Hitachi Unified Compute Platform Select for Oracle Database with Hitachi Unified Storage VM and Hitachi Accelerated Flash Using a N+1 Cold Standby Server and Oracle Real Application Clusters

Reference Architecture Guide

By Kishore Daggubati

December 11, 2013
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

- Hitachi Compute Blade 2000 ..................................................7
- Hitachi Unified Storage VM .....................................................7
- Hitachi Dynamic Provisioning ..................................................8
- Hitachi Storage Navigator .......................................................8
- Hitachi Compute Systems Manager .........................................9
- Oracle Linux ............................................................................9
- Oracle Database with the Real Application Cluster Option .......9

**Solution Design** ..................................................................10

- Storage Architecture ............................................................10
- Server and Application Architecture ......................................15
- SAN Architecture ................................................................18
- Network Architecture ...........................................................19

**Engineering Validation** .......................................................23

- Test Methodology ...............................................................23
- Test Results ...........................................................................25

**Conclusion** ..........................................................................28
Hitachi Unified Compute Platform Select for Oracle Database with Hitachi Unified Storage VM and Hitachi Accelerated Flash Using a N+1 Cold Standby Server and Oracle Real Application Clusters

Reference Architecture Guide

This reference architecture guide shows how the medium size configuration option for Hitachi Unified Compute Platform Select for Oracle Database with the Oracle Real Application Clusters provides a high performance, integrated solution for an Oracle infrastructure.

The environment uses a N+1 cold standby server with Hitachi Compute Systems Manager and Hitachi Unified Storage VM with Hitachi Accelerated Flash. This is a resource to design an infrastructure for your requirements and budget.

The use of a N+1 cold standby server allows using Oracle licenses only for the active blades with the same possible power.

This validated solution integrates servers, storage systems, and storage software. The environment provides reliability, high availability, scalability, and performance while processing small-scale to large-scale OLTP workloads. The dedicated server runs Oracle Database 11g R2 with the Oracle Real Application Clusters option. The operating system is Oracle Linux using the Unbreakable Enterprise Kernel 6, Update 1.

You benefit from using this reference architecture if you are one of the following:

- Database administrator
- Storage administrator
- A person with the responsibility to plan and deploy a Oracle Database 11g R2 solution using the Oracle Real Application Clusters option
To use this reference architecture guide, you need familiarity with the following:

- Hitachi Unified Storage VM using Hitachi Accelerated Flash
- Hitachi Compute Systems Manager
- Hitachi Compute Blade 2000
- Storage area networks
- Oracle RAC Database 11g Release 2
- Oracle Automatic Storage Management

**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.
Solution Overview

This reference architecture discusses implementing Hitachi Unified Compute Platform Select for Oracle Database with the Real Application Clusters Option solution using the following:

- N+1 cold standby server with Hitachi Compute Systems Manager
- Hitachi Unified Storage VM with Hitachi Accelerated Flash

It includes high availability, performance, and scalability benefits for typical on-line transaction processing workloads. Tailor your implementation of this solution to meet your specific needs.

This reference architecture guide includes a discussion of the following:

- Hitachi Compute Blade 2000 using three server blades
  - Two server blades, each hosts one Oracle database server
  - One server blade hosts the N+1 cold standby server
- Hitachi Unified Storage VM with Hitachi Accelerated Flash
- SAN infrastructure

Figure 1 on page 4 shows the infrastructure for this solution.
Figure 1
Key Solution Components

Table 1, starting on page 6, and Table 2 on page 7 have the key hardware and software components used in this reference architecture.

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 Chassis</td>
<td>Hitachi Compute Blade 2000</td>
<td>Firmware Version A0335-E-7174</td>
<td>1</td>
</tr>
<tr>
<td>Server Blade</td>
<td>Model GVAX57A2 (X57-A2), each configured as follows:</td>
<td>EFI BIOS Version 4.6.3.7</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>▪ Intel Xeon E7-8870 at 2.40 GHz, two 10-core physical processors</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>▪ 128 GB RAM using 8 GB DIMMs</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>▪ Two dual port 8 Gb/sec PCIe Fibre Channel cards</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>▪ One onboard dual port 1 Gb/sec Ethernet NIC</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>▪ One quad port 1 Gb/sec Ethernet NIC (Mezzanine Slot 0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Storage System</td>
<td>Hitachi Unified Storage VM configured as follows:</td>
<td>73-02-03-00/00</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>▪ 1 Flash chassis with:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>▪ 16 × 1.6 TB flash modules in 2 flash units with 8 flash modules in each unit, for about 25 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 MP blades, each with two 8-core Intel XEON processor; 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>▪ One front-end port on each controller connects to Hitachi Compute Blade 2000</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
This reference architecture uses two Brocade 6510, 8 Gb/sec Fibre Channel switches for SAN connectivity. You may choose to use specific Fibre Channel switches to meet your needs. In addition, you can design a larger infrastructure by adding additional cache and front-end connectivity modules on Hitachi Unified Storage VM.

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>SAN Connectivity</td>
<td>Brocade 6510, 8 Gb/sec Fibre Channel switch</td>
<td>Firmware Version 7.0.1a</td>
<td>2</td>
</tr>
<tr>
<td>Network Connectivity</td>
<td>1/10 GigE L3 LAN switch modules</td>
<td>Firmware Version 10.7</td>
<td>4</td>
</tr>
<tr>
<td>Web Console</td>
<td>Microsoft® Windows Server® Datacenter configured as follows:</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>· Quad Core AMD Opteron 8347 HE 1.90 GHz, two processors</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>· 2 GB RAM</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>· 32-bit O/S</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

This reference architecture uses N+1 cold standby server software on Hitachi Compute Blade 2000. You may choose to use this feature on Hitachi Compute Blade 500, instead.

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>Oracle Linux using Unbreakable Enterprise Kernel</td>
<td>6 Update 1</td>
</tr>
<tr>
<td>Multipath Software</td>
<td>Linux 6 Device Mapper</td>
<td>0.4.9</td>
</tr>
<tr>
<td>Volume Manager and File System</td>
<td>Oracle Automatic Storage Management</td>
<td>11g R2, 11.2.0.1.0</td>
</tr>
<tr>
<td>Database Software</td>
<td>Oracle</td>
<td>11g R2, 11.2.0.1.0</td>
</tr>
<tr>
<td>Cluster Software</td>
<td>Oracle Real Application Cluster</td>
<td>11g R2, 11.2.0.1.0</td>
</tr>
<tr>
<td>Storage Management Software</td>
<td>Hitachi Storage Navigator</td>
<td>Microcode dependent</td>
</tr>
<tr>
<td></td>
<td>Hitachi Dynamic Provisioning</td>
<td>Microcode dependent</td>
</tr>
<tr>
<td>I/O Calibration Software</td>
<td>Orion</td>
<td>11.1</td>
</tr>
<tr>
<td>N+1 Cold Standby Server Software</td>
<td>Hitachi Compute Systems Manager</td>
<td>7.5.1-01</td>
</tr>
<tr>
<td>Database Client Communication Software</td>
<td>Oracle Net Services</td>
<td>11g R2, 11.2.0.1.0</td>
</tr>
</tbody>
</table>
Hitachi Compute Blade 2000

Hitachi Compute Blade 2000 is an enterprise-class blade server platform. It features the following:

- A balanced system architecture that eliminates bottlenecks in performance and throughput
- Configuration flexibility
- Eco-friendly power-saving capabilities
- Fast server failure recovery using a N+1 cold standby design that allows replacing failed servers within minutes

Hitachi Unified Storage VM

Hitachi Unified Storage VM is an entry-level enterprise storage platform. It combines storage virtualization services with unified block, file, and object data management. This versatile, scalable platform offers a storage virtualization system to provide central storage services to existing storage assets.

Unified management delivers end-to-end central storage management of all virtualized internal and external storage on Unified Storage VM. A unique, hardware-accelerated, object-based file system supports intelligent file tiering and migration, as well as virtual NAS functionality, without compromising performance or scalability.

The benefits of Unified Storage VM are the following:

- Enables the move to a new storage platform with less effort and cost when compared to the industry average
- Increases performance and lowers operating cost with automated data placement
- Supports scalable management for growing and complex storage environment while using fewer resources
- Achieves better power efficiency and with more storage capacity for more sustainable data centers
- Lowers operational risk and data loss exposure with data resilience solutions
- Consolidates management with end-to-end virtualization to prevent virtual server sprawl
Hitachi Dynamic Provisioning

On Hitachi storage systems, Hitachi Dynamic Provisioning provides wide striping and thin provisioning functionalities.

Using Dynamic Provisioning is like using a host-based logical volume manager (LVM), but without incurring host processing overhead. It provides one or more wide-striping pools across many RAID groups. Each pool has one or more dynamic provisioning virtual volumes (DP-VOLs) of a logical size you specify of up to 60 TB created against it without allocating any physical space initially.

Deploying Dynamic Provisioning avoids the routine issue of hot spots that occur on logical devices (LDEVs). These occur within individual RAID groups when the host workload exceeds the IOPS or throughput capacity of that RAID group. Dynamic provisioning distributes the host workload across many RAID groups, which provides a smoothing effect that dramatically reduces hot spots.

When used with Hitachi Virtual Unified Storage Platform VM, Hitachi Dynamic Provisioning has the benefit of thin provisioning. Physical space assignment from the pool to the dynamic provisioning volume happens as needed using 42 MB pages, up to the logical size specified for each dynamic provisioning volume. There can be a dynamic expansion or reduction of pool capacity without disruption or downtime. You can rebalance an expanded pool across the current and newly added RAID groups for an even striping of the data and the workload.

Hitachi Storage Navigator

Hitachi Storage Navigator enables essential management and optimization functions. Using Java agents, Storage Navigator runs on most browsers. A command line interface is available.

Use Storage Navigator for the following:

- Pool creation and expansion
- LUN creation and expansion
- Online microcode updates and other system maintenance functions
- Performance metrics

You need Storage Navigator to take advantage of the full features of Hitachi Unified Storage VM.
Hitachi Compute Systems Manager

Hitachi Compute Systems Manager provides centralized enterprise infrastructure management. Use it to monitor, manage, and configure Hitachi server hardware, physical hosts, and virtual hosts.

Compute Systems Manager has a Web-based interface. The agentless architecture allows taking advantage of available all computing resources. Compute Systems Manager can manage and monitor up to 10,000 physical or virtual servers.

Fully integrated within Hitachi Command Suite, you have seamless management of server and storage resources with one simple interface using Compute Systems Manager.

Oracle Linux

Oracle Linux is an enterprise-class operating system built and tested to run Oracle hardware, databases, and middleware. It is fully compatible with the Red Hat Enterprise Linux kernel.

Oracle Database with the Real Application Cluster Option

Oracle Database is optimized for use with Oracle products. This solution uses Oracle Real Application Cluster and Oracle Automatic Storage Management, which are part of the grid infrastructure component in Oracle Database.

- **Real Application Cluster** (RAC) scales the database across multiple servers and protects against server failure.

- **Automatic Storage Management** (ASM) combines the features of a volume manager and an application-optimized file system for database files.
Solution Design

This is the detailed description of the reference architecture environment implementing Hitachi Unified Compute Platform Select for Oracle Database with the Oracle Real Application Clusters option using the following:

- N+1 cold standby server with Hitachi Compute Systems Manager
- Hitachi Unified Storage VM with Hitachi Accelerated Flash

Specific infrastructure configuration includes the following:

- **Server** — This is a single server node, consisting of one X57-A2 server blade for one logical Oracle database server.
- **N+1 Cold Standby Server** — This is a single server node, consisting of one X57-A2 server blade for one N+1 cold standby server.
- **Storage System** — There are LDEVs mapped to each port that are presented to the server as LUNs.
- **SAN Fabric** — There are six zones created on each Fibre Channel switch module to zone the two PCIe Fibre Channel dual ports on each server blade and the twelve storage host ports.

Storage Architecture

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

Storage Configuration

Create the RAID groups and storage pools with Hitachi Dynamic Provisioning on Hitachi Unified Storage VM shown in Figure 2 on page 12.
Table 3 has the details for the dynamic provisioning pool.

Table 3. Dynamic Provisioning Pool

<table>
<thead>
<tr>
<th>Dynamic Provisioning Pool ID</th>
<th>Parity Group</th>
<th>RAID Level</th>
<th>Drive Type</th>
<th>No of Drives</th>
<th>Pool Capacity</th>
</tr>
</thead>
<tbody>
<tr>
<td>ora_ssdpool</td>
<td>2-23 – 2-26</td>
<td>RAID-10 (2D+2D)</td>
<td>1.6 TB Flash Module Drives</td>
<td>16</td>
<td>11.95 TB</td>
</tr>
</tbody>
</table>

Table 4 on page 13 has the details for the virtual volumes created from the dynamic provisioning pool.
Table 4. Dynamic Provisioning Pool LUN Information

<table>
<thead>
<tr>
<th>Dynamic Provisioning Pool ID</th>
<th>LDEVs</th>
<th>LDEV Size (GB)</th>
<th>Purpose</th>
<th>Storage Port</th>
</tr>
</thead>
<tbody>
<tr>
<td>ora_ssdpool</td>
<td>00:00:24</td>
<td>200</td>
<td>▪ Operating system and Oracle software binaries for the first Oracle database server in a two node Oracle RAC Database</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>▪ Operating system and Oracle software binaries for the N+1 cold standby server in the event of unavailability of the first Oracle database server in a two node Oracle RAC Database</td>
<td>1A, 3A, 5A, 1B, 3B, 5B, 2A, 4A, 6A, 2B, 4B, 6B</td>
</tr>
<tr>
<td></td>
<td>00:00:25</td>
<td>200</td>
<td>▪ Operating system and Oracle software binaries for the second Oracle database server in a two node Oracle RAC Database</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>▪ Operating system and Oracle software binaries for the N+1 cold standby server in the event of unavailability of the second Oracle database server in a two node Oracle RAC Database</td>
<td></td>
</tr>
</tbody>
</table>
|                              | 00:00:26 00:00:2A | 200   | ▪ Oracle Archived Redo Logs  
▪ Oracle Cluster Registry  
▪ Voting Disk |              |
|                              | 00:00:2B | 200           | ▪ Oracle Online Redo Logs  
▪ Oracle Control Files |              |
|                              | 00:00:2C | 200           | ▪ Oracle Online Redo Logs  
▪ Oracle Control Files |              |
|                              | 00:00:2D – 00:00:3B | 200 | ▪ Flash Recovery Area |              |
|                              | 00:00:3C – 00:00:5F | 200 | ▪ Oracle System  
▪ Sysaux  
▪ Undo  
▪ Temp  
▪ OLTP Application Tablespaces |              |
Database Layout

The database layout design uses recommended practices from Hitachi Data Systems for Hitachi Unified Storage VM using Hitachi Accelerated Flash for small random I/O traffic, such as those in OLTP transactions. The layout also takes into account the Oracle ASM 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 Data ASM diskgroup for the data and index tablespaces. 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 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 UNDO tablespace in this configuration in the Data ASM diskgroup. Assign one UNDO tablespace for each database instance in a two node Oracle RAC database. 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. Quite a number of redo logs are created for each database instance in a two node Oracle RAC database.

- **Archive Redo Logs** — Assign one ASM diskgroup for the Oracle RAC database.

- **Flash Recovery Area** — Assign one ASM diskgroup for the Oracle RAC database.

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

- **Size Settings** — Set the database block size to 8 KB. Set ASM allocation unit to 1 MB.

Table 5 on page 15 lists the disk mappings from the LUNS to the operating system devices and to the ASM disk groups.
<table>
<thead>
<tr>
<th>LDEV's</th>
<th>LUN</th>
<th>Operating System Device</th>
<th>ASM Disk</th>
<th>ASMDG</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>00:00:24</td>
<td>0000</td>
<td>/dev/mapper/vg_oraracnode1</td>
<td>N/A</td>
<td>N/A</td>
<td>- Operating system and Oracle software binaries for the first Oracle database server in a two node Oracle RAC Database</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Operating system and Oracle software binaries for the N+1 cold standby server in the event of unavailability of the first Oracle database server in a two node Oracle RAC Database</td>
</tr>
<tr>
<td>00:00:25</td>
<td></td>
<td>/dev/mapper/vg_oraracnode2</td>
<td>N/A</td>
<td>N/A</td>
<td>- Operating system and Oracle software binaries for the second Oracle database server in a two node Oracle RAC Database</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Operating system and Oracle software binaries for the N+1 cold standby server in the event of unavailability of the second Oracle database server in a two node Oracle RAC Database</td>
</tr>
<tr>
<td>00:00:26 -</td>
<td>0001 -</td>
<td>/dev/mapper/mpathaap1 -</td>
<td>ARDISK01 -</td>
<td>ARCHDG</td>
<td>- Oracle Archived Redo Logs, Oracle Cluster Registry and Voting Disk</td>
</tr>
<tr>
<td>00:00:2A</td>
<td>0005</td>
<td>/dev/mapper/mpathaep1</td>
<td>ARDISK05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>00:00:2B</td>
<td>0006</td>
<td>/dev/mapper/mpathafp1</td>
<td>RGDISK01</td>
<td>REDODG01</td>
<td>- Online REDO log group and control file</td>
</tr>
<tr>
<td>00:00:2C</td>
<td>0007</td>
<td>/dev/mapper/mpathagp1</td>
<td>RGDISK11</td>
<td>REDODG11</td>
<td>- Online REDO log group and control file</td>
</tr>
<tr>
<td>00:00:2D -</td>
<td>0008 -</td>
<td>/dev/mapper/mpathahp1 -</td>
<td>FRDISK01 -</td>
<td>FRADG</td>
<td>- Oracle Flash Recovery Area</td>
</tr>
<tr>
<td>00:00:3B</td>
<td>0022</td>
<td>/dev/mapper/mpathavp1</td>
<td>FRDISK15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>00:00:3C -</td>
<td>0023 -</td>
<td>/dev/mapper/mpathawp1 -</td>
<td>DADISK01 -</td>
<td>DATADG</td>
<td>- Sys</td>
</tr>
<tr>
<td>00:00:5F</td>
<td>0058</td>
<td>/dev/mapper/mpathywp1</td>
<td>DADISK35</td>
<td></td>
<td>- Undo</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Temp</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Application Data</td>
</tr>
</tbody>
</table>
Server and Application Architecture

The reference architecture uses three server blades in the Hitachi Compute Blade 2000 chassis.

- Two server blades, each hosts one Oracle database server
  - The Oracle database server includes a single logical database instance
- One server blade hosts N+1 cold standby server

There are 20 combined CPU cores with 128 GB RAM for each server blade. This provides the compute power for Oracle RAC Database to handle complex database queries and a large volume of transaction processing in parallel. See “Key Solution Components” on page 5 for hardware details.

Figure 3 on page 17 shows the server infrastructure for the reference architecture.
Figure 3

The N+1 cold standby server is brought online using Hitachi Compute Systems Manager for one of the following reasons:

- Unexpected hardware failure of the database server
- Scheduled maintenance of the database server
- Maintain Oracle RAC environment when one of the database servers is offline

The N+1 cold standby server boots from the boot LUN of the failed database server and resumes its environment. This provides availability, serviceability, and reliability of Oracle data in a two node Oracle RAC database during the unavailability of the failed database server.
Figure 4 shows the failover architecture of one of the database server to the N+1 cold standby server and maintaining Oracle RAC environment.

This reference architecture does not use a third party volume manager. Instead, it uses Oracle RAC Database 11g R2 ASM volume manager.
To monitor and manage the database, the database servers, the N+1 cold standby server, and the storage, this solution uses the following:

- Oracle Database 11g R2 Automatic Workload Repository (AWR) reports
- Hitachi Storage Navigator
- O/S IOSTAT and VMSTAT
- Hitachi Compute Systems Manager

**SAN Architecture**

Map the provisioned LDEVs to multiple ports on Hitachi Unified Storage VM using Hitachi Accelerated Flash. These LDEV port assignments provide multiple paths to the storage system from the host for high availability.

The environment used two Fibre Channel switch modules. This provides scalability and high availability. "Key Solution Components" on page 5 has host configuration details.

Each server in this reference architecture used four Fibre Channel ports, with two ports from each of the PCIe HBA Emulex cards. This provides an eight-path connection for all LUNs mapped to each of the database servers in a two node Oracle RAC Database.

The environment used two Fibre Channel switch modules to provide redundant paths for all Hitachi Unified Storage VM using Hitachi Accelerated Flash LUNs.

Table 6 provides the zoning details for the SAN.

<table>
<thead>
<tr>
<th>Server</th>
<th>HBA Ports</th>
<th>Switch Zone</th>
<th>Storage Port</th>
<th>Switch Bay ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Database server 1</td>
<td>B0-PCI_0_HBA1-1</td>
<td>BS2K_16_B0_PCI_0_HBA1_1_ASE45_190_1A</td>
<td>1A</td>
<td>Bay 1</td>
</tr>
<tr>
<td></td>
<td>B0-PCI_0_HBA1-2</td>
<td>BS2K_16_B0_PCI_0_HBA1_2_ASE45_190_2A</td>
<td>2A</td>
<td>Bay 2</td>
</tr>
<tr>
<td></td>
<td>B0-PCI_1_HBA1-1</td>
<td>BS2K_16_B0_PCI_1_HBA1_1_ASE45_190_1B</td>
<td>1B</td>
<td>Bay 1</td>
</tr>
<tr>
<td></td>
<td>B0-PCI_1_HBA1-2</td>
<td>BS2K_16_B0_PCI_1_HBA1_2_ASE45_190_2B</td>
<td>2B</td>
<td>Bay 2</td>
</tr>
<tr>
<td>Database server 2</td>
<td>B1-PCI_2_HBA1-1</td>
<td>BS2K_16_B1_PCI_2_HBA1_1_ASE45_190_3A</td>
<td>3A</td>
<td>Bay 1</td>
</tr>
<tr>
<td></td>
<td>B1-PCI_2_HBA1-2</td>
<td>BS2K_16_B1_PCI_2_HBA1_2_ASE45_190_4A</td>
<td>4A</td>
<td>Bay 2</td>
</tr>
<tr>
<td></td>
<td>B1-PCI_3_HBA1-1</td>
<td>BS2K_16_B1_PCI_3_HBA1_1_ASE45_190_3B</td>
<td>3B</td>
<td>Bay 1</td>
</tr>
<tr>
<td></td>
<td>B1-PCI_3_HBA1-2</td>
<td>BS2K_16_B1_PCI_3_HBA1_2_ASE45_190_4B</td>
<td>4B</td>
<td>Bay 2</td>
</tr>
<tr>
<td>N+1 cold standby server</td>
<td>B2-PCI_4_HBA1-1</td>
<td>BS2K_16_B2_PCI_4_HBA1_1_ASE45_190_5A</td>
<td>5A</td>
<td>Bay 1</td>
</tr>
<tr>
<td></td>
<td>B2-PCI_4_HBA1-2</td>
<td>BS2K_16_B2_PCI_4_HBA1_2_ASE45_190_6A</td>
<td>6A</td>
<td>Bay 2</td>
</tr>
<tr>
<td></td>
<td>B2-PCI_5_HBA1-1</td>
<td>BS2K_16_B2_PCI_5_HBA1_1_ASE45_190_5B</td>
<td>5B</td>
<td>Bay 1</td>
</tr>
<tr>
<td></td>
<td>B2-PCI_5_HBA1-2</td>
<td>BS2K_16_B2_PCI_5_HBA1_2_ASE45_190_6B</td>
<td>6B</td>
<td>Bay 2</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 required redundancy to critical applications for the following:

- 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 Hitachi Unified Storage VM using Hitachi Accelerated Flash 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 for Oracle Linux using the Unbreakable Enterprise Kernel Device Mapper.

**Network Architecture**

This architecture requires the following separate networks:

- **Private Network** (also called **Cluster Interconnect**) — This network must be scalable. In addition, it must meet the low latency needs of the network traffic generated by cache synchronization of Oracle Real Application Clusters and inter-node communication among the nodes in the cluster.
- **Public Network** — This network provides client connections to the applications and Oracle Real Application Clusters.

Hitachi Data Systems recommends using a pair of 1 Gb/sec NICs for the cluster interconnect and public network.
Each server in this reference architecture has one dual port 1 Gb/sec on-board NIC and one quad port 1 Gb/sec Ethernet NIC (Mezzanine Slot 0) for different types of data traffic. The NIC ports on each server blade connect to four internal 1/10 Gb/sec Ethernet switches in the chassis.

Observe these points when configuring private and public networks in your Oracle RAC environment.

- For each server in the Oracle RAC Clusterware configuration, use at least two identical, high bandwidth, low-latency NICs for the interconnection.
- Use NIC bonding to provide fail over and load balancing of interconnections within a server.
- Set all NICs to full duplex mode.
- Use at least two public NICs for client connections to the application and database.
- Use at least two private NICs for the cluster interconnection.

Figure 5 on page 22 shows the network configuration for the reference architecture environment.
Table 7. Network Architecture

<table>
<thead>
<tr>
<th>Server</th>
<th>NIC Ports</th>
<th>Switch Bay ID</th>
<th>Switch Port (Internal)</th>
<th>VLAN</th>
<th>Subnet</th>
<th>IP Address</th>
<th>Network</th>
</tr>
</thead>
<tbody>
<tr>
<td>Database server 1</td>
<td>B0-NIC-0</td>
<td>0</td>
<td>5</td>
<td>170</td>
<td>170</td>
<td>172.17.170.189</td>
<td>Public</td>
</tr>
<tr>
<td></td>
<td>B0-NIC-1</td>
<td>1</td>
<td>5</td>
<td>-</td>
<td>208</td>
<td>10.10.208.2</td>
<td>Private</td>
</tr>
<tr>
<td></td>
<td>B0-MNIC-0</td>
<td>2</td>
<td>5</td>
<td>170</td>
<td>170</td>
<td>172.17.170.190</td>
<td>Public</td>
</tr>
<tr>
<td></td>
<td>B0-MNIC-2</td>
<td>3</td>
<td>5</td>
<td>-</td>
<td>209</td>
<td>10.10.209.2</td>
<td>Private</td>
</tr>
</tbody>
</table>

**Figure 5**

Table 7 has the network details for this solution.
Table 7. Network Architecture (Continued)

<table>
<thead>
<tr>
<th>Server</th>
<th>NIC Ports</th>
<th>Switch Bay ID</th>
<th>Switch Ports (Internal)</th>
<th>VLAN</th>
<th>Subnet</th>
<th>IP Address</th>
<th>Network</th>
</tr>
</thead>
<tbody>
<tr>
<td>Database server 2</td>
<td>B1-NIC-0</td>
<td>0</td>
<td>6</td>
<td>170</td>
<td>170</td>
<td>172.17.170.191</td>
<td>Public</td>
</tr>
<tr>
<td></td>
<td>B1-NIC-1</td>
<td>1</td>
<td>6</td>
<td>-</td>
<td>208</td>
<td>10.10.208.3</td>
<td>Private</td>
</tr>
<tr>
<td></td>
<td>B1-MNIC-0</td>
<td>2</td>
<td>6</td>
<td>170</td>
<td>170</td>
<td>172.17.170.192</td>
<td>Public</td>
</tr>
<tr>
<td></td>
<td>B1-MNIC-2</td>
<td>3</td>
<td>6</td>
<td>-</td>
<td>209</td>
<td>10.10.209.3</td>
<td>Private</td>
</tr>
<tr>
<td>N+1 cold standby server</td>
<td>B2-NIC-0</td>
<td>0</td>
<td>7</td>
<td>170</td>
<td>170</td>
<td>The IP address of the failed database server is assigned</td>
<td>Public</td>
</tr>
<tr>
<td></td>
<td>B2-NIC-1</td>
<td>1</td>
<td>7</td>
<td>-</td>
<td>208</td>
<td>Private</td>
<td></td>
</tr>
<tr>
<td></td>
<td>B2-MNIC-0</td>
<td>2</td>
<td>7</td>
<td>170</td>
<td>170</td>
<td>Public</td>
<td></td>
</tr>
<tr>
<td></td>
<td>B2-MNIC-2</td>
<td>3</td>
<td>7</td>
<td>-</td>
<td>209</td>
<td>Private</td>
<td></td>
</tr>
</tbody>
</table>
Engineering Validation

This is the functionality and performance validation of the reference architecture for Hitachi Unified Compute Platform Select for Oracle Database with the Oracle Real Application Clusters option using the following:

- N+1 cold standby server with Hitachi Compute Systems Manager
- Hitachi Unified Storage VM with Hitachi Accelerated Flash

Testing involved the following:

- Configuration for a two node Oracle RAC Database
- Fail over and fail back operations using Hitachi Compute Systems Manager

Test Methodology

This is the methodology used to validate this solution.

Workload

Testing included test iterations to run simulated and synthetic workloads using the following:

- **Oracle Orion** — This test software simulates the Oracle database server I/O patterns without installing Oracle or creating a database. Unlike other I/O calibration tools, Orion simulates Oracle database I/O workloads by design. When used for testing, Orion determined the I/O throughput and bandwidth for the server configuration and the workloads.

A storage capacity of approximately 11.60 TB is used for Oracle Orion testing.

These were the simulated workloads for Oracle Orion testing:

- 100% read transactions
- 90% read transactions, 10% write transactions
Testing Procedure
This is the testing procedure used to test the environment:

- **Small Random Reads Test** — Measured the IOPS using Orion
- **Large Sequential I/O Test** — Measured the throughput of large sequential I/Os using Orion
- **N+1 Cold Standby Server** — These were the steps followed to test the fail over and fail back operations using Hitachi Compute Systems Manager:
  1. Build N+1 cold standby server using Hitachi Compute Systems Manager
  2. For a given network bandwidth, measure Recovery Time Objective (RTO) for the following functions:
     - Fail over to the N+1 cold standby server caused by the abrupt failure of one of the database server in a two node Oracle RAC database
     - Fail back from the N+1 cold standby server to the database server, that is operationally resumed after taking corrective actions, in a two node Oracle RAC database

A test passes if each function works as intended in the reference architecture.

Data Gathering
Collection of the performance statistics was at the following levels:

- **Storage**
  - Hitachi Storage Navigator collected storage performance data

- **Operating System**
  - “iostat” and “vmstat” collected operating system statistics

- **Standalone tests on the drives/LUNS**
  - Oracle Orion for the throughput and IOPS performance for a given set of workload

Collection of the fail over and fail back statistics for N+1 cold standby was at the following levels:

- Oracle database instance alert logs
- Hitachi Compute Systems Manager to verify each function worked
Test Results

This summarizes the key observations from the test results for the reference architecture for Hitachi Unified Compute Platform Select for Oracle Database with the Oracle Real Application Clusters option using the following:

- N+1 cold standby server with Hitachi Compute Systems Manager
- Hitachi Unified Storage VM with Hitachi Accelerated Flash

Small Random Reads Test
Oracle Orion testing demonstrated increased IOPS for different workload distribution on a two node Oracle RAC Database.

Table 8 has the details of Orion IOPS for different workload distributions.

<table>
<thead>
<tr>
<th>Workload Distribution</th>
<th>IOPS</th>
</tr>
</thead>
<tbody>
<tr>
<td>100% read transactions</td>
<td>431,423</td>
</tr>
<tr>
<td>90% read transactions, 10% write transactions</td>
<td>279,932</td>
</tr>
</tbody>
</table>

Large Sequential I/O Test
Oracle Orion testing demonstrated increased throughput for different workload distribution on a two node Oracle RAC Database.

Table 9 has the details of ORION throughput for different workload distributions.

<table>
<thead>
<tr>
<th>Workload Distribution</th>
<th>Throughput in MBPS</th>
</tr>
</thead>
<tbody>
<tr>
<td>100% read transactions</td>
<td>6,045</td>
</tr>
<tr>
<td>90% read transactions, 10% write transactions</td>
<td>5,835</td>
</tr>
</tbody>
</table>

N+1 Cold Standby Server Test
This summarizes the key observations from the test results for N+1 cold standby server on Hitachi Unified Compute Platform Select for Oracle RAC Database reference solution. All functions described in “Test Methodology” on page 24 worked as intended. Using the Oracle database instance alert logs and Hitachi Compute Systems Manager confirmed the N+1 cold standby server had the correct status for each function tested.

The Service Processor (SVP) web console of Hitachi Compute Blade 2000 is used to enable the N+1 cold standby server.
Table 10 has the configuration details of the N+1 cold standby server in the Hitachi Compute Blade 2000 SVP.

**Table 10. N+1 Cold Standby Configuration in the Hitachi Compute Blade 2000 SVP**

<table>
<thead>
<tr>
<th>Partition</th>
<th>Server</th>
<th>N+1 Cold Standby Server</th>
<th>Pre-Configure Method</th>
<th>WWN Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Partition 0</td>
<td>Database server 1</td>
<td>Enable</td>
<td>Inline</td>
<td>Additional</td>
</tr>
<tr>
<td>Partition 1</td>
<td>Database server 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Partition 2</td>
<td>N+1 cold standby server</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 11 has the Hitachi Compute Blade status in the Hitachi Compute Systems Manager between the failed database server and the N+1 cold standby server before performing the fail over.

**Table 11. Hitachi Compute Blade Status in the Hitachi Compute Systems Manager Before the Fail Over**

<table>
<thead>
<tr>
<th>Blade</th>
<th>Server Name</th>
<th>Power Status</th>
<th>Chassis Name</th>
<th>Slot No.</th>
<th>Blade Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Active Blade</td>
<td>Database server 1</td>
<td>ON</td>
<td>4601L1000461</td>
<td>0</td>
<td>Unavailable for fail over</td>
</tr>
<tr>
<td></td>
<td>Database server 2</td>
<td>OFF</td>
<td></td>
<td>1</td>
<td>Available for fail over</td>
</tr>
<tr>
<td>Standby Blade</td>
<td>N+1 cold standby server</td>
<td>OFF</td>
<td></td>
<td>2</td>
<td>Available for fail over</td>
</tr>
</tbody>
</table>

Figure 6 shows the N+1 group configuration status between the failed database server and the N+1 cold standby server before performing the fail over.

**Figure 6**

Figure 7 on page 28 shows the fail over results from the failed database server to the N+1 cold standby server.
Figure 7

Figure 8 shows the fail back results from the N+1 cold standby server to the database server, that is operationally resumed after taking corrective actions, in a two node Oracle RAC database.

Figure 8

The measured RTO is less than 15 minutes for the functions as intended in the reference architecture.
Conclusion

Testing this reference architecture validated that the Hitachi Unified Compute Platform Select for Oracle Database with the Oracle Real Application Clusters option can achieve up to the following:

- 431K random Oracle reads
- 6.0 GB/sec sequential Oracle reads

The N+1 cold standby server testing demonstrated an RTO less than 15 minutes for the functions described in this reference architecture.

Using an N+1 cold standby server provides multiple high availability techniques for different business continuity goals, depending on the product requirements. This makes sure that you have a protected environment against unforeseen hardware failures.

The integrated and coordinated effort provides IT administrators with multiple options to match the desired RTO.

If necessary, you can use this reference architecture as a guide to design a larger infrastructure.
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.