



L'offre Websphere Data Integration Solution (ex Ascential)

▪ L'ETL sur zSeries

– « WebSphere DataStage MVS »

- Génération de code COBOL

– « WebSphere DataStage USS »

- Moteur de transformation
- Parallélisation active des traitements

▪ La transformation de messages

– « WebSphere DataStage TX »

- La transformation de messages en environnement MainFrame

▪ La qualité de données

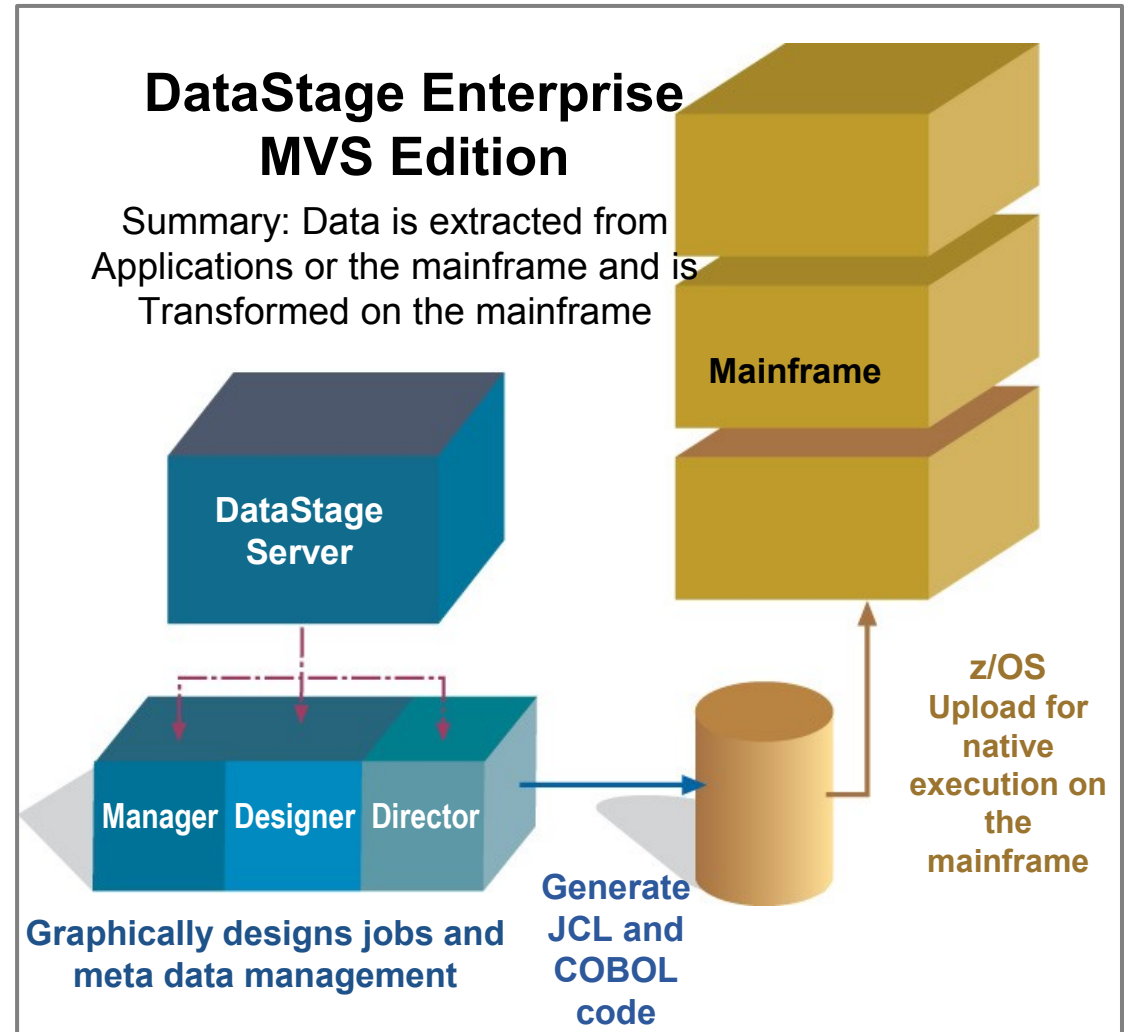
– « WebSphere QualityStage »

- Solution de cleansing

WebSphere DataStage Enterprise MVS Edition

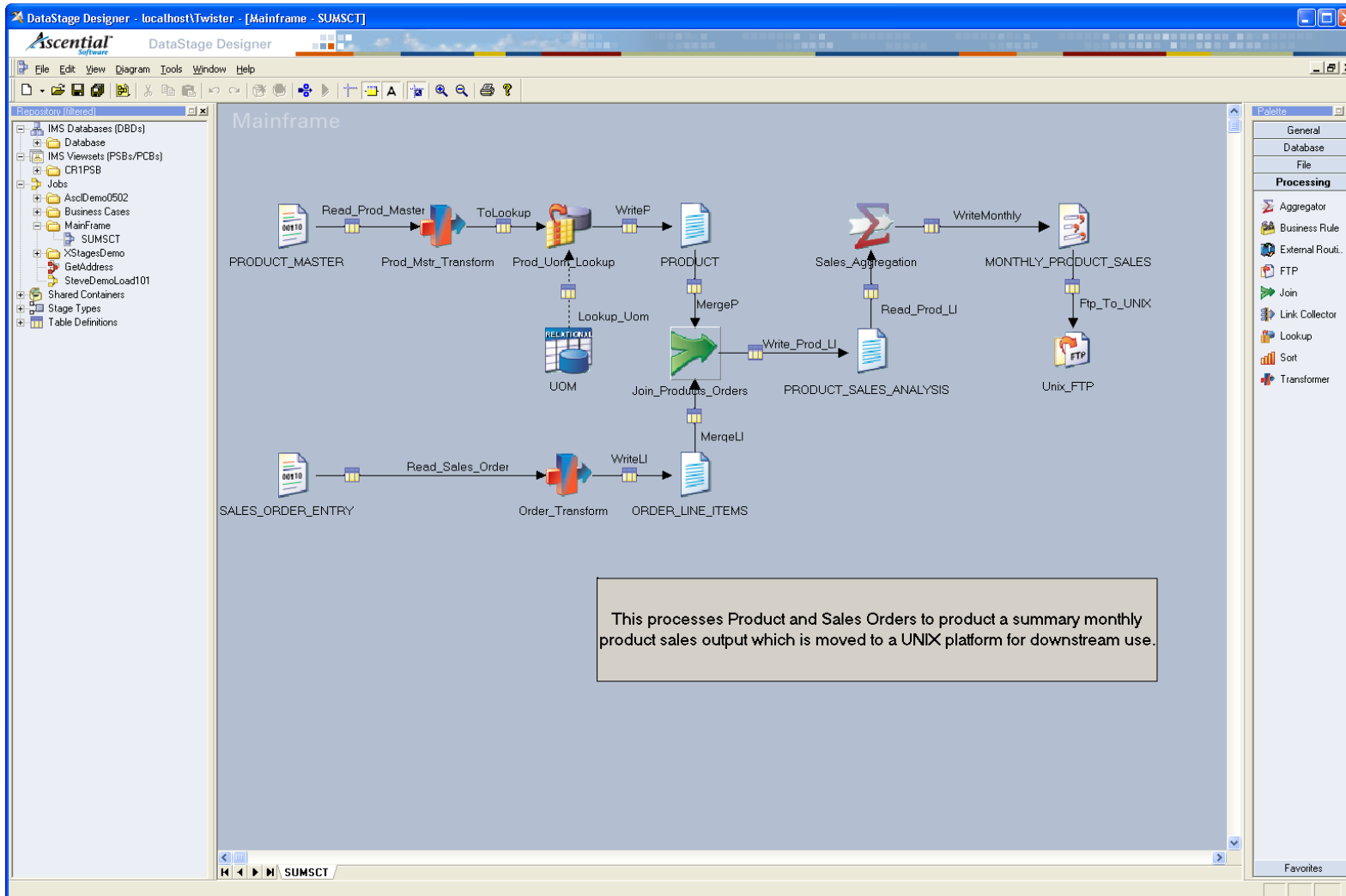


- Support natif OS/390 et z/OS
- Environnement de développement client-serveur
- Développement de flux de données MainFrame via la même interface graphique
 - Approche “Stages” et “liens”
 - Traitement ligne à ligne
 - Design like you think
 - Utilisation des méta-données des sources et cibles Mainframe





Exemple de flux de données avec WDS-MVS





L'offre Websphere Data Integration Solution (ex Ascential)

▪ L'ETL sur zSeries

– « WebSphere DataStage MVS »

- Génération de code COBOL

– « WebSphere DataStage USS »

- Moteur de transformation
- Parallélisation active des traitements

▪ La transformation de messages

– « WebSphere DataStage TX »

- La transformation de messages en environnement MainFrame

▪ La qualité de données

– « WebSphere QualityStage »

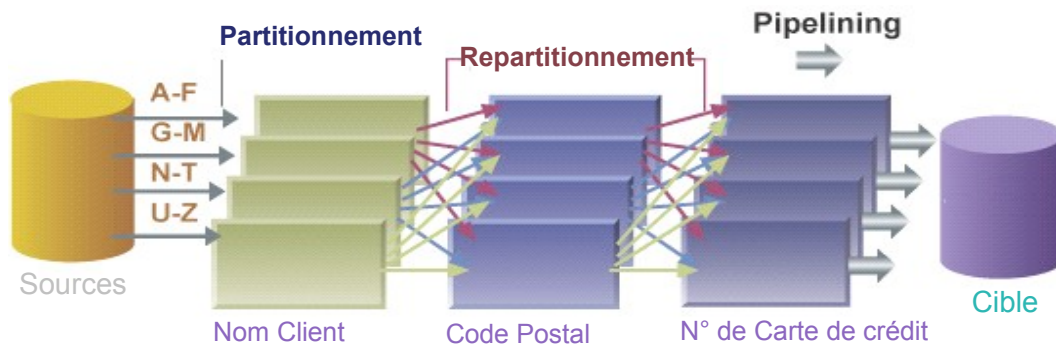
- Solution de cleansing



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
- Etapes 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

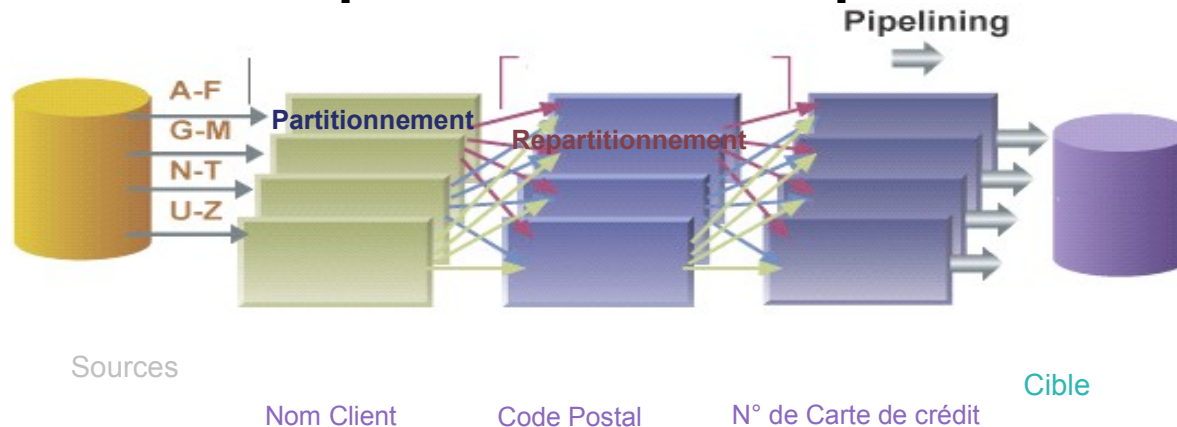
- Fonctionner en mode 'single-thread'
- Décharger les données sur disque entre chaque opération
- Modifier les programmes lorsque l'on change de matériels
- Coder manuellement le parallélisme des bases et des thèmes

IBM Software et IBM System & Technology Groups



Partitionnement + Pipelining = Puissance

L'entreprise a besoin de performance



Le Pipelining

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

Le Partitionnement

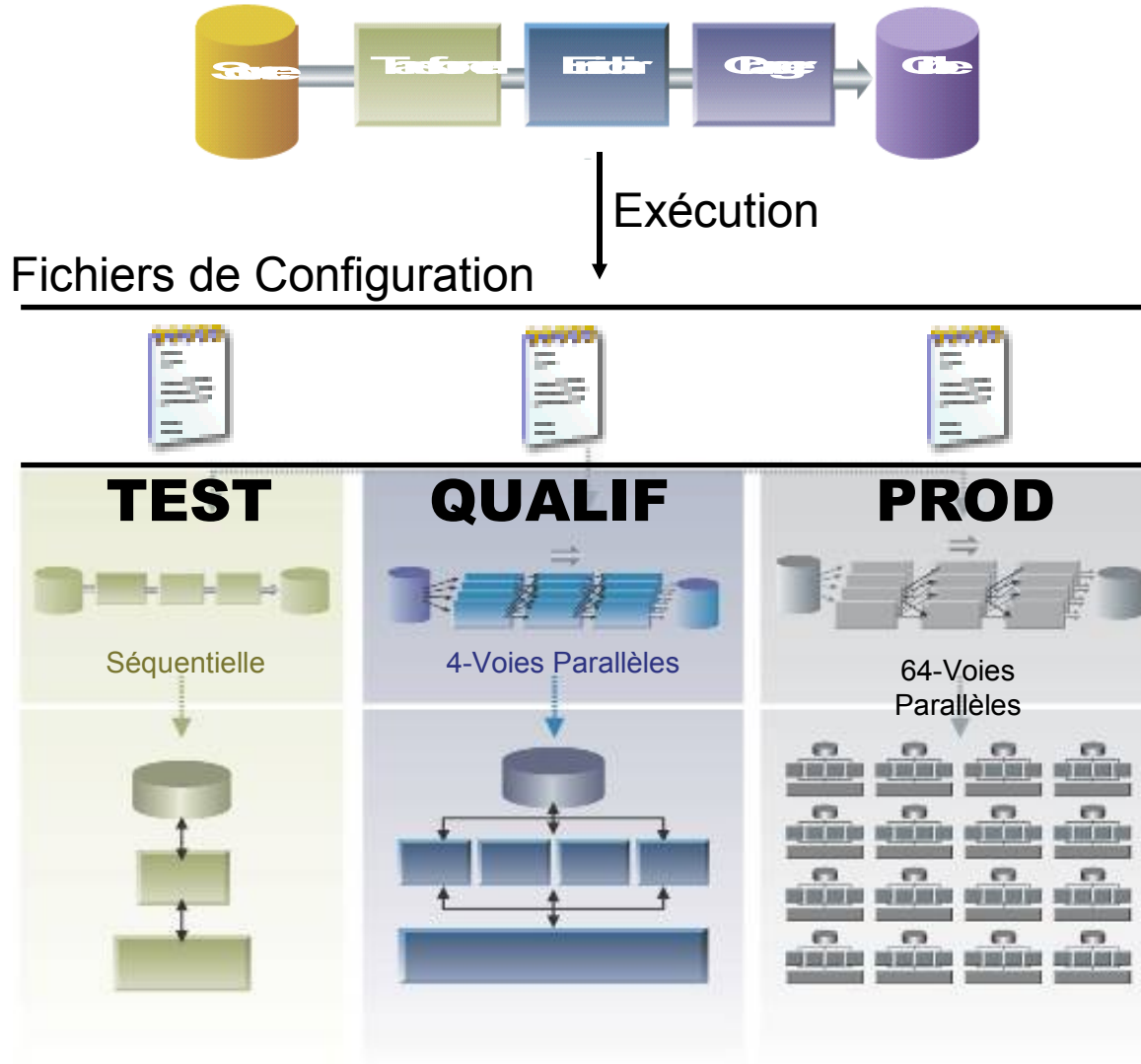
- Une autoroute multivoies, bout en bout.

Le Repartitionnement

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



Déploiement sans retouche des développements





DataStage USS – Connecteurs

▪ 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'offre Websphere Data Integration Solution (ex Ascential)



▪ L'ETL sur zSeries

– « WebSphere DataStage MVS »

- Génération de code COBOL

– « WebSphere DataStage USS »

- Moteur de transformation
- Parallélisation active des traitements

▪ La transformation de messages

– « WebSphere DataStage TX »

- La transformation de messages en environnement MainFrame

▪ La qualité de données

– « WebSphere QualityStage »

- Solution de cleansing



IBM WebSphere Datastage TX

- Transformation et routage des données basés sur le contenu
 - Lecture, transformation, routage et écriture des données dans leur format natif
- Conception graphique sans codage manuel
 - Glisser-déposer de composants prédéfinis
 - Tests conditionnels, Lookups, Fonctions mathématiques, Chaînes de caractères
 - Réutilisation possible de code existant (C, Java, ...)
- Adaptateurs intelligents
 - Fichiers, Bases de données, Messagerie, Bus de messages, Applications
- Options de déploiement flexible
 - Intégration sans modification des systèmes sources et cibles
 - Support multi-plateformes (Windows, AIX, Linux, z/OS)
 - Déclenchement batch ou événementiel : ligne de commande, API, WAS
- Support natif des formats métiers
 - EDIFACT, ANSI X12, EANCOM, TRADACOMS, HL7, HIPAA, SWIFT





Qu'est que IBM® WebSphere® DataStage® TX fait?

Lecture des données dans leur format natif

TOUTES structures des données

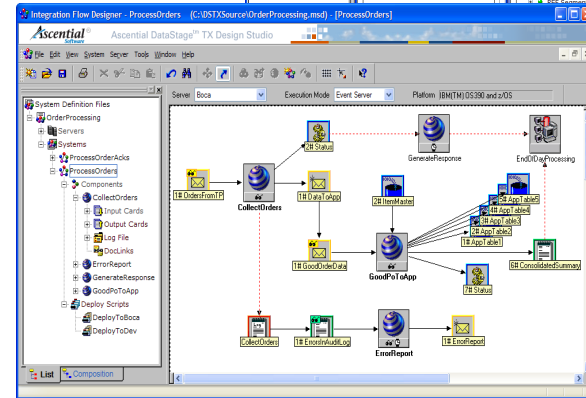
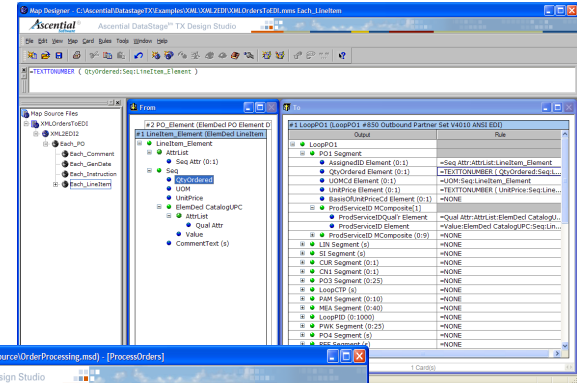


Transformation et routage sans programmation

Intégration N à N

Ecriture des données dans leur format cible

TOUTES structures des données

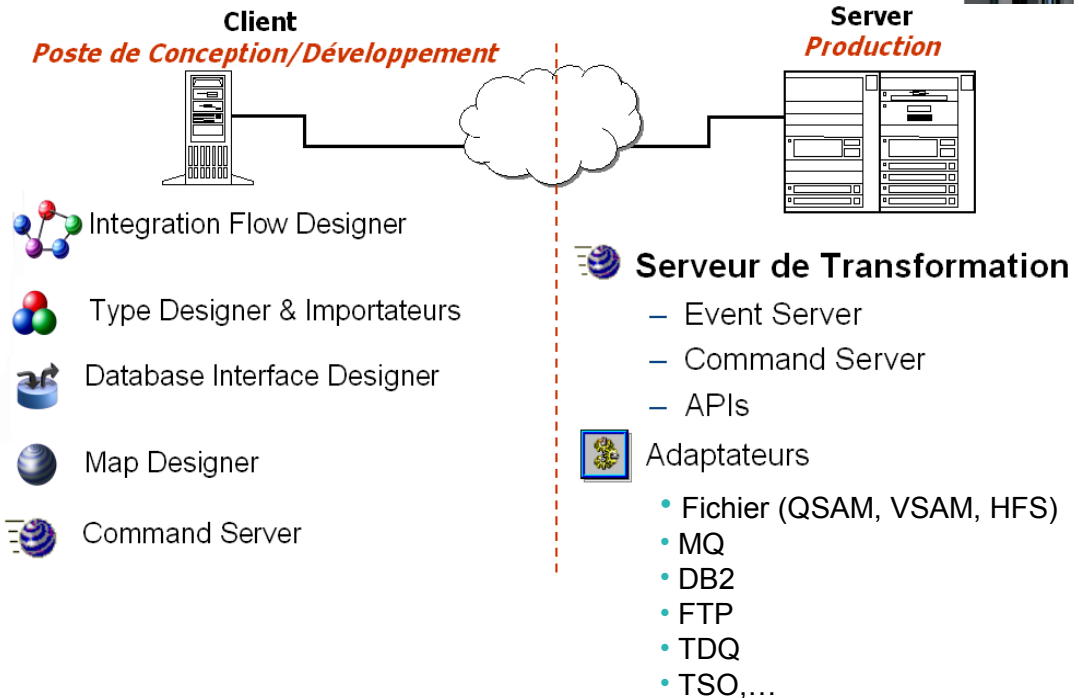




Du développement à l'exploitation...

■ Développement : le « Design Studio »

- Outil totalement graphique sous Windows
- Définition des :
 - Structures de données à manipuler
 - Transformations (interfaces) à assurer
 - Processus à mettre en œuvre
 - Adaptateurs techniques pour accéder aux données
- Exécuter les tests unitaires



■ Exploitation : le « Integration Server »

- Moteur qui embarque les objets définis dans le Design Studio – événementiel Temps réel et/ou batch
- 3 modes d'exécution sur Z/OS:
 - BATCH
 - CICS – Transaction DSTX
 - USS – moteur événementiel (Event Server)

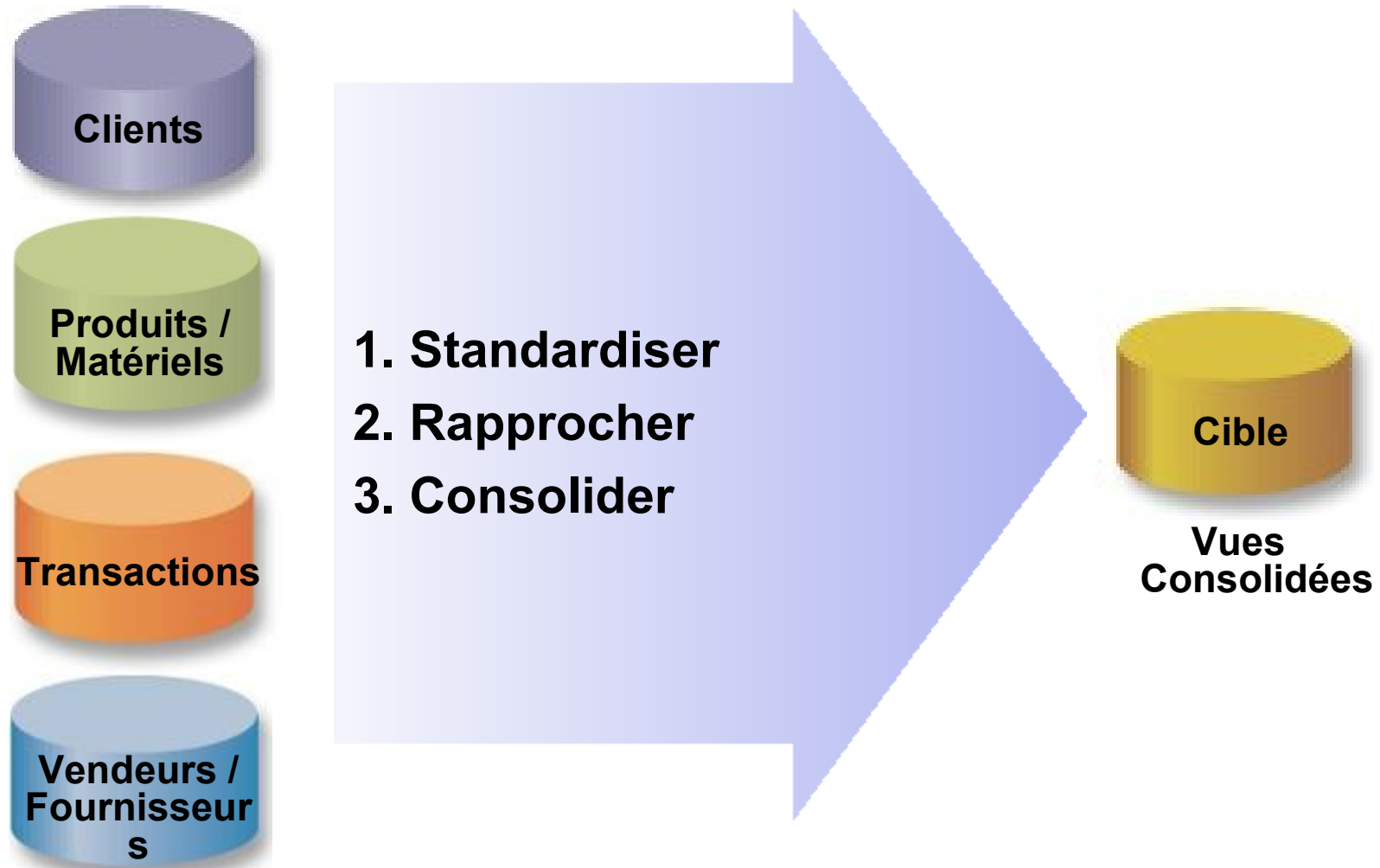


L'offre Websphere Data Integration Solution (ex Ascential)

- **WebSphere DataStage pour Z-Series**
 - ETL Mainframe
- **WebSphere DataStage USS**
 - Le parallélisme actif sur Z-Séries
- **WebSphere DataStage TX**
 - La transformation de messages en environnement MainFrame
- **WebSphere QualityStage**
 - Solution de cleansing



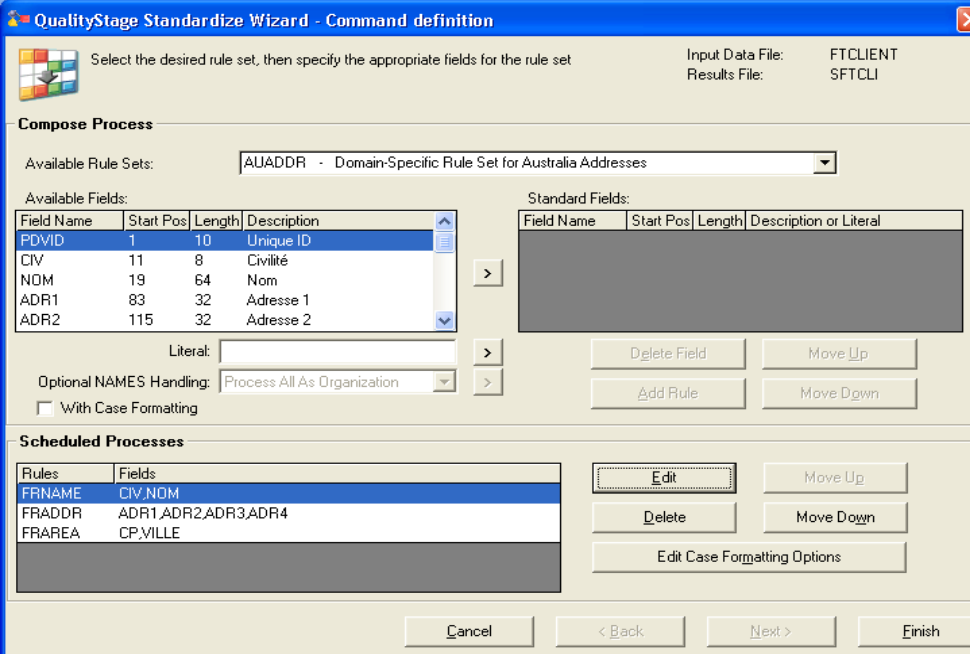
Un processus en 3 étapes





Standardiser

- 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



QualityStage Standardize Wizard - Command definition

Select the desired rule set, then specify the appropriate fields for the rule set

Input Data File: FTCLIENT
Results File: SFTCLI

Compose Process

Available Rule Sets: AUADDR - Domain-Specific Rule Set for Australia Addresses

Available Fields:

Field Name	Start Pos	Length	Description
PDVID	1	10	Unique ID
CIV	11	8	Civilité
NOM	19	64	Nom
ADR1	83	32	Adresse 1
ADR2	115	32	Adresse 2

Standard Fields:

Field Name	Start Pos	Length	Description or Literal
------------	-----------	--------	------------------------

Literal:

Optional NAMES Handling: Process All As Organization

With Case Formatting

Scheduled Processes

Rules	Fields
FRNAME	CIV,NOM
FRADDR	ADR1,ADR2,ADR3,ADR4
FRAREA	CP,VILLE

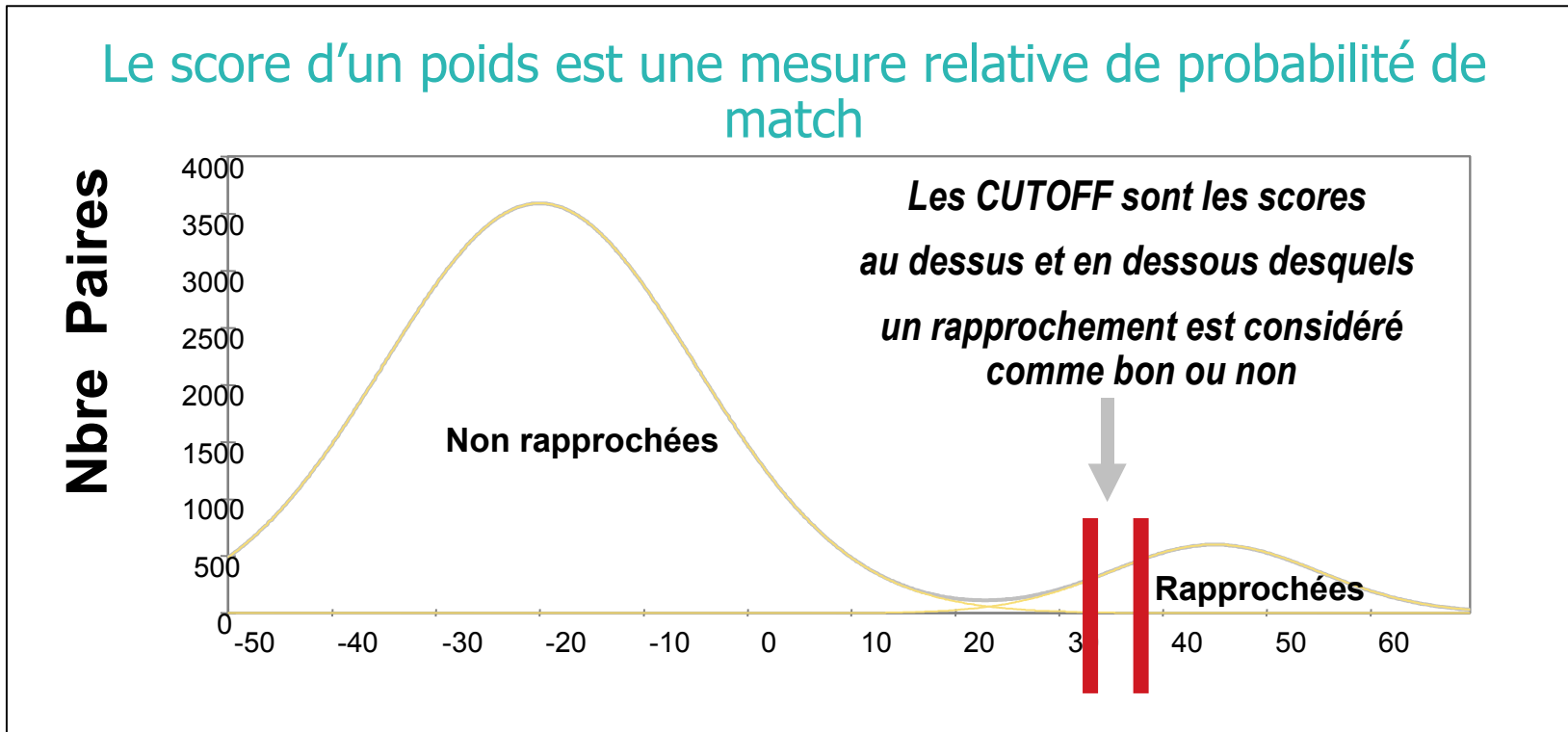
Buttons: Edit, Move Up, Delete, Move Down, Edit Case Formatting Options

Cancel < Back Next > Finish



Rapprocher

<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





Consolider ... et ne rien perdre

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