Oracle Diamond Partner

HITACHI
Inspire the Next

WHITE PAPER

Hitachi Unified Compute Platform 6000 for Oracle Real Application Clusters using Hitachi Virtual Storage Platform G1500, Hitachi Accelerated Flash, and Hitachi Compute Blade 2500 with Intel Xeon 2699v4 Processors

Reference Architecture Guide

By Libo Jiang

August 2017

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.

Revision History

Revision	Changes	Date
AS-614-00	Initial release	August 31, 2017

Table of Contents

Hitachi Unified Compute Platform 6000 for Oracle Real Application Clusters using Hitachi Virtual Storage Platform G1500, Hitachi Accelerated Flash, and Hitachi Compute Blade 2500 with Intel Xeon 2699v4 Processors

Reference Architecture Guide

Use this reference architecture guide to see how Hitachi Unified Compute Platform (UCP) 6000 for Oracle Real Application Clusters provides a high performance, integrated converged solution for Oracle. The environment uses Hitachi Virtual Storage Platform G1500 (VSP G1500), Hitachi Compute Blade 2500 (CB 2500), with Intel Xeon 2699v4 processors (Broadwell) and Hitachi Accelerated Flash (HAF). With this reference architecture guide, design a converged solution for your Oracle environment to meet your requirements and budget.

This proven solution to optimize a converged solution for Oracle databases, integrating servers, storage systems, network, and storage software. The environment provides reliability, high availability, scalability, and performance while processing small-scale to large-scale on-line transaction processing (OLTP) and online analytical processing (OLAP) workloads. The two dedicated servers run Oracle Database 12c R1 with the Oracle Real Application Cluster (RAC) option. The operating system is Red Hat Enterprise Linux 7.3.

This reference architecture document is for you if you are in one of the following roles:

- Database administrator
- Storage administrator
- Database performance analyzer
- IT professional with the responsibility of planning and deploying an Oracle Database solution

To use this reference architecture guide, you need familiarity with the following:

- Hitachi Virtual Storage Platform G1500 using Hitachi Accelerated Flash
- Hitachi Compute Blade 2500
- Storage area networks
- Oracle RAC Database 12c Release 1
- Oracle Automatic Storage Management (Oracle ASM)
- Red Hat Enterprise Linux
- Hitachi Dynamic Link Manager (HDLM)

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 implements a Hitachi Unified Compute Platform 6000 for Oracle RAC system on two nodes using Hitachi Virtual Storage Platform G1500 with Hitachi Accelerated Flash. This environment addresses the high availability, performance, and scalability requirements for OLTP and OLAP workloads. Tailor your implementation of this solution to meet your specific needs.

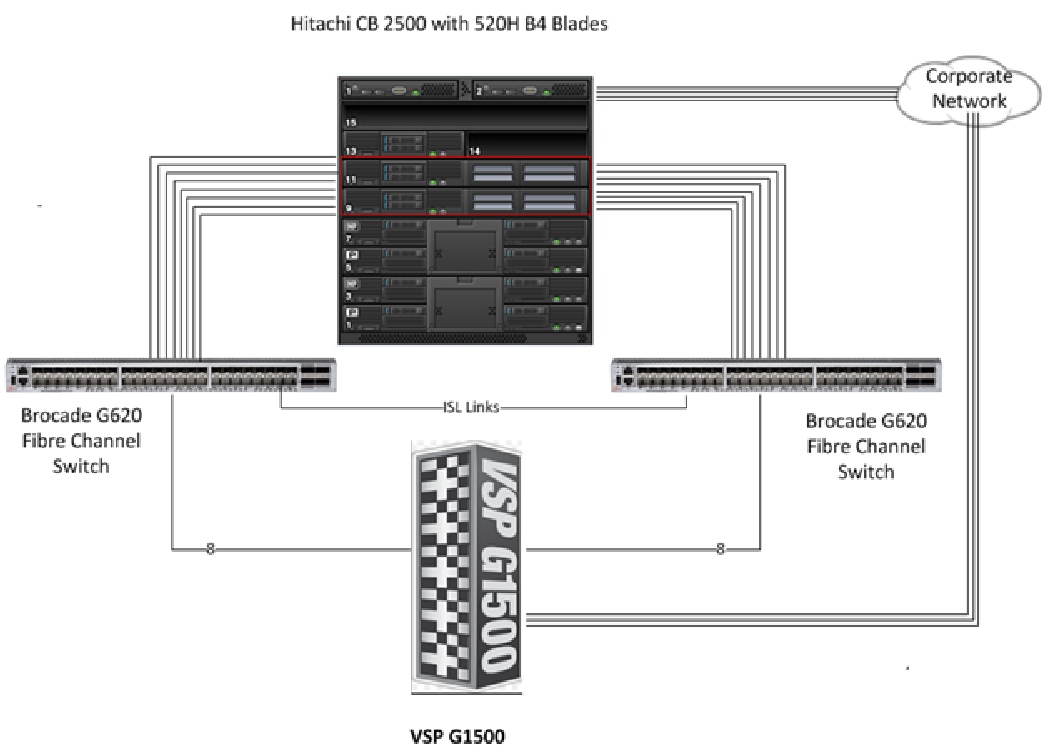
This reference architecture includes the following:

- **Hitachi Compute Blade 2500** with two 520H B4 server blades and two PCIe expansion blades
  - **Server Blade 9** — Oracle RAC NODE 1
  - **Server Blade 11** — Oracle RAC NODE 2
- **Hitachi Virtual Storage Platform G1500** with Hitachi Accelerated Flash
- 16 Gb/sec switched SAN infrastructures
- 10 GbE LAN infrastructure

This reference architecture uses a switched-SAN configuration. A direct-connect or switchless SAN configuration is also supported.

Figure 1 diagrams the high-level infrastructure for this UCP 6000 for Oracle RAC solution.

Figure 1



The configuration of Virtual Storage Platform G1500 and Compute Blade 2500 has the following:

- Fully redundant hardware
- Dual fabric connectivity between hosts and storage

To avoid any performance impact to the production database, Hitachi Data Systems recommends using a configuration with the following:

- A dedicated storage system for the production database
- A dedicated storage system for storing backup data, if needed

Key Solution Components

Table 1 lists the key hardware components used by this reference architecture.

TABLE 1. KEY HARDWARE COMPONENTS

Hardware	Detail Description	Version	Quantity
Hitachi Virtual Storage Platform G1500 (VSP G1500)	1 controller 2 virtual storage director pairs 2 back-end director pairs 2 front-end director pairs 16 × 16 Gb/s Fibre Channel ports 512GB mirrored cache memory	80-05-42-00/00	1
	7 TB high-density FMDs		32 + 2 spares
Hitachi Compute Blade 2500 (CB 2500)	10 Fan modules 2 Management modules 4 Power supply modules 2 × 10 GbE data center bridging (DCB) local area network (LAN) switch module	Management module firmware version A0185-B-1476 Dictionary version A0036 DCB switch version 4.0.1_hit1	1

TABLE 1. KEY HARDWARE COMPONENTS (CONTINUED)

Hardware	Detail Description	Version	Quantity
520H B4 Half-width Server Blade	2 Intel Xeon E5-2699 v4 CPU 512 RAM, 16 × 32 GB DDR4 memory 1 × 4-port 10 GbE converged network adapter (CNA) LAN on motherboard (LOM)	10-09	2
	2 × 300 GB, 15k RPM SAS Drives (RAID-1 for the Red Hat Enterprise Linux operating system and Oracle binaries)		2
	PCIe Expansion blade		2
	Hitachi 16 Gb/s 2-port PCIe Fibre Channel HBA	50-04-00	8
Brocade G620 Fibre Channel switch	32 GB-capable 64-port Fibre Channel switch 32 × 16 Gb/s SFPs	8.0.1	2

Table 2 lists the key software components used in this reference architecture.

TABLE 2. KEY SOFTWARE COMPONENTS

Software	Version	Function
Red Hat Enterprise Linux	7.3	Operating system for Oracle RAC
Oracle ASM	12c Release 1 (12.1.0.2.0)	Volume manager and file system
Oracle Database	12c Release 1 (12.1.0.2.0)	Database system with multitenant architecture
Hitachi Storage Virtualization Operating System	Hitachi Virtual Storage Platform G1500	Storage virtualization, storage service level controls, dynamic provisioning, and performance instrumentation
Hitachi Dynamic Link Manager	8.5.2	Multipath software

Hitachi Compute Blade 2500

Hitachi Compute Blade 2500 delivers enterprise computing power and performance with unprecedented scalability and configuration flexibility. Lower your costs and protect your investment.

Flexible I/O architecture and logical partitioning allow configurations to match application needs exactly with Hitachi Compute Blade 2500. Multiple applications easily and securely co-exist in the same chassis.

Add server management and system monitoring at no cost with Hitachi Compute Systems Manager. Seamlessly integrate with Hitachi Command Suite in Hitachi storage environments.

In this solution, Hitachi Compute Blade 2500 with 520H B4 server blades and PCIe expansion blades provides scalability and flexibility for an Oracle RAC configuration while providing redundancy.

Hitachi Compute Blade can be expanded to accommodate up to four additional PCIe cards by the addition of the PCIe expansion blade. The PCIe expansion blade occupies the slot, adjacent to its host compute blade. This effectively creates a combined double-wide blade.

Hitachi Virtual Storage Platform G1500

Hitachi Virtual Storage Platform G1500 (VSP G1500) is an enterprise platform. These enable continuous operations, self-managing policy-driven management, and agile IT that for cloud applications. Global storage virtualization enables an always-on infrastructure with enterprise-wide scalability. An ideal solution for applications that require zero recovery point and recovery time objectives, Virtual Storage Platform G1500 redefines mission-critical storage virtualization to reset expectations for the data center.

Virtual Storage Platform G1500 is equipped with virtual storage directors. These directors use the latest generation of Intel Xenon 2.3GHz 8-core microprocessors to efficiently manage the following:

- Front-end and back-end directors
- PCI-Express interface
- Local memory
- Communication between the service processors

Virtual Storage Platform G1500, used in this reference architecture, supports Oracle Real Application Clusters.

Hitachi Accelerated Flash

Hitachi Accelerated Flash features a flash module built specifically for enterprise-class workloads. Developed for Hitachi Virtual Storage Platform, Accelerated Flash is available for Hitachi Unified Storage VM and Hitachi Virtual Storage Platform family.

Accelerated Flash features innovative Hitachi-developed embedded flash memory controller technology. Hitachi flash acceleration software speeds I/O processing to increase flash device throughput.

Hitachi Accelerated Flash provides a reliable data storage for the Oracle database file placement with fast data retrieval for the OLTP workload.

Hitachi Storage Virtualization Operating System

Hitachi Storage Virtualization Operating System (SVOS) spans and integrates multiple platforms. It integrates storage system software to provide system element management and advanced storage system functions. Used across multiple platforms, Storage Virtualization Operating System includes storage virtualization, thin provisioning, storage service level controls, dynamic provisioning, and performance instrumentation.

Storage Virtualization Operating System includes standards-based management software on a Hitachi Command Suite (HCS) base. This provides storage configuration and control capabilities for you.

Hitachi Dynamic Link Manager Advanced

Hitachi Dynamic Link Manager Advanced combines all the capabilities of Hitachi Dynamic Link Manager and Hitachi Global Link Manager into a comprehensive multipathing solution. It includes capabilities such as the following:

- Path failover and failback
- Automatic load balancing to provide higher data availability and accessibility

This solution uses the **extended round-robin load balancing policy** configuration in Hitachi Dynamic Link Manager Advanced. This policy determines which path to allocate based on whether the data of the I/O to be issued is sequential with the immediately preceding data issued of the I/O.

If the data is sequential with the preceding data, the extended round-robin load balancing policy normally issues the same path for distributing the I/O that was issued to the immediately preceding I/O. However, after issuing a specified number of I/Os to a path, this policy then switches processing to the next path.

If the data is not sequential with the preceding data, the extended round-robin load balancing policy issues a path for distributing the I/O in order from among all the connected paths. Balancing the load across all available paths optimizes IOPS and response time.

Brocade Networking

Brocade and Hitachi Data Systems partner to deliver storage networking and data center solutions. These solutions reduce complexity and cost, as well as enable virtualization and cloud computing to increase business agility.

The solution using the following Brocade products:

- Brocade VDX 2746 10 GbE switch module
- Brocade 32 Gb/s capable switch with 16Gb/s SFPs

Red Hat Enterprise Linux

Red Hat Enterprise Linux delivers military-grade security, 99.999% uptime, support for business-critical workloads, and so much more. Ultimately, the platform helps you reallocate resources from maintaining the status quo to tackling new challenges.

Oracle Database

Oracle Database has a multitenant architecture so you can consolidate many databases quickly and manage them as a cloud service. Oracle Database also includes in-memory data processing capabilities for analytical performance. Additional database innovations deliver efficiency, performance, security, and availability. Oracle Database comes in two editions: Enterprise Edition and Standard Edition 2.

Oracle Real Application Clusters (Oracle RAC) is a clustered version of Oracle Database. It is based on a comprehensive high-availability stack that can be used as the foundation of a database cloud system, as well as a shared infrastructure. This ensures high availability, scalability, and agility for any application.

Oracle Automatic Storage Management (Oracle ASM) is a volume manager and a file system for Oracle database files. This supports single-instance Oracle Database and Oracle Real Application Clusters configurations. Oracle ASM is the recommended storage management solution that provides an alternative to conventional volume managers, file systems, and raw devices.

Solution Design

This describes the reference architecture environment to implement a quarter-rack environment for Hitachi Unified Compute Platform 6000 for Oracle RAC. The environment uses Hitachi Virtual Storage Platform G1500 with Hitachi Accelerated Flash.

The infrastructure configuration includes the following:

- **Server** — There are two servers configured in an Oracle Real Application Cluster.
- **Storage System** — 16 storage ports are connected to the Brocade G620 switches. LDEVs are mapped to all 16 storage ports and presented to the server as LUNs.
- **SAN Connection** — There are 16 SAN connections to connect the Fibre Channel HBA ports to the Brocade G620 switches.

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.

This reference architecture uses the a one-controller Hitachi Virtual Storage Platform G1500 with the following:

- Two pairs of virtual storage directors
- Two pairs of back-end directors
- Two pairs of front-end directors

The size of this Virtual Storage Platform G1500 configuration is about 25% of the maximum Virtual Storage Platform G1500 configuration. Use a larger configuration than the one in this guide if your implantation of this reference architecture requires higher storage performance.

Storage Layout

The storage layout design uses recommended practice from Hitachi Data Systems for Hitachi Virtual Storage Platform G1500 using Hitachi Accelerated Flash for better IOPS and throughput for OLTP operations.

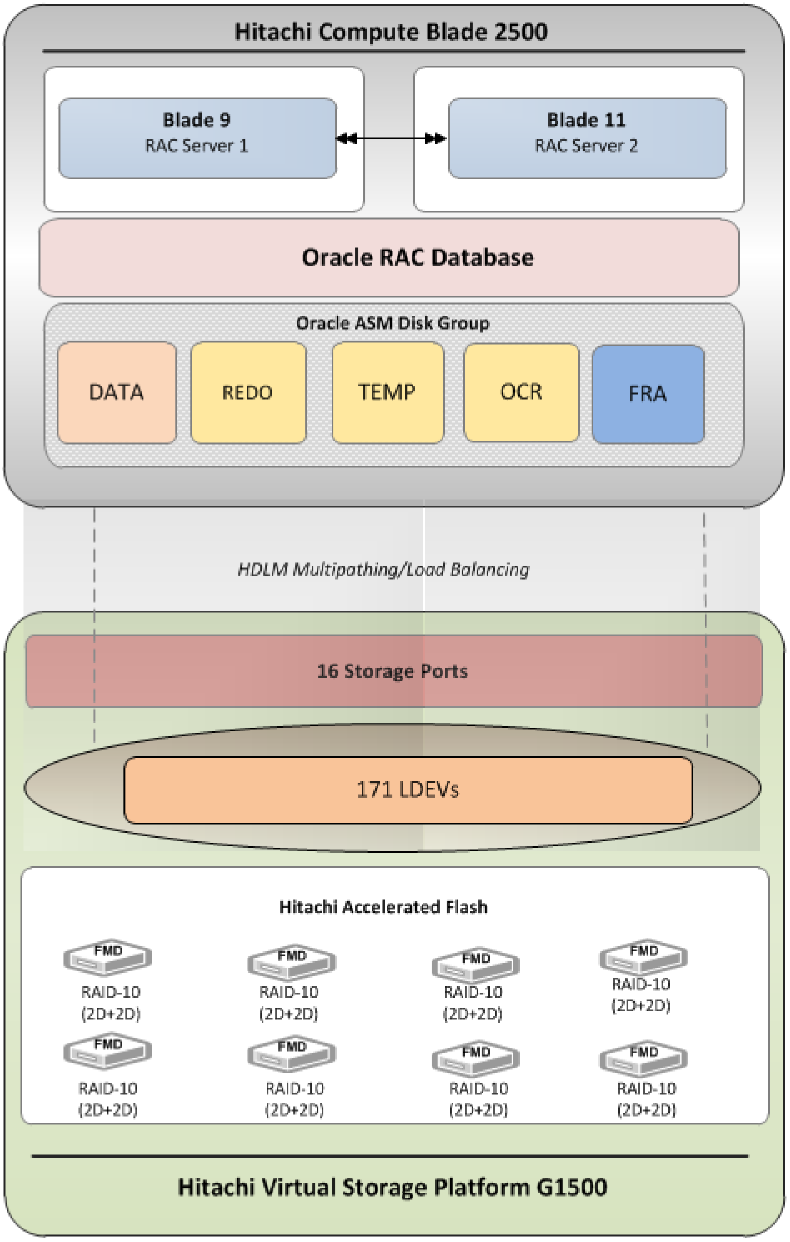
- **Data Disks** — This structure uses 7 TB high-density flash module drives
- **Disk Configuration** — Use RAID-10 (2D+2D) parity groups

Storage Configuration

This is the high-level storage configuration diagram of this solution.

Figure 2 on page 8 describes the storage configuration for this solution. This environment uses Hitachi Dynamic Link Manager for multipathing.

Figure 2



This reference architecture uses the extended least I/O load balancing algorithm. This is the default algorithm when first installing Hitachi Dynamic Link Manager. If you upgrade a Dynamic Link Manager installation, the algorithm used when first installing the system is inherited by the upgraded installation.

Select the load balancing algorithm most suitable for the data access patterns of your system environment. However, if there are no recognizable data access patterns, Hitachi Data Systems recommends using the default algorithm, extended least I/O.

Table 3 shows a sample storage configuration.

TABLE 3. STORAGE CONFIGURATION

Device Type	Non-Dynamic Provisioning
Number of RAID Groups	8
RAID Level	RAID-10 (2D+2D)
Drive Type	7 TB Flash Module Drive (FMD) high density
Number of Drives per Parity Group	4, 1 per tray
Capacity per Parity Group	12.8TB

Table 4 shows the logical storage configuration used in the tested configuration.

TABLE 4. LOGICAL STORAGE CONFIGURATION

Parity Group ID	1-1 -1-8				
Number of LDEVs	128 (16 per Parity Group)	16 (2 per Parity Group)	16 (2 per Parity Group)	8 (1 per Parity Group)	3 (one per Parity Group on first 3 Parity Groups)
LDEV Size	120 GB	10 GB	1200 GB	100 GB	5 GB
ASM Disk Group	DATA	REDO	FRA	TEMP	OCR
Purpose	OLTP Application Tablespaces System Sysaux Undo	Online Redo Logs Control Files	Incremental Backups Archived Redo Logs Control File Autobackups	Temp	OCR VOTING
Storage Port	5A, 5B, 5C, 5D, 6A, 6B, 6C, 6D,7A, 7B, 7C, 7D, 8A, 8B, 8C, 8D				

Database Layout

The database layout design uses recommended practices from Hitachi Data Systems for Hitachi Virtual Storage Platform G1500 using Hitachi Accelerated Flash for small random I/O traffic, such as OLTP transactions. The layout also considers 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 ASM diskgroup DATA for the data and index tablespaces. The small file table space consists of 1,224 data files that are each 8 GB. Set the tablespace to a small initial size with auto extend enabled to maximize storage utilization.
- **TEMP Tablespace** — Place TEMP tablespace in this configuration in ASM diskgroup TEMP. Many small file tempfiles are created within one single small TEMP tablespace. Limit the size of each small file tempfile to 31 GB.
- **Undo Tablespace** — Place UNDO tablespace in this configuration in ASM diskgroup DATA. Assign one BIGFILE UNDO tablespace for each database instance in a two node Oracle RAC database.
- **Online Redo Logs** — Assign ASM diskgroup REDO for online redo logs. Four redo logs are created for each database instance in a two node Oracle RAC database. Set the size of each redo log file to 8 GB.
- **Oracle Cluster Registry and Voting Disk** — Place each of these files in this configuration in ASM diskgroup OCR.
- **Size Settings** — Set the database block size to 8 KB. Set the ASM allocation unit to 1 MB.
- **ASM FILE SYSTEM I/O Settings** — Set the Oracle ASM I/O operations for database files as follows:
  - FILESYSTEMIO_OPTIONS = setall

Table 5 shows the Oracle RAC database configuration.

TABLE 5. ORACLE RAC DATABASE SETTINGS

For This	Use This
RAC configuration	Yes
ASM	Yes - Oracle RAC Database

Table 6 shows the Oracle environment parameters.

TABLE 6. ORACLE ENVIRONMENT PARAMETERS

For This	Use This
SGA_TARGET	128 GB
PGA_AGGREGATE_TARGET	64 GB
DB_CACHE_SIZE	64 GB
DB_KEEP_CACHE_SIZE	32 GB
DB_RECYCLE_CACHE_SIZE	8 GB
LOG_BUFFER	536870912
USE_LARGE_PAGES	TRUE

Table 7 has the details for the disk mappings from the LUNs to the operating system devices and to the ASM disk groups for Oracle RAC Database tablespaces.

TABLE 7. ORACLE ASM DISK CONFIGURATION

LDEV	LUN	LUN Size	Number of LUNs	ASM Disk	ASM Disk Group	Purpose
00:01:01 - 00:01:80	1-128	120 GB	128	DATA001 - DATA128	DATA	Application Data Index, SYS, SYSAUX
00:01:81 - 00:01:90	129-144	10 GB	16	REDO01 – REDO16	REDO	Online REDO Log Group Control File
00:01:91-00:01:A0	145-160	1200 GB	16	FRA01 – FRA16	FRA	Fast Recovery Area
00:01:A1 - 00:01:A8	161-168	100 GB	8	TEMP01 – TEMP08	TEMP	TEMP Files
00:01:A9 - 00:01:AB	169-171	5 GB	3	OCR01 – OCR03	OCR	Oracle Cluster Registry and Voting Disk

Server and Application Architecture

This reference architecture uses a single Hitachi Compute Blade 2500 with two server blades. Use two server blades and two PCIe expansion blades for a two-node Oracle RAC configuration. Install two dual-port Hitachi HBAs on each expansion blade.

This configuration provides the compute power for Oracle RAC database to handle complex database queries and a large volume of transaction processing in parallel.

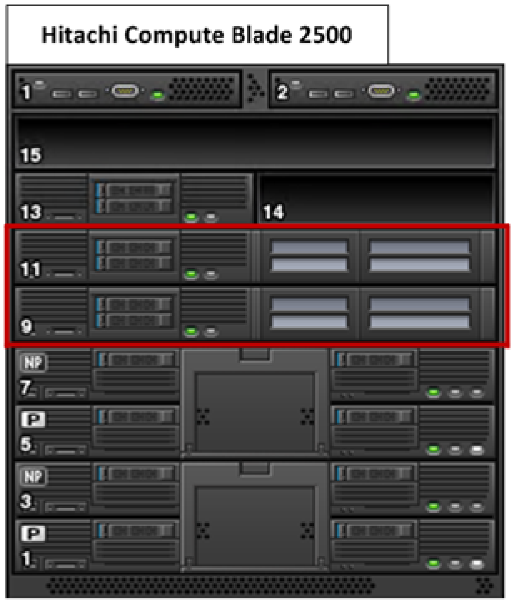
Table 8 describes the details of the server configuration for this solution.

TABLE 8. SERVER DETAILS

Server	Server Blade ID	Form Size	Role	Total Number of CPU Cores	RAM
Server 1	9	Half-width size with an PCIe expansion blade	RAC Node	44	512 GB
Server 2	11	Half-width size with an PCIe expansion blade	RAC Node	44	512 GB

Figure 3 shows the server infrastructure for the reference architecture. This diagram shows three blade configurations but only the half-width blades with PCIe expansion blades in the red rectangle were used for this environment.

Figure 3



SAN Architecture

Each of the database servers uses four HBAs, with two ports from each of the PCIe HBA cards. Zone each HBA port to a dedicated storage port on Hitachi Virtual Storage Platform G1500. Map all the LDEVs to 16 storage ports for multiple paths, high availability and high performance.

Table 9 shows connections from the HBAs to the Virtual Storage Platform G1500 storage ports.

TABLE 9. FIBRE CHANNEL CONNECTIONS TO HITACHI VIRTUAL STORAGE PLATFORM G1500

Server Blade	Host	HBA Ports	Storage Port	Storage Host Group	Zone Names
Server Blade 9	Pub106	HBA1-1	CL5-A	CB2500_80_B9_HBA1_1	CB2500_80_B9_HBA1_1_ASE47_60_5A
		HBA1-2	CL6-A	CB2500_80_B9_HBA1_2	CB2500_80_B9_HBA1_2_ASE47_60_6A
		HBA2-1	CL7-A	CB2500_80_B9_HBA2_1	CB2500_80_B9_HBA2_1_ASE47_60_7A
		HBA2-2	CL8-A	CB2500_80_B9_HBA2_2	CB2500_80_B9_HBA2_2_ASE47_60_8A
		HBA3-1	CL5-B	CB2500_80_B9_HBA3_1	CB2500_80_B9_HBA3_1_ASE47_60_5B
		HBA3-2	CL6-B	CB2500_80_B9_HBA3_2	CB2500_80_B9_HBA3_2_ASE47_60_6B
		HBA4-1	CL7-B	CB2500_80_B9_HBA4_1	CB2500_80_B9_HBA4_1_ASE47_60_7B
		HBA4-2	CL8-B	CB2500_80_B9_HBA4_2	CB2500_80_B9_HBA4_2_ASE47_60_8B

TABLE 9. FIBRE CHANNEL CONNECTIONS TO HITACHI VIRTUAL STORAGE PLATFORM G1500 (CONTINUED)

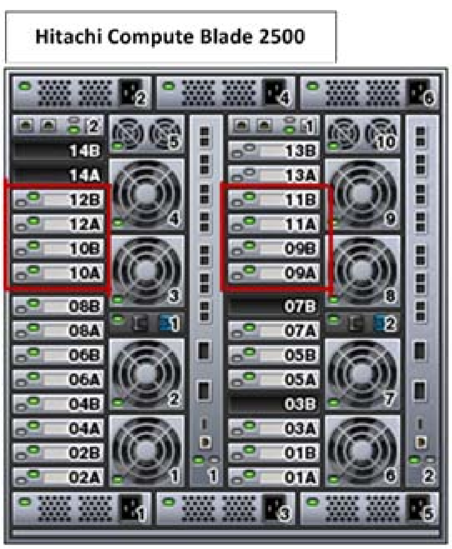
Server Blade	Host	HBA Ports	Storage Port	Storage Host Group	Zone Names
Server Blade 11	Pub107	HBA1-1	CL5-C	CB2500_80_B11_HBA1_1	CB2500_80_B11_HBA1_1_ASE47_60_5C
		HBA1-2	CL6-C	CB2500_80_B11_HBA1_2	CB2500_80_B11_HBA1_2_ASE47_60_6C
		HBA2-1	CL7-C	CB2500_80_B11_HBA2_1	CB2500_80_B11_HBA2_1_ ASE47_60_7C
		HBA2-2	CL8-C	CB2500_80_B11_HBA2_2	CB2500_80_B11_HBA2_2_ ASE47_60_8C
		HBA3-1	CL5-D	CB2500_80_B11_HBA3_1	CB2500_80_B11_HBA3_1_ ASE47_60_5D
		HBA3-2	CL6-D	CB2500_80_B11_HBA3_2	CB2500_80_B11_HBA3_2_ ASE47_60_6D
		HBA4-1	CL7-D	CB2500_80_B11_HBA4_1	CB2500_80_B11_HBA4_1_ ASE47_60_7D
		HBA4-2	CL8-D	CB2500_80_B11_HBA4_2	CB2500_80_B11_HBA4_2_ ASE47_60_8D

PCIe HBA Card Configuration

This describes the configuration for the 16 Gb/sec PCIe HBA cards from Hitachi that are used on the server blades.

Figure 4 shows the 16 Gb/s HBA PCIe cards that are installed in Hitachi Compute Blade 2500.

Figure 4



Set the HBA PCIe card parameters using Table 10.

TABLE 10. HBA PCIE CARD PARAMETERS

For This	Set This
Boot Function	Disable
Link Speed	16Gbps
Connection Type	Point-to-Pont
Multiple Port ID	Disable
Select Boot Device	Disable
Multipath Function	Disable

Network Architecture

This reference 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 RAC and inter-node communication between the nodes in the cluster.
- **Public Network** — This network provides client connections to the applications and Oracle RAC.

Hitachi Data Systems recommends using a pair of 10 Gb/s NICs for the cluster interconnect and public network.

Each server blade in this reference architecture has a quad port 10 Gb/s onboard NIC. The NIC ports have interconnected links to the two internal 10 Gb/s Ethernet switches in the chassis.

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

- Use NIC bonding to provide fail over and load balancing of interconnections within a server.
- Set all NICs to full duplex mode.
- Use two NICs for client connections to the application and database.
- Use two NICs for the cluster interconnection.

Figure 5 shows the IP network configuration for the reference architecture environment.

Figure 5

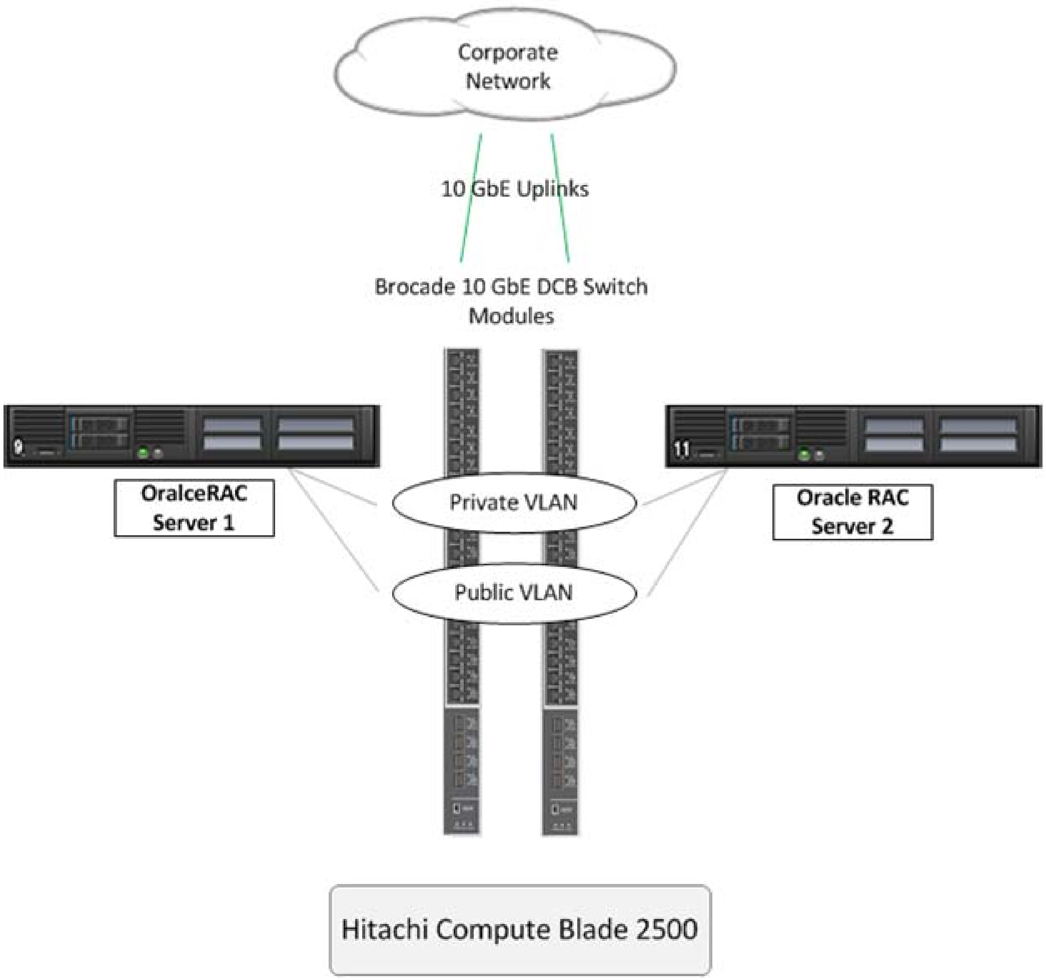


Table 11 shows the network configuration for this solution. Configure the VLAN to fit your network environment.

TABLE 11. NETWORK CONFIGURATION

Server	NIC Ports	NIC Bond	VLAN	IP Address	Network
Server 1	B9-NIC-0	Bond1	208	192.168.208.106	Private
	B9-NIC-1				
	B9-NIC-2	Bond2	167	172.17.167.106	Public
	B9-NIC-3				

TABLE 11. NETWORK CONFIGURATION (CONTINUED)

Server	NIC Ports	NIC Bond	VLAN	IP Address	Network
Server 2	B11-NIC-0	Bond1	208	192.168.208.107	Private
	B11-NIC-1				
	B11-NIC-2	Bond2	167	172.17.167.107	Public
	B11-NIC-3				

Engineering Validation

This describes how test methodology and test results.

Test Methodology

This summarizes Oracle performance testing tools and brief configuration used to validate this solution.

- Oracle Calibrate I/O

  The I/O calibration feature of Oracle Database enables you to assess the performance of the storage subsystem, and determine whether I/O performance problems are caused by the database or the storage subsystem. It issues I/Os randomly using Oracle data files to access the storage media, producing results that more closely match the actual performance of the database. Testing used 128 as the number of disks in the procedure.

- **Oracle ORION**

  Oracle ORION is a tool for predicting the performance of an Oracle database without having to install Oracle or create a database. Unlike other I/O calibration tools, Oracle ORION is expressly designed for simulating Oracle database I/O workloads using the same I/O software stack as Oracle. ORION can also simulate the effect of striping performed by Oracle Automatic Storage Management.

  For more information about ORION, see “I/O Configuration and Design” in the Oracle Database Performance Tuning Guide.

We used an automated ORION test tool adapted at Hitachi Data Systems for this solution architecture.

- **Peakmarks**

  Peakmarks is a commercial software product for various performance benchmark tests for Oracle database platform. Validation testing ran peakmarks storage benchmark tests with an 8 TB 2-node RAC database for this solution architecture.

- **SLOB**

  SLOB is an SGA-intensive Oracle I/O workload generation tool kit. It can be used to test I/O subsystem with genuine Oracle SGA-buffered physical I/O. Validation testing loaded a 9 TB two-instance RAC database with 384 users and tested database update rates between 5% and 50%.

Test Results

This is a summary of the key test results from evaluating this reference architecture with Oracle ORION, Oracle Calibrate I/O, peakmarks, and SLOB.

Figure 12 lists the Oracle ORION I/O test cases and results for this solution architecture. There were 128 × 120 GB LUNs used for the tests.

TABLE 12. ORACLE ORION TEST RESULTS

Test Case	Metric	Value
100% 8 KB Random Read	IOPS	1.95M
	Latency	0.91 ms
100% 8 KB Random Write	IOPS	776K
	Latency	0.79 ms
100% 1 MB Sequential Read	Throughput	25,128 MB/s

Table 13 lists Oracle Calibrate I/O test results for this solution.

TABLE 13. ORACLE CALIBRATE I/O TEST RESULTS

Test Case Name	IOPS	Average Latency (ms)	Throughput (MB/s)
Oracle Calibrate I/O	2,295,269	<0.50	23,686

Figure 6 shows execution and results of Oracle Calibrate I/O test.

Figure 6

```
Connected to:
Oracle Database 12c Enterprise Edition Release 12.1.0.2.0 - 64bit Production
With the Partitioning, Real Application Clusters, Automatic Storage Management, OLAP,
Advanced Analytics and Real Application Testing options

-----------------------------------------
welcome to  peakmarks  benchmark software
swiss precision in performance management
copyright (c) 2011-2017 www.peakmarks.com
-----------------------------------------

SYS@orcl1 SQL> @oracle_io_cal.sql
max_iops = 2295269
latency  = 0
max_mbps = 23686

PL/SQL procedure successfully completed.

SYS@orcl1 SQL>
```

Table 14 lists peakmarks storage test cases and results with a two-instance 8 TB Oracle RAC database. Three peakmarks test cases have been tested with this solution:

- Storage 8 KB random read (STO-RR)
- Storage 1,024 KB sequential read (STO-SR)
- Storage 32 KB random write foreground (STO-RWF)

TABLE 14. PEAKMARKS STORAGE TEST CASES AND RESULTS

Test Name	IOPS	Average Latency (ms)	Throughput (MB/s)
STO-RR	1,207,101	1.0	9,430
STO-SR	21,519	N/A	21,380
STO-RWF	579,971	0.42	18,129

Table 15 lists SLOB test cases and results with a 9 TB two-instance Oracle RAC database. Database update rates were in the range between 5% and 50%.

TABLE 15. SLOB DATABASE UPDATE RATES AND RESULTS WITH 384 USERS

DB Update Rate (%)	Total Physical IOPS	Total Physical Read IOPS	Total Physical Write IOPS	Wait Time (DB Sequential Read, ms)	DB File Parallel Write Latency (ms)	Log File Parallel Write Latency (ms)
5	2,032,604	1,928,803	103,801	1.15	0.34	0.58
10	1,818,581	1,643,708	174,873	1.04	0.28	0.46
20	1,489,230	1,219,742	269,488	1.00	0.24	0.45
25	1,386,302	1,090,780	295,529	0.96	0.26	0.48
30	1,297,581	975,376	322,205	0.96	0.24	0.40
40	1,179,759	818,577	361,182	0.95	0.16	0.41
50	1,102,075	720,990	381,085	0.93	0.15	0.37

Figure 7 shows IOPS results from SLOB tests with 384 users and different database update rates.

Figure 7

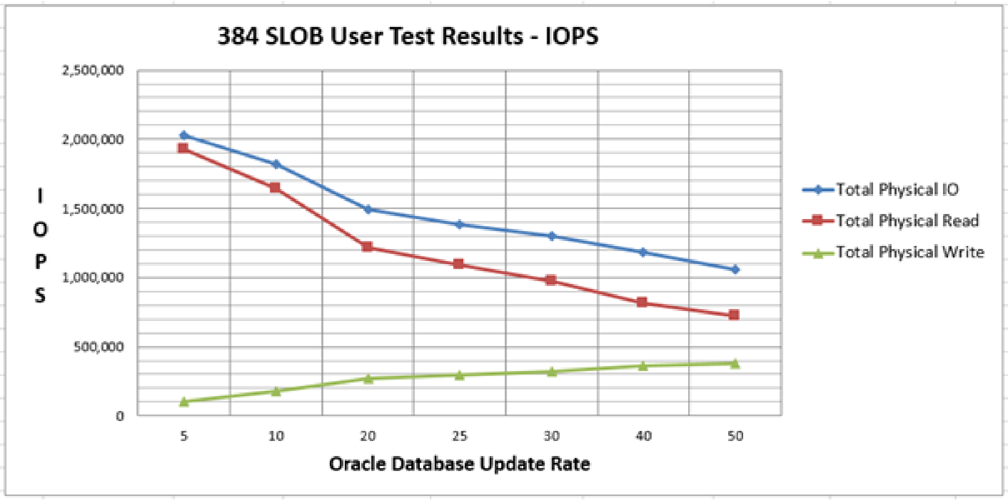


Figure 8 shows I/O latency results from SLOB tests with 384 users and different database update rates.

Figure 8

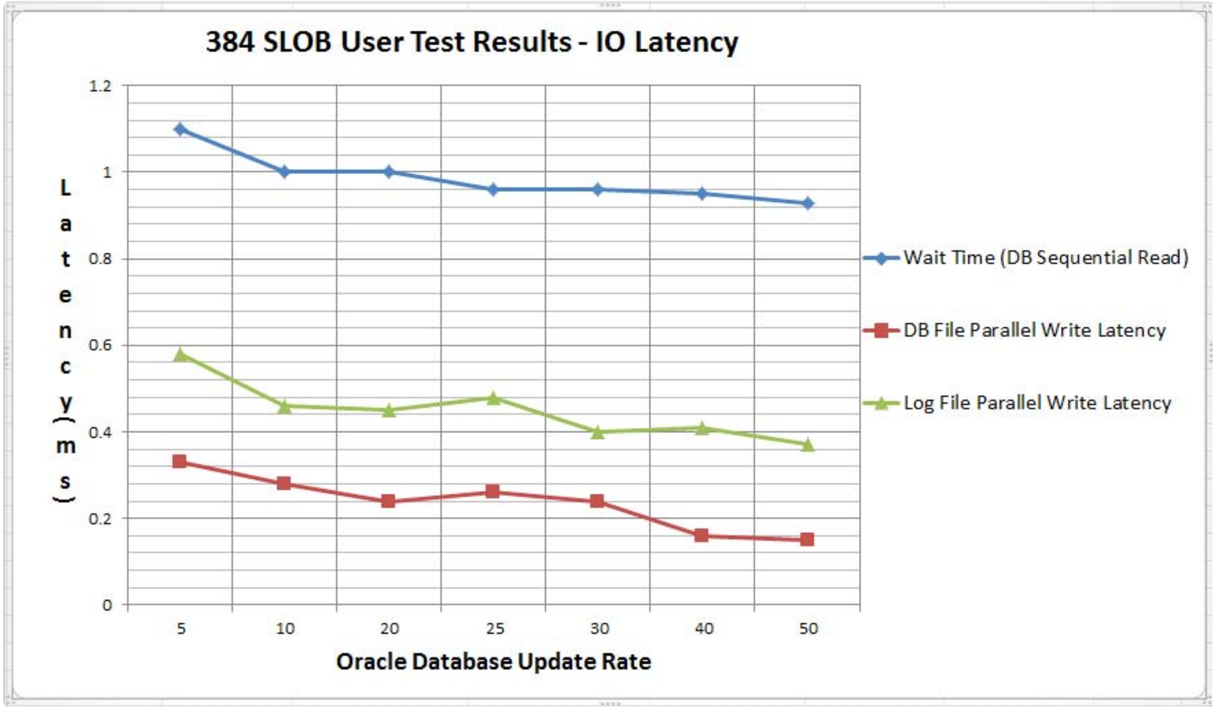


Figure 9 and Figure 10 show the top 10 timed events from the Oracle RAC AWR for SLOB test with 5% database update rate.

Figure 9

	Wait		Event		Wait Time			Summary Avg Wait Time (ms)				
I#	Class	Event	Waits	%Timeouts	Total(s)	Avg(ms)	%DB time	Avg	Min	Max	Std Dev	Cnt
*	User I/O	db file sequential read	234,817,561	0.00	269,301.92	1.15	46.19	1.15	1.13	1.16	0.02	2
	User I/O	db file parallel read	44,247,870	0.00	238,270.44	5.38	40.87	5.39	5.24	5.54	0.21	2
		DB CPU			153,896.86		26.39					2
	Other	latch: gc element	3,029,994	0.00	4,989.58	1.65	0.86	1.65	1.42	1.88	0.33	2
	System I/O	log file parallel write	2,610,853	0.00	1,511.36	0.58	0.26	0.59	0.55	0.62	0.05	2
	Other	latch: gcs resource hash	841,162	0.00	1,253.86	1.49	0.22	1.49	1.30	1.67	0.26	2
	Other	latch free	633,785	0.00	880.49	1.39	0.15	1.38	1.18	1.58	0.28	2
	Other	LGWR intra group sync	997,795	0.00	454.96	0.46	0.08	0.25	0.05	0.46	0.29	2
	Other	target log write size	527,680	0.23	402.60	0.76	0.07	0.76	0.76	0.76		2
	System I/O	db file parallel write	944,570	0.00	313.80	0.33	0.05	0.33	0.32	0.34	0.01	2

Figure 10

physical read IO requests	2,928,418,161	1,928,795.76	1,258.86	964,397.88	30,522.93	942,814.91	985,980.85
physical read bytes	23,989,645,860,864	15,800,724,015.10	10,312,627.18	7,900,362,007.55	250,064,149.99	7,723,539,951.36	8,077,184,063.74
physical read total IO requests	2,928,429,721	1,928,803.37	1,258.87	964,401.69	30,524.88	942,817.34	985,986.03
physical read total bytes	23,989,845,121,024	15,800,855,257.37	10,312,712.84	7,900,427,628.69	250,080,834.42	7,723,593,774.82	8,077,261,482.55
physical read total multi block requests	46	0.03	0.00	0.03		0.03	0.03
physical reads	2,928,423,567	1,928,799.32	1,258.87	964,399.66	30,525.42	942,814.93	985,984.39
physical reads cache	2,928,423,326	1,928,799.16	1,258.87	964,399.58	30,525.41	942,814.86	985,984.30
physical reads cache prefetch	2,649,347,327	1,744,986.41	1,138.90	872,493.20	27,614.66	852,966.69	892,019.71
physical reads direct	116	0.08	0.00	0.04	0.00	0.04	0.04
physical reads direct (lob)	8	0.01	0.00	0.01		0.01	0.01
physical write IO requests	150,129,275	98,882.30	64.54	49,441.15	1,678.91	48,253.98	50,628.32
physical write bytes	1,284,777,754,624	846,215,856.78	552,298.02	423,107,928.39	13,471,134.75	413,582,397.66	432,633,459.12
physical write total IO requests	157,597,454	103,801.19	67.75	51,900.60	1,800.17	50,627.69	53,173.51
physical write total bytes	1,417,894,141,952	933,892,653.28	609,521.86	466,946,326.64	14,863,320.05	456,436,372.24	477,456,281.04
physical write total multi block requests	181,098	119.28	0.08	59.64	2.60	57.80	61.48
physical writes	156,833,246	103,297.85	67.42	51,648.92	1,644.41	50,486.15	52,811.70
physical writes direct	8,461	5.57	0.00	2.79	2.48	1.04	4.54
physical writes direct (lob)	93	0.06	0.00	0.03	0.00	0.03	0.03
physical writes direct temporary tablespace	5,310	3.50	0.00	1.75	2.47	0.00	3.50
physical writes from cache	156,824,775	103,292.27	67.42	51,646.14	1,641.94	50,485.11	52,807.16
physical writes non checkpoint	156,078,620	102,800.82	67.09	51,400.41	1,627.35	50,249.70	52,551.12

Figure 11 shows the top 10 timed events from the Oracle RAC AWR for SLOB test with 25% database update rate.

Figure 11

	Wait		Event		Wait Time			Summary Avg Wait Time (ms)				
I#	Class	Event	Waits	%Timeouts	Total(s)	Avg(ms)	%DB time	Avg	Min	Max	Std Dev	Cnt
*	User I/O	db file sequential read	599,187,789	0.00	576,306.99	0.96	75.64	0.96	0.96	0.96	0.00	2
	User I/O	db file parallel read	26,081,245	0.00	139,513.33	5.35	18.31	5.35	5.31	5.38	0.05	2
		DB CPU			101,184.67		13.28					2
	System I/O	log file parallel write	5,587,585	0.00	2,664.12	0.48	0.35	0.48	0.47	0.48	0.01	2
	System I/O	db file parallel write	1,678,684	0.00	443.34	0.26	0.06	0.26	0.26	0.27	0.01	2
	Configuration	log file switch completion	4,068	0.00	92.71	22.79	0.01	22.78	21.81	23.76	1.38	2
	Other	latch: gc element	244,534	0.00	81.19	0.33	0.01	0.32	0.24	0.41	0.12	2
	Other	LGWR wait for redo copy	2,535,530	0.00	63.79	0.03	0.01	0.03	0.02	0.03	0.00	2
	Cluster	gc current grant 2-way	332,117	0.00	60.00	0.18	0.01	0.77	0.18	1.36	0.83	2
	Cluster	gc cr multi block request	32,384	0.00	24.92	0.77	0.00	0.73	0.69	0.77	0.06	2

Figure 12 shows the I/O statistics from the Oracle RAC AWR report for SLOB test with 25% database update rate.

Figure 12

Statistic	Total	per Second	per Trans	per Second			
				Average	Std Dev	Min	Max
physical read IO requests	2,187,584,836	1,090,772.93	251.79	545,386.47	2,280.89	543,773.63	546,999.30
physical read bytes	17,920,704,593,920	8,935,616,679.14	2,062,661.30	4,467,808,339.57	18,688,379.81	4,454,593,659.48	4,481,023,019.66
physical read total IO requests	2,187,598,468	1,090,779.73	251.79	545,389.87	2,280.89	543,777.03	547,002.70
physical read total bytes	17,920,945,825,792	8,935,736,962.10	2,062,689.07	4,467,868,481.05	18,688,618.66	4,454,653,632.06	4,481,083,330.04
physical read total multi block requests	12	0.01	0.00	0.01		0.01	0.01
physical reads	2,187,586,096	1,090,773.56	251.79	545,386.78	2,281.32	543,773.65	546,999.92
physical reads cache	2,187,585,643	1,090,773.34	251.79	545,386.67	2,281.31	543,773.54	546,999.80
physical reads cache prefetch	1,562,439,453	779,063.12	179.84	389,531.56	1,632.28	388,377.36	390,685.76
physical reads direct	447	0.22	0.00	0.11	0.01	0.11	0.12
physical reads direct (lob)	15	0.01	0.00	0.01		0.01	0.01
physical write IO requests	566,015,649	282,226.57	65.15	141,113.29	1,047.85	140,372.35	141,854.23
physical write bytes	4,836,954,423,296	2,411,800,871.91	556,730.27	1,205,900,435.95	5,936,055.25	1,201,703,011.03	1,210,097,860.88
physical write total IO requests	592,693,382	295,528.62	68.22	147,764.31	1,125.34	146,968.58	148,560.04
physical write total bytes	5,312,877,830,144	2,649,105,666.86	611,508.74	1,324,552,833.43	6,424,200.92	1,320,010,237.39	1,329,095,429.47
physical write total multi block requests	201,709	100.58	0.02	50.29	4.15	47.36	53.22
physical writes	590,448,712	294,409.37	67.96	147,204.69	724.57	146,692.34	147,717.04
physical writes direct	5,366	2.68	0.00	1.34	0.34	1.09	1.58
physical writes direct (lob)	6	0.00	0.00	0.00		0.00	0.00
physical writes direct temporary tablespace	979	0.49	0.00	0.24	0.34	0.00	0.49
physical writes from cache	590,443,274	294,406.66	67.96	147,203.33	724.24	146,691.21	147,715.45
physical writes non checkpoint	579,238,691	288,819.83	66.67	144,409.92	652.12	143,948.80	144,871.03

Figure 13 and Figure 14 show the top 10 timed events from the Oracle RAC AWR for SLOB test with 30% database update rate.

Figure 13

	Wait		Event		Wait Time			Summary Avg Wait Time (ms)				
I#	Class	Event	Waits	%Timeouts	Total(s)	Avg(ms)	%DB time	Avg	Min	Max	Std Dev	Cnt
*	User I/O	db file sequential read	477,967,425	0.00	458,191.35	0.96	79.06	0.96	0.95	0.97	0.02	2
	User I/O	db file parallel read	16,500,605	0.00	86,755.20	5.26	14.97	5.26	5.18	5.34	0.12	2
		DB CPU			71,563.83		12.35					2
	System I/O	log file parallel write	7,973,969	0.00	3,198.95	0.40	0.55	0.40	0.40	0.41	0.01	2
	Other	target log write size	2,661,187	0.21	1,432.79	0.54	0.25	0.54	0.52	0.56	0.02	2
	Other	latch: gc element	197,748	0.00	1,082.57	5.47	0.19	5.12	2.24	8.01	4.09	2
	System I/O	db file parallel write	3,103,675	0.00	747.28	0.24	0.13	0.24	0.24	0.25	0.01	2
	Other	LGWR intra group sync	3,394,161	0.00	525.75	0.15	0.09	0.16	0.15	0.16	0.01	2
	Other	LGWR intra group IO completion	987,309	0.00	202.43	0.21	0.03	0.21	0.19	0.22	0.02	2
	Other	latch: gcs resource hash	53,811	0.00	106.46	1.98	0.02	1.85	0.87	2.83	1.38	2

Figure 14

physical read IO requests	1,483,385,664	975,368.45	209.83	487,684.22	12,406.76	478,911.32	496,457.13
physical read bytes	12,151,896,850,432	7,990,219,306.18	1,718,957.03	3,995,109,653.09	101,635,530.33	3,923,242,480.38	4,066,976,825.79
physical read total IO requests	1,483,397,017	975,375.91	209.84	487,687.96	12,405.99	478,915.59	496,460.32
physical read total bytes	12,152,095,085,568	7,990,349,651.43	1,718,985.07	3,995,174,825.72	101,629,322.53	3,923,312,042.59	4,067,037,608.86
physical reads	1,483,385,854	975,368.57	209.83	487,684.29	12,406.68	478,911.44	496,457.13
physical reads cache	1,483,385,472	975,368.32	209.83	487,684.16	12,406.67	478,911.32	496,457.00
physical reads cache prefetch	988,865,138	650,207.08	139.88	325,103.54	8,265.37	319,259.04	330,948.04
physical reads direct	378	0.25	0.00	0.12	0.01	0.12	0.13
physical write IO requests	468,437,061	308,010.75	66.26	154,005.37	4,069.85	151,127.56	156,883.19
physical write bytes	3,998,099,120,128	2,628,864,377.34	565,554.55	1,314,432,188.67	31,644,603.68	1,292,056,074.82	1,336,808,302.52
physical write total IO requests	490,023,579	322,204.50	69.32	161,102.25	4,340.36	158,033.15	164,171.35
physical write total bytes	4,391,689,227,264	2,887,661,116.10	621,230.18	1,443,830,558.05	35,046,498.65	1,419,048,941.20	1,468,612,174.90
physical write total multi block requests	312,062	205.19	0.04	102.59	6.51	97.99	107.20
physical writes	488,049,230	320,906.31	69.04	160,453.15	3,862.88	157,721.69	163,184.62
physical writes direct	3,448	2.27	0.00	1.13	0.01	1.13	1.14
physical writes direct (lob)	42	0.03	0.00	0.01	0.00	0.01	0.01
physical writes from cache	488,045,772	320,904.04	69.04	160,452.02	3,862.86	157,720.56	163,183.47
physical writes non checkpoint	477,099,724	313,706.70	67.49	156,853.35	3,797.27	154,168.27	159,538.43

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 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 Resources website.

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 Training and Certification website.

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

Hitachi Data Systems

Corporate Headquarters
2845 Lafayette Street
Santa Clara, CA 95050-2639 USA
www.HDS.com community.HDS.com

Regional Contact Information
Americas: +1 866 374 5822 or info@hds.com
Europe, Middle East and Africa: +44 (0) 1753 618000 or info.emea@hds.com
Asia Pacific: +852 3189 7900 or hds.marketing.apac@hds.com

AS-614-00, August 2017