

IBM Information Management



Optimize your System z Data Warehouse with IBM ETL and Analytic Tools

Burt Corwin, Pgm Mgr DB2 z/OS Initiatives Karen Durward, Information Integration Mike Biere, Mkt Mgr DW on DB2 z/OS August 21, 2007



Disclaimer

The information contained in this presentation has not been submitted to any formal IBM review and is distributed on an "As Is" basis without any warranty either expressed or implied. The use of this information is a customer responsibility.

The materials in this presentation are also subject to

- enhancements at some future date,
- a new release of DB2, or
- a Programming Temporary Fix (PTF)

IBM MAY HAVE PATENTS OR PENDING PATENT APPLICATIONS COVERING SUBJECT MATTER IN THIS DOCUMENT. THE FURNISHING OF THIS DOCUMENT DOES NOT IMPLY GIVING LICENSE TO THESE PATENTS.

TRADEMARKS: THE FOLLOWING TERMS ARE TRADEMARKS OR ® REGISTERED TRADEMARKS OF THE IBM CORPORATION IN THE UNITED STATES AND/OR OTHER COUNTRIES: AIX, AS/400, DATABASE 2, DB2, e-business logo, Enterprise Storage Server, ESCON, FICON, OS/390, OS/400, ES/9000, MVS/ESA, Netfinity, RISC, RISC SYSTEM/6000, iSeries, pSeries, xSeries, SYSTEM/390, IBM, Lotus, NOTES, WebSphere, z/Architecture, z/OS, zSeries, System z.

THE FOLLOWING TERMS ARE TRADEMARKS OR REGISTERED TRADEMARKS OF THE MICROSOFT CORPORATION IN THE UNITED STATES AND/OR OTHER COUNTRIES: MICROSOFT, WINDOWS, WINDOWS NT, ODBC and WINDOWS 95.

For additional information visit the URL http://www.ibm.com/legal/copytrade.phtml for "Copyright and trademark information"



Agenda

- Review of DW on System z (8/7 Teleconference)
- IBM Information Server Supporting DW on System z
- IBM Analysis & Reporting Tools Supporting DW on System z
- Q & A



Review of DW on System z

- Today's Teleconference is designed to build upon 8/7
 Teleconference
- "Leveraging DB2 z/OS and System z for Dynamic Warehousing"



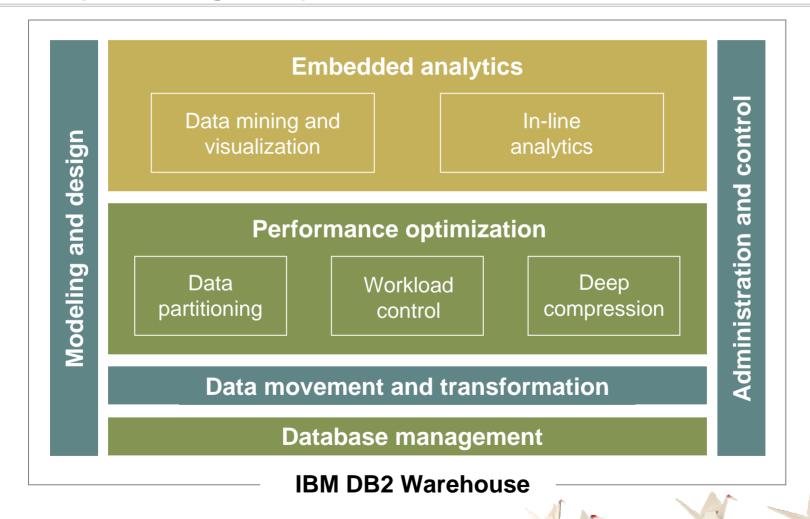
8/7 Teleconference Summary

- Introduced Dynamic Warehousing & Established DW on DB2 z/OS as a full participant in that IBM Initiative
- Introduced Layered Data Architecture as an approach that seeks to consolidate the logical components of a DW environment (ODS, DW & DM's) within a single database environment
- Reviewed What is Available Today to Support DW on System z
- Recommended a Solution Architecture for DW on System z

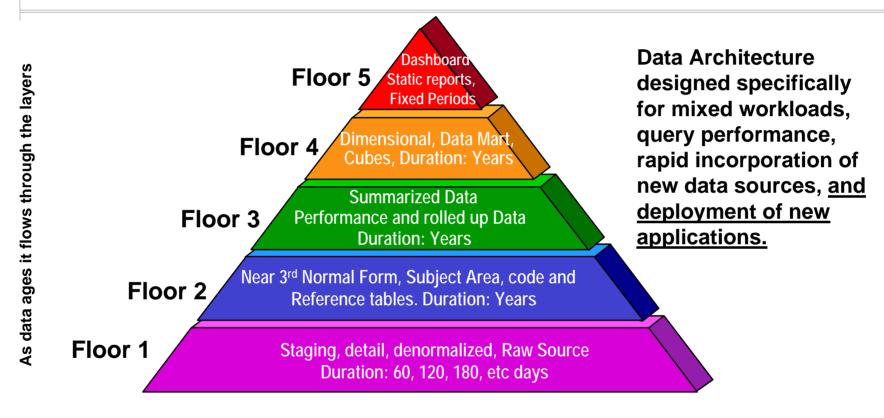


IBM DB2 Warehouse software

A complete, integrated platform



Layered "Data Architecture"





What DW/BI Capabilities Exist Today for System z

DB2 z/OS V8 & V9 EDW Base

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

Integration (EDW ETL)

- –WS DataStage EE, WS Classic Federation Server, WS Classic Data Event Publishers, Distributed DBMS Event Publishers, BatchPipes for OS/390, DB2 Unload/Load Utilities
- -IBM Information Server (WS Information Analyzer, WS QualityStage)

Analyze/Report

- -Alphablox, QMF, DataQuant
- -Partner offerings from Cognos, Hyperion (Oracle), Business Objects, SAS, IBI

Performance Management

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

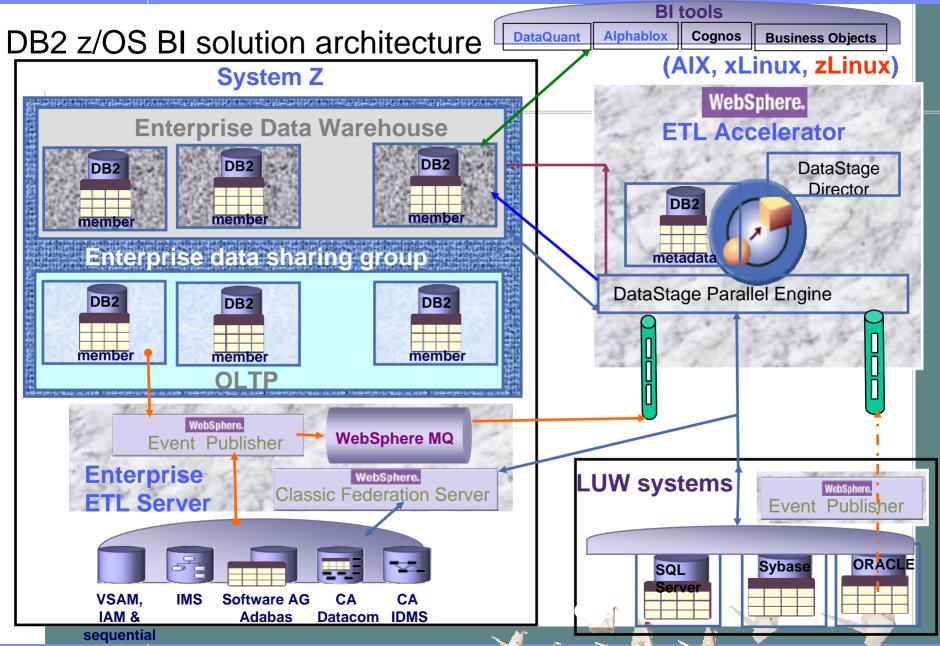
Security and Regulatory Compliance

 –DB2 Data Archive Expert, DB2 Test Database Generator, DB2 Audit Manager Expert, IBM Encryption for DB2 and IMS Databases

Application Management

-DB2 Table Editor, DB2 Web Query Tool, DB2 Automation Tool, DB2 Query Monitor

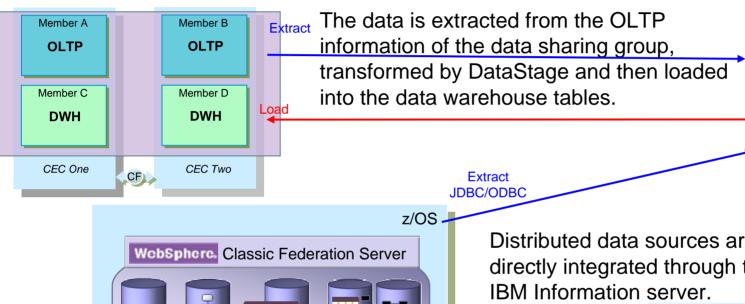






DataStage Parallel Engine

Initial load of the Data Warehouse



IDMS

Datacom

Legacy data sources are integrated through the Classic Federation Server. Data is extracted by DataStage directly out of mainframe data sources like IMS, VSAM...

Software AG

Adabas

Distributed data sources are directly integrated through the

Extract JDBC/ODBC

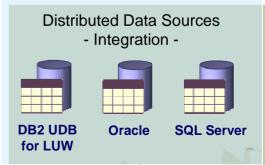
WebSphere.

ETL Accelerator

pSeries

xSeries

zl inux



VSAM.

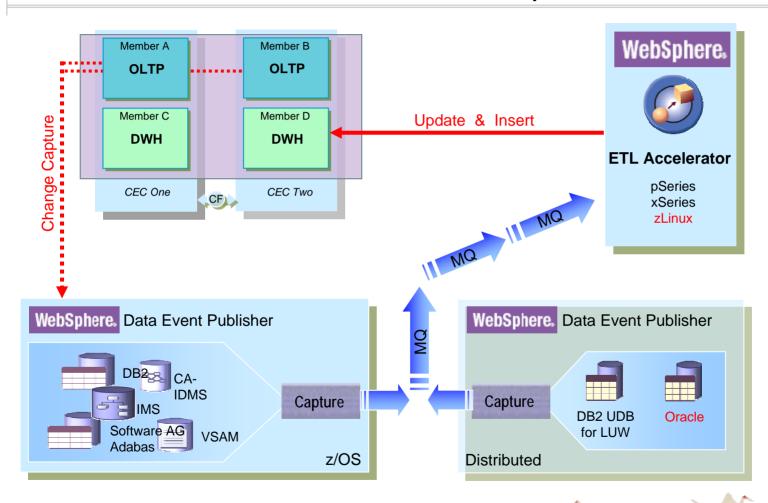
IAM &

sequential

IMS

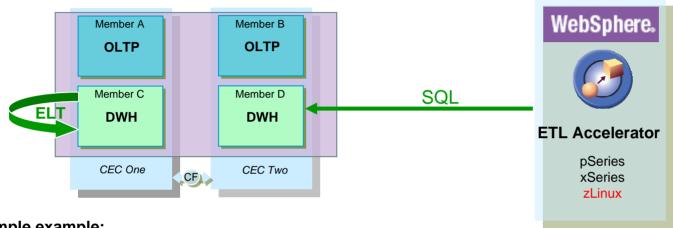


At runtime, the Data Warehouse is updated incrementally





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
```

Wherever possible, "In Database" transformations (ETL) are used to spare the transport of the data to the accelerator. But the used SQL is still sent from the ETL Accelerator to the database to have one place of documentation for all ETL steps.

This can also be used to shift the data up the hierarchy within the Layered Data Architecutre.



8/21 Teleconference Intent

- Expand on "building the warehouse" Components of BI Solution Architecture delivered by IBM Information Server Products
- Expand on Analysis & Reporting Component of BI Solution Architecture delivered by DataQuant & AlphaBlox Products



The IBM Information Server Delivering information you can trust

IBM Information Server

Unified Deployment

Understand



Discover, model, and govern information structure and content

Cleanse



Standardize, merge, and correct information

Transform



Combine and restructure information for new uses

Deliver



Synchronize, virtualize and move information for in-line delivery

Unified Metadata Management

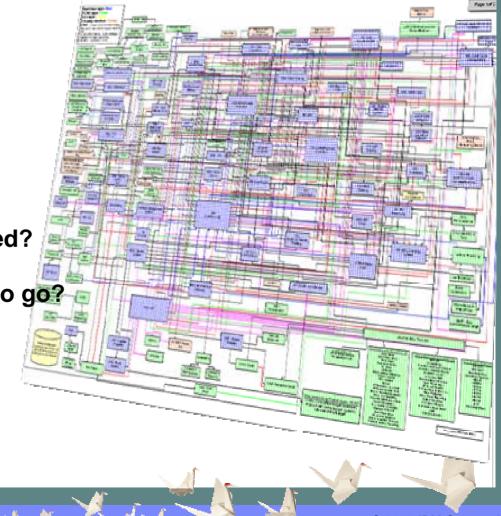
Parallel Processing

Rich Connectivity to Applications, Data, and Content



Why is it Important to Begin with Understanding?

- Where is my information?
- How do I get it when I need it?
- What does it mean?
- Can I trust it?
- How do I get it in the form I need?
- How do I get it where it needs to go?
- How do I control it?





WebSphere Business Glossary

Business Meta Data

- Web-based authoring, managing & sharing of business metadata
- Aligns the efforts of IT with the goals of the business
- Provides business context to information technology assets
- Establishes responsibility and accountability

Database = DB2

Schema = NAACCT

Table = DLYTRANS

Column = ACCT_NO

data type = char(11)



Technical



Business

GL Account Number

The ten digit account number. Sometimes referred to as the account ID. This value is of the form L-FIIIIVVVV.



Subject Matter Experts



Business Users

WebSphere Business Glossary

Create and manage business vocabulary and relationships, while linking to physical sources



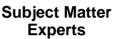


Rational Data Architect

Logical Meta Data

- Data modeling for data structures and federations
- Federated data discovery
- Metadata relationship discovery and mapping
- Impact analysis, and synchronization across models
- SQL & XML generation capabilities



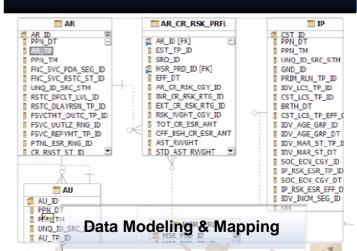




Architects

Rational Data Architect

Create and manage business vocabulary and relationships, while linking to physical sources





WebSphere Information Analyzer

Physical Meta Data

- Data-centric analysis of application, database and file-based sources
- Secure, detailed profiling of fields, across fields and across sources
- Creation of metadata from profiling results
- Results instantly promotable across IBM Information Server

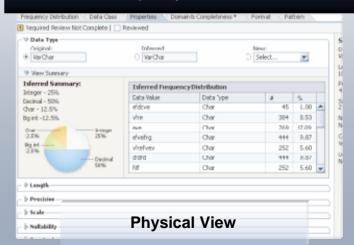


Subject Matter Experts

Data Analysts

WebSphere Information Analyzer

Analyze source data structures, and monitor adherence to integration and quality rules





Why Care About Cleansing Information?

- Lack of information standards
 - Different formats & structures across different systems
- Data surprises in fields
 - Data misplaced in the database
- Information is buried
 - Free-form fields
- Data myopia
 - Lack of consistent identifiers inhibit a single view of:
 - customers products ...
- The redundancy nightmare
 - Duplicate records with a lack of standards

```
416 Columbus Ave #2, Boston, Mass 02116
Kate A. Roberts
Catherine Roberts Four sixteen Columbus APT2, Boston, MA 02116
Mrs. K. Roberts
                  416 Columbus Suite #2, Suffolk County 02116
Name
                                        Telephone
                         Tax ID
J Smith DBA Lime Cons.
                         228-02-1975
                                        6173380300
Williams & Co. C/O Bill
                         025-37-1888
                                        415-392-2000
1st Natl Provident
                         34-2671434
                                        3380321
HP 15 State St.
                         508-466-1200
                                       Orlando
WING ASSY DRILL 4 HOLE USE 5J868A HEXBOLT 1/4 INCH
WING ASSEMBY, USE 5J868-A HEX BOLT .25" - DRILL FOUR HOLES
USE 4 5J868A BOLTS (HEX .25) - DRILL HOLES FOR EA ON WING ASSEM
RUDER, TAP 6 WHOLES, SECURE W/KL2301 RIVETS (10 CM)
19-84-103
            RS232 Cable 6' M-F CandS
CS-89641
            6 ft. Cable Male-F, RS232 #87951
C&SUCH6
            Male/Female 25 PIN 6 Foot Cable
90328574
                                   187 N.Pk. Str. Salem NH 01456
            TRM
90328575
            I.B.M. Inc.
                                  187 N.Pk. St. Salem NH 01456
90238495
            Int. Bus. Machines
                                  187 No. Park St Salem NH 04156
90233479
            International Bus. M. 187 Park Ave Salem NH 04156
            Inter-Nation Consults 15 Main Street Andover MA 02341
90233489
90345672
            I.B. Manufacturing
                                   Park Blvd. Bostno MA 04106
```



Cleanse

WebSphere QualityStage

- Specialized data quality functions seamlessly integrated with DataStage
- Visual tools for defining complex matching and survivorship logic
- Ensures clean, standardized, deduplicated information
- Enables a single version of the truth







Data Analysts

WebSphere QualityStage™

Standardize and correct source data fields, and match records together across sources to create a single view

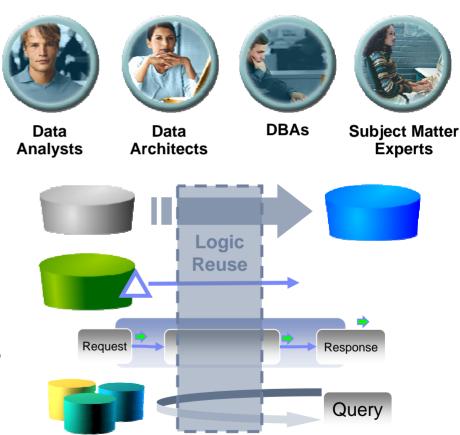


Visual Match Rule Design



What Is Important About Transformation & Delivery?

- Transformation is key to enabling information to be used in new business contexts – it needs to be metadata-driven
- Designed for use by information experts using the understanding imparted by the metadata
- Transformation and Delivery can be reused across multiple mechanisms
 - Large volume batch movement
 - Real-time event-driven response
 - Service-oriented architecture
 - Federated query





Transform

Deliver





WebSphere DataStage

- Codeless visual design of data flows with hundreds of built-in transformation functions
- Optimized reuse of data integration objects
- Leverages parallel processing without requiring design changes
- Capable of supporting batch and real-time operations



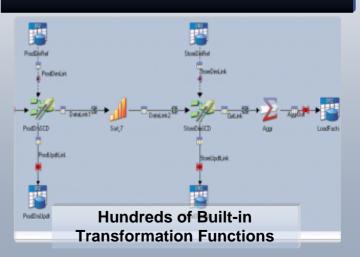


Developers

Architects

WebSphere DataStage®

Transform and aggregate any volume of information in batch or real time through visually designed logic





Deliver



WebSphere Federation Server

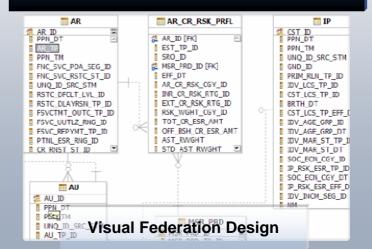
SQL Access to distributed data

- Access diverse & distributed information as if it were in one system
- Industry leading query optimization with single sign-on, unified views, and function compensation
- Transactional write capabilities across heterogeneous sources
- Visual tools for federated data discovery & data modeling

WebSphere Federation Server

Access and integrate heterogeneous information across multiple sources as if they were a single source

Extend value of existing analytical applications by providing real-time access to integrated information





Deliver



WebSphere Classic Federation Server

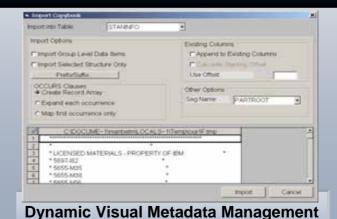
Read/Write mainframe data using SQL

- Metadata-driven, so there's no mainframe programming required
- Works with existing mainframe infrastructure and "modern" applications and tools you need
- Deliver mainframe data to
 - Self-service portals
 - e-commerce solutions
 - Reporting and analytical tools
 - IBM's own data profiling, cleansing and transformation solutions
- Sources:
 - VSAM, IMS, CA-IDMS, CA-Datacom, Software AG Adabas, DB2

IBM WebSphere Classic Federation Server

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

Empowers mainframe data integration with Information Server components, your applications as well as IBM and 3rd party tools and applications





Deliver



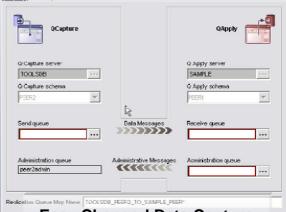
WebSphere Data Event Publishers

Push Changed Data "Events"

- Feed data to DataStage for change-only updating of the warehouse
- Loosely-coupled, no impact on applications
- Low-latency or scheduled data capture
- Multiple publication formats:
 - Consistent relational format for ease of use
 - XML for ease of consumption
 - Delimited values minimizes message size
- Recoverability, assured delivery
- Sources:
 - DB2, IMS, CA-IDMS & Software AG Adabas

IBM WebSphere Data Event Publisher

Detect and respond to data changes in source systems, and publish changes to subscribed systems, to ETL or to other modules for eventbased processing



Easy Changed Data Capture



The IBM Information Server What makes it a PLATFORM?

IBM Information Server

Unified Deployment

Understand



Discover, model, and govern information structure and content

Cleanse



Standardize, merge, and correct information

Transform



Combine and restructure information for new uses

Deliver



Synchronize, virtualize and move information for in-line delivery

Unified Metadata Management

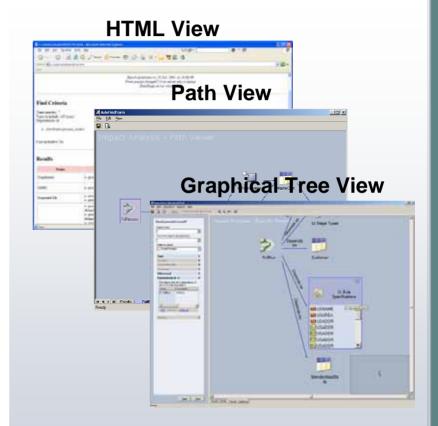
Parallel Processing

Rich Connectivity to Applications, Data, and Content



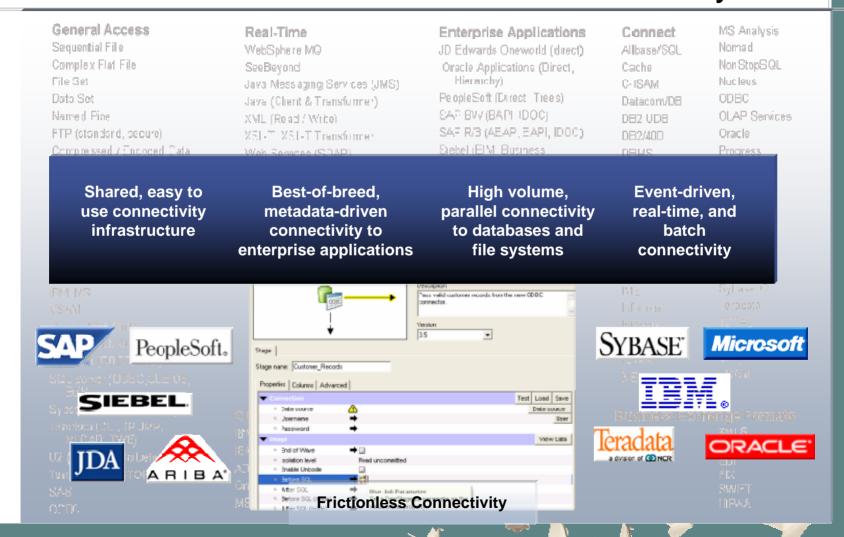
IBM Information Server Platform Features

- Role-Based Tools with Integrated Metadata
 - Simplify Integration
 - Promote reuse
 - Increase compliance to standards
 - ▶ Open to 3rd party integration
- Visual lineage reporting increases trust
- Graphical impact analysis facilitates change management





Information Server Platform: Rich Connectivity





WebSphere Information Services Director

Rapid SOA Deployment:

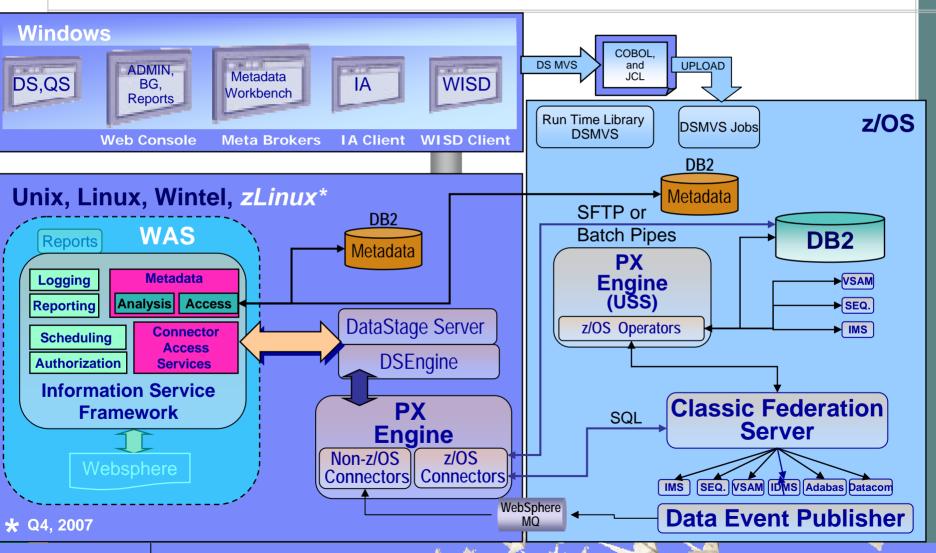
- Packages information integration logic as services that insulate developers from underlying sources
- Allows these services to be invoked as Enterprise Java Beans or Web services
- Provides load balancing & fault tolerance for requests across multiple Information Servers
- Provides foundation infrastructure for Information Services





IBM Information Server Implementation Architecture

Leveraging the power of a multi-platform Environment





IBM Analysis & Reporting Products

IBM DB2 DataQuant

- Out of the box query, reporting, analysis and dashboards
- Server-based licensing and deployment
- Rich client and thin client architectures
- Easy to use, programmer free



IBM DB2 Alphablox

- In-depth analytical tool
- Rich development environment for imbedded analytics
- OLAP analysis and 'cube engine'



Overview – IBM DataQuant

Provides

- a comprehensive query, reporting and data visualization package
- support for both web and workstation-based environments.

Introduces

- a wide variety of powerful business intelligence capabilities, from executive dashboards
- interactive visual applications to information-rich graphical reports and ad-hoc querying and data analysis.

Together with over 100 built-in analytical functions, IBM DataQuant allows organizations to derive maximum value from their data and rapidly build and distribute comprehensive data visualization solutions across the enterprise.



Product Components

DataQuant for Workstation

- Java-based, rich client application based upon the award-winning Eclipse platform
- Provides a graphical environment for the development of query, report, dashboard and database visualization solutions
- Supports multiple workstation environments, including Windows, Linux, HP-UX and Solaris

DataQuant for WebSphere

- Provides a comprehensive runtime environment, capable of displaying all DataQuant content, under a HTML, 'thin client' model
- Runs under IBM WebSphere Application Server (WAS) on all platforms supported by WAS, including Windows, Linux, Solaris, HP-UX, iSeries, zLinux and z/OS



Highlights

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

SOA Layer

 Allowing queries, reports & BI solutions to be shared with both internal users and external organizations outside of the enterprise

Enhanced Analytics

Adds over 100 analytical functions to queries, reports and visual solutions

Backward-compatible with existing QMF infrastructure





Visual Dashboards



Provides a rapid development environment for:

- Executive dashboards
- Information portals
- What if'-style analytical applications

Key benefits

- Highly graphical, visually intuitive solutions
- Rapid development model permits solutions to be developed and deployed within days, rather than weeks, using a zero-coding, drag and drop authoring model
- Optionally leverages existing QMF queries

- **Deployment** via web browsers (thin client) and/or the DataQuant for Workstation application (rich

client)





Visual Dashboards (cont.)

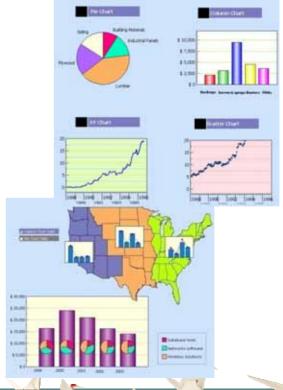
Supports over 20 'stock' visual layouts

- Line and bar charts
- Geographic maps
- Pie charts
- Horizon charts
- Hierarchical charts
- Tabular layouts

Flexible rendering

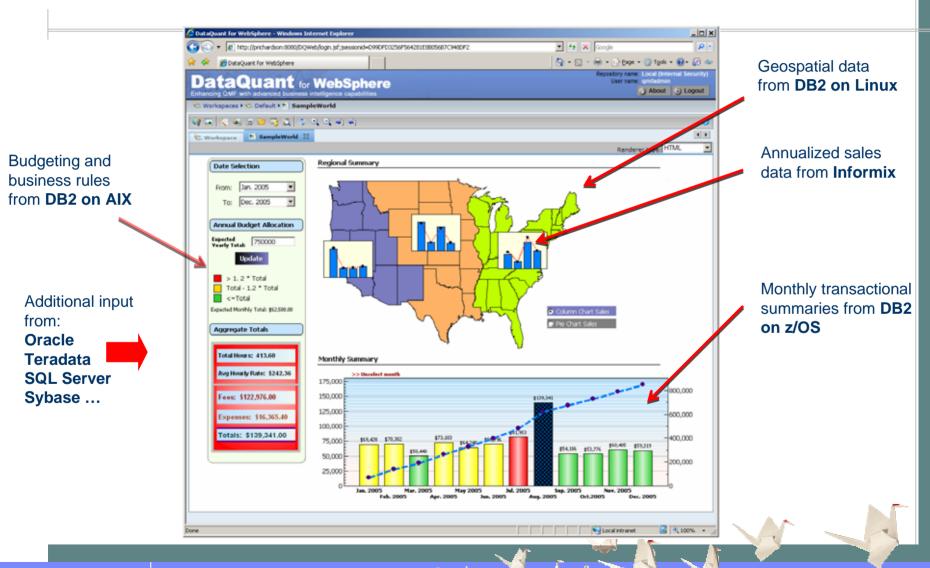
- Motifs can be nested indefinitely
 (e.g. pie charts upon bars within maps etc.)
- Customized graphical 'templates' users can build their own chart types







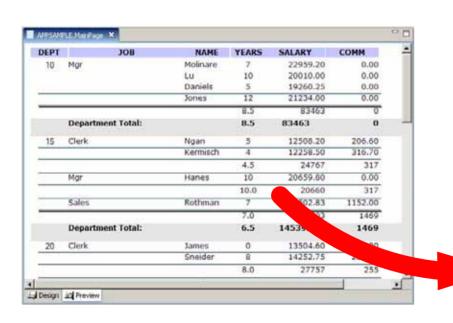
Enterprise Business Intelligence

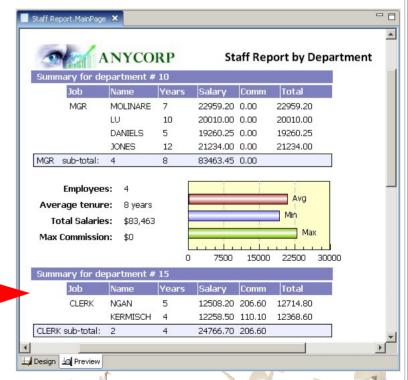




Enhanced Graphical Reporting

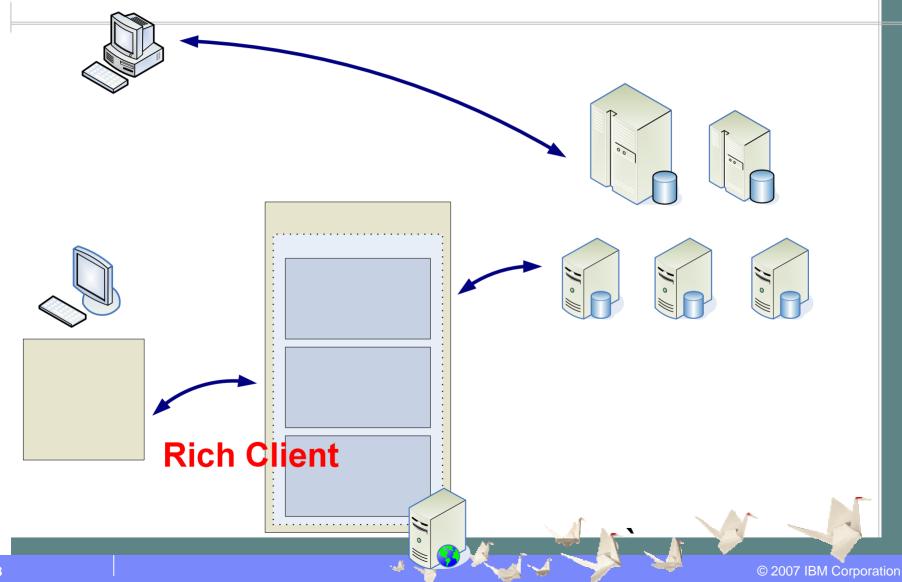
- Transforms text-based reports into highly graphical, information-rich documents (HTML or PDF output)
- Familiar data-driven 'sectional report' model, with ability to 'nest' an unlimited number of subordinate queries







DataQuant Architecture





Security & Personalization

Robust security and support for single sign-on

- Introduces a comprehensive security foundation
- Optionally interfaces with existing LDAP directories or RACF etc.
- Built-in 'internal security' model available for customers that do not operate LDAP-based directories
- Optionally logs onto data sources using accounts mapped from LDAP/internal security directory

Personalization

- Tailors 'look and feel', including available reports, visualizations and data on a per user/group basis
- Ability to store user and group-specific offline views of database contents
- Ability to create an unlimited number of workspaces, each containing differing views of available data sources, reports and visual dashboards



Backward Compatible with QMF Infrastructure

- Queries, forms and procedures remain object-compatible with prior QMF versions
- Retains full support for QMF Catalogs
 - Ability to create, run, edit and share QMF objects with QMF for TSO¹ and prior QMF distributed products
 - Optionally use resource limits defined in QMF catalogs
- Similar 'database explorer' provides familiar look and feel for prior QMF users
- Built-in ability to import QMF Server Definitions Files (SDF) and QMF Visionary solutions
- Built-in ability to migrate visual reports into the new graphical format





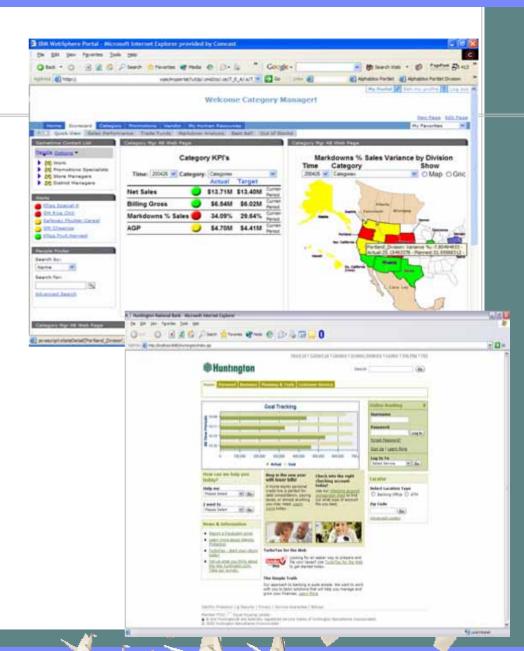
IBM DataQuant for z/OS V1.1 - 5697-N64

- OTC offering, IPLA product
- Processor pricing -- \$1500/VU
- Pre-reqs QMF Enterprise Edition (either V8 or V9)
- Supports DB2, Informix, and other major relational databases
 - It offers access to DB2 on any platform (including i-series), plus other relational DBs
- DataQuant for WebSphere runs on any WebSphere supported OS
 - Including zOS and zLinux
- Try & Buy available



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



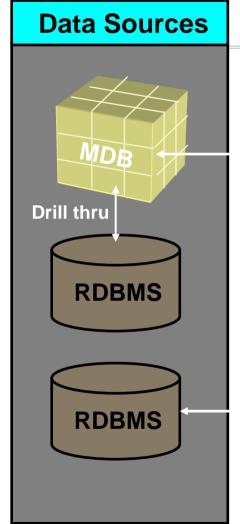


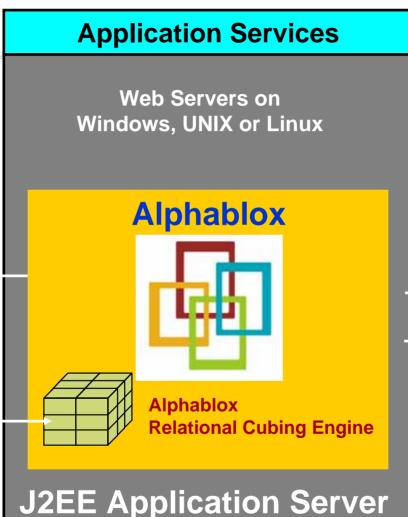
Alphablox Components

- •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.



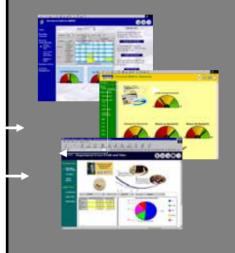
Alphablox Platform Architecture





Presentation

Internet Explorer, Mozilla, FireFox



Alphablox Analytical Applications



DataQuant / AlphaBlox Positioning

Where IBM DataQuant fits best

- Where there's a need to distribute data using turn key graphical reports and information dashboards
- Where quick prototyping and rapid development is more important than sophisticated analytical features
- In QMF and/or z-based environments where tracking, governing and z-based deployment are valued factors
- For customers that seek an 'out of the box' packaged BI solution for the development and distribution of query, reporting and dashboard content

Where DB2 Alphablox fits best

- Where there is a need for customized analytical solutions that are tightly integrated and embedded within the customer's existing BI infrastructure
- Where "in-line" analytics is required—that is, where analytical application content is embedded in a business process
- For customers that require full control of the solution's appearance and behavior, from data gathering and analysis to the user interface 'look and feel'
- Where there is a need to include sophisticated analytical capabilities
- As a development environment for customized Web applications instead of packaged apps



Thank You for Joining Us today!

Go to www.ibm.com/software/systemz/events to:

- Replay this teleconference
- Replay previously broadcast teleconferences
- Register for upcoming events



IBM Information Management



Backup Charts





Data Warehousing for DB2 on System z – more info

- Whitepaper: Why Data Warehousing on System z available on the web http://www.ibm.com/software/data/db2bi/systemz.html
- DW on system z **Demo** available in the Technical Marketing Competence Center, Böblingen, Germany, <u>TMCC@de.ibm.com</u>
- DW on system z Customer Briefings possible in the Technical Marketing Competence Center, Böblingen,Germany or at Silicon Valley Lab TMCC@de.ibm.com
- DW on system z 'Best Practices' paper planned to be available 2Q2007

• ...

