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

John Thompson
IBM

March 13th, 2014
Session 14615
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

• Requirements
• Concepts and Configurations
• Components
• Scenarios
• Recent Enhancements
• Summary
Suite of GDPS products to meet various availability and disaster recovery requirements

**Continuous Availability of Data within a Data Center**
- **GDPS/PPRC HM**
  - RPO=0
  - [RTO secs]
  - for disk only
- **Single Data Center**
  - Applications remain active
  - Continuous access to data in the event of a storage outage

**Disaster Recovery Extended Distance**
- GDPS/MGM & GDPS/MzGM
  - RPO=0
  - RTO<1h
  - RPO secs, RTO<1h
- **Three Data Centers**
  - High availability for site disasters
  - Disaster recovery for regional disasters

**Continuous Availability with DR within Metropolitan Region**
- GDPS/PPRC
  - RPO=0
  - RTO mins / RTO<1h
  - (<20km) / (>20km)
- **Two Data Centers**
  - Systems remain active
  - Multi-site workloads can withstand site and/or storage failures

**Disaster Recovery**
- GDPS/GM & GDPS/XRC
  - RPO secs, RTO<1h
  - RPO – recovery point objective
  - RTO – recovery time objective
- **Two Data Centers**
  - Rapid Systems D/R w/ “seconds” of data loss
  - Disaster Recovery for out of region interruptions

**Completion**

Complete your session evaluations online at www.SHARE.org/Anaheim-Eval
Evolving Customer Requirements

- Shift focus from failover model to *near-continuous availability* model (RTO near zero)
- Access data from *any site* (unlimited distance between sites)
- Multi-sysplex, multi-platform solution
  - “Recover *my business rather than my platform* technology”
- Ensure successful recovery via *automated processes* (similar to GDPS technology today)
  - Can be handled by less-skilled operators
- Provide *workload distribution between sites* (route around failed sites, dynamically select sites based on ability of site to handle additional workload)
- Provide *application level granularity*
  - Some workloads may require immediate access from every site, other workloads may only need to update other sites every 24 hours (less critical data)
  - Current solutions employ an all-or-nothing approach (complete disk mirroring, requiring extra network capacity)
From High Availability to Continuous Availability

<table>
<thead>
<tr>
<th>GDPS/PPRC</th>
<th>GDPS/XRC or GDPS/GM</th>
<th>GDPS/Active-Active</th>
</tr>
</thead>
<tbody>
<tr>
<td>Near Continuous Availability model</td>
<td>Failover model</td>
<td>Near Continuous Availability model</td>
</tr>
<tr>
<td>Recovery time = 2 minutes</td>
<td>Recovery time &lt; 1 hour</td>
<td>Recovery time &lt; 1 minute</td>
</tr>
<tr>
<td>Distance &lt; 20 KM</td>
<td>Unlimited distance</td>
<td>Unlimited distance</td>
</tr>
</tbody>
</table>

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
Terminology

• **Active/Active Sites**
  • This is the overall concept of the shift from a failover model to a continuous availability model.
  • Often used to describe the overall solution, rather than any specific product within the solution.

• **GDPS/Active-Active**
  • The name of the GDPS product which provides, along with the other products that make up the solution, the capabilities mentioned in this presentation such as workload, replication and routing management and so on.
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 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

This is a fundamental paradigm shift from a failover model to a continuous availability model.
Active/Standby Configuration – Both Sites Active for Individual Workloads

Static Routing
Automatic Failover

Transactions

Workload Distributor

Application A standby, Application B active

Application A active, Application B standby

Replication

site1

site2

DB2

IMS

Complete your session evaluations online at www.SHARE.org/Anaheim-Eval
Active/Query Configuration

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

Appl 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

Transactions

Read-only or query transactions to be routed to both sites, while update transactions are routed only to the active site

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

[A] latency>5; as “max latency” policy has been exceeded. route all queries to site2

[A] latency<3; as latency is less than “reset latency” policy, route more queries to site1

Replication

Workload Distributor

Complete your session evaluations online at www.SHARE.org/Anaheim-Eval
What Is 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)
Conceptual View

Transactions

Workload Routing to active sysplex

Workload Distribution

Active Update Workload & Active Query Workload

S/W Data Replication

Standby Update Workload & Active Query Workload

Controllers

Any load balancer or workload distributor that supports the Server Application State Protocol (SASP)

Control information passed between systems and workload distributor

Workload Lifeline, Tivoli NetView, System Automation, …
What S/W Makes Up a GDPS/Active-Active Environment?

• GDPS/Active-Active
• IBM Tivoli NetView Monitoring for GDPS
  • IBM Tivoli NetView for z/OS
  • IBM Tivoli NetView for z/OS Enterprise Management Agent (NetView agent)
• System Automation for z/OS
• IBM Multi-site Workload Lifeline for z/OS
• 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

Integration of a number of software products
GDPS/Active-Active (the product)

- **Automation code** is an extension on many of the techniques tried and tested in other GDPS products and with many client environments for management of their mainframe CA & DR requirements
- **Control code** runs on Controller systems and application systems
- **Workload management** - start/stop components of a workload in a given Sysplex
- **Replication management** - start/stop replication for a given workload between sites
- **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
- **Planned/Unplanned situation management and control** - planned or unplanned site or workload switches; automatic actions such as automatic workload switch (policy dependent)
- **Powerful scripting capability** for complex/compound scenario automation
What is 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 (eg: 10.10.10.1:80)
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

* 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)

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

Replication latency (E2E)
Connectivity – Deeper Insight

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

WAN & SASP-compliant Routers used for workload distribution

Active Production
- z/OS
- Lifeline Agent
- Workload
- IMS/DB2/VSAM
- Replication Capture
- TCPIP
- MQ
- NetView
- SA
- Other Automation Product

Primary Controller
- z/OS
- Lifeline Advisor
- NetView
- SA & BCPii
- GDPS/A-A
- Tivoli Monitoring

Backup Controller
- z/OS
- Lifeline Advisor
- NetView
- SA & BCPii
- GDPS/A-A
- Tivoli Monitoring

Standby Production
- z/OS
- Lifeline Agent
- Workload
- IMS/DB2/VSAM
- Replication Apply
- TCPIP
- MQ
- NetView
- SA
- Other Automation Product

SE/HMC LAN
Disk Replication Integration

Active Sysplex A

DB2, IMS, VSAM
System Volumes
Batch, Other

SW replication managed by GDPS/A-A

Transactions

Workload Distributor

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

Standby Sysplex B

DR Sysplex A

RTO < 1 hour
HW replication for all data in region

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

DB2, IMS, VSAM

Two switch decisions for Sysplex A problems …
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
Planned site switch

Note: multiple workloads and needed infrastructure resources are not shown for clarity sake.
Planned site switch (cont)

• **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

Note: multiple workloads and needed infrastructure resources are not shown for clarity sake.
Unplanned Region Switch with Disk Replication Integration

Region A
- Sysplex A
  - WKLD-1 active
  - WKLD-2 active
  - WKLD-3 standby
  - (System, Batch, other)
  - SW replication for DB2, IMS and/or VSAM

Region B
- Sysplex B
  - WKLD-1 standby
  - WKLD-2 standby
  - WKLD-3 active
  - (System, Batch, other)

Transactions

High Availability in Region & DR Protection in other Region

Complete your session evaluations online at www.SHARE.org/Anaheim-Eval
Unplanned Region Switch with Disk Replication Integration (cont)

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
Summary

- Manages availability at a workload level
- Provides a central point of monitoring & control
- Manages replication between sites
- Provides the ability to perform a controlled workload site switch
- Provides near-continuous data and systems availability and helps simplify disaster recovery with an automated, customized solution
- Reduces recovery time and recovery point objectives – measured in seconds
- Facilitates regulatory compliance management with a more effective business continuity plan
- Simplifies system resource management

GDPS/Active-Active is the next generation of GDPS
Testing results*

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

**Test1:**
Planned site switch

<table>
<thead>
<tr>
<th>GDPS Active/Active</th>
<th>GDPS/XRC GDPS/GM</th>
</tr>
</thead>
<tbody>
<tr>
<td>20 seconds</td>
<td>≈ 1-2 hour</td>
</tr>
</tbody>
</table>

**Test2:**
Unplanned site switch
After a site failure (Automatic)

<table>
<thead>
<tr>
<th>GDPS Active/Active</th>
<th>GDPS/XRC GDPS/GM</th>
</tr>
</thead>
<tbody>
<tr>
<td>15 seconds</td>
<td>≈ 1 hour</td>
</tr>
</tbody>
</table>

* IBM laboratory results; actual results may vary.
Suite of GDPS products to meet various availability and disaster recovery requirements

<table>
<thead>
<tr>
<th>Continuous Availability of Data within a Data Center</th>
<th>Continuous Availability with DR within Metropolitan Region</th>
<th>Disaster Recovery Extended Distance</th>
<th>CA Regionally and Disaster Recovery Extended Distance</th>
<th>CA, DR, &amp; Cross-site Workload Balancing Extended Distance</th>
</tr>
</thead>
<tbody>
<tr>
<td>GDPS/PPRC HM RPO=0 [RTO secs] for disk only</td>
<td>GDPS/PPRC RPO=0 RTO mins / RTO&lt;1h (&lt;20km)</td>
<td>GDPS/GM &amp; GDPS/XRC RPO secs, RTO&lt;1h</td>
<td>GDPS/MGM &amp; GDPS/MzGM RPO=0, RTO mins/&lt;1h &amp; RPO secs, RTO&lt;1h</td>
<td>GDPS/Active-Active RPO secs, RTO secs</td>
</tr>
</tbody>
</table>

**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 D/R w/ “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 or more Active Data Centers**
- Automatic workload switch in seconds; seconds of data loss

RPO – recovery point objective  RTO – recovery time objective
QR Code for Evaluations
Backup Charts
Pre-requisite products

- **IBM Multi-site Workload Lifeline v2.0**
  - Advisor – runs on the Controllers & provides information to the external load balancers on where to send transactions and information to GDPS on the health of the environment
    - There is one primary and one secondary advisor
  - Agent – runs on all production images with active/active workloads defined and provide information to the Lifeline Advisor on the health of that system

- **IBM Tivoli NetView Monitoring for GDPS v6.2**
  - Runs on all systems and provides automation and monitoring functions. This new product requires IBM Tivoli NetView for z/OS. The NetView Enterprise Master runs on the Primary Controller

- **IBM Tivoli Monitoring v6.3 FP1**
  - Can run on zLinux, or distributed servers – provides monitoring infrastructure and portal plus alerting/situation management via Tivoli Enterprise Portal, Tivoli Enterprise Portal Server and Tivoli Enterprise Monitoring Server
Pre-requisite products (continued)

• IBM InfoSphere Data Replication for DB2 for z/OS v10.2
  • Runs on production images where required to capture (active) and apply (standby) data updates for DB2 data. Relies on MQ as the data transport mechanism (QREP)

• IBM InfoSphere Data Replicator for IMS for z/OS v11.1
  • Runs on production images where required to capture (active) and apply (standby) data updates for IMS data. Relies on TCPIP as the data transport mechanism

• IBM Infosphere Data Replicator for VSAM for z/OS v11.1
  • Runs on production images where required to capture (active) and apply (standby) data updates for VSAM data. Relies on TCP/IP as data transport mechanism. Requires CICS TS or CICS VR

• System Automation for z/OS v3.3 or higher
  • Runs on all images. Provides a number of critical functions:
    • BCPii
    • Remote communications capability to enable GDPS to manage sysplexes from outside the sysplex
    • System Automation infrastructure for workload and server management

• Optionally the OMEGAMON suite of monitoring tools to provide additional insight
# Pre-requisite software matrix

<table>
<thead>
<tr>
<th>Pre-requisite software [version/release level]</th>
<th>GDPS Controller</th>
<th>A-A Systems</th>
<th>non A-A Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Operating Systems</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>z/OS 1.13 or higher</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td><strong>Application Middleware</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DB2 for z/OS V9 or higher</td>
<td>NO</td>
<td>YES wkld dependent</td>
<td>as required</td>
</tr>
<tr>
<td>IMS V11</td>
<td>NO</td>
<td>YES wkld dependent</td>
<td>as required</td>
</tr>
<tr>
<td>Websphere MQ V7.0.1</td>
<td>NO</td>
<td>MQ is only req’d for DB2 data replication</td>
<td>as required</td>
</tr>
<tr>
<td>CICS Transaction Server for z/OS V5.1</td>
<td>NO</td>
<td>YES ¹)</td>
<td>as required</td>
</tr>
<tr>
<td>CICS VSAM Recovery for z/OS V5.1</td>
<td>NO</td>
<td>YES ¹)</td>
<td>as required</td>
</tr>
</tbody>
</table>

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)

<table>
<thead>
<tr>
<th>Pre-requisite software [version/release level]</th>
<th>GDPS Controller</th>
<th>A-A Systems</th>
<th>non A-A Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Management and Monitoring</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GDPS/A-A V1.4</td>
<td>YES</td>
<td>YES 4)</td>
<td>YES 4)</td>
</tr>
<tr>
<td>IBM Tivoli NetView Monitoring for GDPS V6.2</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Tivoli System Automation for z/OS V3.3 + SPE APARs</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>IBM Multi-site Workload Lifeline Version for z/OS 2.0</td>
<td>YES</td>
<td>YES</td>
<td>NO</td>
</tr>
<tr>
<td>IBM Tivoli Monitoring V6.3 Fix Pack 1</td>
<td>YES 6)</td>
<td>YES 7)</td>
<td>NO</td>
</tr>
<tr>
<td>IBM Tivoli Management Services for z/OS V6.3</td>
<td>YES 8)</td>
<td>NO</td>
<td>NO</td>
</tr>
</tbody>
</table>

4) GDPS/A-A requires the installation of the GDPS satellite code in production systems where A-A workloads run.

5) IBM Tivoli NetView Monitoring for GDPS V6.2 requires IBM Tivoli NetView for z/OS V6.2.

6) Tivoli Enterprise Monitoring Server can optionally run on zLinux or on distributed server accessible by the NetView agent; in addition, the Tivoli Enterprise Portal Server runs on a distributed platform.

7) Optional IBM Tivoli Monitoring agents might be required on production systems for purposes other than GDPS/A-A.

8) IBM Tivoli Management Services for z/OS is available separately, or is shipped with the OMEGAMON products.

**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