Extending z/OS Mainframe Workload Availability with GDPS/Active-Active

John Thompson

IBM

jgthomps@us.ibm.com

March 5th, 2015

Session 17022

CELEBRATING 60 YEARS OF SHARE
Influencing IT Since 1955

SHARE in Seattle 2015

Disclaimer

- IBM's statements regarding its plans, directions, and intent are subject to change or withdrawal without notice at IBM's sole discretion.

- Information regarding potential future products is intended to outline our general product direction and it should not be relied on in making a purchasing decision.

- The information mentioned regarding potential future products is not a commitment, promise, or legal obligation to deliver any material, code or functionality. Information about potential future products may not be incorporated into any contract. The development, release, and timing of any future features or functionality described for our products remains at our sole discretion.

- Performance is based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput or performance that any user will experience will vary depending upon many factors, including considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve results similar to those stated here

Agenda

- Background

- Concepts and configurations

- Components

- Scenarios

- Disk replication integration
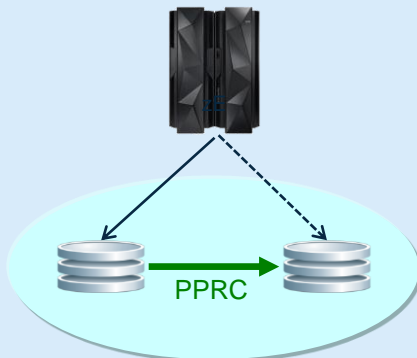
- Recent enhancements

- Summary

Multiple products meeting various availability objectives

GDPS®/PPRC HM[1]

Near-continuous availability of data within a data center

Single data center

Applications can remain active
Near-continuous access to data in the event of a storage subsystem outage

PPRC

Recovery Point Objective equals 0
and
Recovery Time Objective equals 0

GDPS/PPRC

Near-continuous availability (CA) and disaster recovery (DR) within a metropolitan region

Two data centers
Systems can remain active
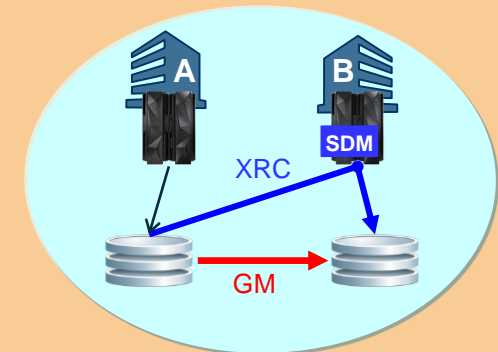Multisite workloads can withstand site and storage failures

A   B

PPRC
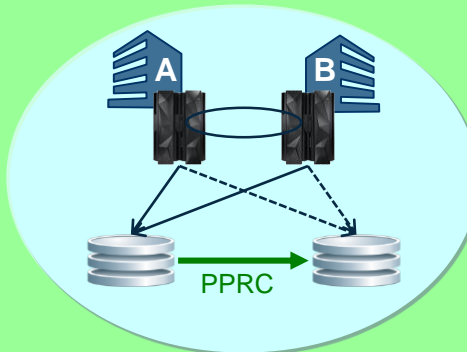
DR RPO equals 0 and RTO is less than 1 hour or
CA RPO equals 0 and RTO minutes

GDPS/GM[2] and GDPS/XRC[3]

Disaster recovery at extended distance

Two data centers
More rapid systems disaster recovery with "seconds" of data loss
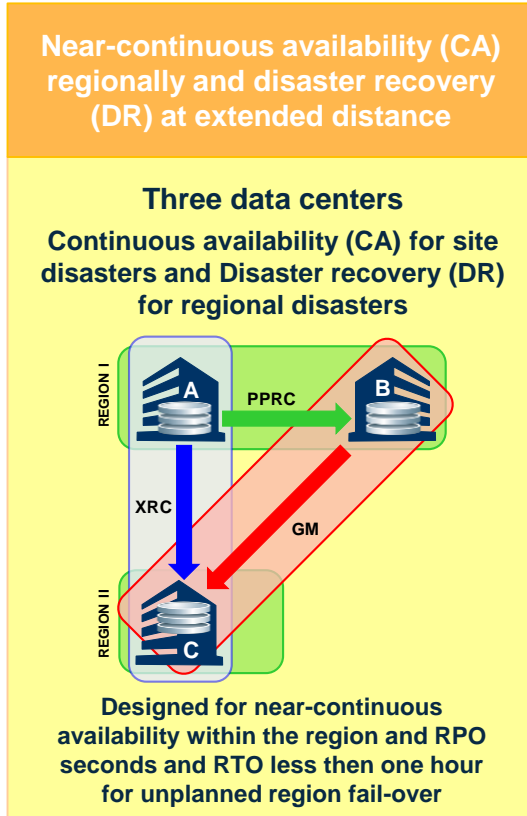Disaster recovery for out-of-region interruptions

A   B
SDM
XRC
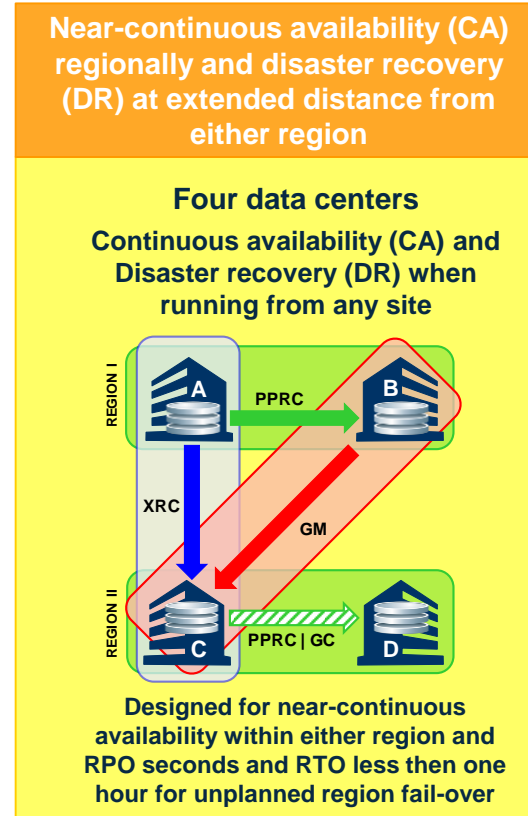GM

RPO seconds and RTO less than 1 hour

[1]Peer-to-peer remote copy (PPRC), Metro Mirror HyperSwap Manager (HM)   [2]Global Mirror (GM)   [3]Extended Remote Copy (XRC), zGlobal Mirror

Multiple products meeting various availability objectives (continued)

GDPS/MGM[1] 3-site & GDPS/MzGM[2] 3-site

Near-continuous availability (CA) regionally and disaster recovery (DR) at extended distance

Three data centers

Continuous availability (CA) for site disasters and Disaster recovery (DR) for regional disasters

REGION I — A — PPRC — B

XRC — GM

REGION II — C

Designed for near-continuous availability within the region and RPO seconds and RTO less then one hour for unplanned region fail-over

GDPS/MGM[1] 4-site & GDPS/MzGM[2] 4-site(SOD)

Near-continuous availability (CA) regionally and disaster recovery (DR) at extended distance from either region

Four data centers

Continuous availability (CA) and Disaster recovery (DR) when running from any site

REGION I — A — PPRC — B

XRC — GM

REGION II — C — PPRC | GC — D

Designed for near-continuous availability within either region and RPO seconds and RTO less then one hour for unplanned region fail-over

[1]Metro Global Mirror (MGM)   [2]Metro z/OS Global Mirror (MzGM)

From high availability to continuous availability

| GDPS/PPRC	GDPS/XRC or GDPS/GM	GDPS/Active-Active
Near Continuous Availability model	Failover model	Near Continuous Availability model
Recovery time = 2 minutes	Recovery time < 1 hour	Recovery time < 1 minute
Distance < 20 KM	Unlimited distance	Unlimited distance

GDPS/Active-Active is for mission critical workloads that have stringent recovery objectives that can not be achieved using existing GDPS solutions.

- RTO approaching zero, measured in seconds for unplanned outages
- RPO approaching zero, measured in seconds for unplanned outages
- Non-disruptive site switch of workloads for planned outages
- At any distance

Active-Active is NOT intended to substitute for local availability solutions such as Parallel SYSPLEX

Active/Active concept

- **Two or more sites, separated by _unlimited_ distances, running the same applications and having the same data to provide:**
  - Cross-site Workload Balancing
  - Continuous Availability
  - Disaster Recovery

- **Data at geographically dispersed sites kept in sync via replication**

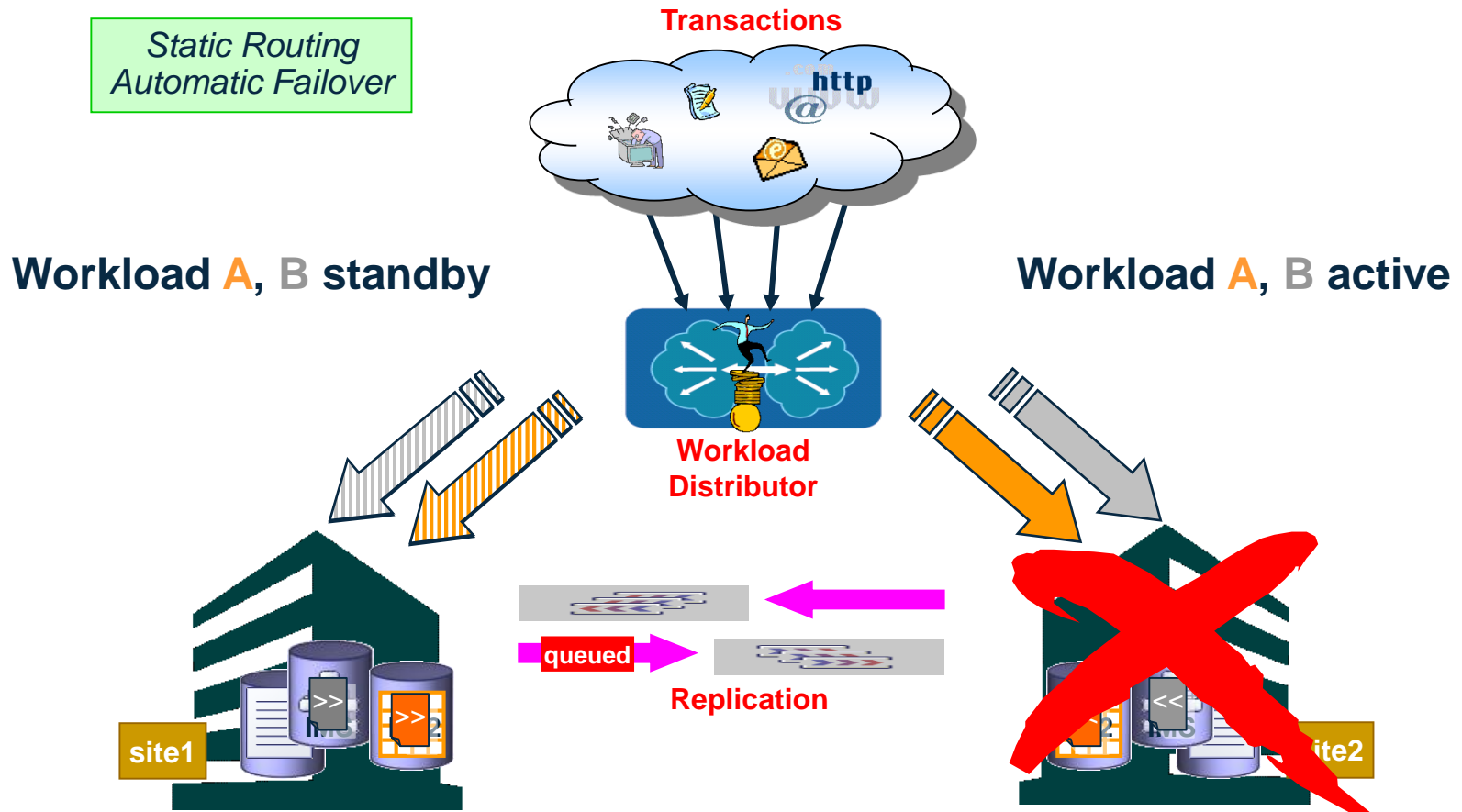**Transactions**

Workload Distributor
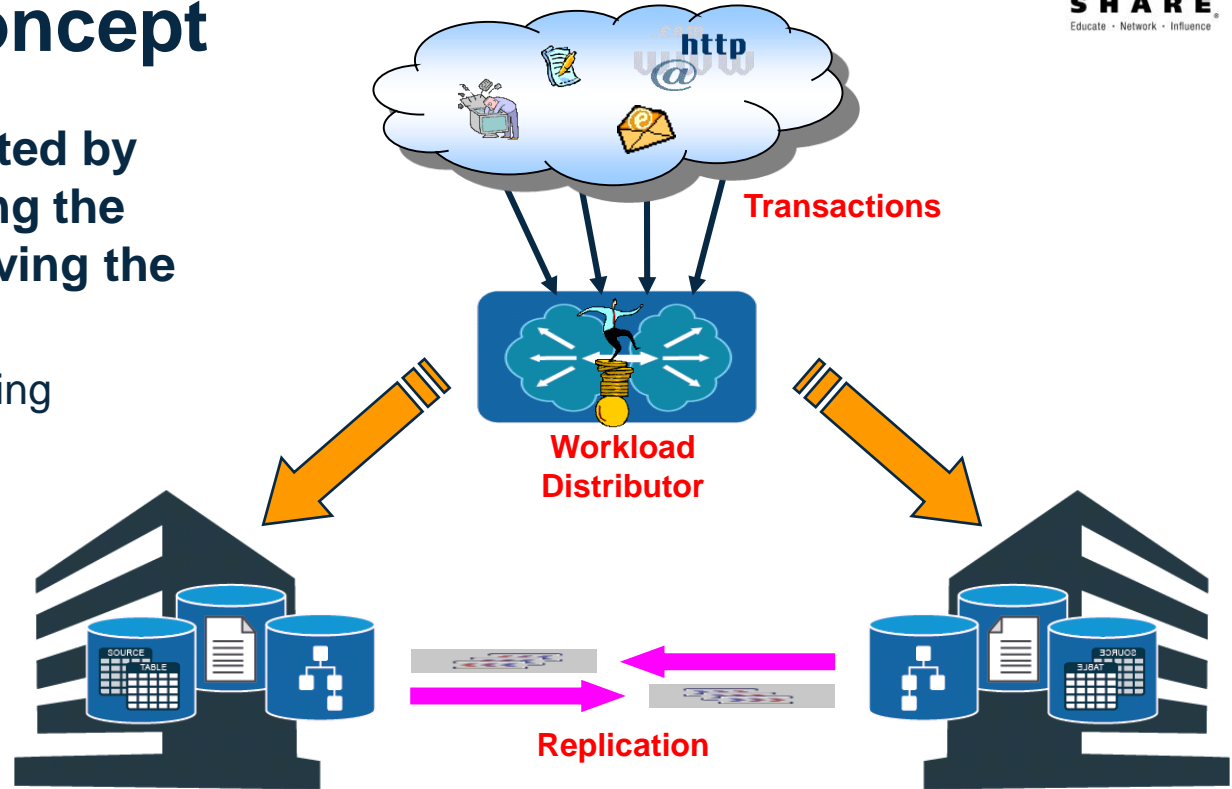
Replication

Transactions are routed to one of many replicas, depending upon workload weight and latency constraints; extends workload balancing to SYSPLEXs across multiple sites

Monitoring spans the sites and now becomes an essential element of the solution for site health checks, performance tuning, etc
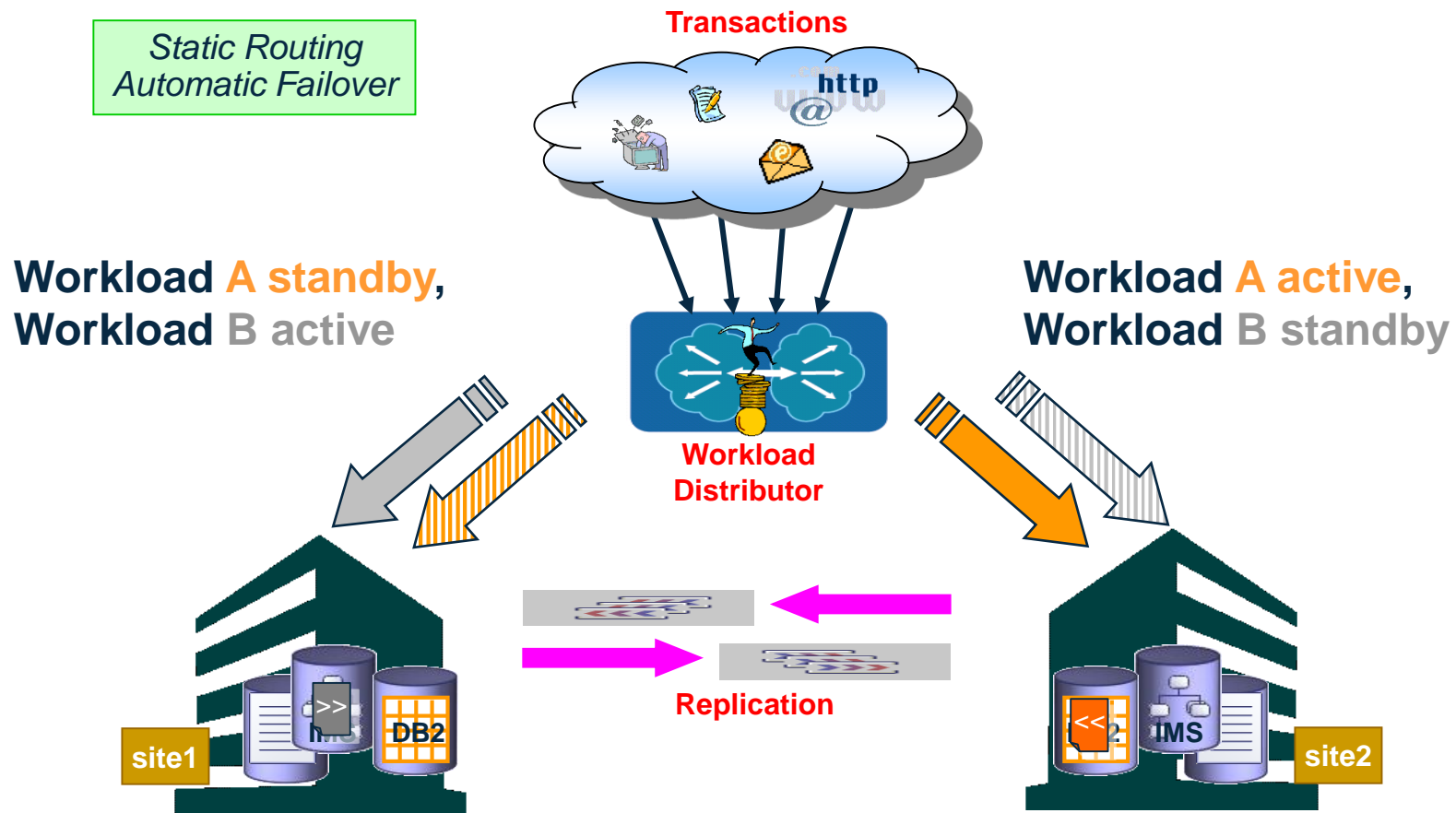
Active/Active Sites configurations

- Configurations
  - Active/Standby – GA date 30th June 2011
  - Active/Query – GA date 31st October 2013
  - Active/Active – intended direction
- A configuration is specified on a workload basis
  - Update workload
    - Currently only run in what is defined as an Active/Standby configuration
    - Some, but not necessarily all, transactions are update transactions
  - Query workload
    - Run in what is defined as an Active/Query configuration
    - Must not perform any updates to the data
    - Associated with / shares data with an update workload

Active/Standby configuration

Static Routing
Automatic Failover

Transactions

http

Workload A, B standby

Workload A, B active

Workload
Distributor

queued

Replication

site1

site2

Active/Standby Configuration – both sites active for individual workloads

Static Routing
Automatic Failover

Transactions

Workload **A standby**,
Workload **B active**

Workload **A active**,
Workload **B standby**

Workload Distributor

Replication

site1

DB2

IMS

site2

Active/Query configuration

Transactions

Workload B (grey) is in **active/query** configuration
- using same data as Appl A
- active to both site1 & site2, but favor site1
- routed according to [A] latency policy
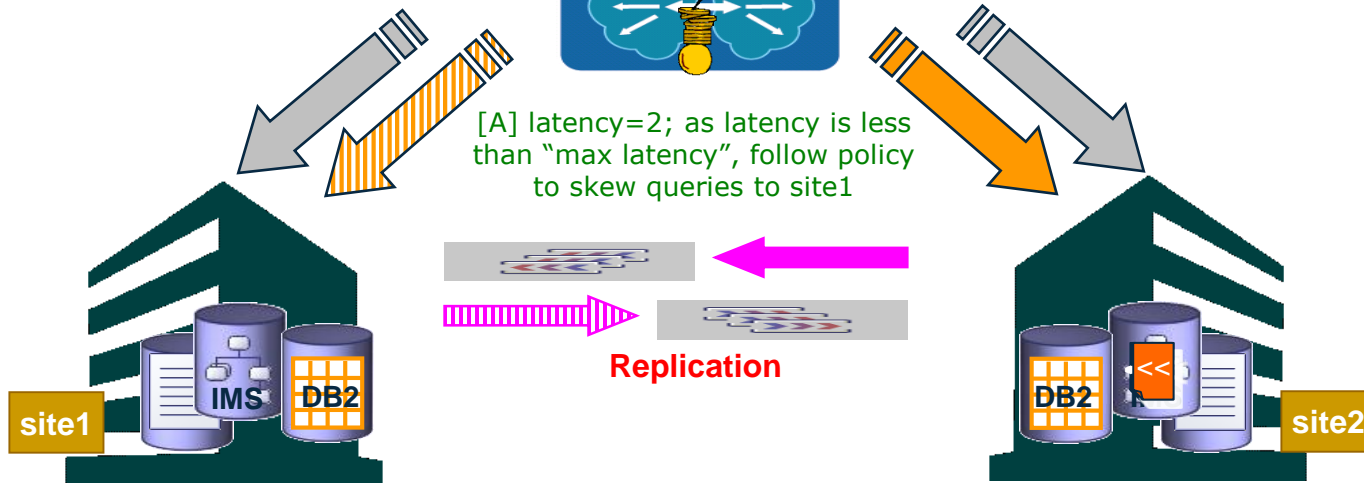- policy for query routing: max latency 5, reset latency 3

Workload A (gold) is in **active/standby** configuration
- performing updates in active site [site2]

Workload Distributor

[A] latency=2; as latency is less than "max latency", follow policy to skew queries to site1

Replication

site1

IMS   DB2

DB2   <<   site2

Components of a GDPS/Active-Active environment

- **Two Production Sysplex environments (also referred to as sites…and now regions) in different locations**
  - One active, one standby – for each defined workload
  - Software-based replication between the two sysplexes/sites
    - DB2, IMS and VSAM data is supported
- **Two Controller Systems**
  - Primary/Backup
  - Typically one in each of the production locations, but there is no requirement that they are co-located in this way
- **Workload balancing/routing switches**
  - Must be Server/Application State Protocol compliant (SASP)
    - RFC4678 describes SASP
  - Examples of SASP-compliant switches/routers
    - Cisco Catalyst 6500 Series Switch Content Switching Module
    - F5 Big IP Switch
    - Citrix NetScaler Appliance
    - Radware Alteon Application Switch (bought Nortel appliance line)

Software in a GDPS/Active-Active environment

- GDPS/Active-Active
  - Workload management - start/stop components of a workload including software replication
  - Routing management - start/stop routing of transactions to a site
  - System and Server management - STOP (graceful shutdown) of a system, LOAD, RESET, ACTIVATE, DEACTIVATE the LPAR for a system, and capacity on demand actions such as CBU/OOCoD
  - Monitoring the environment and alerting for unexpected situations

- IBM Tivoli NetView Monitoring for GDPS
  - provides automation and monitoring functions
  - provides an execution environment for GDPS/Active-Active

- System Automation for z/OS
  - Remote communications capability to enable GDPS to manage sysplexes from outside the sysplex
  - System Automation infrastructure for workload and server management

# Software in a GDPS/Active-Active environment (continued)

- IBM Multi-site Workload Lifeline for z/OS
  - provides information to the external load balancers on where to send connections
  - provides information to GDPS on the health of the environment
- Middleware – DB2, IMS, CICS…
- Replication Software
  - IBM InfoSphere Data Replication for DB2 for z/OS (IIDR for DB2)
  - IBM InfoSphere Data Replicator for IMS for z/OS (IIDR for IMS)
  - IBM InfoSphere Data Replicator for VSAM for z/OS (IIDR for VSAM)

- Optionally the Tivoli OMEGAMON XE suite of monitoring products

Definition of an Active/Active workload

- A workload is the aggregation of these components

  - *Software:* user written applications (eg: COBOL programs) and the middleware run time environment (eg: CICS regions, InfoSphere Replication Server instances and DB2 subsystems)

  - *Data:* related set of objects that must preserve transactional consistency and optionally referential integrity constraints (eg: DB2 Tables, IMS Databases)

  - *Network connectivity*: one or more TCP/IP addresses & ports

Software – deeper insight

- All components of a Workload should be defined in SA* as
  - One or more Application Groups (APG)
  - Individual Applications (APL)

- The Workload itself is defined as an Application Group

- SA z/OS keeps track of the individual members of the Workload's APG and reports a "compound" status to the A/A Controller
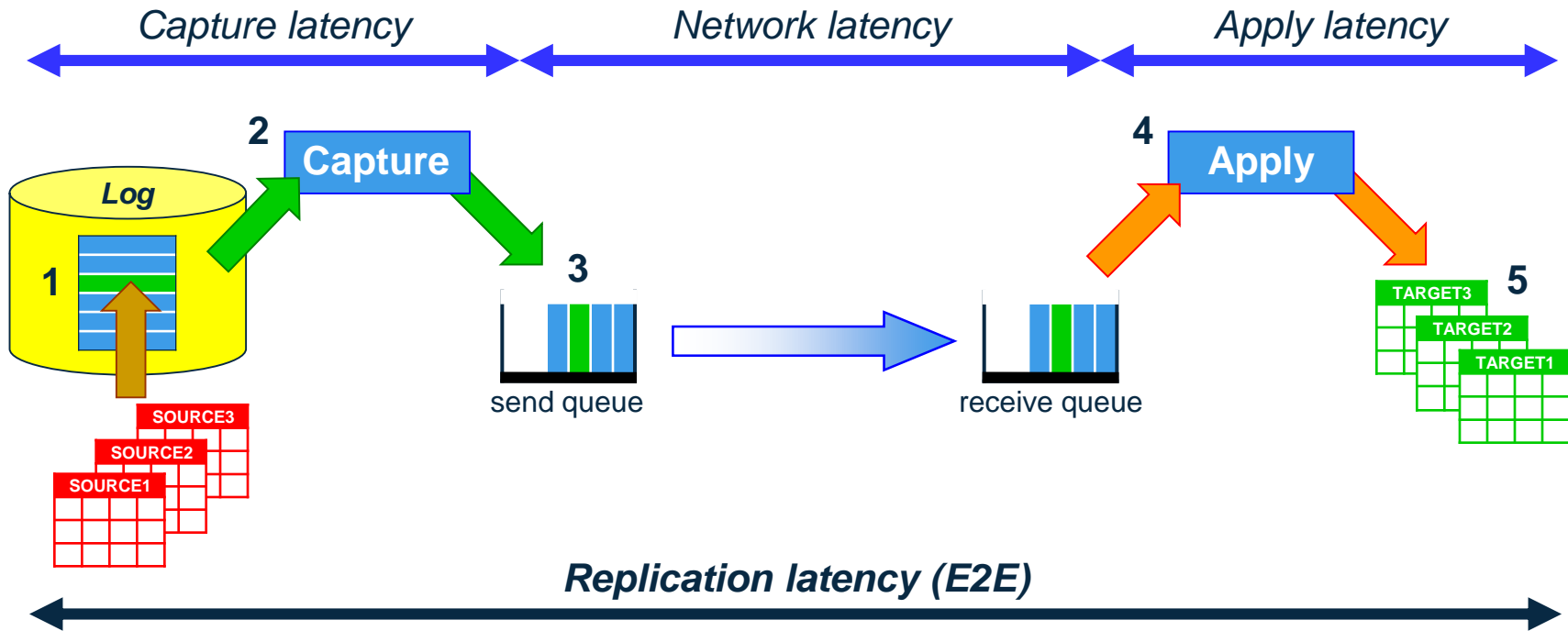

DDS=OnDemand | Asis
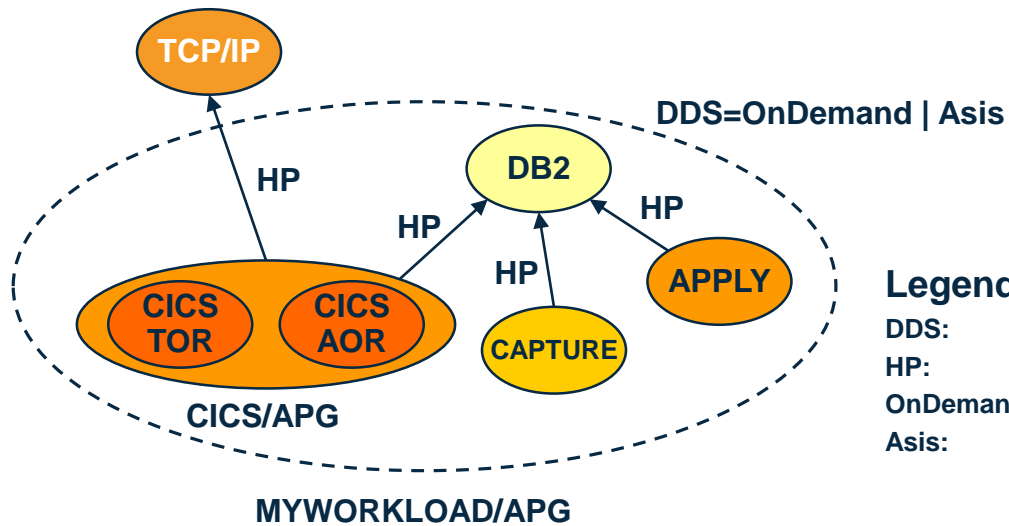
Legend

DDS:	Default Desired Status
HP:	HasParent
OnDemand:	Resource is UNAVAILABLE at IPL time
Asis:	Resource is kept in the state it is at IPL time

* Note that although SA is required on all systems, you can be using an alternative automation product to manage your workloads.
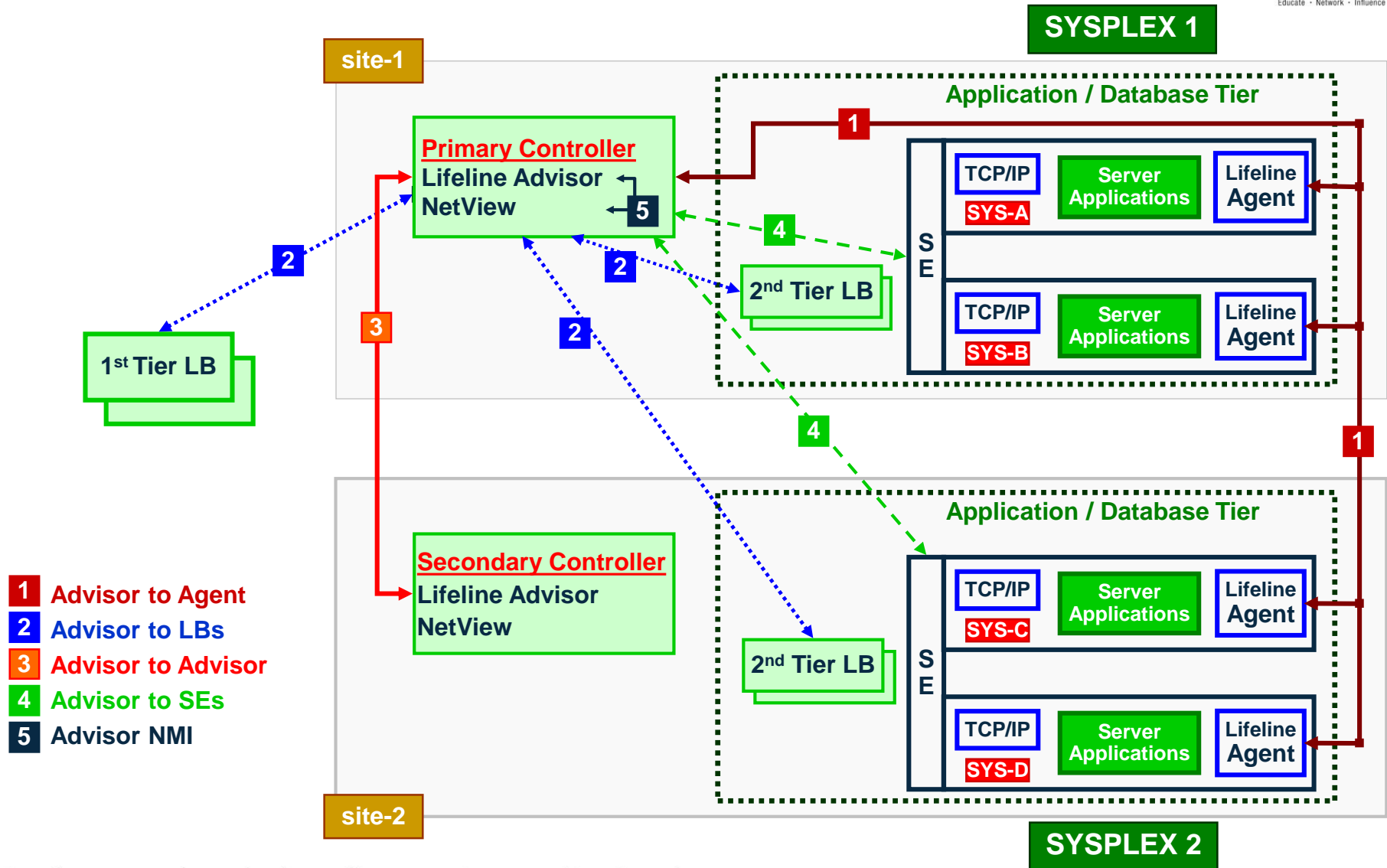
Data – deeper insight (S/W replication)

Capture latency   Network latency   Apply latency

Log

2 Capture

1

SOURCE3
SOURCE2
SOURCE1

3 send queue

receive queue

4 Apply

5 TARGET3
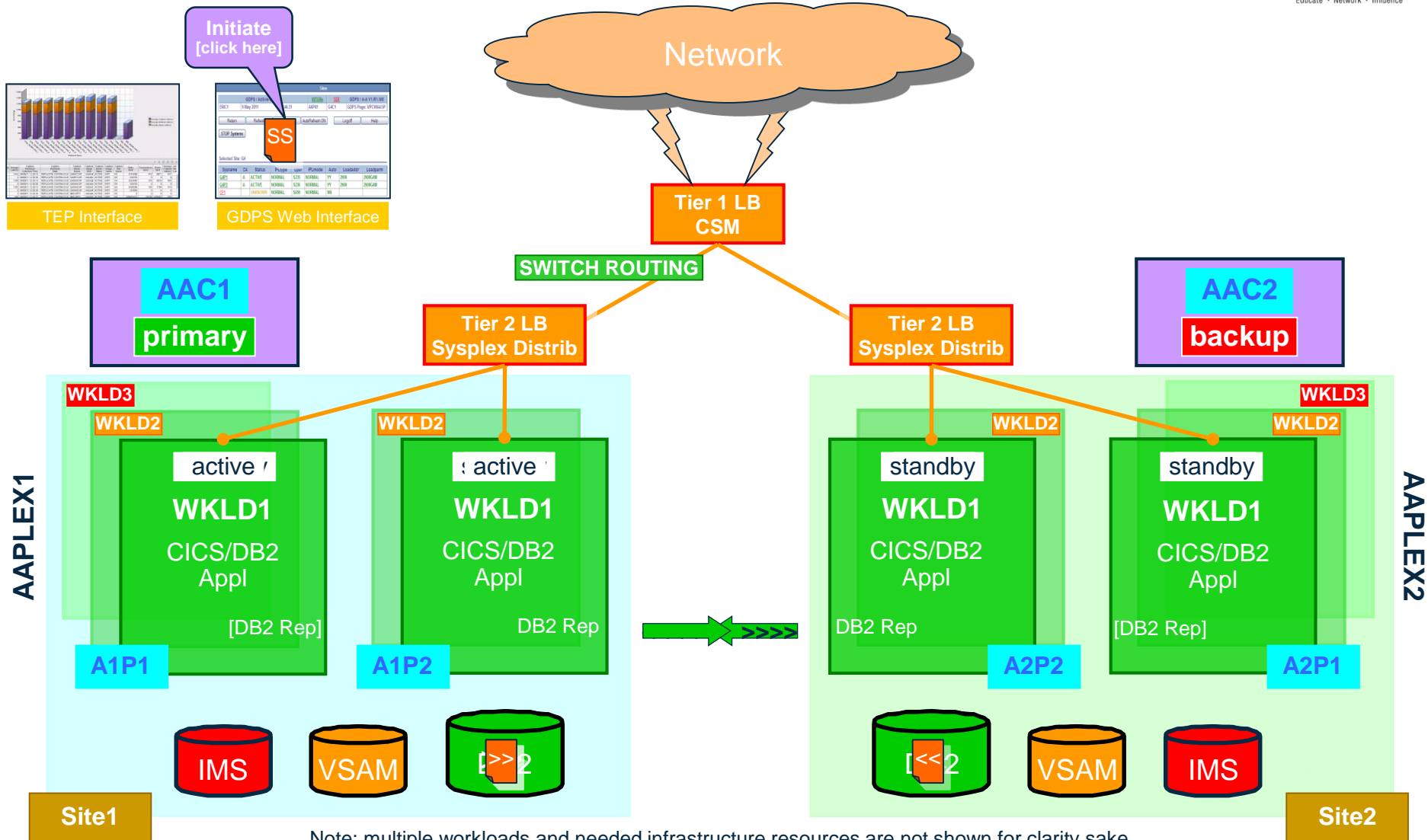TARGET2
TARGET1

Replication latency (E2E)

1. Transaction committed
2. Capture reads the DB updates from the log
3. Capture puts the updates on the send-queue
4. Apply receives the updates from the receive-queue
5. Apply copies the DB updates to the target databases

Connectivity – deeper insight

Legend:

1. Advisor to Agent
2. Advisor to LBs
3. Advisor to Advisor
4. Advisor to SEs
5. Advisor NMI

SYSPLEX 1

- site-1
- Primary Controller — Lifeline Advisor / NetView
- Application / Database Tier
  - TCP/IP SYS-A — Server Applications — Lifeline Agent
  - TCP/IP SYS-B — Server Applications — Lifeline Agent
- 2nd Tier LB
- 1st Tier LB

SYSPLEX 2

- site-2
- Secondary Controller — Lifeline Advisor / NetView
- Application / Database Tier
  - TCP/IP SYS-C — Server Applications — Lifeline Agent
  - TCP/IP SYS-D — Server Applications — Lifeline Agent
- 2nd Tier LB

Planned site switch

Planned site switch (continued)

> COMM = 'Switch all workloads to SITE2'
>
> ROUTING = 'SWITCH WORKLOAD=ALL SITE=AAPLEX1'

- **Switch routing** for all workloads active to Sysplex AAPLEX1 in Site1
  - quiesce batch, prevent new connections, quiesce OLTP and terminate persistent connections, allow replication to drain, and start routing to the newly active site
- **Note:** Replication is expected to be active in both directions at all times

> **The workloads are now processing transactions in Site2 for all workloads with replication from Site2 to Site1**

Unplanned site failure

Network
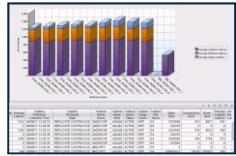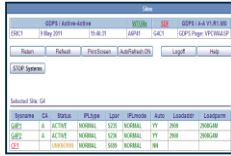
TEP Interface

GDPS Web Interface

LB 1° Tier
CSM

START ROUTING   STOP ROUTING

AAC1
primary

LB 2° Tier
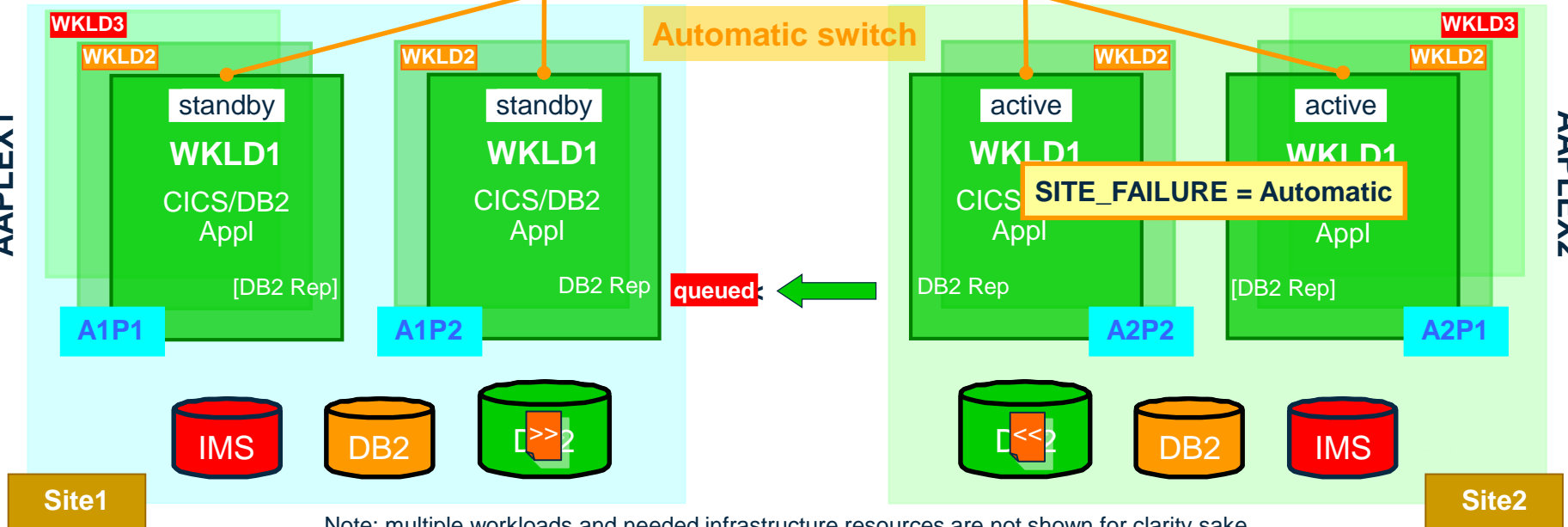Sysplex Distrib

LB 2° Tier
Sysplex Distrib

AAC2
backup

Automatic switch

AAPLEX1

WKLD3
WKLD2

standby
WKLD1
CICS/DB2
Appl

[DB2 Rep]

A1P1

WKLD2

standby
WKLD1
CICS/DB2
Appl

DB2 Rep

A1P2

queued

WKLD2

active
WKLD1
CICS
Appl

DB2 Rep

A2P2

SITE_FAILURE = Automatic

WKLD3
WKLD2

active
WKLD1
Appl

[DB2 Rep]

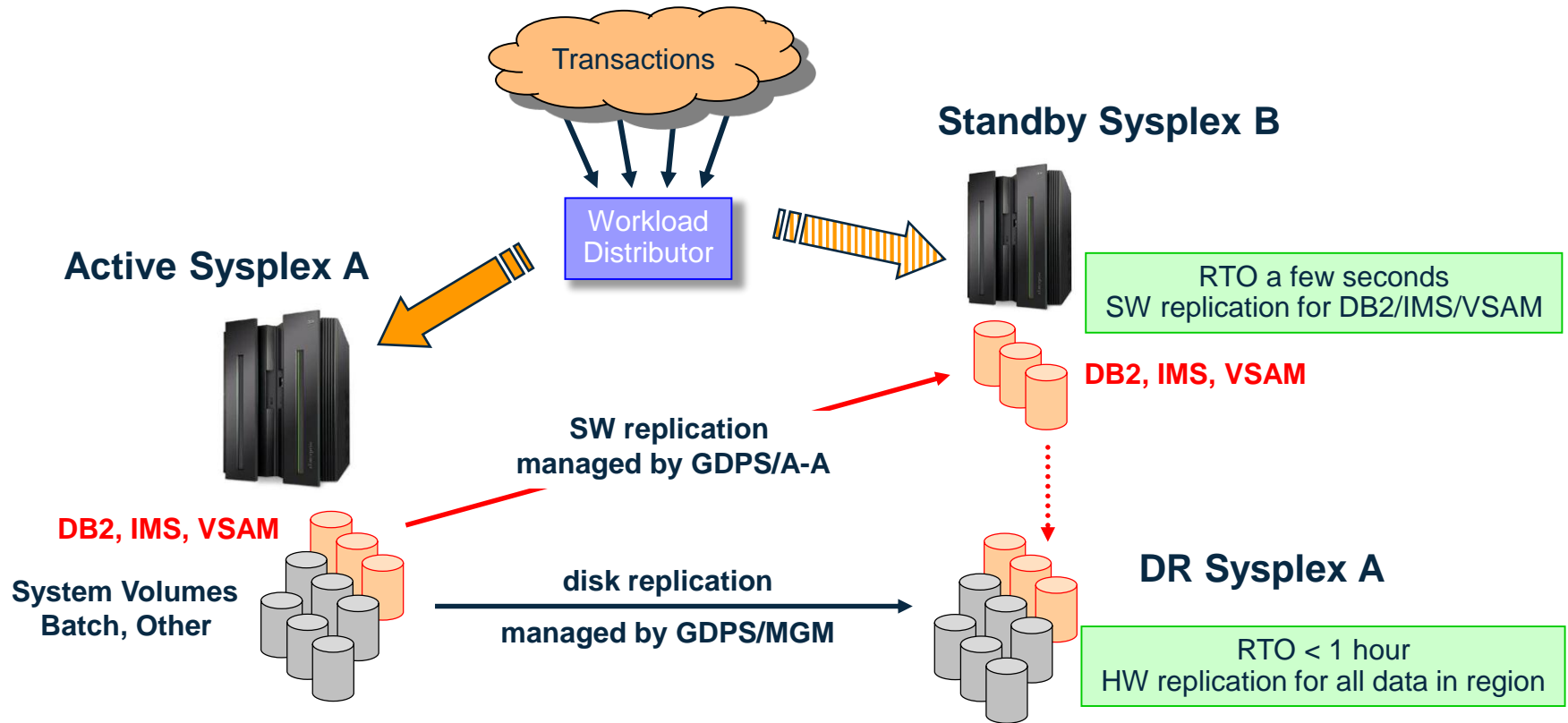A2P1

AAPLEX2

IMS   DB2   D>>2

D<<2   DB2   IMS

Site1

Site2

Note: multiple workloads and needed infrastructure resources are not shown for clarity sake

Complete your session evaluations online at www.SHARE.org/Seattle-Eval

03/05/2015   21

Disk replication integration

Transactions

Standby Sysplex B

Workload Distributor

Active Sysplex A

RTO a few seconds
SW replication for DB2/IMS/VSAM

DB2, IMS, VSAM

SW replication
managed by GDPS/A-A

DB2, IMS, VSAM

**System Volumes
Batch, Other**

disk replication

managed by GDPS/MGM

DR Sysplex A
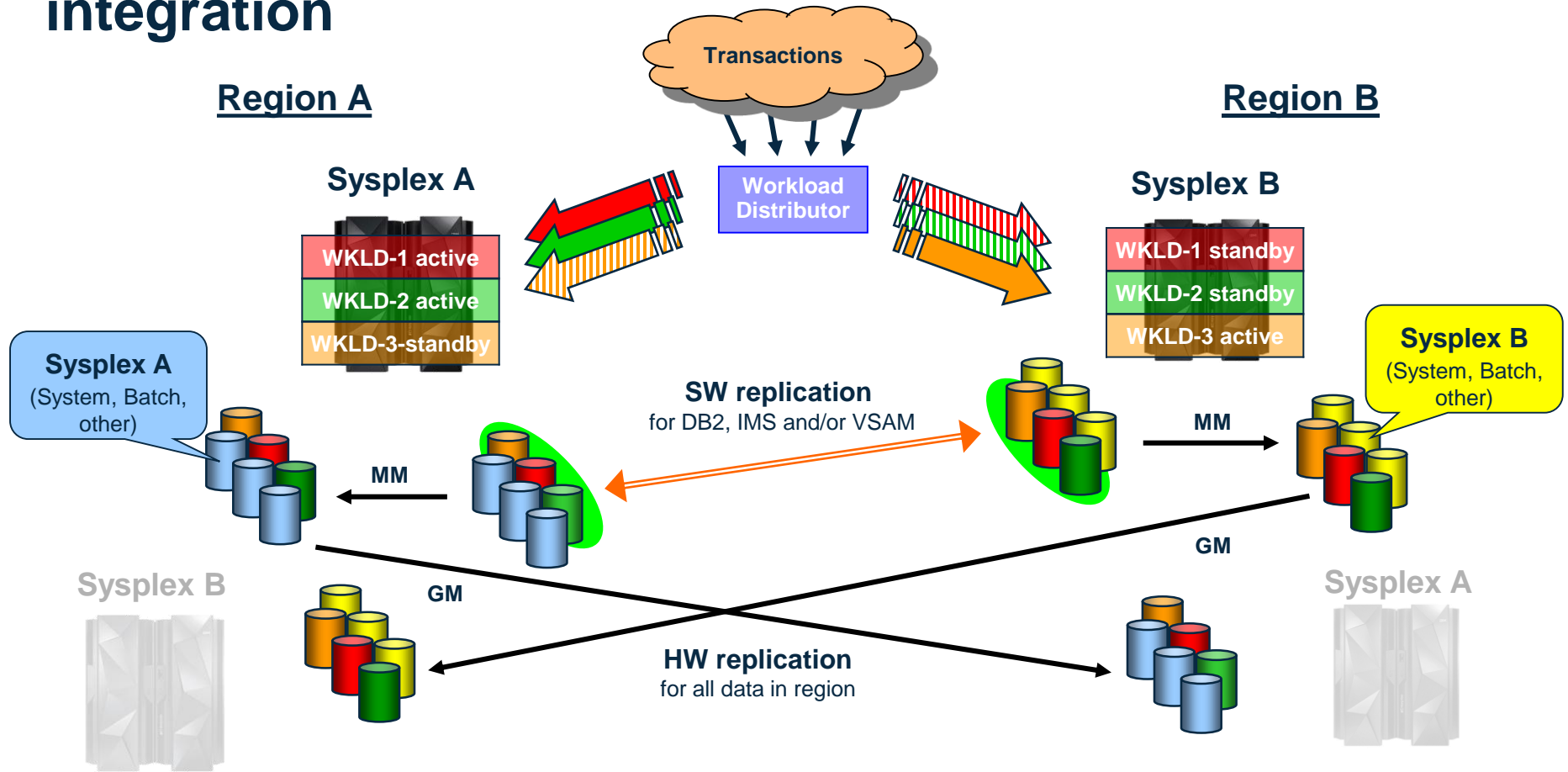
RTO < 1 hour
HW replication for all data in region

Two switch decisions for Sysplex A problems …

Workload Switch – switch to SW copy (B); once problem is fixed, simply restart SW replication

Site Switch – switch to SW copy (B) and restart DR Sysplex A from the disk copy
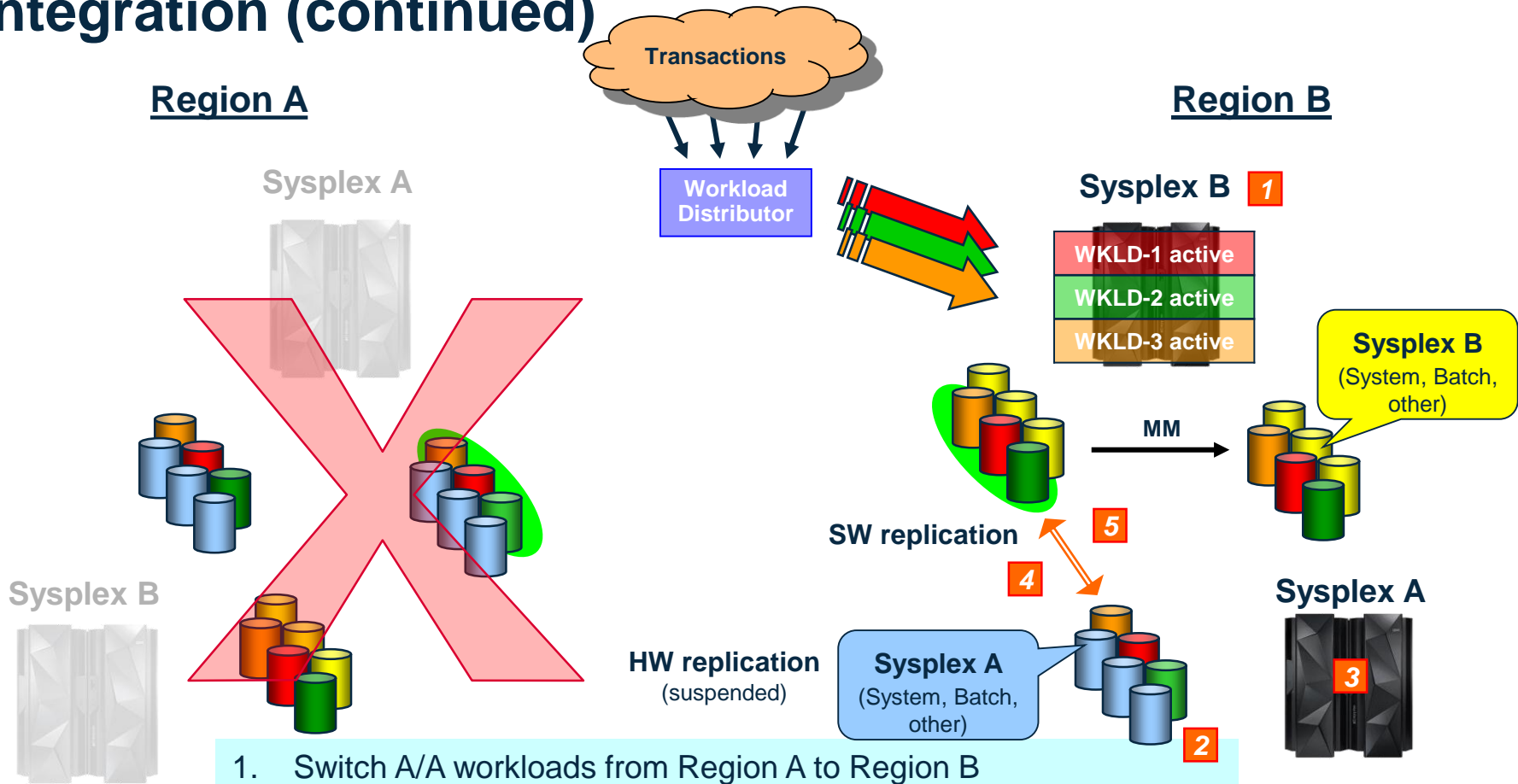
Disk replication integration (cont)

- Provide DR for whole production sysplex (AA workloads & non-A/A workloads)

- Restore A/A Sites capability for A/A Sites workloads after a planned or unplanned region switch

- Restart batch workloads after the prime site is restarted and re-synced

- The disk replication integration is optional

Unplanned region switch with disk replication integration

Region A

Region B

Sysplex A

Sysplex B

Transactions

Workload Distributor

WKLD-1 active
WKLD-2 active
WKLD-3-standby

WKLD-1 standby
WKLD-2 standby
WKLD-3 active

Sysplex A (System, Batch, other)

Sysplex B (System, Batch, other)

SW replication
for DB2, IMS and/or VSAM

MM

MM

Sysplex B

GM

GM

Sysplex A

HW replication
for all data in region

Unplanned region switch with disk replication integration (continued)

Region A

Sysplex A

Sysplex B

Transactions

Workload Distributor

Region B

Sysplex B 1

WKLD-1 active
WKLD-2 active
WKLD-3 active

MM

Sysplex B (System, Batch, other)

SW replication 5

4

HW replication (suspended)

Sysplex A (System, Batch, other)

2

Sysplex A

3

1. Switch A/A workloads from Region A to Region B
2. Recover Sysplex A secondary/tertiary disk
3. Restart Sysplex A in Region B
Potential manual tasks … (not automated by GDPS)
4. *Start software replication from B to A using adaptive (force) apply*
5. *Start software replication from A to B with default (ignore) apply*
6. *Manually reconcile exceptions from force (step 4)*

GDPS/A-A 1.4 new function summary

- **Active /Query configuration**
  - Fulfills SoD made when the Active/Standby configuration was announced
- **VSAM Replication support**
  - Adds to IMS and DB2 as the data types supported
  - Requires either CICS TS V5 for CICS/VSAM applications or CICS VR V5 for logging of non-CICS workloads
- **Support for IIDR for DB2 (Qrep) Multiple Consistency Groups**
  - Enables support for massive replication scalability
- **Workload switch automation**
  - Avoids manual checking for replication updates having drained as part of the switch process
- **GDPS/PPRC Co-operation support**
  - Enables GDPS/PPRC and GDPS/A-A to coexist without issues over who manages the systems
- **Disk replication integration**
  - Provides tight integration with GDPS/MGM for GDPS/A-A to be able to manage disaster recovery for the entire sysplex

Testing results*

Configuration:

- **9 * CICS-DB2** workloads + **1 * IMS** workload
- Distance between site 300 miles (≈500kms)

Test 1: Planned site switch

GDPS Active/Active	GDPS/XRC GDPS/GM
20 seconds	≈ 1-2 hour

Test 2: Unplanned site switch after a site failure (Automatic)

GDPS Active/Active	GDPS/XRC GDPS/GM
15 seconds	≈ 1 hour

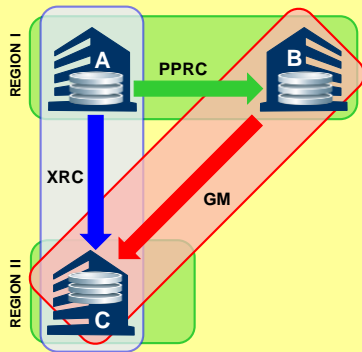*** IBM laboratory results; actual results may vary.**

Complete your session evaluations online at www.SHARE.org/Seattle-Eval

Multiple products meeting various availability objectives

GDPS/MGM[1] 3-site & GDPS/MzGM[2] 3-site

Near-continuous availability (CA) regionally and disaster recovery (DR) at extended distance

Three data centers

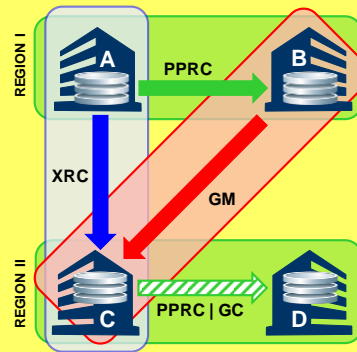Continuous availability (CA) for site disasters and Disaster recovery (DR) for regional disasters

Designed for near-continuous availability within the region and RPO seconds and RTO less then one hour for unplanned region fail-over

GDPS/MGM[1] 4-site & GDPS/MzGM[2] 4-site(SOD)

Near-continuous availability (CA) regionally and disaster recovery (DR) at extended distance from either region

Four data centers

Continuous availability (CA) and Disaster recovery (DR) when running from any site
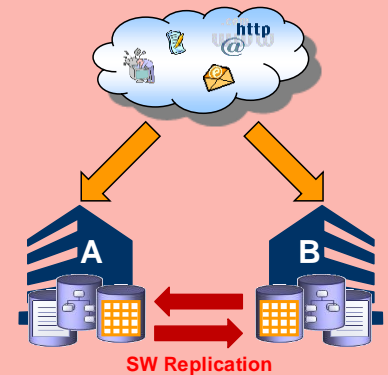

Designed for near-continuous availability within either region and RPO seconds and RTO less then one hour for unplanned region fail-over

GDPS/Active-Active

Near-continuous availability, disaster recovery, and cross-site workload balancing at extended distance

Two or more data centers

Disaster recovery for out-of-region interruptions

All sites active

RPO seconds and RTO seconds

[1]Metro Global Mirror (MGM)   [2]Metro z/OS Global Mirror (MzGM)

Backup Charts

#SHAREorg

Pre-requisite software matrix

Pre-requisite software [minimum version/release level]	GDPS Controller	A-A Systems	non A-A Systems
Operating Systems			
z/OS 1.13 or higher	YES	YES	YES
Application Middleware			
DB2 for z/OS V9 or higher	NO	YES wkld dependent	as required
IMS V12	NO	YES wkld dependent	as required
Websphere MQ V7.0.1	NO	MQ is only req'd for DB2 data replication	as required
CICS Transaction Server for z/OS V5.1	NO	YES [1]	as required
CICS VSAM Recovery for z/OS V5.1	NO	YES [1]	as required
[1] CICS TS and CICS VR are required when using VSAM replication for A-A workloads			
Replication			
InfoSphere Data Replication for DB2 for z/OS 10.2 and SPE	NO	YES wkld dependent	as required [2]
InfoSphere Data Replication for IMS for z/OS V11.1	NO	YES wkld dependent	as required [2]
InfoSphere Data Replication for VSAM for z/OS V11.1	NO	YES wkld dependent	as required [2]
[2] Non-Active/Active systems & their workloads can, if required, use Replication Server instances, but not the same instances as the A-A workloads			

Pre-requisite software matrix (continued)

Pre-requisite software [minimum version/release level]	GDPS Controller	A-A Systems	non A-A Systems
Management and Monitoring			
GDPS/A-A V1.4	YES	YES [3]	YES [3]
[3] GDPS/A-A requires the installation of the GDPS satellite code in production systems where A-A workloads			
IBM Tivoli NetView Monitoring for GDPS v6.2.1 [4]	YES	YES	YES [3]
[4] IBM Tivoli NetView Monitoring for GDPS v6.2.1 requires IBM Tivoli NetView for z/OS V6.2.1			
IBM Tivoli Management Services for z/OS V6.3 Fixpack 3 or later	YES [5]	YES [6]	YES [6]
[5] IBM Tivoli NetView Management Services for z/OS is required for the NetView for z/OS Enterprise Management Agent to monitor the A-A solution. [6] IBM Tivoli NetView Management Services for z/OS is optionally required to run where the NetView for z/OS Enterprise Management Agent runs to monitor NetView itself or where OMEGAMON XE products are deployed.			
IBM Tivoli Monitoring V6.3 Fix Pack 3 or later	NO	NO	NO
Tivoli System Automation for z/OS V3.5	YES	YES	YES
IBM Multi-site Workload Lifeline Version for z/OS 2.0	YES	YES	NO
Optional Monitoring Products			
Additional products such as Tivoli OMEGAMON XE on z/OS, Tivoli OMEGAMON XE for DB2, and Tivoli OMEGAMON XE for IMS may optionally be deployed to provide specific monitoring of products that are part of the Active/Active sites solution			

Note: Details of cross product dependencies are listed in the PSP information for GDPS/A-A which can be found by selecting the **Upgrade:GDPS** and **Subset:AAV1R4** at the following URL: http://www14.software.ibm.com/webapp/set2/psearch/search?domain=psp&new=y

Complete your session evaluations online at www.SHARE.org/Seattle-Eval

Additional Information

- **Web sites:**
GDPS	**www.ibm.com/systems/z/gdps**
Parallel Sysplex	**www.ibm.com/systems/z/pso**
Bus Resiliency z	**www.ibm.com/systems/z/resiliency**
Bus Resiliency	**www.ibm.com/systems/business_resiliency**

- System z   **www.ibm.com/systems/z/hardware**
  Storage   **www.ibm.com/systems/storage**
  Redbooks®GDPS Family: An Introduction to Concepts and Capabilities
              www.redbooks.ibm.com/abstracts/sg246374.html?Open

- **GDPS Web Site White Papers and Presentations**
  - GDPS: The Ultimate e-business Availability Solution
  - IBM Implementation Services for GDPS/Global Mirror
  - GDPS Business Continuity Solutions
  - Consistency Groups in a Nutshell
  - DS8000™ Data Replication
  - GDPS Solutions

- **e-mail:**   **gdps@us.ibm.com**