (simplifié), combiné SBCS/DBCS
1390	Japonais
1392	Chine GB18030
1394	Shift JIS X0213
1399	Japon EBCDIC
4909	Grèce/Latin ASCII
4930	Corée DB EBCDIC
4933	Chine EBCDIC
4971	Grèce EBCDIC
5026	Japonais Katakana

5035	Japonais Latin
5039	Japonais, combiné SBCS/DBCS
5210	Chine SB PC
5346	Windows Latin-2
5347	Windows Cyrillique
5348	Windows Latin-1
5349	Windows Grèce
5350	Windows Turquie
5351	Windows Hébreu avec l'Euro
5352	Windows Arabe
5353	Windows Baltique
5354	Vietnamien
9044	Latin-2 PC
9048	Hébreu PC
9049	Turquie PC
9061	Grèce PC
9238	Arabe - PC
12712	Hébreu EBCDIC
13121	Corée SB EBCDIC
13488	UCS2
17248	Arabe PC
17584	UCS-2
18030	caractère chinois standard
21427	Taïwan BIG-5
33722	IBMecJP
61955	

Annexe E. Langues prises en charge

Voici les paramètres de langue que vous pouvez spécifier dans DB2 Extension Net Search. Dans l'Extension Net Search, le seul traitement important relatif à la langue concerne les documents écrits en thaïlandais ou, en cas de demande, pendant le traitement des mots vides. Pour plus de détails, reportez-vous à la section «Mots vides» à la page 235.

AR_AA	Arabe/Pays arabes
BE_BY	Biélorusse/Biélorussie
BG_BG	Bulgare/Bulgarie
CA_ES	Catalan/Espagne
CS_CZ	Tchèque/République Tchèque
DA_DK	Danois/Danemark
DE_CH	Allemand/Suisse
DE_DE	Allemand/Allemagne
EL_GR	Grec/Grèce
EN_AU	Anglais/Australie
EN_BE	Anglais/Belgique
EN_GB	Anglais/Royaume-Uni
EN_US	Anglais/Etats-Unis
EN_ZA	Anglais/Afrique du Sud
ES_ES	Espagnol/Espagne
ET_EE	Estonien/Estonie
FI_FI	Finnois/Finlande
FR_BE	Français/Belgique
FR_CA	Français/Canada
FR_CH	Français/Suisse
FR_FR	Français/France
HE_IL	Hébreu/Israël
HI_IN	Hindi/Inde
HR_HR	Croate/Croatie
HU_HU	Hongrois/Hongrie
ID_ID	Indonésien/Indonésie
IT_CH	Italien/Suisse
IW_IL	Hébreu/Israël
IT_IT	Italien/Italie
JA_JP	Japonais/Japon

Langues

KO_KR	Coréen/Corée
LT_LT	Lituanien/Lituanie
LV_LV	Letton/Lettonie
MK_MK	Macédonien/FYR Macédoine
MS_MY	Malais/Malaisie
NB_NO	Norvégien Bokmal/Norvège
NL_BE	Hollandais/Belgique
NL_NL	Hollandais/Pays-Bas
NN_NO	Norvégien Nynorsk/Norvège
NO_NO	Norvégien/Norvège
PT_BR	Portugais/Brésil
PL_PL	Polonais/Pologne
PT_PT	Portugais/Portugal
RO_RO	Roumain/Roumanie
RU_RU	Russe/Russie
SH_SP	Serbe (Latin)/Serbie
SK_SK	Slovaque/Slovaquie
SL_SI	Slovène/Slovénie
SQ_AL	Albanais/Albanie
SR_SP	Serbe (Cyrillique)/Serbie
SV_SE	Suédois/Suède
TA_IN	Tamoul/Inde
TE_IN	Télougou/Inde
TH_TH	Thaïlandais/Thaïlande
TR_TR	Turc/Turquie
UK_UA	Ukrainien/Ukraine
VI_VN	Vietnamien/Vietnam
ZH_CN	Chinois/RPC
ZH_TW	Chinois/Taïwan

Annexe F. Messages DB2 Extension Net Search

DB2 Extension Net Search fournit les types de messages suivants :

- Messages d'information et d'avertissement
- Messages d'erreur

Notez que les états SQL renvoyés de la fonction de recherche sont 38600 plus le numéro d'erreur CTE.

Messages d'information et d'avertissement

CTE0001	Exécution de l'opération réussie.
CTE0002	Les services de verrouillage et de mise à jour sont activés et sont en cours de fonctionnement.
CTE0003	La mise à jour de l' a démarré.
CTE0004	La mise à jour de l' est terminée.
CTE0005	Validation de la mise à jour : "%1", "%2", "%3" insertion, mise à jour et suppression des documents réussies.
CTE0006	Incident lors de l'accès à l' de texte. Pour plus de détails, veuillez consulter db2diag.log.
CTE0007	La section "%1" n'apparaît dans aucun des documents ou est un nom de section de modèle de document incorrect.
CTE0008	Réorganisation de l' commencée.
CTE0009	Réorganisation de l' terminée.
CTE0010	L'attribut "%1" n'est pas correct.
CTE0011	Activation de la mémoire cache commencée.
CTE0012	Activation de la mémoire cache terminée.
CTE0013	Suppression de la mémoire cache persistante.
CTE0014	Désactivation de la mémoire cache.

Messages d'erreur

CTE0100	Une opération DB2 a échoué. Informations DB2 : "%2" "%4".
---------	--

Explication : Une erreur DB2 entrave la poursuite des opérations.

Solution : Pour plus de détails sur cette erreur DB2, reportez-vous à db2 ? SQLxxx.

CTE0101	Une opération du moteur de recherche a échoué. Code anomalie : "%2", "%3", "%4", "%5", "%6".
---------	---

Explication : Une erreur du moteur de recherche est survenue, elle empêche la poursuite des opérations.

Solution : Pour plus de détails, reportez-vous aux descriptions des codes de raison du moteur de recherche.

CTE0102	Une fonction générale du système a échoué. Erreur : "%2".
---------	--

Explication : Une erreur système entrave la poursuite des opérations.

Solution : Des informations complémentaires se trouvent dans le fichier d'en-tête errno.h sous UNIX.

Messages d'erreur

CTE0103 Une erreur interne s'est produite.
Emplacement : "%1", "%2".

Explication : Une erreur de traitement interne entrave la poursuite des opérations. Essayez de démarrer et d'arrêter la mise à jour, les services de verrouillage et DB2.

Solution : Si l'erreur persiste, exécutez un programme trace et vérifiez db2diag.log.

CTE0104 Erreur d'allocation de mémoire (moteur de recherche).

Explication : Le système manque de mémoire.

Solution : Augmentez la taille de la mémoire allouée au propriétaire d'instance ou arrêtez les autres processus en exécution parallèle.

CTE0105 Erreur d'allocation de mémoire.

Explication : Le système manque de mémoire.

Solution : Augmentez la taille de la mémoire allouée à l'utilisateur ou arrêtez les autres processus en exécution parallèle.

CTE0106 La table "%1"."%2" ne contient pas de clé primaire.

Explication : Vous avez tenté de créer une table qui ne possède pas de clé primaire.

Solution : Appelez db2 alter table pour vous assurer de l'existence d'une clé primaire. Essayez ensuite de recréer l'.

CTE0107 Le répertoire "%1" n'existe pas.

Explication : Vous avez indiqué un répertoire qui n'existe pas.

Solution : Créez le répertoire en accordant l'accès au propriétaire d'instance. Essayez ensuite de respecifier l'. Notez que dans un environnement DB2 réparti, ce répertoire doit exister sur chaque nœud physique.

CTE0108 La taille interne "%4" des colonnes de la clé sur l'objet "%1"."%2" est supérieure à la taille maximale autorisée "%3".

Explication : La représentation interne des colonnes de clé dépasse la taille maximale.

Solution : Modifiez la présentation de la table avant de recréer l'. Utilisez de petites colonnes de clé qui améliorent les performances.

CTE0109 Le nombre de colonnes de clé "%3" sur l'objet "%1"."%2" est supérieur à la taille maximale "%4".

Explication : Un nombre maximal de 14 colonnes de clé est pris en charge.

Solution : Modifiez la présentation de la table avant de recréer l'.

CTE0111 Le fichier "%1" ne peut pas être lu.

Explication : Le fichier ne peut être lu.

Solution : Vérifiez les droits d'accès du fichier. Tenez compte du fait que la procédure enregistrée fonctionne comme un ID utilisateur isolé qui requiert les droits d'accès pour traiter ce fichier.

CTE0112 Le fichier "%1" ne peut pas être ouvert.

Explication : Impossible d'ouvrir le fichier spécifié.

Solution : Vérifiez que le fichier est correctement spécifié.

CTE0113 Erreur lors de la conversion du fichier modèle "%1" en code UTF8.

Explication : Le CCSID que vous avez indiqué ou le CCSID de base de données par défaut ne correspond pas au CCSID du fichier modèle.

Solution : Entrez les indications correctes pour le CCSID du fichier modèle.

CTE0114 Impossible d'enregistrer le modèle de document "%1" dans le fichier "%2".

Explication : Le fichier modèle n'a pas pu être utilisé.

Solution : Vérifiez que la syntaxe du fichier modèle est correcte.

CTE0115 Un incident de verrouillage s'est produit. Informations concernant le gestionnaire de verrouillage : "%1" "%2".

Explication : Un incident de verrouillage interne s'est produit.

Solution : Vérifiez les verrouillages courants à l'aide de la commande db2text control. A l'aide de cette commande, supprimez les verrous en attente. Si cela ne suffit pas, arrêtez, puis recommencez le verrouillage et la mise à jour.

CTE0116 Opération incompatible avec le verrou existant.

Explication : Vous avez lancé une commande qui actuellement n'est pas autorisée car d'autres commandes ont déjà été lancées dans cet .

Solution : Vérifiez les verrous mis en suspens dans cet pour savoir quelles commandes sont en cours. Patientez jusqu'à la fin des autres commandes. Si cette opération n'est plus en cours d'exécution mais si le verrou est toujours actif, rafraîchissez les verrous de cet et réessayez.

CTE0117 Tout l'espace de verrouillage disponible pour les bases de données est utilisé. Veuillez modifier la configuration.

Explication : Vous avez essayé de traiter une ou plusieurs bases de données configurées dans votre fichier de verrouillage.

Solution : Modifiez le nombre de bases de données que vous voulez traiter en parallèle dans votre configuration de verrouillage db2extlm.cfg. Recommencez la mise à jour et redémarrez les services de verrouillages à l'aide des commandes db2text stop et db2text start.

CTE0118 Tout l'espace de verrouillage disponible pour les index d'une base de données est utilisé. Veuillez modifier la configuration.

Explication : Vous avez essayé de traiter plus d'index sur une base de données que la configuration de votre fichier de verrouillage ne le prévoit.

Solution : Modifiez le nombre d'index que vous voulez traiter en parallèle dans votre fichier de configuration de verrouillage db2extlm.cfg. Recommencez la mise à jour et redémarrez les services de verrouillages à l'aide des commandes db2text stop et db2text start.

CTE0119 Tout l'espace disponible pour les verrous dans un index est utilisé.

Explication : Les opérations en cours demandent plus de verrous pour un index que la configuration de votre fichier de configuration de verrouillage n'a prévu.

Solution : Modifiez le nombre de verrous que vous voulez traiter en parallèle dans votre configuration de verrouillage db2extlm.cfg. Recommencez la mise à jour et redémarrez les services de verrouillages à l'aide des commandes db2text stop et db2text start.

CTE0120 Erreur dans le fichier de configuration des services de verrouillage et de mise à jour.

Explication : Le fichier de configuration est db2extlm.cfg erroné.

Solution : Vérifiez le fichier db2extlm.cfg et corrigez l'erreur. Recommencez la mise à jour et redémarrez les services de verrouillage à l'aide de la commande db2text start.

CTE0121 Impossible d'ouvrir le fichier de configuration des services de verrouillage et de mise à jour.

Explication : Impossible d'ouvrir le fichier db2extlm.cfg.

Solution : Vérifiez que le fichier existe réellement et que vous pouvez y accéder. Si le fichier n'est pas accessible, essayez de mettre à jour votre instance db2 à l'aide de db2iupdt.

CTE0122 Erreur de syntaxe dans le fichier de configuration des services de verrouillage et de mise à jour.

Explication : Erreur de syntaxe dans le fichier de configuration des services de verrouillage et de mise à jour.

Solution : Identifiez les erreurs dans le fichier de configuration des services de mise à jour et de verrouillage.

CTE0126 Le fichier d'entrée du service de verrouillage et de mise à jour "%1" est endommagé.

Explication : Le fichier requis pour le service de mise à jour et de verrouillage est peut-être endommagé.

Solution : Vérifiez que le fichier existe réellement et que vous pouvez y accéder. Si vous pouvez accéder au fichier, renommez-le et redémarrez le service de mise à jour et de verrouillage. Le fichier doit être créé de nouveau. Cependant toutes les mises à jour indiquées pour l'index de création sont supprimées.

CTE0127 Une erreur interne au service de mise à jour et de verrouillage s'est produite. Code anomalie : "%1".

Explication : Une erreur interne s'est produite dans la zone du service de mise à jour et de verrouillage.

Solution : Arrêtez DB2 et l'Extension Net Search, puis supprimez vos ressources partagées. Réessayez et redémarrez. Si le problème persiste, avertissez votre représentant IBM.

Messages d'erreur

CTE0129 Les valeurs NULL ne peuvent pas être transmises comme paramètres.

Explication : DB2 a transmis une valeur NULL à une fonction interne définie par l'utilisateur.

Solution : Assurez-vous tout d'abord que la table de base spécifiée possède une clé primaire. Modifiez votre instruction SELECT pour éviter ce problème. Revenez à la fonction trace et transmettez les informations renvoyées aux services d'assistance IBM.

CTE0130 L'argument de recherche spécifié dépasse la longueur maximale autorisée. La longueur des critères de recherche est actuellement "%1". Or, la longueur maximale acceptée est "%2".

Explication : La longueur des critères de recherche indiqués est "%1". La longueur maximale ne doit pas dépasser "%2".

Solution : Diminuez la longueur de vos critères de recherche à "%2".

CTE0131 La fonction définie par l'utilisateur (fonction UDF) "%1". "%2" n'existe pas.

Explication : La fonction UDF que vous avez indiquée n'existe pas dans la base de données.

Solution : Vérifiez le nom indiqué à cette fonction UDF ou enregistrez la fonction UDF dans la base de données utilisée.

CTE0132 L'index de texte "%1". "%2" n'existe pas.

Explication : L'index de texte que vous avez indiqué n'existe pas dans la base de données.

Solution : Vérifiez le nom indiqué et la base de données utilisée. Utilisez la vue db2ext.textcolumns pour afficher les index de texte existants.

CTE0133 L'index de texte "%1". "%2" existe déjà.

Explication : L'index de texte indiqué existe déjà dans cette base de données.

Solution : Vérifiez le nom indiqué et la base de données utilisée. Utilisez la vue db2ext.textcolumns pour afficher les index de texte existants.

CTE0135 L'objet "%1". "%2" n'existe pas.

Explication : Le nom de l'objet indiqué n'existe pas dans cette base de données.

Solution : Vérifiez le nom de l'objet indiqué et la base de données que vous utilisez.

CTE0136 La colonne "%1" n'existe pas dans "%2"."%3".

Explication : La colonne spécifiée n'existe pas.

Solution : Vérifiez le nom de la colonne indiqué. Vérifiez la table, la vue ou la base de données que vous utilisez.

CTE0137 L'espace table "%1" n'existe pas.

Explication : L'espace table indiqué n'existe pas dans cette base de données.

Solution : Vérifiez le nom indiqué et la base de données utilisée.

CTE0138 L'espace table "%1" est incorrect.

Explication : L'espace table indiqué n'est pas correct. La table des événements ne peut être créée que dans un espace table normal.

Solution : Réutilisez cette commande avec un espace table normal.

CTE0139 La variable d'environnement "%1" n'est pas paramétrée.

Explication : Une variable requise de l'environnement n'a pas été définie.

Solution : Vérifiez votre environnement, indiquez la variable requise et lancez de nouveau la commande.

CTE0140 La base de données "%1" est déjà activée pour le texte.

Explication : La base de données indiquée est déjà activée pour le texte.

Solution : Vérifiez le nom que vous avez spécifié. Vérifiez également la variable DB2DBDFT qui comprend la connexion implicite.

CTE0141 La base de données "%1" n'est pas activée pour le texte.

Explication : La base de données indiquée n'est pas activée pour le texte.

Solution : Vérifiez le nom indiqué et la variable DB2DBDFT. Si le nom de la base de données est correct, utilisez la commande db2text enable database for text.

CTE0142 La commande requiert un droit de mise à jour dans la table "%1". "%2" pour l'utilisateur "%3".

Explication : Vous ne disposez pas des droits pour utiliser cette commande.

Solution : Seul le propriétaire de cette table peut

utiliser cette commande ou vous accorder l'autorisation voulue.

CTE0143 La commande requiert un droit d'administration des bases de données pour l'utilisateur "%1".

Explication : Vous ne disposez pas des droits pour utiliser cette commande.

Solution : Seul le propriétaire de cette base de données peut utiliser cette commande ou vous accorder l'autorisation requise.

CTE0144 Au moins un index de texte actif se trouve dans la base de données "%1".

Explication : Vous ne pouvez pas désactiver votre base de données jusqu'à ce que tous les index de texte soient supprimés.

Solution : Reportez-vous à la vue `db2ext.textcolumns` pour connaître les index existants. Supprimez les index existants à l'aide de la commande `DROP INDEX` ou spécifiez l'option `FORCE` avec la commande `DISABLE DATABASE`.

CTE0145 Le CCSID "%1" n'est pas pris en charge.

Explication : Le CCSID indiqué n'est pas pris en charge.

Solution : Spécifiez un CCSID valide.

CTE0146 La langue "%1" n'est pas prise en charge.

Explication : La langue spécifiée n'est pas prise en charge.

Solution : Spécifiez une langue valide.

CTE0147 Le format "%1" n'est pas pris en charge.

Explication : Le format spécifié n'est pas pris en charge.

Solution : Spécifiez un format valide.

CTE0148 Le format spécifié "%1" ne prend pas en charge ce fichier modèle.

Explication : Le format "%1" ne prend pas en charge les fichiers modèles.

Solution : Utilisez un format qui prend en charge un fichier modèle ou supprime le fichier modèle à partir de votre commande.

CTE0149 Trop de termes (commençant par "%1") sont spécifiés pour la fréquence de mise à jour des index.

Explication : La syntaxe de la fréquence de mise à jour n'est pas correcte.

Solution : Assurez-vous que les paramètres `DAY`, `HOURL` et `MINUTE` ne sont indiqués qu'une seule fois.

CTE0150 Fin de commande inattendue. Vérifiez la syntaxe de la commande.

Explication : La syntaxe de la commande est incorrecte.

Solution : Vérifiez la syntaxe de la commande. Vérifiez que vous avez bien indiqué les paramètres requis.

CTE0151 Marque "%1" inattendue. Vérifiez la syntaxe de la commande.

Explication : La syntaxe de la commande n'est pas correcte.

Solution : Vérifiez la syntaxe de la commande et vérifiez que la marque utilisée est acceptée dans cette commande.

CTE0152 Marque "%1" trop longue.

Explication : La marque est trop longue.

Solution : Vérifiez la syntaxe de la commande et que la marque a été réduite à la taille maximale autorisée.

CTE0153 La marque "%1" apparaît deux fois dans la fréquence de mise à jour.

Explication : Vous avez indiqué une syntaxe non correcte pour la fréquence de mise à jour.

Solution : Assurez-vous que les paramètres `DAY`, `HOURL` et `MINUTE` ne sont indiqués qu'une seule fois.

CTE0154 La valeur "%1" pour "%2" est hors plage. La plage valide est "%3" - "%4".

Explication : Vous avez indiqué une valeur qui n'est pas correcte. La valeur doit être comprise dans la plage autorisée.

Solution : Mettez à jour la commande. Modifiez la valeur afin qu'elle soit comprise dans la plage autorisée.

CTE0155 La chaîne de recherche est vide.

Explication : Vous avez indiqué une chaîne de recherche vide.

Solution : Vérifiez que la chaîne de recherche comprend des caractères alphanumériques valides.

Messages d'erreur

CTE0157 Erreur de syntaxe en regard de "%1".

Explication : Vous avez indiqué une syntaxe de recherche incorrecte.

Solution : Vérifiez la syntaxe située près de %1. Corrigez et réessayez.

CTE0158 La chaîne de recherche texte libre est manquante.

Explication : Indiquez une chaîne de texte libre.

Solution : Vérifiez que la chaîne de recherche située après "is about" comprend des caractères alphanumériques valides.

CTE0159 La chaîne de recherche dépasse la longueur de "%1" autorisée.

Explication : La chaîne de recherche est trop longue.

Solution : Diminuez la chaîne de recherche et réessayez.

CTE0160 Aucune nom de section n'est spécifié dans la chaîne de recherche.

Explication : Vous devez indiquer un nom de section valide.

Solution : Ajoutez un nom de section valide et réessayez.

CTE0162 La commande ESCAPE n'a pas pu être traitée.

Explication : Votre chaîne de recherche comprend trop de caractères spéciaux pouvant être utilisés comme caractères génériques.

Solution : Réduisez le nombre de caractères spéciaux dans votre terme recherché ou évitez la commande d'échappement. Les caractères spéciaux suivants peuvent être utilisés : ! * + , _ . : ; { } ~ | ? [] ` = \

CTE0163 Aucun nom de thésaurus spécifié dans la clause du thésaurus.

Explication : Une recherche de thésaurus ne demande pas d'entrer un nom de thésaurus.

Solution : Définissez un nom de thésaurus dans vos critères de recherche.

CTE0164 Erreur de syntaxe dans la relation du thésaurus "%1".

Explication : La syntaxe indiquée pour le thésaurus n'est pas correcte.

Solution : Adaptez la relation du thésaurus aux caractéristiques de la syntaxe.

CTE0166 Freetext doit être la dernière instruction de la recherche.

Explication : D'autres opérateurs ne sont pas autorisés après la marque "is about".

Solution : Réécrivez la chaîne de requête. Le dernier opérateur est "is about".

CTE0167 Erreur de syntaxe dans la recherche de texte libre "%1".

Explication : La syntaxe de la chaîne de texte libre n'est pas correcte.

Solution : Adaptez la chaîne de texte libre aux caractéristiques de la syntaxe.

CTE0168 Une parenthèse ouvrante est manquante dans une instruction de section.

Explication : La syntaxe de l'instruction de la section n'est pas correcte.

Solution : Adaptez l'instruction de la section aux caractéristiques de la syntaxe.

CTE0169 Une virgule ou une parenthèse fermante est manquante dans une instruction de section.

Explication : La syntaxe de l'instruction de la section n'est pas correcte.

Solution : Adaptez l'instruction de la section aux caractéristiques de la syntaxe.

CTE0170 Des guillemets de fermeture sont manquants.

Explication : La syntaxe indiquée pour le terme de syntaxe n'est pas correcte.

Solution : Mettez à jour le terme recherché selon la syntaxe indiquée.

CTE0171 Des guillemets d'ouverture associés à un nom de section sont manquants.

Explication : La syntaxe de l'instruction de la section n'est pas correcte.

Solution : Adaptez l'instruction de la section aux caractéristiques de la syntaxe.

CTE0172 Des guillemets de fermeture associés à un nom de section sont manquants.

Explication : La syntaxe de l'instruction de la section n'est pas correcte.

Solution : Adaptez l'instruction de la section aux caractéristiques de la syntaxe.

CTE0173 **Un seul caractère d'échappement doit être défini dans une clause d'échappement.**

Explication : Vous ne pouvez pas définir plusieurs caractères dans une clause d'échappement.

Solution : Supprimez les caractères supplémentaires dans la clause d'échappement.

CTE0174 **Le caractère blanc ne peut pas être utilisé comme caractère d'échappement.**

Explication : Une clause d'échappement n'accepte pas de caractère blanc.

Solution : Corrigez la clause d'échappement.

CTE0175 **Une clause d'échappement est définie mais aucun caractère générique n'a été trouvé dans la phrase faisant l'objet d'une recherche.**

Explication : Une clause d'échappement est indiquée sans caractère masqué.

Solution : Supprimez la clause d'échappement.

CTE0176 **Le caractère placé, dans la phrase, après un caractère d'échappement n'est ni un caractère du même type, ni un caractère masqué.**

Explication : Le caractère suivant le caractère d'échappement doit être un caractère de masquage ou le caractère d'échappement lui-même.

Solution : Modifiez la chaîne de recherche pour utiliser correctement le caractère d'échappement.

CTE0177 **La valeur numérique "%1" est incorrecte.**

Explication : La valeur numérique indiquée dans les critères de recherche n'est pas correcte.

Solution : Vérifiez la valeur dans la documentation. Mettez à jour la valeur dans les critères de recherche.

CTE0178 **Les caractères génériques dans la phrase à correspondance partielle doivent être précédés d'un caractère d'échappement.**

Explication : Les caractères génériques ne sont pas autorisés avec une recherche floue.

Solution : Adaptez la chaîne de recherche au caractère d'échappement.

CTE0179 **Le nom du thésaurus "%1" dépasse la longueur autorisée de "%2".**

Explication : Les clés primaires plus longues que 60 octets ne sont pas prises en charge.

Solution : Modifiez la présentation de la table avant de recréer l'index.

CTE0180 **Le thésaurus "%1" est introuvable.**

Explication : Le thésaurus spécifié est introuvable.

Solution : Vérifiez que les fichiers du thésaurus se trouvent dans le répertoire du thésaurus.

CTE0181 **La bibliothèque "%1" n'a pas pu être chargée.**

Explication : La bibliothèque est introuvable.

Solution : Vérifiez que la bibliothèque se trouve bien au chemin d'accès de la bibliothèque et qu'elle est disponible. Démarrez et arrêtez DB2 pour vérifier que les paramètres courants sont utilisés.

CTE0182 **La fonction "%1" n'a pas pu être chargée à partir de la bibliothèque "%2".**

Explication : Le point d'entrée de la bibliothèque ne peut pas être chargé.

Solution : La bibliothèque accédée semble être incorrecte. Vérifiez que la bibliothèque n'est spécifiée qu'une seule fois.

CTE0183 **Une erreur s'est produite lors de l'utilisation des ressources partagées du système.**

Explication : Une requête effectuée dans les ressources du système partagé comme la mémoire partagée ou les sémaphores est impossible.

Solution : Vérifiez l'état actuel du système et la configuration. Sous UNIX, utilisez la commande `ipcs` pour vérifier les ressources. Arrêtez toutes les applications, telles DB2 et l'Extension Net Search. Si d'autres ressources sont répertoriées, supprimez-les à l'aide de la commande `ipcrm`.

CTE0184 **Aucune commande db2text n'a été exécutée.**

Explication : Une commande nécessitant les services de verrouillage et de mise à jour a été appelée.

Solution : Démarrez les services de verrouillage et de mise à jour avec la commande `db2text start`.

Messages d'erreur

CTE0185 Les services de verrouillage et de mise à jour sont déjà actifs.

Explication : Une commande db2text start est exécutée mais les services de mise à jour et de verrouillage fonctionnent déjà.

Solution : Aucune autre action.

CTE0186 Une erreur concernant le service de verrouillage et de mise à jour s'est produite, vérifiez db2diag.log.

Explication : Une erreur interne au service de mise à jour et de verrouillage s'est produite.

Solution : Reportez-vous à db2diag.log pour obtenir davantage d'informations ou nettoyez les ressources partagées. Voir aussi CTE0183.

CTE0187 Les services de mise à jour et de verrouillage sont toujours actifs, utilisez l'option FORCE pour les arrêter.

Explication : La commande db2text stop n'a pas arrêté les services de verrouillage, certaines opérations sont toujours en cours d'exécution.

Solution : Vérifiez à l'aide de la commande db2text control quelles opérations sont en cours de fonctionnement et patientez avant qu'elles s'achèvent. Si vous voulez les arrêter, utilisez l'option FORCE.

CTE0188 Un incident temporaire s'est produit lors de l'utilisation des services de verrouillage et de mise à jour. Veuillez recommencer.

Explication : La commande db2text stop n'a pas arrêté les services de verrouillage. Des programmes sont en cours de fonctionnement ou un incident a été détecté.

Solution : Vérifiez à l'aide de la commande db2text control quelles opérations sont en cours de fonctionnement et patientez avant qu'elles s'achèvent. Pour les arrêter, utilisez l'option FORCE.

CTE0189 Programme exécutable "%1" introuvable.

Explication : Impossible de situer le fichier du programme et d'y accéder.

Solution : Vérifiez si le fichier du programme se trouve dans la corbeille ou dans le répertoire adm du serveur DB2. L'installation est corrompue si le fichier n'a pas été trouvé.

CTE0190 Le programme exécutable "%1" n'a pas pu démarrer.

Explication : Impossible de démarrer le programme.

Solution : Vérifiez si le programme se trouve dans la corbeille ou dans le répertoire adm du serveur DB2 et que les bibliothèques appropriées sont installées. Pour plus d'informations, démarrez manuellement le programme sur le serveur.

CTE0191 L'opération drop index (suppression d'index) a été réalisée de manière incomplète. Pour plus de détails, veuillez consulter db2diag.log.

Explication : L'opération drop index (suppression d'index) a été réalisée de manière incomplète, peut-être en raison de l'option FORCE.

Solution : L'option FORCE supprime tout sans tenir compte des erreurs. Vérifiez s'il y a des fichiers temporaires dans le répertoire index et supprimez-les manuellement.

CTE0192 Des erreurs se sont produites lors de l'opération update index (mise à jour d'index). Vérifiez la table index "%1". "%2" et db2diag.log pour plus de détails.

Explication : Pendant la mise à jour des index, les erreurs détectées dans les documents sont ajoutées à la table d'événements.

Solution : Vérifiez la table d'événements pour plus d'informations sur les erreurs détectées dans le document. Rafraîchissez le journal des événements après résolution des incidents.

CTE0194 Le type "%1" de la colonne "%2" n'est pas pris en charge.

Explication : Vous avez utilisé une colonne qui ne figure pas dans la liste des colonnes prises en charge.

Solution : Vérifiez la liste des colonnes appropriées pour les clés primaires et l'indexation dans l'index de création. Apportez les modifications appropriées à votre commande et réessayez.

CTE0195 "%1" n'est pas un chemin d'accès absolu.

Explication : Un chemin d'accès absolu au serveur est nécessaire.

Solution : Vérifiez le chemin d'accès et écrivez un chemin d'accès absolu dans la commande.

CTE0198 Pas d'index de texte correspondant.

Explication : Aucun index de texte ne figure dans la colonne.

Solution : Vérifiez que l'index de texte existe toujours.

CTE0199 Il n'y a pas d'index de texte correspondant à la colonne "%1" de la table "%2".

Explication : Vous avez essayé de chercher une colonne sans index de texte.

Solution : Vérifiez la colonne dans laquelle vous effectuez la recherche ou créez un index de texte dans la colonne.

CTE0200 Une commande doit être au moins spécifiée.

Explication : La commande ALTER INDEX modifie les caractéristiques d'un index, comme les options de mise à jour et de stockage. Aucune des caractéristiques à modifier n'a été indiquée.

Solution : Indiquez au moins une option de commande. Veuillez vous reporter à la syntaxe de la commande pour obtenir toutes les options possibles.

CTE0201 Il y a un conflit avec l'index de texte existant sur la même colonne.

Explication : Un index de texte défini dans la même colonne a été créé avec des paramètres différents à partir de la commande de création d'index.

Solution : Corrigez les valeurs des paramètres dans la commande de création d'index. Assurez-vous que les paramètres suivants ont la même valeur pour l'index existant et l'index à créer : ccsid, langue, format, modèle de document, configuration d'index, fonction des colonnes et attributs.

CTE0202 L'objet "%1"."%2" doit être une vue lorsque les colonnes clés sont spécifiées.

Explication : L'objet spécifié ne figure pas dans une vue. La clause KEY COLUMNS FOR INDEX ON VIEW n'est autorisée que lors de l'indexation d'une colonne d'une vue.

Solution : Supprimez la clause KEY COLUMNS FOR INDEX ON VIEW (liste-nomcolonne-SQL).

CTE0203 L'index de texte "%1"."%2" n'a pas été créé avec l'option CACHE TABLE. Cette opération est nécessaire à l'exécution de la commande.

Explication : Cette commande ne peut être exécutée que si l'index indiqué a été créé avec l'option CACHE TABLE.

Solution : Créez un index avec l'option CACHE TABLE. Reportez-vous à la documentation pour plus d'informations sur la syntaxe de la commande.

CTE0204 Le nom de l'attribut est manquant. Ajoutez "AS <nom attribut>" à l'expression de l'attribut.

Explication : A chaque utilisation d'une expression de colonne dans l'expression d'attribut, un nom d'attribut doit être indiqué. Par exemple : (C1+C2 AS monnom).

Solution : Ajoutez "AS <nom attribut>" à l'expression de l'attribut.

CTE0205 Les expressions CACHE TABLE ne sont pas correctes.

Explication : La liste des colonnes dans l'expression de la table de la mémoire cache n'est pas correcte.

Solution : Corrigez la liste des colonnes de la table de la mémoire cache dans la commande de création d'index. Assurez-vous que les colonnes existent dans la table indiquée. Si une fonction est appliquée dans une colonne, vérifiez qu'elle est correctement utilisée.

CTE0206 Les expressions ATTRIBUTE ne sont pas correctes.

Explication : La liste de colonnes dans l'expression des attributs n'est pas correcte.

Solution : Corrigez la liste des colonnes des attributs dans la commande de création d'index. Assurez-vous que les colonnes existent dans la table indiquée. Si une fonction est appliquée dans une colonne, vérifiez qu'elle est correctement utilisée.

CTE0207 KEY COLUMNS FOR INDEX ON VIEW non spécifiés pour l'index de la vue "%1"."%2".

Explication : Si les index présents dans les vues sont créés, la clause KEY COLUMNS FOR INDEX ON VIEW (liste-nomcolonne-SQL) doit être spécifiée. La liste des noms de colonne indique les colonnes qui identifient une ligne dans la vue.

Solution : Incluez la clause KEY COLUMNS FOR INDEX ON VIEW (liste-nomcolonne-SQL) dans la commande de création d'index.

CTE0208 Les colonnes INITIAL SEARCH RESULT ORDER ne sont pas correctes.

Explication : La liste des colonnes dans l'expression INITIAL SEARCH RESULT ORDER (liste-par-ordre-SQL) n'est pas correcte.

Solution : Corrigez l'ordre de la liste de colonne dans la commande create index (création d'index). Vérifiez que la syntaxe est correcte et que les colonnes existent

Messages d'erreur

dans la table spécifiée. Si une fonction est appliquée dans une colonne, vérifiez qu'elle est correctement utilisée.

CTE0209 **Le type "%1" de l'attribut de colonne "%2" n'est pas pris en charge, le type DOUBLE est requis.**

Explication : Dans le cas des colonnes d'attributs, le seul type de données pris en charge est DOUBLE.

Solution : Assurez-vous que les colonnes d'attributs de la table avec la colonne texte à indexer sont de type DOUBLE. Il est possible d'utiliser des opérateurs de transtypage dans les expressions de colonnes d'attributs. Reportez-vous aux références SQL pour obtenir les types de données qui peuvent être transtypés en double.

CTE0210 **La valeur "%1" du paramètre de configuration d'index "%2" n'est pas correcte. "%3" est une valeur correcte.**

Explication : La valeur indiquée pour le paramètre de configuration est incorrecte. Pour obtenir les valeurs correctes des paramètres, reportez-vous à la syntaxe des commandes.

Solution : Corrigez la valeur des paramètres de configuration de l'index dans la commande de création d'index.

CTE0211 **"%1" n'est pas un paramètre de configuration d'index valide.**

Explication : L'option de configuration d'index n'est pas connue.

Solution : Vérifiez la syntaxe de la commande create index (création d'index). Les options de configuration d'index correctes sont TreatNumbersAsWords et IndexStopWords. Elles doivent être séparées par des virgules : configuration d'index (treatnumberaswords 1, indexstopwords 1).

CTE0212 **Le fichier de configuration d'index interne "%1" n'a pas pu être enregistré.**

Explication : Le fichier de configuration d'index interne n'a pas pu être enregistré.

Solution : Assurez-vous que le propriétaire d'instance possède les droits d'écriture sur le répertoire dans lequel le fichier doit être enregistré. Si un fichier portant le même nom existe déjà, assurez-vous que le propriétaire d'instance possède les droits d'écriture de ce fichier.

CTE0213 **Le modèle de fichier de configuration d'index interne "%1" n'a pas pu être enregistré.**

Explication : Le modèle de fichier de configuration d'index ne doit pas être lu.

Solution : Assurez-vous que le fichier existe à l'emplacement correct et qu'il peut être lu.

CTE0214 **Erreur interne à la définition d'une nouvelle entrée "[%1],%2=%3" dans le fichier de configuration d'index.**

Explication : Erreur interne lors de l'écriture du fichier de configuration interne pour l'index.

Solution : Si le fichier existe, vérifiez que le propriétaire d'instance possède les droits d'écriture et de lecture de ce fichier. Vérifiez qu'il y a assez d'espace sur l'unité où se trouve le fichier.

CTE0215 **La création d'index avec l'alias "%1"."%2" n'est pas prise en charge. Utilisez la table de base "%3"."%4" à la place.**

Explication : Impossible de créer l'index dans l'alias.

Solution : Entrez la commande de création d'index avec la table de base.

CTE0217 **Le planificateur est déjà actif.**

Explication : Le service est déjà actif, vous ne devez pas le démarrer.

Solution : Aucune action n'est requise.

CTE0218 **La fonction "%1" a échoué et génère le code d'erreur "%2".**

Explication : Une fonction Windows a échoué, le code d'erreur précise que la poursuite des opérations est entravée.

Solution : Utilisez le code d'erreur du système Windows pour obtenir des informations détaillées sur l'anomalie.

CTE0219 **Le service "%1" n'a pas pu être ouvert. Code d'erreur "%2".**

Explication : Le service spécifié est introuvable dans le système Windows.

Solution : Vérifiez que le service indiqué est bien installé sur votre système Windows. Utilisez le code d'erreur du système Windows pour obtenir des informations détaillées sur l'anomalie.

CTE0220 Le chemin d'accès au profil de l'instance DB2 n'a pas pu être trouvé.

Explication : La fonction interne de DB2 permettant d'obtenir le chemin du profil de l'instance DB2 a échoué.

Solution : Créez une instance DB2 sans indiquer le chemin du profil de l'instance et réessayez.

CTE0221 La fréquence de mise à jour "%1" n'est pas correctement indiquée.

Explication : La syntaxe de l'instruction de la fréquence de mise à jour n'est pas correcte.

Solution : Corrigez l'instruction de la fréquence de mise à jour selon la syntaxe indiquée.

CTE0222 Le fichier d'entrée du planificateur "%1" est corrompu.

Explication : Le fichier du planificateur contenant des informations de mise à jour de l'index est corrompu.

Solution : Utilisez votre éditeur système et essayez d'apporter une solution à l'incident. Une entrée a peut-être été tronquée ou le caractère de fin de ligne a été supprimé. Si le contenu du fichier n'est pas restauré, essayez la procédure suivante :

- Appelez la commande db2text stop pour arrêter le planificateur.
- Supprimez le fichier des services du planificateur.
- Appelez la commande db2text start pour démarrer le planificateur.
- Utilisez la commande db2text alter index ... pour recréer les entrées de fréquence de mise à jour pour tous les index concernés.

CTE0223 Le fichier "%1" n'a pas pu être fermé.

Explication : Le fichier indiqué n'a pas pu être fermé.

Solution : Vérifiez que le fichier est correctement spécifié.

CTE0224 Le fichier "%1" n'a pas pu être copié sur "%2".

Explication : Le premier fichier ne peut être copié sur le second fichier.

Solution : Vérifiez que les fichiers ont été correctement indiqués. Vérifiez que le second fichier existe déjà et est en lecture seule. Vérifiez également que l'espace disque est suffisant.

CTE0225 Le fichier "%1" n'a pas pu être supprimé.

Explication : Le fichier indiqué n'a pas pu être supprimé du système.

Solution : Vérifiez que le fichier a été correctement indiqué et vérifiez les droits d'accès au fichier.

CTE0225 Le fichier "%1" n'a pas pu être supprimé.

Explication : Le fichier indiqué n'a pas pu être supprimé du système.

Solution : Vérifiez que le fichier a été correctement indiqué et vérifiez les droits d'accès au fichier.

CTE0227 Echec d'écriture sur le fichier "%1".

Explication : Le fichier indiqué ne peut être lu.

Solution : Vérifiez que le fichier a été correctement indiqué et vérifiez les droits d'accès au fichier. Vérifiez également que l'espace disque est suffisant.

CTE0228 L'utilisateur ne dispose pas de droits d'accès suffisants au niveau du système d'exploitation.

Explication : La commande doit disposer des droits d'administrateur au niveau du système d'exploitation.

Solution : Assurez-vous que vous disposez des droits d'administrateur sur le système d'exploitation. Vérifiez que vous êtes membre du groupe administrateur.

CTE0231 "%1" n'est pas défini dans le même groupe de noeuds ("%4") comme l'espace table de "%2"."%3".

Explication : L'espace des tables d'administration est requis pour être distribué aux différents noeuds de la même manière que la table comprise dans la colonne de texte à indexer. Pour mettre en oeuvre cette mesure, l'espace table indiqué est vérifié afin d'observer s'il est défini dans le même groupe de noeuds.

Solution : Indiquez un espace table qui est défini dans le même groupe de noeuds que la table comprenant la colonne de texte à indexer.

CTE0232 L'espace table indiqué ou par défaut "%1" ne possède pas qu'un seul noeud. C'est ce qui est demandé dans le cas d'un index de vue ou lorsque l'option CACHE TABLE est indiquée.

Explication : Un index dans une vue ou bien doté de l'option CACHE TABLE activée n'est pris en charge que pour les tables dans un seul noeud.

Solution : Mettez la table dans un espace table à un

Messages d'erreur

seul noeud si l'espace table par défaut a provoqué cette erreur. Ou bien précisez un autre espace table à un seul noeud si vous avez indiqué un espace table à plusieurs noeuds.

CTE0233 Une commande d'administration conflictuelle est en cours de fonctionnement. Relancez cette commande ultérieurement.

Explication : Une autre commande d'administration est toujours en cours d'exécution ou a été arrêtée de manière anormale sans déverrouiller la commande.

Solution : Vérifiez à l'aide de la commande CONTROL LIST les verrous qui sont toujours actifs. Si un verrou est actif mais qu'aucune commande n'est en cours d'exécution, libérez le verrou manuellement à l'aide de la commande CONTROL CLEAR. Assurez-vous que quelqu'un d'autre puisse exécuter la commande d'administration en maintenant le verrou.

CTE0234 Une commande d'administration conflictuelle est en cours de fonctionnement sur l'index de texte. Relancez cette commande ultérieurement ou indiquez l'option FORCE d'une commande DISABLE DATABASE.

Explication : Une autre commande d'administration est toujours en cours d'exécution ou a été arrêtée de manière anormale sans déverrouiller la commande.

Solution : Vérifiez à l'aide de la commande CONTROL LIST les verrous qui sont toujours actifs. Si un verrou est actif mais qu'aucune commande n'est en cours d'exécution, libérez le verrou manuellement à l'aide de la commande CONTROL CLEAR. Assurez-vous que quelqu'un d'autre puisse exécuter la commande d'administration en maintenant le verrou. Dans le cas d'une commande DISABLE DATABASE, vous pouvez indiquer l'option FORCE qui arrête toutes les autres commandes dans cette base de données.

CTE0235 Impossible de trouver une licence valide pour DB2 Extension Net Search.

Explication : Impossible de trouver une licence valide pour DB2 Extension Net Search.

Solution : Vérifiez que la licence a été correctement installée avec db2lic. Vérifiez que les instances existantes sont mises à jour après l'installation du produit.

CTE0236 Seul Node0 est pris en charge par les instances MPP.

Explication : Les index de texte ne peuvent être créés que sur les instances MPP si la table comprenant la colonne de texte à indexer se trouve dans Node0.

Solution : Vérifiez le groupe de noeuds de l'espace table dans lequel l'espace table est défini.

CTE0237 Erreur interne : la table de journalisation "%1"."%2" comprend une opération non valide "%3".

Explication : La table de journalisation garde trace des opérations exécutées sur la table contenant la colonne d'index indexée. Il se peut que cette table soit corrompue car elle contient une entrée qui n'a pas été écrite par l'Extension Net Search.

Solution : Vérifiez la table de journalisation et supprimez l'entrée corrompue.

CTE0238 Erreur interne : la table "%1"."%2" comprend une expression incorrecte dans la colonne "%3".

Explication : La liste des expressions contient une erreur dans la colonne de texte indiquée.

Solution : Vérifiez le délimiteur de début et de fin.

CTE0239 Erreur interne : la longueur totale des propriétés d'index "%1" ne doit pas dépasser "%2".

Explication : La taille maximale des propriétés d'index (1016 octets) est dépassée. Les propriétés contiennent l'instance, l'index, le répertoire de travail ainsi que d'autres informations.

Solution : Assurez-vous que ces noms de chemin ne sont pas trop longs.

CTE0240 Erreur interne : la définition de la variable d'environnement "%1" a échoué.

Explication : La définition de la variable d'environnement indiquée a échoué. Il y a peut-être un problème avec la configuration de l'environnement.

Solution : Vérifiez les instructions spécifiques à votre système d'exploitation.

CTE0241 Erreur interne : la fonction UDF Datalink "%1"."%2" renvoie le type "%3". Le type attendu est : "%4".

Explication : La fonction UDF utilisée pour extraire le contenu d'une référence Datalink est définie avec un type de données imprévu.

Solution : Extrayez la définition de la fonction UDF Datalink à partir de SYSCAT.FUNCTIONS et signalez l'erreur aux services IBM compétents.

CTE0242 **La valeur "%1" du paramètre "%2" est incorrect.**

Explication : La procédure mémorisée de recherche ou la fonction de valeur table DB2EXT.TEXTSEARCH a été appelée avec des paramètres incorrects.

Solution : Corrigez les valeurs des paramètres de la procédure mémorisée de recherche ou la fonction de valeur table. Pour obtenir des paramètres corrects, reportez-vous à la documentation.

CTE0243 **La mémoire cache de l'index de texte "%1"."%2" n'a pas été activée.**

Explication : Une opération de l'Extension Net Search demande une mémoire cache activée. La mémoire cache n'est pas activée. Voici les raisons possibles :

- La mémoire cache n'a jamais été activée après la dernière commande DB2TEXT START.
- La mémoire cache a été explicitement désactivée avec la commande DB2TEXT DEACTIVATE CACHE.

Solution : Exécutez une commande DB2TEXT ACTIVATE CACHE pour l'index et relancez l'opération de l'Extension Net Search.

CTE0244 **Erreur interne : appel de "%1" renvoie rc="%2", SQLCODE="%3".**

Explication : Une erreur de traitement interne entrave la poursuite des opérations.

Solution : Si l'erreur persiste, exécutez un programme trace et vérifiez db2diag.log. Signalez l'erreur.

CTE0245 **La taille requise de la mémoire cache est supérieure à la taille de la mémoire cache disponible. Augmentez la taille de la mémoire cache maximale à une valeur > "%1" ou diminuez la valeur de pctfree.**

Explication : La taille de la mémoire cache nécessaire pour charger toutes les données dépasse la valeur MAXIMUM CACHE SIZE pour cet index. Elle peut être identifiée pendant l'activation de la mémoire cache (la commande DB2TEXT ACTIVATE) ou par la mise à jour de l'index lorsque la mémoire cache est activée.

Solution : Si l'erreur a été signalée dans la commande DB2EXT ACTIVATE, recalculez la taille de la mémoire cache maximale à l'aide de la fonction DB2EXT.MAXIMUM_CACHE_SIZE et modifiez le paramètre MAXIMUM CACHE SIZE de l'index. Diminuez la valeur PCTFREE. Si le nombre maximal de documents est dépassé pendant la mise à jour incrémentielle, reconstruisez la mémoire cache à l'aide des commandes db2 deactivate cache et db2text activate cache recreate.

CTE0246 **Le fichier "%1" est vide.**

Explication : Une commande DB2TEXT CREATE INDEX a échoué car le fichier du modèle de document indiqué dans la commande est vide.

Solution : Spécifiez un fichier de modèles de documents dans la commande.

CTE0247 **Une procédure enregistrée de DB2 Extension Net Search n'a pas pu être créée.**

Explication : Une commande DB2TEXT ENABLE DATABASE n'a pas pu créer la procédure mémorisée interne DB2EXT.CTESRVSP.

Solution : Vérifiez le message d'erreur DB2 supplémentaire qui est associé à une instruction CREATE PROCEDURE. Si l'erreur ne peut être corrigée en supprimant la procédure mémorisée existante avec un nom identique, lancez un programme trace et signalez l'erreur.

CTE0248 **La chaîne de recherche générée est trop longue. Réduisez la complexité de la requête de recherche.**

Explication : Une requête de l'Extension Net Search est trop longue ou trop complexe pour être traitée par le moteur de recherche de base. La complexité est touchée par les extensions du thésaurus, les expressions FUZZY FORM OF et les caractères de masquage.

Solution : Réduisez la complexité ou la longueur de la requête.

CTE0249 **Le programme exécutable "%1" a été terminé anormalement.**

Explication : Lors de l'exécution de la commande de l'Extension Net Search, le programme exécutable "%1" a été appelé mais arrêté de manière anormale.

Solution : Vérifiez que le programme exécutable a été arrêté explicitement par l'utilisateur, par exemple, un signal. Si ce n'est pas le cas, lancez une trace, réexécutez la commande et signalez l'erreur.

CTE0250 **Le type de retour "%1" de la fonction de transformation du type de colonne "%2"."%3" n'est pas pris en charge.**

Explication : Dans une commande DB2TEXT CREATE INDEX, une transformation de type colonne a été indiquée et renvoie un type de données qui n'est pas pris en charge. Les types de données pris en charge sont : CHARACTER, VARCHAR, LONG VARCHAR, CLOB, GRAPHIC, VARGRAPHIC, LONG VARGRAPHIC, DBCLOB, BLOB et DATALINK.

Solution : Choisissez une autre fonction de transformation du type de colonne.

Messages d'erreur

CTE0251 Erreur interne : le type de colonne "%1" n'est pas pris en charge.

Explication : Vous avez utilisé une colonne qui ne figure pas dans la liste des colonnes prises en charge.

Solution : Vérifiez la liste des colonnes appropriées pour les clés primaires et l'indexation dans l'index de création. Apportez les modifications appropriées à votre commande et réessayez. Si l'erreur persiste, exécutez un programme trace et vérifiez db2diag.log. Signalez l'erreur aux services d'assistance IBM.

CTE0252 Le paramètre "%1" est manquant.

Explication : Erreur interne - lors de l'exécution d'une commande de l'Extension Net Search, un programme exécutable d'administration a été appelé avec un paramètre manquant "%1".

Solution : Essayez de modifier les commandes de paramètre de l'Extension Net Search pour éviter un incident. Si l'erreur persiste, revenez à la fonction trace et transmettez l'erreur aux services d'assistance IBM.

CTE0253 Le document indiqué dans la vue journal n'a pas été trouvé.

Explication : Le contenu d'un document texte indiqué dans la vue journal a changé et son accès est impossible.

Solution : Vérifiez que le document existe et les droits d'écriture et de lecture des documents texte à inclure dans l'index.

CTE0254 La mémoire cache de l'index "%1" est déjà activée.

Explication : L'index a déjà été activé avec la commande ACTIVATE CACHE.

Solution : Vérifiez le nom de l'index indiqué et la base de données utilisée.

CTE0255 Un nom de colonne de l'expression de la colonne de résultats de la mémoire cache est manquant. Ajoutez "AS <nom colonne mémoire cache>" à l'expression.

Explication : Une expression de colonne de résultats de la mémoire cache doit porter un nom. Par exemple : 'C1+C2 AS monrésultat'.

Solution : Ajoutez "AS <nom colonne mémoire cache>" à l'expression.

CTE0256 La requête nécessaire à la sélection des données pour l'indexation a échoué. Réduisez la complexité de l'attribut, de la table de la mémoire cache ou des expressions sur ordre de recherche des résultats.

Explication : L'Extension Net Search crée une requête à partir des expressions dans votre commande afin de sélectionner des données pour l'indexation à partir de la base de données. La requête a échoué car elle était trop complexe.

Solution : Réduisez la complexité de l'attribut, de la table de la mémoire cache ou des expressions de l'ordre des résultats de la recherche.

CTE0257 Erreur lors de la création de la mémoire partagée.

Explication : Les ressources de la mémoire partagée n'ont pas pu être créées en raison d'une erreur précédente ou d'un problème de droits d'accès.

Solution : Reportez-vous à db2diag.log pour obtenir davantage d'informations ou nettoyez les ressources partagées. Reportez-vous à l'erreur CTE0183.

CTE0258 Erreur de la version de la mémoire partagée.

Explication : Impossible d'accéder aux ressources de la mémoire partagée car elles sont corrompues ou un conflit existe entre les versions.

Solution : Pour plus de détails, veuillez consulter db2diag.log. Désactivez et réactivez la base de données et réessayez.

CTE0259 Impossible d'insérer l'entrée dans la mémoire partagée globale. L'entrée existe déjà.

Explication : Une entrée à insérer dans la mémoire partagée globale existe en raison d'une erreur précédente.

Solution : Pour plus de détails, veuillez consulter db2diag.log. Recommencez la mise à jour et redémarrez les services de verrouillage à l'aide des commandes db2text stop et db2text start.

CTE0260 Impossible d'accéder à l'entrée dans la mémoire partagée globale. L'entrée est introuvable.

Explication : Une entrée à supprimer de la mémoire partagée globale n'existe pas en raison d'une erreur précédente.

Solution : Pour plus de détails, veuillez consulter db2diag.log. Tentez de recommencer la mise à jour et de redémarrer les services de verrouillage à l'aide des

commandes db2text stop et db2text start.

CTE0261 Il existe au moins une mémoire cache activée pour un index de texte dans cette instance. Désactivez la mémoire cache de tout index activé à l'aide de la commande DEACTIVATE CACHE ou utilisez l'option FORCE pour arrêter.

Explication : La commande db2text stop ne peut être utilisée que si vous exécutez une commande DEACTIVATE CACHE sur tous les index de texte qui ont été activés avec la commande ACTIVATE CACHE.

Solution : Désactivez la mémoire cache de tout index activé à l'aide de la commande DEACTIVATE CACHE ou utilisez l'option FORCE pour arrêter.

CTE0262 La valeur du paramètre "%1" est trop longue.

Explication : La valeur dépasse la taille autorisée maximale.

Solution : Vérifiez la taille maximale.

CTE0263 L'index de texte "%1"."%2" a été créé à l'aide de l'option RECREATE INDEX ON UPDATE. Dans ce contexte, UPDATE MINIMUM ou COMMITCOUNT FOR UPDATE ne doivent pas être spécifiés.

Explication : Le minimum de mises à jour et le nombre de validation pour la mise à jour ne sont effectifs que si l'index est mis à jour de manière incrémentielle.

Solution : Si vous voulez créer un index à chaque mise à jour, supprimez les paramètres UPDATE MINIMUM et COMMITCOUNT FOR UPDATE. Si vous voulez utiliser les paramètres UPDATE MINIMUM et COMMITCOUNT FOR UPDATE, ne précisez pas RECREATE INDEX ON UPDATE.

CTE0264 Des erreurs se sont produites lors de l'opération d'activation de l'index. Vérifiez la table d'événements "%1"."%2" et db2diag.log pour obtenir plus de détails.

Explication : Pendant l'activation des index, les erreurs détectées sont ajoutées à la table d'événements et au fichier db2diag.log.

Solution : Vérifiez la table d'événements pour plus d'informations sur les erreurs détectées dans le document. Rafraîchissez le journal des événements après résolution des incidents.

CTE0265 L'espace table de la table utilisateur ou l'espace table de l'administration ("%1") n'est pas définie sur le noeud 0 uniquement.

Explication : Si les index de texte sont créés dans les instances MPP, l'espace table de la table utilisateur doit se trouver dans Node0.

Solution : Utilisez une table dans laquelle l'espace table se trouve dans Node0.

CTE0266 ValueFrom "%1" doit être inférieur à ValueTo "%2".

Explication : Les valeurs indiquées lors de la recherche d'attribut ne sont pas corrects. Si la syntaxe de recherche est 'BETWEEN ValueFrom AND ValueTo', la limite la plus basse (ValueFrom) doit être inférieure à la limite la plus élevée (ValueTo).

Solution : Modifiez les limites de la clause 'BETWEEN ValueFrom AND ValueTo'.

CTE0267 Les objets de la base de données de l'Extension Net Search qui se trouvent dans la base de données "%1" sont dans un état incohérent.

Explication : Un objet de l'Extension DB2 Net Search est absent ou corrompu. La migration de la base de données ultérieure à l'installation d'une nouvelle version de l'Extension DB2 Net Search n'a pas eu lieu ou un utilisateur de la base de données a modifié ou annulé un ou des objets internes de l'Extension Net Search. Dans ce cas, tous les index de texte sont perdus et le mode texte doit être désactivé pour la base de données.

Solution : Pour procéder à la migration d'une base de données vers la version en cours, reportez-vous à la description du processus de migration mentionné dans la documentation de l'Extension DB2 Net Search. Il est également possible de lancer une commande DB2TEXT DISABLE DATABASE à l'aide de l'option FORCE. Vous pouvez ainsi réactiver le mode texte pour la base de données avec la commande DB2TEXT ENABLE DATABASE.

CTE0270 La table de journalisation "%1"."%2" n'a pas pu être modifiée après la mise à jour incrémentielle. Les entrées doivent être traitées au cours de la prochaine MISE A JOUR.

Explication : Un horodatage est créé lorsqu'une mise à jour de l'indexation incrémentée est lancée. Cet horodatage délimite ensuite le traitement des enregistrements de modifications. Les modifications en concurrence avec la mise à jour incrémentielle sont traitées ultérieurement, à l'occasion de la mise à jour suivante. Dans certaines situations, il est possible que

Messages d'erreur

des modifications apportées à certaines transactions ne soient pas validées au début de la mise à jour et le soient durant la mise à jour de l'index. Cette situation risque d'engendrer des incohérences.

Pour éviter une telle situation, les enregistrements de modifications effectués avant la limite indiquée par l'horodatage ne sont pas supprimés de la table de journalisation, même lorsqu'ils ont été partiellement traités. Lors de la mise à jour incrémentielle suivante, ces modifications seront appliquées une nouvelle fois à l'index.

Solution : Lors de la mise à jour d'index suivante, ces modifications sont appliquées une nouvelle fois à l'index. En cas d'opérations de suppression, le message d'erreur suivant peut s'afficher :

CTE0101:ItlEnReasonCode_Docmap_docid _not_found.

Notez que cette erreur peut être ignorée car le document a déjà été supprimé. Si les erreurs CTE0270 se produisent fréquemment, il est conseillé de supprimer et de re-crée l'index en lui attribuant un seuil d'horodatage différent pour la mise à jour de l'index. Par exemple : db2text "CREATE INDEX ... INDEX CONFIGURATION(UPDATEDELAY 30)"

Ce message signifie que le traitement effectué durant une mise à jour incrémentielle prend en compte uniquement les enregistrements de modifications datant de plus de 30 secondes et évite ainsi les interférences avec les transactions de modifications concurrentes datant de moins de 30 secondes.

CTE0273 **La mémoire cache de l'index "%1", "%2" est déjà activée.**

Explication : L'index a déjà été activé avec la commande ACTIVATE CACHE.

Solution : Vérifiez le nom de l'index indiqué et la base de données utilisée.

CTE0274 **Le système de base de données cible "%1" pour la connexion n'est pas pris en charge.**

Explication : Vous avez tenté d'exécuter une commande DB2TEXT alors qu'une connexion à un système de bases de données non pris en charge par l'Extension DB2 Net Search était active.

CTE0275 **Les informations concernant le type et la version du serveur "%1" n'ont pas été trouvées.**

Explication : Les informations concernant le type et la version du serveur n'ont pas été détectées dans la vue de catalogue DB2 'SERVEURS'.

Solution : Assurez-vous que l'environnement fédéré DB2 est correctement configuré.

CTE0276 **Impossible de localiser la fonction UDF Datalink.**

Explication : La fonction UDF Datalink n'est pas enregistrée en tant que fonction définie par l'utilisateur au sein de la base de données.

Solution : Exécutez la commande 'db2 -tvf ctedlcud.ddl' dans le répertoire db2ext/ddl.

CTE0277 **Impossible d'attacher un segment de mémoire cache.**

Explication : Le système ne peut pas allouer un espace mémoire suffisant pour charger un segment de mémoire cache de taille importante ou le segment de mémoire cache ne peut pas être ouvert car il a été supprimé.

Solution : Vérifiez les paramètres de votre système et augmentez l'espace de pagination ainsi que la mémoire disponible. Une préparation du système peut s'avérer nécessaire lorsque la mémoire cache est de taille importante. Reportez-vous à la documentation de l'Extension DB2 Net Search. Utilisez les commandes DEACTIVATE et ACTIVATE [RECREATE] pour re-crée la mémoire cache. Si le problème persiste, consultez le fichier db2diag.log pour obtenir des informations supplémentaires.

CTE0278 **Sur un système AIX 32 bits, il convient de modifier le paramètre MAXDATA avant d'activer une mémoire cache de taille importante.**

Explication : Lorsque vous utilisez la procédure mémorisée de recherche sur un système AIX 32 bits, il est possible que vous deviez remplacer le paramètre MAXDATA par l'exécutable db2fmp.

Solution : Reportez-vous à la documentation de l'Extension DB2 Net Search pour plus d'informations sur la modification du paramètre MAXDATA.

CTE0279 **La taille des données placées en mémoire cache a atteint la limite fixée par le système.**

Explication : En réduisant la valeur PCTFREE, il est possible d'augmenter la taille de données au cours de l'activation de la mémoire cache. Ainsi, le système peut réduire l'espace disponible réservé dans la mémoire cache.

Solution : Réduisez la valeur PCTFREE ou diminuez la taille des données de texte à placer en mémoire cache. Utilisez les commandes DEACTIVATE et ACTIVATE [RECREATE] pour re-crée la mémoire cache.

CTE0280 Espace disque insuffisant pour procéder à l'écriture des fichiers cache persistants.

Explication : Le système ne peut pas écrire un fichier de taille suffisante pour la mémoire cache persistante du répertoire cache.

Solution : Remplacez le répertoire cache persistant par un système de fichiers vide à l'aide de la commande ALTER INDEX. Il est également possible de réduire la taille de la mémoire cache en diminuant les valeurs PCTFREE ou MAXIMUM CACHE SIZE ou en ayant recours à une mémoire cache temporaire.

CTE0281 Echec de la suppression du fichier cache persistant "%1".

Explication : Le fichier n'existe pas ou il n'est pas possible d'y accéder.

Solution : Vérifiez que le fichier existe et supprimer-le manuellement. ";

CTE0282 Le nombre de documents placés en mémoire cache a atteint la limite fixée par le système.

Explication : En réduisant la valeur PCTFREE, il est possible d'augmenter le nombre maximal d'entrées de document à placer en mémoire cache au cours de l'activation de la mémoire cache. Ainsi, le système peut réduire l'espace disponible réservé dans la mémoire cache.

Solution : Réduisez la valeur PCTFREE ou le nombre d'entrées de document placées en mémoire cache. Utilisez les commandes DEACTIVATE et ACTIVATE [RECREATE] pour re-crée la mémoire cache.

CTE0283 Impossible de créer un segment de mémoire cache.

Explication : Le système ne peut pas allouer un espace mémoire suffisant pour charger un segment de mémoire cache de taille importante dans la mémoire cache. En réduisant la valeur PCTFREE value, vous diminuez la taille du segment de mémoire cache.

Solution : Vérifiez les paramètres de votre système et augmentez l'espace de pagination ainsi que la mémoire disponible. Vous pouvez également réduire la taille de la mémoire cache en utilisant une valeur PCTREE inférieure. Une préparation du système peut s'avérer nécessaire lorsque la mémoire cache est de taille importante. Reportez-vous à la documentation de l'Extension DB2 Net Search. Utilisez les commandes DEACTIVATE et ACTIVATE [RECREATE] pour re-crée la mémoire cache. Si le problème persiste, consultez le fichier db2diag.log pour obtenir des informations supplémentaires.

CTE0284 L'index de texte se situe sur le noeud "%1", alors que la fonction de recherche a été appliquée au niveau du noeud "%2".

Explication : La procédure mémorisée de recherche ou la fonction de valeur table DB2EXT.TEXTSEARCH n'a pas été appelée sur le noeud où se trouve l'index. La fonction de recherche ne sera pas transmise automatiquement au noeud appropriée.

Solution : Définissez la variable d'environnement DB2NODE sur le noeud où l'index est connecté avant d'établir une connexion avec la base de données.

CTE0285 La fonction de recherche n'est pas autorisée dans le cas d'index de texte transmis à de multiples noeuds.

Explication : La fonction de valeur table DB2EXT.TEXTSEARCH ne doit pas être appelée avec des index transmis à plusieurs noeuds. En effet la fonction ne sera pas transmise automatiquement aux noeuds appropriés, mais exécutée au niveau du noeud coordinateur.

Solution : Utilisez la fonction CONTAINS, SCORE ou NUMBEROFMATCHES dans un environnement comportant plusieurs noeuds.

CTE0286 Aucune ligne détectée au niveau de "%1". "IBMSNAP_REGISTER" pour la table source "%2". "%3" et la table de modification des enregistrements "%4". "%5".

Explication : Aucune entrée valable n'a été détectée dans la table IBMSNAP_REGISTER pour les caractéristiques de la table de capture de réplication spécifiée dans la commande DB2TEXT CREATE INDEX. Une entrée correcte doit contenir la table source correspondant à l'index dans les colonnes SOURCE_OWNER et SOURCE_NAME, avec SOURCE_VIEW_QUAL=0 et la table de capture de réplication indiquée dans les colonnes PHYS_CHANGE_OWNER et PHYS_CHANGE_TABLE.

Causes possibles : La table source spécifiée n'était pas enregistrée en tant que source de réplication pour la table de capture de réplication.

Solution : Enregistrez la table source correctement pour DB2 Replication ou indiquez une table de capture de réplication correcte pour la table source.

CTE0287 Valeur incorrecte "%1" pour "%2" dans "%3". "IBMSNAP_REGISTER" pour la table source "%4". "%5" et la table de modification des enregistrements "%6". "%7".

Explication : Un paramètre de réplication détecté dans la table IBMSNAP_REGISTER n'est pas autorisé.

Messages d'erreur

Causes possible : 1.La colonne CHG_UPD_TO_DEL_INS ne contient pas la valeur 'Y'.
2.La colonne CCD_CONDENSED contient la valeur 'Y'.

Solution : Lors de l'enregistrement de la table source pour DB2 Replication, assurez-vous que les opérations de mise à jour ont été transformées en paires d'opérations de suppression et d'insertion. Vérifiez, en outre, qu'aucune table de capture de réplication n'est utilisée.

CTE0288 **La table source "%1"."%2" et la table de modification des enregistrements "%3"."%4" se trouvent sur des serveurs différents ("%5" et "%6").**

Explication : La table source et la table de capture de réplication indiquées doivent résider sur un même serveur.

CTE0289 **L'encapsuleur "%1" n'est pas pris en charge.**

Explication : L'encapsuleur n'est pas pris en charge. Reportez-vous à la documentation de l'Extension DB2 Net Search pour obtenir la liste des encapsuleurs pris en charge. Pour plus d'informations, voir page 122.

CTE0290 **L'alias "%1"."%2" n'est pas autorisé au sein de la clause de réplication.**

Explication : Vous n'avez pas l'autorisation d'indiquer un alias pour alias de type nickname dans une clause de réplication.

Solution : Indiquez l'alias de type nickname au lieu de l'alias ou créez un nouvel alias de type nickname pour la table éloignée.

CTE0360 **Un message d'erreur spécifique s'affiche**

Explication : Ce numéro de message est réservé aux erreurs de l'Extension Net Search.

Solution : Répondre au message d'erreur spécifique.

CTE0451 **Le format de document "%1" spécifié n'est pas pris en charge par la fonction UDF de mise en évidence.**

Explication : Le format de document "%1" ne prend pas en charge la mise en évidence.

Solution : Utilisez un format de document prenant en charge la fonction UDF de mise en évidence.

CTE0452 **Erreur de syntaxe à proximité de l'option "%1" dans la fonction UDF de mise en évidence.**

Explication : Vous avez indiqué une syntaxe incorrecte à proximité de l'option spécifiée.

Solution : Vérifiez la syntaxe à proximité de l'option %1. Corrigez-la et faites une nouvelle tentative.

CTE0453 **La taille de renvoi de la fonction UDF de mise en évidence est trop réduite.**

Explication : Les éléments requis du document mis en évidence ne correspondent pas au paramètre renvoyé de la fonction UDF de mise en évidence.

Solution : Réduisez le nombre de fenêtres, la taille des fenêtre et/ou le nombre de sections à partir desquelles les occurrences doivent être affichées. De cette manière, le nombre de documents renvoyés à l'utilisateur sera réduit.

CTE0454 **Erreur de conversion des paramètres de la fonction UDF de mise en évidence de la page de codes "%1" vers UTF8.**

Explication : Les paramètres de la fonction UDF de mise en évidence du CCSID spécifié (il peut s'agir du CCSID de la base de données par défaut), ne peuvent pas être convertis en UTF8.

Solution : Entrez les indications correctes pour le fichier modèle CCSID.

CTE0455 **La page de codes de la base de données "%1" n'est pas prise en charge par la fonction UDF de mise en évidence.**

Explication : La page de code de la base de données n'est pas prise en charge par la fonction UDF de mise en évidence.

CTE0456 **La fonction UDF de mise en évidence prend uniquement en charge les documents enregistrés avec la page de codes UTF8.**

Explication : Seuls les documents en page de codes UTF8 prennent en charge la fonction UDF de mise en évidence.

CTE0457 **La valeur "%1" du paramètre "%2" n'est pas valable dans la fonction UDF de mise en évidence.**

Explication : L'une des valeurs du paramètre de mise en évidence n'est pas valable.

Solution : Vérifiez la valeur du paramètre et assurez-vous que cette valeur est autorisée dans la plage de données.

CTE0841 **L'option de commande "%1" est manquante.**

Explication : Une option de commande requise n'a pas été spécifiée.

Solution : Vérifiez les paramètres spécifiés et ajoutez le paramètre manquant.

CTE0842 **Aucune valeur n'est indiquée pour l'option de commande "%1".**

Explication : Une valeur requise pour une option de commande n'a pas été spécifiée.

Solution : Vérifiez les paramètres spécifiés et ajoutez l'option manquante.

CTE0843 **Aucune valeur numérique n'est indiquée pour l'option de commande "%1".**

Explication : Une chaîne a été spécifiée à la place d'un nombre.

Solution : Vérifiez les paramètres spécifiés et remplacez la chaîne par le nombre approprié.

CTE0844 **Le chemin d'accès du fichier de définitions "%1" est trop long.**

Explication : Le chemin d'accès est trop long et n'a pas pu être traité.

Solution : Utilisez un chemin plus court puis réessayez.

CTE0845 **Aucun fichier de définition spécifié.**

Explication : Le fichier de définitions doit être spécifié.

Solution : Ajoutez un nom de définition correct et réessayez.

CTE0846 **Le nom du fichier de définitions "%1" est trop long.**

Explication : Le nom du fichier de définitions est trop long.

Solution : Diminuez la longueur du nom du fichier de définition pour qu'il soit conforme aux critères de taille autorisés.

CTE0847 **Le fichier de définitions "%1" n'existe pas.**

Explication : Le fichier de définitions n'a pas pu être utilisé.

Solution : Vérifiez que le fichier de définitions se trouve à l'emplacement approprié et que l'utilisateur en cours peut y accéder.

CTE0849 **Le fichier dictionnaire "%1" n'a pas pu être verrouillé.**

Explication : Le processus n'est pas parvenu à verrouiller le fichier dictionnaire. Vous ne bénéficiez pas des droits en écriture ou le fichier a été ouvert par un autre processus afin de procéder à l'écriture de données.

Solution : Contrôlez les processus afin de détecter les processus susceptibles de verrouiller le fichier dictionnaire et vérifiez les droits d'accès qui vous sont attribués.

CTE0850 **Le fichier de sortie "%1" existe déjà.**

Explication : Le fichier de sortie n'a pas pu être écrasé.

Solution : Vérifiez que vous pouvez créer le thésaurus dans le répertoire spécifié.

CTE0851 **Le dictionnaire %1 est entièrement perdu.**

Explication : Les fichiers du dictionnaire Thésaurus sont endommagés.

Solution : Nettoyez le répertoire et compilez une nouvelle fois le fichier de définitions.

CTE0852 **Erreur de version dans le fichier dictionnaire "%1".**

Explication : Le fichier dictionnaire a été généré à l'aide d'une version plus ancienne du compilateur du thésaurus.

Solution : Compilez votre fichier de définitions une nouvelle fois à l'aide de la version en cours du compilateur du thésaurus.

CTE0853 **Le dictionnaire existant "%1" ne peut pas être écrasé.**

Explication : Il est impossible d'écraser le dictionnaire existant.

Solution : Vérifiez que vous disposez des droits en écriture sur le fichier dictionnaire, l'emplacement du répertoire et celui du sous-répertoire.

CTE0855 **Un terme du thésaurus est spécifié de manière incorrecte.**

Explication : Le fichier de définitions contient une erreur de syntaxe.

Solution : Consultez la documentation de l'Extension DB2 Net Search pour obtenir des informations sur le fichier de définitions du thésaurus et la prise en charge sur thésaurus.

Messages d'erreur

CTE0856 **Le fichier de définition "%1" est vide.**

Explication : Le fichier de définitions ne doit pas être vide.

Solution : Consultez la documentation de l'Extension DB2 Net Search pour obtenir des informations sur le fichier de définitions du thésaurus et la prise en charge sur thésaurus.

CTE0857 **Aucune ligne de départ de bloc n'a été trouvée dans le fichier "%1", à la ligne "%2".**

Explication : Le fichier de définitions contient une erreur de définition.

Solution : Un bloc doit débuter par 'WORDS'. Consultez la documentation de l'Extension DB2 Net Search pour obtenir des informations sur les concepts du thésaurus.

CTE0858 **Une relation incorrecte a été indiquée dans le fichier "%1", à la ligne "%2".**

Explication : Le fichier de définitions contient une erreur de définition.

Solution : Examinez la "définition du terme associé". Consultez la documentation de l'Extension DB2 Net Search pour obtenir des informations sur le fichier de définitions du thésaurus.

CTE0859 **Le numéro de relation est hors plage dans le fichier "%1", à la ligne "%2".**

Explication : Toutes les relations définies par l'utilisateur sont basées sur le type associatif. Elles sont identifiées par des numéros uniques compris entre 1 et 128.

Solution : Vérifiez les nombres associés à la relation.

CTE0861 **Aucun terme n'est défini dans le fichier "%1", à la ligne "%2".**

Explication : Les termes requis n'ont pas été spécifiés.

Solution : Consultez la documentation de l'Extension DB2 Net Search pour obtenir des informations sur le fichier de définitions du thésaurus.

CTE0861 **Le terme du thésaurus qui se trouve dans le fichier "%1", à la ligne "%2" est trop long.**

Explication : La longueur du terme du thésaurus est limitée à 64 octets.

Solution : Modifiez le terme du thésaurus et réessayez.

CTE0862 **La valeur du paramètre STRENGTH est spécifiée de manière incorrecte dans le fichier "%1" à la ligne "%2".**

Explication : Le fichier de définitions contient une erreur de syntaxe.

Solution : Consultez la documentation de l'Extension DB2 Net Search pour obtenir des informations sur le fichier de définitions du thésaurus et la prise en charge sur thésaurus.

CTE0863 **La valeur du paramètre STRENGTH est hors plage dans le fichier "%1" à la ligne "%2".**

Explication : La valeur STRENGTH doit être comprise entre 1 et 100.

Solution : Modifiez la valeur STRENGTH de sorte que sa valeur numérique soit comprise entre 1 et 100.

CTE0864 **Erreur interne : Le compilateur du thésaurus a échoué avec le code anomalie "%1".**

Explication : Une erreur de traitement interne entrave la poursuite des opérations. Essayez de démarrer et d'arrêter la mise à jour, les services de verrouillage et DB2.

Solution : Si l'erreur persiste, exécutez un programme trace et vérifiez db2diag.log.

CTE0865 **Le répertoire "%1" n'a pas pu être créé.**

Explication : Le répertoire spécifié n'a pas pu être créé.

Solution : Assurez-vous que le répertoire existe déjà et vérifiez les droits d'accès au répertoire.

CTE0866 **Le répertoire "%1" n'a pas pu être supprimé.**

Explication : Le répertoire n'a pas pu être supprimé.

Solution : Vérifiez que vous disposez des droits d'accès au répertoire spécifié.

Annexe G. Référence du modèle de document

DB2 Extension Net Search fournit des informations de référence sur les modèles de document :

- Définition de type de document pour des modèles de document
- Sémantique des expressions du pointeur (XPath)
- Limite pour les zones de texte et les attributs de document
- Valeurs de l'attribut de marque Outside-In

Définition de type de document pour des modèles de document

Voici une description formelle de la syntaxe des modèles de document sous forme d'une définition de type de document:

```
<!ELEMENT GPPModel (GPPFieldDefinition|GPPAttributeDefinition)+>
<!ELEMENT HTMLModel (HTMLFieldDefinition|HTMLAttributeDefinition)+>
<!ELEMENT XMLModel (XMLFieldDefinition|XMLAttributeDefinition)+>

<!ELEMENT GPPFieldDefinition EMPTY>
<!ATTLIST GPPFieldDefinition name CDATA #REQUIRED>
<!ATTLIST GPPFieldDefinition start CDATA #REQUIRED>
<!ATTLIST GPPFieldDefinition end CDATA #IMPLIED>
<!ATTLIST GPPFieldDefinition exclude (YES|NO) NO>

<!ELEMENT GPPAttributeDefinition EMPTY>
<!ATTLIST GPPAttributeDefinition name CDATA #REQUIRED>
<!ATTLIST GPPAttributeDefinition start CDATA #REQUIRED>
<!ATTLIST GPPAttributeDefinition end CDATA #REQUIRED>
<!ATTLIST GPPAttributeDefinition type NUMBER #REQUIRED>

<!ELEMENT HTMLFieldDefinition EMPTY>
<!ATTLIST HTMLFieldDefinition name CDATA #REQUIRED>
<!ATTLIST HTMLFieldDefinition tag CDATA #REQUIRED>
<!ATTLIST HTMLFieldDefinition meta-qualifier CDATA #IMPLIED>
<!ATTLIST HTMLFieldDefinition exclude (YES|NO) NO>

<!ELEMENT HTMLAttributeDefinition EMPTY>
<!ATTLIST HTMLAttributeDefinition name CDATA #REQUIRED>
<!ATTLIST HTMLAttributeDefinition tag CDATA #REQUIRED>
<!ATTLIST HTMLAttributeDefinition meta-qualifier CDATA #IMPLIED>
<!ATTLIST HTMLAttributeDefinition type NUMBER #REQUIRED>

<!ELEMENT XMLFieldDefinition EMPTY>
<!ATTLIST XMLFieldDefinition name CDATA #REQUIRED>
<!ATTLIST XMLFieldDefinition locator CDATA #REQUIRED>
<!ATTLIST XMLFieldDefinition ignore (YES|NO) NO>
<!ATTLIST XMLFieldDefinition priority CDATA #IMPLIED>
<!ATTLIST XMLFieldDefinition exclude (YES|NO) NO>

<!ELEMENT XMLAttributeDefinition EMPTY>
<!ATTLIST XMLAttributeDefinition name CDATA #REQUIRED>
<!ATTLIST XMLAttributeDefinition locator CDATA #REQUIRED>
<!ATTLIST XMLAttributeDefinition ignore (YES|NO) NO>
<!ATTLIST XMLAttributeDefinition priority CDATA #IMPLIED>
<!ATTLIST XMLAttributeDefinition type NUMBER #REQUIRED>
```

Sémantique des expressions du pointeur (XPath)

Selon le modèle de données XML, les documents XML sont affichés comme des arborescences contenant les types de noeuds suivants :

- Le noeud racine (root)
- Noeuds d'élément
- Noeuds de type texte
- Noeuds d'attribut
- Noeuds d'espace nom
- Noeuds d'instruction de traitement
- Noeuds de commentaire

Les liaisons entre ces noeuds ou, en d'autres termes, la relation formant l'arborescence, reflète la relation de confinement immédiate dans le document XML.

Le **noeud racine (root)** peut uniquement apparaître à la racine et nulle part ailleurs dans l'arborescence. Il comporte, tout comme ses enfants, l'élément de document, les commentaires facultatifs et les instructions de traitement.

Les **noeuds d'éléments** peuvent comporter tous les types de noeuds, sauf le noeud racine (root). Les autres types de noeuds sont uniquement autorisés sur les noeuds terminaux de l'arborescence.

Il existe trois types de **liaisons de confinement** : 'enfant', 'attribut' et 'espace de nom'. Les liaisons de confinement 'attribut' et 'espace de nom' doivent respectivement mener aux noeuds d'attribut et d'espace nom. En d'autres mots, pour accéder aux enfants d'un noeud d'élément (en termes de théorie graphique), vous devez suivre les liaisons 'attribut' pour trouver tous les attributs contenus, suivre les liaisons 'espace nom' pour trouver toutes les déclarations d'espace nom contenues et suivre les liaisons 'enfant' pour trouver les éléments contenus, les noeuds de type texte, les instructions de traitement et les commentaires.

Une expression XPath doit être interprétée conformément à un noeud de contexte et elle décrit un ensemble de noeuds. Lorsqu'il est utilisé comme modèle de sélecteur Extension Net Search, le noeud de contexte est libre, c'est-à-dire qu'un modèle de chemin relatif p est interprété comme étant //p.

Voici les modèles de sélecteurs XPath Extension Net Search :

- Pattern '|' LocationPathPattern dans un contexte N indique l'union des noeuds qui correspondent à Pattern et LocationPathPattern, tous deux dans un contexte N.
- '/' RelativePathPattern dans un contexte N a la même signification que RelativePathPattern dans le contexte racine (root).
- '//' RelativePathPattern dans un contexte N indique l'union des significations de ce RelativePathPattern interprété dans tout contexte descendant de la racine (sur l'axe enfant).
- RelativePathPattern '/' StepPattern correspond à un noeud dans un contexte N, si et seulement si ce noeud correspond à StepPattern dans le contexte de ses parents et si son noeud parent correspond à RelativePathPattern dans le contexte N.

- RelativePathPattern '/' StepPattern correspond à un noeud dans un contexte N, si et seulement si ce noeud correspond à StepPattern dans le contexte de ses parents, et s'il a un noeud ancêtre qui correspond à RelativePathPattern dans un contexte N.
- 'child'::NodeTest (syntaxe abrégée : NodeTest) dans un contexte N correspond à un noeud qui est un enfant de N (sur l'axe enfant) et qui satisfait NodeTest.
- 'attribute'::NodeTest (syntaxe abrégée : @NodeTest) dans un contexte N correspond à un noeud qui est un attribut de N et qui satisfait NodeTest.
- NodeType '(' ')' est satisfait pour un noeud si et seulement si il est du type spécifié.
- 'processing-instruction' '(' Literal ')' est satisfait pour tout noeud de type instruction-traitement (processing-instruction-type) qui possède Literal dans son nom.
- '*' est satisfait pour tout élément ou noeud d'attribut (masque de nom pour un nom d'élément).
- NCName ':' '*' est satisfait pour tout noeud d'élément possédant NCName comme préfixe de nom.
- QName est satisfait pour tout noeud avec un nom spécifié.

Remarque

Un NodeTest de la forme NameTest suppose que le noeud est de type principal sur l'axe sélectionné, soit le type d'attribut sur l'axe d'attribut ou le type enfant sur l'axe enfant. Par conséquent, NameTest ne peut pas être utilisé pour choisir des commentaires ou des noeuds d'instruction-traitement mais seulement des noeuds enfant et attribut. Par ailleurs, les modèles permettent la sélection de tous les types de noeuds, sauf les noeuds d'espace de nom, car le spécificateur d'axe 'espace de nom' n'est pas autorisé.

Exemples de modèles :

- chapter | appendix désigne tous les éléments de chapitres et les éléments d'annexes
- table désigne tous les éléments de table
- * désigne tous les éléments (notez qu'il s'agit de l'abréviation pour enfant::*)
- ulist/item désigne tous les éléments d'articles possédant un parent ulist
- appendix//subsection désigne tous les éléments de sous-section avec un ancêtre d'annexe
- / désigne le singleton contenant uniquement le noeud racine (root)
- comment() désigne tous les noeuds de commentaires
- processing-instruction() désigne toutes les instructions de traitement
- attribute::* (ou @*) désigne tous les noeuds d'attribut

Voici la syntaxe d'un élément de releveur de coordonnées :

```
Locator      ::= LocationPathPattern
              | Locator '|' LocationPathPattern
LocationPathPattern ::= '/' RelativePathPattern ?
              | '//'? RelativePathPattern
RelativePathPattern ::= StepPattern
                    | RelativePathPattern '/' StepPattern
                    | RelativePathPattern '//' StepPattern
StepPattern      ::= ChildOrAttributeAxisSpecifier NodeTest
ChildOrAttributeAxisSpecifier ::=
```

```
      ('child' | 'attribute') '::'
      | '@'?
NodeTest ::= NameTest
          | NodeType '(' ')'
          | 'processing-instruction' '(' Literal ')'
NameTest ::= '*' | NCName ':' '*' | QName
NodeType ::= 'comment' | 'processing-instruction'
```

NCName et QName sont tels que définis dans les Recommandations de dénomination :

NCName

Nom XML ne contenant pas de signe deux-points.

QName

NCName qui peut être précédé d'un NCName suivi de deux-points. Par exemple : NCName:NCName

Limites des zones de texte et des attributs de document

Voici une liste des limites pour les zones de texte et les attributs de document :

- Nombre maximal de zones dans un index : 32767
- Nombre maximal de valeurs pour un attribut de type STRING dans un document : 1024
- Nombre maximal d'attributs de type STRING : 253
- Le nombre de caractères dans une valeur d'attribut STRING est tronqué à 128
- Nombre maximal d'attributs de types DATE et NUMBER : 32766
- Le nombre de caractères dans une valeur d'attribut DATE ou NUMBER est tronqué à 128
- Pour des attributs NUMBER, un nombre à virgule flottante à double précision est accepté comme valeur.
- Nombre maximal de valeurs pouvant être spécifiées pour un attribut de type DATE ou NUMBER dans un document : sans limite

Des balises peuvent être incluses dans un modèle de document HTML :

- <A>
- <ADDRESS>
- <AU>
- <AUTHOR>
- <H1>
- <H2>, <H3>, <H4>, <H5>
- <H6>
- <TITLE>

Les balises, telles que <HEAD> et <BODY> pouvant comporter d'autres balises ne peuvent pas être spécifiées dans un modèle de document HTML comme zone de texte.

Valeurs de l'attribut de marque Outside-In

Valeurs possibles pour l'attribut de marque relatif aux types de balises des propriétés du document Outside-In :

```
SCCCA_ABSTRACT
SCCCA_ACCOUNT
SCCCA_ADDRESS
```


SCCC_A ATTACHMENTS
 SCCC_A AUTHORIZATION
 SCCC_A BACKUPDATE
 SCCC_A BASEFILELOCATION
 SCCC_A BILLTO
 SCCC_A BLINDCOPY
 SCCC_A CARBONCOPY
 SCCC_A CATEGORY
 SCCC_A CHECKEDBY
 SCCC_A CLIENT
 SCCC_A COMPANY
 SCCC_A COMPLETEDDATE
 SCCC_A COUNTCHARS
 SCCC_A COUNTPAGES
 SCCC_A COUNTWORDS
 SCCC_A CREATIONDATE
 SCCC_A DEPARTMENT
 SCCC_A DESTINATION
 SCCC_A DISPOSITION
 SCCC_A DIVISION
 SCCC_A DOCCOMMENT
 SCCC_A DOCTYPE
 SCCC_A EDITMINUTES
 SCCC_A EDITOR
 SCCC_A FORWARDTO
 SCCC_A GROUP
 SCCC_A KEYWORD
 SCCC_A LANGUAGE
 SCCC_A LASTPRINTDATE
 SCCC_A LASTSAVEDBY
 SCCC_A MAILSTOP
 SCCC_A MANAGERSCCCA_MATTER
 SCCC_A OFFICE
 SCCC_A OPERATOR
 SCCC_A OWNER
 SCCC_A PRIMARYAUTHOR
 SCCC_A PROJECT
 SCCC_A PUBLISHER
 SCCC_A PURPOSE
 SCCC_A RECEIVEDFROM
 SCCC_A RECORDEDDBY
 SCCC_A RECORDEDDATE
 SCCC_A REFERENCE
 SCCC_A REVISIONDATE
 SCCC_A REVISIONNOTES
 SCCC_A REVISIONNUMBER
 SCCC_A SECONDARYAUTHOR
 SCCC_A SECTION
 SCCC_A SECURITY
 SCCC_A SOURCE
 SCCC_A STATUS
 SCCC_A SUBJECT
 SCCC_A TITLE
 SCCC_A TYPIST
 SCCC_A USERDEFINEDPROP
 SCCC_A VERSIONDATE
 SCCC_A VERSIONNOTES
 SCCC_A VERSIONNUMBER

Valeurs possibles pour l'attribut de marque relatif aux sous-types de balises de début et de fin Outside-In :

SCCC_A ALTFONTDATA
 SCCC_A ANNOTATIONREFERENCE
 SCCC_A CAPTIONTEXT
 SCCC_A CHARACTER
 SCCC_A COMPILEDFIELD

Langues

SCCCA_COUNTERFORMAT
SCCCA_CUSTOMDATAFORMAT
SCCCA_DATEDEFINITION
SCCCA_DOCUMENTPROPERTYNAME
SCCCA_ENDNOTEREFERENCE
SCCCA_FONTANDGLYPHDATA
SCCCA_FOOTNOTEREFERENCE
SCCCA_FRAME
SCCCA_GENERATEDFIELD
SCCCA_GENERATOR
SCCCA_HYPERLINK
SCCCA_INDEX
SCCCA_INDEXENTRY
SCCCA_INLINEDATAFORMAT
SCCCA_LISTENTRY
SCCCA_MERGEENTRY
SCCCA_NAMEDCELLRANGE
SCCCA_REFERENCEDTEXT
SCCCA_STYLE
SCCCA_SUBDOCTEXT
SCCCA_TOA
SCCCA_TOAENTRY
SCCCA_TOC
SCCCA_TOCENTRY
SCCCA_TOF
SCCCA_VECTORSAVETAG
SCCCA_XREF

Veuillez noter que les tables incluent toute propriété du document ainsi que tous les sous-types de balises reconnus par les filtres INSO. Il y a deux exceptions du sous-type : SCCCA_DOCUMENTPROPERTY et SCCCA_BOOKMARK.

Annexe H. Moteur de recherche de texte

DB2 Extension Net Search fournit les informations suivantes sur le moteur de recherche de texte :

- Tokenisation
- Mots vides
- Configuration

Tokenisation

Lors de l'indexation, l'Extension Net Search traite le texte du document de la manière suivante, en le divisant en plusieurs entités.

Mots

Tous les caractères alphanumériques ("a".."z,"A".."Z", "0".."9") sont utilisés pour créer un index en texte intégral. Les caractères de séparation sont les caractères vides et les caractères décrits dans la rubrique relative à la reconnaissance de phrases ci-après. Les caractères non imprimables, tels que les retours ligne (aussi appelé caractère de retour de ligne) et les caractères vides sont interprétés de la manière suivante : les caractères non imprimables (inférieurs à 0x20) en milieu de ligne sont considérés comme des caractères vides. Les caractères vides et les caractères non imprimables situés avant et après un retour de ligne (0x0A) sont ignorés. Les retours de ligne situés avant et après un caractère d'un octet sont considérés comme des caractères vides et les caractères de 2 octets pour le même caractère sont toujours considérés comme étant les mêmes caractères. Les lettres majuscules et minuscules pour le même caractère, par exemple, "A" et "a" sont considérés comme étant les mêmes caractères si aucune spécification particulière n'a été faite lors de la recherche ou ils sont considérés comme caractères différents si vous recherchez des correspondances exactes.

Phrases

L'Extension Net Search détecte les fins de phrases de la manière suivante : ".", "!", "?", suivis par des caractères vides et le point final japonais et chinois à la fin d'une ligne.

Paragraphes

La reconnaissance des paragraphes dépend du format du document. Dans le format Texte en clair, deux caractères de nouvelle ligne consécutifs (éventuellement avec un retour chariot) sont considérés comme une limite de paragraphe. En HTML, le code de paragraphe <p> est interprété comme limite de paragraphe. Les autres formats de document ne prennent pas en charge la reconnaissance de paragraphe.

Mots vides

Les mots vides sont des mots dont la fréquence d'apparition est trop élevée et dont le contenu n'est pas pertinent pour le processus d'extraction. Généralement, tous les mots de liaisons et prépositions sont considérés comme étant des mots vides, par exemple, "et", "ou" et "dans".

DB2 Extension Net Search assure un traitement de mots vides pour une liste de langues, où les mots vides ne sont pas indexés et par conséquent, la recherche ne portera pas sur eux. Cependant, le résultat du traitement de mots vides est un index de texte plus petit et plus rapide.

Notez que les mots vides qui ne sont pas indexés sont traités de la même façon que les mots normaux lors de la recherche. Cependant, si un mot vide a été indexé, il sera ignoré lors de la recherche.

Langues prenant en charge les mots vides

Les langues suivantes assurent un traitement de mots vides.

AR_AA	Arabe parlé dans les pays arabes
CA_ES	Catalan parlé en Espagne
DA_DK	Danois parlé au Danemark
DE_CH	Allemand parlé en Suisse
DE_DE	Allemand parlé en Allemagne
EL_GR	Grec parlé en Grèce
EN_GB	Anglais parlé au Royaume-Uni.
EN_US	Anglais parlé aux Etats-Unis
ES_ES	Espagnol parlé en Espagne
FI_FI	Finnois parlé en Finlande
FR_CA	Français parlé au Canada
FR_FR	Français parlé en France
HE_IL	Hébreu parlé en Israël
IS_IS	Islandais parlé en Islande
IT_IT	Italien parlé en Italie
IW_IL	Hébreu parlé en Israël
NB_NO	Norvégien Bokmal parlé en Norvège
NL_BE	Néerlandais parlé en Belgique
NN_NO	Norvégien Nynorsk parlé en Norvège
PT_BR	Portugais parlé au Brésil
PT_PT	Portugais parlé au Portugal
RU_RU	Russe parlé en Russie
SV_SE	Suédois parlé en Suède

Configuration

L'Extension Net Search peut rechercher des mots formés de diverses combinaisons de caractères comme par exemple, des caractères alphanumériques, des nombres et des caractères spéciaux. Pour cela, l'Extension Net Search offre les configurations suivantes :

Normalisation d'un caractère

La normalisation des caractères garantit que les mots qui s'écrivent de deux façons différentes puissent être recherchés indifféremment à partir de leurs deux orthographes. Le terme allemand 'über', par exemple, peut également s'écrire 'ueber'. La normalisation permet une recherche à partir de l'une ou l'autre orthographe, 'über' ou 'ueber'. Vous pouvez également utiliser cette fonction avec des accents, comme par exemple á.

Utilisation de caractères spéciaux dans un même mot

L'utilisation de caractères spéciaux dans un même mot permet de rechercher des noms de produit pouvant comporter une série de caractères alphanumériques, de caractères spéciaux ou de nombres comme on le ferait pour un seul mot. En autorisant, par exemple, le caractère spécial '/', les plateformes AS/400 ou OS/390 peuvent être recherchées comme un mot unique au lieu d'être divisées en plusieurs segments comme 'AS' et '400', ou 'OS' et '390'.

Pour ces deux types de configuration, des modules sont disponibles. Pour les personnaliser, modifiez le modèle de fichier .ini avant de créer un index.

Le modèle de fichier .ini est enregistré sous sqllib/db2ext/cteixcfg.ini. Vous pouvez également modifier le modèle de fichier en utilisant la commande CREATE INDEX, il est donc recommandé de ne modifier que les éléments suivants :

UnicodeNormalization	- dans le cadre de la normalisation d'un caractère
AccentNormalization	- dans le cadre de la normalisation d'un caractère
UmlautNormalization	- dans le cadre de la normalisation d'un caractère
AdditionalAlphanumCharacters	- dans le cadre de l'utilisation de caractères spéciaux dans un même mot

Vous ne pouvez modifier que les paramètres figurant dans le fichier .ini.

Notez qu'une fois la modification effectuée, chaque nouvel index fonctionnera avec la nouvelle configuration.

Pour activer les modules, supprimez le ";" en début de ligne. Pour plus d'informations, reportez-vous au fichier cteixcfg.ini.

Annexe I. Codes anomalie du moteur de recherche de texte

0	Opération réussie - aucune erreur ne s'est produite.
1	Un descripteur incorrect a été transmis à la fonction.
2	La fonction n'a pas pu allouer assez de mémoire.
3	La fonction n'a pas pu s'exécuter en raison des limites d'accès ou des restrictions de sécurité.
4	L'opération n'est pas prise en charge pour cette version exécutable du Moteur de recherche de texte.
5	Cette opération n'est actuellement pas activée.
6	L'application a violé le protocole du Moteur de recherche de texte en appelant les fonctions du Moteur de recherche de texte dans un ordre incorrect.
7	Une erreur inattendue s'est produite. Veuillez signaler cette erreur à votre technicien de maintenance.
8	Une langue incorrecte a été spécifiée.
9	La langue spécifiée est correcte mais pas prise en charge par la phase d'exécution du Moteur de recherche de texte.
10	Un CCSID incorrect a été spécifié.
11	Le CCSID spécifié est correct mais pas pris en charge par la phase d'exécution du Moteur de recherche de texte.
12	Un ID document incorrect a été spécifié.
13	Le format de document spécifié est correct mais pas pris en charge par la phase d'exécution du Moteur de recherche de texte.
14	Un format de document incorrect a été spécifié.
15	L'opération n'a pas abouti en raison des limites d'accès lors de l'entrée-sortie du fichier.
16	L'opération n'a pas abouti en raison des erreurs de lecture lors de l'entrée-sortie du fichier.
17	L'opération n'a pas abouti en raison des erreurs de lecture lors de l'entrée du fichier.
18	L'opération n'a pas abouti en raison des erreurs d'écriture lors de la sortie du fichier.
19	L'opération n'a pas abouti en raison des erreurs de recherche lors de l'entrée-sortie du fichier.
20	Cette opération n'a pas abouti en raison des erreurs d'indication lors de l'entrée-sortie du fichier.
21	L'opération n'a pas abouti en raison des erreurs de fermeture lors de l'entrée-sortie du fichier.
22	L'opération n'a pas abouti en raison des erreurs survenues lors de la phase de changement de nom.

Codes anomalie

23	L'opération n'a pas abouti en raison des erreurs survenues lors de la suppression.
24	L'opération n'a pas abouti en raison des erreurs survenues lors de l'exécution de mkdir.
25	Un ou plusieurs arguments de fonction avaient une valeur incorrecte (par exemple un pointeur NULL ou une valeur de type énumératif incorrecte).
26	Le répertoire spécifié n'existe pas.
27	Une erreur imprévue de Moteur de recherche de texte s'est produite. Pour plus de détails, veuillez consulter le code d'erreur dans les informations relatives à l'erreur.
28	Une erreur de convention de service IBM s'est produite. Veuillez signaler cette erreur.
29	Il y a eu tentative de mise à jour d'un document vide.
30	L'argument spécifié n'est pas pris en charge pour cette opération.
31	L'analyseur de l'attribut de date a détecté une valeur incorrecte lors de la tentative d'analyse d'un attribut de date.
32	L'analyseur de l'attribut de nombre a détecté un nombre incorrect lors de la tentative d'analyse d'un attribut de nombre.
33	Le nom de l'attribut est incorrect, probablement trop long.
35	Réservez un nombre pour une utilisation ultérieure.
36	Le document d'entrée contient un attribut (DATE, NUMBER ou STRING) qui dépasse la limite fixée pour la longueur des attributs. L'attribut de texte a été tronqué pour être conforme à cette limite.
38	Le seuil d'avertissement défini par l'utilisateur a été dépassé. Par conséquent, cette erreur a été générée.
39	Le document d'entrée n'a pas pu être indexé. Il comporte trop de zones imbriquées.
40	La limite des différents attributs pour l'un des types d'attribut a été dépassée pour cet index.
46	L'itérateur n'est pas (ou plus) correct, car sa liste est vide ou a été supprimée.
47	La fonction n'est pas prise en charge pour le type de descripteur transmis. Cette erreur se produit, par exemple, lorsque vous tentez d'utiliser itlQueryResultEntryGetData sur une liste d'itérateurs qui ne représente pas un itérateur de résultat de requête.
48	Cet avertissement est émis si un fichier de mots vides est introuvable pour la langue et le chemin de ressources spécifiés.
49	Cet avertissement est émis si un fichier de mots vides ne comporte pas de mots vides.
50	Cet avertissement est émis si un fichier de mots vides comporte des données incorrectes.
100	Cet index n'a pas pu être ouvert car il n'existe pas avec le nom et/ou le répertoire spécifié(s).
101	Le nom d' indiqué est incorrect.

102	Le nom de répertoire indiqué est incorrect.
103	L'opération n'a pas pu être réalisée car le Moteur de recherche de texte a détecté une corruption dans la structure de l' et/ou l'ensemble de fichiers d'.
104	L' spécifié ne peut pas être créé car il existe déjà avec le nom et le répertoire donnés.
109	Avant d'exécuter toute autre opération, il faut effectuer une annulation.
110	Le fichier de configuration de l' ne comporte pas de section obligatoire comme indiqué dans le contexte de l'erreur.
111	Le fichier de configuration de l' ne comporte pas d'option obligatoire comme indiqué dans le contexte de l'erreur.
112	Le fichier de configuration de l' comporte des données incorrectes dans l'option comme indiqué dans le contexte d'erreur.
113	Le fichier de configuration de l' ne correspond pas à la version du Moteur de recherche de texte.
200	Le nom du modèle de document spécifié est incorrect.
201	Le nom de zone du modèle de document spécifié est incorrect.
202	Le modèle de document spécifié est inconnu.
203	Le modèle de document spécifié existe déjà et ne peut pas être redéfini.
204	Trop de modèles de document ou des modèles de document de taille trop grande ont été ajoutés à l'.
205	Le modèle de document contient trop d'éléments.
206	L'élément du modèle de document contient un paramètre (attribut XML) non autorisé pour ce type d'élément.
207	L'élément du modèle de document contient une valeur de paramètre non autorisée pour ce type de paramètre (attribut XML).
208	L'élément du modèle de document ne comporte pas de paramètre obligatoire (attribut XML), par exemple "nom".
209	Le modèle de document ne semble pas être au format XML ou il commence avec un élément XML imprévu.
210	La valeur XPath donnée (valeur du releveur de coordonnées) contient une marque imprévue.
211	La valeur XPath donnée (valeur du releveur de coordonnées) contient un spécificateur d'axe imprévu.
212	La valeur XPath donnée (valeur du releveur de coordonnées) contient un test de noeud imprévu.
213	Le fichier répertoire du modèle de document (extension .mdx) est corrompu.
214	Le fichier d' du modèle de document (extension .mox) est corrompu.
215	Le document contient un élément XML qui est mappé sur un

	attribut de document et qui contient un autre attribut de document. L'attribut interne est ignoré.
216	La valeur de paramètre donnée est trop longue en tant que code GGP ou HTML.
217	Le modèle de document contient une définition de zone en double.
218	Le modèle de document contient une définition d'attribut en double.
300	L'opération ne peut pas être réalisée car le Moteur de recherche de texte a détecté une corruption dans les fichiers d' utilisés pour le mappage du nom de document.
301	L'opération ne peut pas être réalisée car le Moteur de recherche de texte a détecté un numéro de document incorrect.
302	L'opération ne peut pas être réalisée car le Moteur de recherche de texte a détecté un identificateur de documents incorrect.
303	L'opération ne peut pas être réalisée car le Moteur de recherche de texte n'a détecté aucune entrée d' pour l'identificateur de document.
304	L'opération ne peut pas être réalisée car le Moteur de recherche de texte n'a détecté aucune entrée d' pour le numéro de document.
305	L'opération ne peut pas être réalisée car le Moteur de recherche de document a détecté un dépassement dans les numéros de documents utilisés.
306	L'ID document que l'application a tenté d'er est déjà apparu dans la liste de documents. Le Moteur de recherche de texte ne prend pas en charge les ID document en double apparaissant dans une séquence d'ation, c'est-à-dire avant que la mise à jour ne soit validée.
340	Le terme STRENGTH est incorrect.
341	Le numéro de relation est incorrect.
342	Le type de relation est incorrect, utilisez l'une des définitions décrites dans l'interface de programme d'application (API).
343	La phrase (terme) est trop longue.
344	Une fin de fichier imprévue est survenue lors de la lecture.
345	Un conflit de version a été détecté lors de la lecture des fichiers d'/thésaurus.
346	Dépassement dans la mémoire tampon du thésaurus.
347	Nom incorrect, probablement trop long pour un nom de fichier ou de répertoire.
348	La recherche n'a pas trouvé le terme (phrase) dans le dictionnaire ou l'entrée dans le fichier de définition ne contient pas de terme obligatoire.
349	Le fichier de définition est vide.
350	Le dictionnaire Thésaurus ou le fichier de définition comme spécifié dans le paramètre d'entrée n'existe pas.
351	Erreurs de syntaxe dans le fichier de définition.

352	La relation a été spécifiée de manière incorrecte.
352	Le numéro de relation n'est pas compris dans la plage autorisée.
360	Un caractère de masquage unique incorrect a été utilisé.
361	Un caractère de masquage multiple incorrect a été utilisé.
362	L'arité de l'opérateur est plus petite que le nombre d'opérandes donné dans la requête.
363	La valeur de l'opérateur non comprise dans la plage autorisée a été définie par l'énumération ItlEnOperator.
364	La valeur de la formule de niveau n'est pas comprise dans la plage d'énumération.
365	Le numéro identifiant le segment de proximité n'est pas compris dans la plage autorisée.
366	La requête est en cours de construction et ne peut pas être redéfinie ni réinitialisée.
367	La portée donnée comme résultat de la recherche précédente dénote un résultat vide.
368	Demande d'appel incorrecte pour ajouter des noms de zones avant de définir le premier nom.
369	L'indicateur de recherche incorrect demandant une comparaison avec le contenu de l' est ignoré. Si, par exemple, une comparaison respectant les majuscules et minuscules est requise pour un créé sans souci de respect des majuscules et minuscules, ce code anomalie est affiché dans les informations relatives aux erreurs.
370	Le masquage des chaînes n'est pas pris en charge pour le Thaïlandais ou les langues à jeu de caractères à deux octets (DBCS).
371	Pas d'entrée de requête correcte. Par exemple, les termes recherchés sont disponibles.
372	Opérations de comparaison incorrectes requises.
373	Opérations de comparaison incorrectes requises.
374	Un descripteur d' d'aide a été requis pour un vide.
375	La combinaison de l'opérateur et du mode opérateur requis n'est pas pris en charge.
380	Le résultat de la recherche est incomplet, la recherche a été interrompue en raison du seuil d'alerte.
381	La recherche dans l' a révélé qu'une requête comporte des mots vides.
401	L'opération ne peut pas être réalisée car le Moteur de recherche de texte a détecté une corruption dans les fichiers d' utilisés pour le mappage du nom de la zone/attribut.
402	L'opération ne peut pas être réalisée car le Moteur de recherche de texte a détecté un nom d'attribut ou de zone incorrect.
403	L'opération ne peut pas être réalisée car le nom d'attribut ou de zone donné est inconnu.

Codes anomalie

404	La limite des différents attributs pour l'un des types d'attribut ou des différentes zones a été dépassée pour cet .
500	Le document/les données comporte(nt) une séquence de caractères incorrecte (dans une source UTF8, UTF16 ou un jeu de caractères à deux octets (DBCS)).
501	Le convertisseur de page de code est erroné.
502	Le document/les données comporte(nt) une séquence de caractères incomplet (dans une source UTF8, UTF16 ou un jeu de caractères à deux octets (DBCS)).
503	Le convertisseur de page de code possède un descripteur incorrect.
600	Le document XML contient une entité asynchrone. Par exemple, une valeur d'attribut XML sans guillemets.
602	Référence incorrecte à un caractère (par exemple, ou).
603	Référence à une entité binaire incorrecte.
604	L'analyseur syntaxique XML Expat n'a pas pu être créé.
605	Un nom d'attribut dans une balise doit être unique.
607	L'analyseur syntaxique XML a détecté une référence à une entité externe incorrecte.
608	Les documents comportent une marque incorrecte, par exemple un < ou un > manquant.
609	Les documents XML doivent comporter une balise de fermeture après laquelle tout texte est interdit.
610	Aucune instruction de traitement n'est pas autorisée à cet endroit. Par exemple, la première instruction de traitement n'est pas le prologue <?xml .. ?>.
611	Un élément est une séquence de balise de début, de contenu et de balise de fin. Cette erreur se produit, par exemple, depuis la séquence "<s> text /s>", car la balise de fin est incorrecte.
612	L'allocation de mémoire a échoué dans l'analyseur syntaxique XML.
614	Référence à une entité de paramètre incorrecte.
615	Un caractère non complet, peut-être seulement le premier octet d'un caractère UTF8 à 2 octets.
616	Référence à une entité récursive.
617	Erreur de syntaxe XML ; par exemple, du texte hors des balises de début et de fin.
618	Toutes les balises de début doivent avoir une balise de fin correspondante.
619	Section cdata non fermée.
620	Marque non fermée ; par exemple, du texte après la dernière marque dans un document.
621	Une entité du document n'a pas pu être résolue.
622	Erreur inattendue.

- 631 Impossible d'analyser les informations de zone ou d'attribut dans une balise méta. La balise doit avoir le format suivant <meta name="abc" content="xyz"> ; il se peut que les noms d'attributs ou le contenu de la balise méta soit incorrect.
- 632 L'entité n'a pas pu être transformée en caractère.
- 650 Différentes définitions de zones commencent avec la même balise de début.
- 651 Une balise de début en comporte une autre, les balises sont donc ambiguës.
- 652 Si une zone et un attribut utilisent la même balise de début, ils doivent tous deux utiliser la même balise de fin ou alors aucune balise de fin.
- 653 Une zone non encore fermée si le document arrive à sa fin.
- 654 Aucun modèle de document n'est spécifié pour le format structuré. Le document sera analysé en tant que document texte en clair sans informations sur les zones ou attributs.
- 670 L'opération ne peut être réalisée car elle requiert les bibliothèques "Outside In" (TM) qui sont introuvables.
- 671 L'opération ne peut être réalisée car elle requiert une procédure des bibliothèques "Outside In" (TM) qui ne peuvent pas être chargées. Les bibliothèques sont probablement périmées ou corrompues.
- 672 Une erreur s'est produite lors du traitement du document avec "Outside In".

Annexe J. Identification et résolution des incidents

DB2 Extension Net Search fournit les informations suivantes sur le traçage des erreurs et sur la suppression des objets DB2 sans utiliser les commandes appropriées.

Traçage des erreurs

Si vous devez signaler une erreur à votre représentant IBM, vous serez invité à activer la fonction de trace afin que les informations puissent être écrites dans un fichier qui sera utilisé pour localiser l'erreur.

Comme les performances du système sont affectées lorsque la fonction de trace est activée, utilisez-la uniquement lorsqu'un représentant du Centre de support IBM ou qu'un support technique interne vous y invite.

Pour activer la fonction de trace, utilisez la fonction DB2 :

```
db2trc on
```

Reportez-vous à la documentation relative à *DB2 UDB Command Reference* (SC09-4828) pour plus d'informations.

Pour obtenir plus d'informations sur Extension Net Search, vous pouvez utiliser un masque avec le composant in 96 :

```
db2trc on -m *.*.*.96.*.*
```

En cas d'erreurs graves, vous pouvez également consulter `db2diag.log` pour vous aider.

Suppression des objets DB2 sans utiliser les commandes de l'Extension Net Search appropriées.

DB2 Extension Net Search fournit les informations suivantes, pour le cas où vous supprimeriez des objets DB2 sans utiliser les commandes appropriées.

Suppression d'une table

Avant de supprimer une table avec un ou plusieurs index de texte, vous devez exécuter la commande suivante pour chaque index de texte :

```
db2text drop index <nom_index> for text
```

Si, par erreur, vous supprimez une table avant de supprimer les index, certaines parties des index existeront encore, comme par exemple, les tables d'administration et les fichiers des index de texte.

Pour effacer ces fichiers, vous devez supprimer les index en utilisant la commande `db2text drop index`, même si la table n'existe plus.

Suppression d'une base de donnée

Avant de supprimer une base de données avec un ou plusieurs index de texte, vous devez exécuter la commande suivante pour chaque index de texte :

```
db2text drop index <nom_index> for text
```

Traçage des erreurs

Si vous n'utilisez pas cette commande, vous devrez effacer manuellement tous les fichiers d'index figurant dans le répertoire_index et dans le répertoire_travail_index.

Notez que si les index de la base de données supprimée ont été créés au cours d'une mise à jour automatique, vous devrez éditer le fichier du planificateur ctedem.dat.

Pour ce faire, saisissez les commandes suivantes :

Pour UNIX :

```
db2text stop force  
cd ~/sql1lib/db2ext
```

ou pour Windows:

```
db2text stop force  
cd <db2_installer_chemin>\sql1lib\<db2_instance_nom>\db2ext
```

Ouvrez le fichier ctedem.dat dans le répertoire et effacer toutes les entrées ayant trait à la base de données supprimée.

Annexe K. Messages Data Link

Les erreurs de Data Link n'entraîneront pas d'erreurs Extension Net Search, et par conséquent, à une fin du processus d'indexation. Veuillez vérifier le journal des événements et prenez soin des documents non indexés. Vous devez vous assurer manuellement que ces documents sont réindexés si nécessaire.

Tableau 18. Messages d'avertissement Data Link

Numéro	Message Data Link
01H90=	CTEDL - Une erreur est survenue lors de la définition de la valeur de retour de l'objet BLOB.
01H91=	CTEDL - Le délai de l'opération d'E-S de DataLink a été dépassé.
01H92=	CTEDL - Le codage de caractère n'est pas pris en charge.
01H93=	CTEDL - Le schéma DataLink n'est pas pris en charge.
01H94=	CTEDL - Une erreur est survenue lors de la création de l'instance du type de données de l'objet BLOB.
01H95=	CTEDL - Une erreur est survenue lors de la définition des informations proxy.
01H96=	CTEDL - Le schéma UNC est uniquement valide sur Windows mais le système d'exploitation n'a pas pu être déterminé.
01H97=	CTEDL - Le schéma UNC est uniquement valide sur le système d'exploitation Windows.
01H98=	CTEDL - Le schéma DFS est uniquement valide sur AIX.
01H99=	CTEDL - Le schéma DFS est uniquement valide sur AIX mais le système d'exploitation n'a pas pu être déterminé.
01H01=	CTEDL - L'erreur détermine le numéro de port dans l'adresse URL.
01H02=	CTEDL - Un schéma DataLink inconnu a été détecté.
01H03=	CTEDL - Impossible d'établir la connexion.
01H00=	CTEDL - Une erreur est survenue lors de l'exécution de la fonction UDF DataLink.
01H80=	CTEDL - Le fichier DataLink "{0}" est introuvable.
01H81=	CTEDL - Accès non autorisé au fichier DataLink à "{0}".
01H82=	CTEDL - Fin de fichier imprévue ou fin de flux atteint pour "{0}".
01H83=	CTEDL - Le fichier DataLink "{0}" ne peut pas être lu.
01H85=	CTEDL - Le schéma DataLink URL "{0}" requiert un nom de fichier.
01H86=	CTEDL - Aucune connexion au serveur de fichiers Datalink "{0}" n'a été établie.
01H60=	CTEDL - Requête HTTP incorrecte - syntaxe incorrecte de DataLink URL.
01H61=	CTEDL - Requête DataLink non autorisée - authentification utilisateur requise.
01H62=	CTEDL - L'accès à DataLink est payant.
01H63=	CTEDL - Accès interdit à DataLink URL.
01H64=	CTEDL - Fichier introuvable sur le serveur DataLink.
01H65=	CTEDL - La méthode requise n'est pas autorisée pour la ressource DataLink.
01H66=	CTEDL - Requête non admise.

Tableau 18. Messages d'avertissement Data Link (suite)

01H67=	CTEDL - Authentification proxy requise.
01H68=	CTEDL - Dépassement de délai de la requête client.
01H69=	CTEDL - Conflit avec l'état courant de la ressource DataLink.
01H10=	CTEDL - La ressource DataLink n'est plus disponible sur le serveur.
01H11=	CTEDL - Vous devez spécifier une longueur de contenu pour que la requête soit acceptée.
01H12=	CTEDL - La précondition fournie dans la zone d'en-tête a donné un résultat faux.
01H13=	CTEDL - L'entité requise DataLink est trop grande.
01H14=	CTEDL - L'URL DataLink requise est trop longue.
01H15=	CTEDL - Type MIME non pris en charge.
01H16=	CTEDL - Requête de plage non satisfaisante.
01H17=	CTEDL - Echec de l'espérance.
01H18=	CTEDL - La requête sur DataLink URL n'a obtenu aucune information sur la longueur.
01H20=	CTEDL - Réponse HTTP incorrecte.
01H70=	CTEDL - Erreur de serveur DataLink interne.
01H71=	CTEDL - Fonction non prise en charge par le serveur DataLink.
01H72=	CTEDL - Passerelle incorrecte.
01H73=	CTEDL - Service indisponible - Le serveur DL est provisoirement surchargé ou en cours de maintenance.
01H74=	CTEDL - Dépassement de délai de la passerelle.
01H75=	CTEDL - Version HTTP non prise en charge.
01H30=	CTEDL - Tentative d'établissement d'une connexion socket - erreur dans le protocole sous-jacent.
01H31=	CTEDL - Impossible d'établir un chemin vers le Serveur DataLink {0}.
01H32=	CTEDL - Connexion socket impossible vers l'adresse éloignée {0}.
01H33=	CTEDL - Impossible de définir les accès du socket sur l'adresse locale.
01H34=	CTEDL - L'adresse IP du serveur DataLink {0} n'a pas pu être déterminée.
01H35=	CTEDL - Exception de services inconnu - Pas de prise en charge du type MIME.
01H36=	CTEDL - Syntaxe incorrecte de l'URL '{0}' - Protocole non pris en charge ou le DL URL n'a pas pu être analysé.

Annexe L. CCSID pris en charge par le thésaurus

Les CCSID suivants sont pris en charge par le thésaurus :

CCSID

819	Latin 1
850	PC Data Latin 1
874	Thaïlandais
932	Japonais combiné
943	Japonais combiné
950	Chinois traditionnel combiné
954	Japonais
970	Coréen combiné
1208	UTF 8
1250	Latin 2
1252	Latin 1
1253	Tchèque
1254	Turc
1255	Hébreu
1256	Arabe
1258	Vietnamien
1363	Coréen combiné
1381	Chinois simplifié combiné
1383	Chinois (simplifié), SBCS/DBCS combiné
1386	Chinois (simplifié), SBCS/DBCS combiné
5039	Japonais (SNCS/DBCS) combiné

Pour compiler le fichier de définition du thésaurus, consultez «DB2EXTTH (Utilitaire)» à la page 132.

Annexe M. Messages renvoyés par les outils du thésaurus

ADM_MSG_INVALID_CCSID

Explication : CCSID spécifié incorrect.

La page de codes requise n'est pas prise en charge.

ITL_THES_MSG_DEFFILE_MISSING

Explication : Erreur de paramètre *nom de fichier*. Le fichier de définition du thésaurus n'existe pas.

ITL_THES_MSG_NONAME_ERROR

Explication : Erreur de paramètre. Aucun nom de fichier de définition de thésaurus n'est spécifié.

ITL_THES_MSG_PATHLEN_ERROR

Explication : Erreur de paramètre *nom de fichier*. Le chemin d'accès au fichier de définition du thésaurus est trop long. La longueur du chemin d'accès ne doit pas dépasser la longueur maximale prise en charge pour les noms de répertoires dans le système d'exploitation.

ITL_THES_MSG_NAMELEN_ERROR

Explication : Erreur de paramètre *nom de fichier*. Le nom du fichier de définition du thésaurus est trop long.

ITL_THES_MSG_NO_TARGET_DIR_ERROR

Explication : Erreur de paramètre. Aucun répertoire cible n'est spécifié.

ITL_THES_MSG_UNEXPECTED_ERROR

Explication : Erreur interne imprévue.

ITL_THES_MSG_PARAMETER_ERROR

Explication : Erreur interne de paramètre.

ITL_THES_MSG_FILE_OPEN_ERROR

Explication : Impossible d'ouvrir de fichier *nom de fichier*.

ITL_THES_MSG_FILE_REACHED_END

Explication : Fin de fichier imprévue dans le fichier de définition de thésaurus.

Il y a une erreur dans le fichier de définition.

ITL_THES_MSG_FILE_READ_ERROR

Explication : Impossible de lire le fichier *nom de fichier*.

ITL_THES_MSG_FILE_WRITE_ERROR

Explication : Impossible d'écrire dans le fichier *nom de fichier*.

ITL_THES_MSG_FILE_ACCESS_ERROR

Explication : Impossible d'accéder au fichier *nom de fichier*.

ITL_THES_MSG_FILE_REMOVE_ERROR

Explication : Impossible de supprimer le fichier *nom de fichier*.

ITL_THES_MSG_FILE_RENAME_ERROR

Explication : Impossible de renommer le fichier *nom de fichier 1* en *nom de fichier 2*.

ITL_THES_MSG_FILE_CLOSE_ERROR

Explication : Impossible de fermer le fichier *nom de fichier*.

ITL_THES_MSG_FILE_EOF_ERROR

Explication : Fin de fichier imprévue dans *nom de fichier*.

Erreur dans le fichier de définition.

ITL_THES_MSG_MEMORY_ERROR

Explication : Erreur de mémoire.

ITL_THES_MSG_BUFFER_OVERFLOW

Explication : Dépassement de la mémoire tampon.

ITL_THES_MSG_LOCKING_ERROR

Explication : Impossible de verrouiller le dictionnaire *nom de fichier*.

ITL_THES_MSG_LOCKED

Explication : Le dictionnaire Thésaurus *nom du dictionnaire* est en cours d'utilisation.

Messages renvoyés par les outils du thésaurus

ITL_THES_MSG_OUTFILE_EXIST

Explication : Le fichier de sortie *nom de fichier* existe déjà.

ITL_THES_MSG_DICT_INTEGRITY_ERROR

Explication : Tout le dictionnaire *nom du dictionnaire* est perdu.

Le fichier du dictionnaire Thésaurus est corrompu.

ITL_THES_MSG_DICT_VERSION_ERROR

Explication : Erreur de version du dictionnaire *nom du dictionnaire*.

Le dictionnaire Thésaurus a été créé avec une version antérieure incompatible.

ITL_THES_MSG_DICT_NOT_EXIST

Explication : Le dictionnaire Thésaurus *nom du dictionnaire* n'existe pas.

ITL_THES_MSG_DICT_EXIST

Explication : Le dictionnaire Thésaurus *nom du dictionnaire* existe déjà.

Impossible d'écraser.

ITL_THES_MSG_NORMALIZE_ERROR

Explication : Erreur de normalisation d'un terme.

Erreur dans le fichier de définition de thésaurus.

ITL_THES_MSG_INTERNAL_ERROR

Explication : Erreur interne.

ITL_THES_MSG_INPUT_ERROR

Explication : Erreur dans le fichier de définition du thésaurus *nom de fichier* à la ligne *numéro de ligne*.

ITL_THES_MSG_ERROR_IN_FILE

Explication : Erreur dans le fichier *nom de fichier*.

ITL_THES_MSG_IE_EMPTY

Explication : Le fichier de définition du thésaurus *nom de fichier* est vide.

ITL_THES_MSG_IE_BLOCK_START

Explication : Aucune ligne de départ de bloc n'a été détectée dans le fichier *nom de fichier* à la ligne *numéro de ligne*.

ITL_THES_MSG_IE_REL_SYNTAX

Explication : La relation est incorrectement spécifiée dans *nom de fichier* à la ligne *numéro de ligne*.

ITL_THES_MSG_IE_USER_DEF

Explication : La relation est incorrectement spécifiée dans *nom de fichier* à la ligne *numéro de ligne*.

ITL_THES_MSG_IE_USER_DEF_DOMAIN

Explication : Le numéro de relation n'est pas compris dans la plage autorisée dans *nom de fichier* à la ligne *numéro de ligne*.

ITL_THES_MSG_IE_NO_TERM

Explication : Aucun terme n'est défini dans *nom de fichier* à la ligne *numéro de ligne*.

ITL_THES_MSG_IE_TERM_LEN

Explication : Un terme de thésaurus dépasse 64 caractères.

ITL_THES_MSG_IE_STRENGTH_SYNTAX

Explication : Une valeur STRENGTH est spécifiée incorrectement.

Syntaxe : Après le terme, tapez [:20] pour une longueur égale à 20.

ITL_THES_MSG_IE_STRENGTH_DOMAIN

Explication : La valeur STRENGTH n'est pas comprise dans la plage autorisée.

Les valeurs correctes sont comprises entre 1 - 100 ; la valeur par défaut est égale à 100.

Annexe N. Erreurs système Windows

Voici une liste des erreurs système Windows :

Erreurs système

1	Fonction incorrecte.
2	Le système ne peut pas détecter le fichier spécifié.
3	Le système ne peut pas localiser le chemin d'accès spécifié.
4	Le système ne peut pas ouvrir le fichier.
5	L'accès est refusé.
6	Le descripteur est incorrect.
8	L'espace mémoire disponible est insuffisant pour traiter cette commande.
14	L'espace mémoire disponible est insuffisant pour terminer cette opération.
15	Le système ne peut pas détecter l'unité spécifiée.
29	Le système ne peut pas écrire sur l'unité spécifiée.
30	Le système ne peut pas lire sur l'unité spécifiée.
32	Le processus ne peut pas accéder au fichier car il est en cours d'utilisation par un autre processus.
36	Trop de fichiers sont ouverts pour le partage.
38	La fin du fichier est atteinte.
39	Le disque est saturé.
80	Le fichier existe.
82	Impossible de créer le répertoire ou le fichier.
100	Impossible de créer un autre sémaphore système.
101	Le sémaphore exclusif appartient à un autre processus.
102	Le sémaphore est défini et ne peut pas être fermé.
103	Le sémaphore ne peut pas être défini à nouveau.
104	Impossible de demander des sémaphores exclusifs au moment de l'interruption.
105	Les précédents droits de propriété de ce sémaphore sont terminés.
110	Le système ne peut pas ouvrir l'unité ou le fichier spécifié.
111	Le nom de fichier est trop long.
112	L'espace est insuffisant sur le disque.
121	Le délai d'attente du sémaphore a expiré.
126	Le module spécifié n'a pas pu être détecté.
127	La procédure spécifiée n'a pas pu être détectée.

147	Les ressources disponibles sont insuffisantes pour traiter cette commande.
155	Impossible de créer une autre unité d'exécution.
161	Le chemin d'accès spécifié est incorrect.
164	Des unités d'exécution ne peuvent plus être créées dans le système.
170	Le ressource demandée est en cours d'utilisation.
183	Impossible de créer un fichier lorsqu'il existe déjà.
187	Le nom du sémaphore système spécifié n'a pas pu être détecté.
206	Le nom de fichier ou l'extension est trop long.
267	Le nom du répertoire est incorrect.
288	Tentative de libération du processus mutex n'appartenant pas au demandeur.
298	Trop de requêtes ont été adressées à un sémaphore.
998	Accès incorrect à l'emplacement de la mémoire.
1051	Une commande d'arrêt a été émise à un service dont dépendent d'autres services en cours d'exécution.
1052	La commande requise n'est pas correcte pour ce service.
1053	Le service n'a pas répondu à la requête de démarrage ou de contrôle de manière périodique.
1054	Une unité d'exécution n'a pas pu être créée pour le service.
1055	La base de données de services est verrouillée.
1056	Une instance du service est déjà en cours d'exécution.
1057	Le nom du compte est incorrect ou il n'existe pas.
1058	Le service ne peut pas être démarré, soit parce qu'il est désactivé, ou parce qu'aucune unité activée ne lui est associée.
1059	Une dépendance circulaire de services a été spécifiée.
1060	Le service spécifié n'existe pas en tant que service installé.
1061	Le service ne peut pas accepter de messages de contrôle à ce stade.
1062	Le service n'a pas été démarré.
1063	Le processus du service n'a pas pu se connecter au contrôleur de service.
1064	Une exception est survenue dans le service lors du traitement de la requête de contrôle.
1066	Le service a renvoyé un code d'erreur spécifique au service.
1067	Le processus s'est terminé de manière inattendue.
1068	Le démarrage du service ou du groupe de dépendance a échoué.
1069	Le service n'a pas démarré en raison d'un échec de la connexion.
1070	Après le démarrage, le service est suspendu en état d'attente de démarrage.
1071	Le verrou spécifié de la base de données de services est incorrect.

1072	Le service spécifié a été marqué pour être supprimé.
1073	Le service spécifié existe déjà.
1078	Ce nom est déjà utilisé en tant que nom de service ou nom d'affichage du service.
1079	Le compte indiqué pour ce service est différent du compte spécifié pour d'autres services exécutés dans le même processus.
1082	Aucun programme de reprise n'a été configuré pour ce service.
1154	L'un des fichiers de bibliothèque requis pour exécuter cette application est endommagé.
1219	Les justificatifs fournissent au conflit un ensemble existant de justificatifs.
1242	Le service est déjà enregistré.
1243	Le service spécifié n'existe pas.
1244	L'opération requise n'a pas été réalisée car l'utilisateur n'a pas été authentifié.
1245	L'opération requise n'a pas été réalisée car l'utilisateur ne s'est pas connecté au réseau. Le service spécifié n'existe pas.
1392	Le fichier ou le répertoire est endommagé et ne peut être lu.
1455	Le fichier de pagination est trop petit pour que cette opération se termine.
1793	Le compte de l'utilisateur a expiré.

Annexe O. 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 brevets couvrant les produits mentionnés 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 EMEA Director of Licensing
IBM Europe Middle-East Africa
Tour Descartes
La Défense 5
2, avenue Gambetta
92066 - Paris-La Défense CEDEX
France

Pour le Canada, veuillez adresser votre courrier à :

IBM Director of Commercial Relations
IBM Canada Ltd.
3600 Steeles Avenue East
Markham, Ontario
L3R 9Z7
Canada

Les informations sur les licences concernant les produits utilisant un jeu de caractères à deux octets peuvent être obtenues par écrit à l'adresse suivante :

IBM World Trade Asia Corporation
Licensing
2-31 Roppongi 3-chome, Minato-ku
Tokyo 106, Japon

Le paragraphe suivant ne s'applique ni au Royaume-Uni ni dans aucun autre pays dans lequel il serait contraire aux lois locales. LE PRESENT DOCUMENT EST LIVRE «EN L'ETAT». IBM DECLINE TOUTE RESPONSABILITE, EXPRESSE OU IMPLICITE, RELATIVE AUX INFORMATIONS QUI Y SONT CONTENUES, Y COMPRIS EN CE QUI CONCERNE LES GARANTIES DE QUALITE 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 Canada Limited
Office of the Lab Director
1150 Eglinton Ave. East
North York, Ontario
M3C 1H7
CANADA

Ces informations peuvent être soumises à des conditions particulières, prévoyant notamment le paiement d'une redevance.

Le logiciel sous licence décrit dans ces message d'information et tous les éléments sous licence disponibles s'y rapportant sont fournis par IBM conformément aux dispositions de l'ICA, des Conditions internationales d'utilisation des logiciels IBM ou de tout autre accord équivalent.

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.

Ce document peut contenir des exemples de données et des 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.

LICENCE DE COPYRIGHT :

Le présent logiciel peut contenir des exemples de programmes d'application en langage source destinés à illustrer les techniques de programmation sur différentes plateformes d'exploitation. Vous avez le droit de copier, de modifier et de distribuer ces exemples de programmes sous quelque forme que ce soit et sans paiement d'aucune redevance à IBM, à des fins de développement, d'utilisation, de vente ou de distribution de programmes d'application conformes aux interfaces de programmation des plateformes pour lesquels ils ont été écrits ou aux interfaces de programmation IBM. Ces exemples de programmes n'ont pas été rigoureusement testés dans toutes les conditions. Par conséquent, IBM ne peut garantir expressément ou implicitement la fiabilité, la maintenabilité ou le fonctionnement de ces programmes.

Toute copie totale ou partielle de ces programmes exemples et des oeuvres qui en sont dérivées doit comprendre une notice de copyright, libellée comme suit :

© (nom de votre société) (année). Des segments de code sont dérivés des Programmes exemples d'IBM Corp. © Copyright IBM Corp. _entrez l'année ou les années_. All rights reserved.

Marques

Les termes qui suivent sont des marques d'International Business Machines Corporation aux Etats-Unis et/ou dans certains autres pays.

AIX	DB2 Universal Database
DB2	IBM DRDA
Extensions DB2	z/OS
Informix	OS/400
OS/390	AS/400
zSeries	

Les termes qui suivent sont des marques d'autres sociétés :

Microsoft, Windows, Windows NT et le logo Windows sont des marques de Microsoft Corporation aux Etats-Unis et/ou dans certains autres pays.

Java ou toutes les marques et logos incluant Java, sont des marques de Sun Microsystems, Inc. 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.

D'autres sociétés sont propriétaires des autres marques, noms de produits ou logos qui pourraient apparaître dans ce document.

Glossaire

Ce glossaire donne une définition des termes et abréviations utilisés dans ce manuel. Si vous ne trouvez pas le terme que vous recherchez, reportez-vous à l'index ou à la *Terminologie du traitement de l'information*, New York : McGraw-Hill, 1994.

A

activation (enable) : Préparation d'une base de données pour que DB2 Extension Net Search puisse l'utiliser.

argument de recherche (search argument) : Conditions spécifiées lors de la recherche, constituées d'un ou de plusieurs termes recherchés et de paramètres de recherche.

C

caractéristiques d'un index (index characteristics) : Propriétés d'un *index de texte* déterminant :

- la fréquence à laquelle l'index est mis à jour
- la date de la première mise à jour d'index

caractère d'échappement (escape character) : Caractère indiquant que le caractère sous-jacent ne doit pas être interprété comme étant un *caractère de masquage*.

caractère de masquage (masking character) : Caractère de remplacement utilisé pour représenter un ou plusieurs caractères au début, au milieu ou à la fin d'une chaîne de recherche. Les caractères de masquage sont généralement employés pour retrouver les variantes d'un terme dans un index.

caractère générique (wildcard character) : Voir *caractère de masquage*.

CCSID (CCSID) : ID de jeu de caractères codés.

colonne de texte (text column) : Colonne contenant des *documents texte*.

D

DBCS (DBCS) : jeu de caractères à deux octets.

déclencheur (trigger) : Mécanisme qui ajoute automatiquement des informations sur des documents devant être indexés dans une *table de journalisation* à chaque ajout, mise à jour ou suppression d'une colonne de texte.

désactivation (disable) : Restauration d'une base de données à l'état dans lequel elle se trouvait avant d'être activée pour DB2 Extension Net Search Text par retrait des éléments créés durant le processus d'activation.

document (document) : Voir *document texte*.

document texte (text document) : Texte de type CHAR, GRAPHIC, VARGRAPHIC, LONG VARGRAPHIC, DBCLOB, VARCHAR, LONG VARCHAR ou types de données CLOB, stocké dans une table DB2.

E

extension (expand) : Action d'ajouter des termes supplémentaires dérivés d'un thésaurus à un terme recherché.

extraire (retrieve) : Rechercher un document texte utilisant un argument de recherche dans une des fonctions de recherche de DB2 Extension Net Search.

F

fonction (function) : Voir *fonction d'accès*.

fonction d'accès (access function) : Fonction utilisateur permettant de convertir le type de données d'un texte stocké dans une colonne en un type pouvant être traité par DB2 Extension Net Search.

fonction définie par l'utilisateur (user-defined function) : Fonction SQL fournie par un utilisateur de DB2, en opposition avec une fonction SQL fournie par DB2.

fonction UDF (UDF) : Fonction définie par l'utilisateur.

format (format) : Type de document, par exemple ASCII ou HTML.

fréquence de mise à jour (update frequency) : Fréquence de mise à jour d'un index, exprimée en jours, heures et minutes, et le nombre minimal de noms de documents devant être répertoriés dans la *table de journalisation* pour l'indexation, avant que celle-ci ne puisse avoir lieu.

G

groupe de noeuds (nodegroup) : Sous-ensemble nommé d'un ou plusieurs serveurs de partitions de bases de données. *Noeud* attribué à une machine physique distincte. Voir *noeud logique*.

I

index (index) : Extraction de termes importants d'un texte et stockage de ces termes dans un *index de texte*.

index de texte (text index) : Ensemble de termes importants extraits de documents texte. Chaque terme est associé au document duquel il est extrait. Le temps de recherche est considérablement amélioré en effectuant la recherche dans l'index plutôt que dans les documents mêmes.

index Ngram (Ngram index) : *Index de texte* prenant en charge les documents à jeu de caractères à deux octets (DBCS) et la recherche floue des documents à jeu de caractères mono-octet (SBCS). Voir aussi *index linguistique* et *index exact*.

indexation périodique (periodic indexing) : Indexation à des intervalles de temps prédéterminés, exprimés en jours, heures et minutes, et nombre minimal de noms de documents devant être répertoriés dans la *table de journalisation* pour l'indexation, avant que celle-ci ne puisse avoir lieu.

interpréteur de commandes (command line processor) : Programme appelé db2text qui :
vous permet d'entrer des commandes DB2
Extension Net Search
traite les commandes
affiche le résultat

M

modèle de document (document model) : Définition de la structure d'un document par rapport aux sections qu'il contient. Grâce au modèle de document, DB2 Extension Net Search tiendra compte des sections des documents lors de l'indexation. Un modèle de document répertorie les balises HTML identifiant les sections. Pour chaque balise, vous pouvez spécifier un nom de section descriptif pour faciliter l'utilisation de cette section dans les requêtes. Vous pouvez spécifier un ou plusieurs modèles de document dans un fichier de modèles de documents.

N

nombre (count) : Mot clé utilisé pour spécifier le nombre de niveaux (la profondeur) des termes du thésaurus, utilisés pour étendre la recherche pour une relation donnée.

O

occurrence (match) : Occurrence d'un terme recherché dans un document texte.

P

page de codes (code page) : Attribution de caractères graphiques et de définitions de fonctions de contrôle à tous les points de code. Par exemple, attributions de caractères et de définitions à 256 points de code pour un code à 8 octets.

R

recherche booléenne (Boolean search) : Recherche impliquant la combinaison d'un ou de plusieurs termes à l'aide d'opérateurs booléens.

recherche de texte libre (free-text search) : Recherche dans laquelle le terme recherché est exprimé sous forme de texte à format libre – une phrase décrivant le sujet recherché en langue naturelle.

recherche floue (fuzzy search) : Recherche pouvant trouver des mots dont l'orthographe est similaire à celle du terme recherché.

recherche hybride (hybrid search) : Combinaison de la *Recherche booléenne* et de la *recherche en texte libre*.

S

SBCS (SBCS) : Jeu de caractères mono-octet

score (Score) : Valeur absolue de type DOUBLE comprise entre 0 et 1, indiquant l'efficacité de la correspondance entre le document et les critères de recherche relatifs aux autres documents trouvés. La valeur indique le nombre d'occurrences détectées dans le document en relation avec la taille du document.

T

table de journalisation (log table) : Table créée par DB2 Extension Net Search contenant des informations sur lesquels les documents texte doivent être indexés. Des *déclencheurs* sont utilisés pour stocker ces informations dans une table de journalisation à chaque ajout, modification ou suppression d'un document dans une colonne de texte activée.

traçage (tracing) : Action de stocker des informations dans un fichier qui peut ensuite être utilisé pour la recherche de la cause d'une erreur.

type défini par l'utilisateur (user-defined type) : Type de données créé par un utilisateur de DB2, en opposition avec un type de données fourni par DB2, tel que LONG VARCHAR.

type UDT (UDT) : Type défini par l'utilisateur.

V

vue catalogue (catalog view) : Vue d'une table système créée par DB2 Extension Net Search Text pour des tâches administratives. Une vue catalogue comporte des informations sur les tables et colonnes activées pour être utilisées par DB2 Extension Net Search Text.

Index

Caractères spéciaux

- & (AND) opérateur dans l'argument de recherche
 - mode d'utilisation 74
- | (OR) opérateur dans l'argument de recherche
 - mode d'utilisation 74

A

- ACTIVATE CACHE, commande
 - syntaxe 106
 - utilisation 42
- administration
 - activation de la mémoire cache 68
 - affichage de l'état de l'index de texte 49
 - affichage de l'état des index 69
 - arrêt de DB2 Extension Net Search 27, 52, 95
 - Centre de contrôle DB2 51
 - création d'un index de texte 34, 55
 - démarrage de DB2 Extension Net Search 27, 52, 94
 - désactivation de la mémoire cache 69
 - effacement des événements d'index 48, 68
 - gestion des index de texte 45, 65
 - mise à jour d'index de texte 46
 - mise à jour des index de texte 68
 - modification des paramètres de l'index de texte 47, 65
 - résumé des commandes d'administrateur de base de données 97
 - résumé des commandes de propriétaire d'instance 91
 - résumé des commandes de propriétaire de table de texte 105
 - sauvegarde et restauration d'index 49
 - suppression d'index de texte 48
 - suppression des index de texte 67
 - traçage des erreurs 247
 - utilisation de services de verrouillage 27
- ALTER INDEX, commande
 - syntaxe 108
 - utilisation 47
- annexes
 - catalogues d'informations 191
 - CCSID 201
 - CCSID pris en charge par le thésaurus 251
 - codes anomalie du moteur de recherche de texte 239
 - Configuration du moteur de recherche de texte 236

- annexes (*suite*)
 - Erreurs système Windows 255
 - identification et résolution des incidents 247
 - langues prises en charge 207
 - messages Data Link 249
 - Messages DB2 Extension Net Search 209
 - messages renvoyés par les outils du thésaurus 253
 - migration 183
 - modèles de document 229
 - moteur de recherche de texte 235
 - mots vides 235
 - tokenisation 235
 - utilisation de grandes quantités de mémoire 187
- argument de recherche
 - description 141
 - nom d'attribut 148
 - opérateurs booléens 144
 - opérateurs de recherche primaire 145
 - recherche avec & et | 74
 - recherche avec NOT 76
 - recherche d'attributs numériques 77
 - recherche de parties d'un terme 75
 - recherche de termes dans des sections de documents 76
 - recherche de termes dans la même phrase 76
 - recherche de termes dans le même paragraphe 76
 - recherche de termes dans n'importe quelle séquence 73
 - recherche de termes dans une séquence fixe 75
 - recherche de texte libre 77
 - recherche de thésaurus 76
 - recherche floue 74, 146
 - spécification 73
 - syntaxe 142
 - utilisation de caractères de masquage 75
 - utilisation de caractères génériques 75
- arrêt de DB2 Extension Net Search 95
- ASCII, format de document 23
- ASCII à plat, format de document 23

B

- base de données
 - activation d'une base de données 32, 52
 - désactivation d'une base de données 33, 53
 - sauvegarde et restauration d'index 49
- Boîte de dialogue Activation de la mémoire cache 68

- boîte de dialogue Désactivation de la mémoire cache 69
- Boîte de dialogue Etat des index 69
- Boîte de dialogue Evénements d'index 68
- Boîte de dialogue Mise à jour de l'index 68
- Boîte de dialogue Modification de l'index 65
- Boîte de dialogue Suppression d'index 67

C

- caractère d'échappement
 - utilisation 75
- caractères de masquage dans un terme recherché 75
- caractères génériques dans un terme recherché 75
- caractéristiques du texte
 - CCSID 23
 - format 23
- catalogues d'information Extension Net Search
 - Voir vues 191
- catalogues d'informations 191
- CCSID
 - liste de 201
 - pages de codes de document 23
 - prise en charge par le thésaurus 251
- Centre de contrôle DB2
 - activation d'une base de données 52
 - administration 51
 - Boîte de dialogue Activation de la mémoire cache 68
 - boîte de dialogue Désactivation de la mémoire cache 69
 - Boîte de dialogue Etat des index 69
 - Boîte de dialogue Evénements d'index 68
 - Boîte de dialogue Mise à jour de l'index 68
 - Boîte de dialogue Modification de l'index 65
 - Boîte de dialogue Suppression d'index 67
 - création d'un index de texte 55
 - démarrage et arrêt de DB2 Extension Net Search 52
 - désactivation d'une base de données 53
 - écran Cible 56
 - écran Mise à jour des caractéristiques 59
 - écran Nom 55
 - écran Propriétés de type texte 58
 - écran Récapitulatif 64
 - écran Table de mémoire cache 61
 - gestion des index de texte 65
 - utilisation de l'assistant 55

CLEAR EVENTS, commande
syntaxe 112
utilisation 48

commandes

ACTIVATE CACHE 106
ALTER INDEX 108
CLEAR EVENTS 112
CONTROL 92
COPYRIGHT 139
CREATE INDEX 114
DB2EXTDL 102
DB2EXTHL 103
DB2EXTTH 132
db2text 91, 97, 105
DEACTIVATE CACHE 128
DISABLE DATABASE 100
DROP INDEX 130
ENABLE DATABASE 98
HELP 138
START 94
STOP 95
UPDATE INDEX 134

COMMITCOUNT

mot clé 122
performances, remarques 45

concepts

clé 3
concepts supplémentaires 8
fonction de transformation de
colonne 8
services de l'instance 9
utilisation d'une fonction de recherche
scalaire SQL 5
utilisation d'une fonction de valeur
table 7
utilisation d'une recherche de
procédure mémorisée 6
vues 9

concepts clés 3

concepts supplémentaires 8

configuration système requise 13

CONTAINS, fonction

exemple 72
syntaxe 150

CONTROL, commande

syntaxe 92
utilisation 29

COPYRIGHT, commande

syntaxe 139

CREATE INDEX, commande

syntaxe 114
utilisation 34

D

Datalink Manager

installation du fichier jar 38
messages d'erreur 249
modification de la taille de renvoi de
la fonction Data Link 102
modification de la taille de renvoi de
la fonction Datalink 38
types de données DATALINK 37

DB2EXTDL, commande

syntaxe 102
utilisation 132

DB2EXTHL, commande

syntaxe 103

DB2TX, interpréteur de commandes

syntaxe 91, 97, 105
utilisation 31

DEACTIVATE CACHE, commande

syntaxe 128
utilisation 42

déclencheurs

création 35
description 5

démarrage de DB2 Extension Net
Search 94

DISABLE DATABASE, commande

syntaxe 100
utilisation 33, 53

document

CCSID 23
conversion de types de données 37
format, description 23
formats pris en charge 23
indexation 3
structure 165

documents généraux (GGP)

définition d'un modèle de
document 169
définition de type de document pour
des modèles de document 229
format de document 23
limites 232

documents HTML

définition d'un modèle de
document 171
définition de type de document pour
des modèles de document 229
documents structurés 165
format de document 23
limites 232
modèle de document par défaut 168

documents structurés

activation de la prise en charge de
section 165
exemple 76
modèles de document par défaut 168
présentation 83
syntaxe de la recherche 145

documents XML

définition d'un modèle de
document 173
définition de type de document pour
des modèles de document 229
documents structurés 165
format de document 23
limites 232
modèle de document par défaut 168
Sémantique de l'expression
XPath 230

données

stockées en externe 9
données stockées en externe 9

DROP INDEX, commande

syntaxe 130
utilisation 48

E

écran Cible 56

écran Mise à jour des caractéristiques 59

écran Nom 55

écran Propriétés de type texte 58

écran Récapitulatif 64

écran Table de mémoire cache 61

ENABLE DATABASE, commande

syntaxe 98
utilisation 32, 52

environnement, client/serveur 11

environnement client/serveur 11

Erreurs système Windows 255

espace disque pour les index 23

espace requis pour les index 23

espace table 35, 56

événements de mise à jour de l'index

enregistrement 36

suppression 48

EXPANSION LIMIT mot clé 144

extension du terme recherché

Voir thésaurus

Extension Net Search

activation d'une base de données 32,
52

activation de la mémoire cache 68

affichage de l'état de l'index de
texte 49

affichage de l'état des index 69

création d'un index de texte 34, 55

création d'un index de texte sur un

alias de type nickname à l'aide de
DB2 Replication 38

création d'une mémoire cache pour
une recherche de procédure
mémorisée 39

démarrage et arrêt 27, 52

désactivation d'une base de
données 33, 53

désactivation de la mémoire cache 69

effacement des événements
d'index 48, 68

gestion des index de texte 45, 65

messages 209

mise à jour d'index de texte 46

mise à jour des index de texte 68

modification des paramètres de
l'index de texte 47, 65

sauvegarder et restaurer 49

services de l'instance 27

services de mise à jour 30

suppression d'index de texte 48

suppression des index de texte 67

utilisation du Centre de contrôle
DB2 51

F

fonction

fonctions de recherche 71
pour la conversion de types de
données 37

fonction de la procédure mémorisée

recherche 78

syntaxe 162

fonction de transformation de colonne 8

- fonction de valeur table
 - création d'un index de texte sur un alias de type nickname à l'aide de la réplication 38
 - index de texte sur une vue 43
 - recherche 79
 - syntaxe 153
 - syntaxe HIGHLIGHT 157
 - utilisation de la fonction HIGHLIGHT 80
- fonctions
 - CONTAINS 150
 - description 71
 - HIGHLIGHT 157
 - NUMBEROFMATCHES 151
 - présentation 149
 - procédure mémorisée 162
 - recherche de texte 72
 - référence 149
 - SCORE 152
 - spécification des arguments de recherche 73
 - valeur table SQL 153
- fonctions clés 10
- fonctions de recherche
 - CONTAINS 150
 - HIGHLIGHT 157
 - NUMBEROFMATCHES 151
 - procédure mémorisée 162
 - SCORE 152
 - valeur table SQL 153
- fonctions exemples
 - exécution 78
- format des documents texte 23
 - description 23
 - liste de prise en charge 23
- fréquence de mise à jour 46
- FUZZY FORM OF mot clé 146

H

- HELP, commande
 - syntaxe 138
- HIGHLIGHT, fonction
 - exemple 158
 - modification de la taille de l'objet CLOB 103
 - syntaxe 157
 - utilisation de la fonction TEXTSEARCH 157

I

- identification et résolution des incidents 247
- index
 - activation de la mémoire cache 68
 - affichage de l'état de l'index de texte 49
 - affichage de l'état des index 69
 - calcul de taille 23
 - Centre de contrôle DB2 51
 - désactivation de la mémoire cache 69
 - effacement des événements d'index 48, 68
 - fréquence de mise à jour 46

- index (*suite*)
 - gestion des index de texte 45, 65, 85
 - mise à jour d'index de texte 46
 - mise à jour des index de texte 68
 - modification des paramètres de l'index de texte 47, 65
 - planification 23
 - présentation 3
 - relations 86
 - sauvegarder et restaurer 49
 - suppression d'index de texte 48
 - suppression des index de texte 67
 - utilisation de documents structurés 83
- initiation 19
- INSO, format
 - Voir le logiciel de filtrage Outside-In 15
- installation 13
- installation AIX 14
- installation du fichier jar Data Links 38
- installation pour un serveur DB2
 - partitionné 14
- installation sous Windows 15
- installation UNIX 14

L

- langues prises en charge 207
- logiciel de filtrage Outside-In
 - attribut de marque 232
 - définition d'un modèle de document 177
 - documents structurés 165
 - format de document 23
 - installation des bibliothèques 15
 - introduction 24
 - modèle de document par défaut 168

M

- migration 183
- modèle de document
 - définition de type de document 229
 - limites 232
 - par défaut 168
- modèle de document par défaut 168
- modèles de document
 - attributs de document 167
 - description 165
 - description de documents structurés 165
 - Exemple 165
 - modification 165
 - nom d'attribut dans la syntaxe de recherche 148
 - présentation 83
 - SECTION mot clé dans la syntaxe de recherche 145
 - zones de texte 167
- modification de la taille de renvoi de la fonction Datalink 38, 53
- mot clé COUNT 147
- mot clé EXPAND 147
- mot clé STOP SEARCH AFTER nombre DOCUMENTS(S) 144

- mot clé TERM OF 147
- Moteur de recherche de texte
 - Codes anomalie 239
 - configuration 236
 - langues prenant en charge les mots vides 236
 - mots vides 235
 - tokensation 235
- mots clé de l'argument de recherche
 - COUNT 147
 - EXPAND 147
 - EXPANSION LIMIT 144
 - FUZZY FORM OF 146
 - PRECISE FORM OF 146
 - RESULT LIMIT 144
 - SECTION 145
 - STEMMED FORM OF 146
 - STOP SEARCH AFTER nombre DOCUMENTS(S) 144
 - TERM OF 147
 - THESAURUS 147
- mots vides 235

N

- noms de répertoires et noms de fichiers 15
- normalisation d'un caractère 237
- NUMBEROFMATCHES, fonction
 - exemples 72
 - syntaxe 151

O

- occurrence
 - dans un résultat de recherche 72
 - Fonction NUMBEROFMATCHES 151
- occurrences du terme recherché 151
- opérateurs booléens
 - & (AND) et | (OR) 74
 - NOT 76
 - syntaxe de la recherche 144
- opérateurs de recherche primaire 145
- OR, opérateur booléen 74

P

- panneaux et boîtes de dialogue
 - Boîte de dialogue Activation de la mémoire cache 68
 - boîte de dialogue Désactivation de la mémoire cache 69
 - Boîte de dialogue Etat des index 69
 - Boîte de dialogue Evénements d'index 68
 - Boîte de dialogue Mise à jour de l'index 68
 - Boîte de dialogue Modification de l'index 65
 - Boîte de dialogue Suppression d'index 67
 - écran Cible 56
 - écran Mise à jour des caractéristiques 59
 - écran Nom 55
 - écran Propriétés de type texte 58

- panneaux et boîtes de dialogue (*suite*)
 - écran Récapitulatif 64
 - écran Table de mémoire cache 61
- performances, remarques
 - pour l'indexation 45
 - pour la recherche 82
- planification 23
- PRECISE FORM OF mot clé 146
- présentation 3
- présentation de DB2 Extension Net Search 3
- procédure mémorisée
 - activation d'un index de texte 42
 - création d'un index de texte sur un alias de type nickname à l'aide de DB2 Replication 38
 - désactivation d'un index de texte 42
 - index de texte sur une vue 43
 - mise à jour d'un index de texte 42
 - présentation 39
- profondeur des termes dans un thésaurus, spécification 147

Q

- quantités de mémoire 187

R

- récapitulatif des commandes
 - pour les administrateurs de base de données 97
 - pour les propriétaires d'instance 91
 - pour tous les propriétaires de table de texte 105
- recherche de texte
 - obtention du nombre d'occurrences trouvées 72
 - obtention du score d'un document trouvé 73
 - présentation 72
 - réalisation d'une requête 72
 - syntaxe 142
 - utilisation d'une fonction de valeur table 79
 - utilisation d'une procédure mémorisée 78
 - utilisation de la fonction HIGHLIGHT 80
- recherche de thésaurus
 - exemple 76
 - syntaxe 147
 - THESAURUS mot clé 147
- recherche des erreurs 247
- recherche floue, exemple 74
- recherche sur plusieurs colonnes, exemple 82
- reconnaissance
 - mot 235
 - mots vides 235
 - normalisation d'un caractère 237
 - paragraphe 235
 - phrase 235
 - Utilisation de caractères spéciaux dans un même mot 237
- recréation d'un index 46

- référence du modèle de document 229
- relation dans un thésaurus 86
- RESULT LIMIT mot clé 144
- rôles des utilisateurs
 - administrateurs de base de données 25
 - propriétaire d'instance DB2 25
 - propriétaires de table de texte 26

S

- sauvegarde et restauration d'index 49
- scénarios des utilisateurs
 - exemple de fonction de valeur table SQL 22
 - exemple de recherche de procédure mémorisée 21
 - exemple de recherche scalaire SQL 19
- SCORE, fonction
 - exemple 73
 - syntaxe 152
- Sémantique de l'expression XPath 230
- serveur
 - arrêt 95
 - démarrage 94
 - traçage des erreurs 247
- services de l'instance 9, 27, 91
- services de mise à jour 30
- services de verrouillage
 - affichage 29
 - commande CONTROL 92
 - utilisation 27
- START, commande
 - syntaxe 94
 - utilisation 27, 52
- STEMMED FORM OF mot clé 146
- STOP, commande
 - syntaxe 95
 - utilisation 27, 52

T

- table de capture de réplication 120, 199
- table de journalisation
 - création 35
 - description 5
- termes clés 3
- termes recherchés étendus
 - Voir thésaurus
- TEXTSEARCH, fonction
 - exemple 155
 - syntaxe 153
 - utilisation de la fonction HIGHLIGHT 153
- thésaurus
 - CCSID pris en charge 251
 - compilation 87
 - concepts 85
 - création 87
 - fichier de définitions 87
 - messages 253
 - structure 85
 - syntaxe de définition de thésaurus 179
 - traçage des erreurs 247

- types de clé primaire 127
- types de document 23
- types de données du document
 - conversion de types de données non pris en charge 37
 - types de données binaires 36
 - types de données DATALINK 37

U

- UPDATE INDEX, commande
 - fréquence de mise à jour 46
 - option RECREATE 46
 - pour une procédure mémorisée 42
 - syntaxe 134
 - utilisation 47
- Utilisation de caractères spéciaux dans un même mot 237
- utilisation de grandes quantités de mémoire 187
- HP-UX 189
- Linux 189
- pour AIX 187
- pour Sun Solaris 188
- pour Windows 188
- utilitaire de compilation de thésaurus 132

V

- vérification de l'installation 16
- vérification de l'installation sous UNIX 16
- vérification de l'installation sous Windows 16
- vue d'événements 197
- vue db2ext.dbdefaults 191
- vue db2ext.indexconfiguration 196
- vue db2ext.proxyinformation 193
- vue db2ext.textindexformats 196
- vue de la table de journalisation 198
- vues
 - db2ext.dbdefaults 33, 191
 - db2ext.indexconfiguration 33, 196
 - db2ext.proxyinformation 33, 193
 - db2ext.textindexes 33, 49, 194
 - db2ext.textindexformats 33, 196
 - présentation 9
 - table de capture de réplication 199
 - vue d'événements 197
 - vue de la table de journalisation 198
 - vues créées 33
- vues de catalogue 33



Référence : CT2XLFR

SH11-1287-01



(1P) P/N: CT2XLFR

