A09: Continuous Availability

using Data Replication and GDPS Active-Active

Karen Durward
Product Manager for System z, Information Integration and Governance
2015-03-18

IMS Technical Symposium 2015

Agenda – Continuous Availability

Why is High Availability no longer enough
- Background
- Usage Scenarios

InfoSphere Data Replication for z/OS
- DB2, IMS, and VSAM
- InfoSphere Data Replication for IMS – In Depth

The GDPS Family of Solutions
- Disaster Recovery and High Availability: GDPS PPRC, XRC, GM, …
- Continuous Availability: GDPS Active-Active

Wrap-Up

How much interruption can your business tolerate?

Business Continuity Spectrum

Cold Standby

Warm Standby

Hot Standby

Active/Active

- Disaster Recovery
  - Restore business after an unplanned outage

- High-Availability
  - 99.9% availability
  - 8.8 hours of down-time a year

- Continuous Availability
  - No downtime
    … planned or unplanned

Global Enterprises that operate across time zones no longer have any 'off-hours' window.

Continuous Availability is mandatory! High-Availability is not enough!

Annual Downtime of 300 to 1,200 hours depending on industry[1]
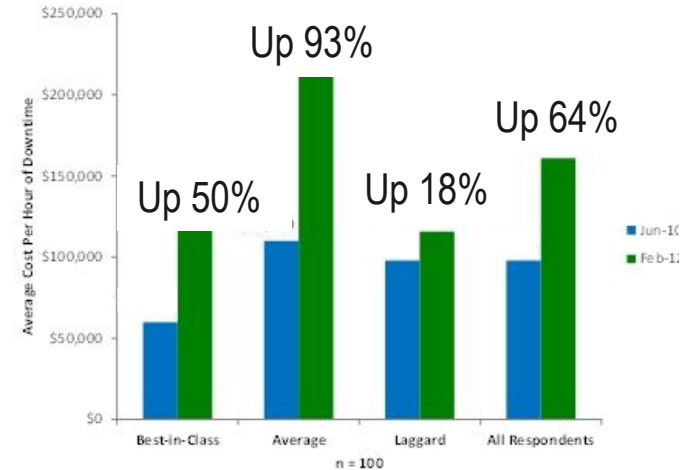
Source: Aberdeen Group, February 2012

Disruptions Also Impact Credibility & Market Position

20 July, 2013

DMV Computers Fail Statewide, Police Can't Access Database

18 August , 2013

Google total eclipse sees 40 percent drop in Internet traffic

22 August, 2013

Nasdaq: 'Connectivity issue' led to three-hour shutdown

16 April, 2013

American Airlines Grounds Flights Nationwide

- Downtime costs can equal up to 16% of revenue[1]

- Four hours of downtime severely damaging for 32% of organizations[2]

- Fines for downtime & inability to meet regulatory compliance

- Data is growing at explosive rates –from 161EB in 2007 to 988EB in 2010[3]

[1] Infonetics Research, *The Costs of Enterprise Downtime: North American Vertical Markets 2005*, Rob Dearborn and others, January 2005.
[2] Continuity Central, "Business Continuity Unwrapped," 2006, http://www.continuitycentral.com/feature0358.htm
[3] The Expanding Digital Universe: A Forecast of Worldwide Information Growth Through 2010, IDC white paper #206171, March 2007

4

Lesson learned from September 11, 2001
Periodic testing and geographic dispersion are critical

1. Identify clearing and settlement activities to provide critical support of financial markets
2. Determine appropriate recovery and resumption objectives for clearing and settlement activities in support of critical markets
3. Maintain sufficient geographically dispersed resources to meet recovery and resumption objectives
4. Routinely use or test recovery and resumption arrangements.

Interagency Paper on Sound Practices to
Strengthen the Resilience of the U.S. Financial System
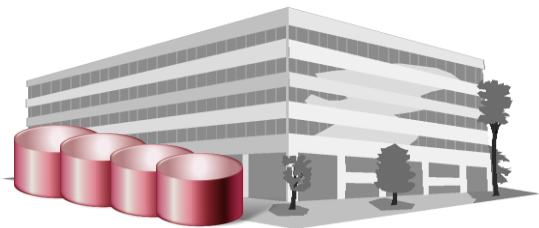[Docket No. R-1128] (April 7, 2003)

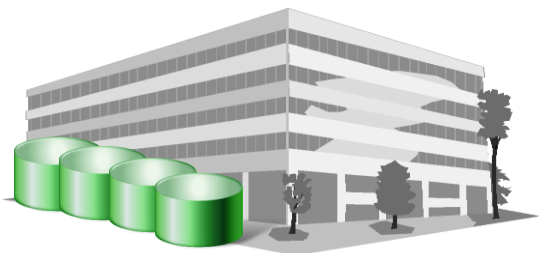# Continuous Availability Concepts

What does is take?

Continuous Availability Concept

New York

Madrid

Two or more sites, separated by _unlimited_ distances, running the same applications and having the same data to provide:

- *Continuous* Availability for both planned and unplanned outages
- Cross-site **Workload Balancing** to leverage all resources

Continuous Availability Concept

New York
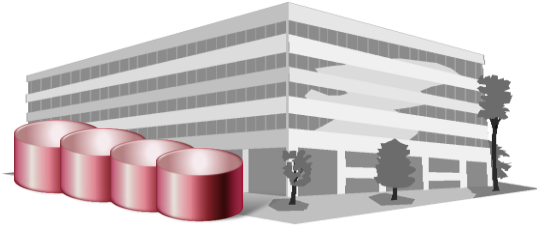
Replication

Data at geographically dispersed sites are kept in sync using very high speed, low latency software-based data replication (DB2, IMS and VSAM)
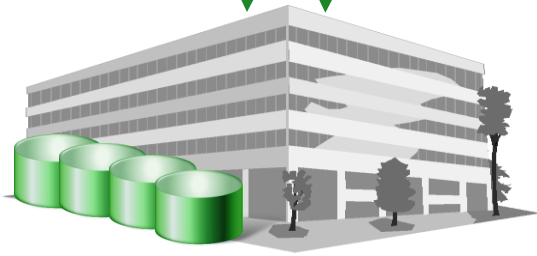
Madrid

Continuous Availability Concept

New York

Workloads are routed to one of the available sites based on workload weight and latency constraints

Transactions

Replication

Workload Distributor

Load Balancing with SASP (z/OS Comm Server)
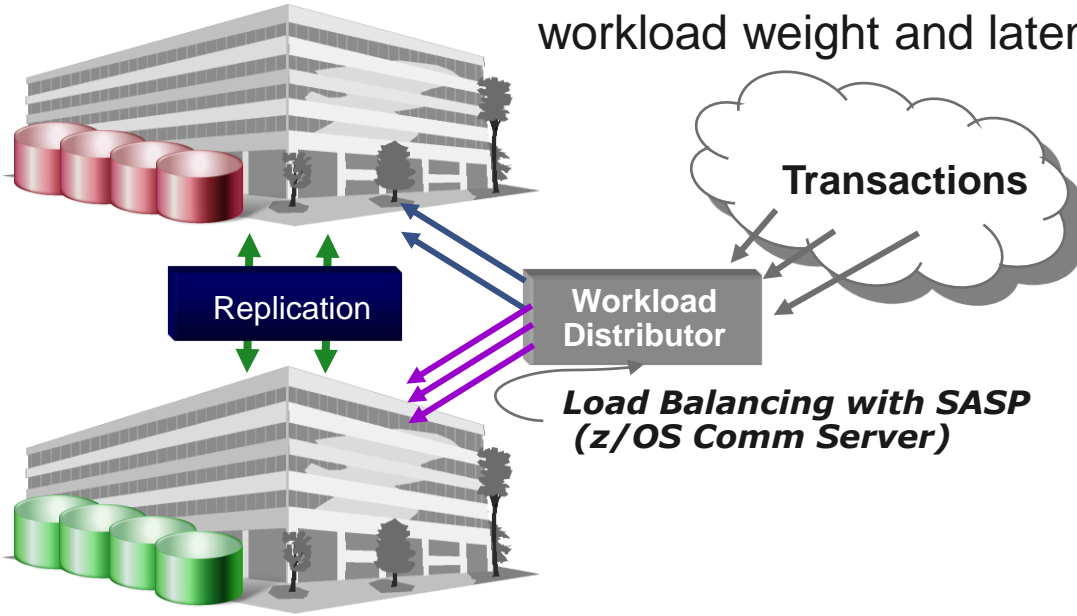
Madrid

Continuous Availability Concept

New York

Monitoring spans the sites and is an integral element of the solution for site health checks, performance tuning, etc.

Replication

Workload Distributor

Transactions

Load Balancing with SASP (z/OS Comm Server)

Tivoli Enterprise Portal

Madrid

Extended Use Cases

- Continuous availability during maintenance
    - Reduce planned outages

- Secondary application environment for Mobile application support
    - Highly query oriented
    - Very unpredictable workloads
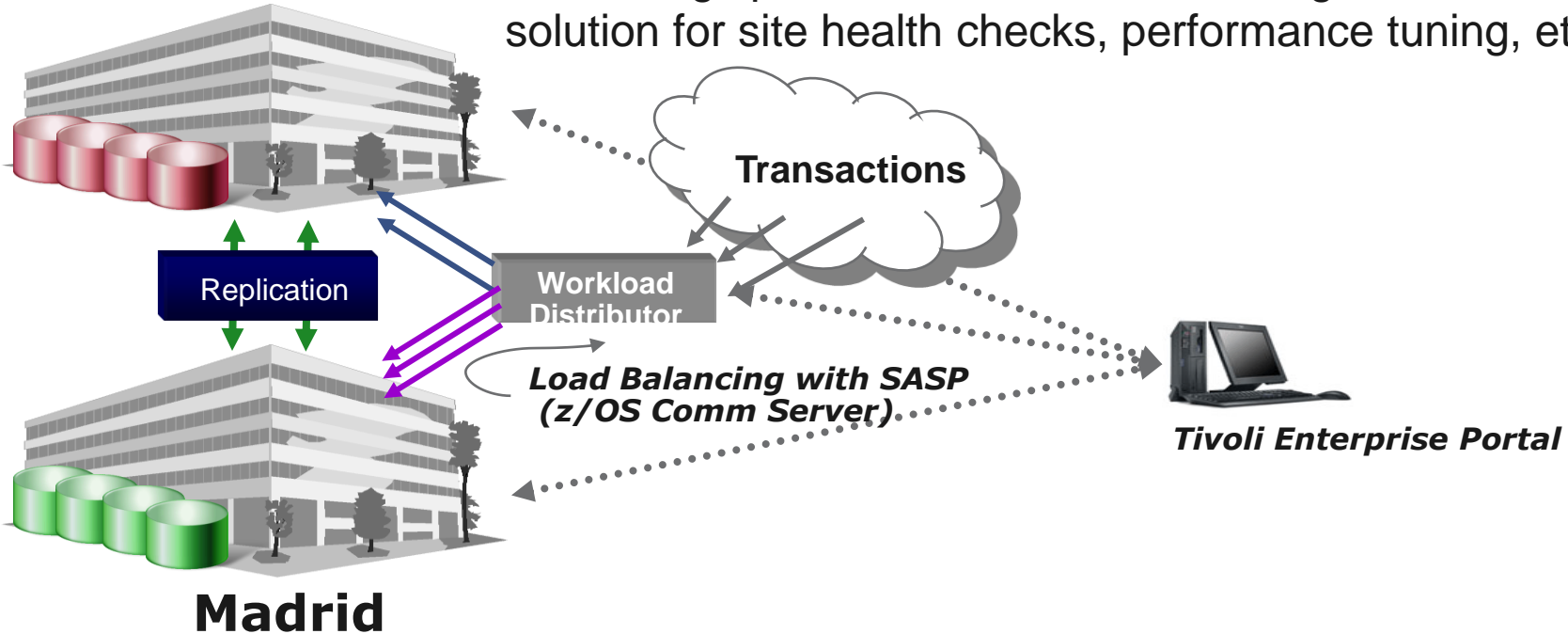   - Limit impact on traditional transaction processing environment

- Low latency replication for Data Warehousing and Analytics
    - Heterogeneous targeting for data warehousing
        - Transaction data to an RDBMS, Hadoop, etc.
    - Off load analytics to dedicated environment
        - Limit impact on transactions while still empowering near real time analytics

IBM's InfoSphere Data Replication (IIDR) Coverage

DB2 z/OS

DB2 z/OS

IMS

IMS

VSAM

VSAM

DB2 (z/OS, i, LUW)
Informix
Oracle
MS SQL Server
Sybase
PD for A (Netezza)
Teradata
Information Server
HDFS/Hive

Message Queues
Files
FlexRep *(JDBC targets)*
Customized Apply

DB2 (i, LUW)
Informix
Oracle
MS SQL Server
Sybase

ESB, Cognos Now, …

MySQL, GreenPlum, …
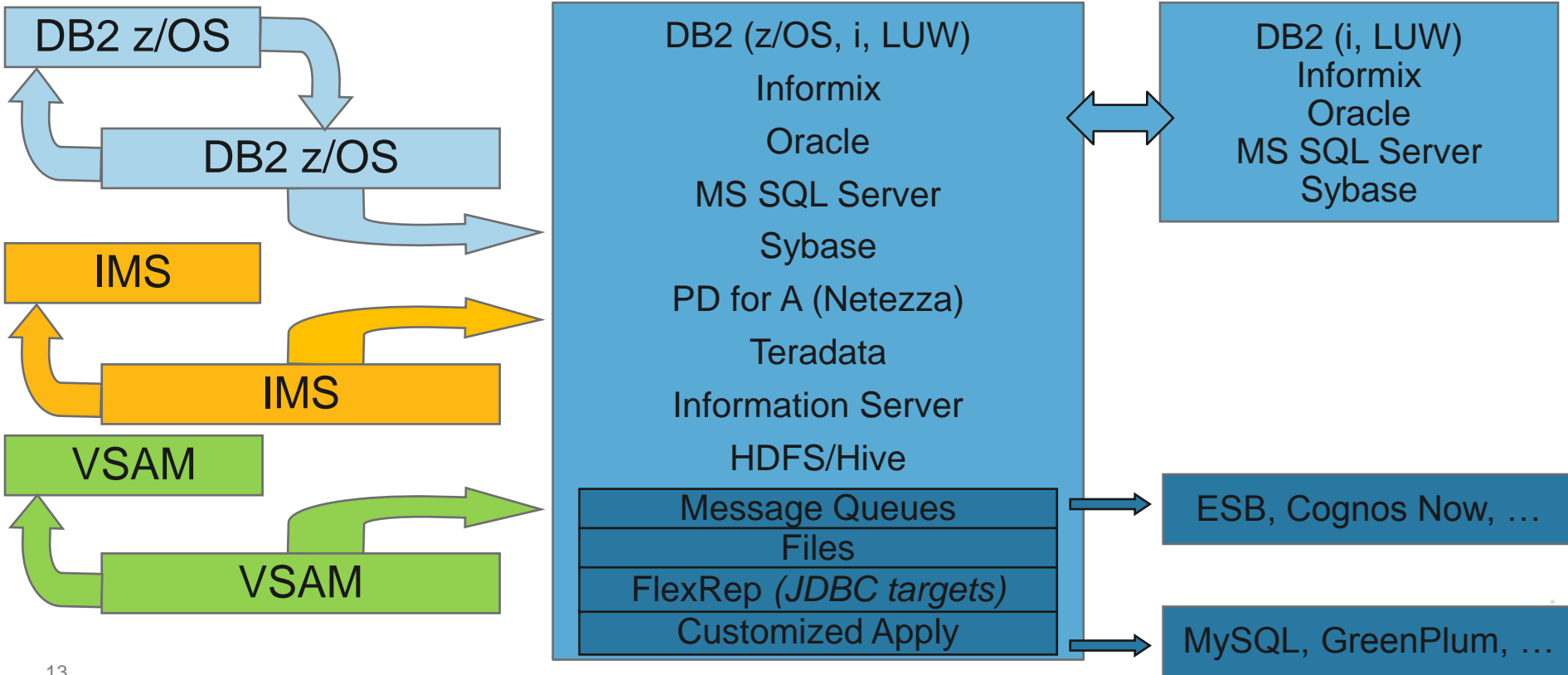
Focus on Continuous Availability Coverage
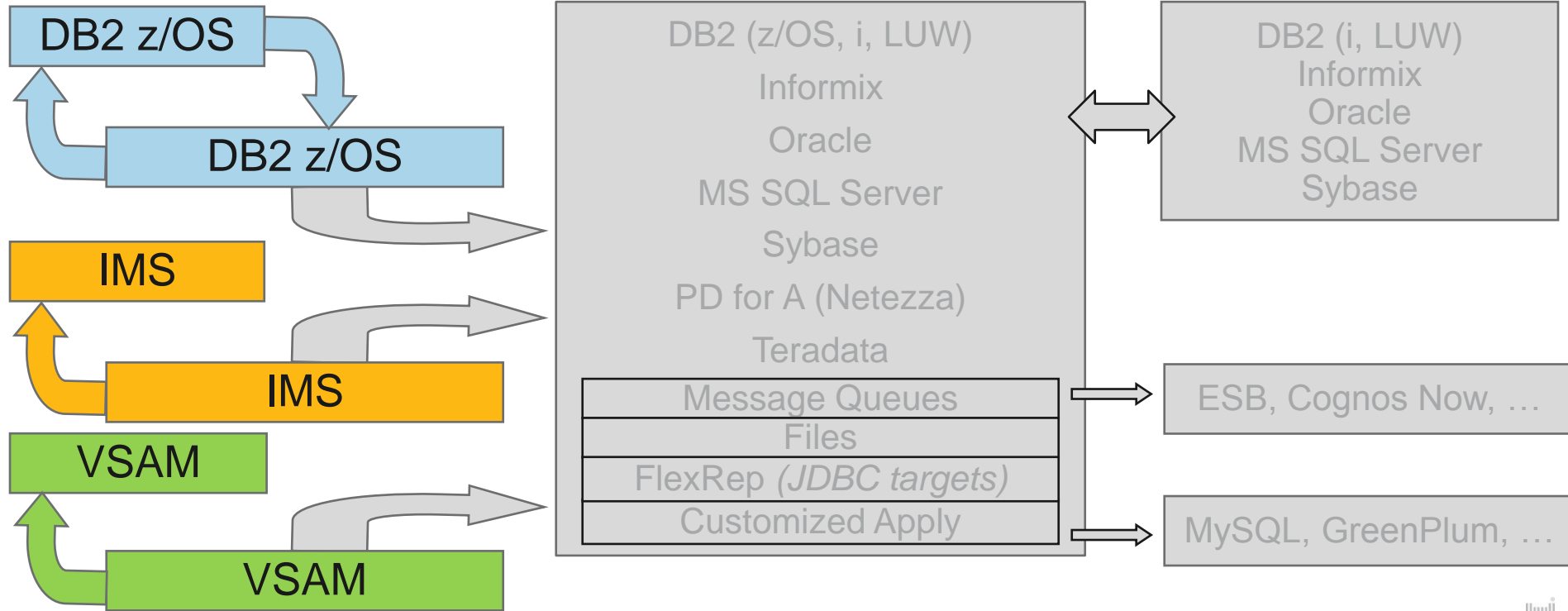
IIDR for DB2 -- IIDR for IMS -- IIDR for VSAM

Capture   **Push**   **Apply**

Log

Source Data → Push Engine → Apply Engine → Target Data
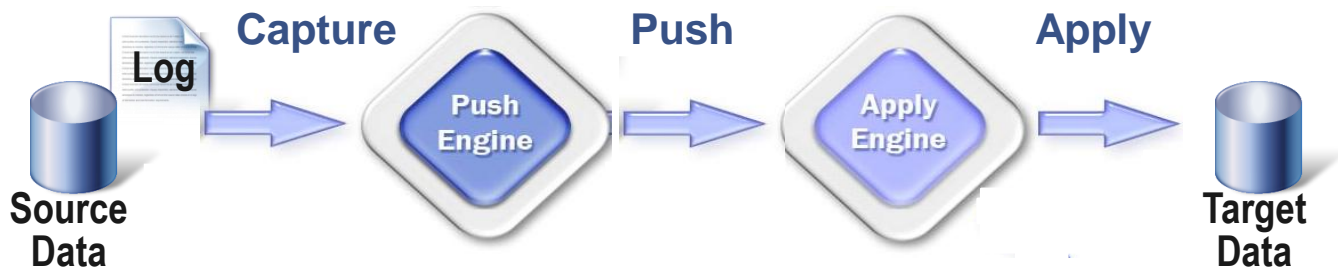
- **Log-based Capture**
  - Minimize impact on source environment and is recoverable

- **Apply using native database/file I/O**
  - No dependence on internal control blocks, storage, etc.

- **All share enterprise characteristics:**
  - Unit-of-Work aware
  - Capable of thousands of updates per second
  - Recoverable

IIDR for IMS: Two models in one product

- High speed, low latency IMS to IMS data replication spanning unlimited distances
  - Replication monitoring is built in as well as integration with Tivoli
  - External initial load of target DB is required
  - Conflicts will be detected and require a manual resolution

IMS Logs

IIDR for IMS

IMS databases

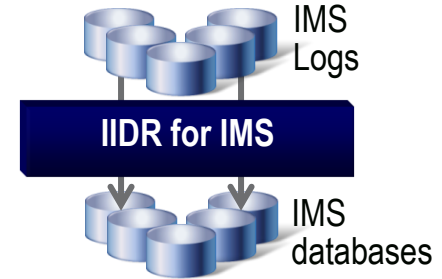- Heterogeneous IMS to non-IMS replication*** (when used with InfoSphere Data Replication's CDC Target Engines)
  - One IMS capture can target IMS and non-IMS
  - Synchronize IMS data with relational data warehouses, Hadoop, packaged apps, MDM, …
  - Leverages highly heterogeneous targeting capabilities of IIDR's CDC

IMS logs

IIDR for IMS

IIDR (CDC)

IMS databases

DBMS

ETL

hadoop

Flat files

Message Queue

SOA

A. IMS to IMS Data Replication
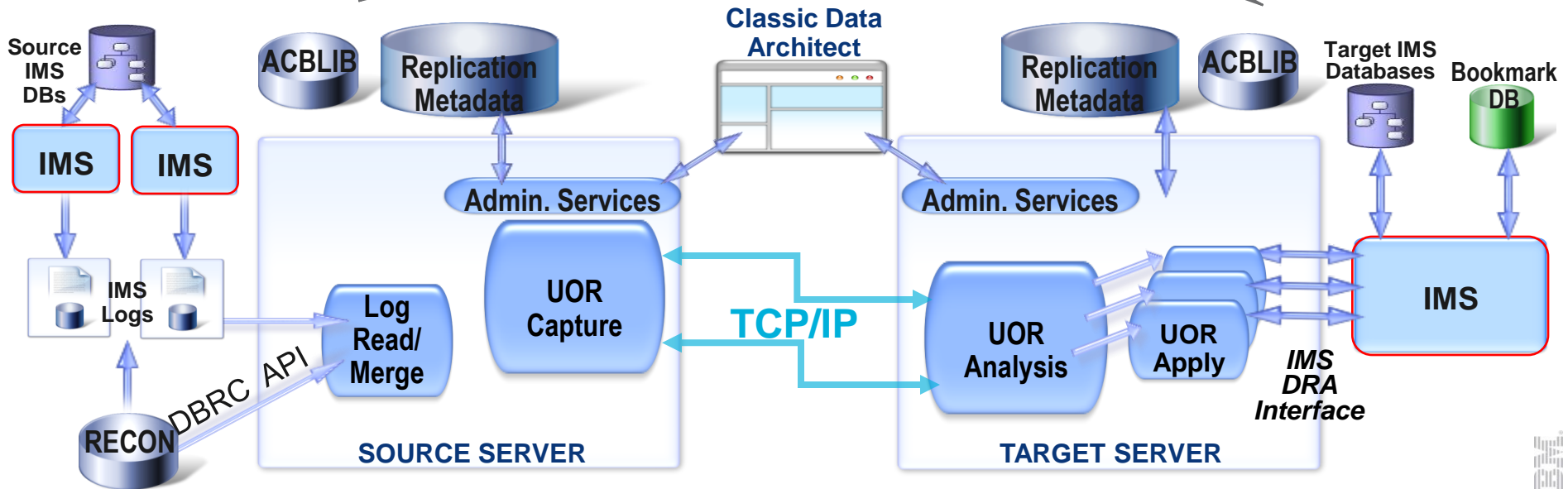
A. IMS to IMS Data

WHAT: **IMS logging**

ACTION: IMS replication logs developed for IMS v10 and higher specifically to support IMS to IMS and IMS to non-IMS data replication

Source IMS DBs

ACBLIB

Replication Metadata

Classic Data Architect

Replication Metadata

ACBLIB

Target IMS Databases

Bookmark DB

IMS   **IMS**

IMS Logs

Admin. Services

Admin. Services

UOR Capture

Log Read/ Merge

UOR Analysis

UOR Apply

IMS

TCP/IP

IMS DRA Interface

RECON   DBRC API

SOURCE SERVER

TARGET SERVER

18
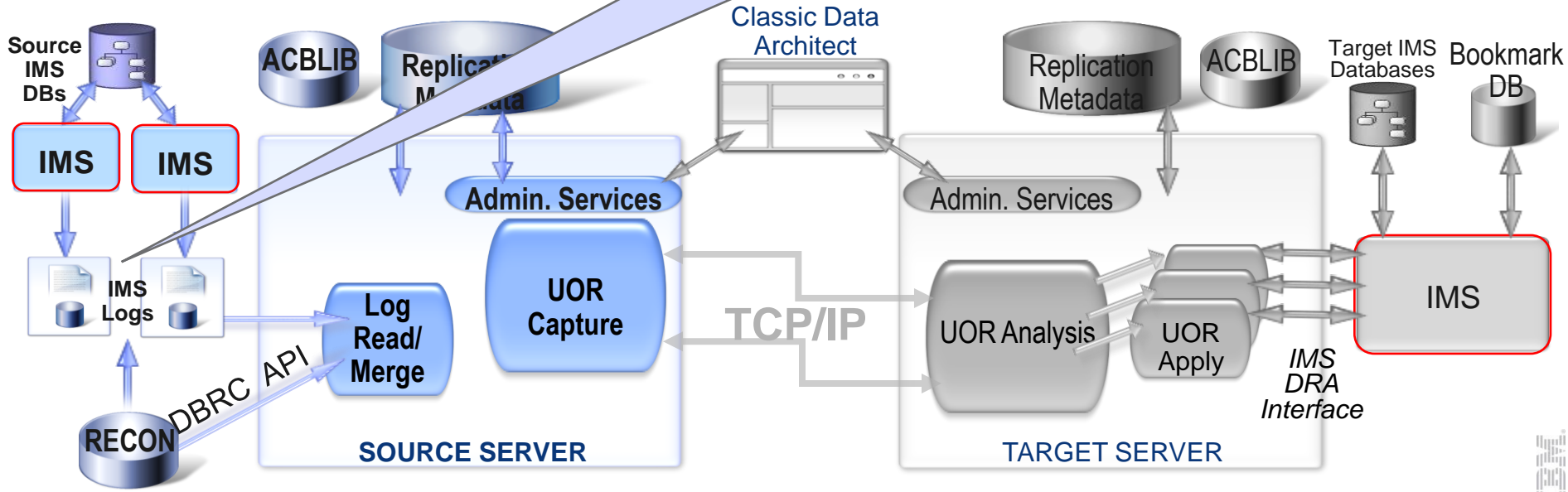
A. IMS to IMS Data Replication

WHAT: **IMS Log Reader**
ACTION: IMS log reader capable of capturing changes from BOTH local and remote logs. It ensures proper sequencing of committed changes for a single local IMS instance or for multiple logs in an IMSPLEX

Source IMS DBs

ACBLIB

Replication Metadata

IMS   **IMS**

IMS Logs

RECON   DBRC API

Log Read/ Merge   **UOR Capture**

SOURCE SERVER

TCP/IP

Replication Metadata   **ACBLIB**

Target IMS Databases   Bookmark DB

Admin. Services

UOR Analysis   UOR Apply

IMS DRA Interface

IMS

TARGET SERVER
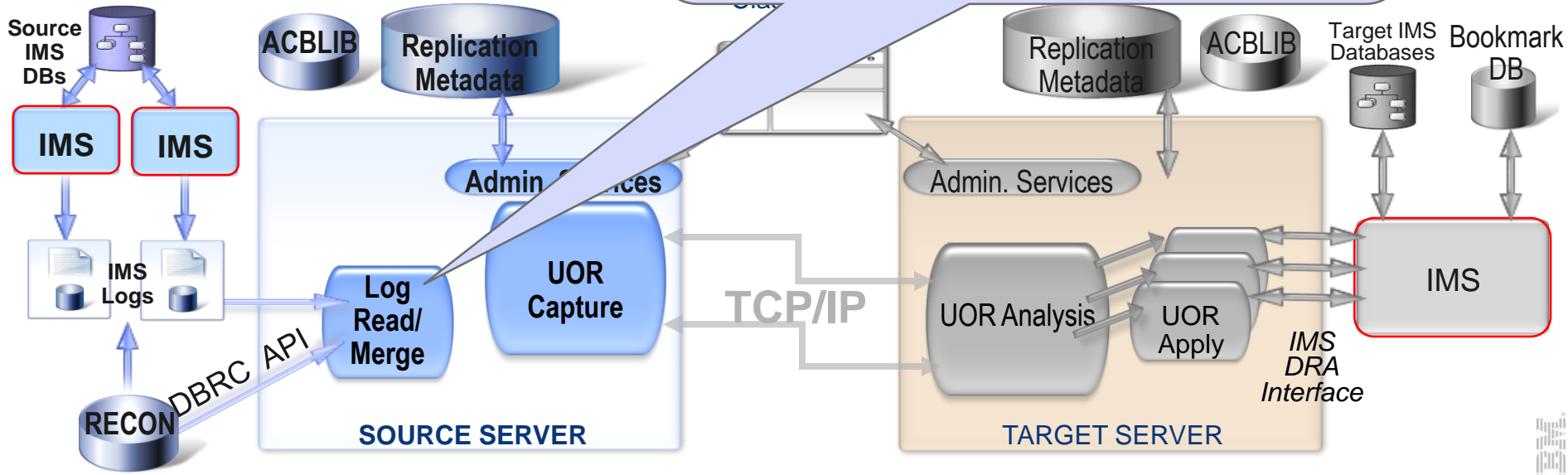
A. IMS to IMS Data Replication
Details of IMS Source Capture

IMS Databases

IMS Logs

IMS TM / DB*

Start Notification Exit Routine

WHAT: Classic routine associated with IMS's Partner Program Exit.
ACTION: Notify Classic Server when an IMS system starts.

TCP/IP Notification

Replication Metadata

Log Reader Service

BATCH DL/I

Batch Start-Stop Exit Routine

TCP/IP Notification

Change Stream Ordering

Capture Services

DBRC API

RECON

SOURCE SERVER

WHAT: Classic routine associated with IMS's Log Exit.
ACTION: Notify Classic Server when a Batch DL/I job starts or stops.

* includes BMP and DBCTL

A. IMS to IMS Data Replication

WHAT: Multi-threaded TCP/IP conversations between Source and Target Servers

ACTION:
- One "control" and one "data" conversation per subscription.
- Each Subscription represents a "consistency group."

21

A. IMS to IMS Data Replication

WHAT: **Dependency Analysis**
ACTION: Leverage multiple connections to the target for parallel writes when possible.

Source IMS DBs

ACBLIB

Replication Metadata

Classic Data Architect

Replication Metadata

ACBLIB

Target IMS Databases

Bookmark DB

IMS

IMS

IMS Logs

Admin. Services

Admin. Services

IMS

Log Read/ Merge

UOR Capture

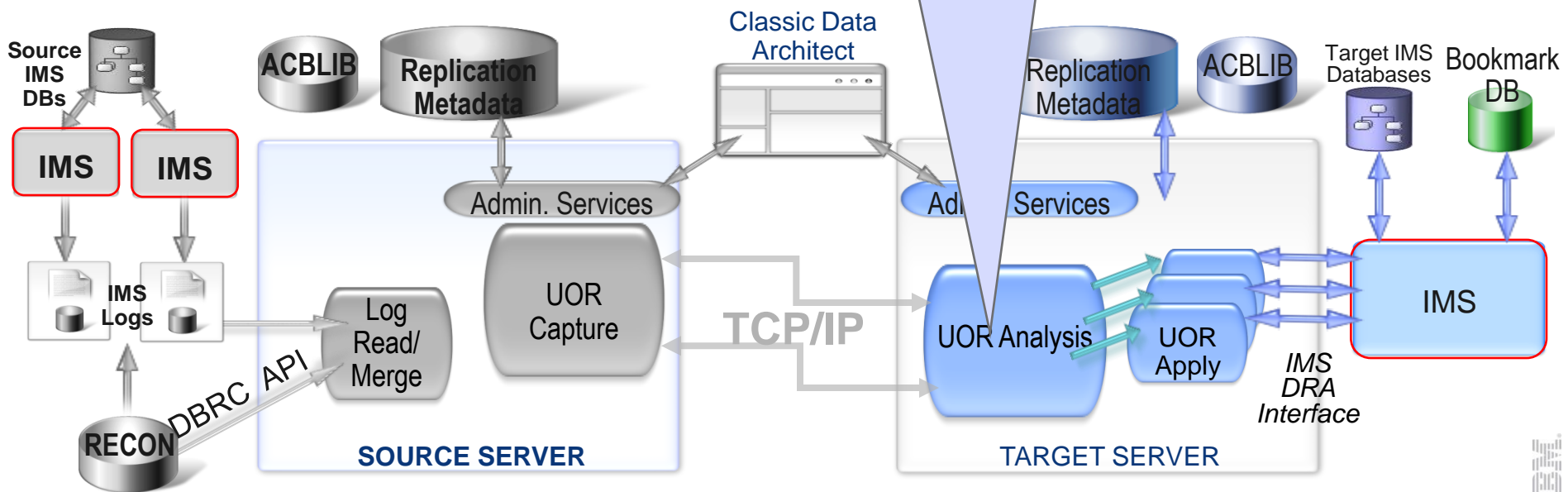UOR Analysis

UOR Apply

IMS DRA Interface

RECON

DBRC API

TCP/IP

SOURCE SERVER

TARGET SERVER

22

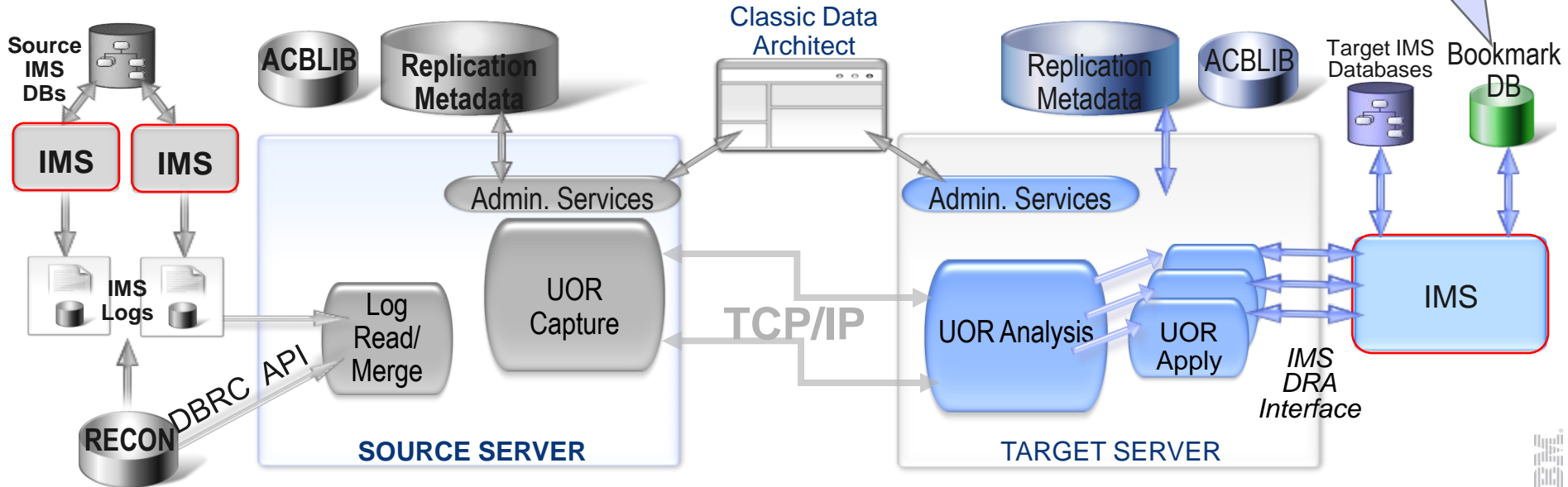# A. IMS to IMS Data Replication
Target Engine Details
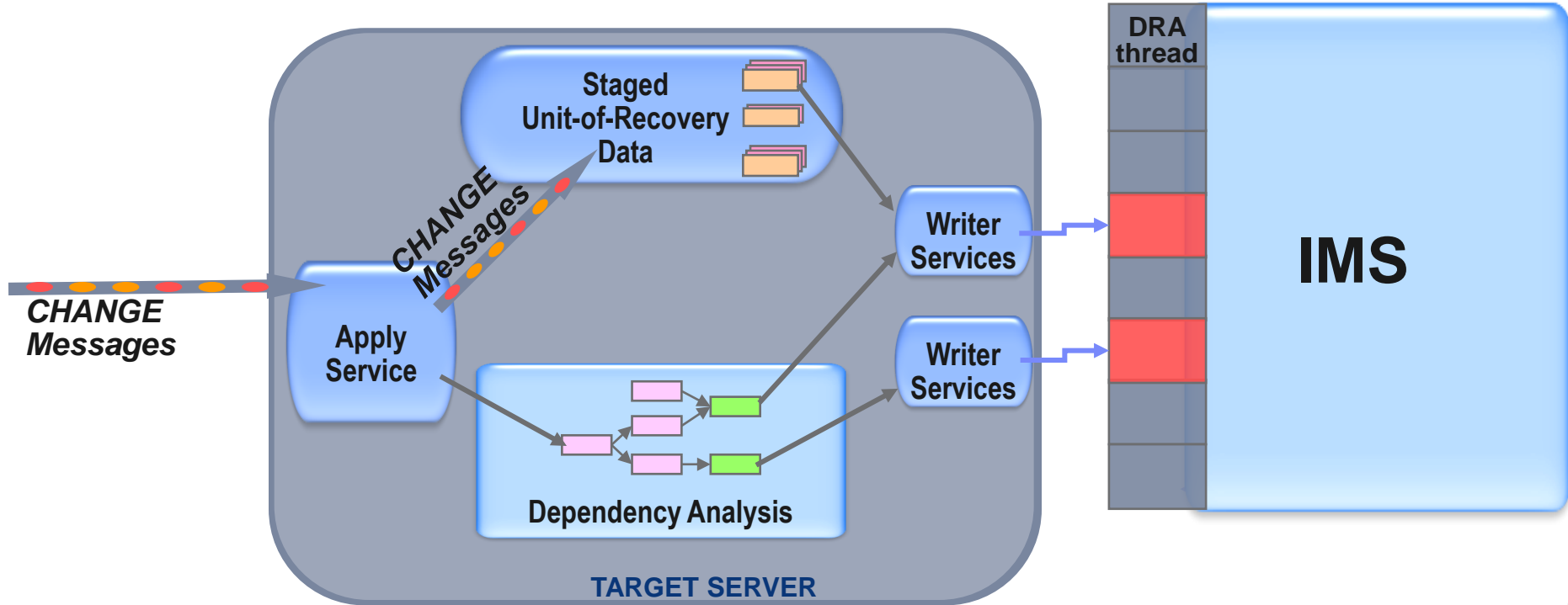
A. IMS to IMS Data Replication

WHAT: **Bookmark DB**
ACTION: New database required to hold bookmarks for each subscription.

Source IMS DBs

ACBLIB

Replication Metadata

Classic Data Architect

Replication Metadata

ACBLIB

Target IMS Databases

Bookmark DB

IMS   IMS

IMS Logs

Admin. Services

Admin. Services

Log Read/ Merge

UOR Capture

UOR Analysis

UOR Apply

IMS

RECON

DBRC API

TCP/IP

IMS DRA Interface

SOURCE SERVER

TARGET SERVER

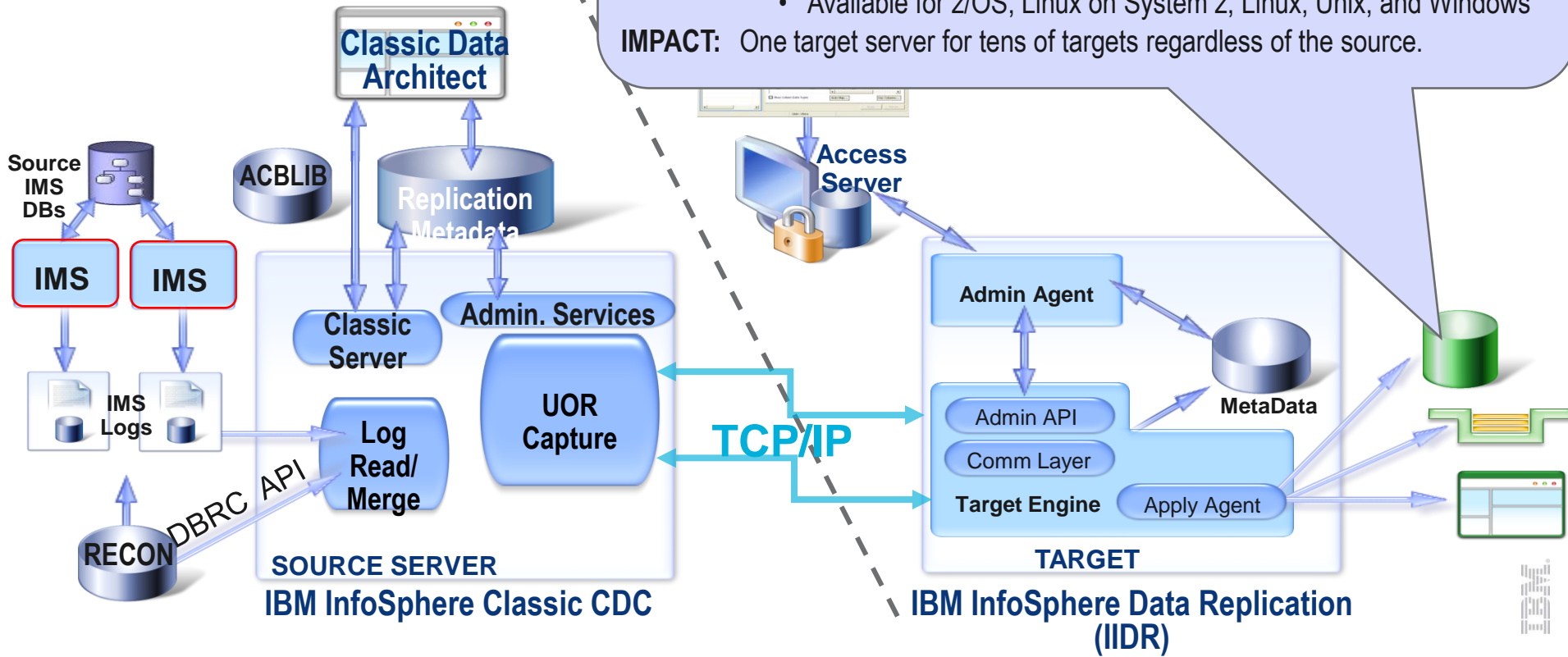# B. IMS to Non-IMS Data Replication
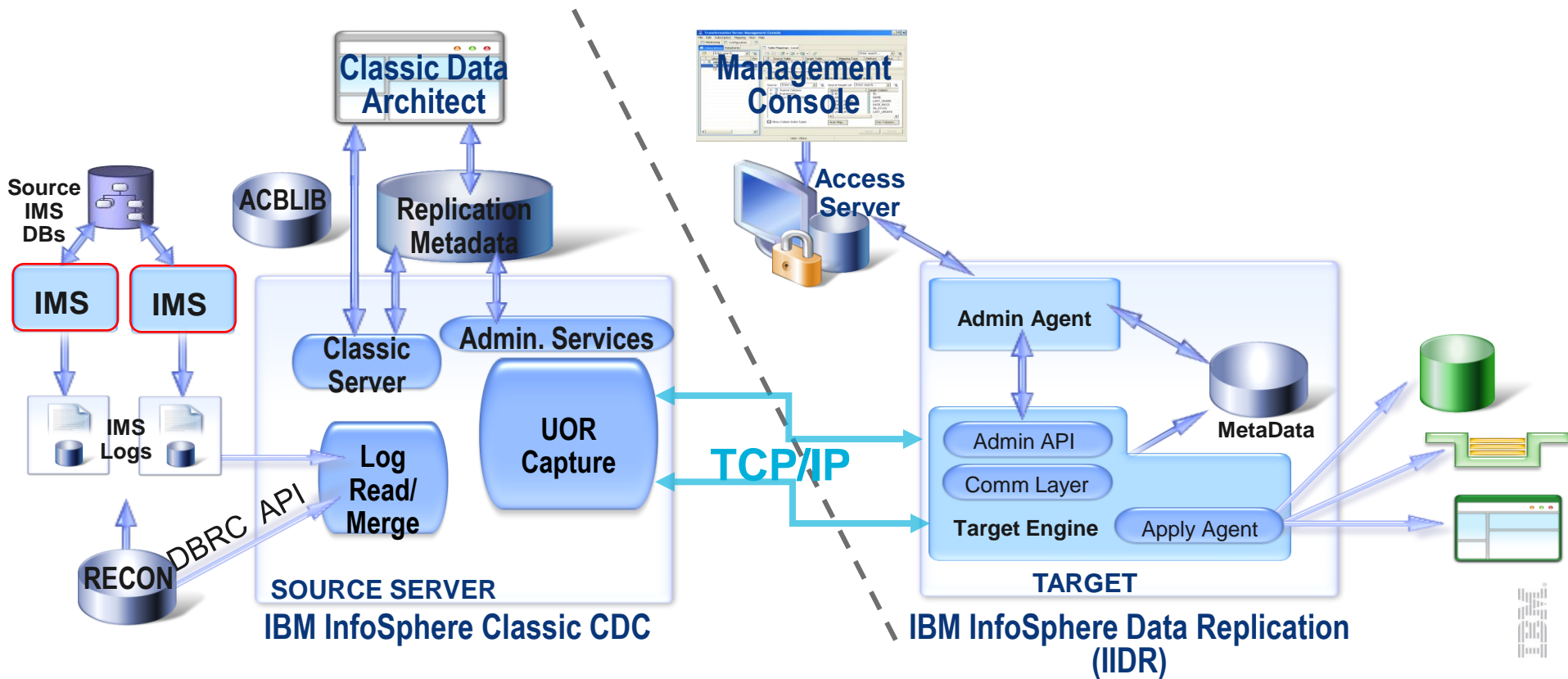
B.   IMS to Non-IMS

WHAT: IIDR Target Server

ACTION: Apply changes to the non-IMS replica(s) while maintaining restart information in a local bookmark for recovery purposes
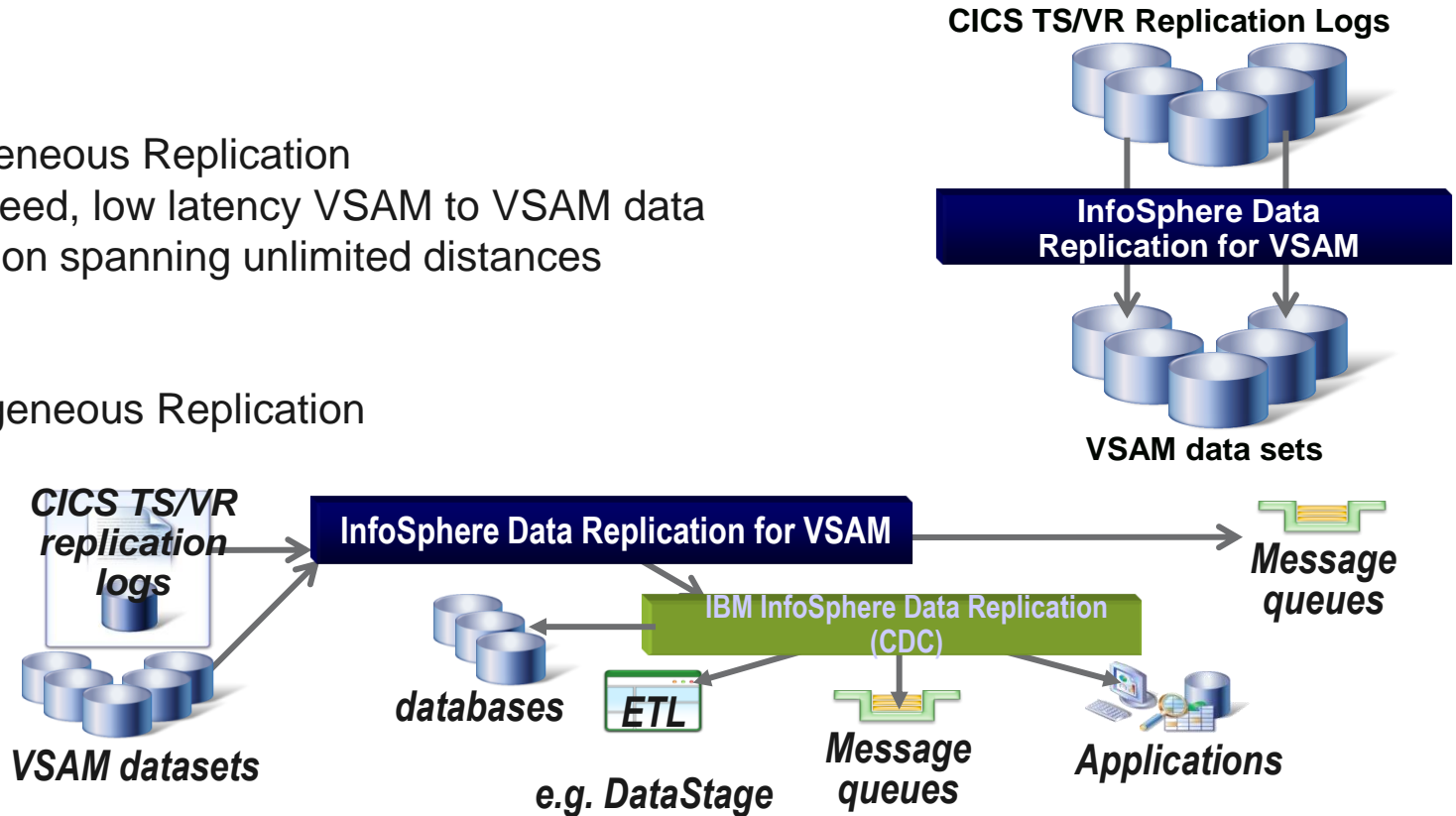- Additional target-based transformations can be applied
- Integration with many other InfoSphere solutions
- Available for z/OS, Linux on System z, Linux, Unix, and Windows

IMPACT: One target server for tens of targets regardless of the source.

Source IMS DBs

Classic Data Architect

ACBLIB

Replication Metadata

Access Server

Admin Agent

IMS   IMS

Classic Server

Admin. Services

MetaData

IMS Logs

Log Read/ Merge

UOR Capture

TCP/IP

Admin API

Comm Layer

Target Engine   Apply Agent

RECON

DBRC API

SOURCE SERVER

TARGET

IBM InfoSphere Classic CDC

IBM InfoSphere Data Replication (IIDR)

InfoSphere Data Replication for VSAM

- Homogeneous Replication
  High speed, low latency VSAM to VSAM data
  replication spanning unlimited distances

- Heterogeneous Replication

CICS TS/VR Replication Logs

InfoSphere Data Replication for VSAM

VSAM data sets

CICS TS/VR replication logs

InfoSphere Data Replication for VSAM

Message queues

IBM InfoSphere Data Replication (CDC)

databases

ETL

e.g. DataStage

Message queues

Applications

VSAM datasets

VSAM Replication Prerequisites
CICS TS and CICS VR Logging for Replication

- CICS v5.1 enhancements to provide a Replication Log
  - CICS Transaction Server provides logging for OLTP updates
  - CICS VSAM Recovery provides logging for BATCH updates

- The Replication Log contains …
  - UNDO records (autocommit bit always on for CICS VR)
  - REDO records (autocommit bit always on for CICS VR)
  - COMMIT/BACKOUT records (for CICS TS)
  - Tie-up records
  - File close records

InfoSphere Data Replication for VSAM for z/OS
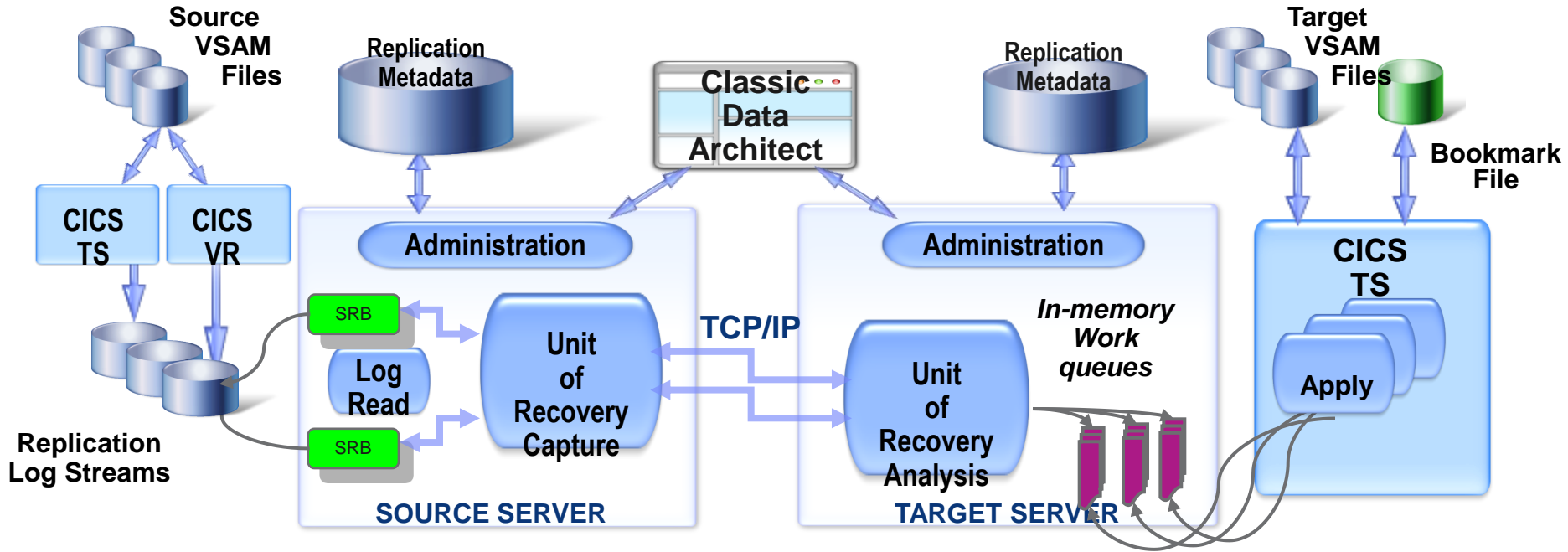
The GDPS Solution Family

Disaster Recovery to High Availability to Continuous Availability

A Roadmap Approach to Continuous Availability
Customers typically evolve over time

1. Continuous availability begins with a remote copy of your data
   - Home-grown procedures to manage planned site switches
   - Manually manage workload distribution for added value
     e.g. Run Mobile workloads on a second instance of data

2. Monitor Your Activities to Trigger Action
   - Drive actions based on actual system performance
     i.e. Replication latency, Response time, Utilization level, …

3. Automate Planned Workload Shifts
   - IBM Workload Lifeline provides a tool-based approach to site switches
   - Stand-alone, it effectively manages planned switches

4. End-to-End Automation
   - Full GDPS Active-Active Continuous Availability
     Manages planned and unplanned outages and workload balancing

GDPS Solutions Spanning the Availability Spectrum

CA of Data within a Data Center	CA with DR within a Metropolitan Region	DR across Extended Distance	Regional CA , DR across Extended Distance	CA, DR, & Workload Balancing across Extended Distance
GDPS/PPRC HM RPO=0 RTO secs for disk only	**GDPS/PPRC** RPO=0 RTO mins / RTO<1h (<20km) (>20km)	**GDPS/GM & GDPS/XRC** RPO secs RTO<1h	**GDPS/MGM & GDPS/MzGM** RPO=0,RTO mins/<1h & RPO secs, RTO<1h	**GDPS/Active-Active** RPO secs RTO secs
Single Data Center Applications remain active Continuous access to data in the event of a storage outage	***Two Data Centers*** Systems remain active Multi-site workloads can withstand site and/or storage failures	***Two Data Centers*** Rapid Systems DR seconds of data loss Disaster Recovery for out of region interruptions	***Three Data Centers*** High availability for site disasters Disaster recovery for regional disasters	***Two Active Data Centers*** Continuous availability Automatic workload switch in seconds; seconds of data loss

32

RPO – recovery point objective   RTO – recovery time objective

Continuous Availability for Mission Critical Workloads

- Shift from failover model to a nearly-continuous availability model
  - Multi-sysplex, multi-platform solution
    - "Recover my business rather than my platform"
    - Non-disruptive site switch for planned outages
    - Geographic dispersion to protect against regional outages

- Minimize cost and Optimize resource utilization
  - Automate recovery processes, minimize operator learning curve
  - Dynamic workload distribution based on resource availability

- Provide application level granularity
  - Match recovery objectives to the service levels of the workload
  - Reduce dependence on all-or-nothing approaches
    - e.g. complete disk mirroring, requiring extra network capacity.

CA, DR, & Workload Balancing across Extended Distance

GDPS/Active-Active

RPO secs
RTO secs

Two Active Data Centers

Continuous availability

Automatic workload switch in seconds; seconds of data loss

GDPS/Active-Active Sites Configurations

- Configurations
  1. **Active-Standby** **– Delivered (2011)**
  2. **Active-Query** **– Delivered (2013)**
  3. *Active-Active* *– Focusing on enablement with partitioned data for Phase 1*
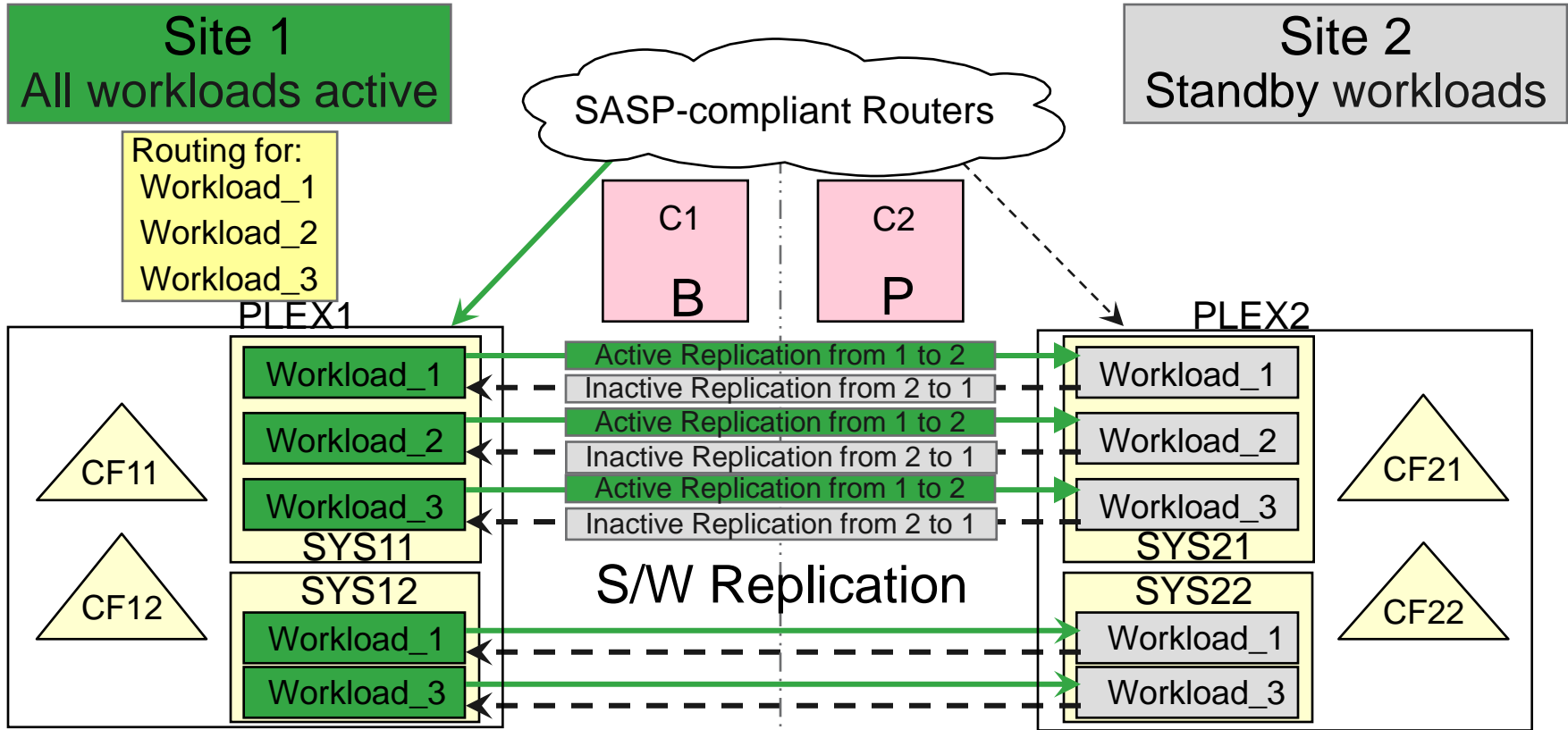
A configuration is specified on a workload basis
- A workload is the aggregation of these components

  - ***Software:*** user written applications (e.g., COBOL program) and the middleware run time environment (e.g., CICS region & DB2 subsystem)

  - ***Data:*** related set of objects that must preserve transactional consistency and optionally referential integrity constraints (e.g., DB2 Tables)

  - ***Network connectivity***: one or more TCP/IP addresses & ports (e.g., 10.10.10.1:80)

Sample Environment 1

Site 1 is all Active, Site 2 is all Stand-by

Site 1
All workloads active

Site 2
Standby workloads

SASP-compliant Routers

Routing for:
Workload_1
Workload_2
Workload_3

C1
B

C2
P

PLEX1

PLEX2

Workload_1
Active Replication from 1 to 2
Inactive Replication from 2 to 1
Workload_1

Workload_2
Active Replication from 1 to 2
Inactive Replication from 2 to 1
Workload_2

Workload_3
Active Replication from 1 to 2
Inactive Replication from 2 to 1
Workload_3

SYS11

SYS21

CF11

CF21

CF12

SYS12

SYS22

CF22

Workload_1

Workload_1

Workload_3

Workload_3

S/W Replication

Scenario 1 –
Start Workload / Replication / Routing

Action :   From GDPS:
        Start the workloads in **both** sites
        Start the replication from site 1 to site 2 (active to stand-by)
        Start the replication from site 2 to site 1 (prep for site switch)
        Start the routing of transactions to site1

We see :   On the GDPS panel:
        The start of Workloads (subsystems)
        Scripts to start replication
        Scripts to start routing transactions to site1
    SDF screen to check the GDPS actions
    TEP interface to check the replication and workload status

Sample Environment 2
## *Distinct Active workloads on each site*
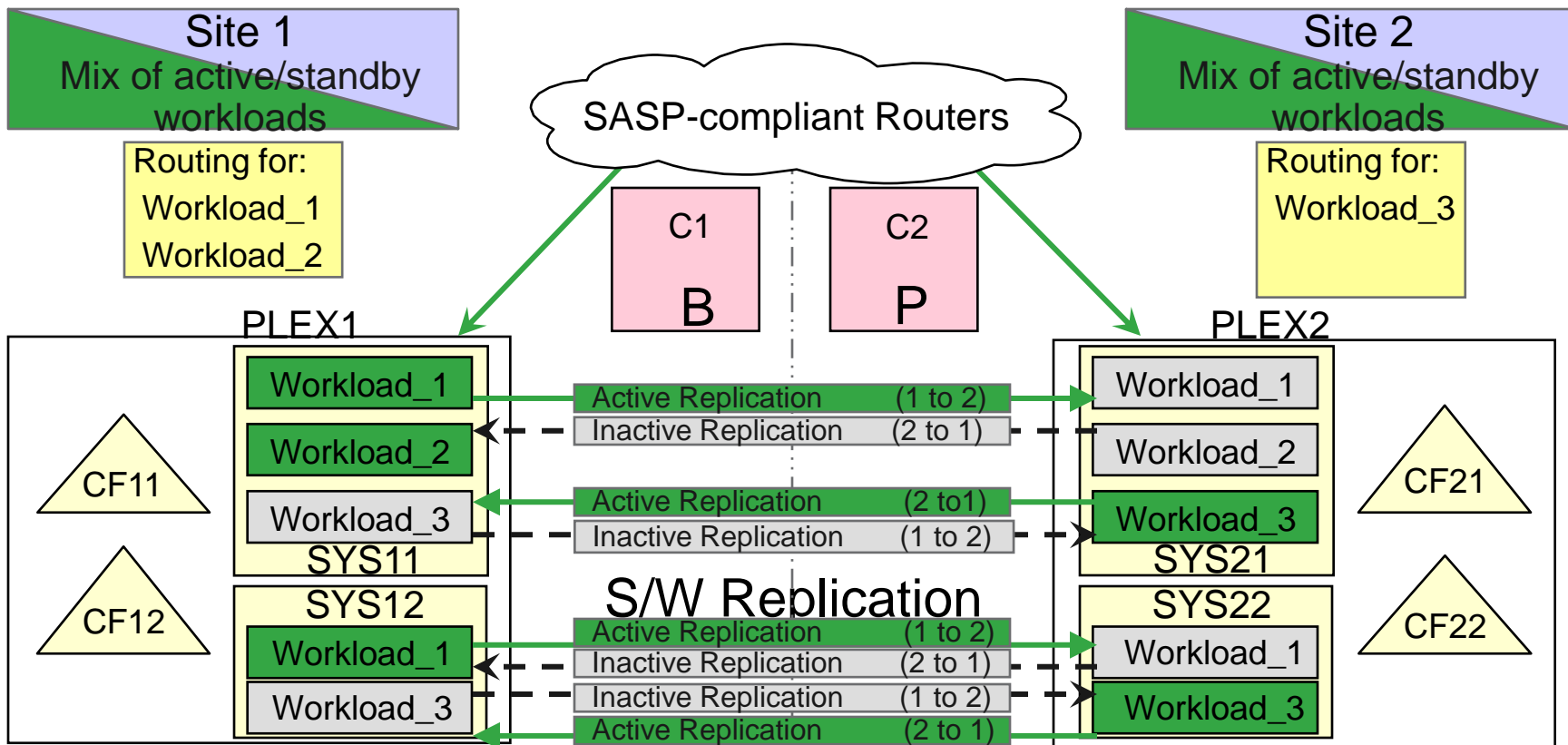
Sample Scenario ---
Both Site1 and Site2 can be "active"!

- For example:
  - DB2 workload Active on one, IMS workload active on another
  - IMS workload using Database "A" active on one, with IMS workload using Database "B" active on another

- Implications:
  - Data will be actively replicating:
    - From Site1 to Site2
    - From Site2 to Site1
  - No conflicts will occur as there is no update overlap in the data replicated

- GDPS recognizes an unavailable site is also a Standby site for other workloads
  - Replication for a second workload may stop during the outage
  - Catch-up for the second workload will occur upon restart after the outage

Sample Environment 3
Add "load balanced" query workloads
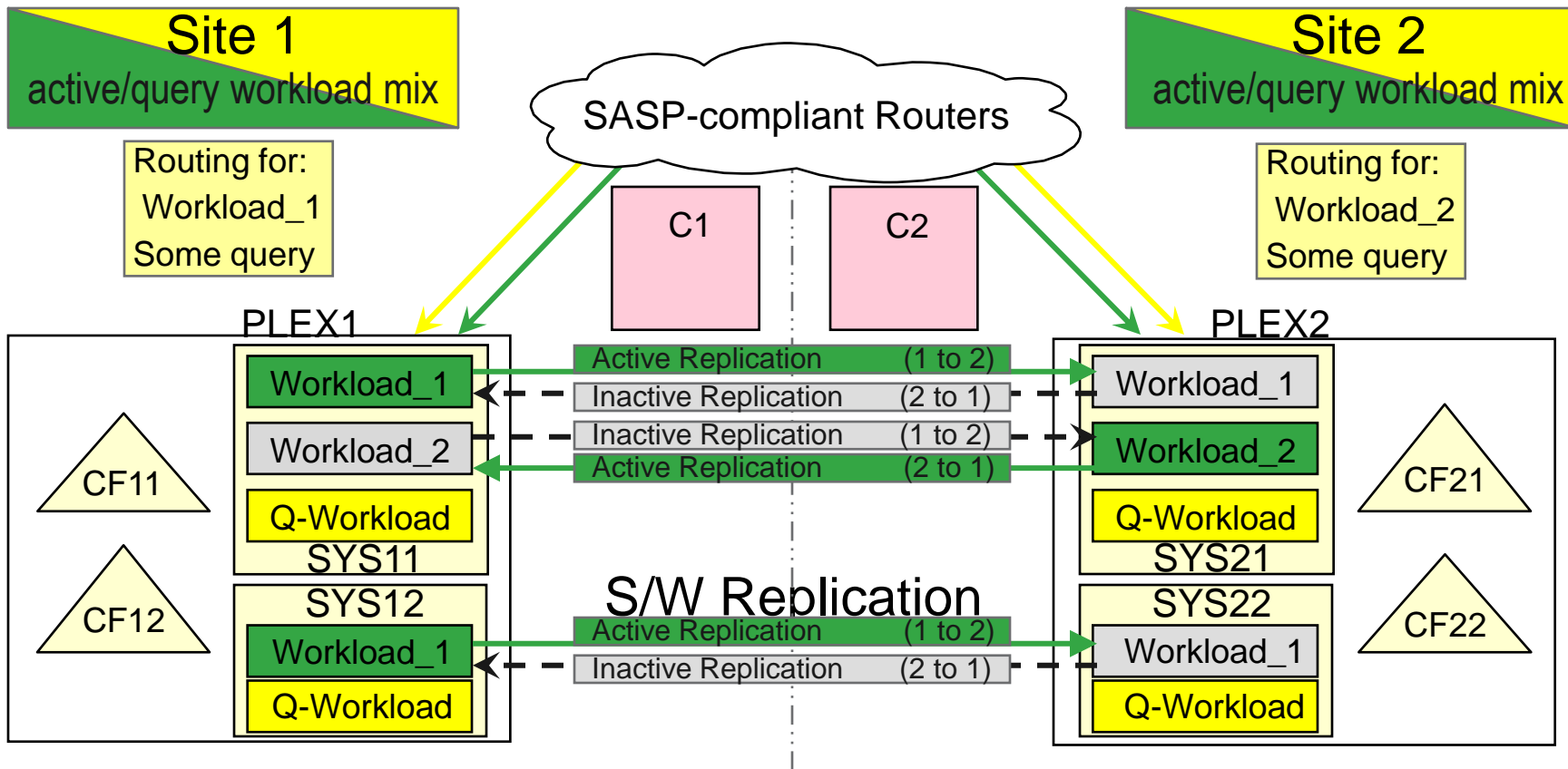
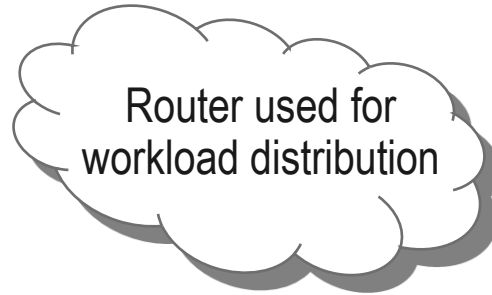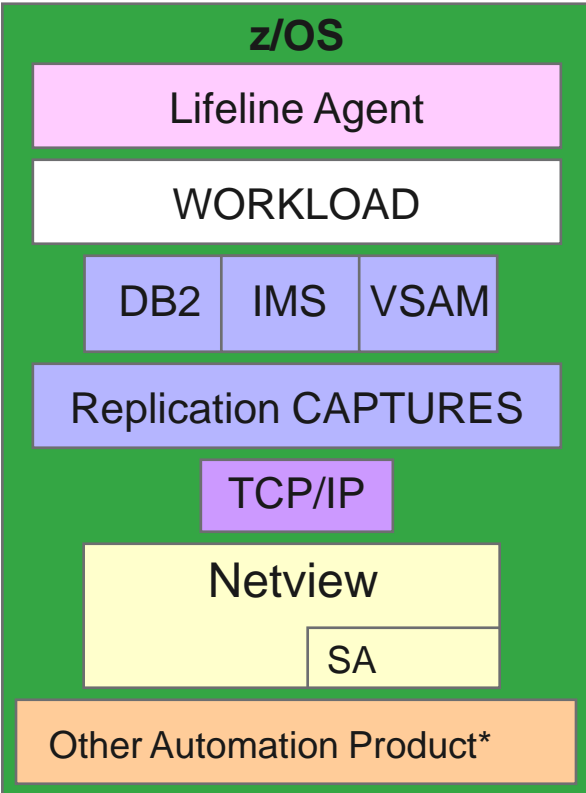# Sample Scenario ---
Add "load balanced" query workloads

- Distribute query transactions based on:
  - Availability of resources
  - Latency of replicated data

- Better utilization of resources for many rapidly growing workloads
  - Ratio of query to update is very high in:
    - Most mobile apps
    - Real-time analytics
    - Self-service applications
  - Optimize the performance of update transactions
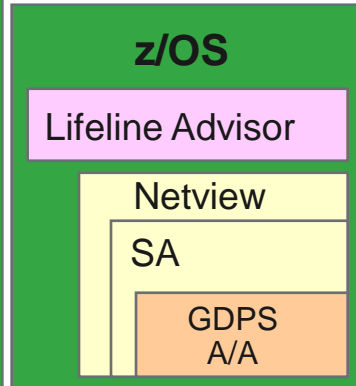   - Ensure resource availability for those transactions that manage data
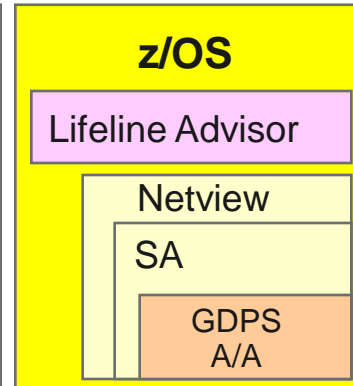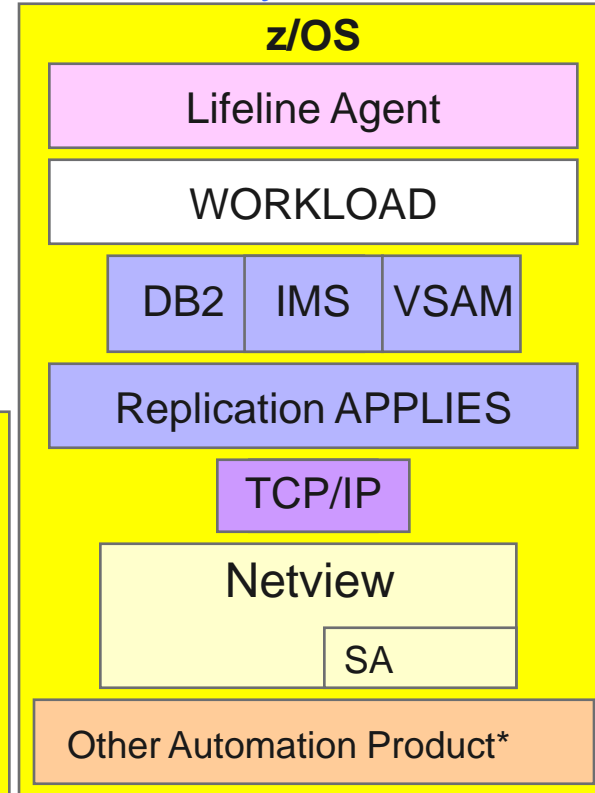
High Level Architecture

Active Production

z/OS
- Lifeline Agent
- WORKLOAD
- DB2 | IMS | VSAM
- Replication CAPTURES
- TCP/IP
- Netview
  - SA
- Other Automation Product*

Standby Production

z/OS
- Lifeline Agent
- WORKLOAD
- DB2 | IMS | VSAM
- Replication APPLIES
- TCP/IP
- Netview
  - SA
- Other Automation Product*

Router used for workload distribution

Primary Controller

z/OS
- Lifeline Advisor
- Netview
  - SA
    - GDPS A/A

Backup Controller

z/OS
- Lifeline Advisor
- Netview
  - SA
    - GDPS A/A

41

Wrap-Up

Thank you!

43