



IBM Software Group

Using IBM IPS products to synchronize your non-relational system Z operational data to relational data targets

WebSphere. Classic Federation
and Classic Data Event Publishers

@.business on demand software

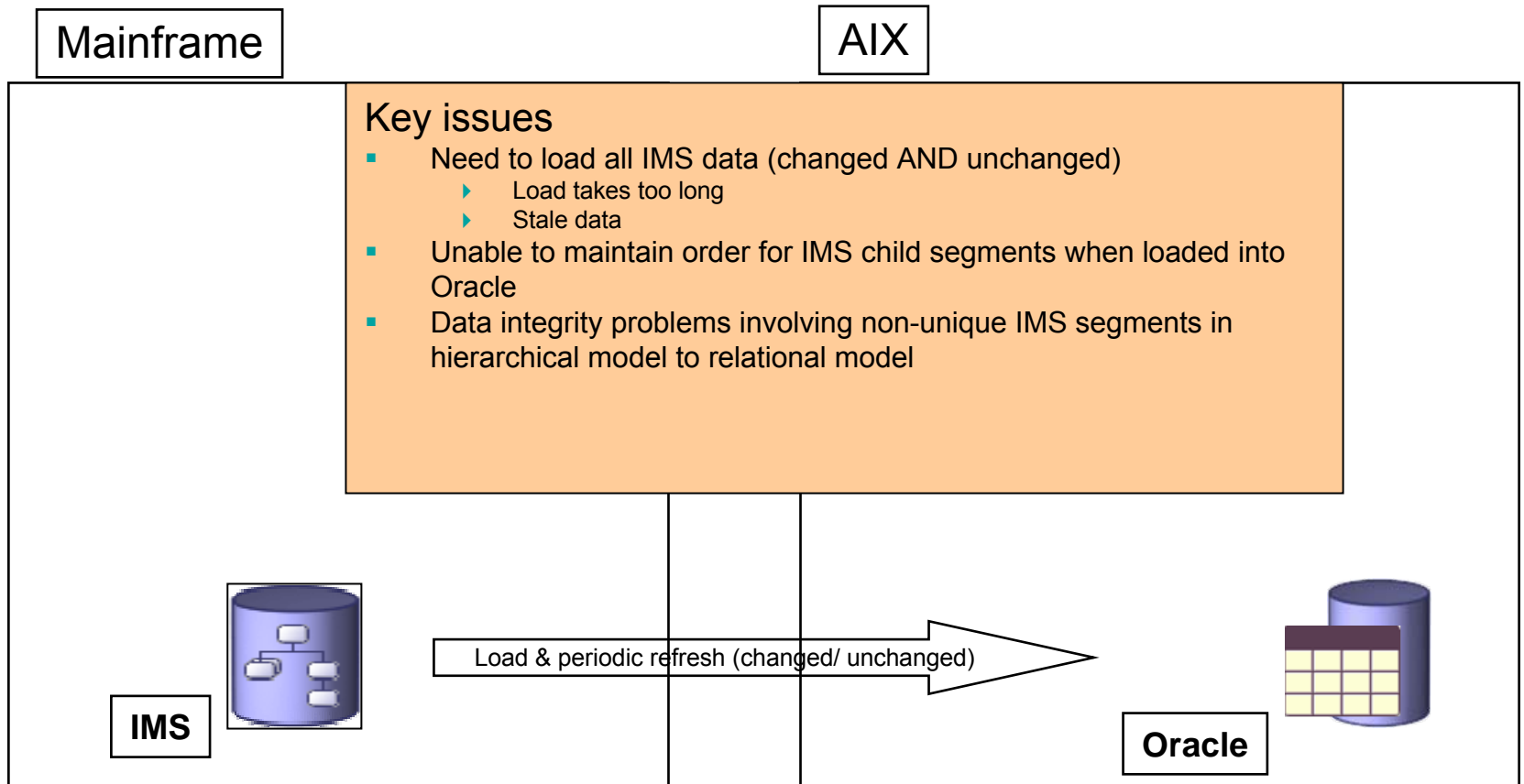
Raj Datta, Sr. Solutions Architect
IPS/z Enablement and Deployment
rajorshi@us.ibm.com

Agenda

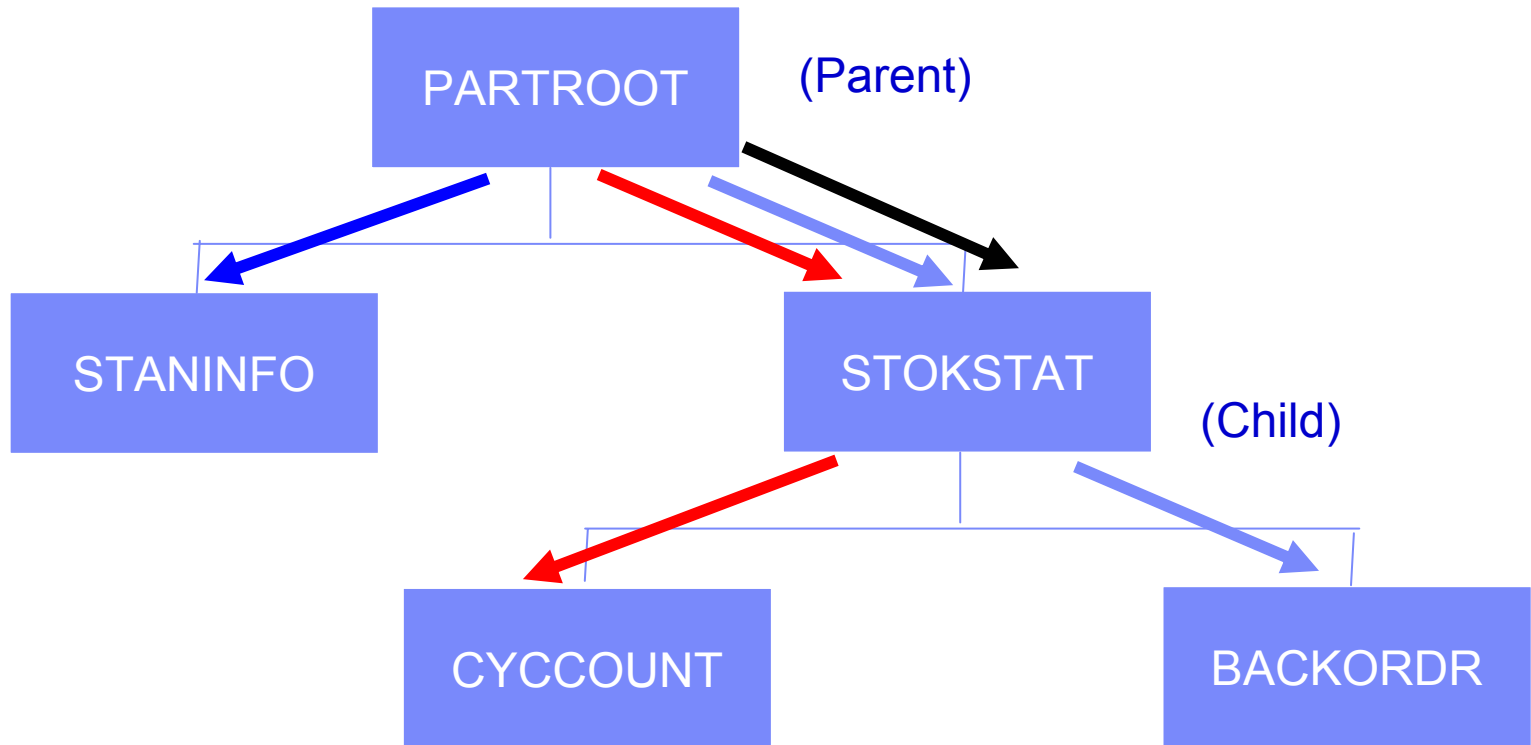
- Introduction to relational mapping of non-relational data
- Federation: Transitioning the applications
- Data Event Publishing: Data Event-based Syncing



DWH Sync problem at US State Govt. Judicial Information Systems



IMS Database Overview DI21PART Sample Database

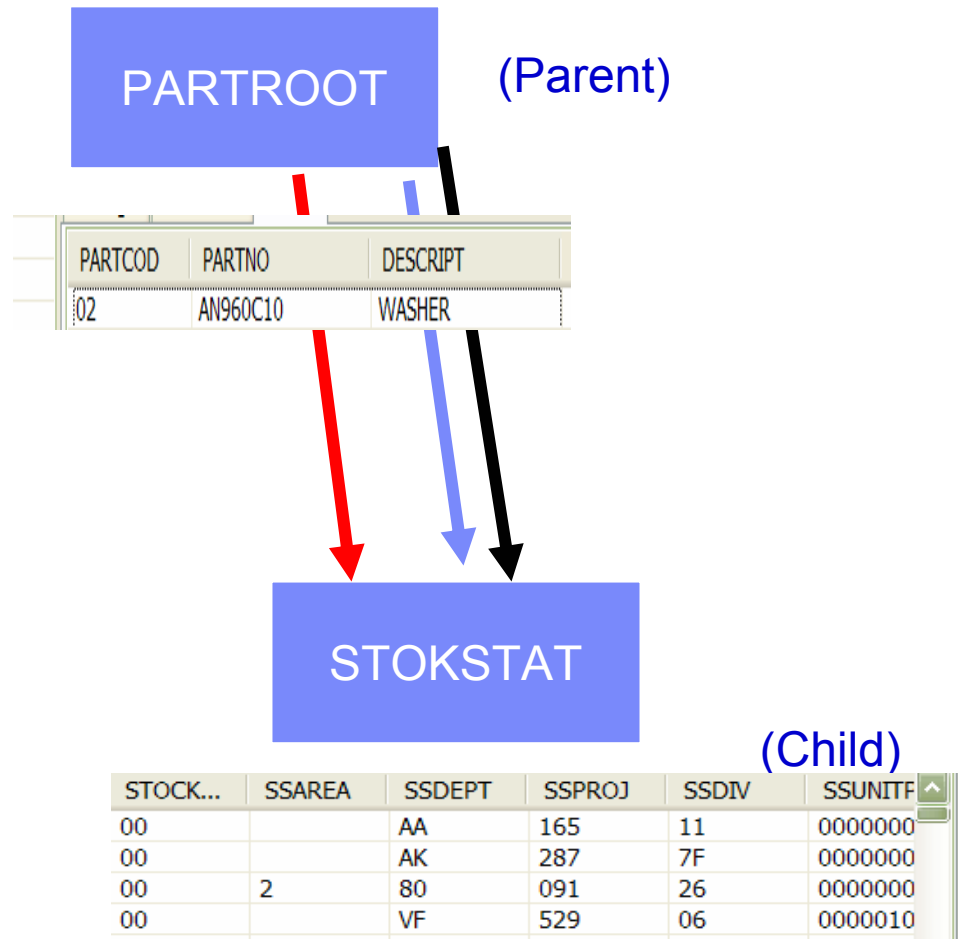


DBD Source - DI21PART Sample Database

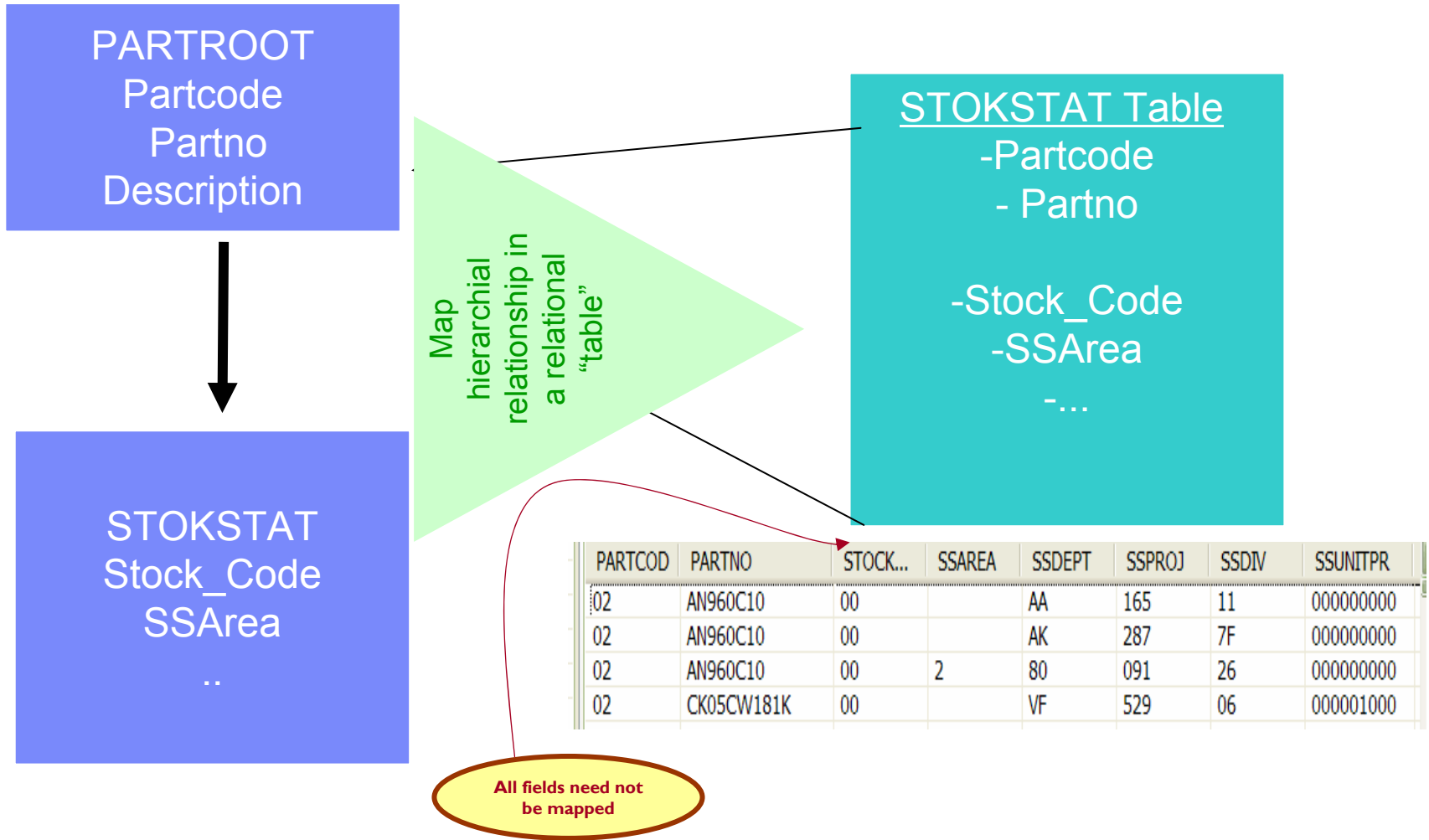
```
DBD  NAME=DI21PART,ACCESS=(HIDAM,VSAM)
      DATASET DD1=DI21PAR4,DEVICE=3380
*
      SEGM  NAME=PARTROOT,PARENT=0,BYTES=50,FREQ=250
      FIELD NAME=(PARTKEY,SEQ),TYPE=C,BYTES=17,START=1
      FIELD NAME=PARTNO,TYPE=C,BYTES=15,START=3
      FIELD NAME=DESCRIPT,TYPE=C,BYTES=20,START=27
      LCHILD NAME=(PARTINDX,DI21PAX4),PTR=INDX
*
      SEGM  NAME=STOKSTAT,PARENT=PARTROOT,BYTES=160,FREQ=2
      FIELD NAME=(STOCKEY,SEQ),TYPE=C,BYTES=16,START=1
      FIELD NAME=SSAREA,TYPE=C,BYTES=1,START=3
      FIELD NAME=SSDEPT,TYPE=C,BYTES=2,START=4
      FIELD NAME=SSPROJ,TYPE=C,BYTES=3,START=6
      FIELD NAME=SSDIV,TYPE=C,BYTES=2,START=9
      FIELD NAME=SSUNITPR,TYPE=C,BYTES=9,START=21
      FIELD NAME=SSUNITMS,TYPE=C,BYTES=4,START=35
      FIELD NAME=SSATTRIT,TYPE=C,BYTES=7,START=51
      FIELD NAME=COAP,TYPE=C,BYTES=3,START=51
      FIELD NAME=PLAN,TYPE=C,BYTES=3,START=54
      FIELD NAME=COAD,TYPE=C,BYTES=1,START=57
      FIELD NAME=SSSTCKDT,TYPE=C,BYTES=3,START=72
      FIELD NAME=LASTTRDT,TYPE=C,BYTES=3,START=75
      FIELD NAME=SSCURRRQ,TYPE=C,BYTES=8,START=90
```



DI21PART Sample Database – unload data from individual segments



IMS – Relational Table Mapping



Agenda

- Introduction to relational mapping of non-relational data
- Federation: Transitioning the applications
- Data Event Publishing: Data Event-based Syncing

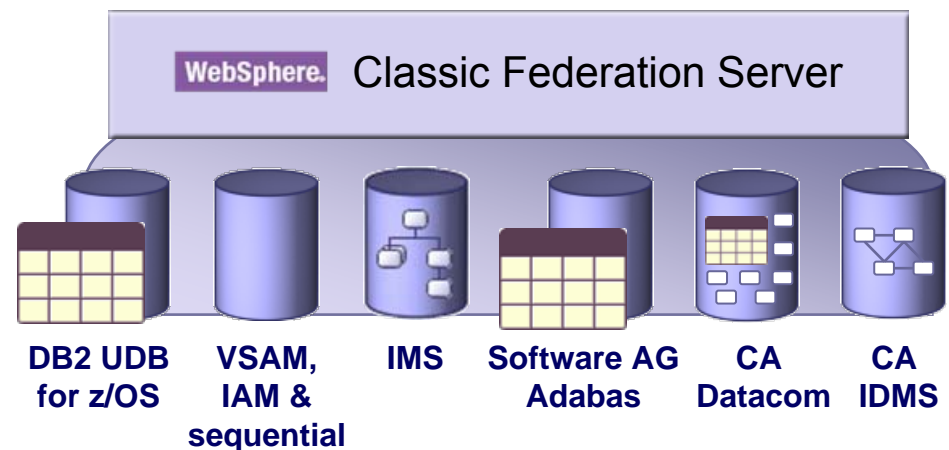


WebSphere II Classic Federation for z/OS

The Basics

Read-from and write-to mainframe data sources using SQL from Unix, Windows and JVM platforms

- **Integrates via standard ODBC/JDBC SQL interfaces**
- **Metadata-driven means:**
 - No mainframe programming required
 - Fast installation & configuration
 - Ease of maintenance
- **Works with existing and new:**
 - Mainframe infrastructure
 - Application infrastructure
 - Toolsets



Agenda

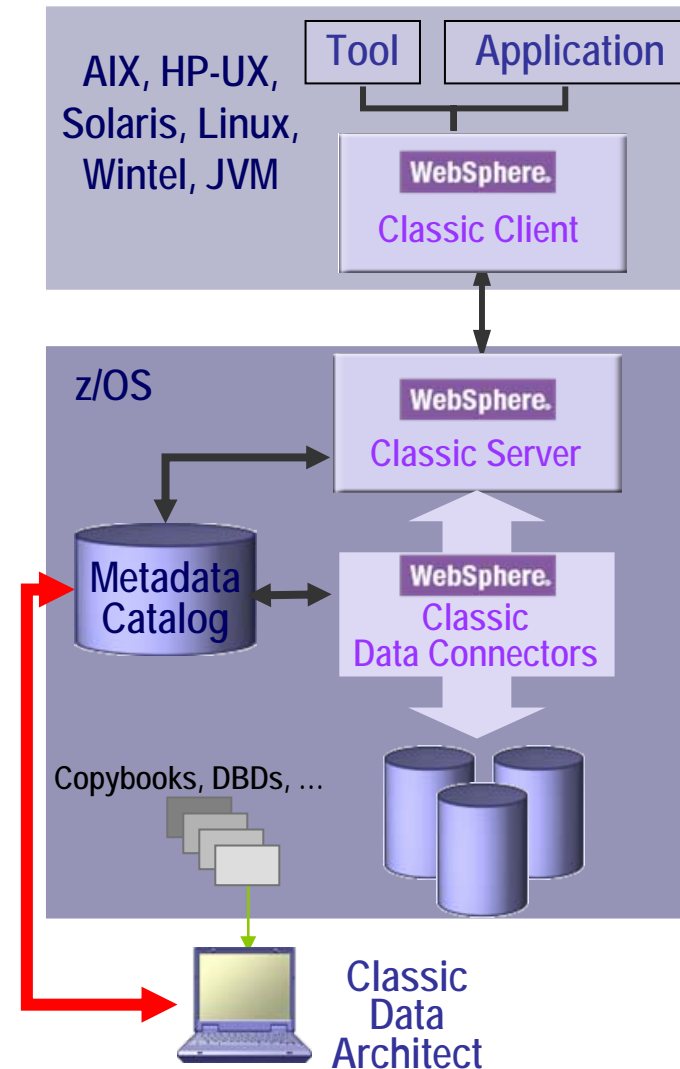
- Introduction to the problem

Federation: Transitioning the applications

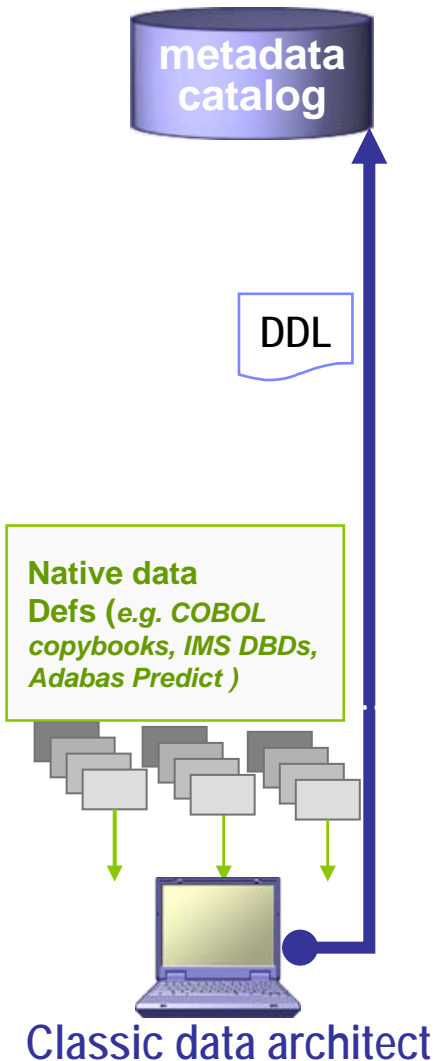
- Data Event Publishing: Data Event-based Syncing

WebSphere II Classic Federation Implementation

- Mainframe SMP/E install
- Configure Mainframe Server and components to act as a relational engine
- Create relational description of mainframe data sources by mapping the physical data definitions to logical tables and views
- Install/ configure JDBC and/or ODBC drivers provide standardized interface for tools and applications

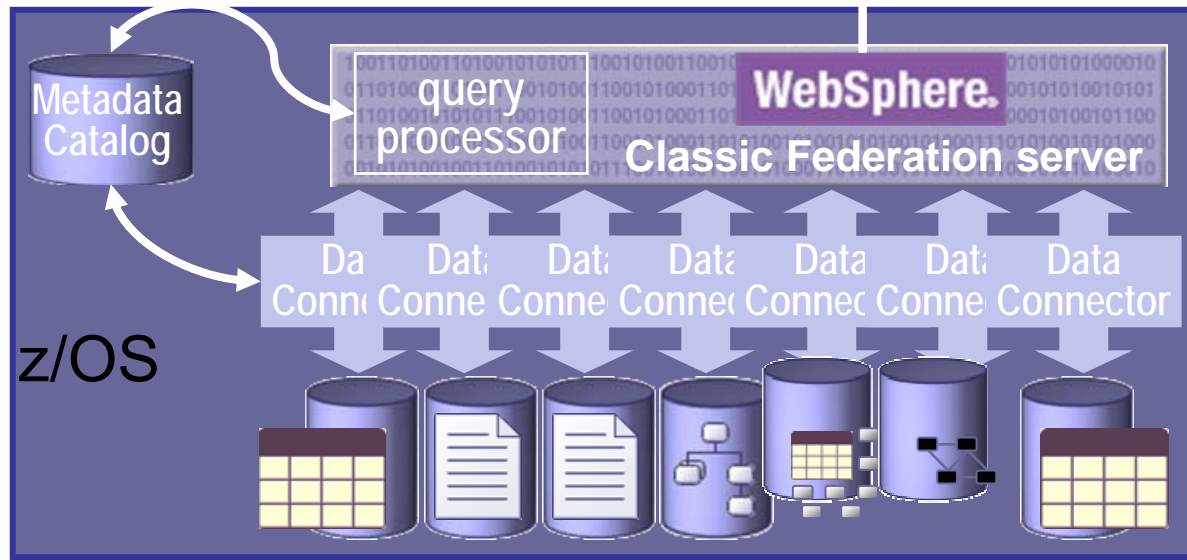
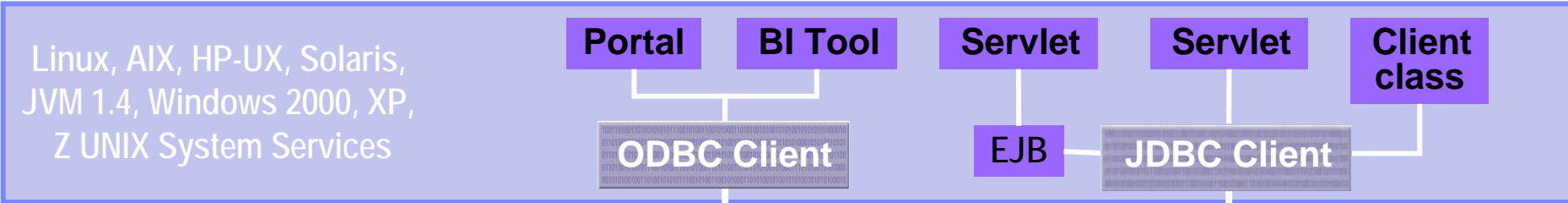


Metadata Management



- Metadata defines business-oriented relational mappings
 - Allows translation of SQL (relational) requests to native calls (e.g. DL/I)
 - Import existing copybooks, IDMS schemas, IMS DBDs, etc.
 - Generate logical relational *reference* table definitions
- Simulated RDBMS catalog and more
 - RDBMS-like catalog support: systables, syscolumns, etc.
 - Query-able tables for non-relational metadata
- Some metadata-driven features
 - Automatic translation of legacy data types
 - Handles legacy constructs like recurring data and redefines
 - Complex tables can span segments, records, etc.
 - Metadata-driven filtering using WHERE clauses
 - Enhances security via schema mapping, views, & DB2-like security
- Metadata Utility (Classic Data Architect)
 - Metadata customization and visual administration
 - Create and update metadata catalog entries
 - Verify metadata against physical (e.g. IMS DBD checks)

Components Overview



- Server address space
- Communication interface (i.e. TCPIP)
- Query Processor
- Database Connectors
- Clients: JDBC, ODBC, CLI

Two broad categories of usage

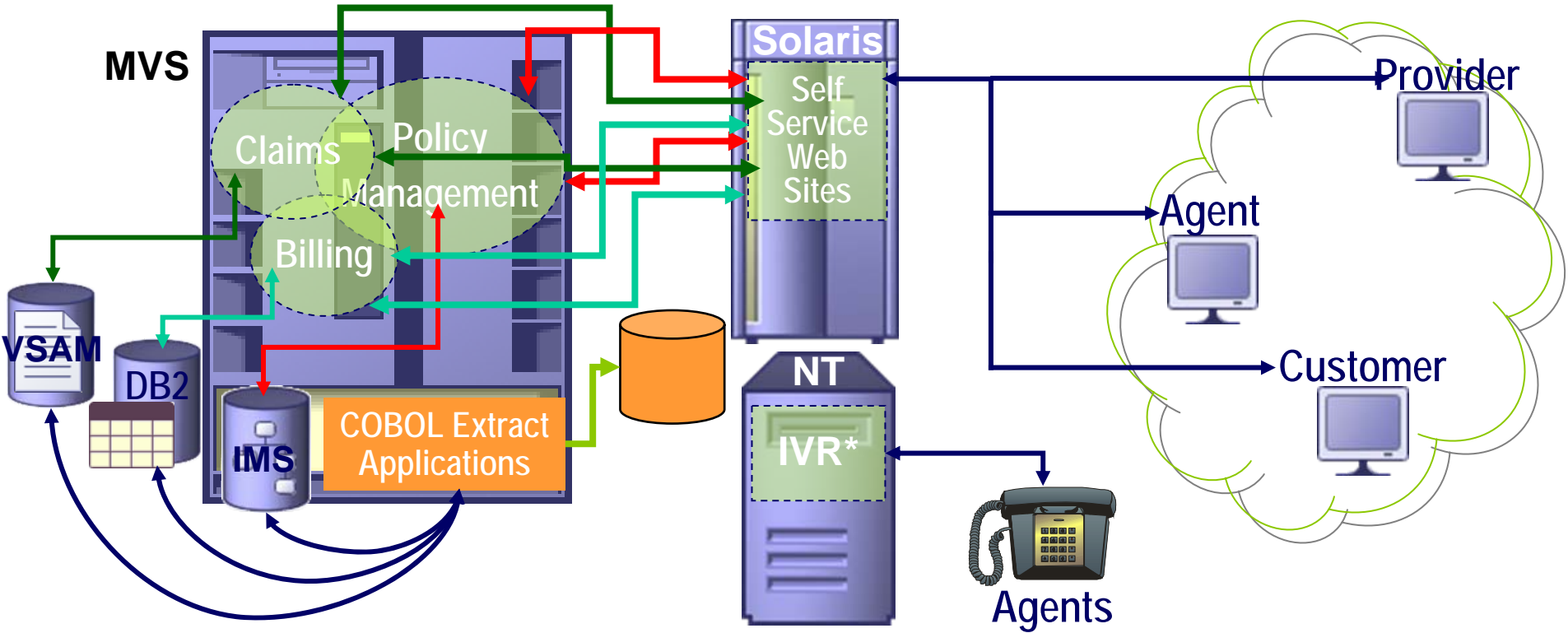
- e-Business
 - Deliver mainframe data to
 - Self-service portals (real-time account details)
 - e-commerce solutions (real-time inventory)
 - Employee portals (real-time claims detail)
 - Web developers become productive with no mainframe skills
 - Eliminates data latency business issues caused by copied data

- Business intelligence
 - Integrates seamlessly with
 - Reporting and analytical tools, e.g. Business Objects
 - Portals, e.g. WebSphere Portal
 - ETL, e.g. WebSphere DataStage

Self-service Application for Insurance Carrier

The Pain associated with “traditional” implementations

- Option a: copy data to non-mainframe environments
 - Estimated cost \$2M
 - Data refreshed every 30 hours or so
- Option b: integrate the IMS transactions
 - Estimated cost 10,000 man-hours per application

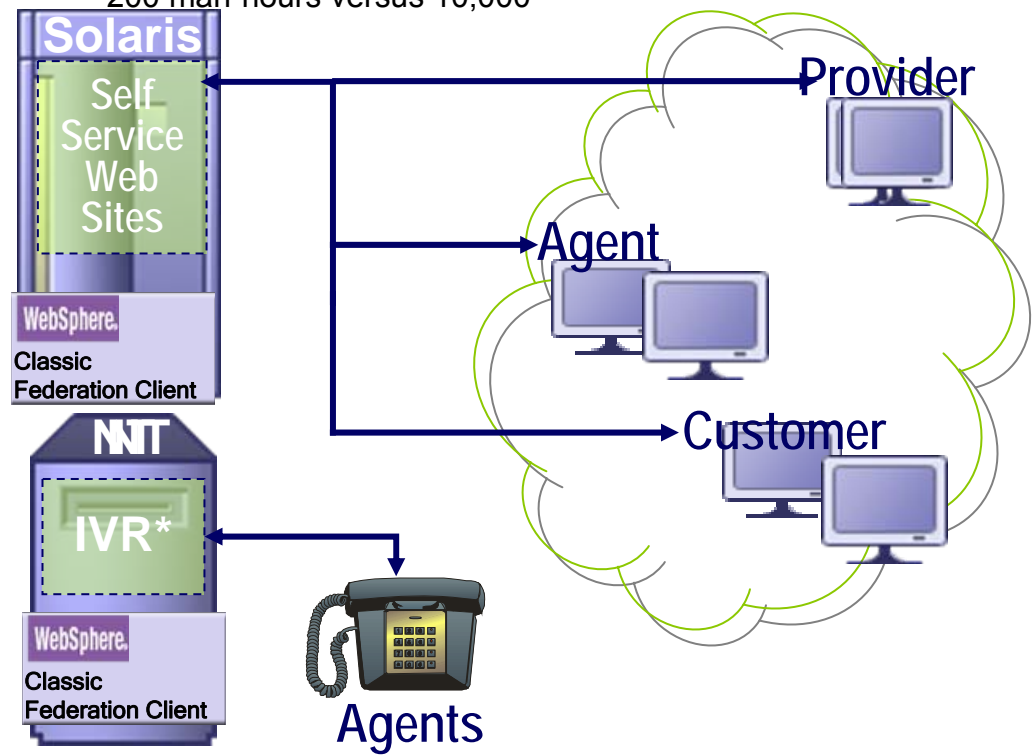
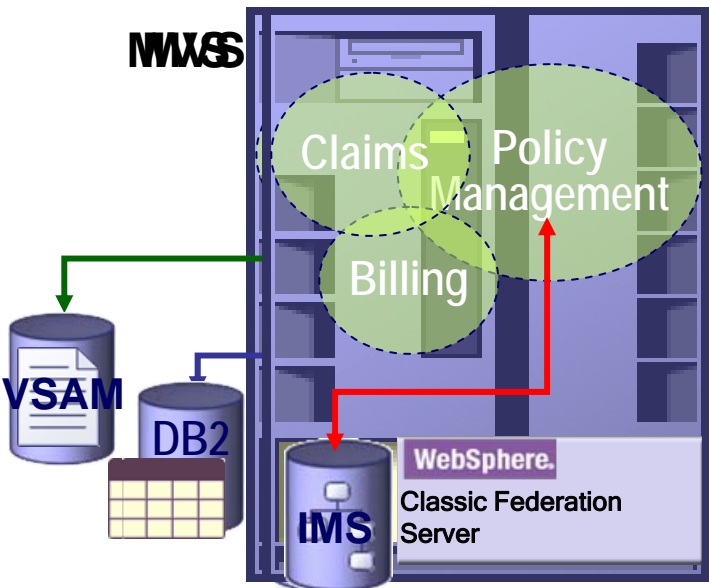


Self-service Application for Insurance Carrier

The IBM solution - *empower self-service environments*

Provide up-to-the-minute policy, claims and accounting information

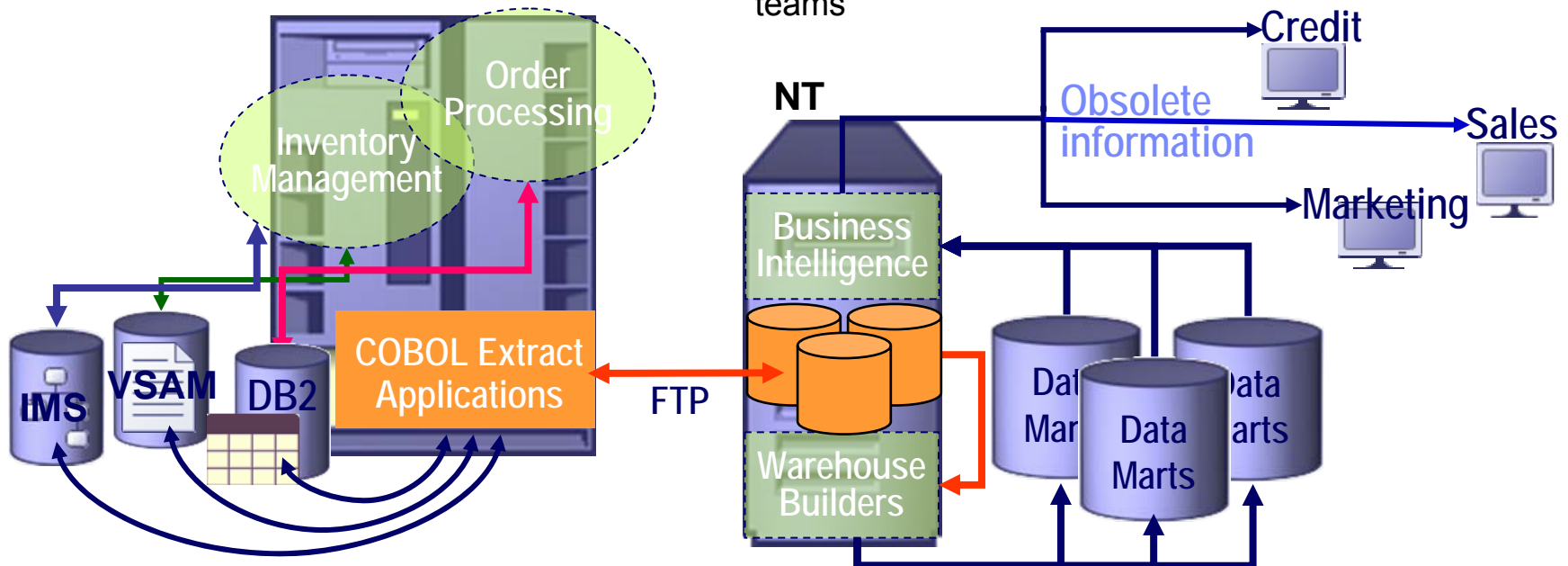
- Connect interactive voice response (IVR) system to IMS, VSAM & DB2
 - \$250K versus \$2M
- Connect operational data with self-service Web sites
 - 200 man-hours versus 10,000



BI solution for Kawasaki

The pain associated with “traditional” implementations

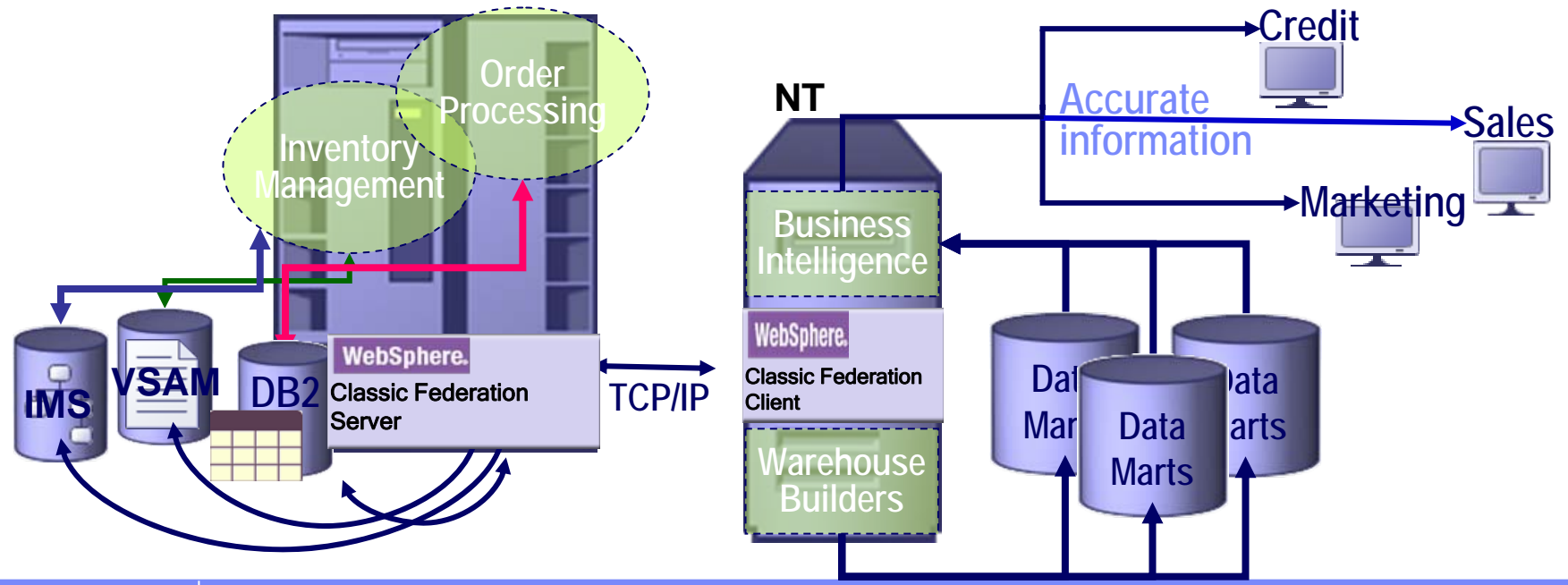
- Disjoint Process
 - Build and maintain mainframe “extract” process
 - Build and maintain distributed data transform & load
- Management challenges lead to increasing costs
 - Multiple skill sets required:
 - Mainframe programming & data warehouse design/build
 - Coordinating multiple components and development teams



BI solution for Kawasaki

IBM solution -- feed operational data to ETL via SQL

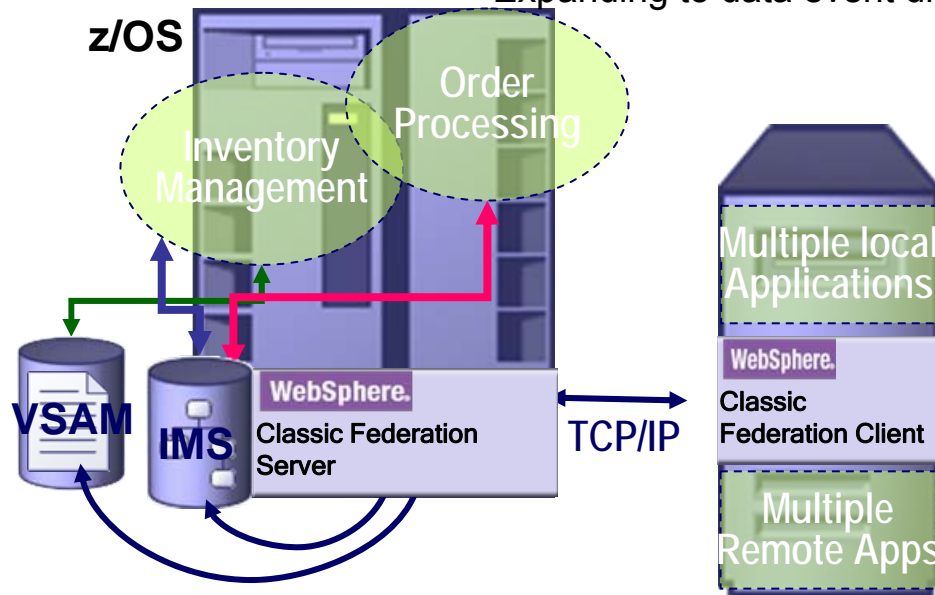
- Dynamically connect data warehouse tool with mainframe data
 - No dependence on mainframe development
 - Dramatically simplified management: One team owns it all
 - One consistent process leverages “power” of ETL tooling
 - Development time “cut in half”
 - Empowers additional uses
 - Dynamic query by business intelligence tools extends the warehouse



E-commerce & infrastructure at international tech distributor

IBM Solution - *Single-source mission critical data*

- Core 24 by 7 transaction processing systems leveraging IICF
 - \$30 billion in annual revenue
 - Up to 1,000 transactions (~ 3,000 DB accesses) per second
 - Operations in EMEA, Asia Pacific and North America
- Technology platform:
 - WebSphere on Solaris
 - IMS for Inventory and Orders, VSAM for Customer data
- Expanding to data event driven models!



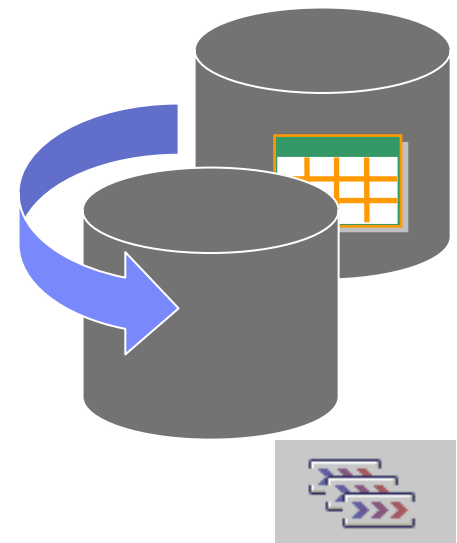
Agenda

- Introduction to the problem
- Federation: Transitioning the applications

Data Event Publishing: Data Event-based Syncing

Why Publish Data?

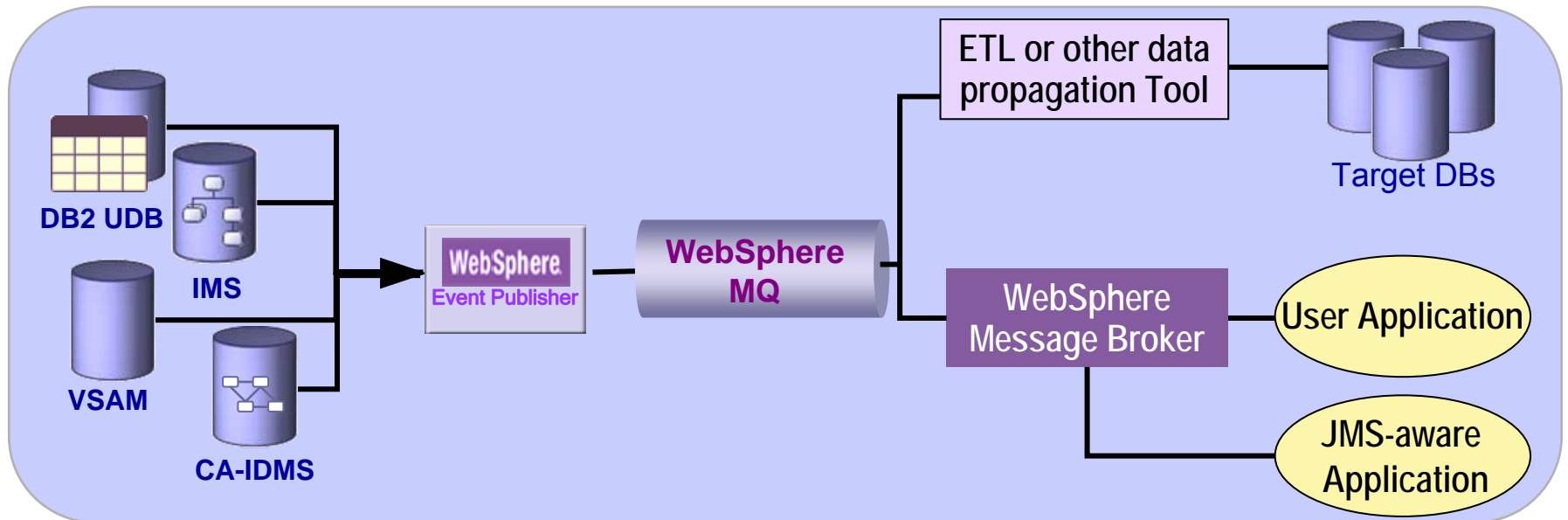
- **Application to Application Messaging**
 - Drive downstream applications or APIs based on the transactional changed data of database events
 - Reduce application development and maintenance, performance impact to source applications, and availability impact to source applications
- **Meet Auditing Requirements**
 - Capture and store information regarding what changes were made to critical business data and by whom
- **Event Notification**
 - Stream changed data information to Web interfaces
 - Stream only particular events of interest (filter data)
- **Warehouse / Business Intelligence**
 - Integrate captured changed data with an ETL tool
 - Perform very complex transformations
 - Use a specific transaction format to update target



WebSphere Data Event Publishers

Facilitating integration

- Capture data events in real time
- Publish these "data events" to
 - Leverage operational data to drive business processes
 - Loosely coupled integration maintains application independence
 - One consistent published data format regardless of source

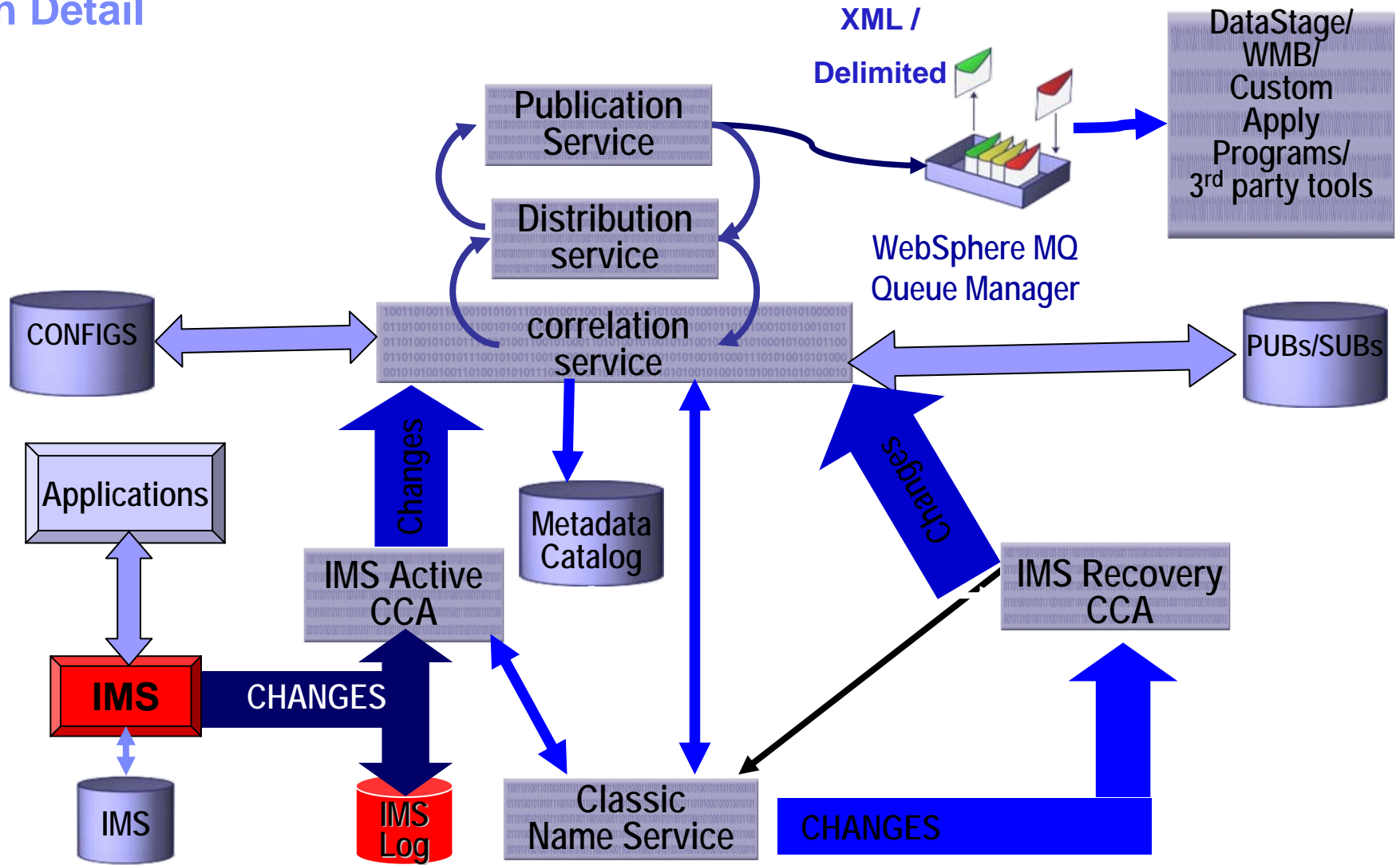


Supported IMS Environments

- IMS Online DBCTL and DB/DC
- IMS Batch (with logging enabled)
- Databases Supported
 - Full function (HDAM, HIDAM, etc.)
 - DEDB (region restrictions apply)
- IMS Releases Supported
 - Versions 7, 8, and 9

Classic Data Event Publisher 9.1

In Detail



Categories of usage

- Application to application messaging
 - Deliver mainframe data event triggers to
 - PoS applications
 - Integration Infrastructure (e.g. Inventory applications)
 - Employee portals (real-time claims detail)
 - Java/ .Net developers become productive with no mainframe skills
 - Eliminates data latency business issues caused by copied data

- Business intelligence – synchronized DataWareHousing
 - Integrates seamlessly with
 - Reporting and analytical tools, e.g. Business Objects
 - Portals, e.g. WebSphere Portal
 - ETL, e.g. WebSphere DataStage

Challenges at leading Spanish retailer

€15.8 Billion revenue, 600+ stores, 80,000 employees, selling 35,000+ products

Business Challenges

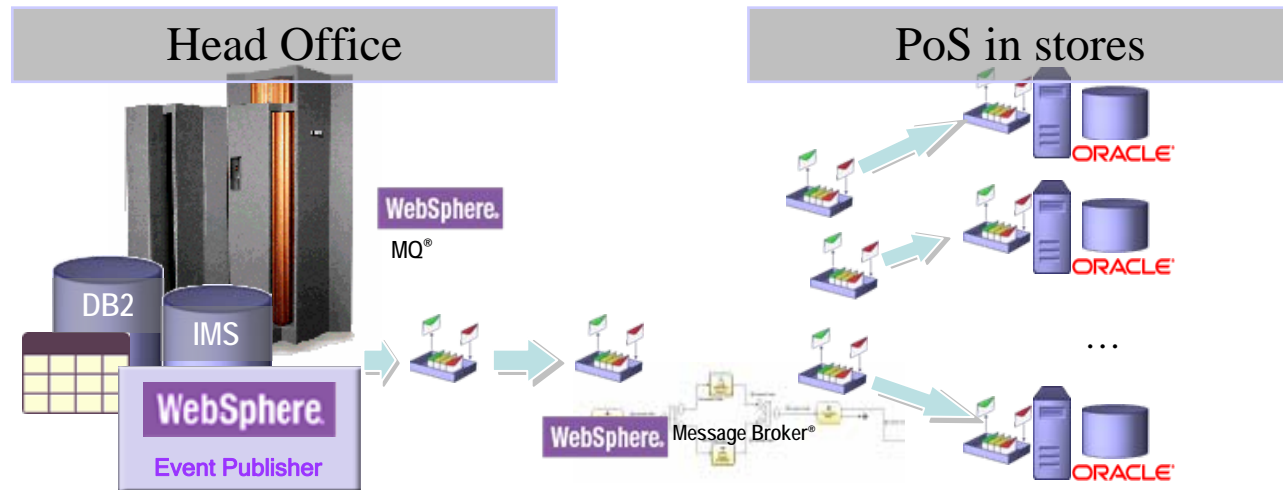
- ▶ How to distribute information around the organization?
 - Price files from centralized application to PoS terminals
 - Real-time environment for business reporting
- ▶ How to handle increasing volumes in the same time window?
- ▶ How to move to a real-time environment?

Technical Challenges

- ▶ Data transformation requirements at both head office and stores
- ▶ Guaranteed delivery of messages
- ▶ Network traffic
- ▶ Support of legacy hardware in stores

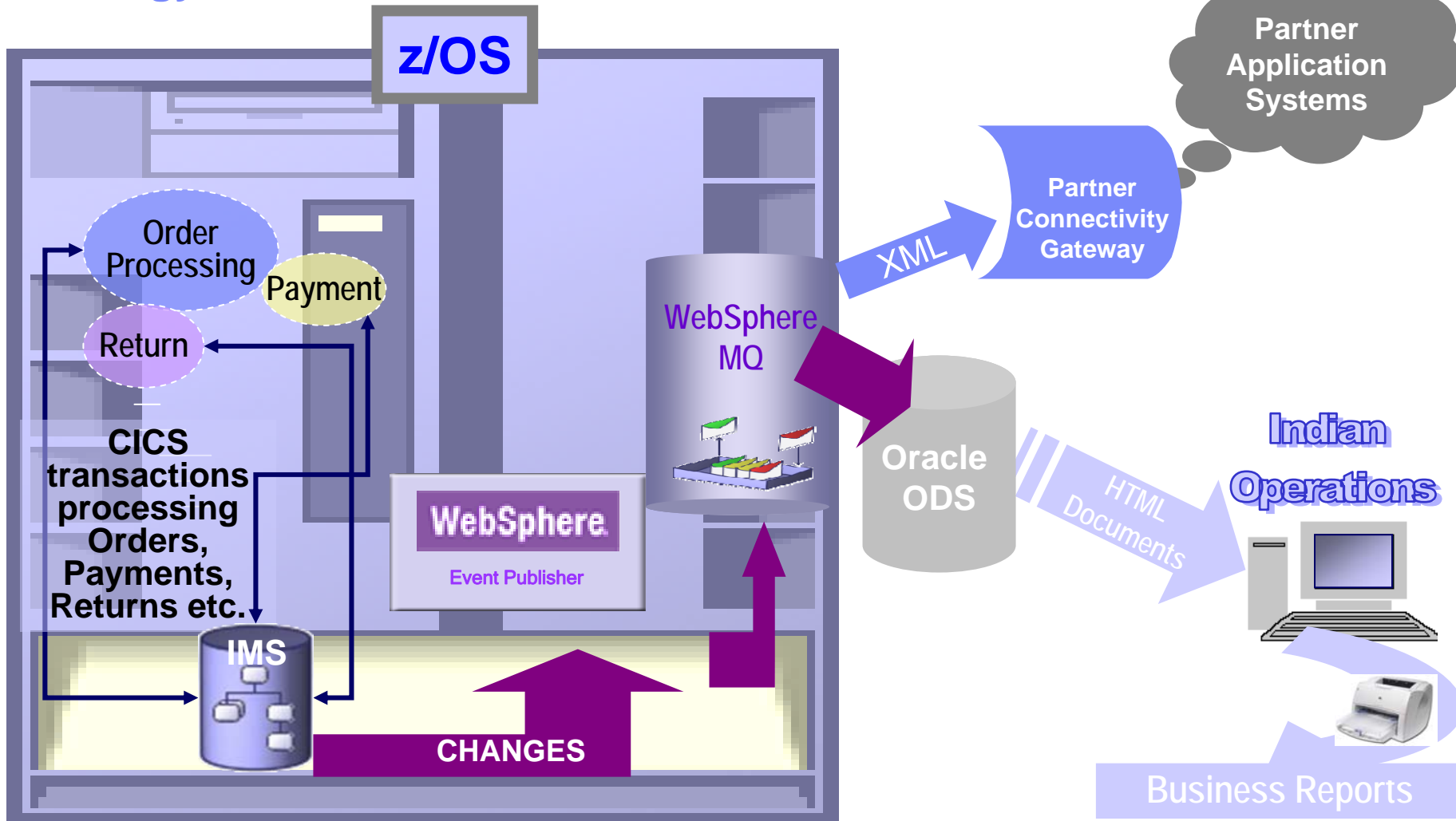


Solution



- ▶ Distribute information which is centralized at the head office around the organization, in a cost-effective manner
 - Minimize mainframe processing overhead as EP reads from logs
 - Easy integration with Message Broker via MQ
 - Multicast capabilities at stores
 - Capability to add DSEE
 - Time window to load data warehouse
 - More advanced data transformation capabilities

Event Monitoring using Classic Event Publisher at large technology distributor



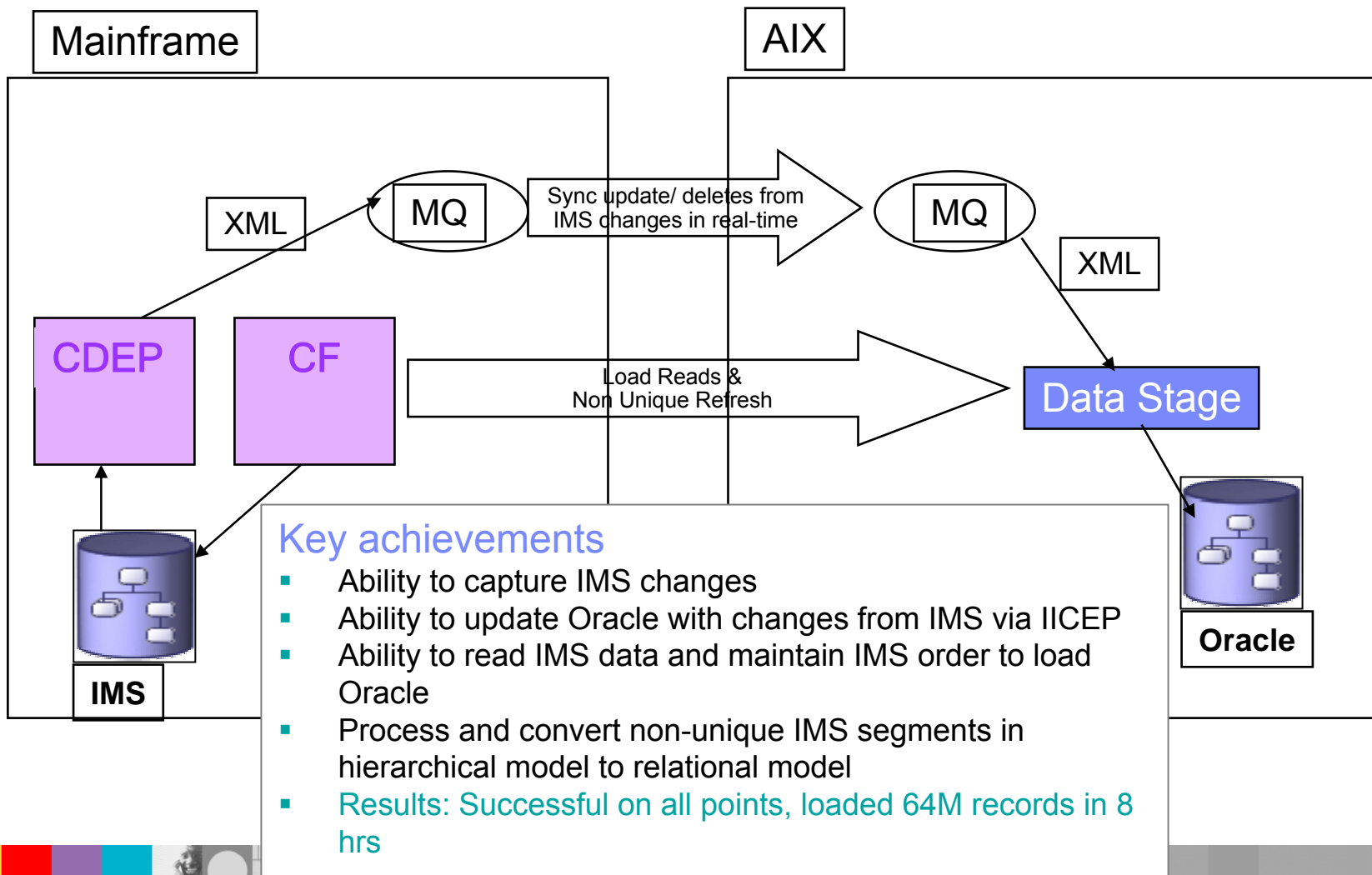
Event Monitoring using Classic Event Publisher at large technology distributor

- ✓ Currently capturing changes from US, AP, CA, EMEA and LA
- ✓ Volume is about 4.5 million transactions per 24-hr period
- ✓ Translates to about 13.5 million database changes per 24-hr period
- ✓ Changes are populating an ODS on Oracle to provide data for operational reports
- ✓ Enables the Partner Connectivity Gateway (PCG) so that partner companies can integrate their applications with Ingram's inventory system



This is the problem we started with!

DWH load and sync for US State Govt. Judicial Information Systems



धन्यवाद

Hindi

多謝

Traditional Chinese

ขอบพระคุณ

Thai

Спасибо

Russian

Gracias

Spanish

ありがとうございました

Japanese

Obrigado

Brazilian Portuguese

شكراً

Arabic

Thank You

English

Grazie

Italian

Merci

French

நன்றி

Tami Tamil

多谢

Simplified Chinese

Danke

German

감사합니다

Korean

Backup slides

Classic Data Event Publisher 9.1 Components

- **CCA (Change Capture Agents)**
- **Correlation Service**
- **Distribution Service**
- **Publication Service**



CCA

- Database Specific - involves changes to Database objects
- Exit based Agent - installed in the database load libraries
- Called when data is written to the IMS log
- Change data sent to Correlation Service using a cross memory data queue
- Minimal latency in busy systems
- May be delayed due to log file buffering (e.g. In testing a small number of changes with an online system, a /CHECKPOINT command may be required to cause data to flush)
- Recovery agent:
 - ▶ Reads one or more log files in chronological order
 - ▶ Uses combination of archive and active logs for online systems
 - ▶ Uses system log for IMS batch jobs (IEFRDER)
 - ▶ Uses combination of archive and active logs for online systems

Correlation Service

- Connects to the DS outbound (Asynchronous TCPIP is the preferred protocol, XM can be used if they are running on the same image. Puts up the CSA.
- Send table metadata and subscription information to the DS, manages control table data
- Gets changes from the CCAs
- Stages changes until COMMIT/ROLLBACK is received
- Sends COMMITTED changes to the DS
- Manages subscription states passing that information to the Distribution Service
- Admin/Recovery message processing
- Agent State maintenance in the CSA
- Opens a pipe to listen to Recovery Agents and forwards the RESTART token requests to the DS.

Distribution Service

- Listener for one or more Correlation Services
- Starts one or more PS threads
- Sequences changes
- Communicates Recovery Tokens to the CS
- Forwards changes to the Publication Service
- Gets error/info/admin messages back from the PS, sending it back up the chain to the CS

Publication Service

- Gets messages from XM Data Queues from the DS
- Loads the appropriate publish module
- Publish messages to MQ or Apply to VSAM
- Handle MQ Admin Q messages for Replication
- Publish Types Supported
 - ▶ XML Publish
 - ▶ CMF (Compact Message Format)
 - ▶ DELIMITED Format
- Manages restart information and sends error/restart/info messages to the DS



One Capture, One Server, Many Outputs

Reusable Parts

- Multiple Distribution Services
 - ▶ XML – for broad applicability
 - ▶ CDV– for custom-delimited-values
 - Simplifies DataStage integration
 - Can be used with other 3rd party tools
 - Informatica
 - Business Objects
 - ▶ Replication – for hook to Q-Rep

