Hitachi Unified Compute Platform Select for Oracle Database with the Oracle Real Application Clusters Option in a Medium Size Configuration Using Hitachi Virtual Storage Platform with Hitachi Accelerated Flash

Reference Architecture Guide

By Kishore Daggubati

July 22, 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.......................................................6
- Hitachi Virtual Storage Platform................................................7
- Hitachi Dynamic Provisioning....................................................7
- Hitachi Storage Navigator........................................................8
- Oracle Linux................................................................................8
- Oracle Database with the Real Application Cluster Option.........8

Solution Design...........................................................................9

- Storage Architecture...............................................................9
- Server and Application Architecture.........................................13
- SAN Architecture.....................................................................15
- Network Architecture..............................................................17

**Engineering Validation**........................................................20

- Test Methodology.................................................................20
- Test Results...............................................................................22

**Conclusion**.............................................................................24
Hitachi Unified Compute Platform Select for Oracle Database with the Oracle Real Application Clusters Option in a Medium Size Configuration Using Hitachi Virtual Storage Platform with Hitachi Accelerated Flash

Reference Architecture Guide

This reference architecture guide shows how the medium size configuration of Hitachi Unified Compute Platform Select for Oracle Database with the Oracle Real Application Clusters option provides a high performance, integrated solution for an Oracle infrastructure. The environment uses Hitachi Virtual Storage Platform with Hitachi Accelerated Flash. This is a resource to design an infrastructure for your requirements and budget.

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 Virtual Storage Platform using Hitachi Accelerated Flash
- 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 a medium size, high performance environment for Hitachi Unified Compute Platform Select for Oracle Database with the Real Application Clusters Option solution using Hitachi Virtual Storage Platform with Hitachi Accelerated Flash. It includes the performance and scalability benefits of this environment for typical on-line transaction processing workloads. Tailor your implementation of this solution to meet your specific needs.

This reference architecture guide includes discussion of the following:

- Hitachi Compute Blade 2000 using eight server blades
  - Two server blades connect using an SMP interface connector hosting one Oracle database server
- Hitachi Virtual Storage Platform 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 5, and Table 2 on page 6 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 A0195-C-6443</td>
<td>1</td>
</tr>
<tr>
<td>Server Blade</td>
<td>Model GVAX57A2 (X57-A2), each server blade configured as follows:</td>
<td>EFI BIOS Version 4.6.3.7</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>- Intel Xeon E7-8870 processor 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>- One dual port Fibre Channel card</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Mezzanine Slot 0 for server blades 0-3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Mezzanine Slot 1 for server blades 4-7</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Two 1 Gb/sec Ethernet NICs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Storage System</td>
<td>Hitachi Virtual Storage Platform configured as follows:</td>
<td>7005020000</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>- 1 Hitachi Accelerated Flash chassis with:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- 16 × 1.6 TB flash modules in 4 flash boxes, with 4 flash modules in each box, for about 25 TB raw capacity</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- 16 × 600 GB SAS 10k RPM drives</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- 8 front-end connectivity modules</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- 8 × 8 Gb/sec Fibre Channel ports for 64 ports total</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- 4 back-end connectivity modules</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- 8 × 6 Gb/sec SAS links each for 32 links total</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- 192 GB cache</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- 42 GB shared memory</td>
<td></td>
<td></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

### 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>■ 6 port, 8 Gb/sec Fibre Channel switch modules</td>
<td>V 642b</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>■ 1/10 GigE L3 LAN switch modules</td>
<td>Firmware Version 10.7</td>
<td></td>
</tr>
<tr>
<td>Symmetric Multiprocessing (SMP)</td>
<td>■ SMP connector creates one server from two GVAX57A2 server blades</td>
<td>N/A</td>
<td>4</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>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 Software</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>Database Client Communication Software</td>
<td>Oracle Net Services</td>
<td>11g R2, 11.2.0.1.0</td>
</tr>
</tbody>
</table>
Hitachi Virtual Storage Platform

Hitachi Virtual Storage Platform is the first 3-D scaling storage platform designed for all data types. Its storage architecture flexibly adapts for performance, capacity, and multi-vendor storage. Combined with the unique Hitachi Command Suite management platform, it transforms the data center.

- **Scale Up** — Meet increasing demands by dynamically adding processors, connectivity, and capacity in a single unit. Provide the highest performance for both open and mainframe environments.

- **Scale Out** — Meet multiple demands by dynamically combining multiple units into a single logical system with shared resources. Support increased demand in virtualized server environments. Ensure safe multi-tenancy and quality of service through partitioning of cache and ports.

- **Scale Deep** — Extend storage value by virtualizing new and existing external storage systems dynamically. Extend the advanced functions of Hitachi Virtual Storage Platform to multivendor storage. Offload less demanding data to external tiers to save costs and to optimize the availability of tier-one resources.

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 Storage Platform 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 Virtual Storage Platform.

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 a medium size environment for Hitachi Unified Compute Platform Select for Oracle Database with the Oracle Real Application Clusters option using Hitachi Virtual Storage Platform with Hitachi Accelerated Flash.

Specific infrastructure configuration includes the following:

- **Server** — This is a single server node, consisting of two X57-A2 server blades connected using an SMP connector for one logical Oracle database server.

- **Storage System** — There are LDEVs mapped to each port that are presented to the server as LUNs.

- **SAN Fabric** — There are four zones created on each switch module to zone the two mezzanine Fibre Channel ports on each server blade and the sixteen 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 Virtual Storage Platform shown in Figure 2 on page 10.
Table 3 has the details of the RAID groups for the Oracle on-line, archived redo logs, Oracle cluster registry, and voting disk.

<table>
<thead>
<tr>
<th>Parity Group</th>
<th>Purpose</th>
<th>RAID Level</th>
<th>Drive Type</th>
<th>No of Drives</th>
<th>Capacity (GB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>14-6</td>
<td>Oracle Online Redo Logs</td>
<td>RAID-10 (2D+2D)</td>
<td>600 GB 10k RPM SAS</td>
<td>4</td>
<td>1000</td>
</tr>
<tr>
<td>14-7</td>
<td>Oracle Online Redo Logs</td>
<td>RAID-10 (2D+2D)</td>
<td></td>
<td>4</td>
<td>1000</td>
</tr>
<tr>
<td>14-8</td>
<td>Oracle Archived Redo Logs, Oracle Cluster Registry and Voting Disk</td>
<td>RAID-10 (2D+2D)</td>
<td></td>
<td>4</td>
<td>1000</td>
</tr>
</tbody>
</table>
Table 4 has the details for the LDEVs created in these RAID groups.

### Table 4. LUNs Used for the Oracle Online, Archived Redo Logs, Oracle Cluster Registry and Voting Disk

<table>
<thead>
<tr>
<th>Parity Group</th>
<th>LDEVs</th>
<th>LDEV Size (GB)</th>
<th>Purpose</th>
<th>Storage Port</th>
</tr>
</thead>
</table>
| 14-6         | 00:01:33 | 1000           | ▪ Oracle Online Redo logs  
▪ Oracle Control File | 1A, 1B, 1C, 1D, 1E, 1F, 1F, 1H  
2A, 2B, 8C, 2D, 2E, 2F, 2G, 2H |
| 14-7         | 00:01:34 | 1000           | ▪ Oracle Online Redo Logs  
▪ Oracle Control File | 1A, 1B, 1C, 1D, 1E, 1F, 1F, 1H  
2A, 2B, 8C, 2D, 2E, 2F, 2G, 2H |
| 14-8         | 00:01:35 | 1000           | ▪ Oracle Archived Redo Logs  
▪ Oracle Cluster Registry  
▪ Voting Disk | 1A, 1B, 1C, 1D, 1E, 1F, 1F, 1H  
2A, 2B, 8C, 2D, 2E, 2F, 2G, 2H |

Table 5 has the details for the dynamic provisioning pool.

### Table 5. 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>
</table>
| ora_ssd_pool_01             | 17-1 – 17-2  
18-1 – 18-2              | RAID-10 (2D+2D) | 1.6 TB Flash Module Drive (FMD) | 16 | 11.95 TB |

Table 6 has the details for the virtual volumes created from the dynamic provisioning pool.

### Table 6. 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>
</table>
| ora_ssd_pool_01             | 00:01:41 – 00:01:7D | 200 | ▪ Oracle System  
▪ Sysaux  
▪ Undo  
▪ Temp  
▪ OLTP Application Tablespaces | 1A, 1B, 1C, 1D, 1E, 1F, 1F, 1H  
2A, 2B, 8C, 2D, 2E, 2F, 2G, 2H |
Database Layout

The database layout design uses recommended practices from Hitachi Data Systems for Hitachi Virtual Storage Platform using Hitachi Accelerated Flash for small random I/O traffic, such as the ones 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 one dynamic provisioning pool for the application data and indexes. The allocated capacity is 11.95 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 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 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 RAID groups for the database instance.

- **Archive Redo Logs** — Assign one RAID group for the database instance.

- **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 7 on page 13 lists the disk mappings from the LUNS to the operating system devices and to the ASM disk groups.
The reference architecture uses eight server blades in the Hitachi Compute Blade 2000 chassis. Each database server includes a single logical database instance running on the two server blades connected through an SMP connector.

There are 40 combined CPU cores with 256 GB RAM for each database server. 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 for hardware details.
This reference architecture does not use a third party volume manager. It uses Oracle RAC Database 11g R2 ASM volume manager software.

To monitor and manage the database, the database servers, 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

Figure 3 shows the server infrastructure for the reference architecture.
SAN Architecture

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

The environment used four Fibre Channel switch modules installed in the Hitachi Compute Blade 2000 chassis. This provides scalability and high availability. Key Solution Components has host configuration details.

Each of the database servers used four Fibre Channel ports, with two ports from the mezzanine card on each server blade. This provides a sixteen-path connection for all LUNs mapped to each of the database server in a four node Oracle RAC Database.

The environment used four Fibre Channel switch modules to provide redundant paths for all Hitachi Virtual Storage Platform using Hitachi Accelerated Flash LUNs.

Table 8 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-HBA1-1</td>
<td>BS2K_16_B0_HBA1_1_ASE45_33_1A</td>
<td>1A</td>
<td>Bay 2</td>
</tr>
<tr>
<td></td>
<td>B0-HBA1-2</td>
<td>BS2K_16_B0_HBA1_2_ASE45_33_2A</td>
<td>2A</td>
<td>Bay 3</td>
</tr>
<tr>
<td></td>
<td>B1-HBA1-1</td>
<td>BS2K_16_B1_HBA1_1_ASE45_33_1B</td>
<td>1B</td>
<td>Bay 2</td>
</tr>
<tr>
<td></td>
<td>B1-HBA1-2</td>
<td>BS2K_16_B1_HBA1_2_ASE45_33_2B</td>
<td>2B</td>
<td>Bay 3</td>
</tr>
<tr>
<td>Database server 2</td>
<td>B2-HBA1-1</td>
<td>BS2K_16_B2_HBA1_1_ASE45_33_1C</td>
<td>1C</td>
<td>Bay 2</td>
</tr>
<tr>
<td></td>
<td>B2-HBA1-2</td>
<td>BS2K_16_B2_HBA1_2_ASE45_33_2D</td>
<td>2D</td>
<td>Bay 3</td>
</tr>
<tr>
<td></td>
<td>B3-HBA1-1</td>
<td>BS2K_16_B3_HBA1_1_ASE45_33_1D</td>
<td>1D</td>
<td>Bay 2</td>
</tr>
<tr>
<td></td>
<td>B3-HBA1-2</td>
<td>BS2K_16_B3_HBA1_2_ASE45_33_8C</td>
<td>8C</td>
<td>Bay 3</td>
</tr>
<tr>
<td>Database server 3</td>
<td>B4-HBA1-1</td>
<td>BS2K_16_B4_HBA1_1_ASE45_33_1E</td>
<td>1E</td>
<td>Bay 4</td>
</tr>
<tr>
<td></td>
<td>B4-HBA1-2</td>
<td>BS2K_16_B4_HBA1_2_ASE45_33_2E</td>
<td>2E</td>
<td>Bay 5</td>
</tr>
<tr>
<td></td>
<td>B5-HBA1-1</td>
<td>BS2K_16_B5_HBA1_1_ASE45_33_1F</td>
<td>1F</td>
<td>Bay 4</td>
</tr>
<tr>
<td></td>
<td>B5-HBA1-2</td>
<td>BS2K_16_B5_HBA1_2_ASE45_33_2F</td>
<td>2F</td>
<td>Bay 5</td>
</tr>
<tr>
<td>Database server 4</td>
<td>B6-HBA1-1</td>
<td>BS2K_16_B6_HBA1_1_ASE45_33_1G</td>
<td>1G</td>
<td>Bay 4</td>
</tr>
<tr>
<td></td>
<td>B6-HBA1-2</td>
<td>BS2K_16_B6_HBA1_2_ASE45_33_2G</td>
<td>2G</td>
<td>Bay 5</td>
</tr>
<tr>
<td></td>
<td>B7-HBA1-1</td>
<td>BS2K_16_B7_HBA1_1_ASE45_33_1H</td>
<td>1H</td>
<td>Bay 4</td>
</tr>
<tr>
<td></td>
<td>B7-HBA1-2</td>
<td>BS2K_16_B7_HBA1_2_ASE45_33_2H</td>
<td>2H</td>
<td>Bay 5</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 Virtual Storage Platform 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 database server in this reference architecture has four on-board 1 Gb/sec NIC ports for different types of data traffic. The two NIC ports on each server blade connect to two internal 1/10 Gb/sec Ethernet switches in the chassis.

Observe these points when configuring private and public network 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 failover 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 4 on page 18 shows the network configuration for the reference architecture environment.
1 Gb/sec Private LAN Ethernet Switch

1/10 Gb/sec LAN Switch Module

1 Gb/sec Public LAN Ethernet Switch

1/10 Gb/sec LAN Switch Module

Database server 4

Database server 3

Database server 2

Database server 1

Onboard 1 Gbps NIC ports connect internally to the 1/10 Gb/sec LAN Switch Module. These connections are used for Public LAN

Onboard 1 Gbps NIC ports connect internally to the 1/10 Gb/sec LAN Switch Module. These connections are used for Private LAN

1 Gbps Ethernet cable for Public LAN

1 Gbps Ethernet cable for Private LAN

Figure 4
Table 9 provides the network details for the reference architecture.

<table>
<thead>
<tr>
<th>Server</th>
<th>NIC Ports</th>
<th>Switch Bay ID</th>
<th>Switch Ports (Internal)</th>
<th>VLAN</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>172.17.170.189</td>
<td>Public</td>
</tr>
<tr>
<td></td>
<td>B0-NIC-1</td>
<td>1</td>
<td>5</td>
<td>170</td>
<td>172.17.170.190</td>
<td>Public</td>
</tr>
<tr>
<td></td>
<td>B1-NIC-0</td>
<td>0</td>
<td>6</td>
<td>208</td>
<td>192.208.0.5</td>
<td>Private</td>
</tr>
<tr>
<td></td>
<td>B1-NIC-1</td>
<td>1</td>
<td>6</td>
<td>209</td>
<td>192.209.0.5</td>
<td>Private</td>
</tr>
<tr>
<td>Database Server 2</td>
<td>B2-NIC-0</td>
<td>0</td>
<td>7</td>
<td>170</td>
<td>172.17.170.191</td>
<td>Public</td>
</tr>
<tr>
<td></td>
<td>B2-NIC-1</td>
<td>1</td>
<td>7</td>
<td>170</td>
<td>172.17.170.192</td>
<td>Public</td>
</tr>
<tr>
<td></td>
<td>B3-NIC-0</td>
<td>0</td>
<td>8</td>
<td>208</td>
<td>192.208.0.6</td>
<td>Private</td>
</tr>
<tr>
<td></td>
<td>B3-NIC-1</td>
<td>1</td>
<td>8</td>
<td>209</td>
<td>192.209.0.6</td>
<td>Private</td>
</tr>
<tr>
<td>Database Server 3</td>
<td>B4-NIC-0</td>
<td>0</td>
<td>9</td>
<td>170</td>
<td>172.17.170.193</td>
<td>Public</td>
</tr>
<tr>
<td></td>
<td>B4-NIC-1</td>
<td>1</td>
<td>9</td>
<td>170</td>
<td>172.17.170.194</td>
<td>Public</td>
</tr>
<tr>
<td></td>
<td>B5-NIC-0</td>
<td>0</td>
<td>10</td>
<td>208</td>
<td>192.208.0.7</td>
<td>Private</td>
</tr>
<tr>
<td></td>
<td>B5-NIC-1</td>
<td>1</td>
<td>10</td>
<td>209</td>
<td>192.209.0.7</td>
<td>Private</td>
</tr>
<tr>
<td>Database Server 4</td>
<td>B6-NIC-0</td>
<td>0</td>
<td>11</td>
<td>170</td>
<td>172.17.170.195</td>
<td>Public</td>
</tr>
<tr>
<td></td>
<td>B6-NIC-1</td>
<td>1</td>
<td>11</td>
<td>170</td>
<td>172.17.170.196</td>
<td>Public</td>
</tr>
<tr>
<td></td>
<td>B7-NIC-0</td>
<td>0</td>
<td>12</td>
<td>208</td>
<td>192.208.0.8</td>
<td>Private</td>
</tr>
<tr>
<td></td>
<td>B7-NIC-1</td>
<td>1</td>
<td>12</td>
<td>209</td>
<td>192.209.0.8</td>
<td>Private</td>
</tr>
</tbody>
</table>
Engineering Validation

This is the functionality and performance validation of the medium size reference architecture for Hitachi Unified Compute Platform Select for Oracle Database with the Oracle Real Application Clusters option using Hitachi Virtual Storage Platform with Hitachi Accelerated Flash. Testing involved a configuration for a four node Oracle RAC Database.

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.

- **Swingbench** — This test software performs a stress test on an Oracle database. The Swingbench load factor is about 313, which fills 1 TB data order entry (OE) application data.

The database size for validating this configuration is 1 TB from a storage capacity of approximately 2 TB.

Based on Oracle best practice in Metalink Note 810484.1, these were the settings used:

- **Database block size** — 8 KB
- **ASM AU setting** — 1 MB

The ASM diskgroup configuration is in "Database Layout."

These were the simulated workloads for Oracle Orion testing:

- 100% read transactions
- 90% read transactions, 10% write transactions

These were the simulated workloads for Swingbench testing:

- 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
- **Scalability Test** — Use Swingbench to run synthetic OLTP read- and write intensive workload testing to identify transactions per second (TPS)

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
- **Database**
  - Oracle Automatic Workload Repository report for database performance
  - Swingbench for application-level statistics for the number of executed transactions
- **Standalone tests on the drives/LUNS**
  - Oracle Orion for the throughput and IOPS performance for a given set of workload
Test Results

This summarizes the key observations from the test results for the medium size reference architecture for Hitachi Unified Compute Platform Select for Oracle Database with the Oracle Real Application Clusters option using Hitachi Virtual Storage Platform with Hitachi Accelerated Flash.

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

Table 10 has the details of Orion IOPS for different workload distribution.

<table>
<thead>
<tr>
<th>Workload Distribution</th>
<th>IOPS</th>
</tr>
</thead>
<tbody>
<tr>
<td>100% read transactions</td>
<td>416,210</td>
</tr>
<tr>
<td>90% read transactions, 10% write transactions</td>
<td>301,581</td>
</tr>
</tbody>
</table>

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

Table 11 has the details of ORION throughput for different workload distribution.

<table>
<thead>
<tr>
<th>Workload Distribution</th>
<th>Throughput in MB/sec</th>
</tr>
</thead>
<tbody>
<tr>
<td>100% read transactions</td>
<td>5,795</td>
</tr>
<tr>
<td>90% read transactions, 10% write transactions</td>
<td>5,009</td>
</tr>
</tbody>
</table>
Scalability Test

Figure 5 illustrates the Swingbench throughput of the system when activating the number of database servers in a four node Oracle RAC Database for 90% read transactions.

![Swingbench throughput for 90% reads - Average TPS](image)

Figure 5
Conclusion

Testing this reference architecture validated that the medium size Hitachi Unified Compute Platform Select for Oracle Database with the Oracle Real Application Clusters option does the following:

Oracle Orion test results show this reference architecture can achieve up to the following:

- 416K random Oracle reads
- 5.7 GB/sec sequential Oracle reads

Swingbench test results show TPS increases with additional nodes in an Oracle RAC Database activated for OLTP workloads.

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.