



IBM STG Technical Conference

Data Warehousing on System z BWDB2UG Presentation 9/12/07

John Partridge, Sr. Software Engineer
BI on Z Swat Team

Agenda

- Understanding Warehousing Terminology
- Data Warehousing (DW) Market Directions
- Target Segments for DW on System z
- What exists today for DW on System z
- Solution Architecture for DW on System z
- What is coming for DW on System z
- Summary

Understanding the terminology

- Business Intelligence (BI) and Data Warehousing (DW) are sometimes used interchangeably
 - Typically BI includes **end user** tools for query, reporting, analysis, dashboarding etc.
 - Both concepts depend on each other
 - BI** almost always assumes a Warehouse (WH), Operational Data Store (ODS) or Data Mart (DM) exists with timely, trusted information
 - A **DW** depends on end user tools that turn data into information.
- Both terms (DW and BI) address desire for timely, accurate, available data delivered when, where and how the end users want it

What is Data Warehousing?

- **Data Warehousing** provides the underlying data storage facilities required to support any BI Solution.
- **An Operational Data Store (ODS)** is a subject oriented database organized by business area. It is up to date (vs. historical) and detailed (vs. summarized).
- **A Data Warehouse (DW or sometimes WH)** is a multi-subject oriented database populated from operational systems and/or ODS's. It is historical (vs. point-in-time) in nature and typically contains detailed data. It is often looked upon as the single source of corporate "truth".
- **A Data Mart (DM)** is a database designed to support the analysis of a particular business subject area. Data has usually been transformed and aggregated from the source DW or operational system. Data Marts can be relational, multidimensional (OLAP) or statistical in nature.

Why Talk about Data Warehousing on z/OS in 2007

- **System z customers are asking for increasing support and capabilities for DWH on z**
 - Protects their investment in System z
 - Keep up or lead in BI and DWH technologies
 - Significant amount of enterprise data is on System z
- **Many z customers have a DWH on DB2 z/OS**
- **System z can capture DW consolidation and modernization opportunities**
- **Demand for embedded analytics in operational applications is growing**
- **The BI operational needs for reliability, availability and security are growing**
 - The z/OS platform is a secured system second to none.
- **Significant DW capabilities have been added to DB2 V8 with even more DB2 9**
- **The new zIIP processor provides TCO and IT optimization**
 - Particularly useful for BI queries (DRDA access and parallel query processing)
- **Newly acquired data integration and analytic components combined with partner capabilities make best of breed integrated solution.**

Market Directions

Dynamic Warehousing

A New Approach to Leveraging Information

Information On Demand to
Optimize Real-Time Processes



*Dynamic
Warehousing*

OLAP & Data Mining to
Understand Why and
Recommend Future Action



*Traditional Data
Warehousing*

Query & Reporting
to Understand
What Happened



Warehousing strategic pillars

Guiding principles for innovation

Simplicity

- Easy to deploy and integrate
- Easy to use
- Easy to manage
- Easy to start and grow as needed

Reliability & Performance

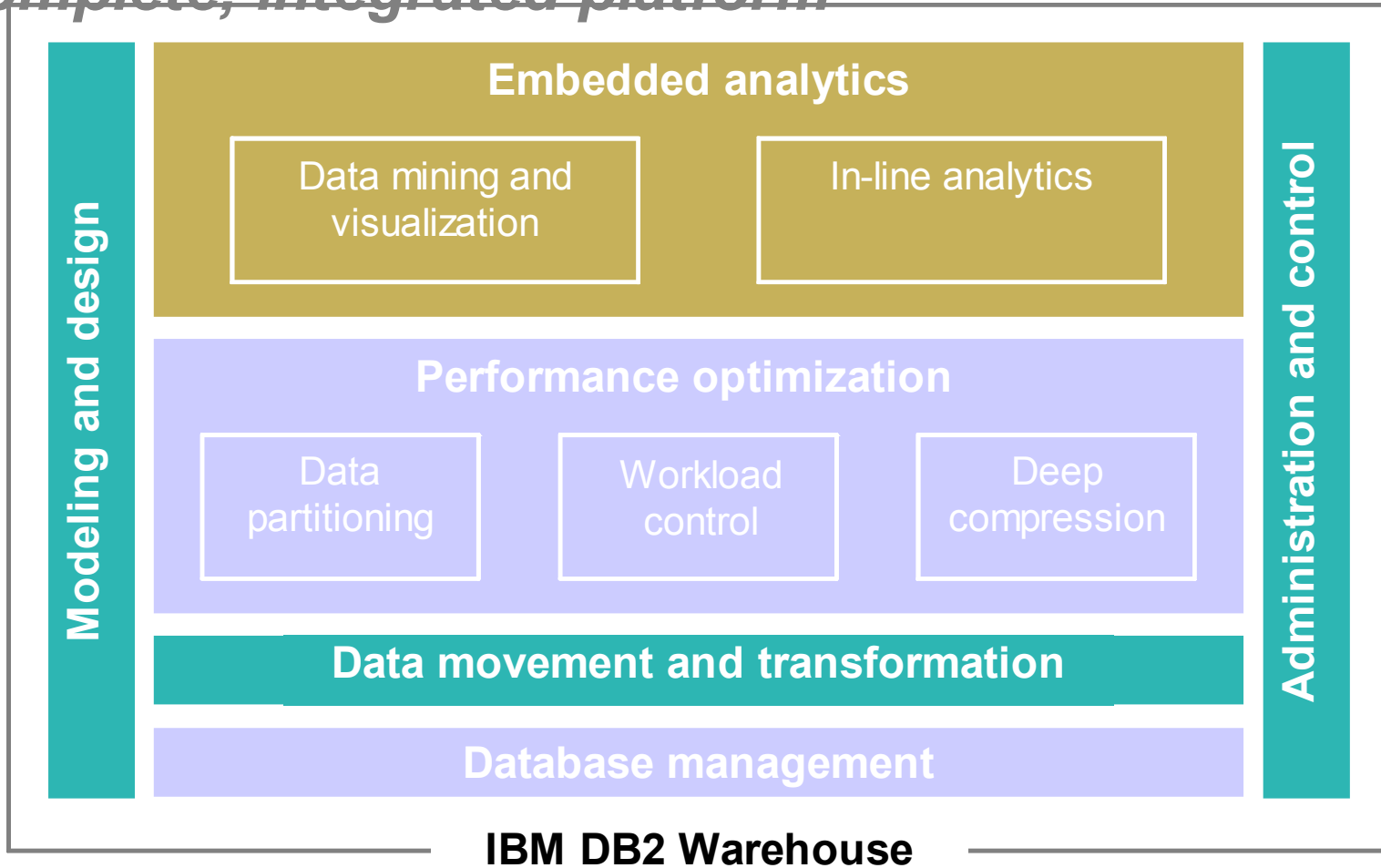
- Reliable access to information
- Highly available
- Real-time performance
- Maximized resource efficiency

Extended Insight

- Beyond traditional capabilities
- Further leverage information
- Extended business insight
- Support broader usage

IBM DB2 Warehouse software

A complete, integrated platform



Gartner perspective on BI and DWH market

System z holds a viable place in the Leaders quadrant

The emergence of issues based on a mix of 4 DW workloads

- Continuous (near-real-time) data loading – similar to an OLTP workload
- Large numbers of standard reports
- An increasing number of true ad hoc query users
- An increasing level of analytics and BI-oriented functionality in OLTP

The mixed workload performance will become the single most important performance issue in DW

The transactional DBMSs have an edge that challenges the DW DBMSs (such as Teradata) to stay ahead

Target Segments for DW on System z

Candidate Environments for DW on System z

- **Focus is on existing System z customers with DB2 z/OS V8**
 - Existing DB2 z/OS skills
 - Established DB2 z/OS commitment
- **Plus, customers with existing ODS and/or DW on System z**
 - Modernization of Reporting & Analytics via DataQuant &/or AlphaBlox
 - Modernization of ETL and ELT via “ETL Accelerator” (periodic batch update to “active EDW”)
- **Plus, customers with perhaps some DM’s (on or off z) but no EDW implemented yet**
 - Enhanced Reporting & Analytics via DataQuant &/or AlphaBlox
 - “Active EDW” environment via ETL Accelerator

Candidate Environments (continued)

- **Plus, customers looking at Embedded Analytics**
 - Deliver BI to customer facing humans and applications
 - Integrate BI Components with Operational Systems & Information Portals
- **Embedded Analytics is a Sweet Spot for DW on System z**
 - Requires massive transaction scalability (z strength)
 - Requires semi-aggregated data, i.e. data aggregated at a low level (z strength)
 - System z provides the OLTP Systems & DW
 - AlphaBlox provides the BI Components

What Exists Today for Data Warehousing on System z

What DW Capabilities Exist Today for System z

▪ **DB2 z/OS V8 & V9 DW Base**

- Functional and performance enhancements
- Easier online reporting and data management capabilities
- DB2 family compatibility

▪ **DW ETL**

- WS DataStage EE, WS II Classic Federation, WS Classic Event Publishers, Distributed DBMS Event Publishers, BatchPipes for OS/390, DB2 Unload/Load Utilities

▪ **Analyze/Report**

- From IBM: Alphablox, QMF--and DataQuant
- From partners: Cognos, Hyperion, Business Objects, SAS, IBI

▪ **Performance Management**

- IBM Tivoli Omegamon XE for DB2 Performance Expert on z/OS, DB2 Query Monitor, Optimization Service Center (V9)

▪ **Other IBM Information Server Components**

- WS Information Analyzer, WS QualityStage

Industry data models

Leverage industry best practices for faster time to market

Over 400 Customers!

Extended Insight

New Offering!

Banking

(Banking Data Warehouse)

- Profitability
- Relationship marketing
- Risk management
- Asset and liability management
- Compliance

Financial Markets

(Financial Markets Data Warehouse)

- Risk management
- Asset and liability management
- Compliance

Health Plan

(Health Plan Data Warehouse)

- Claims
- Medical management
- Provider and network
- Sales, marketing and membership
- Financials

Insurance

(Insurance Information Warehouse)

- Customer centricity
- Claims
- Intermediary performance
- Compliance
- Risk management

Enhanced Capabilities!

Retail

(Retail Data Warehouse)

- Customer centricity
- Merchandising management
- Store operations and product management
- Supply chain management
- Compliance

Telco

(Telecommunications Data Warehouse)

- Churn management
- Relationship management and segmentation
- Sales and marketing
- Service quality and product lifecycle
- Usage profile

DB2 V8 enhancements benefiting DW

- **Usability, Availability & Scalability**
 - Online Partitioning Changes, More Partitions
 - Schema Evolution
 - VSCR with System z 64-bit Architecture
- **Data Warehousing Support**
 - Star Join Improvements
 - Materialized Query Tables
- **Overall System & Query Performance**
 - Locking Improvements
 - Multi-row INSERT & FETCH
 - Index only Access for VARCHAR
 - DDF Performance Improvements

DB2 V9 enhancements benefiting DW

- **Usability, Availability & Scalability**
 - Partition by Growth Tablespace
 - Clone Tables (Online Load Replace)
- **Data Warehousing Support**
 - Star Join Improvements (Dynamic Index ANDing)
 - SQL Enhancements (RANK, DENSE_RANK)
- **Overall System & Query Performance**
 - TRUNCATE Statement for fast delete
 - Index Compression with minimum overhead
 - Significant CPU Reduction for DB2 Utilities
 - DDF Performance Improvements

IBM zIIP optimizes DW workloads

1. Business Intelligence applications via DRDA[®] over a TCP/IP connection



<http://www.ibm.com/systems/z/ziip>

What is DataQuant

- DataQuant provides a comprehensive query, reporting and data visualization platform for both web and workstation-based environments.
- DataQuant introduces a wide variety of powerful business intelligence capabilities, from executive dashboards and interactive visual applications to information-rich graphical reports and ad-hoc querying and data analysis
- DataQuant provides 2 components
 - DataQuant for Workstation – An Eclipse based environment for the development of query, report and dashboard solutions
 - DataQuant for WebSphere – A runtime environment capable of displaying DataQuant content using a “thin client” model

DataQuant Features

- **Visual Dashboards**
 - Interactive dashboards and database visualization applications
 - Offers dozens of interactive charts, layouts, controls and symbols
- **Enhanced graphical reporting**
 - Rapidly produce and disseminate rich graphical BI reports
- **Security & Personalization**
 - Comprehensive LDAP or ad-hoc security model
 - Ability to define role-specific BI workspaces on a user/group basis
- **Enhanced Analytics**
 - Adds over 100 analytical functions to queries, reports and visual solutions
- **Backward-compatible with existing QMF infrastructure**

What is Alphablox?

- Platform for Customized Analytic Applications and Inline Analytics
- Pre-built components (Blox) for analytic functionality
- Allows you to create **customized** analytic components that are **embedded** into existing business processes and web applications

The image displays two screenshots illustrating the use of Alphablox. The top screenshot shows the IBM WebSphere Portal's Category Manager interface. It features a navigation menu with options like 'Home', 'Scorecard', 'Category', 'Promotions', 'Vendor', and 'My Human Resources'. The main content area is divided into several sections:

- Category KPI's:** A table showing performance metrics for 200426.

Category	Actual	Target
Net Sales	\$13.71M	\$13.40M
Billing Gross	\$6.54M	\$6.02M
Markdowns % Sales	34.09%	29.64%
AGP	\$4.70M	\$4.41M
- Markdowns % Sales Variance by Division:** A map of the United States with color-coded regions (Alaska, Seattle, Vancouver, Winnipeg, Denver, Portland, Nor. California, Hawaii, So. California (Vons), Phoenix, Texas, Casa Ley) representing different divisions. A summary below the map shows: 'Portland_Division: Variance %: -7.80494933 Actual: 25.154933378 - Planned: 32.95988312 -'.
- People Finder:** A search interface with fields for 'Name' and 'Search for:'.
- Alerts:** A list of alerts including 'Kilgus Special k', 'GM Rice_CHX', 'Safeway PButter Cereal', 'GM Cheerios', and 'Kilgus Fruit Harvest'.

The bottom screenshot shows the Huntington National Bank website. It features a navigation menu with 'Home', 'Personal', 'Business', 'Planning & Tools', and 'Customer Service'. The main content area includes:

- Goal Tracking:** A horizontal bar chart showing progress from 0 to 7000. The chart has two bars, one labeled 'Actual' and one labeled 'Goal', with the 'Actual' bar reaching approximately 6000.
- Online Banking:** A login section with fields for 'Username' and 'Password', and buttons for 'Log In', 'Forgot Password?', 'Sign Up', and 'Learn More'.
- Locator:** A section for finding a branch, with fields for 'Select Location Type', 'Zip Code', and 'Advanced Locator'.
- How can we help you today?:** A section with 'Help me' and 'I want to' buttons.
- News & Information:** A list of links including 'Report a fraudulent email', 'Learn more about Identity Theft', 'Taxidermy - what you really need to know', and 'Tell us what you think about the new Huntington.com'. Take our survey.
- TurboTax for the Web:** A section promoting TurboTax software.
- The Simple Truth:** A section with the text: 'Our approach to banking is quite simple. We want to work with you to tell us solutions that will help you manage and grow your finances. Learn More.'

At the bottom of the Huntington website, there is a footer with links for 'Identity Protection', 'Security', 'Privacy', 'Service Guarantee', and 'Notices'. It also includes the text: 'Member FDIC. Equal Housing Lender. Huntington is a federally registered service mark of Huntington Bancshares Incorporated. © 2005 Huntington Bancshares Incorporated'.

Alphablox Components

- Blox components, which are modular, reusable components
- An application framework
- A powerful programming model
- ... and a variety of development tools for assembling analytic applications

- Run-time environment via standard J2EE application servers
- Can be installed on leading commercial J2EE application servers such as IBM WebSphere® software and others.
- It enhances the end-user analytics functionality, including support for:
 - full DHTML drag-and-drop
 - Histogram charts
 - 80/20 analysis
 - textual traffic lights
 - ... and much more

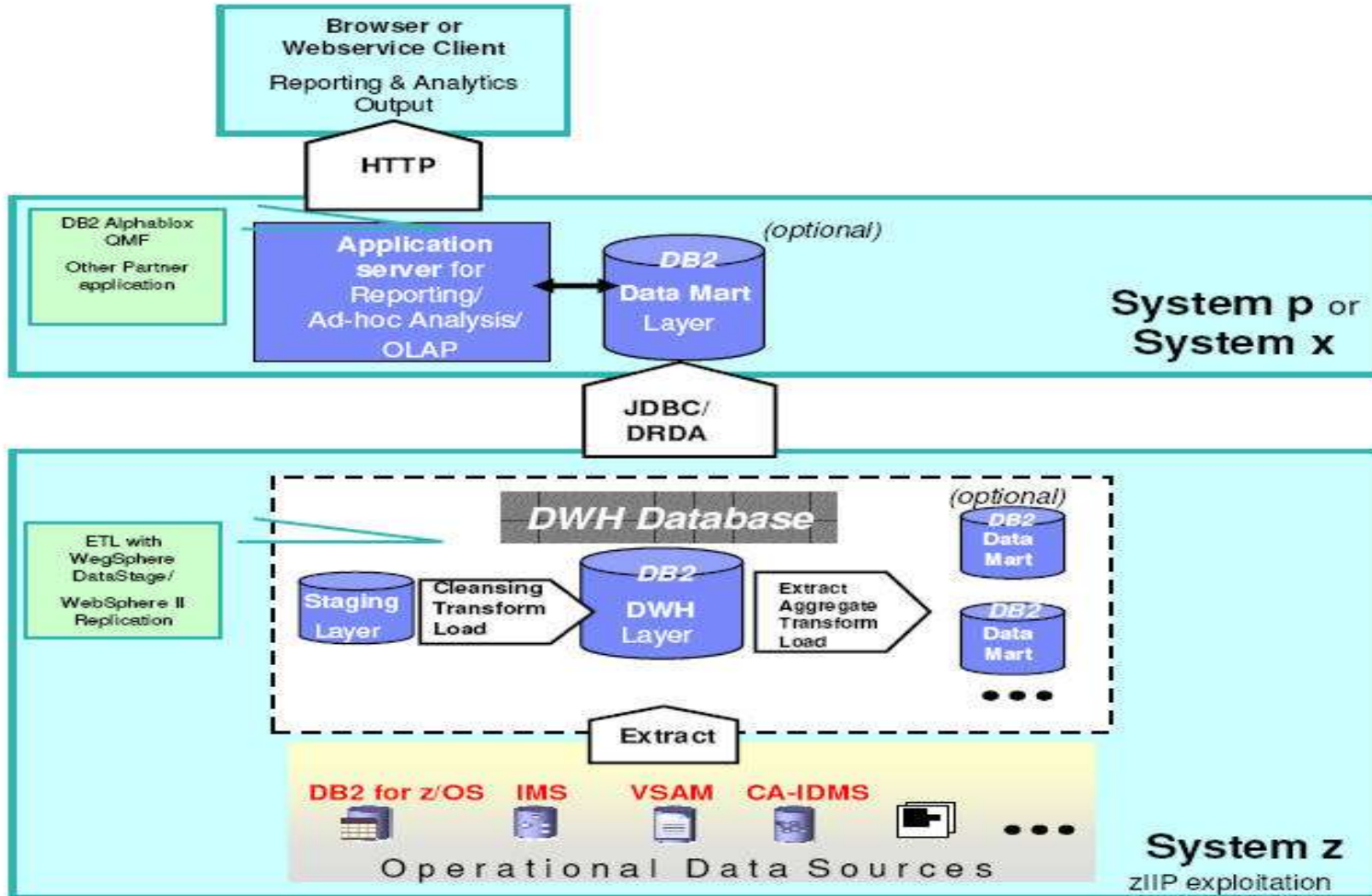
- It can help you and partners maximize the value of information assets by delivering contextual business insight to the right people at the point of decision.

Alternative Architectures (Reporting & Analytics)

- **“Pure” System z BI Solution from a Data Perspective**
 - ODS, Data Warehouse & Datamart(s) in DB2 z/OS
 - End User Tools (e.g. QMF, DataQuant, Business Objects, Cognos) access DB2 z/OS directly (fat client implementation) or via browser (web server implementation)
 - Reporting solution may run on distributed WAS, e.g. Alphablox, QMF, DataQuant, Cognos ReportNet, Business Objects Server

- **“Hybrid” BI Solution from a Data Perspective**
 - ODS & Data Warehouse in DB2 z/OS
 - Relational, Multidimensional (OLAP) and Statistical Datamarts on pSeries and/or xSeries supporting End User Tools, e.g. DB2 DWE, Hyperion Essbase, Cognos PowerPlay

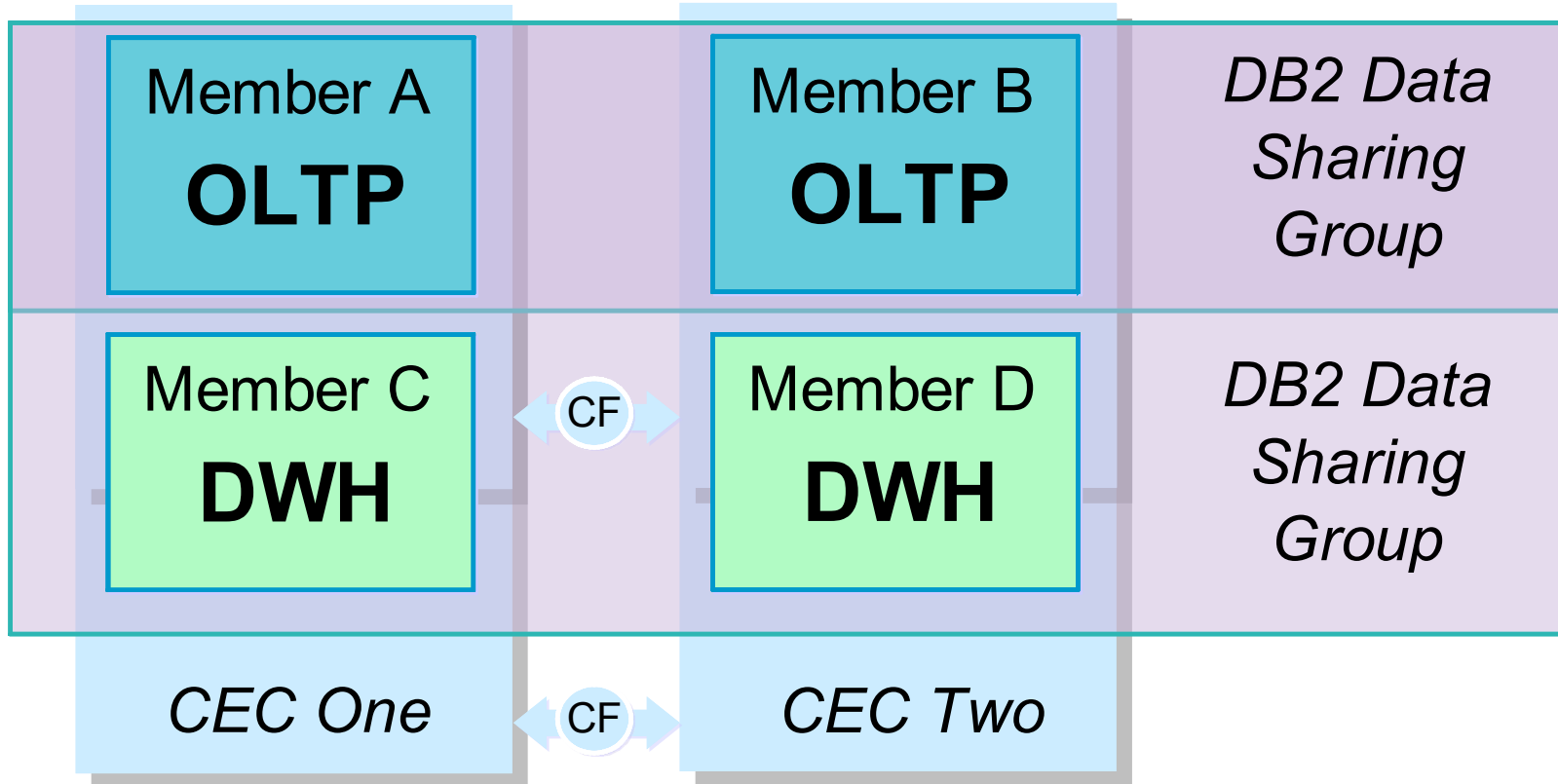
DW on System z – Hybrid Architecture



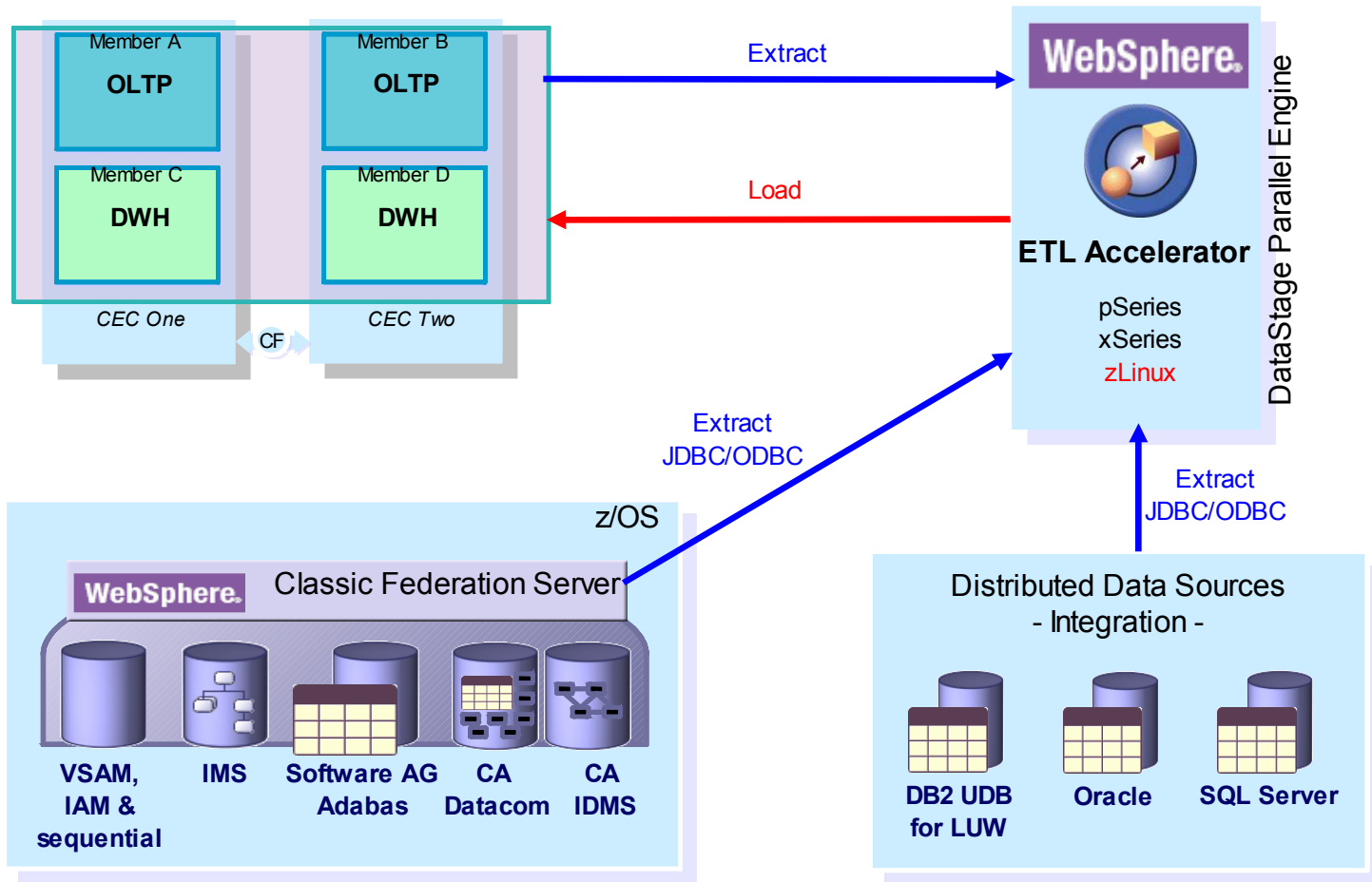
Solution Architecture for DW on System z

Solution Architecture

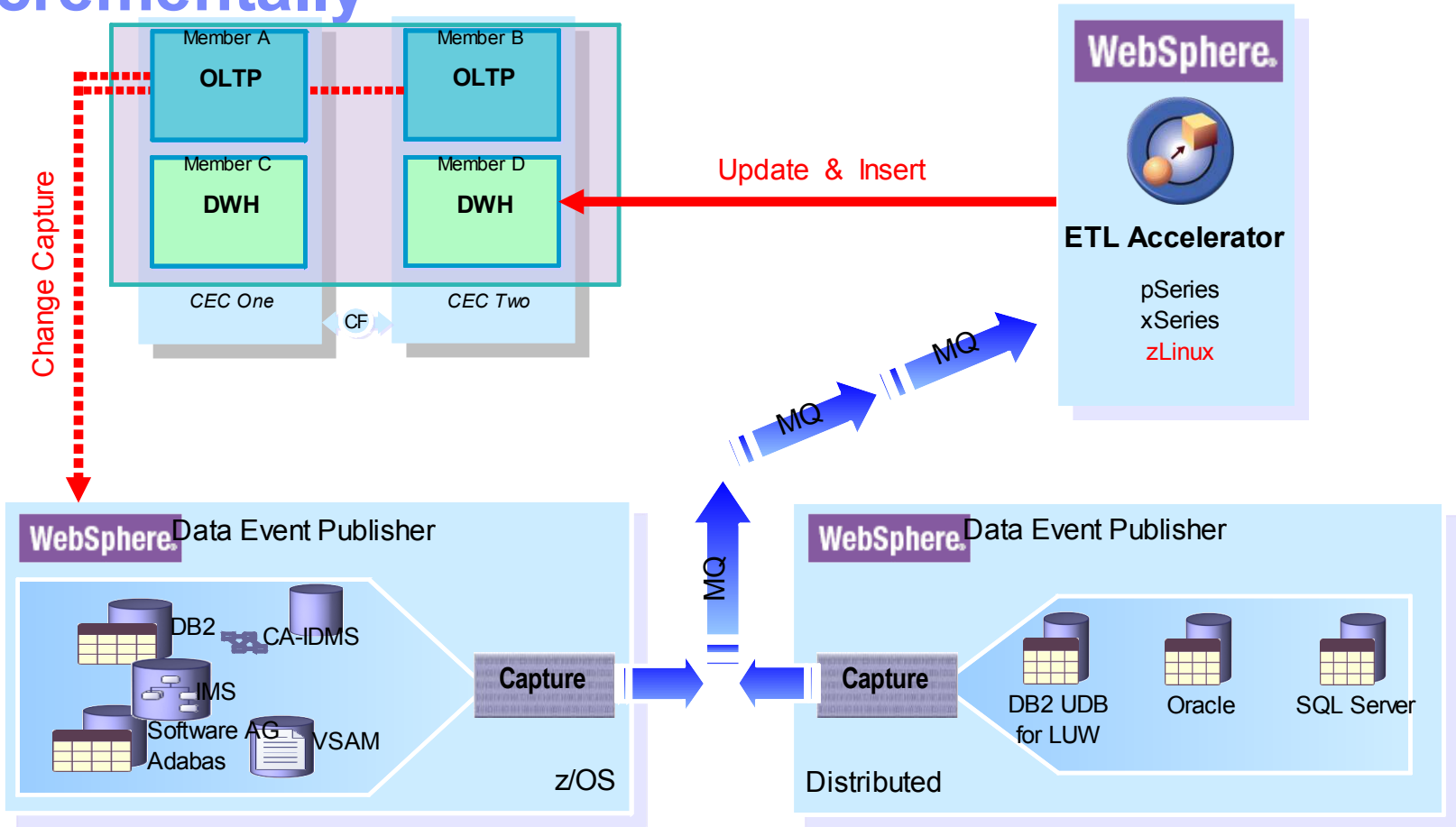
Within a data sharing environment, the data warehouse resides in the same group as the transactional data.



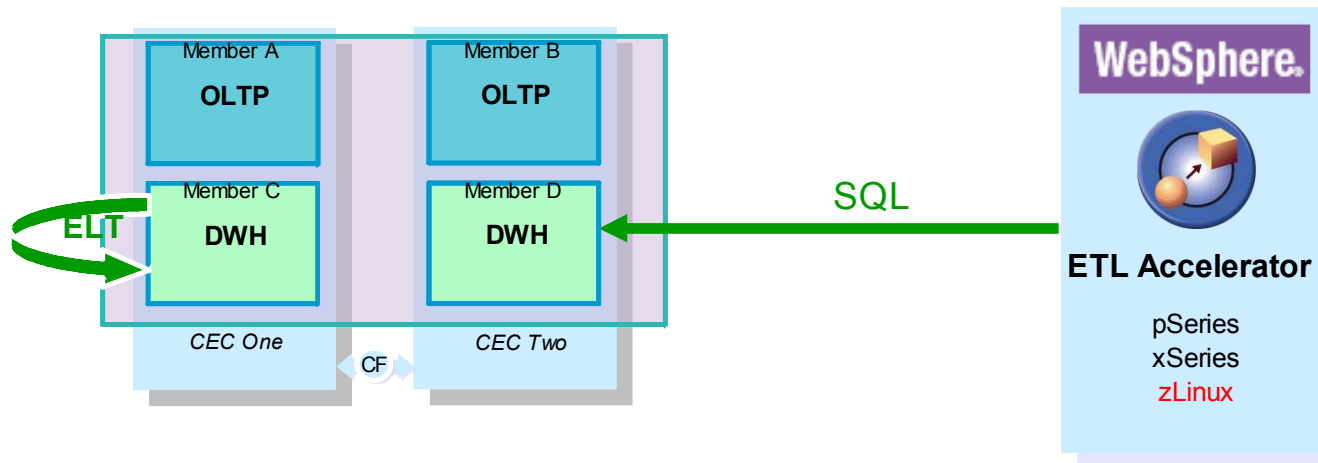
Initial load of the Data Warehouse



At runtime, the Data Warehouse is updated incrementally



In Database ELT is triggered by DataStage



Simple example:

```
-- Aggregate by salary by department into AGGRSALARY
```

```
INSERT INTO AGGRSALARY ( DEPTCODE, AVGBAND, AVGSALARY )
  SELECT DEPTCODE, AVG( BAND ) AS AVGBAND, AVG( SALARY ) AS AVGSALARY
  FROM STAFF
  GROUP BY DEPTCODE
```

But wait...

There's more...

What else is coming for DW on z?

- **SHORT TERM -**
 - **DB2 9**
 - Even more DB2 compatibility
 - Warehouse, reporting & optimizer enhancements
 - **BCU Methodology** (Sizing Tool, Best Practices)
 - **AlphaBlox on z/Linux**
- **MEDIUM TERM – Approx. 1 Year (Not Committed)**
 - **DB2 z/OS Enhancements**
 - Workload Based MQT Design Advisor
 - Common Web-based Admin Console
 - **DWE IuW BI Tooling Enabled for DB2 z/OS Data Access**
 - Cubing Services (Cube Design & Metadata Services)
 - Data Mining (Model Build & Model Deployment)
 - DWE Design Studio

What else is coming for DW on z?

- **LONGER TERM – Greater than 1 Year (Not Committed)**
 - **SHARK OFFLOAD**
 - Large table scans pushed down to Shark
 - Better price/performance
 - **Port of DWE BI Tooling to DB2 z/OS (Native)**
 - Cubing Services (Cube Design & Metadata Services)
 - Data Mining (Model Deployment)
 - DWE Design Studio

DW on System z Support Initiatives

- **BI Sizing Tool for System z (Internal Use)**
 - Target availability early 2H07
 - Input: Txn rates for 4 different txn profiles & DW raw data size
 - Output: CPU requirements (including zIIP estimate) plus guidelines for memory and storage

- **DW on System z Best Practices Guidance (Internal Use)**
 - Target V1 end of 2Q07 & V2 end of 4Q07
 - Recommended Architecture & Components
 - DB2 Configuration & Tuning Guidance

- **DW on System z SWAT Team**
 - Technical Sales & POC Support (as appropriate)

Summary

Value proposition of DW on System z

- **Qualities of Service**
 - Superior Quality
 - Super Availability
 - Security and Regulatory Compliance
 - Scalability
 - Backup and recovery
- **Positioned for the future**
 - Web-based applications
 - XML support
 - Service Oriented Architecture (SOA)
- **Operational data and the ODS together means**
 - Reduced complexity
 - Reduced cost
 - Shared processes, tools, procedures
 - Streamlined compliance and security
- **zIIP specialty engine improves TCO**
- **Better leverage System z skills and investment**

System z Enterprise Hub for Mission Critical Data

- ❖ With a strong foundation for transaction processing, built on 40+ years of technology innovation, System z servers with z/OS and DB2 can provide a premier platform for data serving (OLTP, Warehousing, Web Services, and more) today and into the future*
- ❖ IBM plans to continue to invest in new solutions to address



Today's Capabilities

- Industry-leading data integrity and security
- Data sharing solution for centralized view of data
- Scalability and availability for enterprise class workloads
- Comprehensive systems and data management environment



Extension of capabilities*

- New specialty engine (zIIP) with DB2 exploitation - for mission critical ERP, CRM, and Data Warehousing workloads *
- Database support improves regulatory compliance and autonomies
- Support of encryption capability (tape subsystem) with z/OS centralized key mgmt
- Data protection to achieve highest levels of security certifications



Future direction*

- Additional zIIP exploitation where it makes sense anywhere in the software stack
- DB2 enhancements to help improve usability and reduce complexity and management costs.
- DB2 query enhancements (table scan acceleration via DS8000 and more)
- Support of encryption capability (disk subsystem) with z/OS centralized key mgmt
- Handle larger volumes of data. with

*All statements regarding IBM's future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only.