HITACHI
Inspire the Next

Hitachi Solution for Databases for Oracle Real Application Clusters with Hitachi Virtual Storage Platform 5500 and Hitachi Advanced Server DS220 Using Intel Cascade Lake CPUs

Reference Architecture Guide

By Libo Jiang

April 2020

Feedback

Hitachi Vantara welcomes your feedback. Please share your thoughts by sending an email message to SolutionLab@HitachiVantara.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
MK-SL-169-00	Initial release	October 9, 2019
MK-SL-169-01	Minor updates to Table 1, Table 15, and Table 18.	October 16, 2019
MK-SL-169-02	Updates for Hitachi Virtual Storage Platform 5500	February 19, 2020
MK-SL-169-03	Minor text and figure updates	April 23, 2020

Table of Contents

Hitachi Solution for Databases for Oracle Real Application Clusters with Hitachi Virtual Storage Platform 5500 and Hitachi Advanced Server DS220 Using Intel Cascade Lake CPUs

Reference Architecture Guide

Use this reference architecture guide to see how Hitachi Solution for Databases provides a high performance, integrated, converged solution for Oracle Real Application Cluster (RAC) databases using Hitachi Virtual Storage Platform 5500 (VSP 5500). This environment uses the following:

- Virtual Storage Platform 5500
- Hitachi Advanced Server DS220 with 2nd generation Intel Xeon scalable processors
- Hitachi Advanced Server DS120 with Intel Xeon scalable processors.

With these products, design your infrastructure for Oracle databases to meet your requirements and budget.

This solution uses the high-performing Virtual Storage Platform 5500 with NVMe storage to boost performance. Dedicated servers run Oracle Database 19c with the Oracle Real Application Clusters (RAC) option and Oracle Linux 7.6 for the operating system.

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 5500
- Hitachi Advanced Server DS220
- Hitachi Advanced Server DS120
- Storage area networks
- Oracle 19c RAC Database
- Oracle Automatic Storage Management (Oracle ASM)
- Hitachi Adapters for Oracle
- Oracle Linux

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 Hitachi Unified Compute Platform architecture for Oracle Database is engineered, pre- tested, and qualified to provide high performance and high reliability in demanding, dynamic Oracle environments. It implements Unified Compute Platform with Oracle Real Application Clusters on four nodes using Hitachi Virtual Storage Platform 5500.

This reference architecture addresses the high availability, performance, and scalability requirements for OLTP and OLAP workloads. This reference architecture has been developed using Hitachi Advanced Server DS220 with second generation Intel Xeon scalable processors and Virtual Storage Platform 5500.

Business Benefits

Here are some benefits of the new reference architecture:

- A high performance Oracle Database on Hitachi Virtual Storage Platform 5500
- If you currently have an existing Oracle RAC converged environment, this provides another route to upgrade your system besides Virtual Storage Platform G900 or VSP G1500.

High Level Infrastructure

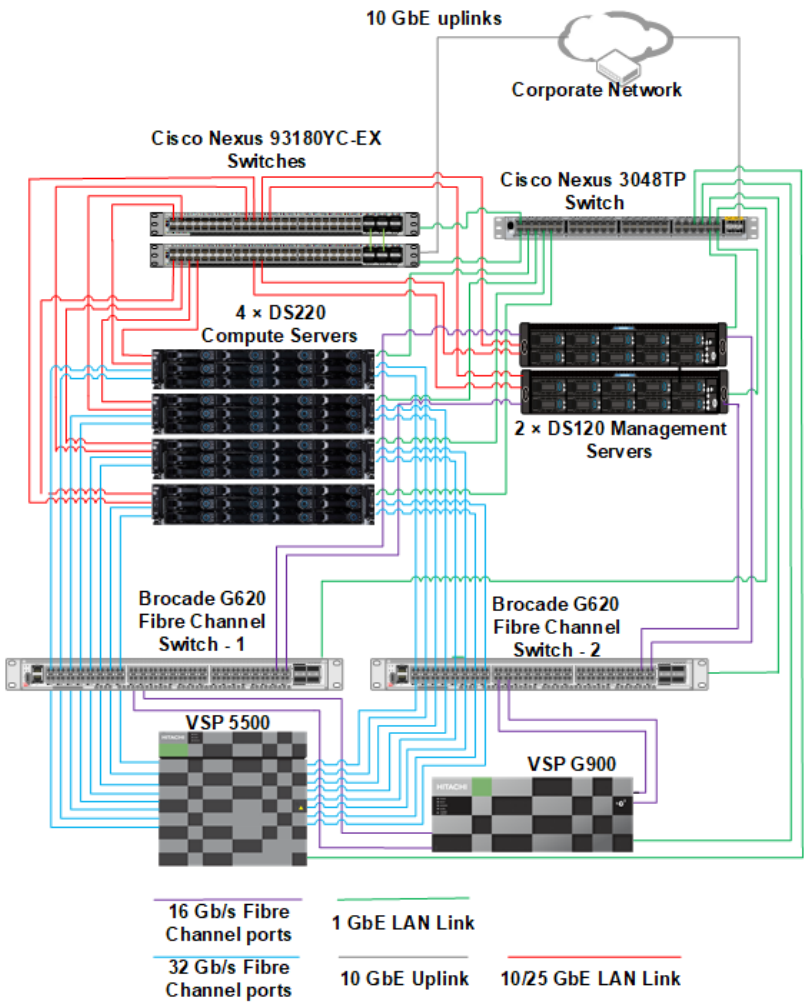
Hitachi Virtual Storage Platform 5500 and Hitachi Advanced DS220 are configured with fully redundant hardware for dual fabric connectivity between hosts and storage.

Figure 1 on page 3 shows the high-level architecture diagram using the following:

- Virtual Storage Platform 5500 and Hitachi Advanced Server DS220 with a 2-socket CPU for a four-node Oracle 19c RAC configuration
- Two Hitachi Advanced Server DS120 units with a 2-socket CPU for management of the architecture.

Configuration of data volumes for the management servers was on Virtual Storage Platform G900 in the lab environment. Virtual Storage Platform G900 is not required to implement this reference architecture. The management server environment can be configured on a Virtual Storage Platform system.

Figure 1



Hitachi Vantara recommends but does not require using a configuration with the following:

- A dedicated storage system for the production databases
- A dedicated storage system for data replication at a different site for business continuity, if needed

Uplink speed to the corporate network will depend on the customer environment and requirements. The Cisco Nexus 93180YC-EX switches can support uplink speeds of 40 GbE or 100 GbE, if higher bandwidth is required.

Key Solution Components

The key solution components for this solution are listed in Table 1, "Hardware Components," on page 4, Table 2, "Software Components for Compute Nodes," on page 6, and Table 3, "Software Components for Management Nodes," on page 6

TABLE 1. HARDWARE COMPONENTS

Hardware	Model	Detailed Description	Version	Quantity
Hitachi Virtual Storage Platform	VSP 5500	1 quad-controller block controllers 4 CHA pairs (16 × 32 Gb/s Fibre Channel ports in use) 1 drive chassis with 4 drive trays 2048 GB cache memory 66 × 1.9 TB NVMe SSDs (two spare drives are recommended for production environments)	90-01-61-00/00	1
	VSP G900	Two controllers 4 × 32 Gb/s Fibre Channel ports 16 × 12 Gb/s backend SAS ports 512 GB cache memory 4 × 6.0 TB 7.2K RPM SAS drives	88-02-03-60/00	1

TABLE 1. HARDWARE COMPONENTS (CONTINUED)

Hardware	Model	Detailed Description	Version	Quantity
Hitachi Advanced Server	DS220	2 Intel Xeon Gold 6254, 18 C CPUs @ 3.10 GHz 768 GB (64 GB × 12) DIMM DDR4 synchronous registered (buffered), 2666 MHz	BIOS: S5BH3B14.H01 BMC: 4.62.06 CPLD: 11	4
		2 Intel XXV710 dual port 25 GbE NIC cards	Driver: i40e Driver Version: 2.7.6-k Firmware: 6.02	
		2 Emulex LightPulse LPe32002-M2 2-Port 32 Gb/s Fibre Channel adapters	Driver: lpfc Driver Version: 12.0.0.10 Firmware: 12.0.261.9	
		2 × 1.5TB Intel Optane SSD DC P4800X	Firmware: E2010435	
	DS120	2 Intel Xeon Silver 4110 CPUs @ 2.10 GHz 256 GB (32 GB × 8) DIMM DDR4 synchronous registered (buffered) 2666 MHz 1 × 64 GB MLC SATADOM for boot	BIOS: 3A10.H8 BMC: 4.23.06 CPLD:10	2
		1 Dual port 25 GbE NIC Intel XXV710 PCIe card	Driver: i40e Driver Version: 1.5.6 Firmware: 6.0.1	
		1 Emulex LightPulse LPe31002-M6 2-port 16 Gb/s Fibre Channel adapters	Firmware: 11.4.204.27 Driver: lpfc Driver version: 11.4.341.0	
Brocade	G620 Fibre Channel Switch	48 port Fibre Channel switch 32 Gb/s SFPs	Kernel: 2.6.34.6 Fabric OS: v8.2.0b	4

TABLE 1. HARDWARE COMPONENTS (CONTINUED)

Hardware	Model	Detailed Description	Version	Quantity
Cisco	Nexus 93180YC-EX	48 × 10/25 GbE fiber ports 6 × 40/100 Gb/s quad SFP (QSFP28) ports	BIOS: version 07.61 NXOS: version 7.0(3)I4(7)	2
	Nexus 3048TP	1 GbE 48-Port Ethernet switch	BIOS: version 4.0.0 NXOS: version 7.0(3)I4(7)	1

Certain components may be optional, depending on the existing infrastructure and required interconnect topology. This may include the SAN, IP switches, and the management servers. However, this reference architecture documents the environment tested in the lab to support a full deployment of the architecture, including supporting components.

TABLE 2. SOFTWARE COMPONENTS FOR COMPUTE NODES

Software	Version	Function
Oracle Linux	7.6 with UEK kernel update 4.14.35-1933.el7uek.x86_64	Operating system
Oracle Database	19c	Database software
Oracle Real Application Cluster	19c	Cluster software
Oracle Grid Infrastructure	19c	Volume management, file system software, and Oracle Automatic Storage Management (ASM)
Oracle ASMlib	2.0.12	Oracle software for ASM

TABLE 3. SOFTWARE COMPONENTS FOR MANAGEMENT NODES

Software	Version	Function
VMware ESXi	6.7.0 Build 10302608	Hypervisor for management server
Vmware vCenter server	6.7.0 Build 10302608	VMware cluster management server
Hitachi Device Manager – Storage Navigator	Microcode dependent	Storage management software
Hitachi Ops Center	10.0.0	Hitachi infrastructure management software

Hitachi Virtual Storage Platform 5000 Series

This enterprise-class, flash array evolution, the Hitachi Virtual Storage Platform 5000 series (VSP) has an innovative, scale-out design optimized for NVMe (non-volatile memory express). It achieves the following:

- **Agility using NVMe** – Speed, massive scaling with no performance slowdowns, intelligent tiering, and efficiency.
- **Resilience** – Superior application availability and flash resilience. Your data is always available, mitigating business risk.
- **Storage simplified** – Do more with less, integrate AI and ML (machine learning), simplify management, and save money and time with consolidation.

Hitachi Virtual Storage Platform 5500 is used in this reference architecture guide. Other models of the Hitachi Virtual Storage Platform 5000 series also supports Oracle Real Application Clusters.

Hitachi Storage Virtualization Operating System RF

Hitachi Storage Virtualization Operating System RF (SVOS RF) 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.

Hitachi Ops Center

Manage, optimize, orchestrate and protect your data with advanced IT analytics and automation using Hitachi Ops Center. Achieve new insights, accelerate resource delivery, eliminate risks, and speed innovation to modernize your data center operations.

Use the power of AI operations with the following:

- Analyzer. Improve IT operations with machine learning (ML) to drive resource service levels, utilization and automation at lower costs. Obtain operational visibility from virtual machines, servers, SAN switches to shared storage resources to optimize an application's full data path.
- Automator. Deliver resources up to 70% faster than manual processes. Free staff to focus on strategic initiatives.

Hitachi Advanced Server DS220

With a combination of two Intel Xeon Scalable processors and high storage capacity in a 2U rack-space package, Hitachi Advanced Server DS220 delivers the storage and I/O to meet the needs of converged solutions and high-performance applications in the data center.

The Intel Xeon Scalable processor family is optimized to address the growing demands on today's IT infrastructure. The server provides 24 slots for high-speed DDR4 memory, allowing up to 3 TB of memory per node when 128 GB DIMMs are used. This server supports up to 12 large form factor storage devices and an additional 2 small form factor storage devices.

Intel has launched the second-generation of Intel Xeon Scalable processors, revolutionary Intel Optane DC persistent memory, and new networking and storage options. Learn more about Second Generation Intel Xeon Scalability Processors.

Hitachi Advanced Server DS120

Optimized for performance, high density, and power efficiency in a dual-processor server, Hitachi Advanced Server DS120 delivers a balance of compute and storage capacity. This rack mounted server has the flexibility to power a wide range of solutions and applications.

The highly scalable memory supports up to 3 TB RAM using 24 slots of 2666 MHz DDR4 RDIMM. DS120 is powered by the Intel Xeon scalable processor family for complex and demanding workloads. There are flexible OCP and PCIe I/O expansion card options available.

These applications were installed in individual virtual machines when testing this architecture. You would install them in most cases:

- VMware vCenter
- Hitachi Ops Center

Other management applications may be installed on additional virtual machines depending on your needs and requirements.

Oracle Linux

Oracle Linux is a Linux distribution packaged and freely distributed by Oracle, available partially under the GNU General Public License since late 2006. It is compiled from Red Hat Enterprise Linux source code, replacing Red Hat branding with Oracle branding.

Oracle Database

Oracle Database has a multi-tenant 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.

Database Smart Flash Cache is a transparent extension of the database buffer cache using solid state device (SSD) technology. The SSD acts as a Level 2 cache to the (Level 1) Oracle system global area (SGA). SSD storage is faster than disk storage, and cheaper than RAM. Database Smart Flash Cache with SSD storage gives you the ability to greatly improve the performance of your Oracle databases by reducing the amount of disk I/O at a lower cost than adding an equivalent amount of RAM. Database Smart Flash Cache is supported on Oracle Solaris and Oracle Linux only.

Validation of this architecture included tests with and without Database Smart Flash Cache.

VMware vSphere

VMware vSphere is a virtualization platform that provides a datacenter infrastructure. It helps you get the best performance, availability, and efficiency from your infrastructure and applications. Virtualize applications with confidence using consistent management.

VMware vSphere components used in this solution are the following:

- VMware vSphere ESXi

  Hypervisor loads directly on a physical server. ESXI provides a robust, high-performance virtualization layer that abstracts server hardware resources and makes them shareable by multiple virtual machines.

- VMware vCenter Server

  This provides a centralized platform for managing your VMware vSphere environments so you can automate and deliver a virtual infrastructure with confidence:

  - VMware vSphere vMotion
  - VMware vSphere Storage vMotion
  - VMware vSphere Distributed Resource Scheduler
  - VMware vSphere High Availability
  - VMware vSphere Fault Tolerance

  This reference architecture uses VMware vCenter Server for the management server cluster only.

Brocade Fibre Channel Switches from Broadcom

Brocade Fibre Channel switches deliver industry-leading performance, simplifying scale-out network architectures. Get the high-performance, availability, and ease of management you need for a solid foundation to grow the storage network you want.

The solution uses Brocade G620 Fibre Channel switches.

SAN switches are optional and direct connect is possible under certain circumstances. Check the support matrix to ensure support prior to implementation.

Cisco Switches

Cisco Nexus data center switches are built for scale, industry-leading automation, programmability, and real-time visibility.

This solution uses the following Cisco switches to provide Ethernet connectivity:

- Cisco Nexus 93180YC-EX is a 1U-sized top-of-rack or leaf switch with 48 ports 10/25 GbE (downlink) and 6 ports 40/100 GbE (uplink) for single-rack and multiple-rack solutions.
- Cisco Nexus 3048TP is a 1U-sized management switch with 48 ports 1 GbE for single-rack and multiple-rack solutions.

Solution Design

This describes the reference architecture environments to implement Hitachi Unified Compute Platform for Oracle with the Real Application Clusters option. The environment uses one Hitachi Virtual Storage Platform 5500.

The infrastructure configuration includes the following:

- **Oracle RAC Servers** – There are four server nodes configured in an Oracle Real Application Cluster.
- **Management Servers** – Two servers are used in the management server cluster.
- **Storage System** – There are virtual volumes mapped to each port that are presented to the server as LUNs.
- **SAN Connectivity** – There are SAN connections to connect the Fibre Channel HBA ports to the storage through Brocade G620 switches.

Storage Architecture

This describes the storage architecture for this solution.

Storage Configuration

The design and deployment of the database storage configuration, follows best practices from Hitachi Vantara and Oracle.

The high-level storage configuration diagram for this solution is shown in Figure 2.

Figure 2

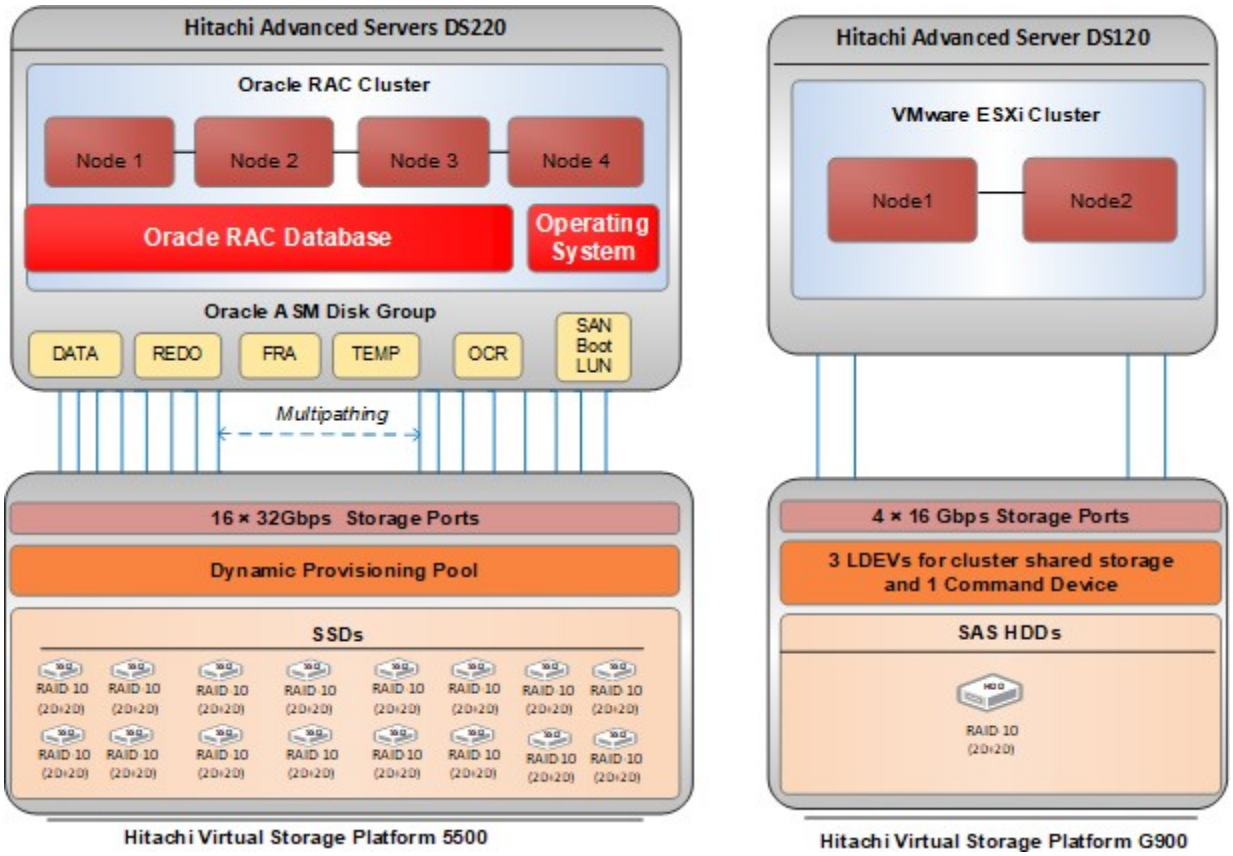


Table 4 shows the storage pool configuration used for this solution and lab verification.

TABLE 4. STORAGE POOL CONFIGURATION

Dynamic Provisioning Pool ID	Oracle-pool
RAID group	1-1 to 1-8, 2-1 to 2-8
RAID level	RAID-10 (2D+2D)
Drive type	1.9 TB SSD
Number of drives	64
Number of spare drives	2
Number of pool volume LDEVs	64
Pool volume LDEV size(s)	880.0 GB
Pool capacity	54.99 TB

Table 5 shows the logical storage configuration used in this solution.

TABLE 5. LOGICAL STORAGE CONFIGURATION

Dynamic Provisioning Pool ID	Oracle-pool					
Total number of DPVols	128	3	16	8	16	4
DPVols sizes (GB)	200	15	20	240	2000	380
ASM disk group	DATA	OCR	REDO	TEMP	FRA	N/A
Purpose	OLTP application tablespaces System Sysaux Undo	Oracle cluster Registry	Online redo Logs Control files	Temp	Incremental backups Archived redo Logs Control file Auto backups	SAN boot OS volumes
Storage port	1A, 1B, 1C, 1D, 2A, 2B, 2C, 2D, 3A, 3B, 3C, 3D, 4A, 4B, 4C, 4D					

Table 6 shows the Hitachi Virtual Storage Platform G900 configuration for management servers.

TABLE 6. HITACHI VIRTUAL STORAGE PLATFORM G900 CONFIGURATION FOR MANAGEMENT SERVERS

Item	Value or Description
Purpose	VMware Datastores Command control interface device
RAID level	RAID-10 (2D+2D)
Drive type	6.0 TB 7.2K RPM SAS
Number of drives	4
Number of spare drives	0
Number of LDEVs	3
LDEV size(s)	3000 GB
Number and size of CCI devices	1 × 100 MB
Storage port for management servers	7A, 7B, 8A, 8B

An additional RAID group consisting of four 3 TB, 7.2k RPM SAS drives configured as RAID-10 (2D+2D) was used as shared storage for the management server cluster. A 3 TB LUN and a command device were mapped to four storage ports.

Additional LUNs can be mapped if required. While the test environment was configured using a dedicated SAS RAID group for the management server cluster, this can be configured as a dedicated SSD RAID group, a dedicated dynamic provisioning pool, or it can use capacity on the dynamic provisioning pool configured for the Oracle environment, depending on your requirements.

Database Layout

The database layout design follows recommended best practices from Hitachi Vantara for Hitachi Virtual Storage Platform 5500 for small random I/O traffic, such as OLTP transactions. The layout also follows Oracle ASM best practices when using Hitachi storage.

The design can vary greatly from one implementation to another, based on the RAID configuration and number of drives used during the implementation. 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.

For Database Smart Flash Cache, install two Intel Optane SSD DC P4800X series drives locally on each Oracle RAC node. Create one partition from each drive. Configure that partition using ASMlib for a local ASM disk group, such as "FLASH1," "FLASH2," "FLASH3," and "FLASH4" for database server 1 through server 4, respectively. Mount each ASM disk group to the ASM instance running on the local Oracle server.

Oracle Configuration

Configure Oracle for this solution doing the following:

- **Data and Indexes Tablespace** – Assign an ASM diskgroup with external redundancy for the data and index tablespaces.
- **TEMP Tablespace** – Place the TEMP tablespace in this configuration in the TEMP ASM diskgroup.
- **Undo Tablespace** – Create an UNDO tablespace in this configuration within the Oracle Data ASM diskgroup. Assign one UNDO tablespace for each node in the Oracle RAC environment.
- **Online Redo Logs** – Create an ASM diskgroup with external redundancy for Oracle online redo logs.
- **Oracle Cluster Registry and Voting Disk** – Create an ASM diskgroup with normal redundancy to contain the OCR and voting disks and to protect against single disk failure to avoid loss of cluster availability. Place each of these files in this configuration in the OCR ASM diskgroups.

Oracle Initial Parameters

Table 7 shows Oracle ASM and database parameters.

TABLE 7. ORACLE ASM AND DATABASE PARAMETERS

Category	Item	Value
Oracle RAC option	RAC configuration	Yes
	ASM	Yes – to support Oracle RAC database
Oracle ASM environment parameters	OCR	3 × 15 GB
	DATA	128 × 200 GB
	REDO	16 × 20 GB
	TEMP	8 × 240 GB
	FRA	16 × 2,000 GB
Oracle Database environment parameters	SGA_TARGET	512 GB
	PGA_AGGREGATE_TARGET	256 GB
	DB_CACHE_SIZE	256 GB
	DB_KEEP_CACHE_SIZE	64 GB
	DB_RECYCLE_CACHE_SIZE	64 GB
	LOG_BUFFER	512 MB
	USE_LARGE_PAGES	TRUE
	FILESYSTEMIO_OPTIONS	SETALL
	DISK_ASYNCH_IO	TRUE

Oracle ASM Disk Mappings

Table 8 shows volumes, LUNs, and ASM diskgroups.

TABLE 8. VOLUMES, LUNS, AND ASM DISKGROUPS

ASM Disk Group	ASM Disk	DM-Multipath LUNs	LUN Details	Purpose
OCR	OCR1 - OCR3	/dev/mapper/mpathaa - /dev/mapper/mpathac	3 × 15 GB	Oracle cluster registry and voting disk
REDO	REDO01 - REDO16	/dev/mapper/mpathca - /dev/mapper/mpathcc /dev/mapper/mpathbn - /dev/mapper/mpathbz	16 × 20 GB	Online REDO log group
FRA	FRA01 - FRA16	/dev/mapper/mpathcl - /dev/mapper/mpathco	16 × 2000 GB	Flash recovery area
TEMP	TEMP1 – TEMP8	/dev/mapper/mpathcd - /dev/mapper/mpathck	8 × 240 GB	Temp
DATA	DATA001 – DATA128	/dev/mapper/mpathdd - /dev/mapper/mpathdp /dev/mapper/mpatheb - /dev/mapper/mpathep /dev/mapper/mpathfa - /dev/mapper/mpathfp /dev/mapper/mpathga - /dev/mapper/mpathgp	128 × 200 GB	Application data

Oracle Server Configuration

Table 9 shows the operating system configurations for the Oracle servers.

TABLE 9. ORACLE SERVER OPERATING SYSTEM CONFIGURATIONS

Server Configuration	Server Operating System Setting Details
RPMs for Oracle Database 19c	binutils-2, compat-libcap1-1, compat-libstdc++-33-3, gcc-4, gcc-c++-4, glibc-2, glibc-devel-2, ksh, libaio-0, mlibaio-devel-0, libgcc-4, libstdc++-4, libstdc++-devel-4, libXi-1, libXtst-1, make-3, sysstat-10
/etc/multipath.conf	user_friendly_names yes find_multipaths yes path_grouping_policy multibus path_selector "service-time 0"

TABLE 9. ORACLE SERVER OPERATING SYSTEM CONFIGURATIONS (CONTINUED)

Server Configuration	Server Operating System Setting Details
/etc/security/limits.conf	oracle soft nofile 1024 oracle hard nofile 65536 oracle soft nproc 16384 oracle hard nproc 16384 oracle soft stack 10240 oracle hard stack 32768 oracle hard memlock 356537484 oracle soft memlock 356537484 grid hard nofile 65536
/etc/sysctl.conf	fs.aio-max-nr = 6291456 fs.file-max = 6815744 kernel.shmall = 2097152 kernel.shmmax = 536870912 kernel.shmmni = 4096 kernel.sem = 250 32000 100 128 net.ipv4.ip_local_port_range = 9000 65500 net.core.rmem_default = 262144 net.core.rmem_max = 4194304 net.core.wmem_default = 262144 net.core.wmem_max = 1048586
Swap space	64 GB

For an OLTP database, Hitachi Vantara recommends setting "service-time 0" for path_selector in /etc/multipath.conf file for the best performance.

Management Server Configuration

Table 10 shows the Hitachi Virtual Storage Platform G900 configuration for the management server cluster.

TABLE 10. Hitachi Virtual Storage Platform 900 Management Server Cluster Configuration

Attribute	Value
Purpose	VMware shared datastores Command device
RAID level	RAID-10 (2D+2D)
Drive type	3 TB 7.2K RPM SAS
Number of drives	4
Number of spare drives	0
Number and size of LDEVs for datastores	1 × 3000 GB
Number and size of command devices	1 × 100 MB
Storage port for management servers	7A, 7B, 8A, 8B

Server and Application Architecture

This reference architecture uses four Hitachi Advanced Server DS220 systems with second generation Intel Xeon scalable processors for each storage system architecture that was tested.

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

Two Hitachi Advanced Server DS120 systems are used for VMware ESXi management server configuration.

Table 11 describes the details of the server configurations for this solution.

TABLE 11. HITACHI ADVANCED SERVER DS220 AND ADVANCED SERVER DS120 SPECIFICATIONS

Server	Server	Host Name	Role	CPU Core	RAM
DS220	Oracle Server 1	rac01	Oracle RAC node 1	36 (2 × 18C)	768 GB (64 GB × 12)
	Oracle Server 2	rac02	Oracle RAC node 2	36 (2 × 18C)	768 GB (64 GB × 12)
	Oracle Server 3	rac03	Oracle RAC node 3	36 (2 × 18C)	768 GB (64 GB × 12)
	Oracle Server 4	rac04	Oracle RAC node 4	36 (2 × 18C)	768 GB (64 GB × 12)
DS120	Management Server 1	VMware ESXi 1	Management virtual machine hosts	16 (2 × 8C)	256 GB (32 GB × 8)
	Management Server 2	VMware ESXi 2		16 (2 × 8C)	256 GB (32 GB × 8)

SAN Architecture

Map the provisioned LDEVs to multiple ports on each Hitachi Virtual Storage Platform 5500. These LDEV port assignments provide multiple paths to the storage system from the host for high availability. This reference architecture uses two dual port Emulex HBAs per Hitachi Advanced Server DS220.

- 16 SAN switch connections are being used for Virtual Storage Platform 5500 Fibre Channel ports.
- 16 SAN switch connections are being used for server HBA ports.
- 4 SAN switch connections are being used for Virtual Storage Platform Fibre Channel ports.
- 4 SAN switch connections are being used for management server HBA ports.

Table 12, "Hitachi Virtual Storage Platform 5500, Virtual Storage Platform G900, Servers, and Fibre Channel Switch Configurations," on page 17 shows details of the Fibre Channel switch connect configuration on Hitachi Virtual Storage Platform 5500, Virtual Storage Platform G900, and database servers.

TABLE 12. HITACHI VIRTUAL STORAGE PLATFORM 5500, VIRTUAL STORAGE PLATFORM G900, SERVERS, AND FIBRE CHANNEL SWITCH CONFIGURATIONS

Server	HBA Ports	Storage Host Group	Switch Zone	Storage System	Storage Port	Brocade G620 Switch
DS220 Server 1	HBA1_1	DS220_59_HBA1_1	DS220_59_HBA1_1_ASE47_108_1A	VSP 5500	1A	SAN-switch 1
	HBA1_2	DS220_59_HBA1_2	DS220_59_HBA1_2_ASE47_108_2A		2A	SAN-switch 2
	HBA2_1	DS220_59_HBA2_1	DS220_59_HBA2_1_ASE47_108_1B		1B	SAN-switch 1
	HBA2_2	DS220_59_HBA2_2	DS220_59_HBA2_2_ASE47_108_2B		2B	SAN-switch 2
DS220 Server 2	HBA1_1	DS220_60_HBA1_1	DS220_60_HBA1_1_ASE47_108_1C		1C	SAN-switch 1
	HBA1_2	DS220_60_HBA1_2	DS220_60_HBA1_2_ASE47_108_2C		2C	SAN-switch 2
	HBA2_1	DS220_60_HBA2_1	DS220_60_HBA2_1_ASE47_108_1D		1D	SAN-switch 1
	HBA2_2	DS220_60_HBA2_2	DS220_60_HBA2_2_ASE47_108_2D		2D	SAN-switch 2
DS220 Server 3	HBA1_1	DS220_61_HBA1_1	DS220_61_HBA1_1_ASE47_108_3A		3A	SAN-switch 1
	HBA1_2	DS220_61_HBA1_2	DS220_61_HBA1_2_ASE47_108_4A		4A	SAN-switch 2
	HBA2_1	DS220_61_HBA2_1	DS220_61_HBA2_1_ASE47_108_3B		3B	SAN-switch 1
	HBA2_2	DS220_61_HBA2_2	DS220_61_HBA2_2_ASE47_108_4B		4B	SAN-switch 2
DS220 Server 4	HBA1_1	DS220_62_HBA1_1	DS220_182_HBA1_1_ASE47_108_3C		3C	SAN-switch 1
	HBA1_2	DS220_62_HBA1_2	DS220_182_HBA1_2_ASE47_108_4C		4C	SAN-switch 2
	HBA2_1	DS220_62_HBA2_1	DS220_182_HBA2_1_ASE47_108_3D		3D	SAN-switch 1
	HBA2_2	DS220_62_HBA2_2	DS220_182_HBA2_2_ASE47_108_4D		4D	SAN-switch 2

TABLE 12. HITACHI VIRTUAL STORAGE PLATFORM 5500, VIRTUAL STORAGE PLATFORM G900, SERVERS, AND FIBRE CHANNEL SWITCH CONFIGURATIONS (CONTINUED)

Server	HBA Ports	Storage Host Group	Switch Zone	Storage System	Storage Port	Brocade G620 Switch
DS120 Server 1	HBA1_1	DS120_68__HBA1_1	DS120_68_HBA1_1_ASE32_50_7A	VSP G900	7A	SAN-switch 1
	HBA1_2	DS120_68_HBA1_2	DS120_68_HBA1_2_ASE32_50_8A		8A	SAN-switch 2
DS120 Server 2	HBA1_1	DS120_69_HBA1	DS120_69_HBA1_1_ASE32_50_7B		7B	SAN-switch 1
	HBA1_2	DS120_69_HBA2	DS120_69_HBA1_2_ASE32_50_8B		8B	SAN-switch 2

Note – In a production environment without a storage system such as Hitachi Virtual Storage Platform G900 for management servers, separate storage ports on the Virtual Storage Platform 5000 series are recommended for the management servers to avoid impact on the database performance. While shared storage ports can be used, monitor port utilization to avoid performance issues in extreme high performance environments.

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 the cache synchronization of Oracle Real Application Clusters and inter-node communication among the nodes in the cluster.
- **Public Oracle Network** – This network provides client connections to Oracle Real Application Clusters and other applications.
- **Management Network** – This network is for hardware management console connections.

Hitachi Vantara recommends using pairs of 25 Gb/s NICs for the cluster interconnect network and public network.

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

- For each server in the 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.

Physical Network Configuration

Figure 3 on page 19 shows the IP network switch connection.

Figure 3

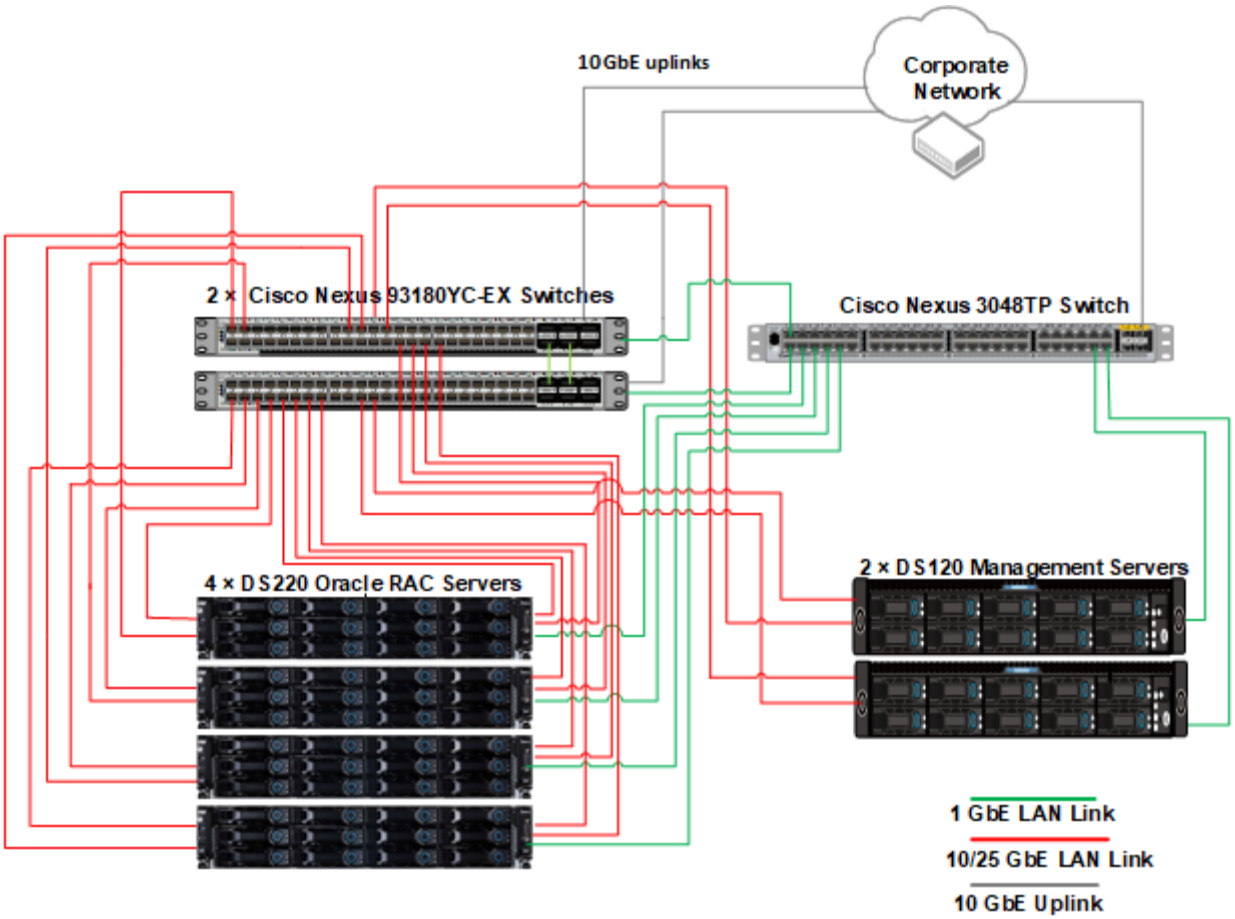


Table 13, "Database Server Network Configuration," on page 20 shows the network configuration and Table 14, "Virtual IP Address and Scan Name Configuration," on page 21 shows the virtual IP address and SCAN name configuration used when testing the environment. Your values may be different.

Configure NIC bonding or NIC teaming with NIC ports across multiple physical NICs to avoid a single point of failure.

TABLE 13. DATABASE SERVER NETWORK CONFIGURATION

Server	NIC Port	VLAN or Subnet	NIC Bond	IP Address	Network	Bandwidth (Gb/s)	Cisco Nexus 93180YC-EX Switch	
							Switch Number	Port
Oracle Database Server1	NIC 1 PORT 1	208	Bond0	192.168.208.91	Private	25	1	31
	NIC 2 PORT 1					25	2	
	NIC 1 PORT 2	242	Bond1	192.168.242.91	Public Oracle	25	1	32
	NIC 2 PORT 2					25	2	
	BMC- Dedicated NIC	242	-	192.168.242.161	Management	1	-	
Oracle Database Server2	NIC 1 PORT 1	208	Bond0	192.168.208.92	Private	25	1	33
	NIC 2 PORT 1					25	2	
	NIC 1 PORT 2	242	Bond1	192.168.242.92	Public Oracle	25	1	34
	NIC 2 PORT 2					25	2	
	BMC- Dedicated NIC	242	-	192.168.242.162	Management	1	-	
Oracle Database Server 3	NIC 1 PORT 1	208	Bond0	192.168.208.93	Private	25	1	31
	NIC 2 PORT 1					25	2	
	NIC 1 PORT 2	242	Bond1	192.168.242.93	Public Oracle	25	1	32
	NIC 2 PORT 2					25	2	
	BMC- Dedicated NIC	242	-	192.168.242.163	Management	1	-	
Oracle Database Server 4	NIC 1 PORT 1	208	Bond0	192.168.208.94	Private	25	1	31
	NIC 2 PORT 1					25	2	
	NIC 1 PORT 2	242	Bond1	192.168.242.94	Public Oracle	25	1	32
	NIC 2 PORT 2					25	2	
	BMC- Dedicated NIC	242	-	192.168.242.164	Management	1	-	
Management Server 1	BMC- Dedicated NIC	242	-	192.168.242.169	Management	1	-	
	NIC 1 PORT 1	244	-	192.168.244.101	Public Oracle	25	1	1
Management Server 2	BMC- Dedicated NIC	242	-	192.168.242.170	Management	1	-	
	NIC 1 PORT 1	244	-	192.168.244.102	Public Oracle	25	1	1

TABLE 14. VIRTUAL IP ADDRESS AND SCAN NAME CONFIGURATION

Server	Virtual IP Address	Scan Name pub-scan
Database Server 1	192.168.242.91	192.168.242.222
Database Server 2	192.168.242.92	192.168.242.223
Database Server 3	192.168.242.93	192.168.242.224
Database Server 4	192.168.242.94	

Table 15 lists the virtual machine configuration running on the management server cluster. Virtual machine configurations listed in Table 15 are required for the versions used for testing in the lab.

TABLE 15. MANAGEMENT SERVER CLUSTER VIRTUAL MACHINE CONFIGURATION

Virtual Machine	vCPU	Virtual Memory	Disk Capacity	IP Address
vCenter	2	10 GB	300 GB	192.168.242.102
Hitachi Oracle Adapters	2	6 GB	40-50 GB	192.168.242.80
Hitachi Ops Center Administrator	4	16 GB	100 GB	192.168.242.81
Hitachi Ops Center Analyzer	4	32 GB	800 GB	192.168.242.194
Hitachi Ops Center Analyzer detail view	4	10 GB	110 GB	192.168.242.197

Engineering Validation

This summarizes the key observations from the test results for the Hitachi Unified Compute Platform architecture to deploy Oracle RAC with Hitachi Virtual Storage Platform 5500.

Test Methodology

This describes the methodology used for testing.

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.

Peakmarks

The test results are created using peakmarks OLTP test cases.

Peakmarks is the leading benchmark software for Oracle platforms that is used for the following:

- Performance verification (quality assurance)
- Evaluation of different infrastructure products, technologies, and solutions (price/performance comparison)
- Performance optimization (improvement in efficiency)

This provides transparency and comparability in price versus performance considerations for Oracle infrastructures. The peakmarks 9.4 tool was used to validate this solution.

Database Configuration

Table 16 shows parameter details for four-node Oracle Real Application Clusters ASM database.

TABLE 16. FOUR-NODE ORACLE RAC ASM DATABASE PARAMETERS

Oracle Database Parameter	Value
compatible	19.3.0.0.0
cluster_database	TRUE
cluster_database_instances	4
Oracle Database size	16 TB
Database storage type	ASM
Database fill factor	80%

Test Environment

Table 17 lists the test environment.

TABLE 17. TEST ENVIRONMENT DETAILS

Item	Description	Value
Operating system		OL 7.6
Workload type		OLTP/OLAP
Database size		16 TB
Number of physical cores	CPU cores	144
Memory	Server memory	768 GB per node
Network	Cluster interconnect	2 × 25 Gb/s NIC bonding

Test Results

Table 18 shows results of the Orion and peakmarks tests.

TABLE 18. ORION AND PEAKMARKS TEST RESULTS

Test Category	Test Case	Test Results			
		Without Database Smart Flash Cache		With Database Smart Flash Cache	
		Throughput	RT (ms)	Throughput	RT (ms)
ORION	8KB Random Read	2,522,547 IOPS	0.8	N/A	N/A
	8KB Random Write	543,975 IOPS	0.9	N/A	N/A
	1MB Sequential Read	27.51 GB/s	N/A	N/A	N/A
	1MB Sequential Write	12.73 GB/s	N/A	N/A	N/A
peakmarks[a]	STO-READ	40.44 GB/s	N/A	N/A	N/A
	STO-Random 100% read	1,542,338 IOPS	0.74	N/A	N/A
	SRV-SCAN	202.80 GB/s	N/A	N/A	N/A
	DBWR-THR	8.63 GB/s	N/A	N/A	N/A
	DA-LOW/ROW	84.42 GB/s	N/A	N/A	N/A
	DA-LOW/STO	26.53 GB/s	N/A	N/A	N/A
	TP-LIGHT 1 RPT 0% update	1,047,591 TPS	0.54	1,600,153 TPS	0.35
	TP-LIGHT 1 RPT 20% update	210,361 TPS	0.96	313,685 TPS	0.91
	TP-LIGHT 1 RPT 30% update	204,361 TPS	1.0	315,276 TPS	0.90
	TP-MEDIUM 25 RPT 0% update	69,164 TPS	4.10	88,805 TPS	3.20
	TP-MEDIUM 25 RPT 20% update	13,472 TPS	9.3	14,180 TPS	8.90
	TP-MEDIUM 25 RPT 30% update	13,025 TPS	7.25	14,241 TPS	8.82

a. Only peakmarks TP test cases are supported with Oracle Database Smart Flash Cache in the peakmarks 9.4 release.

Feedback

Hitachi Vantara 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.

Demonstrations and other resources are available for many Hitachi products. To schedule a live demonstration, contact a sales representative or partner. To view on-line informational resources, see the Resources website.

Hitachi Academy is your education destination to acquire valuable knowledge and skills on Hitachi products and solutions. Our Hitachi Certified Professional program establishes your credibility and increases your value in the IT marketplace. For more information, see the Hitachi Vantana Training and Certification website.

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

Hitachi Vantara

Corporate Headquarters
2845 Lafayette Street
Santa Clara, CA 95054 USA
www.HitachiVantara.com | community.HitachiVantara.com

Regional Contact Information
USA: 1-800-446-0744
Global: 1-858-547-4526
HitachiVantara.com/contact

MK-SL-169-03, April 2020