IBM Software

IBM Information Integration

Q-Replication and Event Publishing – Data is getting ACTIVE!

Christian Lenke
IBM Software Group
Information Management & Integration
clenke@de.ibm.com

DB2. Information Management Software

WebSphere software

ON DEMAND BUSINESS™

Juni 2005

Agenda

- **IBM's View of Information Integration**

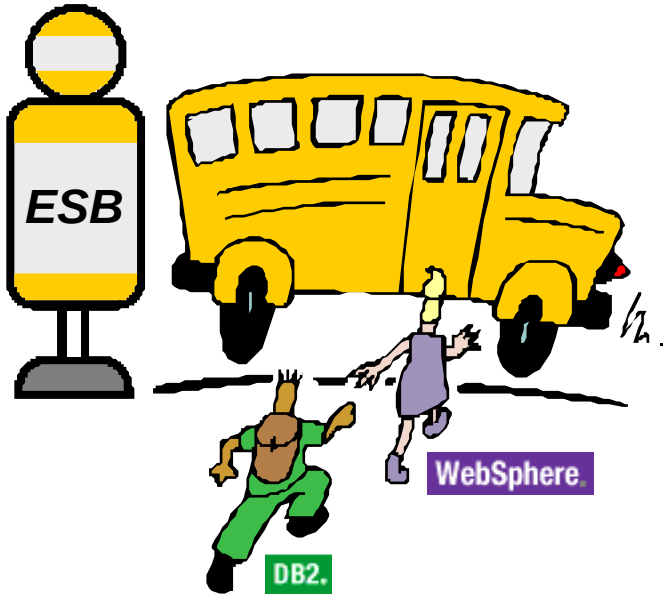
IBM Data Replication Architectures:
Introducing High-Volume and Low-Latency Replication with MQ-based Replication Technologies

Publishing of Data Events:
Integration of Processes, Applications, and Information

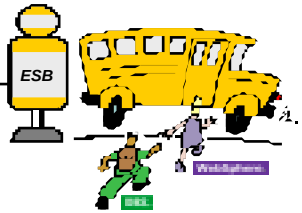
Summary, Collateral, Demo

One Idea of the Enterprise Service Bus: Integrated Information Distribution

Many companies need a high-speed scalable information notification system that can deliver information to the right place, at the right time, in the right format, potentially anywhere in the enterprise and beyond.



Common Usage Examples for Data Replication and Information & Process Integration



Availability

- Scheduled Outage, Failover, Disaster Recovery
- Move Query or Reporting Work to a separate System
- Peer to Peer - Split Workload

Data Distribution & Consolidation

- Move data between Central to Branches, Branches to Central, or Both
- Exchange Data between home-grown and packaged Applications

Data Warehouse & Business Intelligence

- Efficiently move data to new Platform/Database, Transform or Cleanse Data

Mobile Workforce

- Occasionally connected Distribution/Consolidation

Information and Process Integration

- Trigger Applications due to changed data content
- Data and Process Integration across the Enterprise and beyond …

Solution:
WebSphere Information Integrator Family of Products

Integrating diverse Business Information across and beyond the Enterprise

Federation
- Read/Write Access across diverse Data and Content Stores
- Database Programming Model
- Content Programming Model

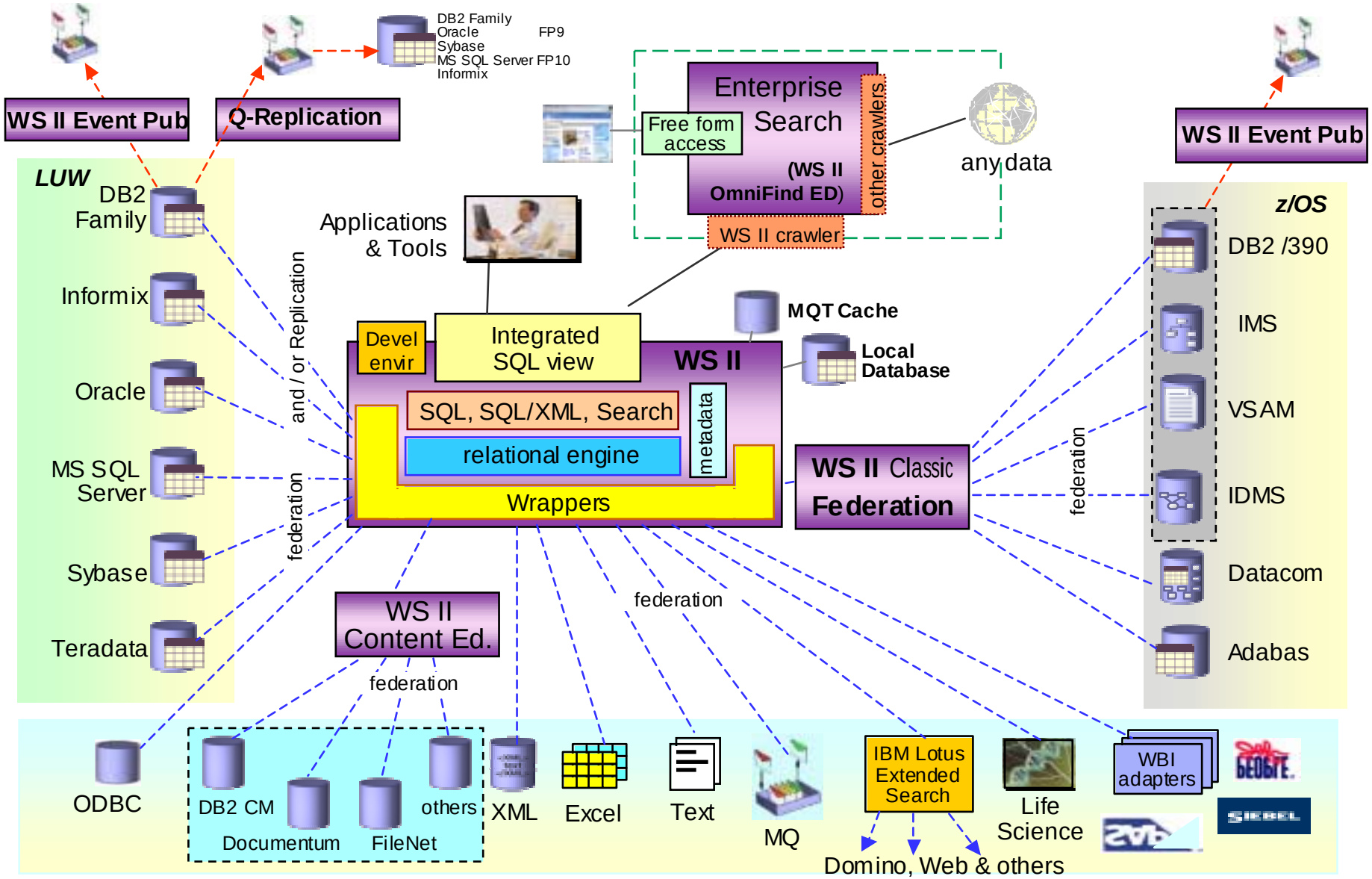
Data Replication and Placement
- Consolidation and Synchronization across complex, multi-vendor IT Environments

Event Publishing
- Capturing of Database Changes
- Formatting of captured Changes into XML Messages
- Publication of Messages into Queues

Enterprise Search
- Indexing and Searching of heterogeneous Enterprise Data

DB2, Informix, Oracle, SQL Server, Sybase, Teradata, XML, OLE DB, ODBC, IMS, VSAM, Message Queues, Web services, Flat Files, Document & Content Repositories, Classic Sources, WWW, Excel, EMail Databases, …

WebSphere Information Integration – *The full Story*



WS II = WebSphere Information Integrator
LUW = Linux / Unix / Windows

Agenda

IBM's View of Information Integration

- **IBM Data Replication Architectures:**
  Introducing High-Volume and Low-Latency Replication with MQ-based Replication Technologies

Publishing of Data Events:
Integration of Processes, Applications, and Information

Summary, Collateral, Demo

IBM's Information Replication Architectures

SQL-Replication (Formally and still known as DataPropagator)

- Log-based asynchronous Change Capture
- Versatile Replication Architecture for Data Replication between all Members of the DB2 Family and beyond (in Combination with WebSphere Information Integrator)
- Relational Staging Concept
- Successfully used for Years by huge User Community

Q-Replication / Event Publishing

- Log-based asynchronous Change Capture
- Captured Transactions immediately sent via Message Queues
- High-Volume, Low-Latency Architecture

HADR (DB2 UDB for Linux, Unix, Windows only)

- Replicates entire Database by Log Buffer Shipping (over IP)
- Initially no Read-Access at Secondary Site

Why Create Another Replication Architecture?

Performance

- Combine high Throughput with low Latency

New Function

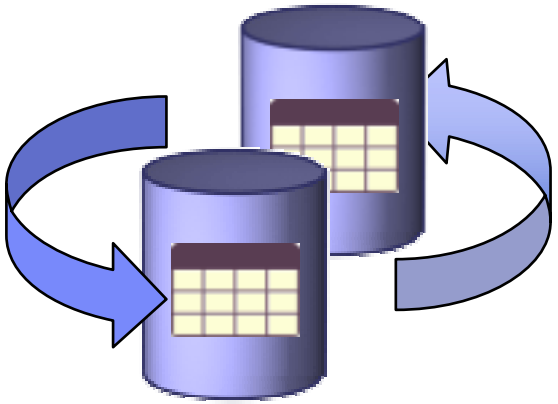
- Event Publishing from DB2 and *Classic* Sources

Capability

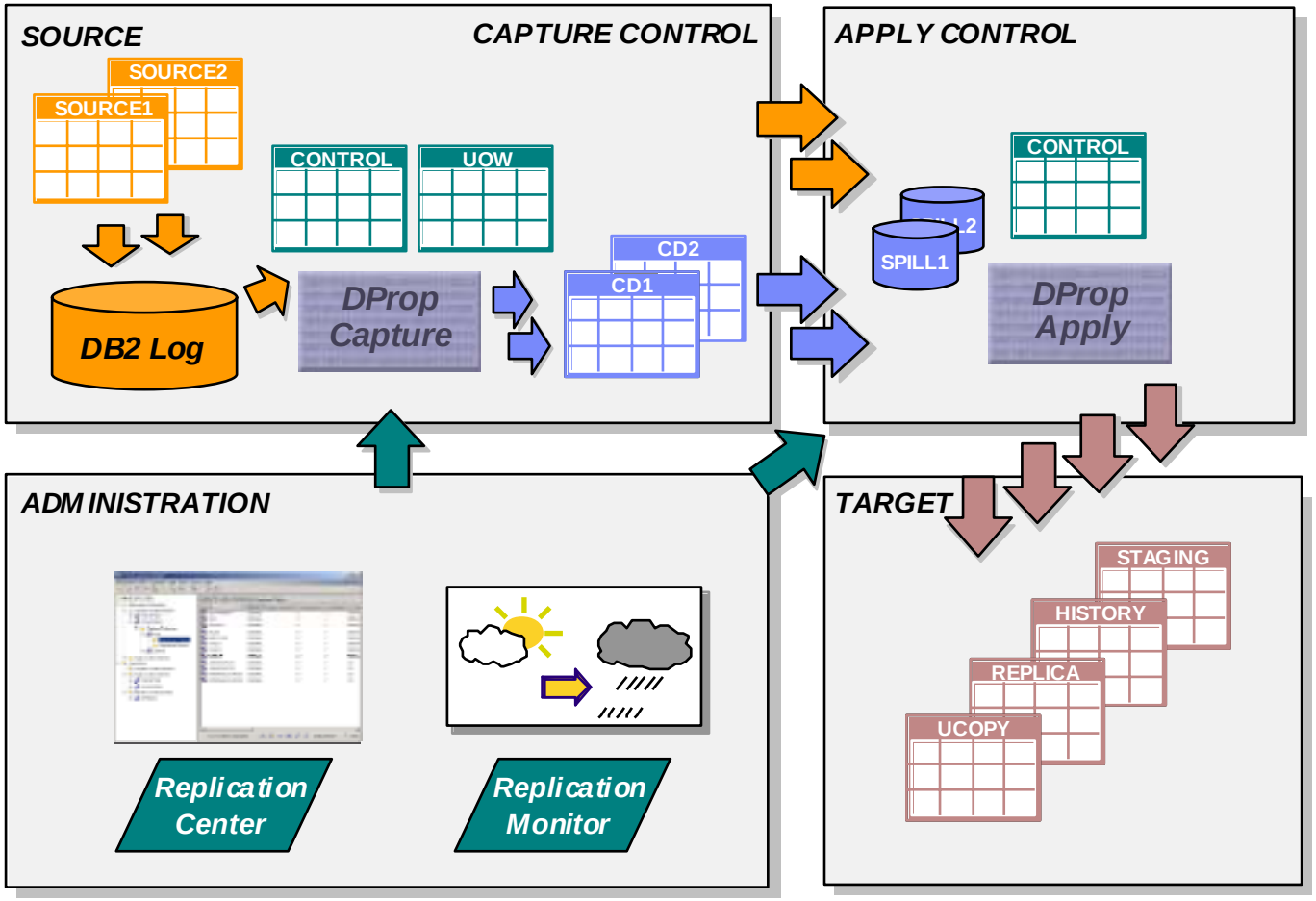
- Significantly improve multi-directional Replication Support

Manageability

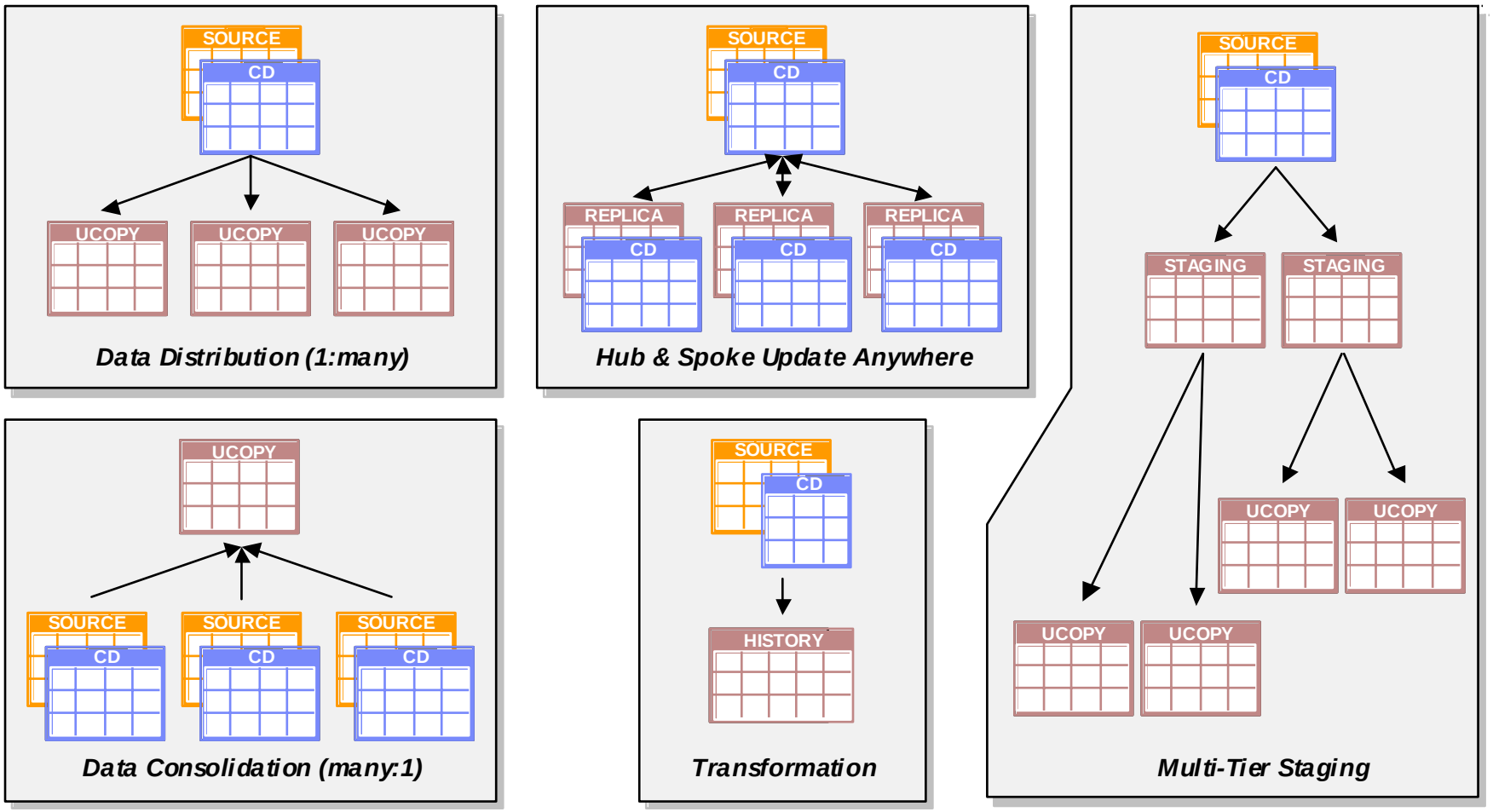
- Reduce the Number of Replication Objects to be defined and managed
- Ease the Definition Process with new Replication Center Wizards



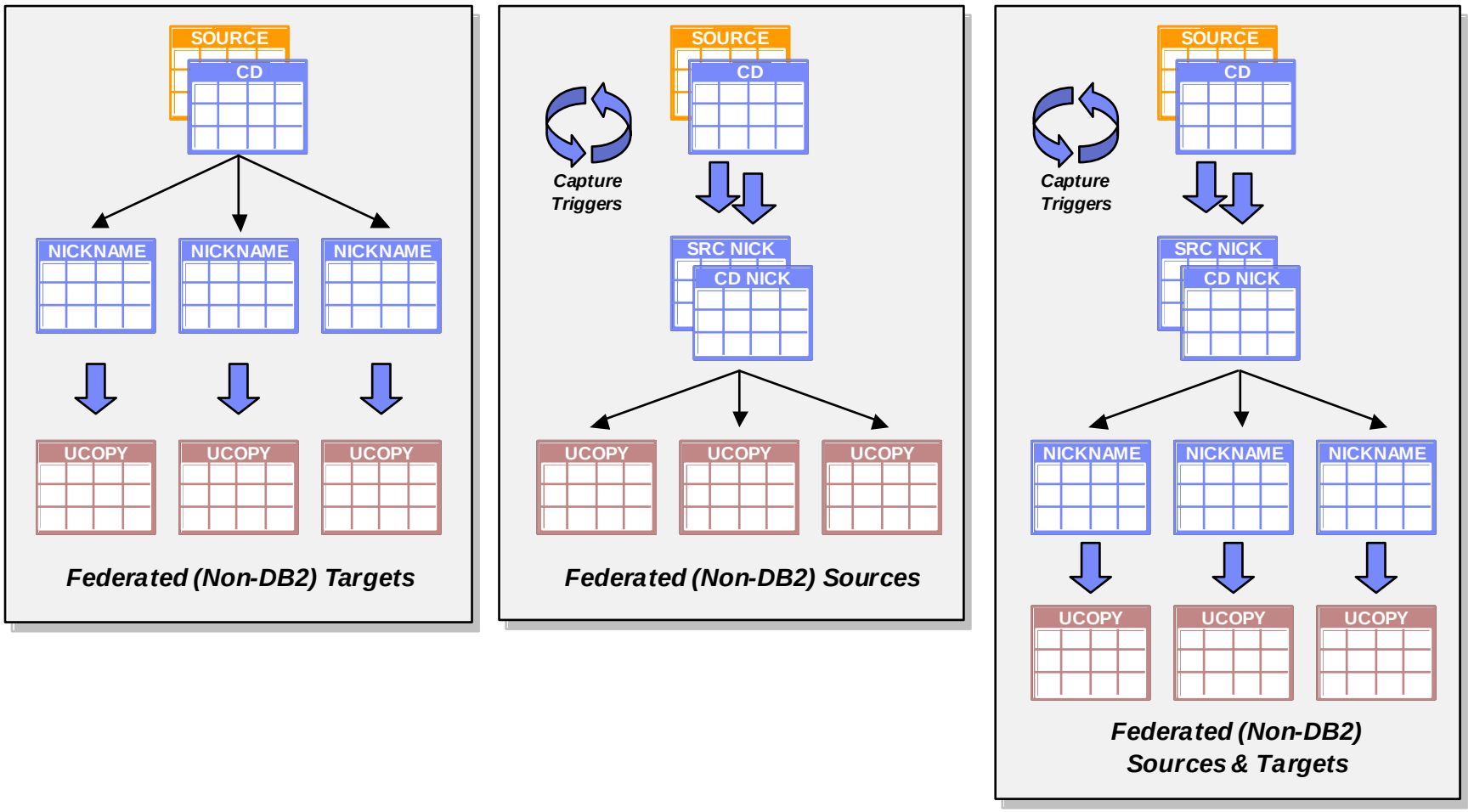
IBM DB2 DataPropagator
SQL-Replication – Component Overview



Sample SQL-Replication Scenarios



Sample SQL-Replication Scenarios (II)



Summary – DProp SQL-Replication Overview

DProp Capture

- Asynchronous DB2 Log Read Interface
- Captures Changes from DB2 Log into Change Data Tables

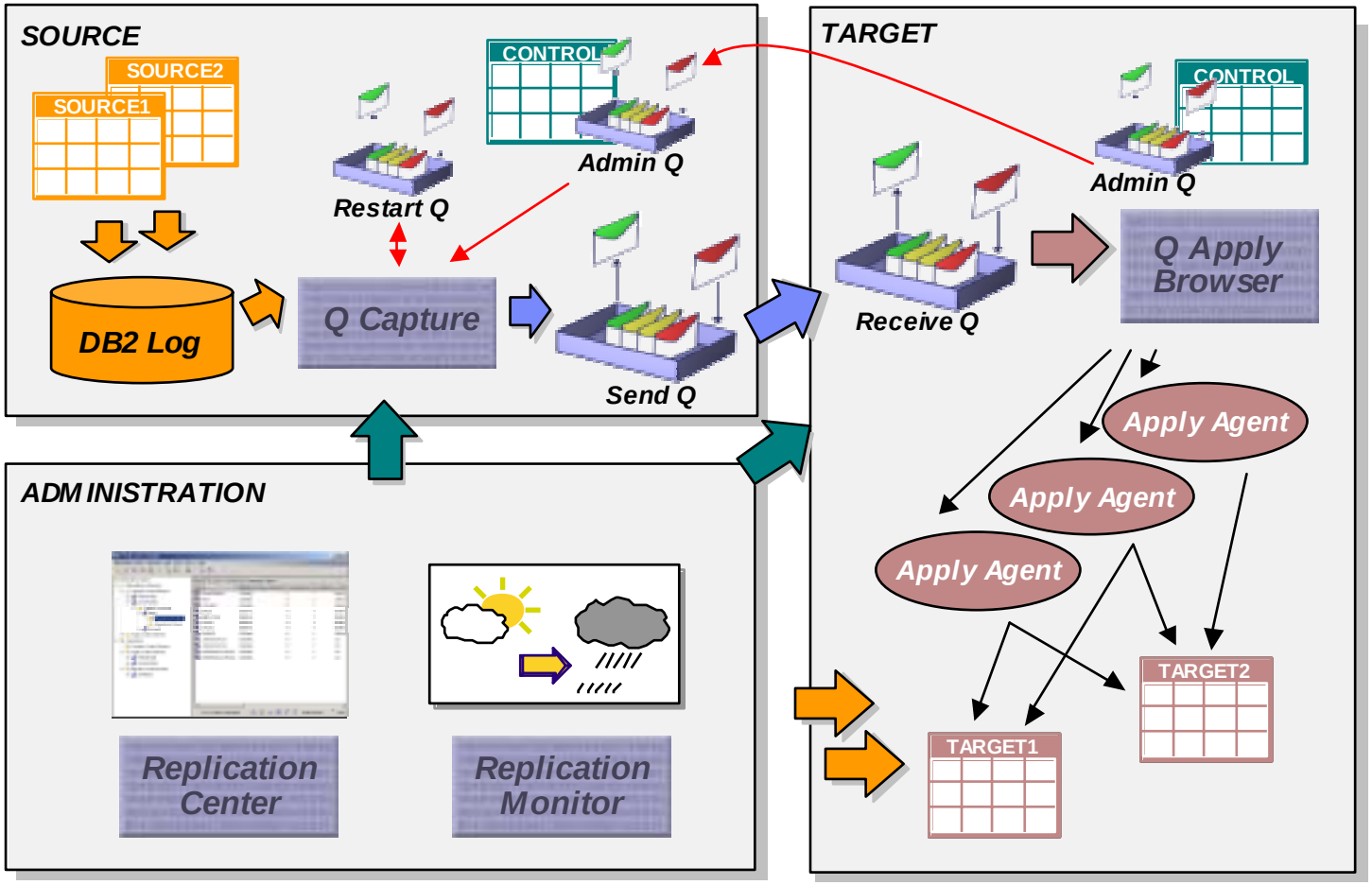
DProp Apply

- Refreshes, Updates or Enhances Target Tables
- Multiple Target Table Types supported

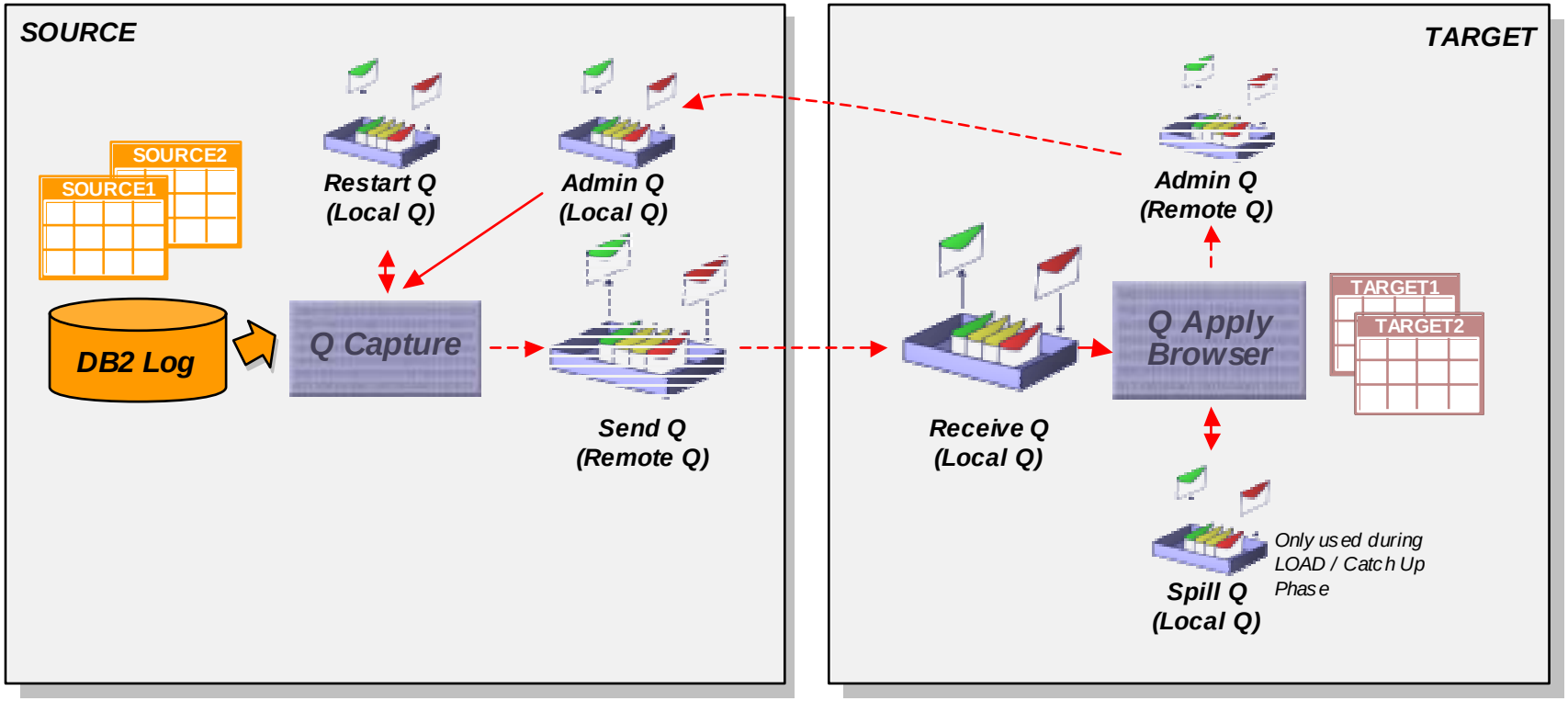
SQL-based Replication Infrastructure

- Data Enhancement through relational Staging Concept
- View and Joins supported as Sources
- Data Distribution Capabilities in huge mobile Environments
- Multi-Vendor Interoperability

WebSphere Information Integrator
Q-Replication – Component Overview



Q-Replication – Queue Setup (Two different Servers)



MQ Series Primer

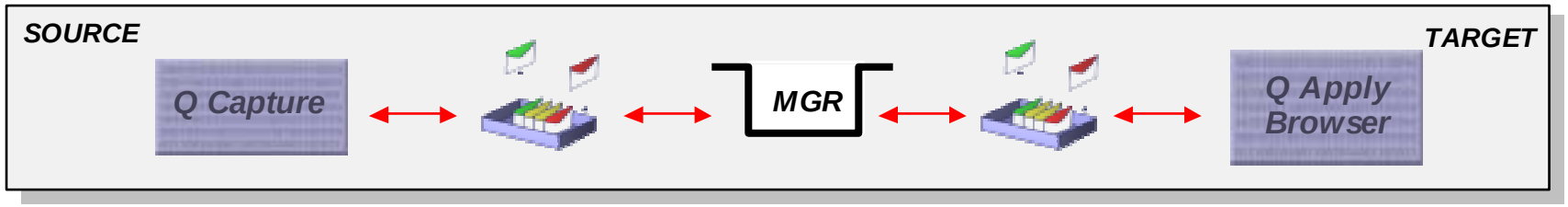
http://publib-b.boulder.ibm.com/Redbooks.nsf/RedbookAbstracts/redp0021.html

Q Replication Setup Fast Starter

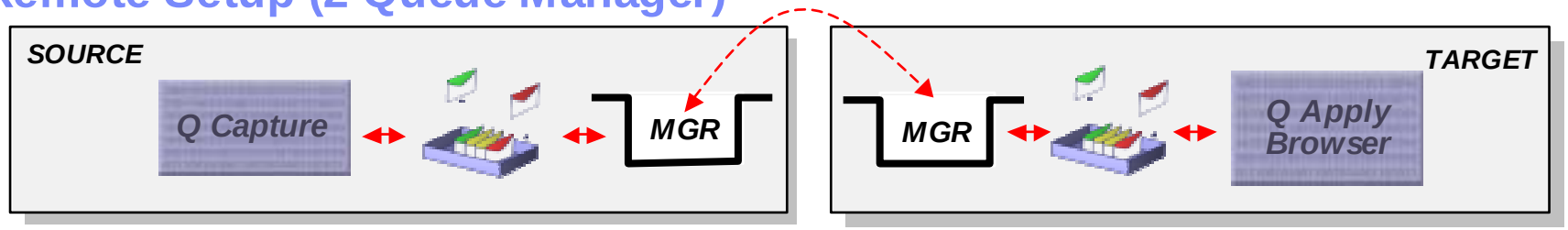
http://www-128.ibm.com/developerworks/edu/dm-dw-dm-0409burner-i.html?S_TACT=104AHW11&S_CMP=LIB

Q-Replication – Supported Queue Manager Topologies

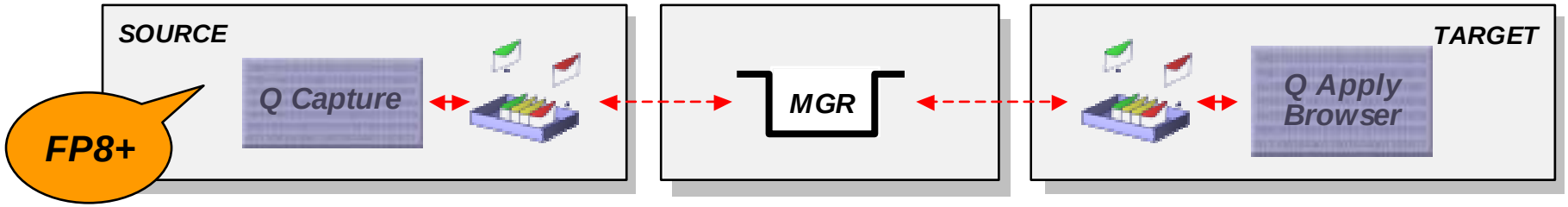
Local Setup (1 Queue Manager)



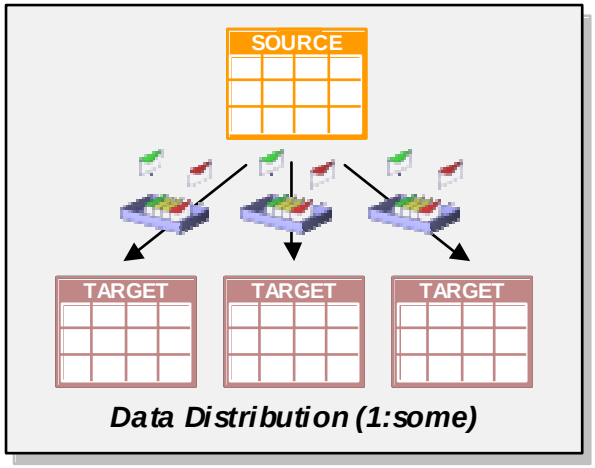
Remote Setup (2 Queue Manager)

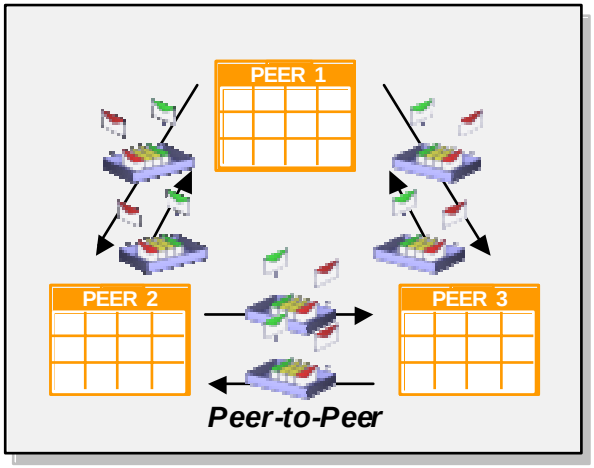


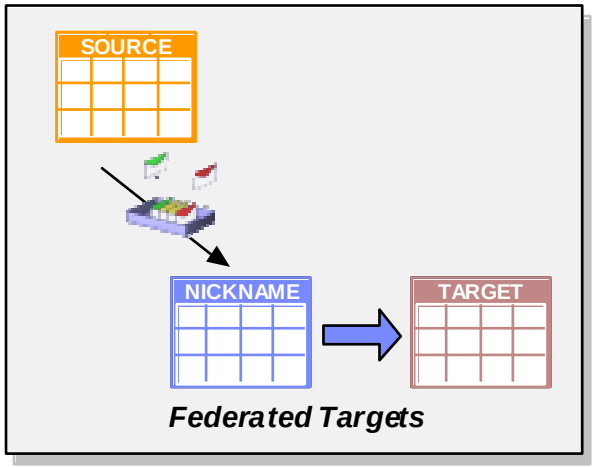
Client Setup (remote Queue Manager(s))

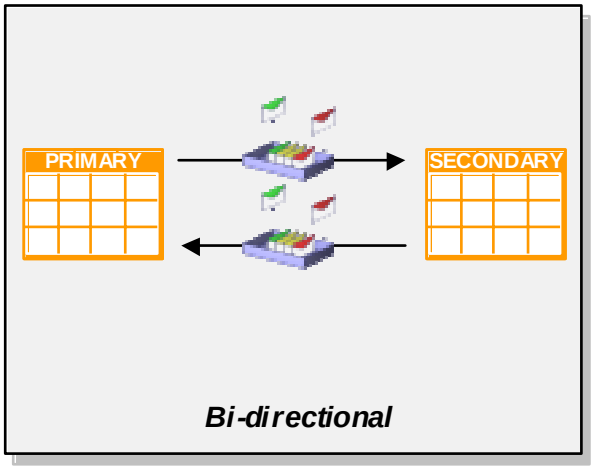


Sample Q-Replication Scenarios





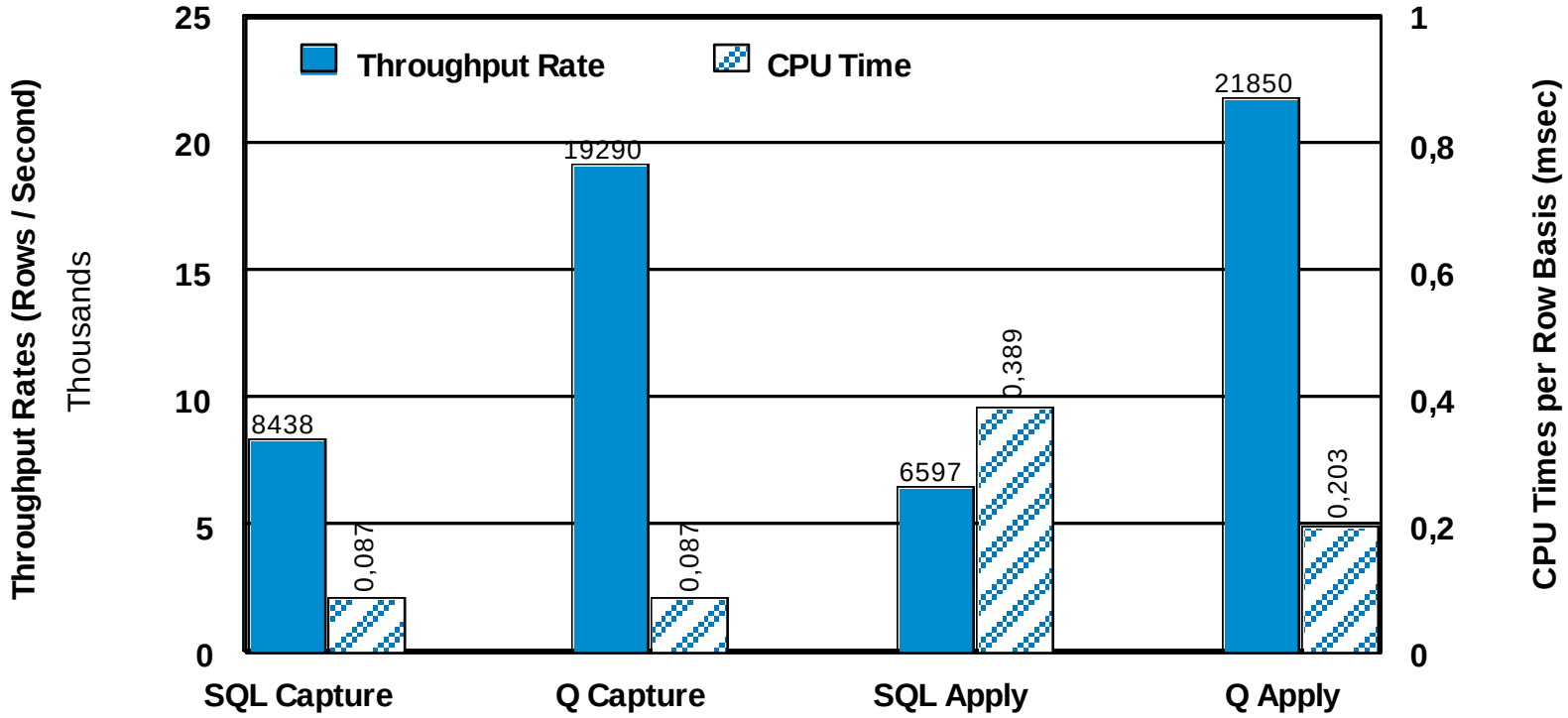




Key Scenarios:

- Low-Latency Replication
- Geographically dispersed Applications with distributed Databases
- Bi-directional Replication with Conflict Checking, Handling, and Notification
- Software-based Hot-Standby
- Cross DB2-Family
- Federated Target Support

Throughput & Performance



- Q Capture max throughput rate is about 2.3 times better in compare with SQL Capture
- Q Apply max throughput rate is about 3.3 times better in compare with SQL Apply

Replication Performance Analysis Summary

http://www-128.ibm.com/developerworks/db2/library/techarticle/dm-0503aschoff/

Summary – Q-Replication Overview

Q-Capture

- Captures Changes from DB2 Log into Message Queue(s)
- Websphere MQ replaces the use of Staging Tables (CD Tables)
- Each Message represents a Transaction
- Very Compact Internal Message Format

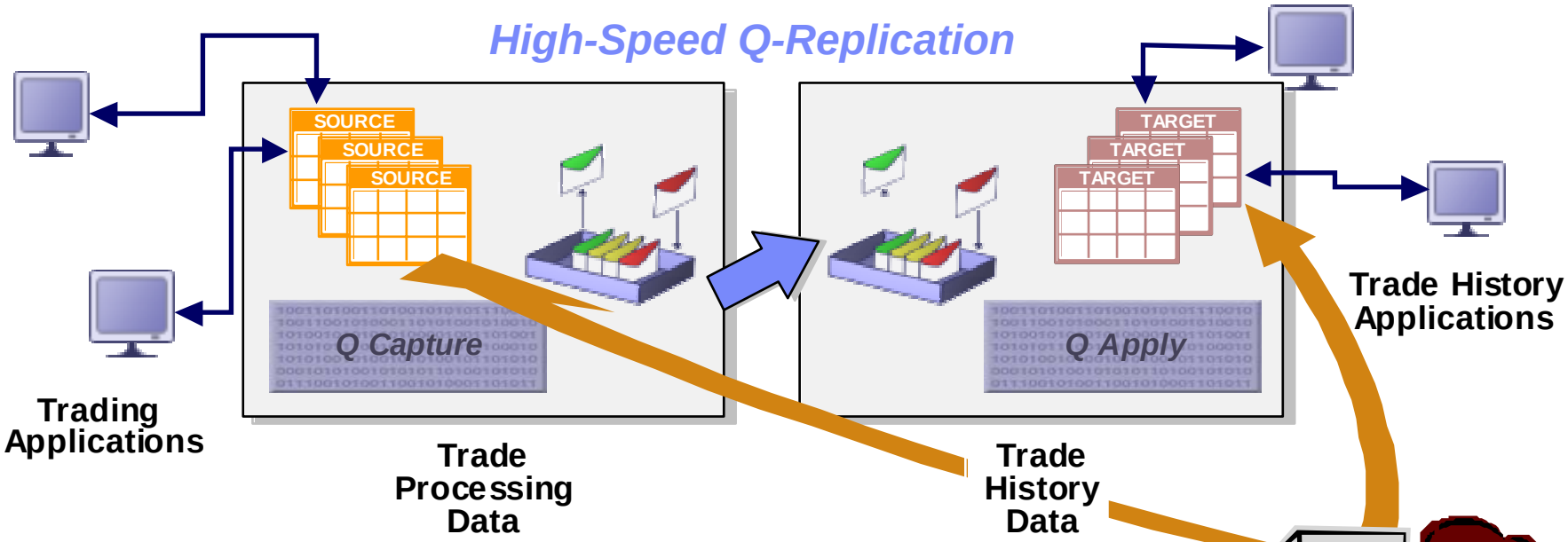
Q-Apply

- Highly-Parallel Apply Processing
- Conflict Detection, Resolution and Documentation

Websphere MQ

- Robust, Secure, and High-Performance Messaging Infrastructure
- Available on all commercially relevant Platforms

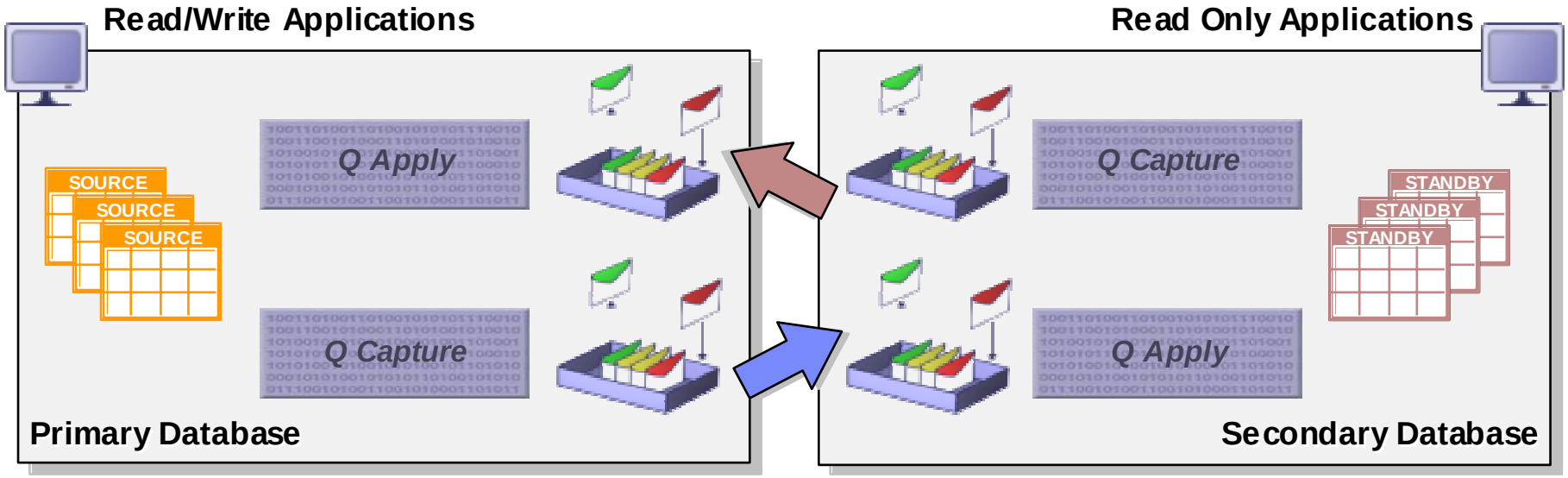
Example 1:
Feeding Trade-History Database with Q-Replication



Business Scenario

- In many Online Environments OLTP Data is kept separately from Query/History Data for better Performance of both Update and Query Applications
- This user has just made an Online Trade – he will keep hitting Enter until he sees that the Trade is complete, in this Case meaning it has been replicated to the Trade History Database

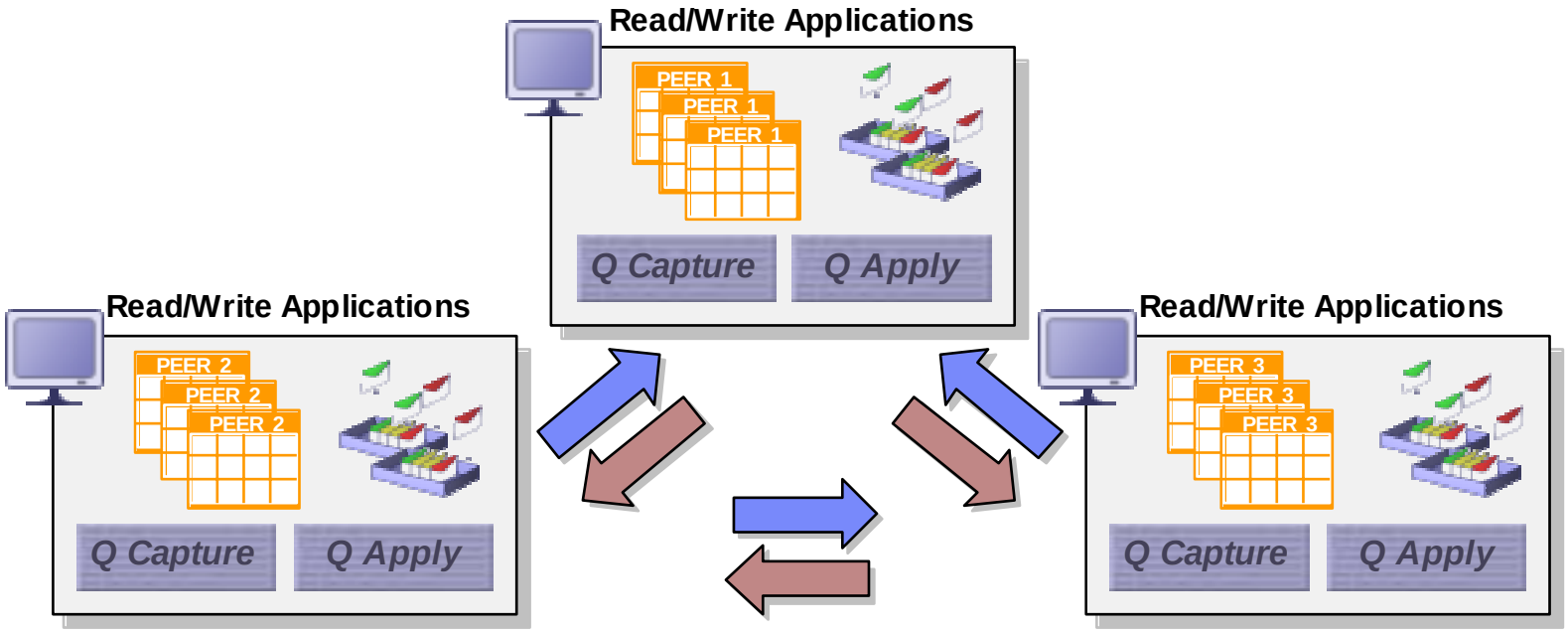
Example 2:
High-Availability Solution built upon Q-Replication



Business Scenario

- Replication Processes and Subscriptions are defined in both Directions, but Data mainly flows in one Direction at a Time
- Recursion is stopped by Capture, which reads special logged Events created by Apply
- Data at the Secondary System is transactionally consistent and is available for “read only” Applications permanently
- Procedures for Failover and Switchback will depend on which Options have been selected for Conflict Detection

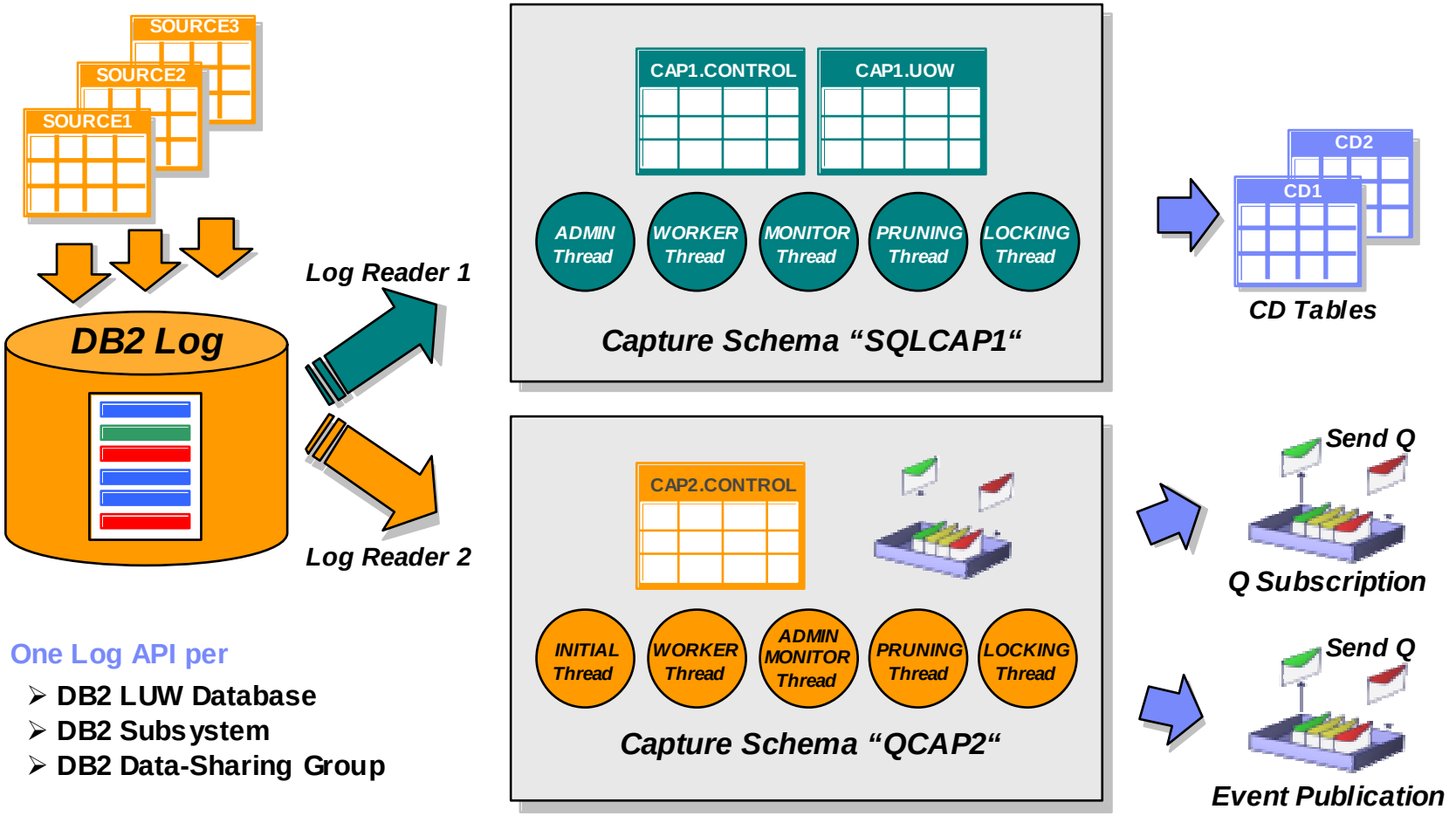
Example 3:
Geographically dispersed Applications (Peer-2-Peer)



Business Scenario

- Each Site operates its own Application and Database
- Each Database holds the unified Data of all Locations
- Conflict Detection and Resolution is typically necessary, unless the Application is carefully designed to completely avoid Conflicts
- In Case of a local Unavailability, the Applications can switch to a Remote Site (DB2 Client Configuration Option)

How does the Q-Replication & Event Publishing Solution complement existing IBM Data Replication Infrastructures



One Log API per

- DB2 LUW Database
- DB2 Subsystem
- DB2 Data-Sharing Group

Reference: CitiStreet

citistreet
A State Street and Citigroup Company

Challenge

- **Support single sign-on access through both Web and IVR applications ensuring 24x7 portal access for plan participants and sponsors**

Solution

- **Support redundant, active single sign-on applications for failover processing replicating profile changes between them in real time.**

"Since nearly 10 million of CitiStreet customers are offered 24-hour access to their retirement accounts, the company can't afford downtime and must be able to replicate data changes when they happen. We fully replicate our database over redundancy data lines, so to us the stability and speed of that asynchronous replication is strategic for us."

Barry Strasnick , CIO
CitiStreet

Overview

- CitiStreet is one of the largest and most experienced global benefits providers servicing over 9 million plan participants across all markets. CitiStreet was formed in partnership between subsidiaries of State Street Corporation and Citigroup

Business benefits

- Ensure application availability for plan participants and sponsors
- The new solutions from IBM will improve data integrity with a reduced level of maintenance

Technology benefits

- Maintain bi-directional synchronization of profile updates (approx 175,000 updates daily) in real time

Agenda

IBM's View of Information Integration

IBM Data Replication Architectures:
Introducing High-Volume and Low-Latency Replication with MQ-based Replication Technologies

Publishing of Data Events:
Integration of Processes, Applications, and Information

Summary, Collateral, Demo

Why Publish Data?

Database to Application Messaging

- Drive downstream Applications or APIs based on the Transactional Data of the changed Database Events

Event Notification

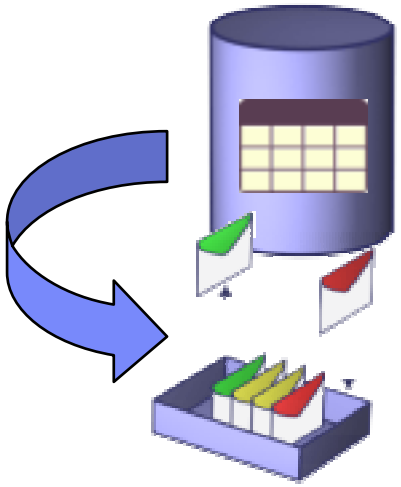
- Stream changed Data Information to Web Interfaces
- Stream only particular Events of Interest (filter Data)

Data Warehouse / Business Intelligence

- Integrate captured Changed Data with an ETL Tool
- Perform complex Transformations with custom Logic
- Use a specific Transaction Format to update Target

MQ provides guaranteed delivery

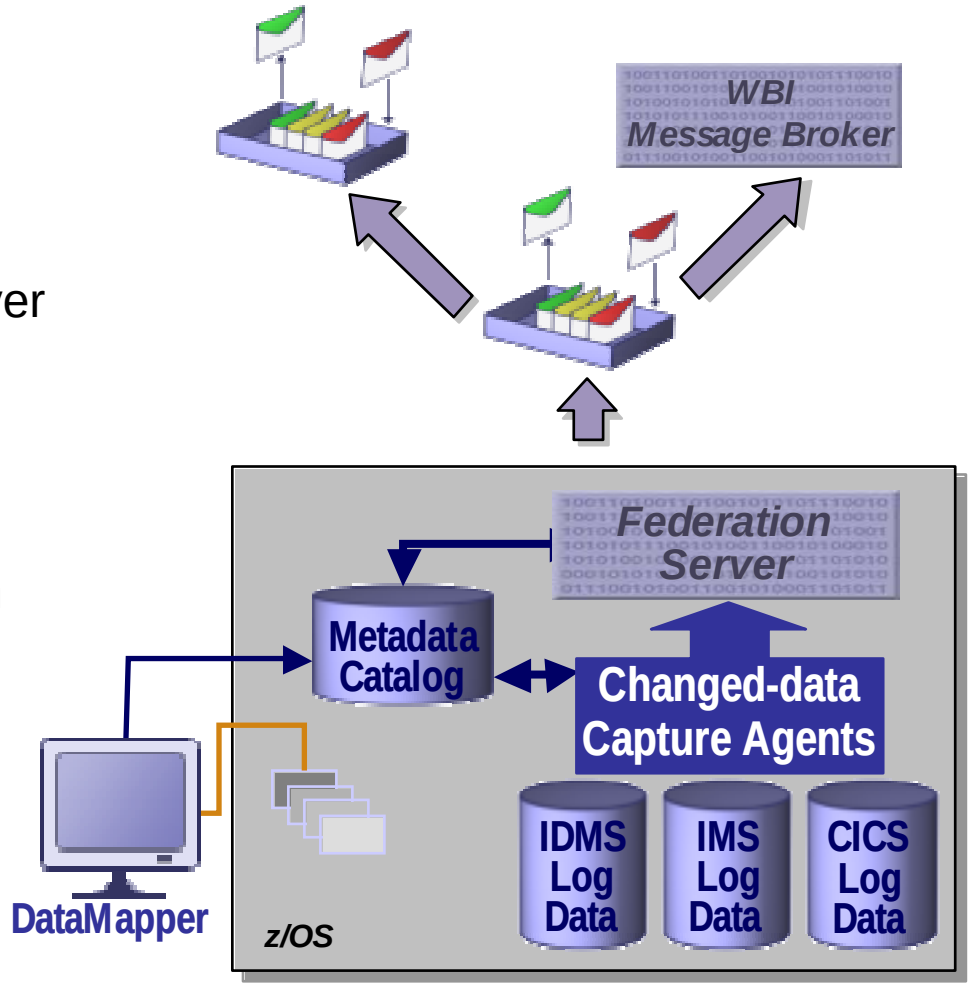
- Avoids the need for 2-Phase-Commit (2PC)
- Works even when the Target is not available



Classic Event Publisher: IMS, VSAM & IDMS Implementation

Architecture & Setup

- Leverages WebSphere Information Integrator Classic Federation Mappings and Server Infrastructure (incl. Relational Mappings)
- Install changed Data Capture Agents for IMS and/or VSAM
- Identify data to be monitored in the Metadata

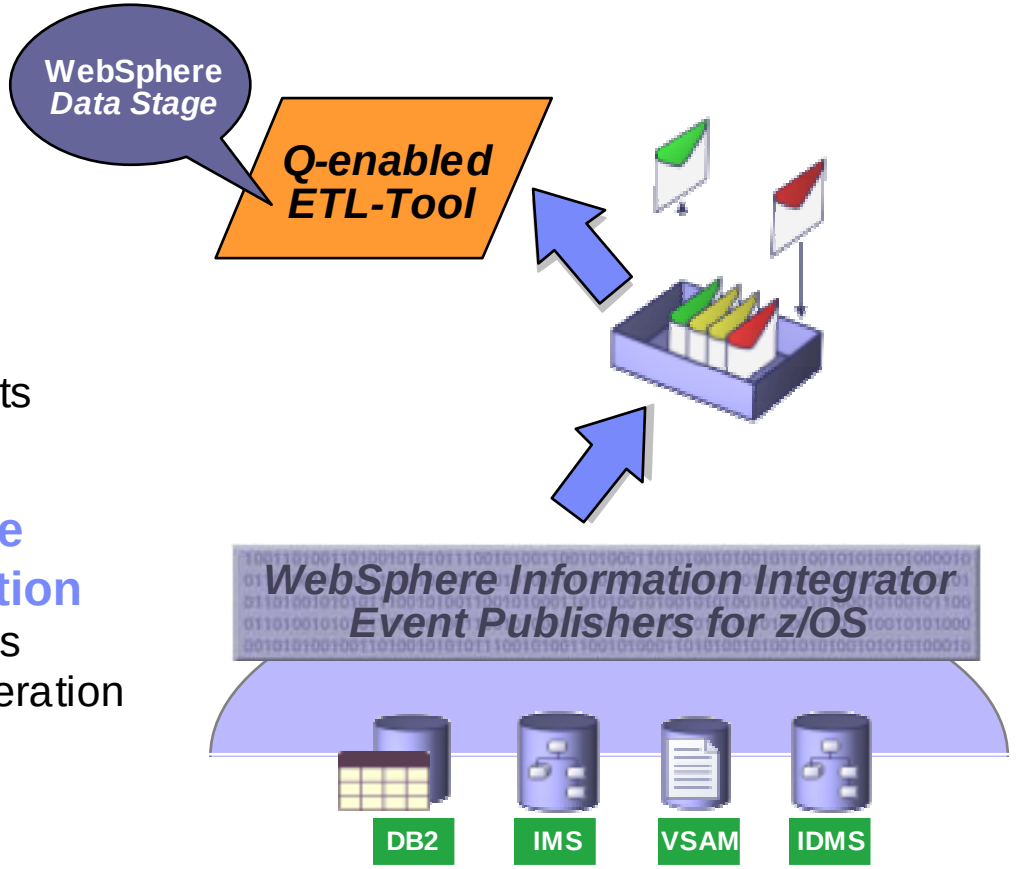


Example 1: WebSphere Information Integrator Event Publisher & Business Intelligence

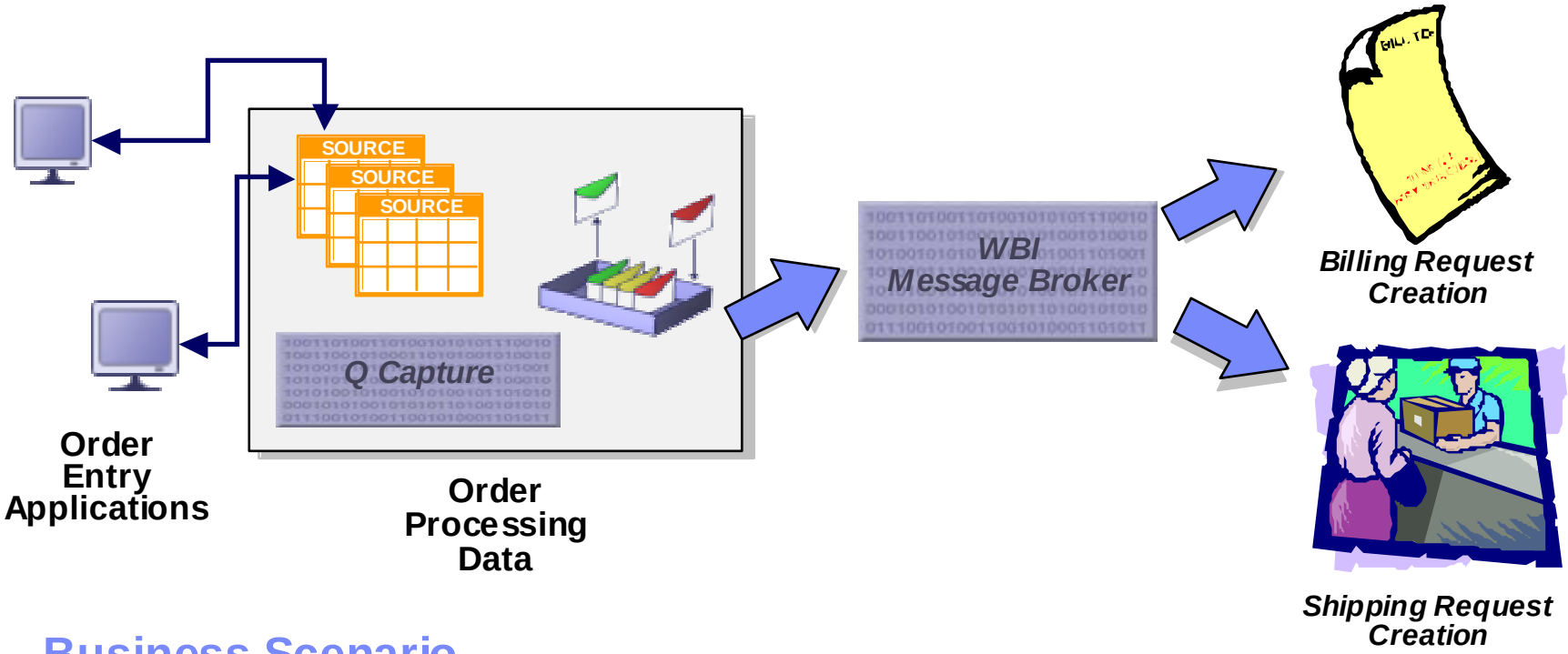
Feeding Changed Data to
- Data Warehouse
- Datamart
- Operational Data Store (ODS)

Optimize Resource Utilization
- Minimize Bandwidth Requirements
- Maximize Data Currency

Complements with WebSphere Information Integrator Federation
- Data Feed using Event Publishers
- Real-Time Extensions using Federation



Example 2:
Application to Application Messaging with Event Publishing



Business Scenario

- As new Orders are entered into the Order Entry System, the pertinent Data is captured and published into a Queue
- The Websphere MQ Integrator Broker processes the queued Data
- A billing Transaction is created and queued in one System and a Shipping Transaction is created and queued in another System

Agenda

IBM's View of Information Integration

IBM Data Replication Architectures:
Introducing High-Volume and Low-Latency Replication with MQ-based Replication Technologies

Publishing of Data Events:
Integration of Processes, Applications, and Information

Summary, Collateral, Demo

Summary

IBM is building out the On Demand Operating Environment

- WebSphere Information Integrator is a Key Component
- IBM is driving the Industry forward in Enterprise Information Integration (EII) with both Vision and Technology

Information Integrator Technologies delivers tangible Business Value

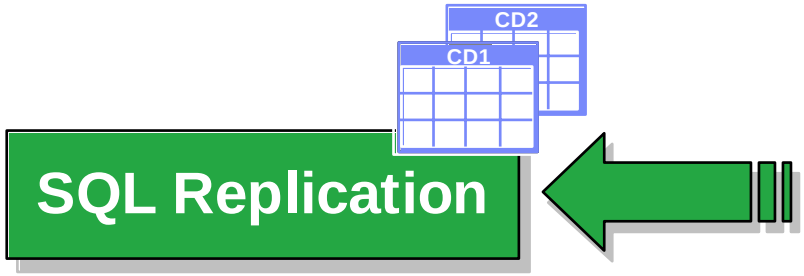
- Speed Time to Market for Integration Projects
- Get more Value and Insight from existing Assets
- Control IT Costs with tailored Business Views and reduced Requirement for new Copies of Data

Masala Features have extended IBM's Leadership in Information Integration

- High-Speed Data Replication based on open, versatile DB2 and Massaging Infrastructure
- Reactivating *Classic* Data
- Integrating seamlessly with Application and Process Integration Technologies

Some Rules – When to choose What (I)

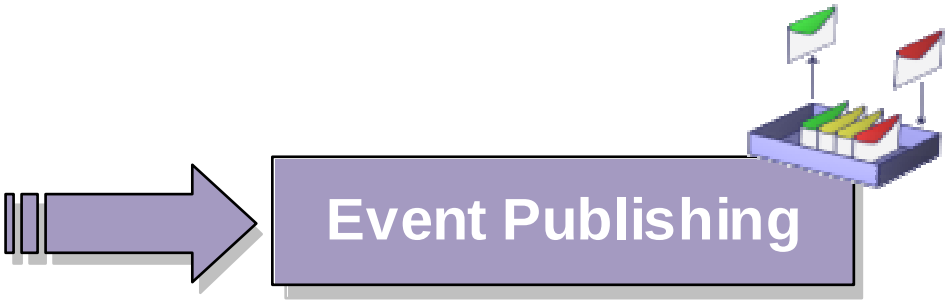
- ✚Low-Latency, High-Volume Replication
- ✚Source and Target Tables of similar Structure
- ✚Bi-directional Replication (e.g. for Hot-Standby Purposes)
- ✚Peer-2-Peer – Splitted Workload
- ✚Huge Number of Tables (e.g. Siebel)
- ✓Replication across DB2-Family
- ✓Replication from DB2 to federated Targets

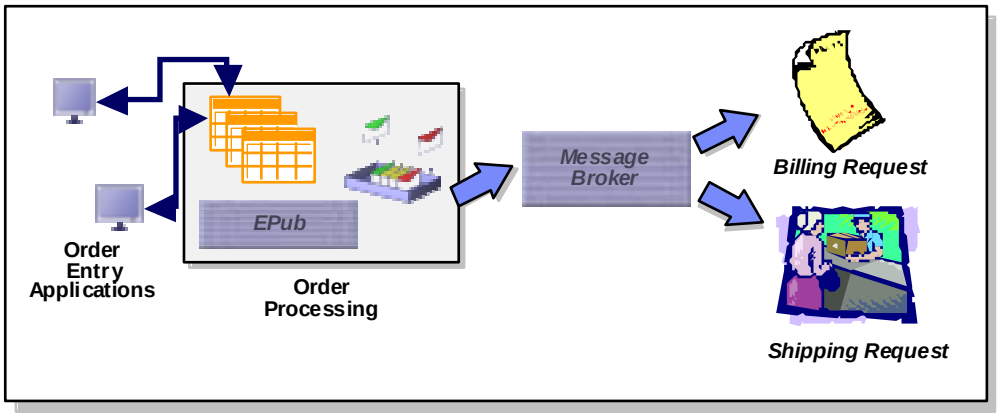


- ✚Fan-Out to huge Number of Targets
- ✚Multi-Tier Staging via CCD
- ✚Source and Target Tables not of the same Structure (which requires Source Views, Joins or SQL Expressions)
- ✓Replication from federated Sources

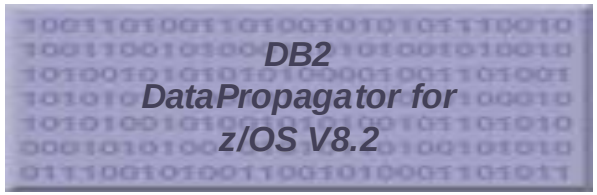
Some Rules – When to choose What (II)

- Processes triggered by Data Events
- Kick-Off of Workflows due to changed Data
- Seamless Integration of DB2 and *Classic* Data (IMS, VSAM, IDMS)
- Delivery of Data Events from various Origins to a single Application or Message Broker
- Real-Time ETL

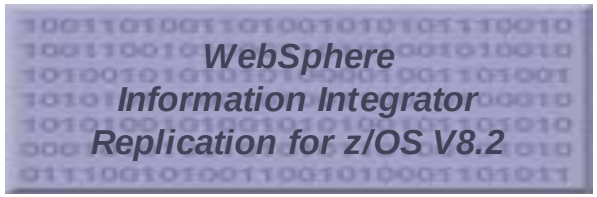




Replication Products: z/OS



- SQL Replication Architecture (DProp Capture and Apply)
- Available for DB2 UDB z/OS V7 and V8



- Includes Q Replication and SQL Replication
- Includes Event Publisher for DB2 UDB z/OS
- Available for DB2 UDB z/OS V7 and V8
- Websphere MQ prerequisite when using Q Replication

Event Publishing Products: z/OS

WebSphere Information Integrator Event Publisher for DB2 z/OS V8.2

- Q-based Event Publisher for DB2 (Q Capture)
- Websphere MQ prerequisite

WebSphere Information Integrator Classic Event Publisher for IMS V8.2

- Q-based Event Publisher for IMS
- Websphere MQ prerequisite

WebSphere Information Integrator Classic Event Publisher for VSAM V8.2

- Q-based Event Publisher for CICS/VSAM
- Websphere MQ prerequisite

WebSphere Information Integrator Classic Event Publisher for CA-IDMS V8.2

- Q-based Event Publisher for CA-IDMS
- Websphere MQ prerequisite

Replication and Event Publishing Products: Linux, Unix, Windows

DB2 UDB V8.2 for Linux, Unix, Windows includes DB2 DataPropagator

- SQL Replication Architecture
- SQL Capture and SQL Apply (for all DB2 UDB V8 Editions incl. Partitioning Feature)
- DB2 Sources and Targets. Informix IDS Sources and Targets supported through Federation Capability

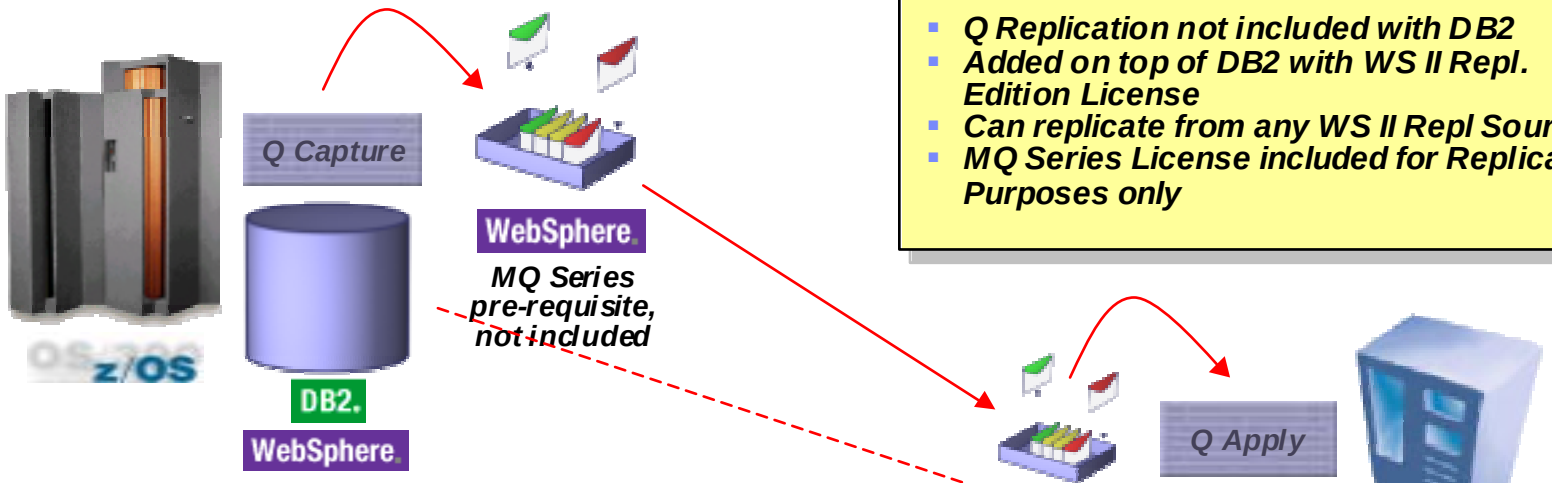
WebSphere Information Integrator Replication Edition V8.2

- SQL Replication: DB2 & Multi-Vendor Sources and Targets (Oracle, Informix, Sybase, MS SQL Server)
- Q Replication: DB2 Sources and Targets
- Multi-Vendor Targets (Oracle, Informix, Sybase, MS SQL Server)
- Includes Event Publisher for DB2 UDB for LUW
- Websphere MQ is bundled with this Product

WebSphere Information Integrator Event Publisher Edition V8.2

- Q Architecture: DB2 Sources
- Websphere MQ is bundled with this Product
- Data Changes pubished through Message Queues in external XML Format

Sample Configuration 1:
Q-Replication from DB2 for z/OS to DB2 for AIX



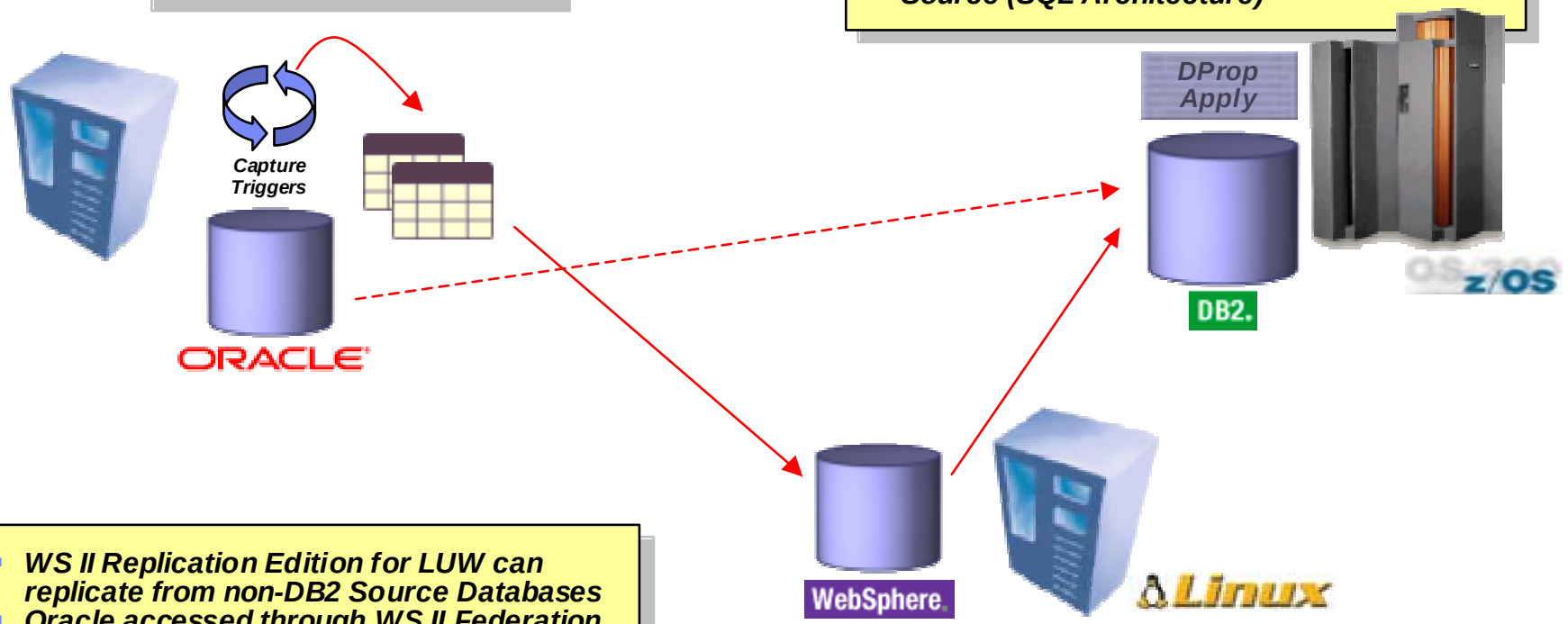
- *Q Replication not included with DB2*
- *Added on top of DB2 with WS II Repl. Edition License*
- *Can replicate from any WS II Repl Source*
- *MQ Series License included for Replication Purposes only*

- *WS II Replication for z/OS contains Q Replication and SQL Replication*
- *MQ License on z/OS required*
- *DB2 z/OS can be Source for any WS II Target*
- *Especially priced Bundles exist which combine z/OS Replication with Replication on LUW*

Sample Configuration 2:
SQL-Replication from Oracle for Solaris to DB2 for Linux

- *Trigger-based Capturing, Triggers created through DB2 Replication Center*

- *DataPropagator still a seperately orderable Product*
- *Can replicate from any DB2 or WS II Repl. Source (SQL Architecture)*



- *WS II Replication Edition for LUW can replicate from non-DB2 Source Databases*
- *Oracle accessed through WS II Federation Layer*
- *No MQ License/Installation required*

Information Integrator Replication Edition for Linux V8.2 + Oracle Connector
PID 5724-C74

Links & Learn ...

IBM Internet

- WebSphere Information Integrator V8.2 Homepage (incl. Federation, Replication, Event Publishing, Classic Federation, OnmiFind):
  http://www.ibm.com/software/data/integration
- WebSphere Information Integrator „Developer Zone“ (incl. White Papers, Hints & Tips, Samples)
  http://www-106.ibm.com/developerworks/db2/zones/db2ii/

Includes Q Replication Fast-Starter

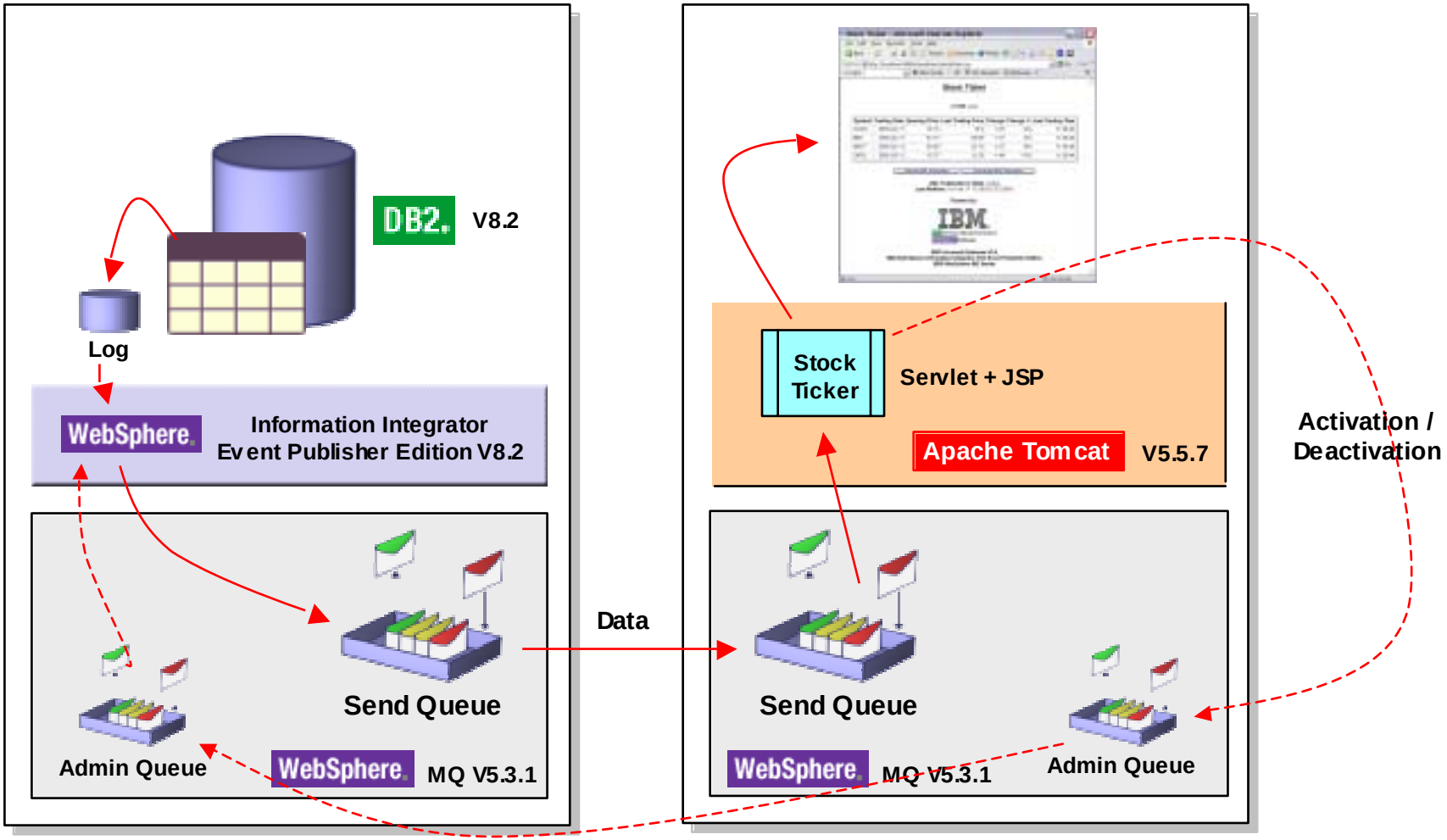
Dokumentation und Redbooks

- Comprehensive WebSphere Information Integrator Online-Documentati
  (incl. Federated Systems Guide, Installation and Coniguration Guid
  http://publib.boulder.ibm.com/infocenter/db2help/index.jsp
- PDF Documentation Download
  http://www-306.ibm.com/software/data/db2/udb/support/manualsv8.html
- Current WebSphere Information Integrator Replication Redbooks
  SQL Replication: http://www.redbooks.ibm.com/abstracts/sg246828.html
  Q Replication: http://www.redbooks.ibm.com/abstracts/sg246487.html

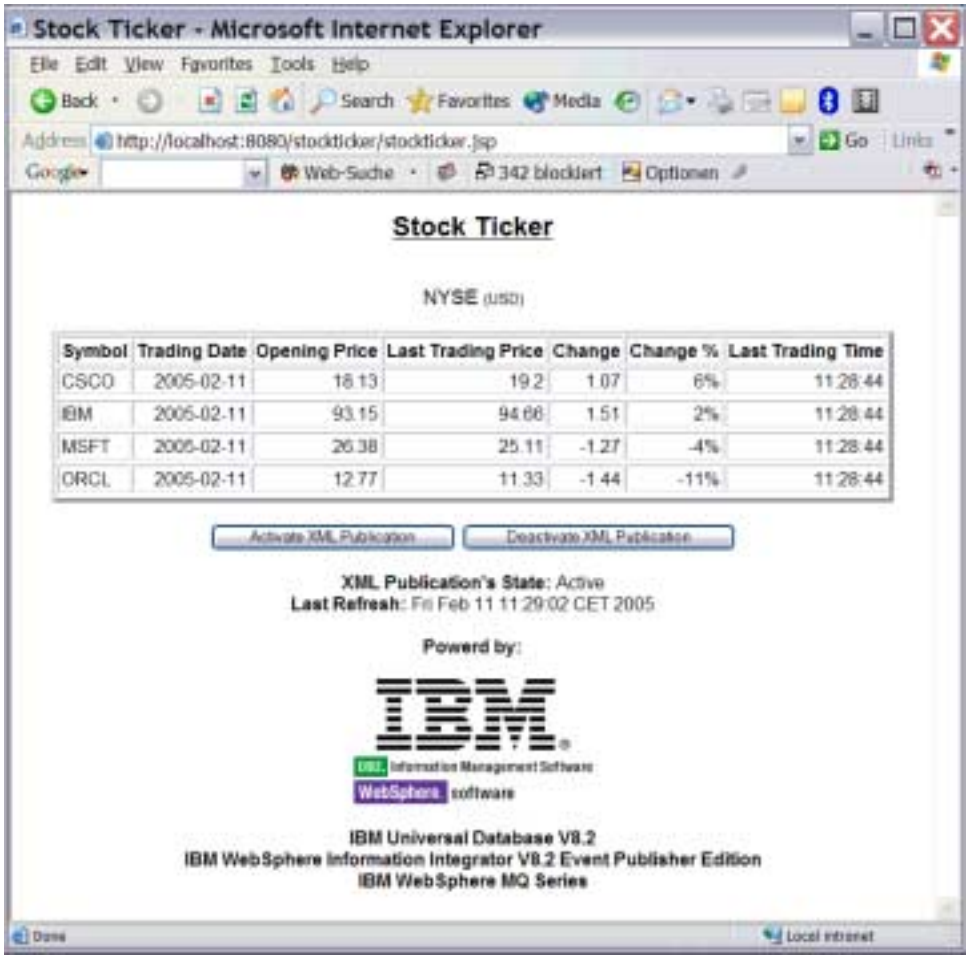
Includes Replication Guides and ASNCLP documentation

NEW

A DEMO, if you like:
Stock Ticker Application powered by II Event Publisher



The Stock Ticker Browser Application – Look and Feel



Finish