



Key Considerations for Mainframe Data Integration



TAKE BACK CONTROL

Agenda

PART I: Information on demand and data delivery

- Delivering Information
 - Data-centric Integration Patterns
 - Sample Implementations
 - The Role of Replication
- Information in a Service Oriented Application (SOA) Environment
- Next Generation Data Warehousing

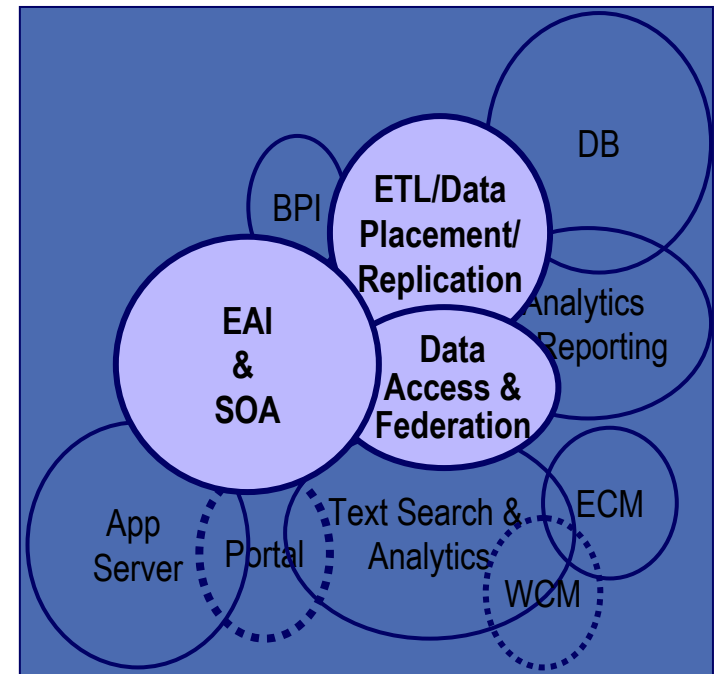
PART II: The IBM Information Server

- The demand for Information Integration
- A platform approach
- Technology
- Information Server for System z
- Wrap-Up: Where do you go from here?



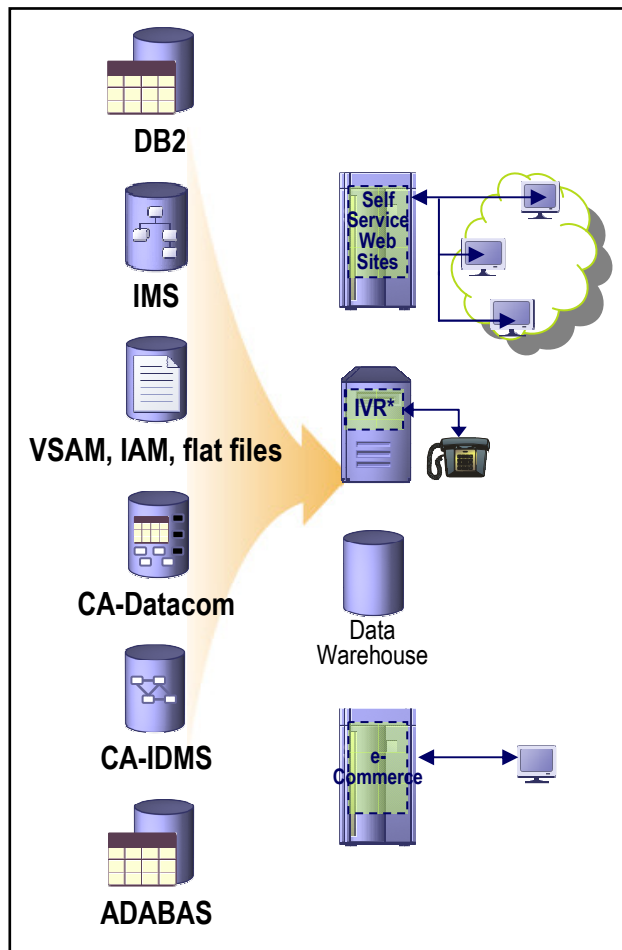
Data Delivery & Integration Landscape

- Process &/or Information
 - EAI “primarily” focuses on application assets
...secondarily on data assets
 - ETL differentiates on transformation
 - Replication focuses on synchronization
 - Federation is standardized data access
 - Changed-data-capture is an enabler
- Emphasis is on complementary uses
 - Incremental, real-time data warehousing
 - ETL fed by changed-data-capture
 - Empower information reuse, flexibility
 - SOA-based data access
 - Information-driven enterprise integration
 - Published “data events” drive EAI
 - Guaranteed high volume deliver
 - Replication to synchronize copies



Mission Critical "z" Data Can Be Difficult

Over 60% of operational data is sourced from System z



Why is system z data delivery an issue?

- Proprietary databases – multiple, complex APIs
- Billions of lines of mainframe code – must leverage
- Integration with data warehousing and SOA not “native”

What are you doing to meet these challenges?

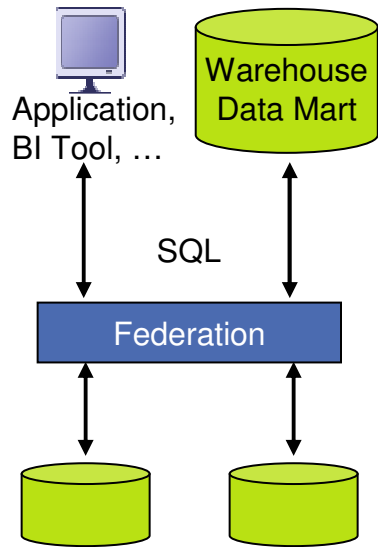
- *Hand-coding* – COBOL/PLI programs for every data need
 - Bottleneck for new initiatives ... slows everything down
 - Dependence on unique skills ... proprietary API experts
- *The One-Size-Fits-All Approaches:*
 - ▶ Integrate the transactions ... this can be:
 - overly complex
 - maintenance intensive due to hard wired integration
 - ill suited to the functional demands of new workloads
 - ▶ A unique "copy" for every requirement... this can lead to:
 - Latency errors
 - Synchronization challenges
 - Quality issues



The Information-Centric Delivery Patterns

Virtualize

All data participates equally



Seamlessly integrate with:

- e-commerce
- Self service
- SOA

Move

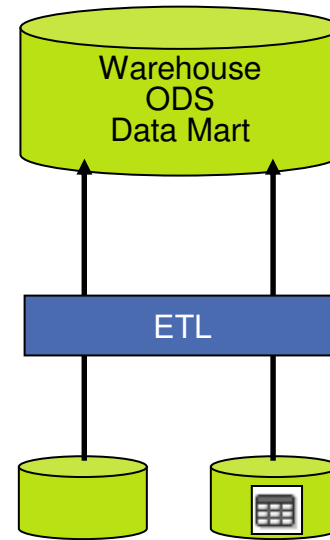
Data changes (events) drive action



Proactively deliver data to:

- Drive integration
- Enable change only updating to:
 - Reduce batch window dependencies
 - Reduce data latency

Transformation rich copies, aggregates, summaries

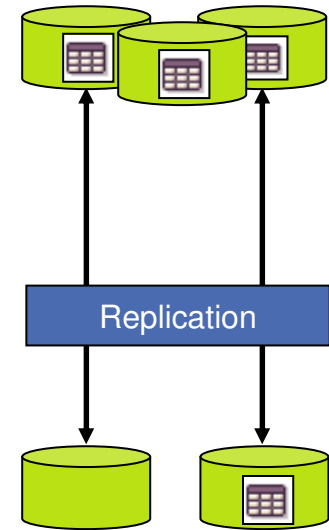


Repackage & repurpose for:

- Data Warehousing
- Operational Data Stores
- MDM
- CDI

Synchronize

Like-to-like copies



Create & synchronize copies for:

- Distributed operations
- High availability data
- Local high performance copies
- DBMS migrations



Choosing an Integration Pattern

A multi-dimensional self service environment

Insurance carrier wants to reduce costs and improve customer and agent satisfaction by implementing self-service environments...BOTH Interactive-voice-response and Web sites

- Initial Environment
 - Call center representatives take calls from agents who interface with clients
 - Mainframe systems consisting of over 1,000 transactions handle:
 - Accounting, Policy management, Claims and more
 - Technology is
 - IMS transactions
 - IMS, VSAM and DB2 data
 - Very complex after decades of application evolution

- Requirements
 - Minimal disruption on existing call-center
 - Staged implementations showing value every 3 – 6 months
 - Performance and accuracy are essential to credibility
 - Unskilled users (agents and customers) mandate an easy-to-use interface

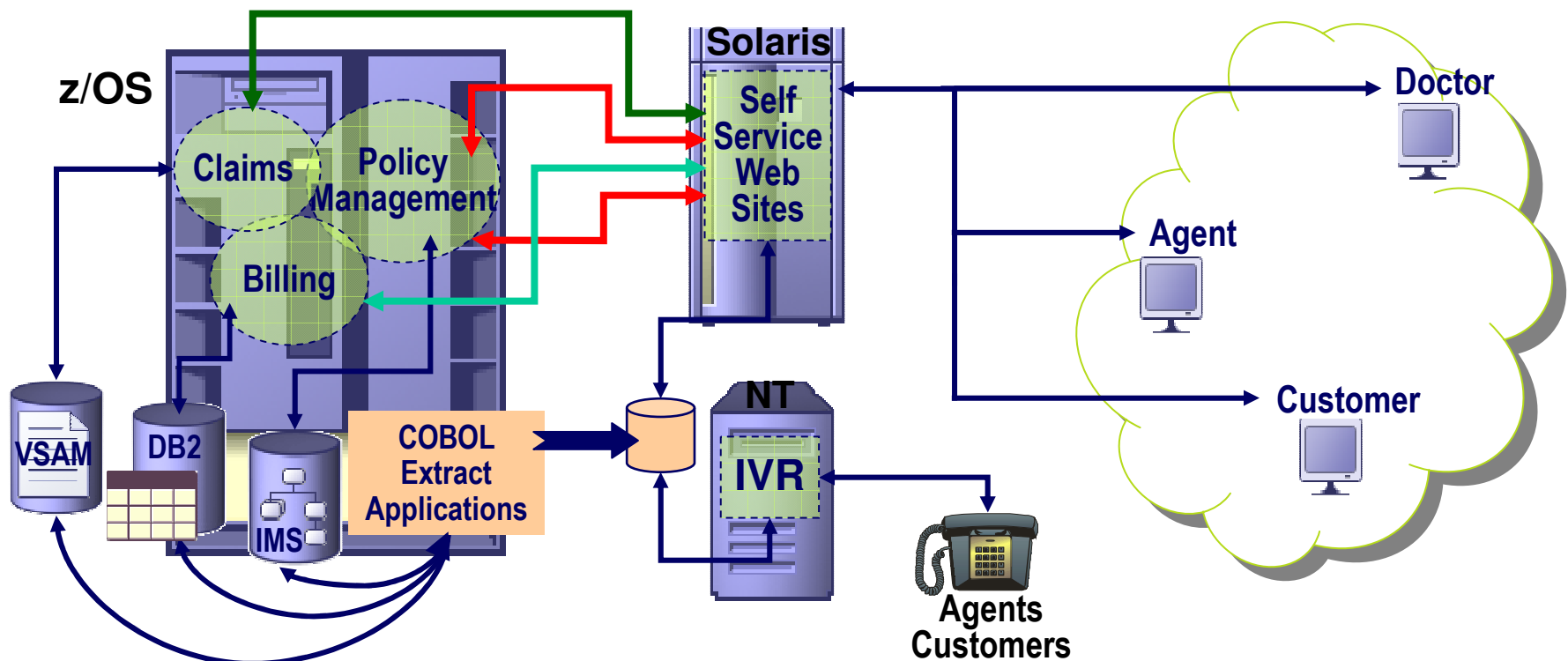


A Multi-Dimensional Self-Service Environment

EAI and ETL are traditional choices

Complexity and volume made them too costly and too time consuming

- Copy the data to Oracle \$2M for hardware and software
- Connect through transactions.... 10,000 man hours per system with little reusability



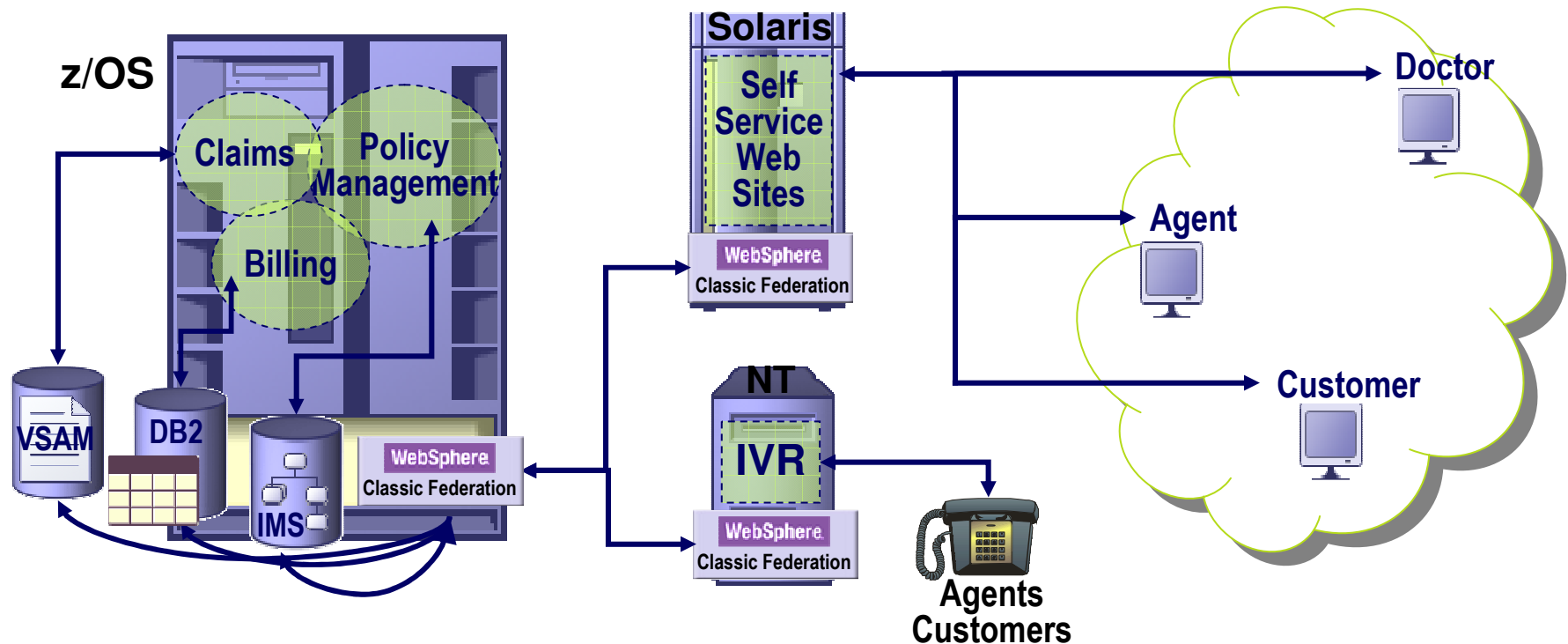
TAKE BACK CONTROL

A Multi-Dimensional Self-Service Environment

Federation delivered in a fraction of the time and cost

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
- No disruption to operational systems with minimal overhead in OLTP environment



Choosing an Integration Pattern

You may need more than one

Catalog retailer is building an e-commerce site

- Initial Environment
 - Call center representatives take orders
 - Mainframe order-processing system used by call center reps
 - CICS transactions
 - CA-IDMS data
 - Highly-optimized for use by skilled employees

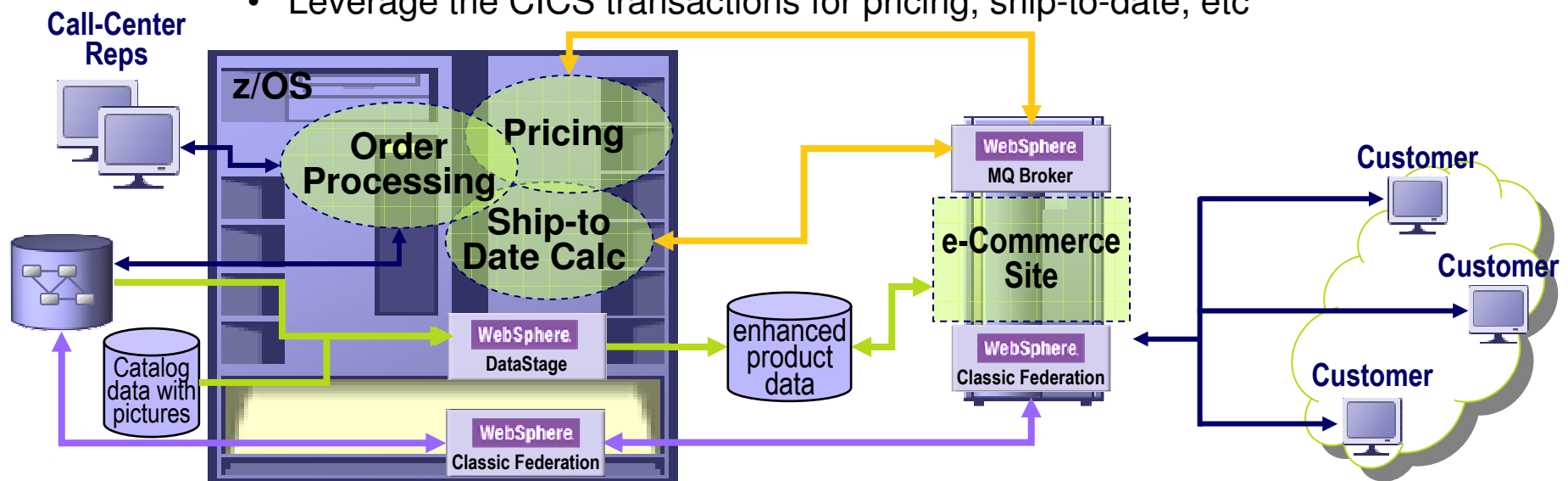
- Requirements
 - Minimal disruption on existing call-center order processing
 - Must support “shopping”, not just order taking
 - E-commerce site must be optimized for unskilled customer usage
 - Performance and accuracy are essential to business continuity and reputation



"Best" solution may leverage multiple patterns

ETL, Federation and EAI

- **ETL:** merge operational product data with image content
 - Product information is copied to “local” RDBMS for performance
 - Changes are infrequent and can be scheduled
 - Value-add image data added as part of the data transformation
- **Federation:** critical “real time” data is dynamically accessed
 - There can only be one version of the “truth” for inventory
- **EAI:** critical processes are shared
 - Leverage the CICS transactions for pricing, ship-to-date, etc



Choosing an Integration Pattern

Federation & Event Publishing as enabling technologies

Global distributor wants to implement a Web e-commerce environment, reduce inventory costs and improve decision making

- Initial Environment
 - Two mainframe-centric transaction processing centers
 - Ensures 24 by 7 operations for their Global business
 - Technology environments
 - Mainframe CICS environment servicing ~50 million transactions/week
 - IMS databases hold most of their data

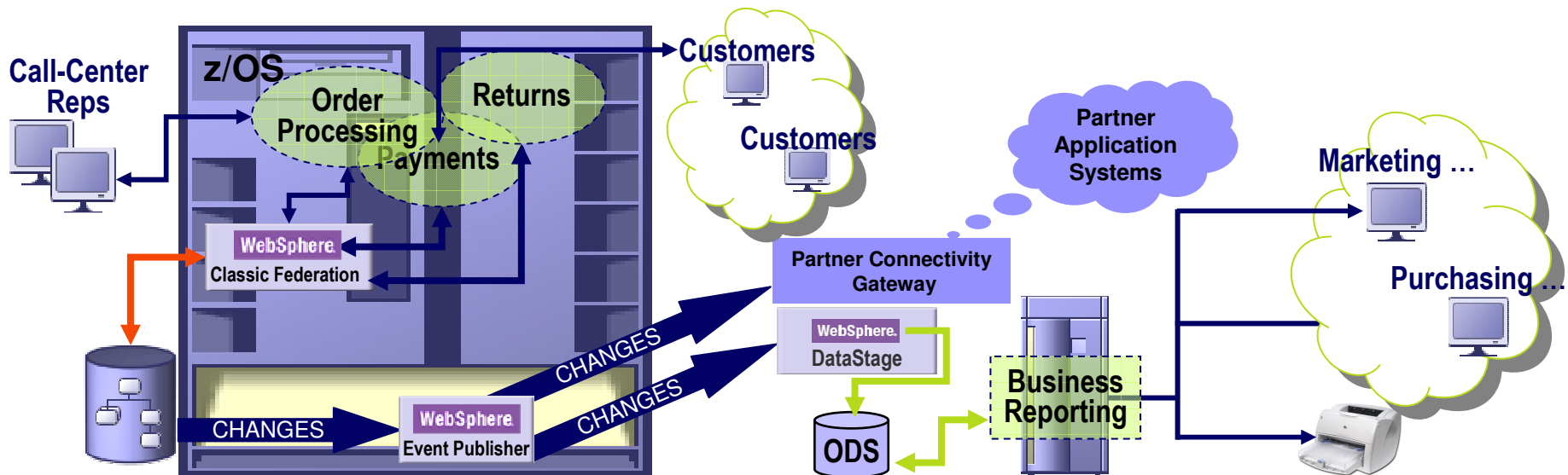
- Requirements
 - Implement their Next-Gen transaction platform on WebSphere
 - Staying with z/OS & IMS data for performance, reliability and security
 - Improve interfaces with their suppliers
 - Implement "real time" interfaces for a just-in-time inventory environment
 - Leverage reduced cost of an "off shore" data center for non-transactional workloads
 - Implement an ODS to support business reporting and decision processing



Federation & Event Publishing Behind ...

Transaction processing, ETL and EAI

- Federation enables Web-based applications
 - Java-based access to IMS data with transactional performance
- Changes delivered to India data center's ODS via ETL
 - Continuous, 24 by 7 delivery:
 - Ensure up-to-the-second data is available
 - Eliminate dependence on batch cycle
 - Steady, predictable usage of communication resources
- Changes delivered to Partner Gateway (EAI)
 - Enables Just-in-Time inventory management



Choosing a Data Movement Solution

Replication or ETL?

Financial Services firm requires 24 by 7 operational environment for financial trading applications

- Starting Environment
 - Two mainframe-centric data centers approximately 1,500 miles apart
 - HADR environment close to primary data center
 - Technology environments
 - Mainframe CICS environment servicing 1,000s of transactions per second
 - DB2 databases hold most of their data
- Requirements
 - Provide secondary operational platform for rapid switch-over
 - Improve TCO profile by leveraging secondary platform for "real" work

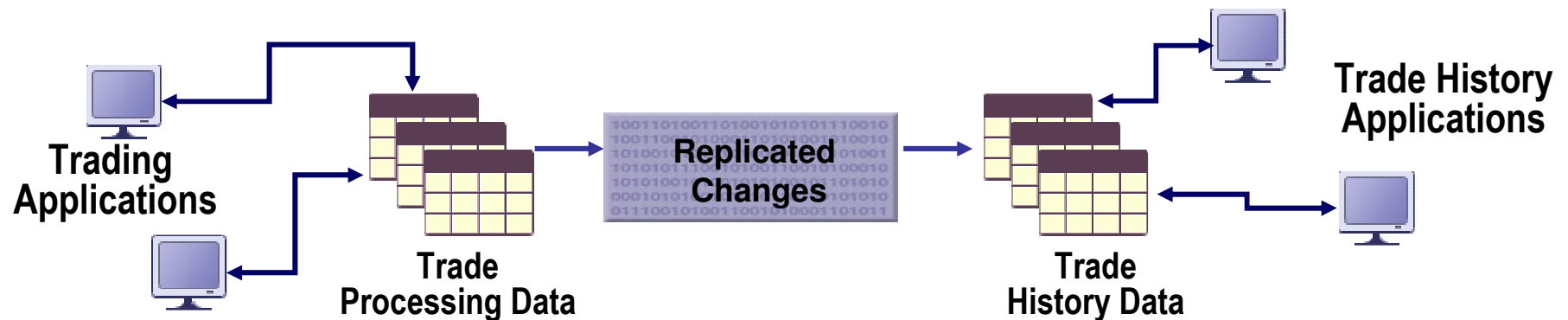
Replication is "best" because:

- *Low latency is critical*
- *No transformations are required*
- *Easy to switch source & target during a failure*



On-line Trading Environment

High Availability Replication



- Replication synchronizes data between the primary & secondary platforms
 - 5-10 million transactions daily replicated to 2ndary platform w/ <1 second delay
 - All query processing routed to 2ndary platform to improve primary throughput
- On primary system failure:
 - Secondary system becomes primary
- On primary recovery:
 - Secondary system replicates to primary until “caught up”
 - Trading users redirected back to primary



Choosing a Data Movement Solution

Replication or ETL?

Rapidly expanding retailer wants to push price changes from central system to hundreds of stores to improve accuracy and competitive position

- Starting Environment
 - Headquarters mainframe data center with DB2 pricing data
 - Store inventories and departments vary ... not all prices apply to every store
 - Communications bandwidth varies from 256k to 1Meg

- Requirements
 - Push price changes out as needed to rapidly react to changing demand
 - Ensure independence of each store's updates
 - Anticipate need to return some data back to HQ

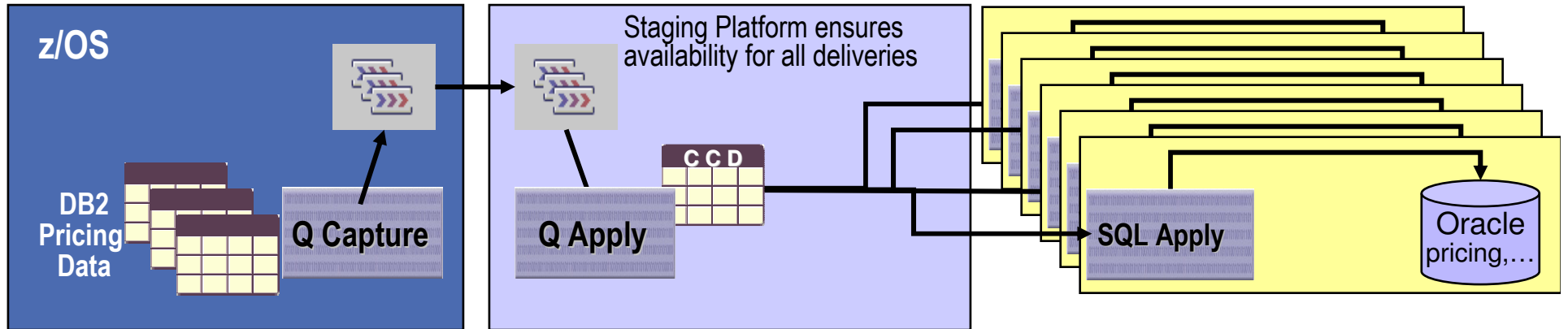
Replication is "best" because:

- *Automatic recovery with guaranteed delivery*
- *Minimal transformations are required*
- *Bidirectional capabilities inherent in Replication*



Rapid, Reliable Data Distribution

Leveraging Multi-faceted Replication



- Customer Pain points
 - Hundreds of thousands of products, with variations from store to store demands automation
 - Guaranteed delivery is essential ... customer satisfaction requires consistency
 - Adding hundreds more stores in next few years
- Value Proposition
 - Capture & guaranteed delivery of "pricing updates"
 - Each store: (1) receives only what it needs (2) is fully recoverable (3) is independently updated
 - Flexible targeting enables changes
 - Enables store-to-HQ replication as needed



Agenda

PART I: Information on demand and data delivery

- Delivering Information
 - Information Integration Patterns
 - Sample Implementations
 - The Role of Replication
- ***Information in a Service Oriented Application (SOA) Environment***
- Next Generation Data Warehousing

PART II: The IBM Information Server

- The demand for Information Integration
- A platform approach
- Technology
- Information Server for System z
- Where do you go from here?



Information On Demand and SOA

Two sides of the same coin...

SOA provides business flexibility

IOD provides the trusted information needed by your services



“You will waste your investment in SOA unless you have enterprise information that SOA can exploit.”

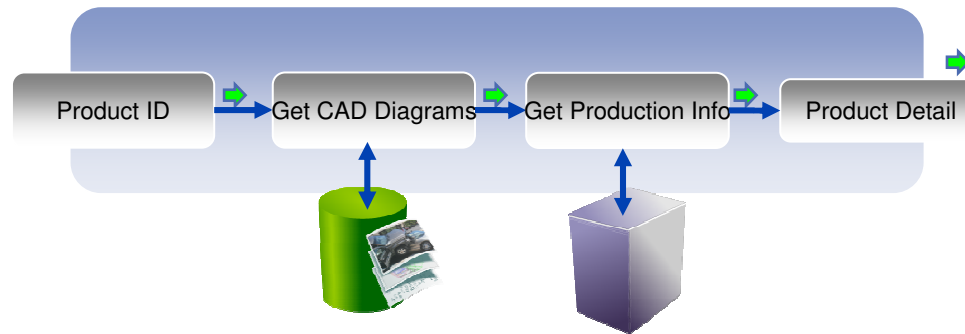
– Gartner Research



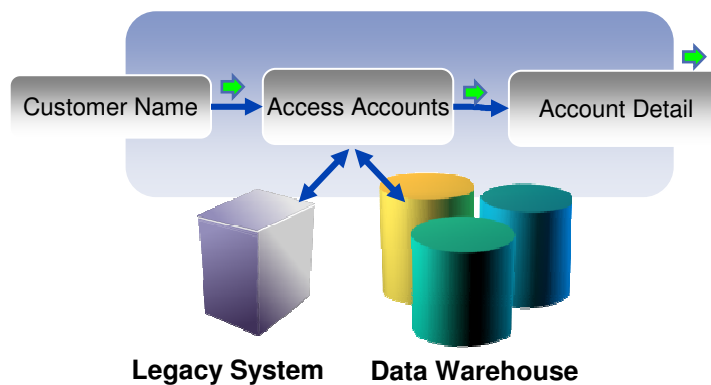
TAKE BACK CONTROL

SOA – Data Service Providers

Data Access Services – Native DB2



Federation Services – Virtualize information from multiple sources



Service-Enabling Mainframe Data

Case Study – Finance, Regulatory Compliance

▶ Challenges:

- Improve risk management across all member institutions
- Meet Basel II compliance deadline
- Access information:
 - in 23 different retail systems
 - from over 2500 branch offices

▶ Solution:

WebSphere Information Integration Platform providing single point access and control of risk-bearing information across many different mainframe systems in different technologies and formats.

▶ Result:

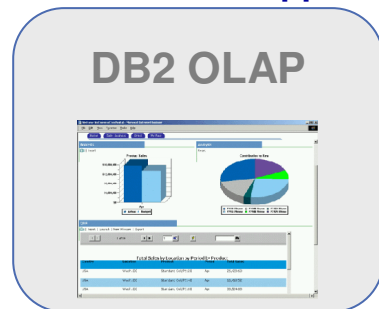
- Enhanced risk management and increased efficiency of data collection for Base II required data
- Ability to view data in operational systems spread across the enterprise including third party information without disrupting retail system



Service-Enabling Heterogeneous Data

Finance: Regulatory Compliance

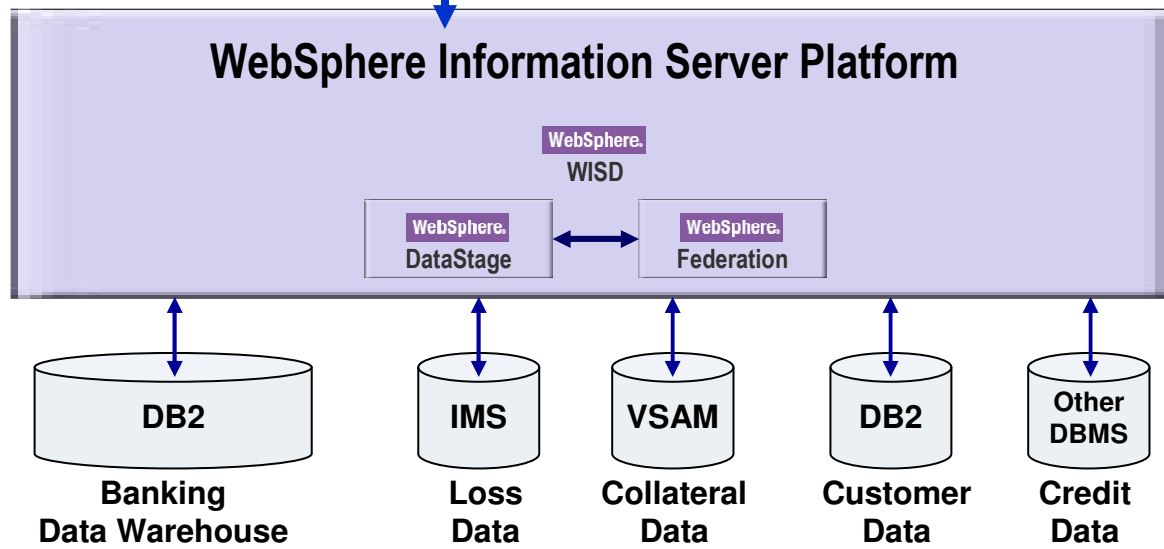
Risk measurement application



Solution – Basel II Exception Monitoring

- ETL packaged as a service provides a single point for access and control of risk-bearing information
- Spans many different mainframe systems
- Federation combines data stored in different technologies & formats

Web Service Interface



SOA – Advanced Service Providers

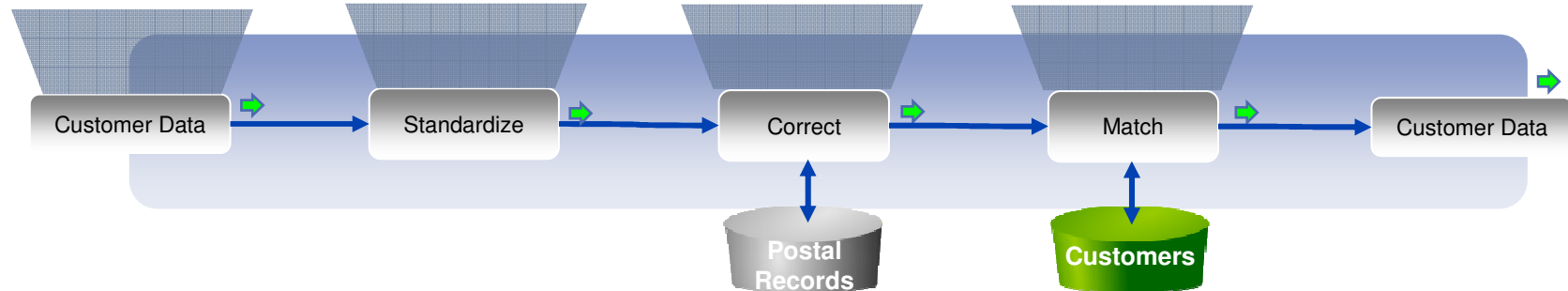
Cleansing Services – Package cleansing job for reuse

Kate A. Roberts
416 Columbus Street #2
Boston, Mass

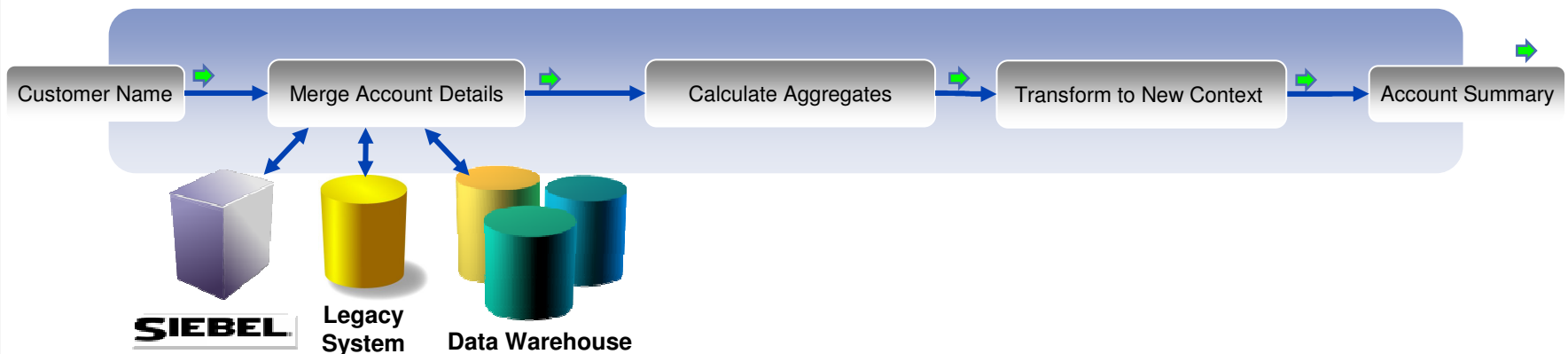
Kate A. Roberts
416 Columbus St #2
Boston, MA

Kate A. Roberts
416 Columbus Ave #2
Boston, MA 02116

Catherine A. Roberts
416 Columbus Ave. Apt. 2
Boston, MA 02116



Transformation Services – Batch and real-time transformations



Service-Enabling Mainframe Data

Case Study – US Manufacturer of Recreational Equipment

▶ Challenges:

- Complex inventory environment is dependent on Japanese parent for parts
 - on-the-boat, in-port, held-in-port, in-US-warehouse, ...
- Manual review of reports needed to provide single view of inventory to Finance, Manufacturing, Sales, etc. --- can take up to 3 weeks!
- Maintaining excessive inventory (high cost)
- Missing product delivery dates (lost revenue)

▶ Solution:

1. Surface inventory "delays" as services
2. Inventory services feed downstream systems (Finance, SCM, ...)
3. Monitoring applications leverage services

▶ Result:

- Reduced inventory overhead
- More efficient use of inventory, accelerating delivery of customer orders
- Consistent, accurate, up-to-date view of inventory for Finance
- Eliminated manual reconciliation – reduced manpower



Service-Enabling Mainframe Data

Case Study – Insurance

▶ Challenges:

- Chronic shortfalls in productivity and customer satisfaction targets
 - Difficult, if not impossible to get correct customer-level information
 - Detailed information available at contract level only
 - Frequent conflicting information at group level
 - Major cause of billing errors and disputes

▶ Solution:

1. Best-of-breed data attributes identified and surfaced via Services
2. Information silos communicate with new "enterprise" applications
3. Complexity of IT environment is "hidden" from business users

▶ Result:

- Enterprise service applications quickly delivered without disrupting existing functional "silo" environments
- Up-to-date information reducing errors, disputes and improving service levels
- Productivity and customer satisfaction steadily improving



Agenda

PART I: Information on demand and data delivery

- Delivering Information
 - Information Integration Patterns
 - Sample Implementations
 - The Role of Replication
- Information in a Service Oriented Application (SOA) Environment
- ***Next Generation Data Warehousing***

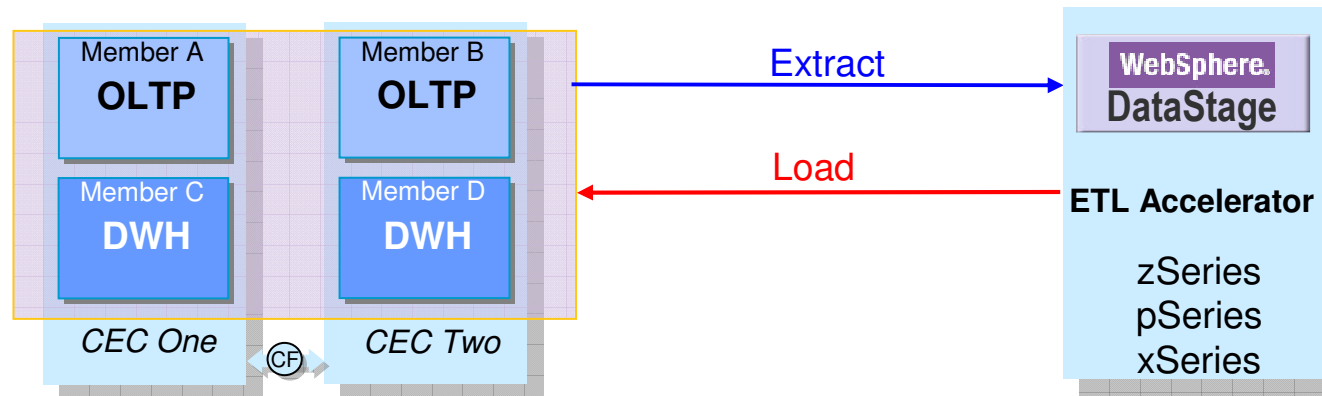
PART II: The IBM Information Server

- The demand for Information Integration
- A platform approach
- Technology
- Information Server for System z
- Where do you go from here?



Managing the Dynamic Data Warehouse Content

ETL Solution Accelerates with Automation & Reuse



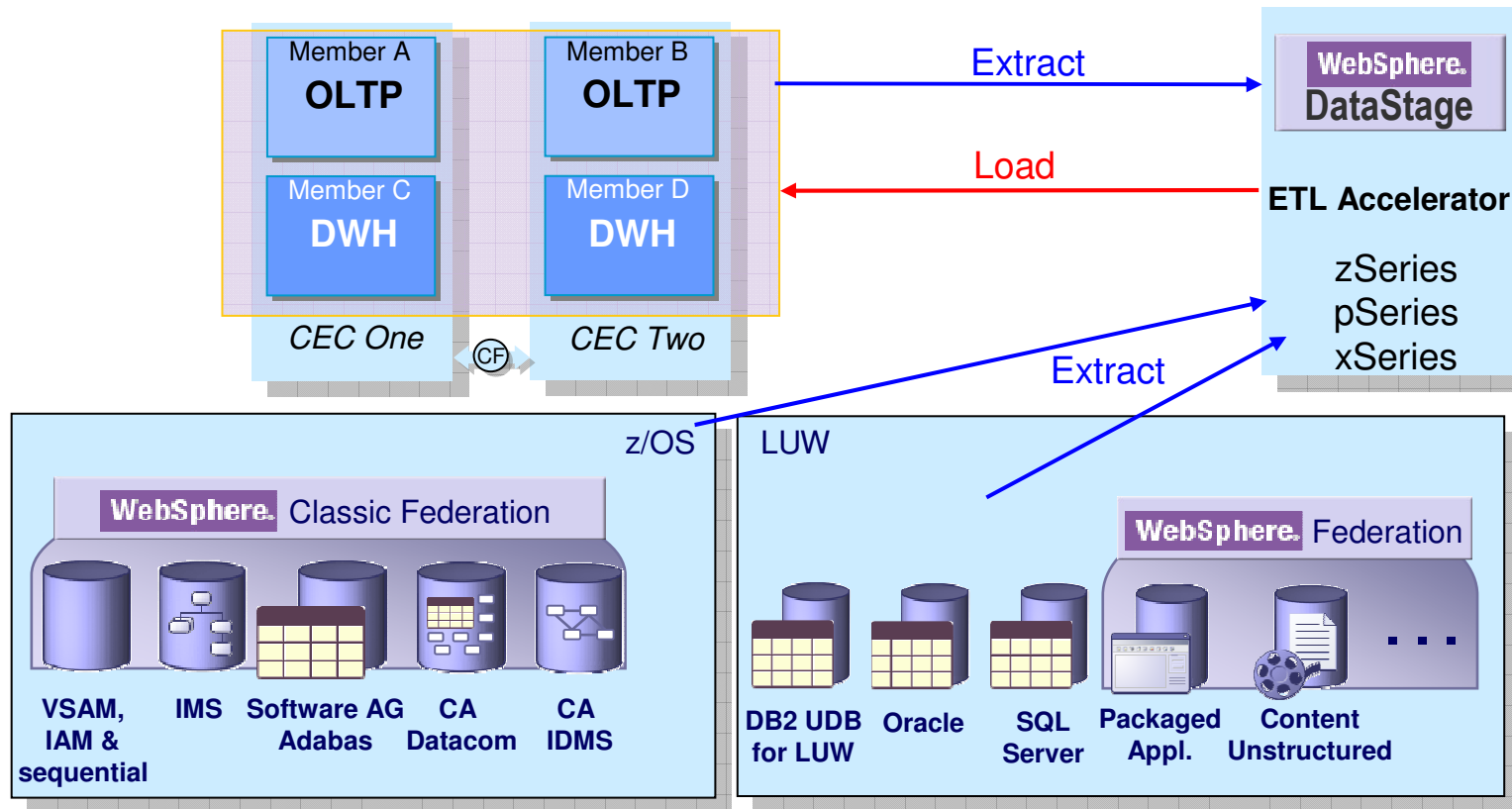
ETL is done using an ETL Accelerator like **WebSphere DataStage**.

- **E** – the data is extracted from the OLTP information sources
- **T** – 100s of embedded transforms & custom built transforms are applied
- **L** – Aggregated and transformed data is loaded into the warehouse tables



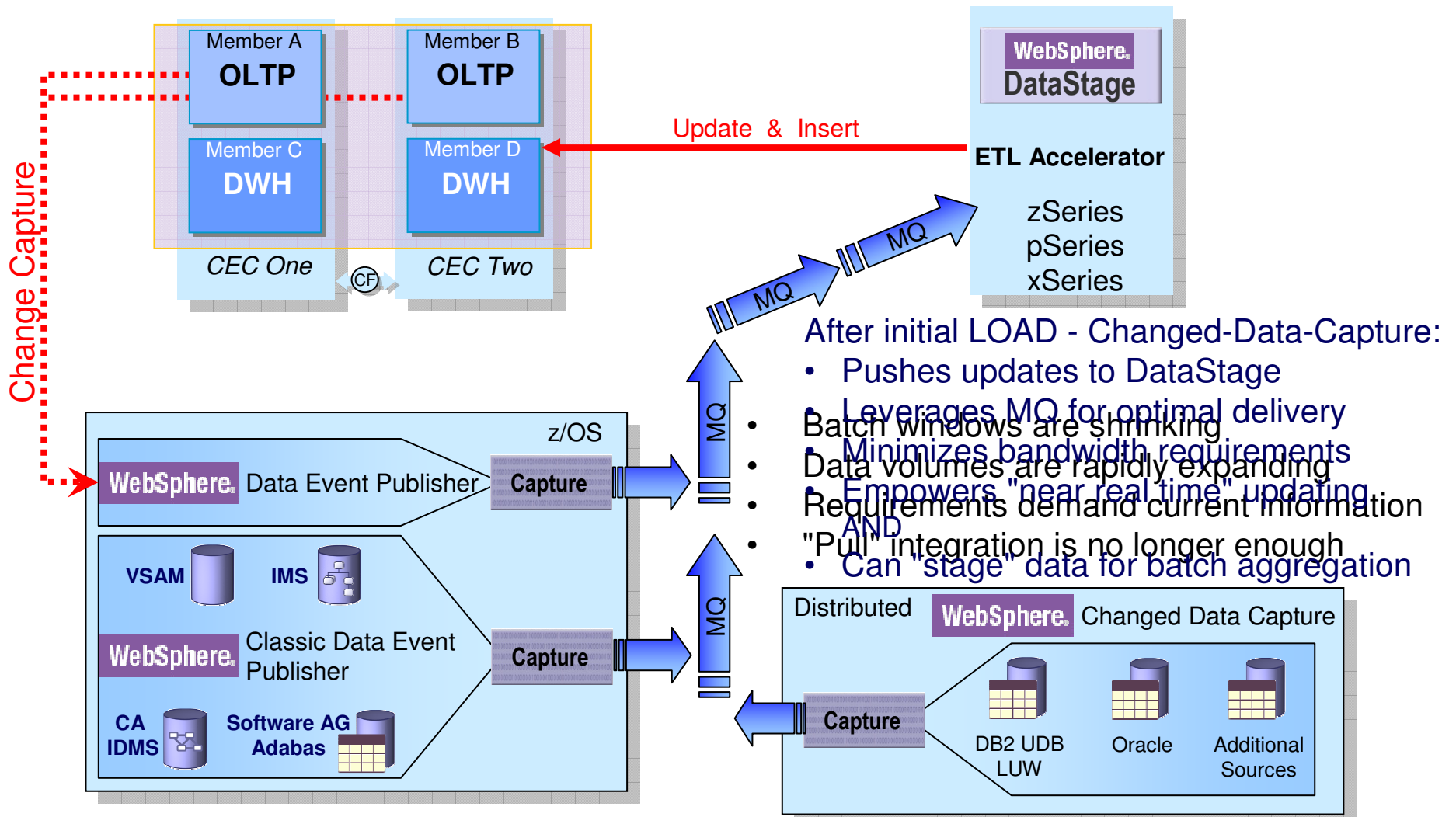
Managing the Dynamic Data Warehouse Content

Native Access and Federation for Universal Data Inclusion



Maintenance of the Dynamic Data Warehouse

Incremental Updating is Essential



After initial LOAD - Changed-Data-Capture:

- Pushes updates to DataStage
- Leverages MQ for optimal delivery
- Batch windows are shrinking
- Minimizes bandwidth requirements
- Data volumes are rapidly expanding
- Empowers "near real time" updating
- Requirements demand current information
- AND
- "Pull" integration is no longer enough
- Can "stage" data for batch aggregation



Agenda

PART I: Information on demand and data delivery

- The Integration Landscape
 - Information Integration Patterns
 - Sample Implementations
 - The Role of Replication
- Information in a Service Oriented Application (SOA) Environment
- 3rd Generation Data Warehousing

PART II: The IBM Information Server

- The demand for Information Integration
- A platform approach
- Technology
- Information Server for System z
- Where do you go from here?



IBM Information on Demand

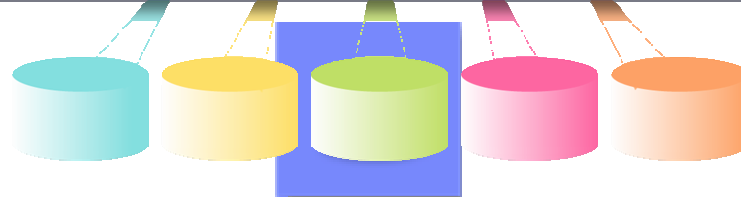
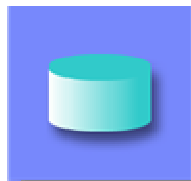
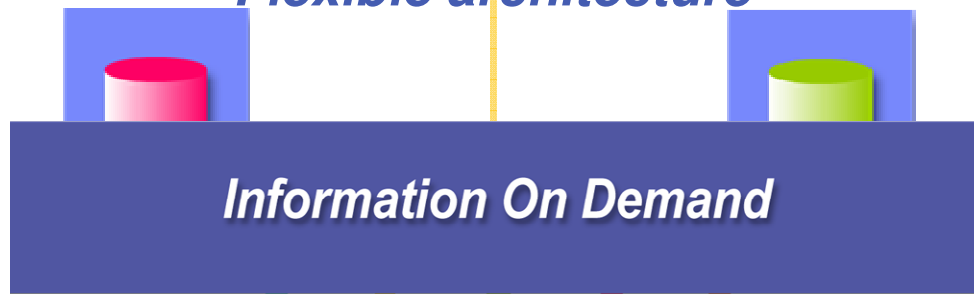
Delivering trusted information in context to optimize business processes, applications and productivity in real time

People Applications & Processes

Applications Applications Applications Applications Applications



**Trusted Information
In Context**
Flexible architecture



Rich media Leverage heterogeneous applications and information sources Documents E-mail

Protect & extend existing investments



TAKE BACK CONTROL

Benefiting from "z" Information on Demand

Addressing Key Customer Requirements

- Workload consolidation
- Integrating data via SOA
- Reduce total cost of ownership
- Next generation warehousing



Dynamic Data Warehousing



- Co-located OLTP & DW
- Move to System z decreased reporting times 80%

Master Data Management



- consolidated customer profile to enable customer centric processing for sales and service

SOA Data Serving



- Exposed legacy IMS through Web services
- Guaranteed integrity on over 4 billion IMS transactions per month

SAP Information Infrastructure



- Significant IT savings
- Improved customer service
- Business growth

Data Governance & Accountability



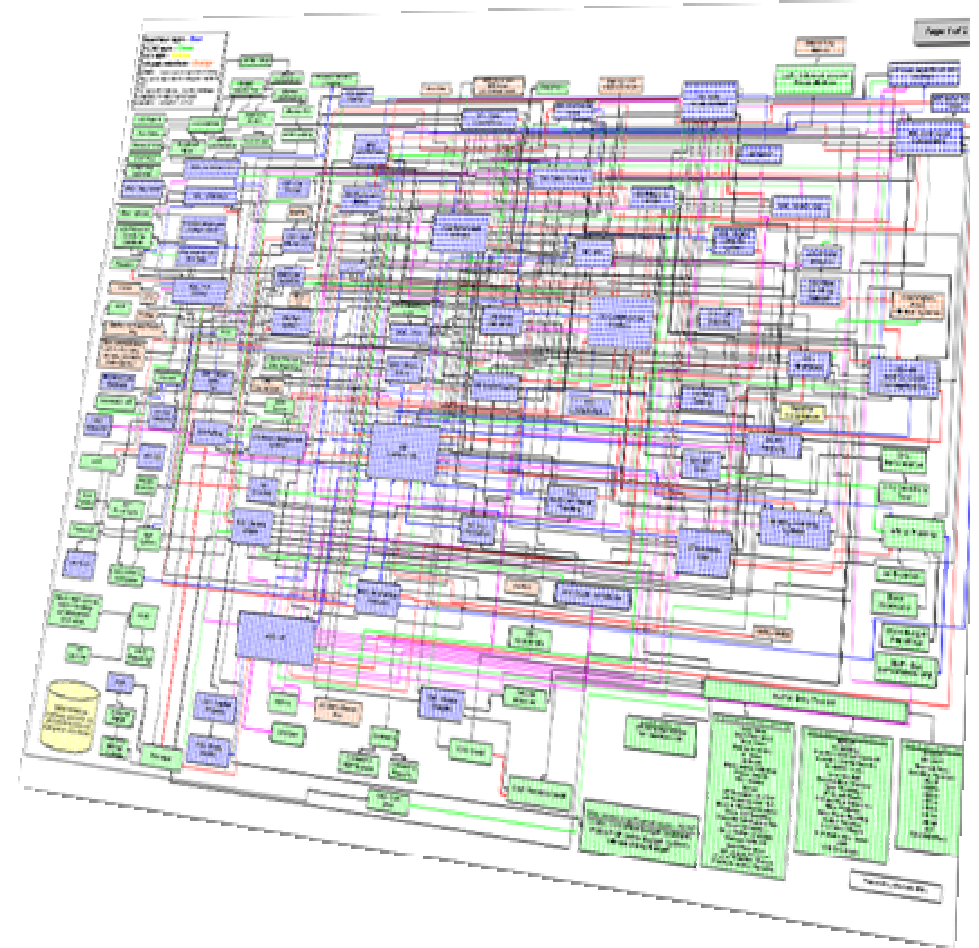
- Automated archiving to manage continuous data growth



TAKE BACK CONTROL

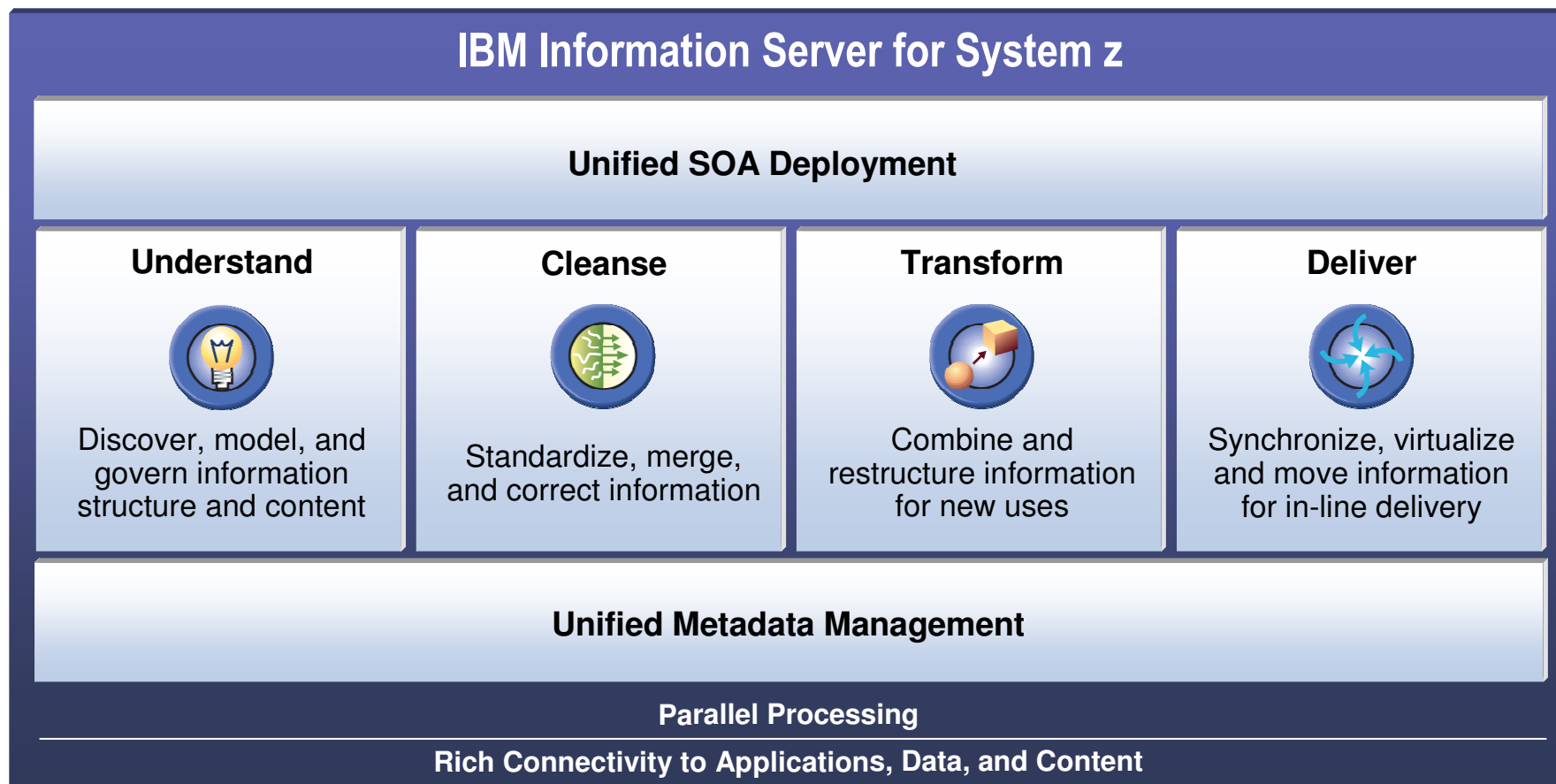
The IBM Information Server

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



IBM Information Server for System z

Delivering information you can trust



TAKE BACK CONTROL

IBM Information Server for System z

Delivering information you can trust

Understand your data and the terms that describe it to ensure consistency, trust and effective communication between the business user and the technology user

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

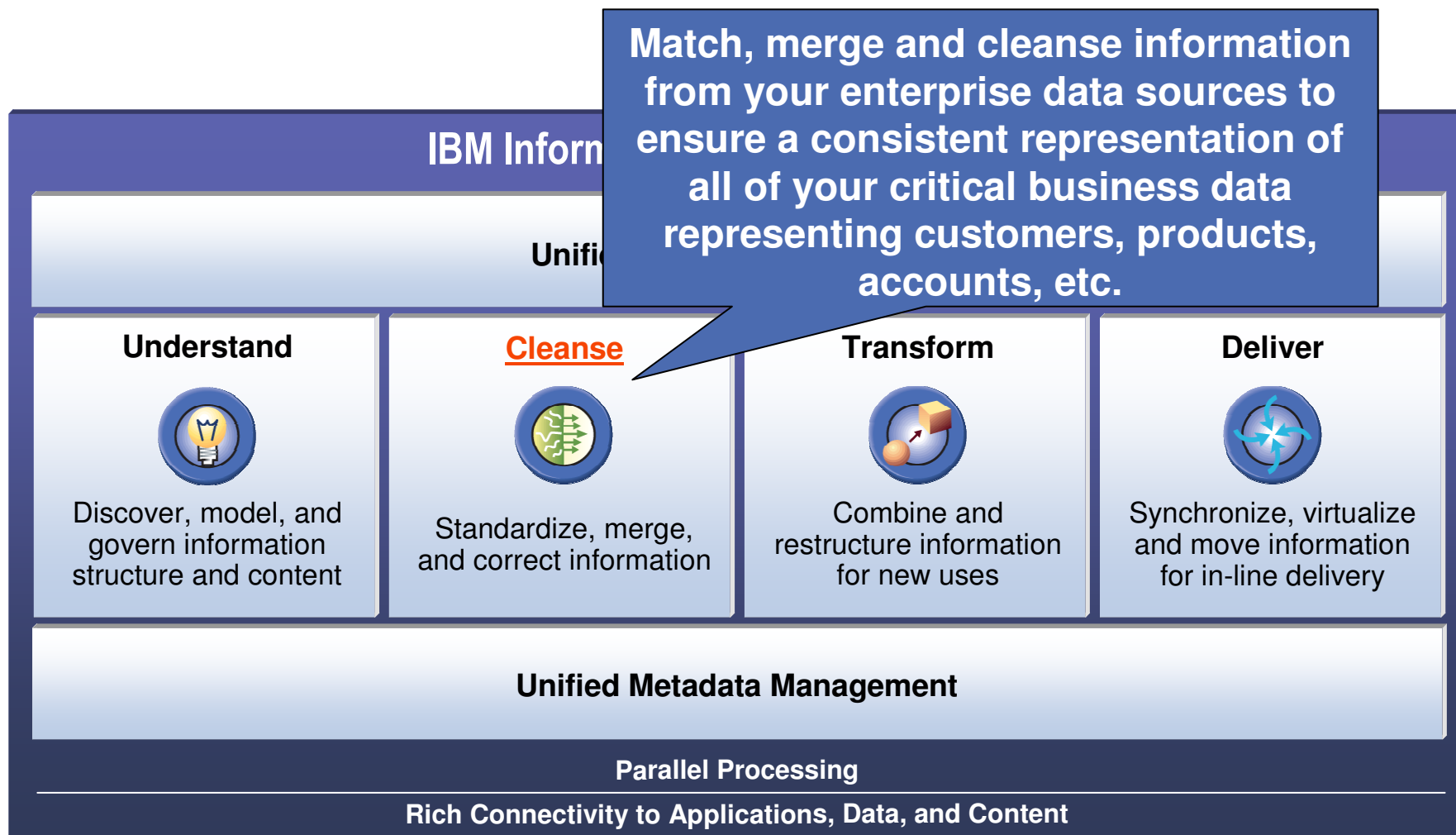
Rich Connectivity to Applications, Data, and Content



TAKE BACK CONTROL

IBM Information Server for System z

Delivering information you can trust

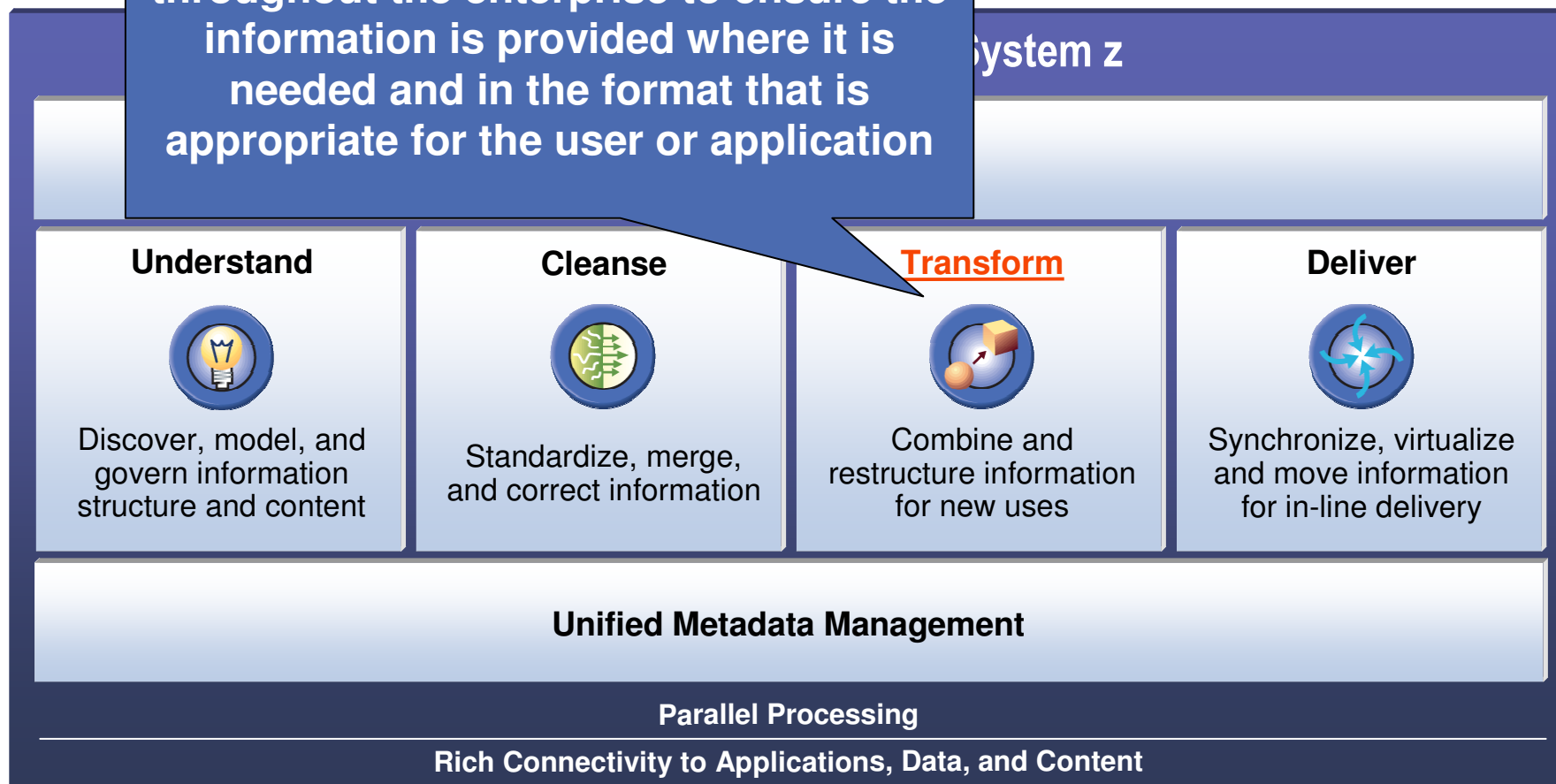


TAKE BACK CONTROL

IBM Information Server for System z

Delivering information you can trust

Repackage and repurpose information throughout the enterprise to ensure the information is provided where it is needed and in the format that is appropriate for the user or application



TAKE BACK CONTROL

IBM Information Server for System z

Delivering information you can trust

Access and/or deliver information when and where needed, ensuring all information is an equal participant ... whether structured or unstructured, distributed or mainframe-based, from within the enterprise or from without

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

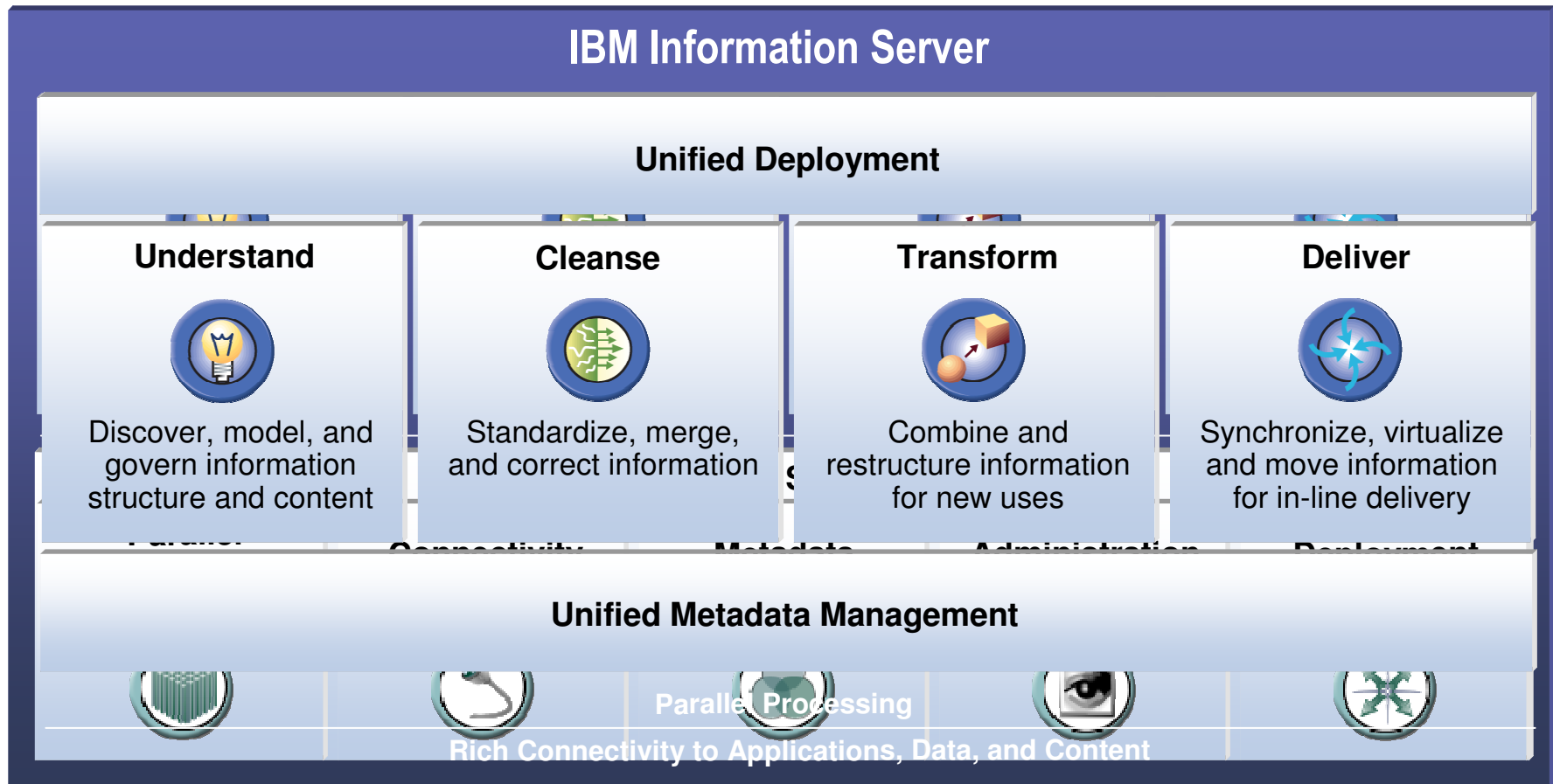
Rich Connectivity to Applications, Data, and Content



TAKE BACK CONTROL

IBM Information Server for System z

The power of a platform ... not just a collection of tools



TAKE BACK CONTROL

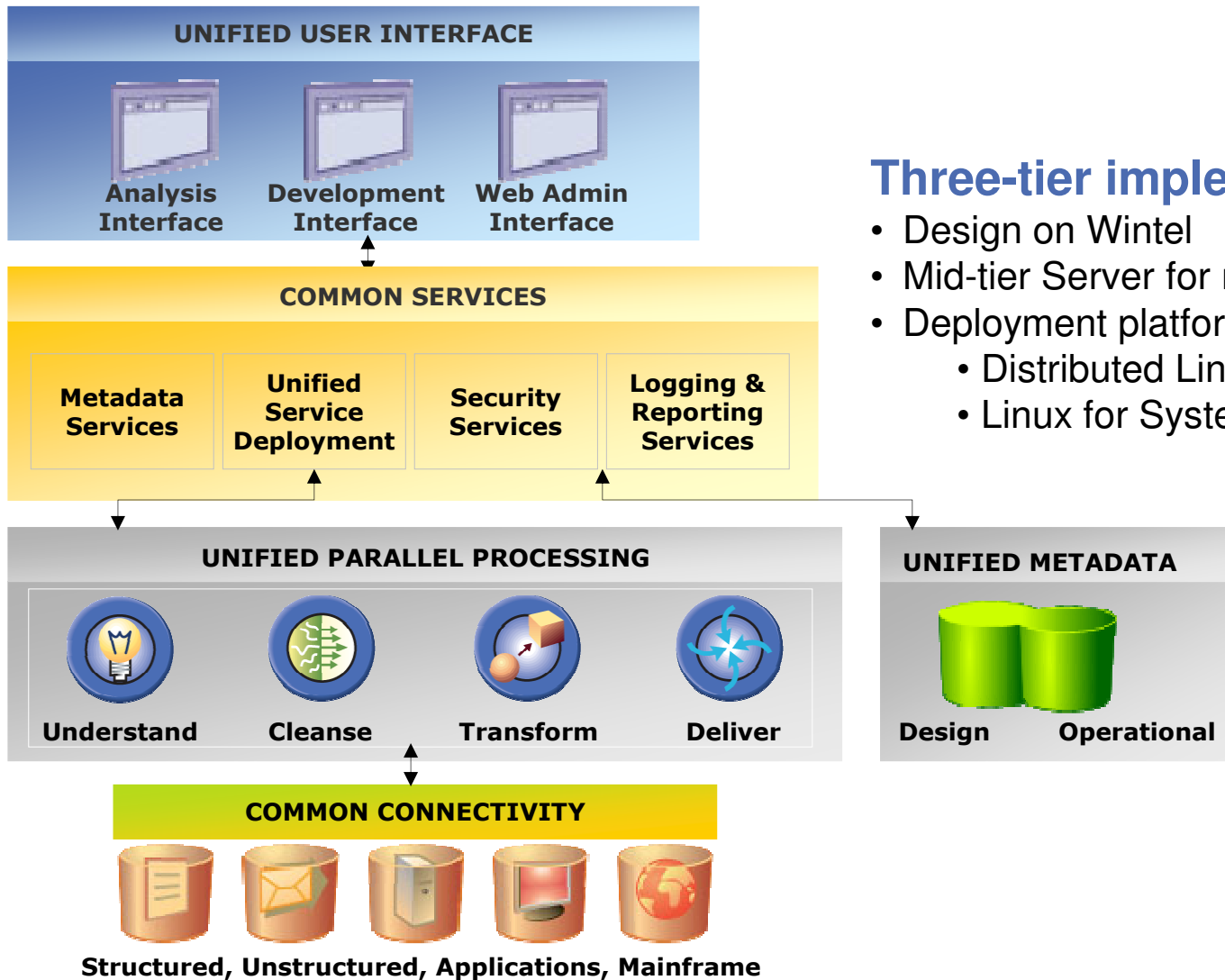
IBM Information Server

A Platform and a set of Products

- IBM Information Server is a *platform*
 - Multiple modules that share a consistent foundation of shared platform services
 - The platform approach increases reusability and creates synergy between and across functional modules
- Companies can purchase one or more of the modules
 - Each module inherently includes the platform services
 - Each module addresses a requirement and provides value on its own
- The modules work together seamlessly, providing increased project leverage
 - As new modules are added, they “snap-in” to the shared platform services, providing immediately leverage
 - Organizations have a flexible entry point ...
Think Strategically, Implement for Immediate Impact



IBM Information Server Architecture

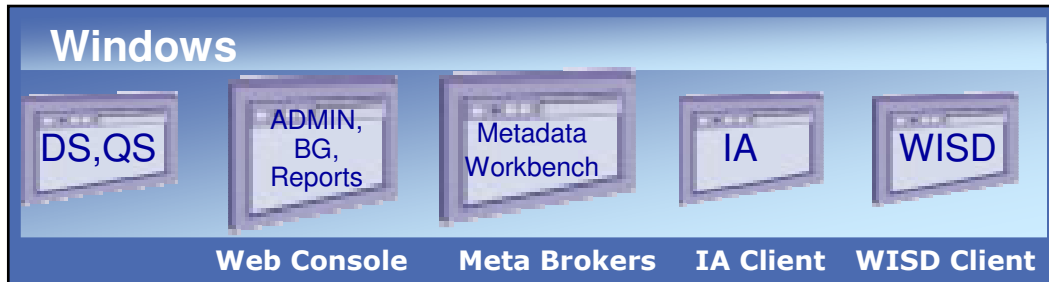


Three-tier implementation:

- Design on Wintel
- Mid-tier Server for metadata & scheduling
- Deployment platforms are:
 - Distributed Linux, Unix and Windows
 - Linux for System z



IBM Information Server for System z

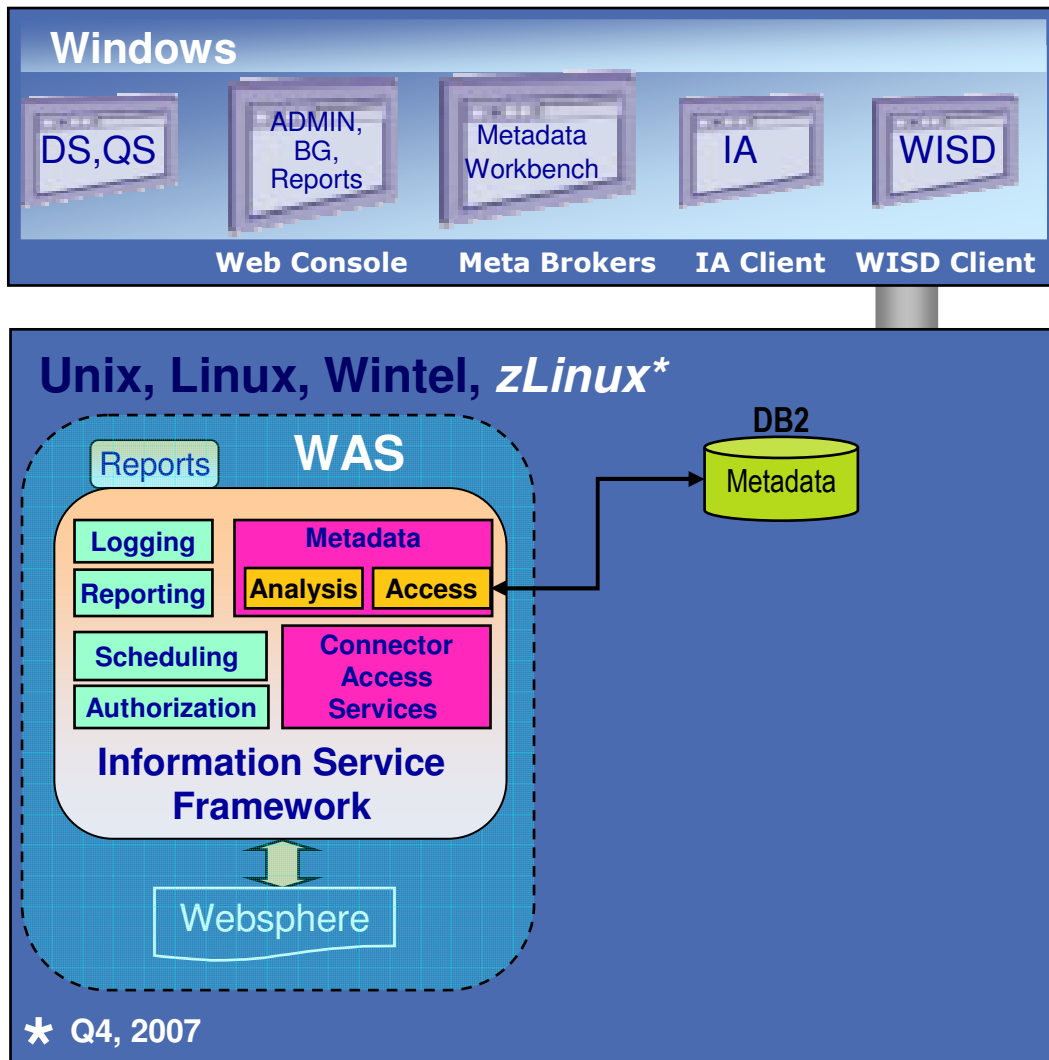


Roles-based GUI Design Tools work the way "you" do



TAKE BACK CONTROL

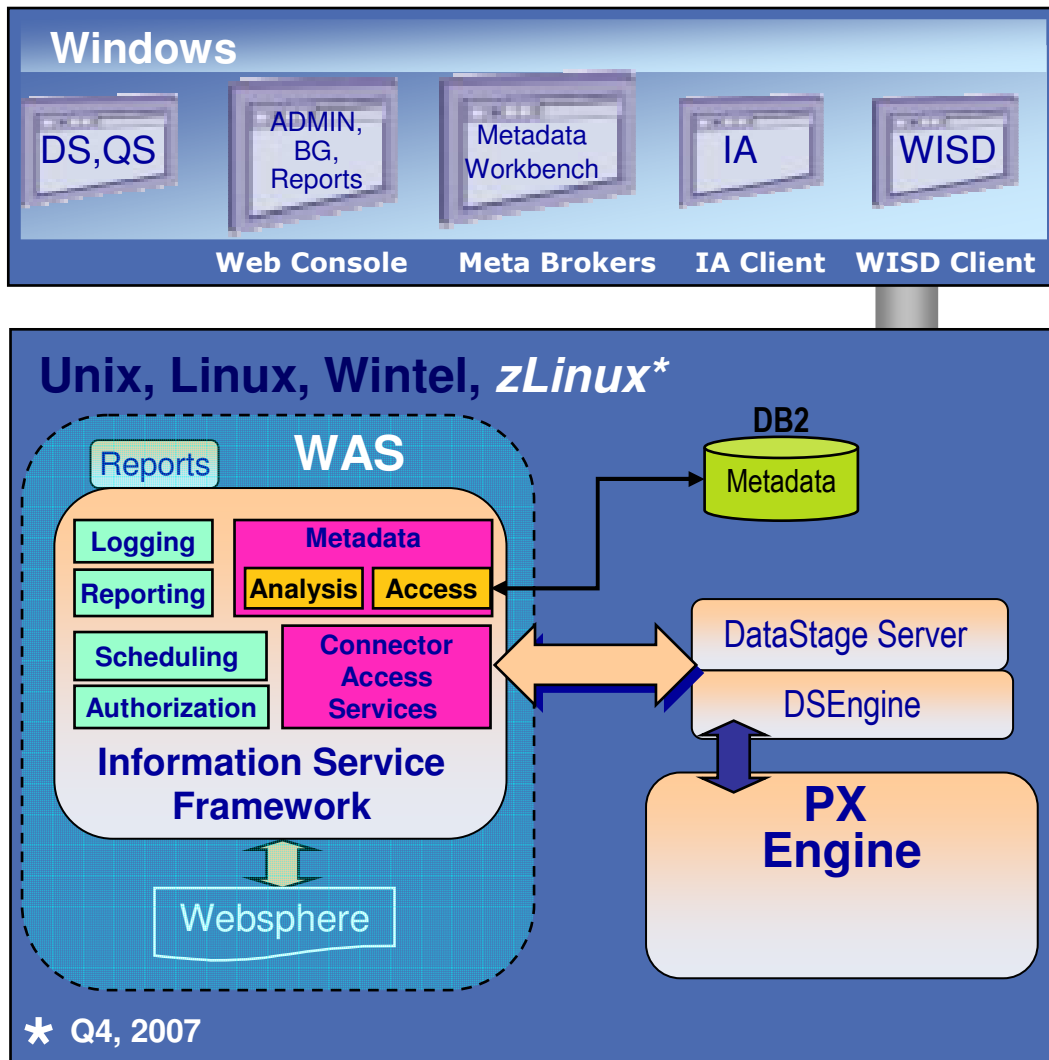
IBM Information Server for System z



- **Common reusable services** framework leverages the power of a SOA environment
- **Meta data repository** promotes:
 - reuse
 - compliance to standards
 - visual lineage
 - impact analysis



IBM Information Server for System z



New Linux for z deployment option

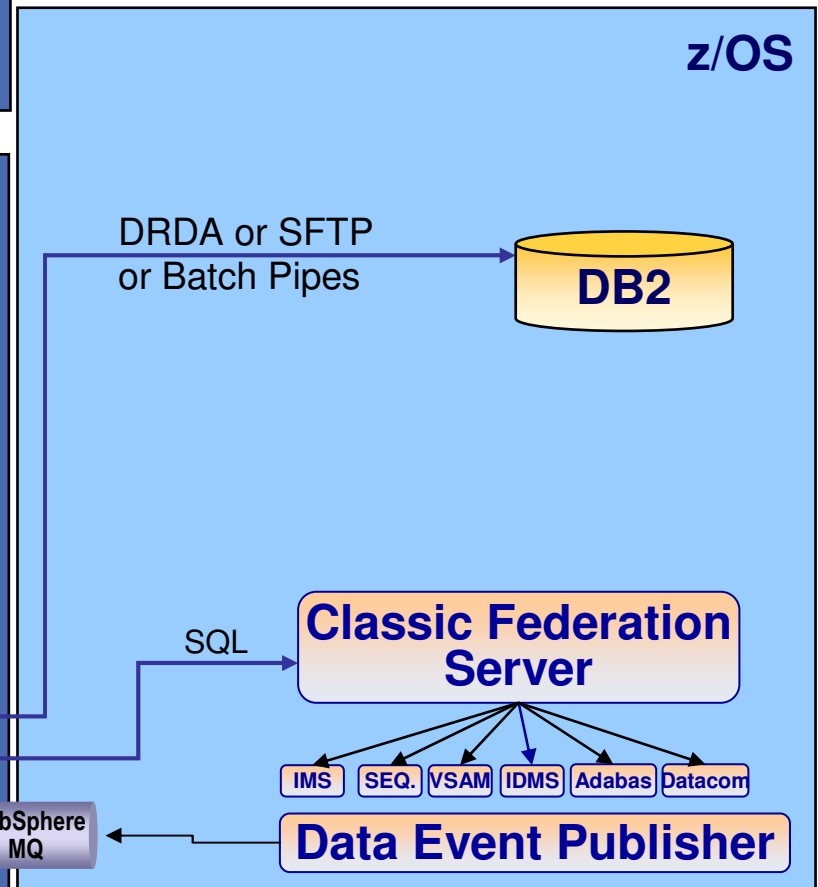
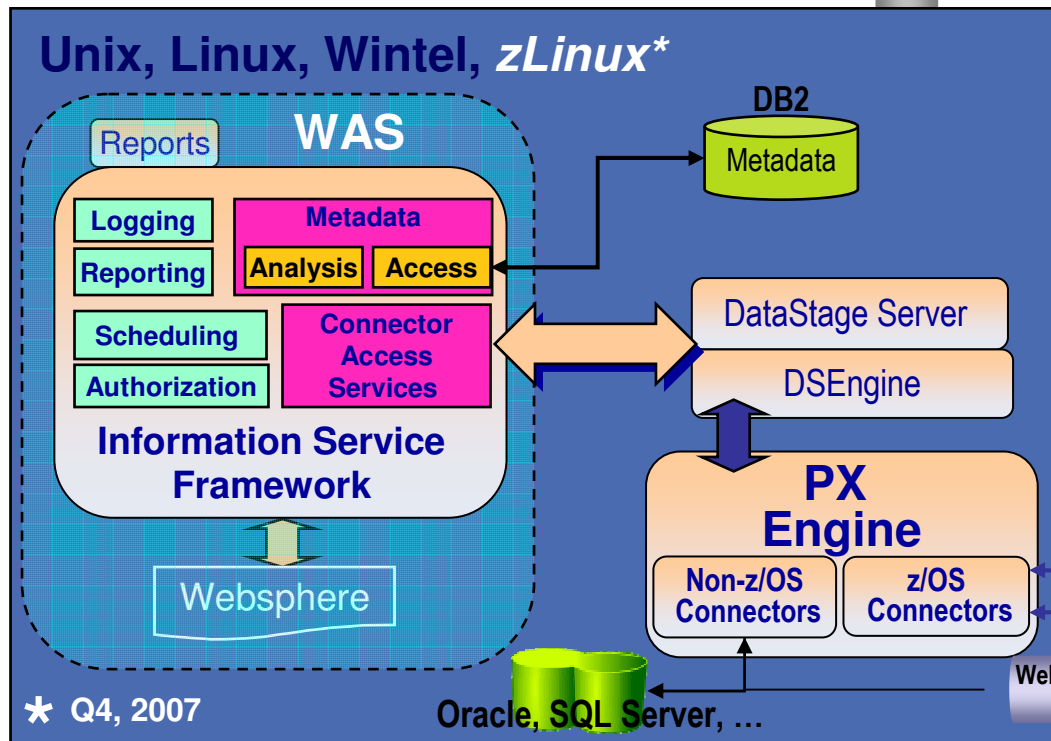
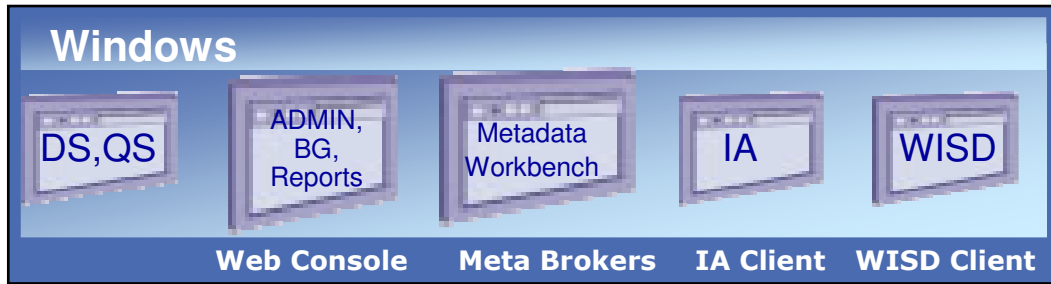
- Robust, parallel processing
- Hipersocket connectivity to z data
- Full Information Server suite: QualityStage, Information Analyzer...
- Minimal impact on z/OS costs: Leverages IFLs and zIIPs



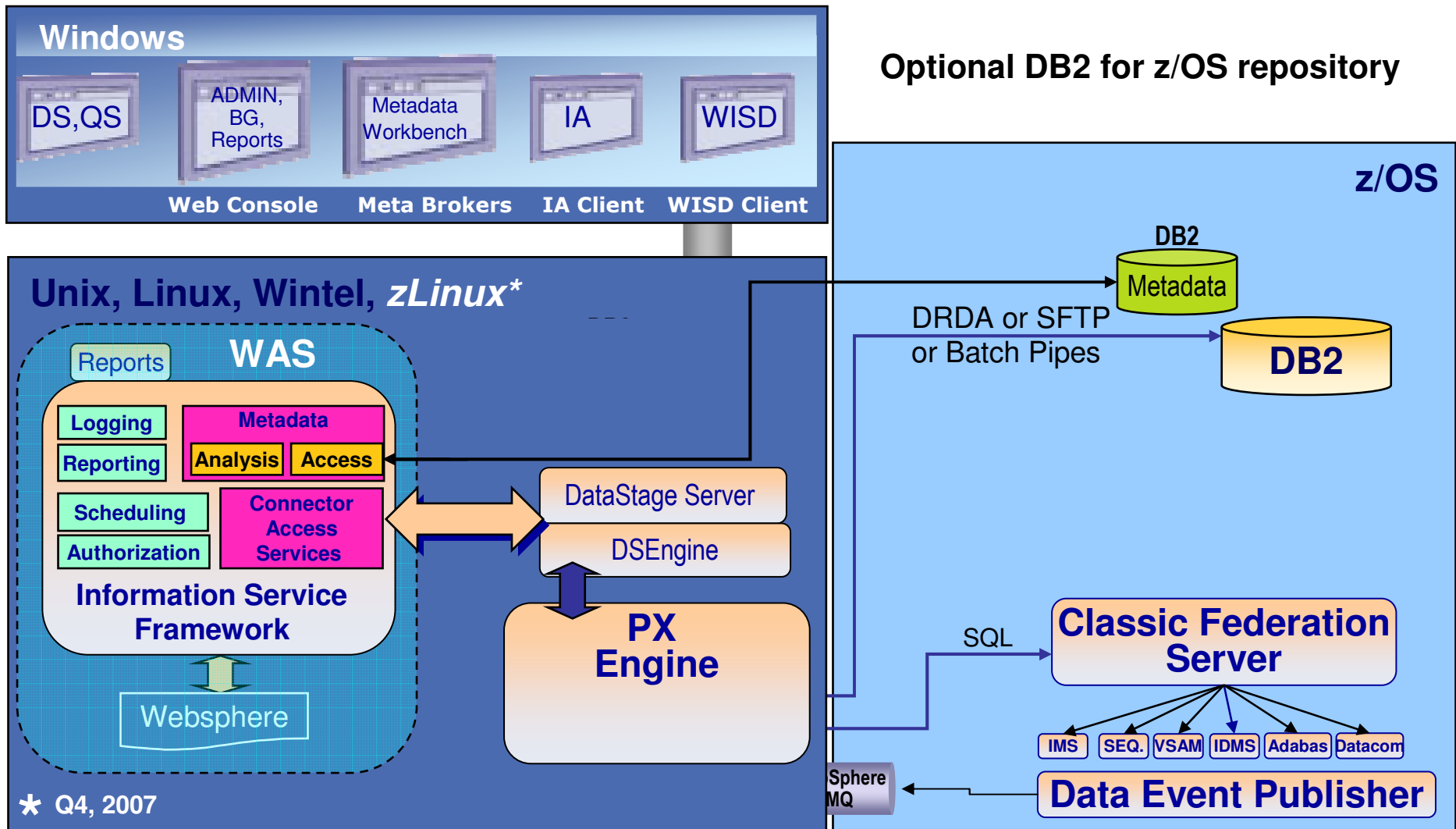
IBM Information Server for System z

Rich connectivity with:

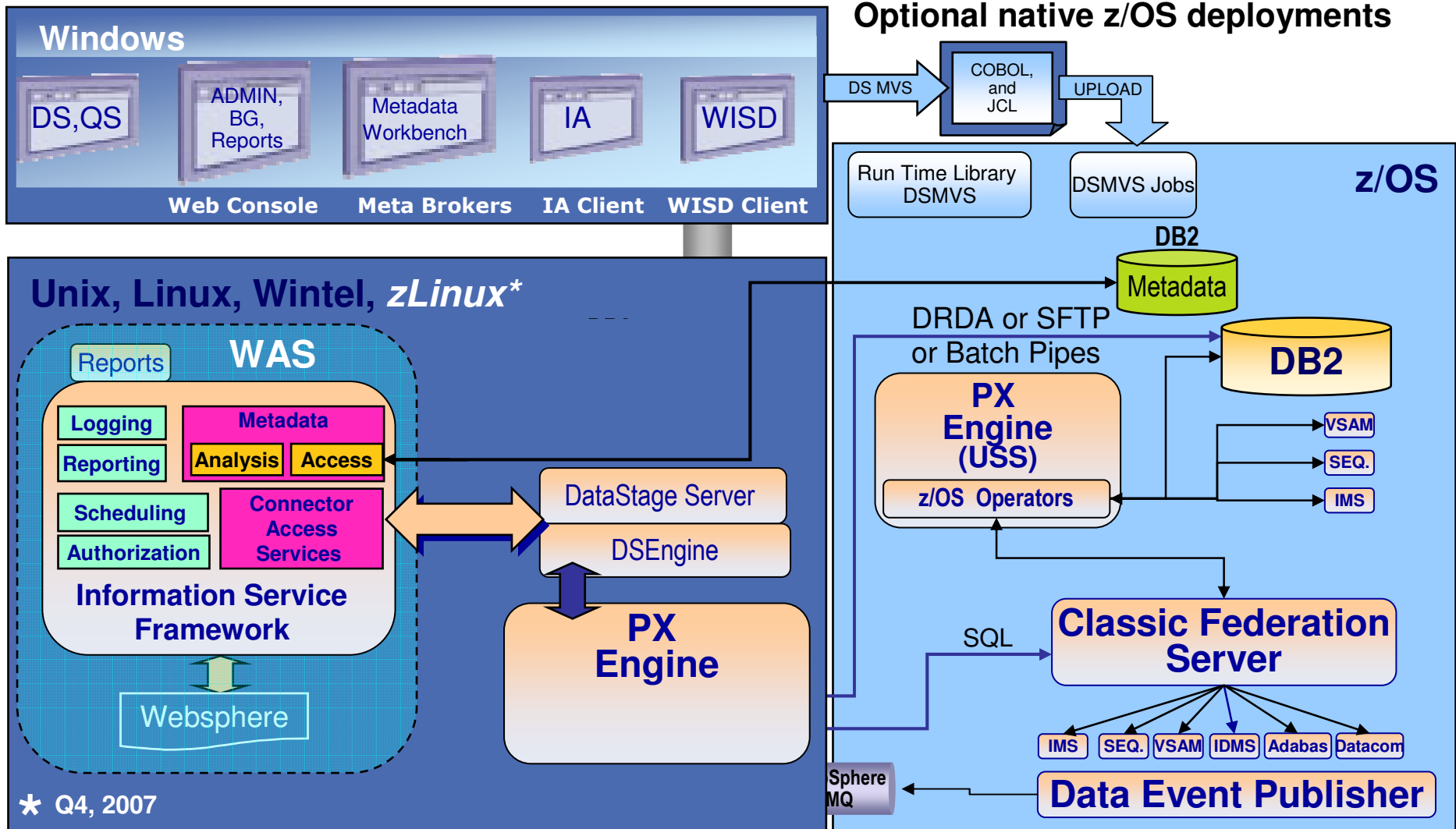
- Application & distributed data
- DB2 for z/OS
- Classic Federation for more z-data
- Data Event Publishers for CDC



IBM Information Server for System z



IBM Information Server for System z



IBM Information Server on System z

Benefits of a hybrid architecture that leverages zLinux

- **Significant cost savings:**
 - MIPs charged at IFL rate ... NOT z/OS rate
 - Minimizes impact on z/OS MIPs
 - All Job Processing is on zLinux, except the z/OS data access
 - DB2 access qualifies for zIIP specialty engine
 - Minimizes impact on other z/OS software costs
- **High performance z data connectivity:**
 - Batch Pipes for DB2 load, DRDA to DB2 over hipersockets
 - SQL to Classic over hipersockets
 - Integration with Data Event Publishers for changed-data-capture
- **Seamless integration with other Information Server platforms**
 - Same operational architecture and meta data Repository
 - Eliminates deployment issues
 - Maintains value of DataStage for z/OS investments



The Bottom Line

IBM provides a platform for . . .

- Gaining access to data at any time, wherever it resides
- Making sure it is complete, correct and up to date, and
- Moving it from any file, database or application to any other destination:
 - Another database
 - Another application
 - Another platform

Implications for System z ...

- System z data must be an equal participant
 - ~60% of enterprise operational data is mainframe based
- System z is a robust, reliable, secure implementation platform for this workload
 - Especially for DB2 on z users who will resist taking mainframe data off platform
- TCO for this **new workload** on z is very attractive ... especially for large IT installations
 - Specialty engines on z (IFLs, zIIPs, zAAPs) provide lower cost MIPs
 - Green messages apply
 - Server consolidation messages apply





Summary

TAKE BACK CONTROL



IBM can help with an *Information Integration Assessment*

On-site workshop for your team providing:

- Information Integration needs assessment
- Report and recommendations

What is it?

An **interactive** session led by an IBM Information Integration Solutions specialist, to help you take advantage of emerging technologies, standards, and current trends in information integration

What will you get as a result?

- Outline of your current information integration environment
- Identification of key projects where information integration could improve IT productivity and speed development
- Exploration of alternative approaches
- Recommendations for architecture enhancements



Workshop Basics

Typical Agenda

- Part One:
Information gathering and sharing
 - Information Integration Overview
 - Discovery –
Extensive interviews and data collection
- Part Two:
Analysis, summary, recommendations
 - Summary of information collected
 - Review of:
challenges, issues, problems, needs
 - Approaches considered
 - Tradeoffs
 - Recommendations

Typical Workshop Participants

- Up to 15 team members:
for example:
 - Architects
 - Data architects, information architects
 - Enterprise architects
 - Data architects or information architects
 - LOB managers in areas like
 - Compliance (financial services and other industries)
 - Supply chain management (CPG)
 - Clinical intelligence (healthcare)
 - Application owners
 - Middleware managers
 - Data warehouse managers and developers



Thank

You



TAKE BACK CONTROL