Hitachi Virtual Storage Platform G1000 with Global-active Device for Oracle Database with the Oracle Real Application Clusters Option

Tech Note

By Kishore Daggubati and Gustavo Castets

February 2, 2015
Feedback

Hitachi Data Systems welcomes your feedback. Please share your thoughts by sending an email message to SolutionLab@hds.com. To assist the routing of this message, use the paper number in the subject and the title of this white paper in the text.
# Table of Contents

Solution Overview.................................................................3

Key Solution Components....................................................5

Solution Design.................................................................8
  Storage Configuration..........................................................8
  Database Layout..............................................................12
  Server Configuration..........................................................15
  SAN Configuration............................................................16

Engineering Validation.......................................................19
  Test Methodology.............................................................19
  Results Summary.............................................................24

Conclusions.............................................................................28

Acknowledgments.....................................................................29
Hitachi Virtual Storage Platform G1000 with Global-active Device for Oracle Database with the Oracle Real Application Clusters Option

Tech Note

This tech note shows how the Hitachi Storage Virtualization Operating System (SVOS) global-active device for Hitachi Virtual Storage Platform G1000 (VSP G1000) provides high availability and business continuity for an Oracle infrastructure by creating active/active data centers in geographically separate locations. The environment uses IBM® Power® 770 and the Oracle Database with Real Application Clusters (RAC) option. Use this resource to design an infrastructure for your requirements and budget.

This validated solution integrates servers, storage systems, network, and storage software. The solution provides reliability, high availability, and scalability in a cost effective manner while processing small-scale to large-scale OLTP workloads. The dedicated server runs Oracle Database 11g R2. The operating system is IBM AIX® 6.1.0.0 with a technology level of 6100-07 and a service pack of 6100-07-03-1207.

You benefit from using this tech note if you are one of the following:

- Database administrator
- Storage administrator
- IT professional with the responsibility of planning and deploying an Oracle Database 11g R2 solution

If necessary, use this tech note to design a larger infrastructure that meets your needs by adding the following:

- Additional IBM Power7 systems
- Additional disks and cache on the VSP G1000 system
To use this tech note, you need familiarity with the following:

- IBM Power 770
- Hitachi Virtual Storage Platform G1000
- ANUE H Series
- Oracle Database 11g Release 2
- Hitachi Dynamic Link Manager

**Note** — Testing of this configuration was in a lab environment. Many things affect production environments beyond prediction or duplication in a lab environment. Follow the recommended practice of conducting proof-of-concept testing for acceptable results in a non-production, isolated test environment that otherwise matches your production environment before your production implementation of this solution. This solution uses ANUE for network simulation. Production environments may not require this component.
Solution Overview

This solution implements Oracle Database with the RAC option using Hitachi Virtual Storage Platform G1000 with Hitachi Storage Virtualization Operating System global-active device. It includes high availability, disaster recovery, and scalability benefits for typical OLTP workloads. The architecture uses a synchronous replication model for zero data loss between two VSP G1000 storage systems that are separated by up to 100 KM for load balancing and active/active availability. Tailor your implementation of this solution to meet your specific needs.

This solution includes Oracle clients that access the database at site 1 (local) and site 2 (remote) simultaneously, in order to store data in and access data from a VSP G1000 storage system.

If there is a storage system failure at the local site, these Oracle clients can access the Hitachi Virtual Storage Platform G1000 storage system at the remote site without any data loss or impact to the Oracle RAC configuration.
Figure 1 shows the infrastructure for this solution.

Note — This solution uses Hitachi Unified Storage VM for quorum device at site 3. You may choose to use any other supported storage system.
Key Solution Components

Table 1 lists the key hardware components and Table 2 lists the software components used in this solution.

Table 1. Key Hardware Components

<table>
<thead>
<tr>
<th>Component</th>
<th>Description</th>
<th>Version</th>
<th>Quantity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Server</td>
<td>IBM Power 770 with two LPARs, each configured as follows: 8 PowerPC_POWER7 processors, 3.1 GHz, 22.25 GB RAM, 2 x 8 Gb/sec Fibre Channel ports</td>
<td>IBM, AM710_114</td>
<td>2</td>
</tr>
</tbody>
</table>
Table 1. Key Hardware Components (Continued)

<table>
<thead>
<tr>
<th>Component</th>
<th>Description</th>
<th>Version</th>
<th>Quantity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Storage System</td>
<td>Hitachi Virtual Storage Platform G1000, configured as follows:</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>▪ 264 × 300 GB SAS 15K RPM hard disk drives</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>▪ 11 trays with 24 disks in each tray</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>▪ About 77 TB raw capacity</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>▪ 2 main storage blades, each with:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>▪ 2 front-end connectivity modules</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>▪ 4 × 8 Gb/sec Fibre Channel ports, 8 ports total</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>▪ 2 back-end connectivity modules</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>▪ 8 × 6 Gb/sec SAS links each, 16 links total</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>▪ 2 controllers, each with 8 MP blades</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>▪ 974 GB cache</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hitachi Unified Storage VM, configured as follows:</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>▪ 24 × 600 GB SAS 10K RPM drives in a single tray 14 TB raw capacity</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>▪ 2 main storage blades, each with:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>▪ 2 front-end connectivity modules</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>▪ 1 × 8 Gb/sec Fibre Channel ports, 2 ports total</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>▪ 2 back-end connectivity modules</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>▪ 8 × 6 Gb/sec SAS links each, 16 links total</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>▪ 2 MP blades, each with two 8-core Intel XEON processors, 2.33 GHz</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>▪ 64 GB cache</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Network Emulator</td>
<td></td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>ANUE H Series FC-8X each with:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>▪ 2 Blades of HS-XGEMOPC1-FC8-B</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SAN Connectivity</td>
<td></td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Brocade 6510</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 2. Key Software Components

<table>
<thead>
<tr>
<th>Component</th>
<th>Description</th>
<th>Version</th>
</tr>
</thead>
<tbody>
<tr>
<td>Operating System</td>
<td>IBM AIX</td>
<td>6.1.0.0</td>
</tr>
<tr>
<td>Database Software</td>
<td>Oracle</td>
<td>11g R2, 11.2.0.3.0</td>
</tr>
<tr>
<td>Cluster Software</td>
<td>Oracle Real Application Cluster</td>
<td>11g R2, 11.2.0.3.0</td>
</tr>
<tr>
<td>Volume Manager and File System software</td>
<td>Oracle Automatic Storage Management</td>
<td>11g R2, 11.2.0.3.0</td>
</tr>
<tr>
<td>Multipath Software</td>
<td>Hitachi Dynamic Link Manager</td>
<td>8.0.1.0</td>
</tr>
<tr>
<td>Storage Management Software</td>
<td>Hitachi Storage Navigator</td>
<td>Microcode dependent</td>
</tr>
<tr>
<td>Network Simulator Software</td>
<td>ANUE</td>
<td>V7.0.2</td>
</tr>
<tr>
<td>Replication Software</td>
<td>Hitachi Storage Virtualization</td>
<td>Microcode dependent</td>
</tr>
<tr>
<td>Database Client Communication Software</td>
<td>Oracle Net Services</td>
<td>11g R2, 11.2.0.3.0</td>
</tr>
<tr>
<td>Load Generator Software</td>
<td>Swingbench</td>
<td>2.5</td>
</tr>
</tbody>
</table>
Solution Design

This describes the Hitachi System for Oracle Database environment using Hitachi Virtual Storage Platform G1000 with Hitachi Storage Virtualization Operating System global-active device.

The hardware components used in this environment are listed in Table 1, “Key Hardware Components,” on page 5. The specific infrastructure configuration includes the following:

- **Server** — This is a single server node consisting of one LPAR in an IBM Power 770 system for one logical Oracle Database server.

- **Storage System** — Command devices and global-active device volume (primary volume P-VOLs and secondary volume S-VOLs) mapped to each port that are presented to the server as LUNs. In this tech note, the storage at the local site is referred as a primary storage system and the storage at the remote site is referred as a secondary storage system.

- **SAN Fabric** — There are four zones created on each Fibre Channel switch module that connect the server and the storage system:
  - One port of the dual-port Fibre Channel cards, from each of the servers
  - One port on each of the storage systems

The other port of the dual-port Fibre Channel cards, from each of the servers, connects directly to the port on the storage system.

Connect three inter-switch links from each of the Fibre Channel switch modules between the following:

- Server and the Storage system
- Storage systems

Storage Configuration

This describes the storage architecture of this environment. The design takes into consideration Hitachi Data Systems and Oracle recommended practices for the deployment of the database storage design.

The environment uses VSP G1000 as the storage subsystem. For the quorum disk, the environment uses Hitachi Unified Storage VM. Only a single quorum disk is required between a primary and secondary storage system. See Figure 1 on page 4 for storage connectivity.

**Note** - Following recommended practice, connect Hitachi Unified Storage VM to each VSP G1000 from a second remote site within metro distance from each site to ensure the quorum disk is accessible from at least one site during a site-wide failure.
Global-active device introduces the concept of a virtual storage machine that is created in the secondary VSP G1000 storage system. A virtual storage machine is created in the secondary array that mimics the VSP G1000 primary array serial number. Volumes (S-VOLs) are created in the virtual storage machine on the secondary array to match those defined in the primary array (P-VOLs). The attached servers see the combination of the physical VSP G1000 at the local site and virtual storage machine in the VSP G1000 at the remote site as a single physical system accessible over multiple Fibre Channel paths.

Figure 2 shows the storage provisioning on each of the two VSP G1000 storage systems for the environment.
Table 3 lists the storage configuration details for the VSP G1000 at the local site.

**Table 3. Storage Configuration of the VSP G1000 at the Local Site**

<table>
<thead>
<tr>
<th>Storage System</th>
<th>Hitachi Virtual Storage Platform G1000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main Microcode Level</td>
<td>80-01-41-00/00</td>
</tr>
<tr>
<td>Cache Memory</td>
<td>974 GB</td>
</tr>
<tr>
<td>Number of Ports Used</td>
<td>9</td>
</tr>
<tr>
<td>Number of SAS 10k RPM Drives</td>
<td>256</td>
</tr>
<tr>
<td>SAS Drive Capacity</td>
<td>300 GB</td>
</tr>
<tr>
<td>RAID Group Type</td>
<td>RAID-6 (6D+2P)</td>
</tr>
<tr>
<td>RAID Groups</td>
<td>1-1 - 1-24; 3-1 - 3-8</td>
</tr>
<tr>
<td>Number of Parity Groups</td>
<td>32</td>
</tr>
<tr>
<td>Size of each Parity Group</td>
<td>1800 GB</td>
</tr>
<tr>
<td>Number of LDEVs in each Parity Group</td>
<td>1</td>
</tr>
<tr>
<td>Size of LDEV in each Parity Group</td>
<td>1800 GB</td>
</tr>
<tr>
<td>Number of LDEVs</td>
<td>32</td>
</tr>
<tr>
<td>LDEVs</td>
<td>00:00:00 - 00:00:17</td>
</tr>
<tr>
<td></td>
<td>00:00:20 - 00:00:27</td>
</tr>
</tbody>
</table>

Table 4 has the storage configuration details for the VSP G1000 at the remote site.

**Table 4. Storage Configuration of the VSP G1000 at the Remote Site**

<table>
<thead>
<tr>
<th>Storage System</th>
<th>Hitachi Virtual Storage Platform G1000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main Microcode Level</td>
<td>80-01-41-00/00</td>
</tr>
<tr>
<td>Cache Memory</td>
<td>974 GB</td>
</tr>
<tr>
<td>Number of Ports Used</td>
<td>9</td>
</tr>
<tr>
<td>Number of SAS 10k RPM Drives</td>
<td>256</td>
</tr>
<tr>
<td>SAS Drive Capacity</td>
<td>300 GB</td>
</tr>
<tr>
<td>RAID Group Type</td>
<td>RAID-6 (6D+2P)</td>
</tr>
<tr>
<td>RAID Groups</td>
<td>1-1 - 1-24; 3-1 - 3-8</td>
</tr>
<tr>
<td>Number of Parity Groups</td>
<td>32</td>
</tr>
<tr>
<td>Size of each Parity Group</td>
<td>1800 GB</td>
</tr>
<tr>
<td>Number of LDEVs in each Parity Group</td>
<td>1</td>
</tr>
<tr>
<td>Size of LDEV in each Parity Group</td>
<td>1800 GB</td>
</tr>
<tr>
<td>Number of LDEVs</td>
<td>32</td>
</tr>
</tbody>
</table>
Table 5 has the details for the virtual volumes created from the dynamic provisioning pool of the VSP G1000 at the local site.

<table>
<thead>
<tr>
<th>Dynamic Provisioning Pool ID</th>
<th>ora_pool_01</th>
</tr>
</thead>
<tbody>
<tr>
<td>VVOLs</td>
<td>00:01:00 - 00:01:1F</td>
</tr>
<tr>
<td>Number of VVOLs</td>
<td>32</td>
</tr>
<tr>
<td>VVOL Size (GB)</td>
<td>24</td>
</tr>
<tr>
<td>Purpose</td>
<td></td>
</tr>
<tr>
<td>Oracle System</td>
<td></td>
</tr>
<tr>
<td>Sysaux</td>
<td></td>
</tr>
<tr>
<td>Undo</td>
<td></td>
</tr>
<tr>
<td>Temp</td>
<td></td>
</tr>
<tr>
<td>Redo Logs</td>
<td></td>
</tr>
<tr>
<td>Archive Logs</td>
<td></td>
</tr>
<tr>
<td>Oracle Cluster Registry</td>
<td></td>
</tr>
<tr>
<td>Voting Disk</td>
<td></td>
</tr>
<tr>
<td>OLTP Application Tablespaces</td>
<td></td>
</tr>
<tr>
<td>Storage Ports</td>
<td>1A, 1J, 2C, 2L, 5D, 6D, 7D, 8D</td>
</tr>
</tbody>
</table>

Table 6 has the details for the virtual volumes created from the dynamic provisioning pool of the VSP G1000 at the remote site.

<table>
<thead>
<tr>
<th>Dynamic Provisioning Pool ID</th>
<th>ora_pool_01</th>
</tr>
</thead>
<tbody>
<tr>
<td>VVOLs</td>
<td>00:01:20 - 00:01:3F</td>
</tr>
<tr>
<td>Number of VVOLs</td>
<td>32</td>
</tr>
</tbody>
</table>
Note — 5D, 6D, 7D and 8D storage ports on each of the VSP G1000 storage system are used for the SVOS global-active device configuration. You may choose to use any other storage ports.

VSP G1000 uses the externalized storage as a quorum disk to ensure that data is consistent between a global-active device pair.

Database Layout

The database layout design uses recommended practices from Hitachi Data Systems for VSP G1000 for small random I/O traffic, such as the ones in OLTP transactions. The layout also takes into account Oracle best practices when using Hitachi storage.
Base the storage design for database layout needs on the requirements of a specific application implementation. The design can vary greatly from one implementation to another. The components in this solution set have the flexibility for use in various deployment scenarios to provide the right balance between performance and ease of management for a given scenario.

- **Data and Indexes Tablespace** — Assign one dynamic provisioning pool for application data and indexes. The allocated capacity is 56 TB. The small file table space consists of several 30 GB data files. Set the tablespace to a small initial size with auto extend enabled to maximize storage utilization from thin provisioning.

- **Temp Tablespace** — Place the TEMP tablespace in this configuration in the Data ASM diskgroup. Quite a number of small file tempfiles are created within one single small TEMP tablespace. Limit the size of each small file tempfile to 30 GB.

- **UNDO Tablespace** — Place an UNDO tablespace in this configuration in the Data ASM diskgroup. Assign one UNDO tablespace for each database server. Quite a number of small file undo datafiles are created within each small UNDO tablespace. Limit the size of each small undo datafile to 30 GB.

- **Online Redo Logs** — Assign two ASM diskgroups for the database instance.

- **Oracle Cluster Registry and Voting Disk** — Place each of the files for this configuration in a separate ASM diskgroup.

- **Size Settings** — Set the database block size to 8 KB. Set the ASM allocation unit to 4 MB.

- **Memory Settings** — Set the Oracle database memory settings as listed below:
  - SGA_TARGET = 8.6 GB.
  - PGA_AGGREGATE_TARGET = 2.875 GB
  - DB_CACHE_SIZE = 6.25 GB
  - LOG_BUFFER = 274877907

The combination of SVOS global-active device with Hitachi Dynamic Link Manager (HDLM) enables the server to see the P-VOLs and S-VOLs as a single LUN.
Table 7 lists the disk mappings from the PVOLs, SVOLs, and LUNS to the ASM disk groups.

<table>
<thead>
<tr>
<th>P-VOLs</th>
<th>S-VOLs</th>
<th>LUN</th>
<th>ASK Disk</th>
<th>ASMDG</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>N/A</td>
<td>N/A</td>
<td>0</td>
<td>N/A</td>
<td>N/A</td>
<td>Operating system and Oracle software binaries for the first Oracle database server in a four node Oracle RAC Database</td>
</tr>
<tr>
<td>N/A</td>
<td>N/A</td>
<td></td>
<td>N/A</td>
<td>N/A</td>
<td>Operating system and Oracle software binaries for the second Oracle database server in a four node Oracle RAC Database</td>
</tr>
<tr>
<td>N/A</td>
<td>N/A</td>
<td></td>
<td>N/A</td>
<td>N/A</td>
<td>Operating system and Oracle software binaries for the third Oracle database server in a four node Oracle RAC Database</td>
</tr>
<tr>
<td>N/A</td>
<td>N/A</td>
<td></td>
<td>N/A</td>
<td>N/A</td>
<td>Operating system and Oracle software binaries for the fourth Oracle database server in a four node Oracle RAC Database</td>
</tr>
<tr>
<td>00:01:1F</td>
<td>00:01:3F</td>
<td>1</td>
<td>/dev/hdisk32</td>
<td>OCRVTDG</td>
<td>Oracle Cluster Registry and Voting Disk</td>
</tr>
<tr>
<td>00:01:00 - 00:01:01</td>
<td>00:01:20 - 00:01:21</td>
<td>2-3</td>
<td>/dev/hdisk01 - /dev/hdisk02</td>
<td>REDODG01</td>
<td>Online REDO log group and control file</td>
</tr>
<tr>
<td>00:01:02 - 00:01:03</td>
<td>00:01:22 - 00:01:23</td>
<td>4-5</td>
<td>/dev/hdisk03 - /dev/hdisk04</td>
<td>REDODG11</td>
<td>Online REDO log group and control file</td>
</tr>
<tr>
<td>00:01:04 - 00:01:1E</td>
<td>00:01:24 - 00:01:3E</td>
<td>6 - 31</td>
<td>/dev/hdisk05 - /dev/hdisk31</td>
<td>DATADG</td>
<td>Sys, Undo, Temp, Application Data</td>
</tr>
</tbody>
</table>
Server Configuration

Table 8 has the server configuration details for the environment.

Table 8. Server Configuration

<table>
<thead>
<tr>
<th>Server Make, Model, Processor</th>
<th>Server Host Name</th>
<th>Role</th>
<th>Operating System</th>
</tr>
</thead>
<tbody>
<tr>
<td>Two IBM Power 770 servers, each with two LPARs.</td>
<td>BCCC-1, BCCC-2, BCCC-3, BCCC-4</td>
<td>Oracle database server with Real Application Cluster</td>
<td>IBM AIX 6.1.0.0 with technology level of 6100-07 and a service pack of 6100-07-03-1207</td>
</tr>
<tr>
<td>Each LPAR configured as</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 PowerPC_POWER7 processors, 3.1 GHz</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22.25 GB RAM</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 × 8 Gb/sec Fibre Channel ports</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note — The operating system for the Oracle database server was booted using the internal drives of the server. You could also use an external drive to perform a SAN boot of the operating system for the Oracle database server.

Figure 3 shows the server infrastructure for the environment.

Figure 3
SAN Configuration

Map the provisioned LDEVs to multiple ports on the VSP G1000. These LDEV port assignments provide multiple paths to the storage system from the Oracle database server for high availability.

The environment used two Fibre Channel switch modules between:

- Server and the storage system
- Storage systems

The Fibre Channel switch modules between the server and the storage system are configured with two inter-switch links (ISL). The Fibre Channel switch modules between the storage systems are configured with three ISLs.

Each database server used one dual-ported HBA. One port was directly connected to the local storage and the other was connected through the switch modules to the remote storage. See Figure 1 on page 4 for the storage connectivity.

Table 9 has the zoning details for the SAN between the server and the storage systems.

Table 9. SAN Switch Architecture

<table>
<thead>
<tr>
<th>Server</th>
<th>HBA Ports</th>
<th>Switch Zone</th>
<th>Storage System</th>
<th>Storage Port</th>
<th>Switch</th>
<th>Use</th>
</tr>
</thead>
<tbody>
<tr>
<td>Database Server 1</td>
<td>P0-HBA1-1</td>
<td>-N/A-</td>
<td>Primary</td>
<td>1J</td>
<td>-N/A-</td>
<td>Database disks</td>
</tr>
<tr>
<td></td>
<td>P0-HBA1-2</td>
<td>P0_HBA1_2</td>
<td>-N/A-</td>
<td>-N/A-</td>
<td>5300-05</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-N/A-</td>
<td>ASE47_96_2J</td>
<td>Secondary</td>
<td>2J</td>
<td>5300-06</td>
<td></td>
</tr>
<tr>
<td>Database Server 3</td>
<td>P0-HBA2-1</td>
<td>-N/A-</td>
<td>Primary</td>
<td>1A</td>
<td>-N/A-</td>
<td></td>
</tr>
<tr>
<td></td>
<td>P0-HBA2-2</td>
<td>P0_HBA2_2</td>
<td>-N/A-</td>
<td>-N/A-</td>
<td>5300-05</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-N/A-</td>
<td>ASE47_96_2A</td>
<td>Secondary</td>
<td>2A</td>
<td>5300-06</td>
<td></td>
</tr>
<tr>
<td>Database Server 2</td>
<td>P1-HBA1-1</td>
<td>-N/A-</td>
<td>Secondary</td>
<td>1L</td>
<td>-N/A-</td>
<td></td>
</tr>
<tr>
<td></td>
<td>P1-HBA1-2</td>
<td>P1_HBA1_2</td>
<td>-N/A-</td>
<td>-N/A-</td>
<td>5300-05</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-N/A-</td>
<td>ASE47_94_2L</td>
<td>Primary</td>
<td>2L</td>
<td>5300-06</td>
<td></td>
</tr>
<tr>
<td>Database Server 4</td>
<td>P1-HBA2-1</td>
<td>-N/A-</td>
<td>Secondary</td>
<td>1C</td>
<td>-N/A-</td>
<td></td>
</tr>
<tr>
<td></td>
<td>P1-HBA2-2</td>
<td>P1_HBA2_2</td>
<td>-N/A-</td>
<td>-N/A-</td>
<td>5300-05</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-N/A-</td>
<td>ASE47_94_2C</td>
<td>Primary</td>
<td>2C</td>
<td>5300-06</td>
<td></td>
</tr>
</tbody>
</table>

To facilitate copying data between storage systems, configure storage replication links between VSP G1000 systems at the primary and secondary sites. Each storage system uses a total of two initiator ports and two target ports.
Table 10 has the zoning details for the SAN between the storage systems.

Table 10. SAN Switch Architecture Between the Storage Systems

<table>
<thead>
<tr>
<th>Storage System</th>
<th>Storage Port</th>
<th>Switch Zone</th>
<th>Switch</th>
<th>Use</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary</td>
<td>5D</td>
<td>ASE47_94_5D</td>
<td>5300-07</td>
<td>Initiator port for SVOS global-active device</td>
</tr>
<tr>
<td></td>
<td>6D</td>
<td>ASE47_94_6D</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>7D</td>
<td>ASE47_94_7D</td>
<td></td>
<td>Target port for SVOS global-active device</td>
</tr>
<tr>
<td></td>
<td>8D</td>
<td>ASE47_94_8D</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Secondary</td>
<td>5D</td>
<td>ASE47_96_5D</td>
<td>5300-08</td>
<td>Target port for SVOS global-active device</td>
</tr>
<tr>
<td></td>
<td>6D</td>
<td>ASE47_96_6D</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>7D</td>
<td>ASE47_96_7D</td>
<td></td>
<td>Initiator port for SVOS global-active device</td>
</tr>
<tr>
<td></td>
<td>8D</td>
<td>ASE47_96_8D</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Hitachi Data Systems recommends the use of dual SAN fabrics, multiple HBAs, and host-based multipathing software when deploying this reference architecture. You need at least two paths to provide the redundancy required for the following critical applications:

- Database hosts connected to two independent SAN fabrics
- SAN fabric to two different controllers of the I/O subsystem
When designing your SAN architecture, follow these recommended practices for a secure, high-performance, and scalable database deployment:

- Use at least two HBAs and place them on different I/O buses within the server. This distributes the workload over the PCI-e bus architecture of the server.

- Use dual SAN fabrics, multiple HBAs, and host-based multipathing software in a business-critical deployment. Connecting two or more paths from the database servers to two independent SAN fabrics provides the redundancy required for critical applications.

- Zone your fabric to handle multiple, unique paths from HBAs to storage ports. Use single initiator zoning. Use at least two Fibre Channel switch fabrics to provide multiple, independent paths to VSP G1000 with SVOS global-active devices to prevent configuration errors from disrupting the entire SAN infrastructure.

- For large bandwidth requirements that exceed the port capability of a single HBA, do the following:
  - Use additional HBAs.
  - Use the round robin load-balancing setting of Hitachi Dynamic Link Manager.
  - Use multiple paths between the primary and remote storage systems and the storage system that contains the quorum disk.
Engineering Validation

This describes the tests performed in the Hitachi Data Systems lab for Oracle Database with the Oracle Real Application Clusters Option Using Hitachi Virtual Storage Platform G1000 with Hitachi Storage Virtualization Operating System global-active device.

Test Methodology

These were the steps followed to test the tech note:

1. Build a four node Oracle RAC Database.
2. Configure SVOS global-active device for synchronous replication between the primary and the secondary VSP G1000 for a zero Recovery Point Objective (RPO) and zero Recovery Time Objective (RTO) solution for an unplanned storage system or site failure.
3. Enable non-preferred path option with Host Mode Option (HMO) 78 on ports 2C and 2L at the primary VSP G1000 and on ports 2A and 2J at the secondary VSP G1000.
4. Enable round trip reduction for True Copy option with HMO 51 on all the SVOS global-active device ports 5D, 6D, 7D and 8D at the primary and secondary VSP G1000.
5. Configure the Quorum disk on Hitachi Unified Storage VM.
6. Configure the global-active device pairs to PAIR state.
7. For a given network bandwidth, using Swingbench to verify the following functions:

- **Active/Active configuration**
  - I/O from Database Server 1 flows to port 1J on the primary VSP G1000.
  - I/O from Database Server 2 flows to port 1L on the secondary VSP G1000.
  - I/O from Database Server 3 flows to port 1A on the primary VSP G1000.
  - I/O from Database Server 4 flows port 1C on the secondary VSP G1000.
  - No I/O on ports 2A and 2J on the secondary VSP G1000.
  - No I/O on ports 2C and 2L on the primary VSP G1000.
  - Verify replication of write I/Os from the initiator SVOS global-active device ports 5D and 6D on the primary VSP G1000, coming from Database Server 1 and Database Server 3, to the SVOLs on the secondary VSP G1000.
  - Verify replication of write I/Os from the initiator SVOS global-active device ports 7D and 8D on the secondary VSP G1000, coming from Database Server 2 and Database Server 4, to the PVOLs on the primary VSP G1000.
  - Using ANUE, simulate the systems at following distance apart, between the local and remote site, to measure the average transactions per second:
    - 0 KM
    - 50 KM
    - 100 KM
Figure 4 shows the transaction flow for this test configuration.

![Figure 4](image-url)
Note — In Active/Active configuration, the non-preferred path will not have an I/O. When the non-preferred path option is enabled, the I/O from the local server goes to the local storage and the I/O from the remote server goes to the remote storage. No I/O goes from the local server to the remote storage unless the local storage system fails. The same also applies to the remote server’s I/O.

- Failover to the secondary VSP G1000 caused by the abrupt failure of the primary VSP G1000 at the local site.
  - All four nodes of the Oracle RAC database are active and functional.
  - I/O from Database Server 1 and Database Server 3 is redirected to ports 2J and 2A on the secondary VSP G1000.
  - I/O from Database Server 2 and Database Server 4 flows to ports 1L and 1C on the secondary VSP G1000.
  - Verify that write I/Os from the initiator SVOS global-active device ports 7D and 8D on the secondary VSP G1000, coming from all four nodes of the Oracle RAC database server, are ready to replicate to the PVOLs when the primary VSP G1000 is restored.
  - Using ANUE, simulate the systems at following distance apart, between the local and remote site, to measure the average transactions per second.
    - 0 KM
    - 50 KM
    - 100 KM
Figure 5 shows the transaction flow for this test configuration.

Note — In a failover configuration, the non-preferred path will have I/O, going from the local server to the active remote storage.

- Failback to the primary VSP G1000.
  - Same as Active/Active configuration.
  - A test passes if each function works as intended in the tech note.
Workload
Testing included test iterations to run simulated and synthetic workloads using Swingbench.

These were the simulated workloads for VSP G1000 with SVOS global-active device testing:

- Order Entry profile configured with minimum delay between transactions of 850 ms and maximum delay between transactions of 1050 ms.
- Order Entry profile configured with no delay between transactions.

Results Summary
This summarizes the key observations from the test results for Hitachi Platform for Oracle Database with the Oracle RAC Option Using VSP G1000 with SVOS global-active device.

Figure 6 shows the average transactions per second for systems at 0 KM apart between the local and the remote site using Swingbench configured with delay between the transactions.

![Graph showing average transactions per second for systems at 0 KM apart with delay.](image)

**Figure 6**
Figure 7 shows the average transactions per second for systems at 50 KM apart between the local and the remote site using Swingbench configured with delay between the transactions.

![Average Transactions/Sec for Systems at 50KM Apart with Delay](image)

Figure 7

Figure 8 shows the average transactions per second for systems at 100 KM apart between the local and the remote site using Swingbench configured with delay between the transactions.

![Average Transactions/Sec for Systems at 100KM Apart with Delay](image)

Figure 8
Figure 9 shows the average transactions per second for systems at 0 KM apart between the local and the remote site using Swingbench configured with no delay between the transactions.

Figure 9

Note — Swingbench does not have an interface to replay the same transactions. The average transactions per second reported from Swingbench for the different test cases may vary. This is more predominant when there is no delay between the transactions.
Figure 10 shows the average transactions per second for systems at 50 KM apart between the local and the remote site using Swingbench configured with no delay between the transactions.

**Figure 10**

Figure 11 shows the average transactions per second for systems at 100 KM apart between the local and the remote site using Swingbench configured with no delay between the transactions.

**Figure 11**
Conclusions

This tech note documents the high level design of Oracle Database with RAC option using global-active device on Hitachi Virtual Storage Platform G1000. Testing this solution validated that this Hitachi platform for Oracle Database environment using Hitachi Storage Virtualization Operating System global-active device for VSP G1000 provides an infrastructure with continuous availability:

- Hitachi Dynamic Link Manager provides host multipathing with deep storage integration.
- Global-active device permits you to use host-based applications to failover between Hitachi Virtual Storage Platform G1000 systems with no disruption.

These components are tightly integrated with Oracle Database infrastructures to protect the data center from various failure scenarios. You can leverage the benefits of synchronous data replication beyond traditional disaster recovery protection to include workload balancing across sites and coordinated site-wide maintenance with no downtime.
Acknowledgments

The following people were part of the team that made this solution possible:

- Sushil Dhilpe
- Catherine Anderson
- Leticia Fernandez
For More Information

Hitachi Data Systems Global Services offers experienced storage consultants, proven methodologies and a comprehensive services portfolio to assist you in implementing Hitachi products and solutions in your environment. For more information, see the Hitachi Data Systems Global Services website.

Live and recorded product demonstrations are available for many Hitachi products. To schedule a live demonstration, contact a sales representative. To view a recorded demonstration, see the Hitachi Data Systems Corporate Resources website. Click the Product Demos tab for a list of available recorded demonstrations.

Hitachi Data Systems Academy provides best-in-class training on Hitachi products, technology, solutions and certifications. Hitachi Data Systems Academy delivers on-demand web-based training (WBT), classroom-based instructor-led training (ILT) and virtual instructor-led training (vILT) courses. For more information, see the Hitachi Data Systems Services Education website.

For more information about Hitachi products and services, contact your sales representative or channel partner or visit the Hitachi Data Systems website.