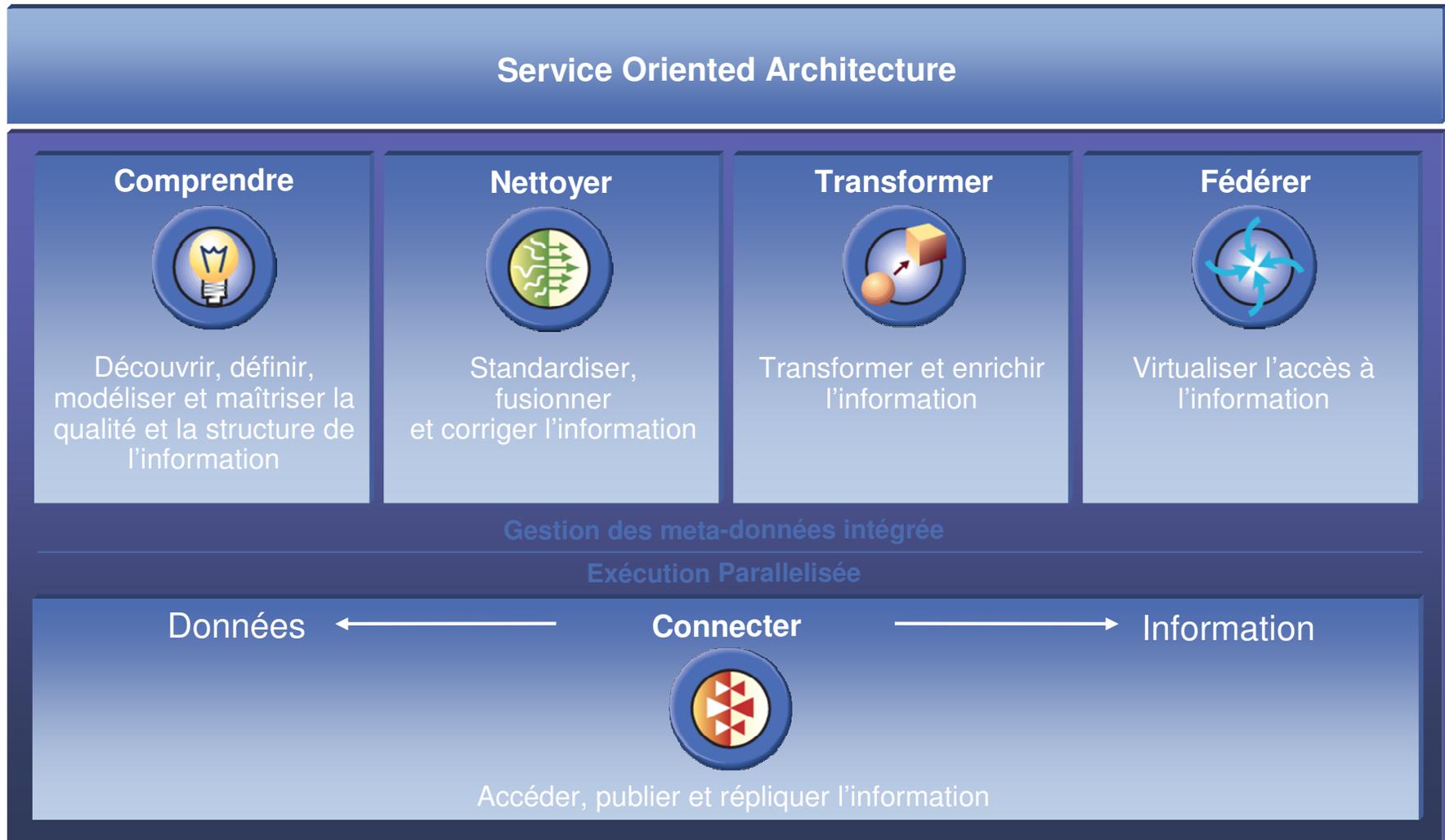


L'offre Information Integration Solutions en environnement Mainframe

Steven Haddad

IT Specialist Information Integration Services





- **L'ETL sur zSeries**
 - « WebSphere DataStage MVS »
 - Génération de code COBOL

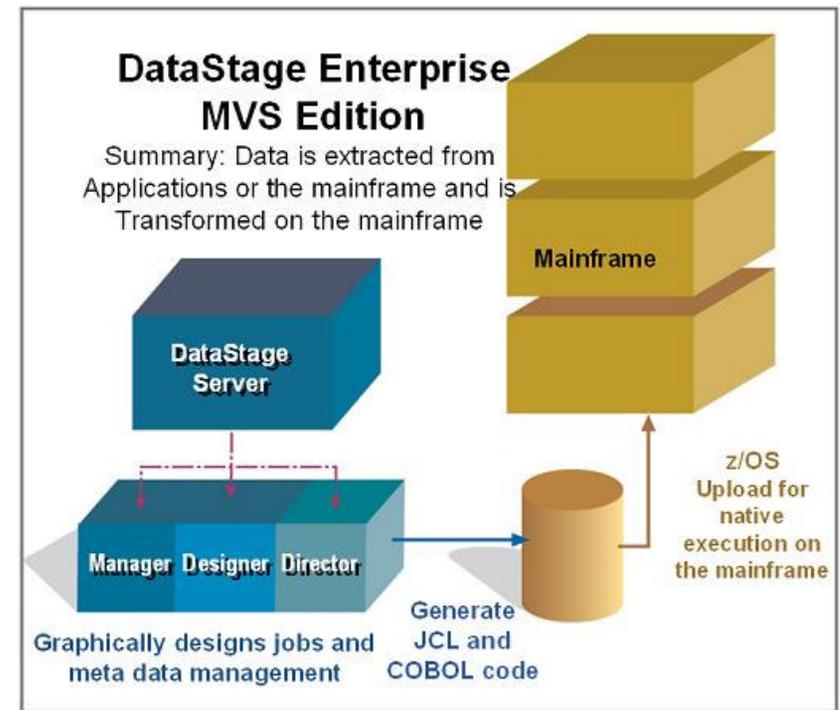
 - « WebSphere DataStage USS »
 - Moteur de transformation
 - Parallélisation active des traitements

- **La qualité de données**
 - WebSphere QualityStage USS
 - Solution de cleansing

- **Intégration en environnement SOA**
 - WebSphere DataStage SOA Edition

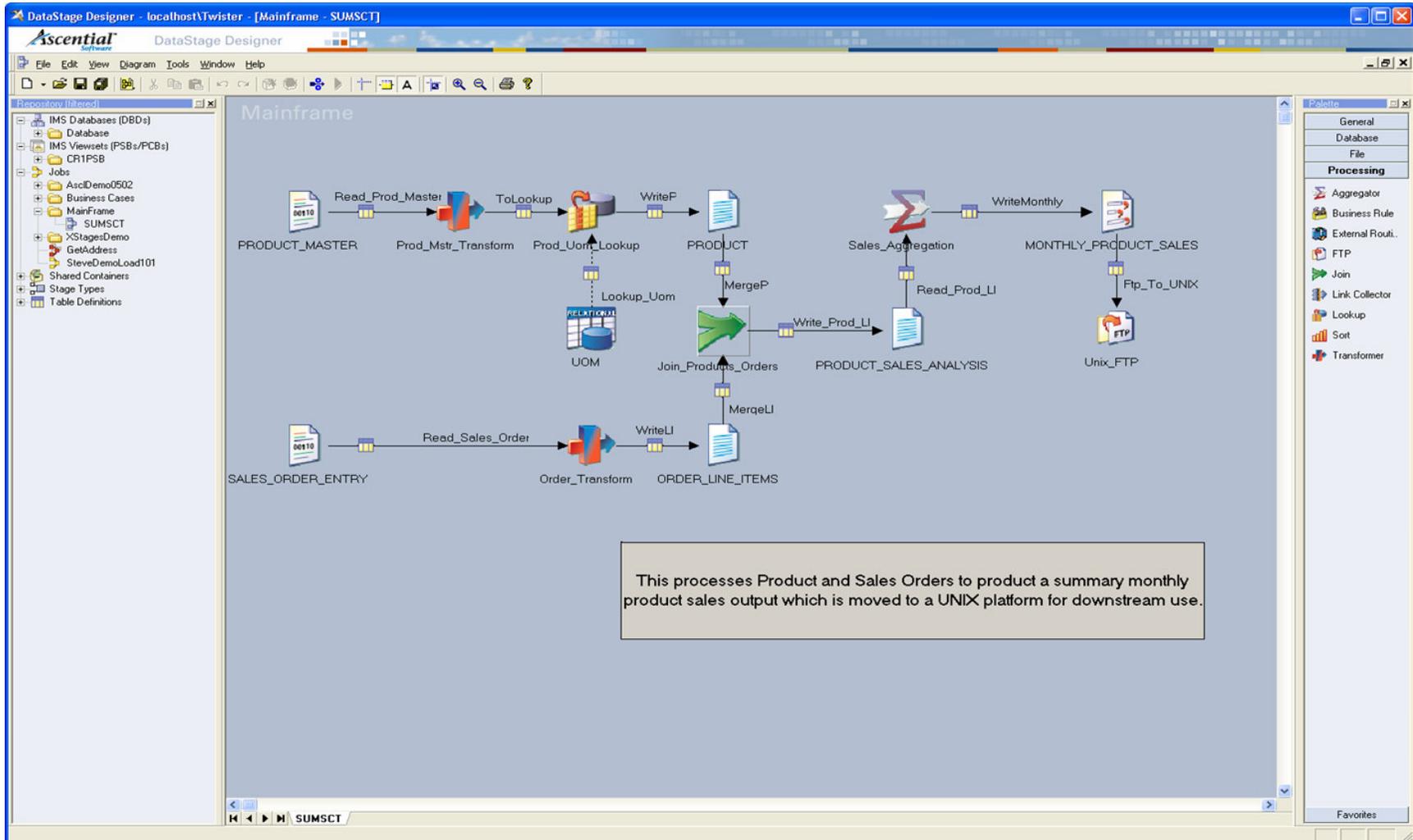


- Support natif OS/390 et z/OS
- Environnement de développement client-serveur
- Développement de flux de données MainFrame via la même interface graphique
 - Approche “Stages” et “liens”
 - Traitement ligne à ligne
 - Design like you think
 - Utilisation des méta-données des sources et cibles Mainframe





Exemple de flux de données avec WDS-MVS



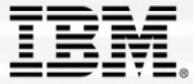


- **L'ETL sur zSeries**

- « WebSphere DataStage MVS »
 - Génération de code COBOL
- « WebSphere DataStage USS »
 - Moteur de transformation
 - Parallélisation active des traitements

- **La qualité de données**

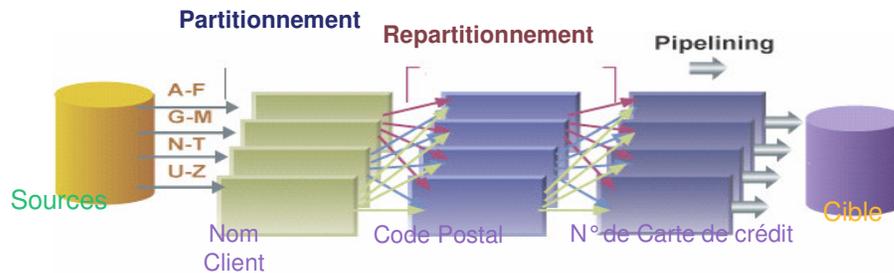
- « WebSphere QualityStage »
 - Solution de cleansing



WebSphere DataStage pour USS: exécution parallélisée



L'entreprise a besoin de performance



- Réalisation des processus d'intégration sans nécessité de prendre en compte les volumes de données et contraintes de temps
- Étapes simplifiées permettant de définir des partitionnement multiples dynamiques dans chaque processus si nécessaire
- Simple fichier de configuration pouvant être modifié pour ajouter de nouveaux processeurs
- Pas de programmation manuelle pour prendre en charge d'avantage de processeurs

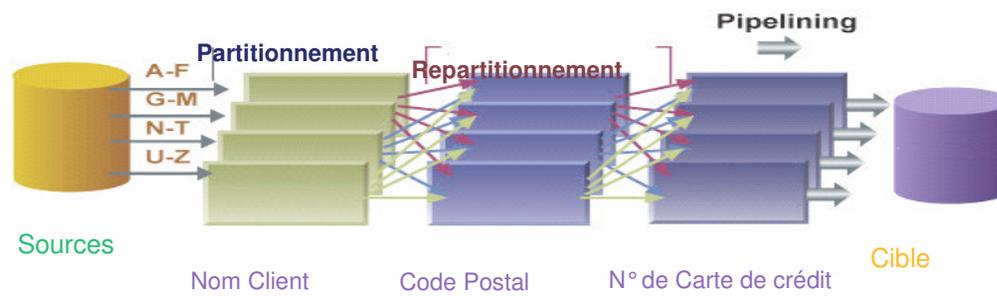
La problématique

- Compromis sur le niveau de service attendu
- Répondre aux demandes inopinées
- Rentabiliser les investissements matériels

Approches alternatives

- Fonctionner en mode 'single-thread'
- Décharger les données sur disque entre chaque opération
- Modifier les programmes lorsque l'on change de matériels
- Coder manuellement le parallélisme des bases et des thèmes

L'entreprise a besoin de performance



Le Pipelining

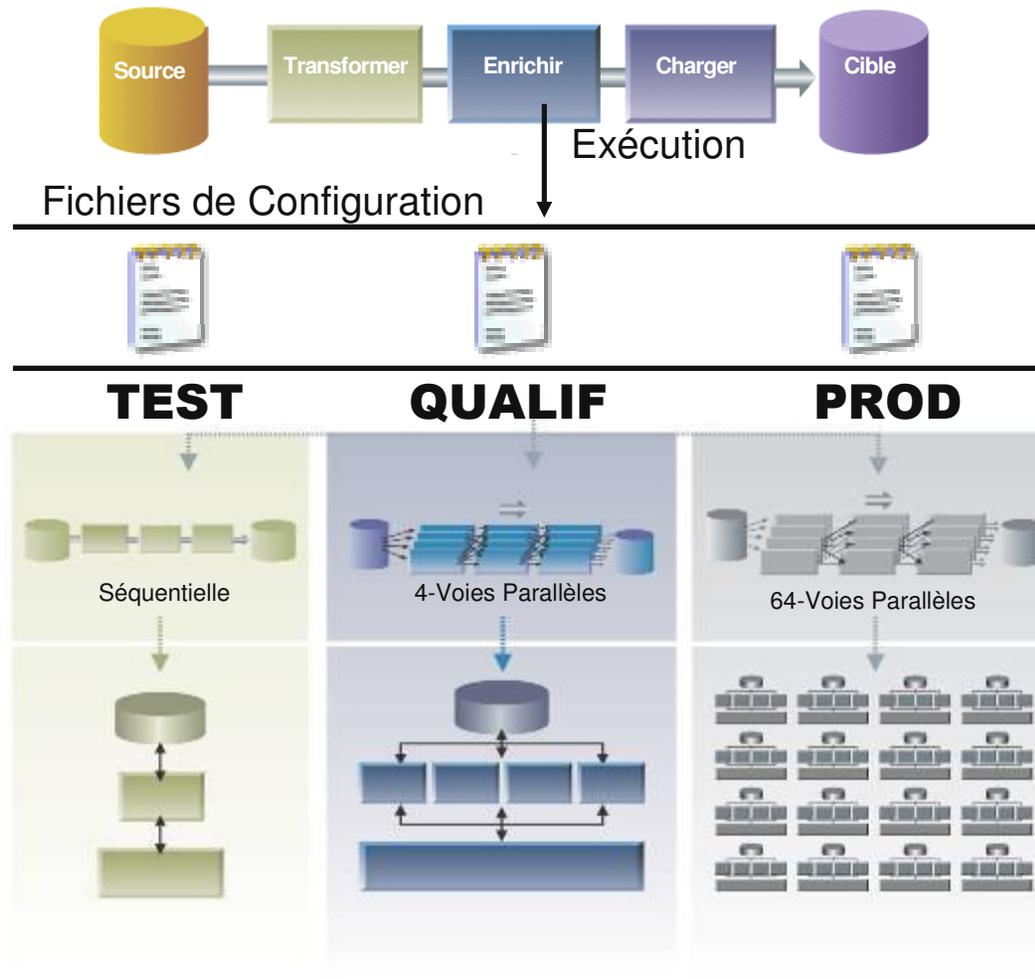
- Mode rapide (mémoire) d'acheminement des Données.

Le Partitionnement

- Une autoroute multivoies, bout en bout.

Le Repartitionnement

- Capacité à changer de voie, à la volée, en fonction des actions à réaliser





- **Fichiers**

- QSAM
 - Lecture séquentielle
 - Structures simples et complexes (CFF)
- VSAM
 - Lecture séquentielle
 - KSDS, ESDS, RRDS
- External Source / External Target
 - Utilisation de scripts comme sources / cibles des flux de données
- DataSet
 - Structure de fichiers répartis (propriétaire)

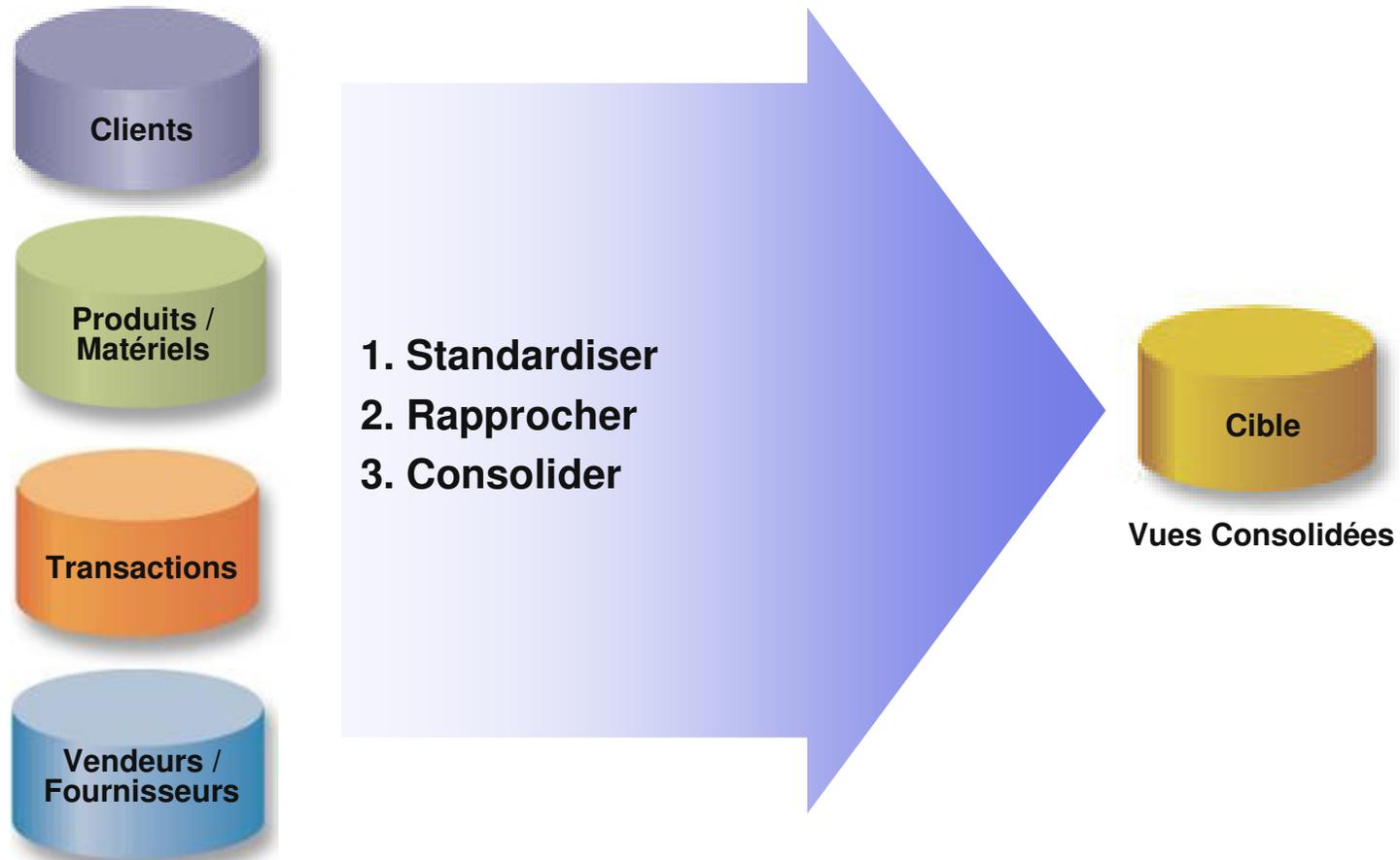
- **SGBDR**

- DB2 sur Z-Series
 - Lecture en parallèle de tables partitionnées
 - Utilisation en source, cible, ou lookup
 - Bulk load
 - Via WII, ouverture vers d'autres sources / cibles non z/OS
- Teradata

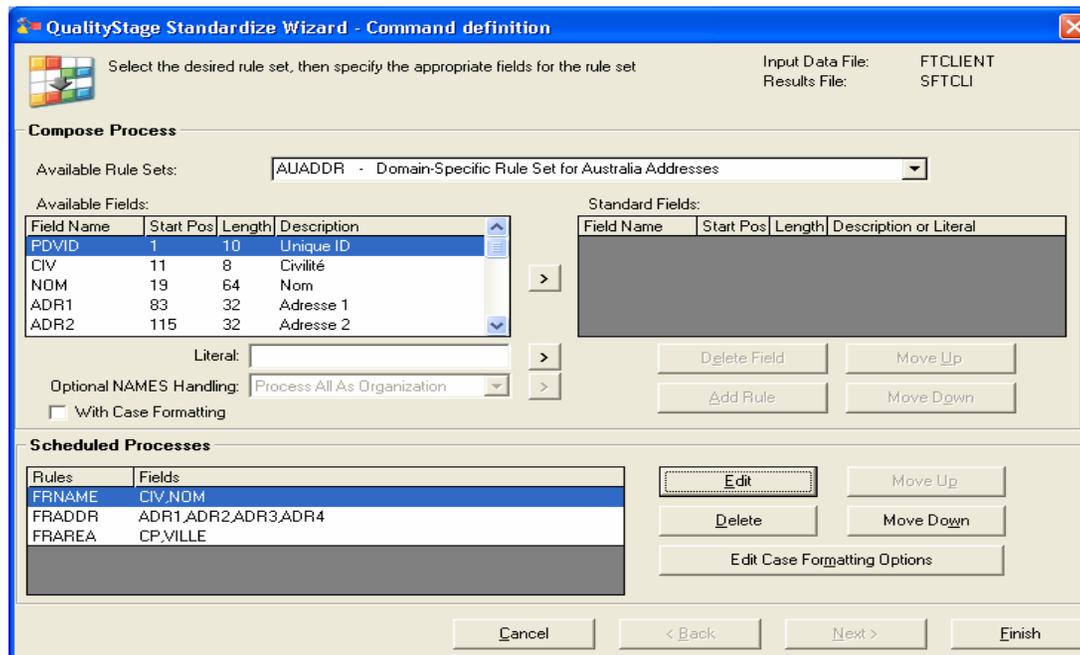


- **L'ETL sur zSeries**
 - « WebSphere DataStage MVS »
 - Génération de code COBOL
 - « WebSphere DataStage USS »
 - Moteur de transformation
 - Parallélisation active des traitements
- **La qualité de données**
 - WebSphere QualityStage USS
 - Solution de cleansing
- **Intégration en environnement SOA**
 - WebSphere DataStage SOA Edition

Un processus en 3 étapes

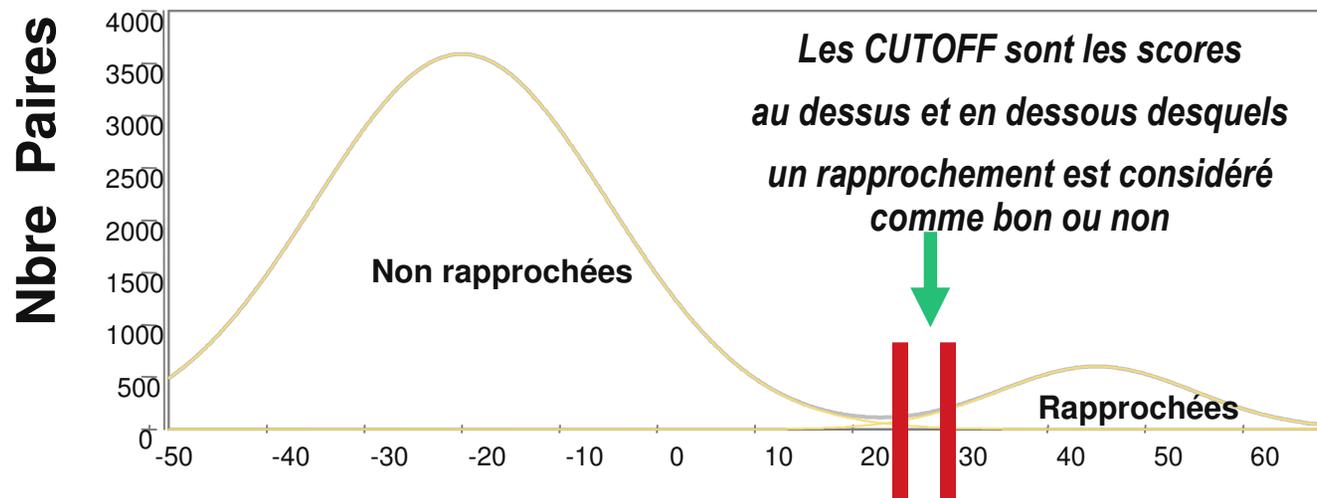


- Application des règles Françaises livrées en standard avec WebSphere QualityStage
 - ▶ FRNAME sur le couple Civilité + Nom
 - ▶ FRADDR sur la concaténation de ADR1+ADR2+ADR3+ADR4
 - ▶ FRAREA sur le couple Code Postal + Ville



<u>Prénom</u>	<u>2ème Prénom</u>	<u>Nom</u>	<u>Fonction</u>	
ALEXANDRE	J	DEMARIA	DG	
ALEXANDRE	JEAN	DEMARA	DG	
+7	+1	+10	+5	= 23

Le score d'un poids est une mesure relative de probabilité de match



Détermine la meilleure donnée à garder “Best of Breed”

- ✓ La plus fraîche
- ✓ La plus fréquente
- ✓ Selon la source
- ✓ La plus complète

....

Résout les conflits de valeur et complète les valeurs manquantes

Crée les formats de sortie:

- ✓ Tables relationnelles avec clés
- ✓ Transactions pour mettre à jour les bases de données
- ✓ Fichiers de références croisées, tables de synonymes
- ✓ Pistes d’audit, rapports divers

555 36,36 MME MARIE-NOELLE CAPDEVILA 3 RUE FOURNARIE 34000 MONTPELLIER 23976
 555 14,09 M JEAN-LOUP CAPDEVILA 3 RUE FOURNARIE 34000 MONTPELLIER 23830



Groupe	Nom	No	Type	Nom	Code	Ville
		Rue	Rue	Postal		
555	CAPDEVILA 3	RUE	FOURNARIE	34000	MONTPELLIER	

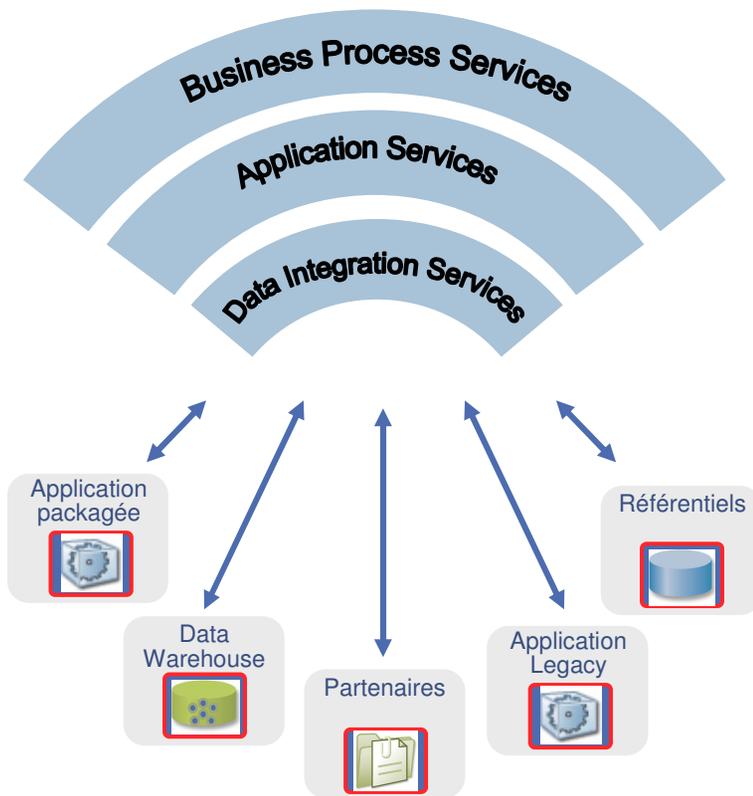
Groupe	Legacy
1	23976 23830

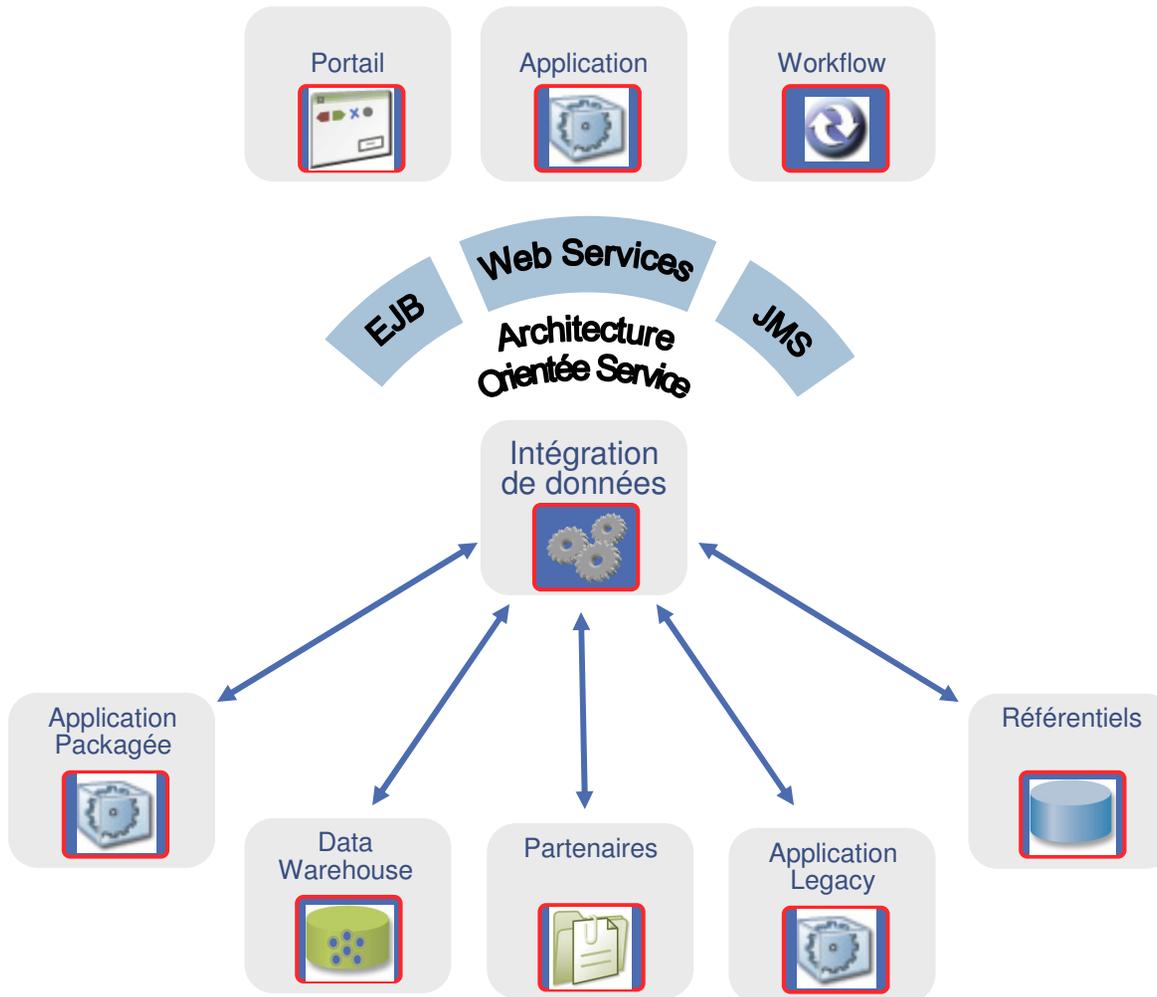


- **L'ETL sur zSeries**
 - « WebSphere DataStage MVS »
 - Génération de code COBOL
 - « WebSphere DataStage USS »
 - Moteur de transformation
 - Parallélisation active des traitements
- **La qualité de données**
 - WebSphere QualityStage USS
 - Solution de cleansing
- **Intégration en environnement SOA**
 - WebSphere DataStage SOA Edition



- SOA Edition met en place une architecture orientée service dans un contexte d'intégration de données
- Il permet de rendre accessible un ensemble de services qui masquent la complexité de l'intégration de données
- Ces services régissent leur propre contrôle sur les données (validation, standardi-sation, transformation)
- Ils laissent le soin aux spécialistes de gérer comment les données sont utilisées, sans en limiter l'accès





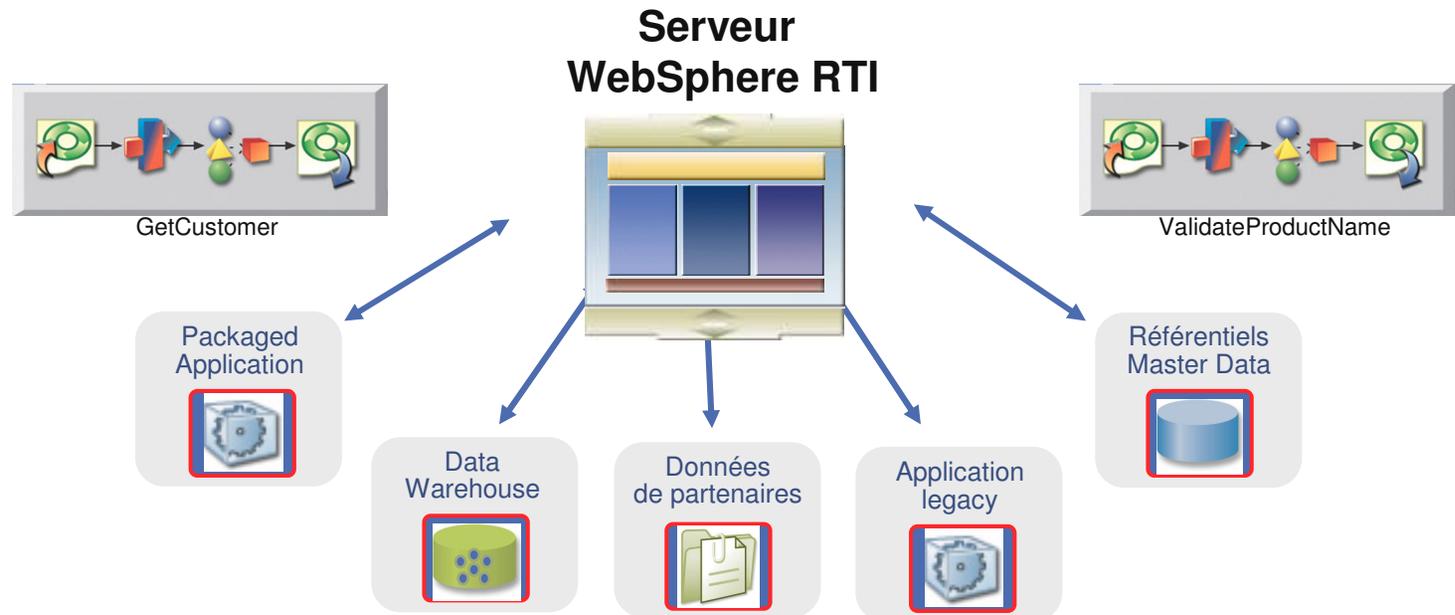
- « Package » la logique d'intégration de donnée (jobs ou maps) en services qui isolent les développeurs des technologies employées
- Permet à ces services d'être invoqués via EJB, JMS, ou Web Services (SOAP/HTTP)
- Apporte une fonction d'équilibrage de charge et de distribution du traitement sur plusieurs serveurs DataStage et DataStage TX
- Intègre les fondations d'une infrastructure de services (sécurité, logging, monitoring)

Service global d'accès aux données

- **Standardisation / harmonisation des données**
- **Lecture / écriture des données disparates et hétérogènes**
- **Rapprochement et consolidation des enregistrements**
- **Logique de transformation pour renvoyer des données métier**

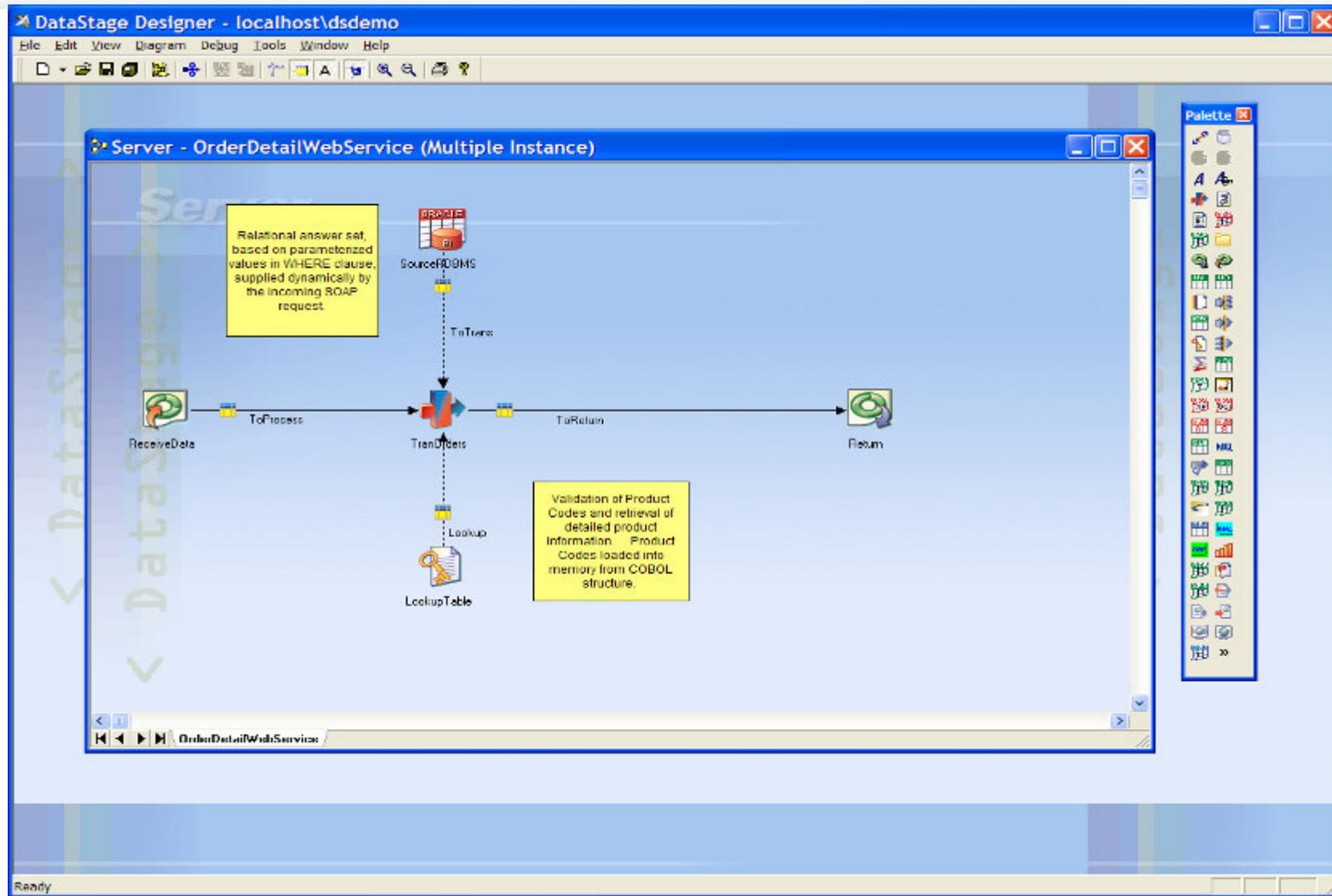
Service de transformation de données

- **Validation, enrichissement, transformation ou standardisation**
- **Exemple : XML vers EDI, validation d'adresse, standardisation de code produit etc.**

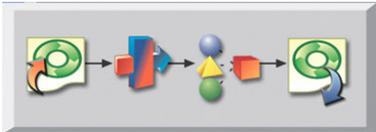




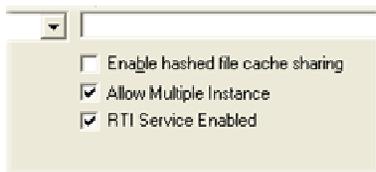
Un job WebSphere DataStage (Services) : un job comme un autre



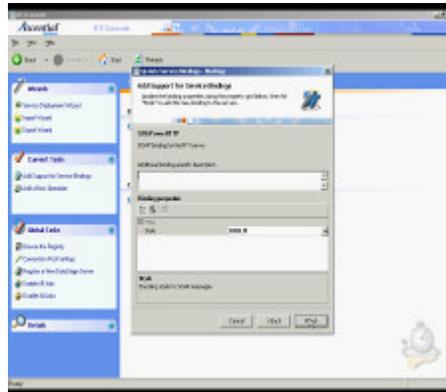
1 Job WebSphere DataStage



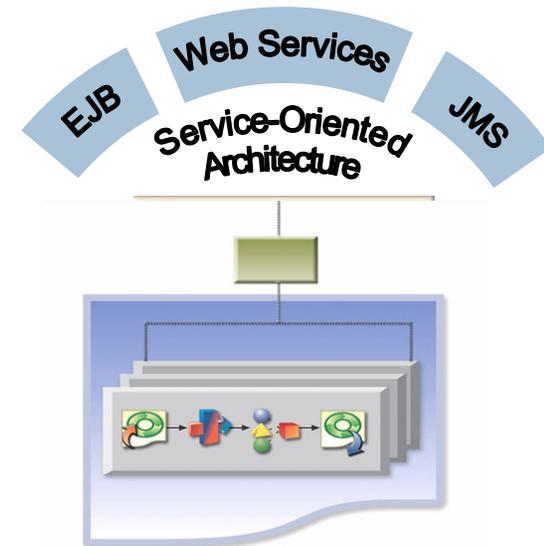
- Crée / réutilise un job avec un SOA Input et/ou Output
- « Déclare » le job comme pouvant être exposé comme un service SOA



2 Console WebSphere RTI

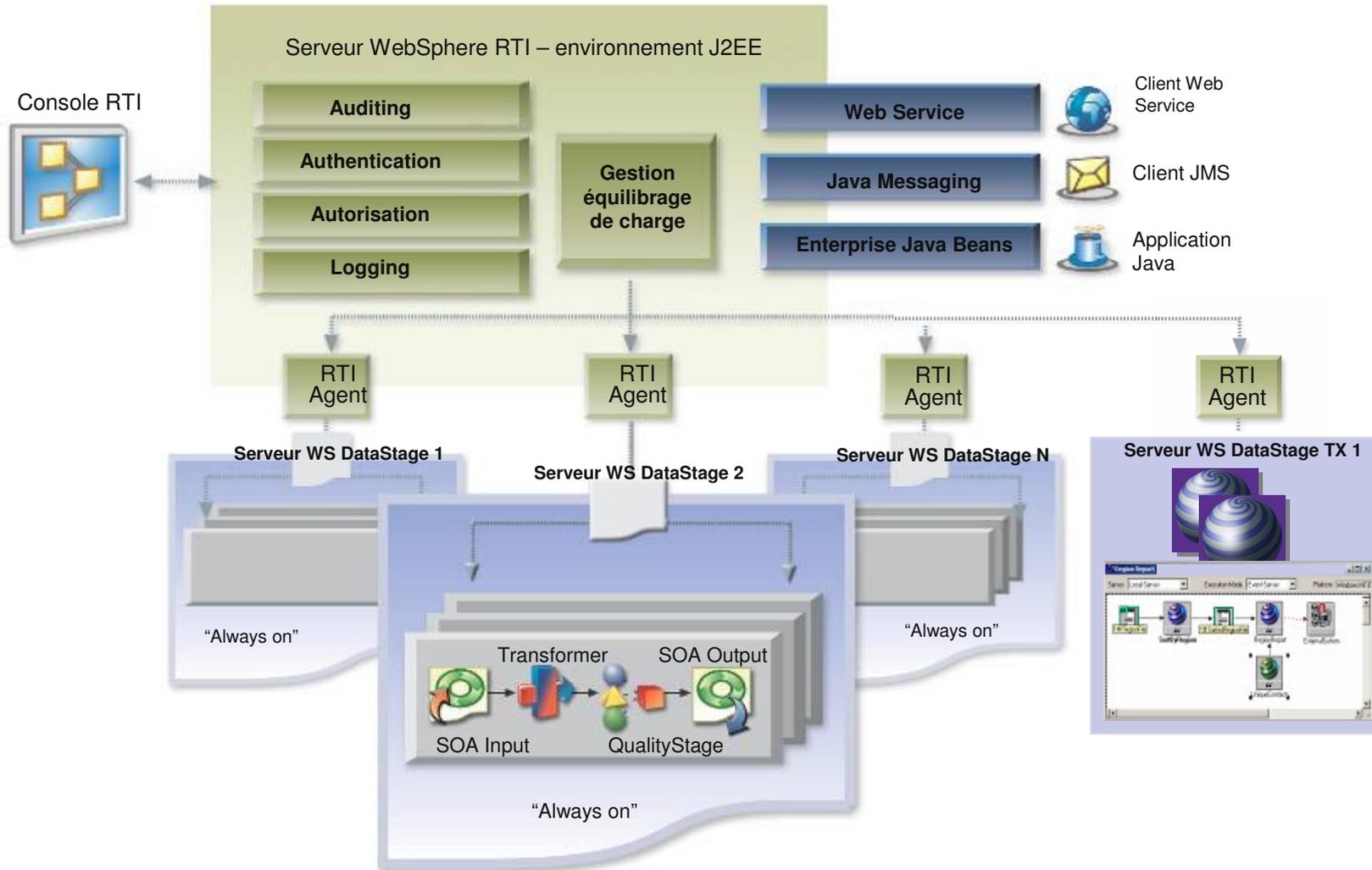


- Choix du binding, paramétrage...
- Par assistants graphiques, avec valeurs par défaut



- Décodage / encodage des messages au fil de l'eau
- Pooling de jobs « always on », load-balancing

**Création d'un service
d'intégration de données
sans codage !**





Merci

steven.haddad@fr.ibm.com