HITACHI
Inspire the Next

ARCHITECTURE GUIDE

Hitachi Solution for Databases - Reference Architecture for Oracle Real Application Clusters Database 12c with Global-Active Device using Hitachi Data Instance Director

Reference Architecture Guide

By Amol Bhoite

August 2019

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-119-00	Initial release	November 7, 2018
MK-SL-119-01	Fix minor errors	November 9, 2019
MK-SL-119-02	Supports the HDID 2 datacenter swap feature available in the HDID 6.7 release.	August 1, 2019

Table of Contents

Solution Overview **2**

Business Benefits 3

High Level Infrastructure 3

Key Solution Components **5**

Hitachi Virtual Storage Platform G and F Series Family 8

Hitachi Storage Virtualization Operating System RF 8

Hitachi Advanced Server DS220 Server 8

Hitachi Advanced Server DS120 Server 9

Red Hat Enterprise Linux 9

Device Mapper Multipathing 9

Hitachi **Data Instance Director** **9**

Hitachi Infrastructure Analytics Advisor 9

Hitachi Storage Advisor 10

Hitachi Storage Adapter for Oracle Enterprise Manager 10

Hitachi Server Adapter for Oracle Enterprise Manager 10

Oracle Database With the Real Application Clusters Option 10

Oracle Enterprise Manager 10

VMware Esxi 11

vCenter Appliance 11

Brocade Switches 11

Cisco Switches 11

Solution Design **11**

Storage Architecture 12

Server and Application Architecture 18

SAN Architecture 19

Network Architecture 25

Global-active Device Setup Pre-configuration 28

Solution Implementation **29**

Deploy the Solution 29

Solution Execution 33

Benefits of using HDID Versus Using Manual Commands for Global-active Device Setup and Configuration **42**

Engineering Validation **43**

Test Methodology 43

Test Results 44

Hitachi Solution for Databases - Reference Architecture for Oracle Real Application Clusters Database 12c with Global-Active Device using Hitachi Data Instance Director

Reference Architecture Guide

Use this reference architecture guide to design a solution with Hitachi Data Instance Director (HDID) to protect Hitachi Unified Compute Platform for non-multitenant Oracle Database 12c. This solution is for Oracle Real Application Clusters on Extended Distance (Stretched) clusters in a two-site environment using global-active device in Hitachi Virtual Storage Platform.

This explains how to use HDID to deploy global-active device to add backup and recovery capabilities in an Oracle environment to achieve zero recovery point objective (RPO) and recovery time objective (RTO). Use global-active device in a two-site replication environment with Virtual Storage Platform storage to provide data protection for Oracle Database. This guide also explains how to use HDID to perform automated 2 datacenter swap global-active device replication on demand and automated recovery of global-active device replication in an error or suspended state.

This Hitachi Unified Compute Platform CI architecture for Oracle Database is engineered, pre-tested, and qualified to provide predictable performance and the highest reliability in demanding, dynamic Oracle environments. This solution is validated to ensure consistent, predictable results.

This proven solution optimizes your Oracle database environment, and integrates servers, storage systems, network, and storage software. This provides reliability, high availability, scalability, and performance while processing small-scale to large-scale OLTP workloads. The dedicated servers run Oracle Database 12c Release 2 with the Oracle Real Application Cluster option. The operating system is Red Hat Enterprise Linux 7.6.

Tailor your implementation of these best practices to meet your specific data backup and recovery needs.

The practices in this guide are valid for all storage systems that support global-active device and are not limited to the storage environment used to validate these best practices.

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 GX00
- Hitachi Advanced Server DS220 servers
- Hitachi Advanced Server DS120 servers
- Storage area networks
- Oracle RAC Database 12c Release 2
- Oracle Automatic Storage Management (Oracle ASM)
- Hitachi Global-active Device
- Hitachi Data Instance Director (HDID)
- Hitachi Adapters for Oracle Database
- Hitachi Storage Adapter for Oracle Enterprise Manager
- Hitachi Server Adapter for Oracle Enterprise Manager
- Red Hat Enterprise Linux
- Red Hat Enterprise Linux Device-Mapper Multipath

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 Hitachi Unified Compute Platform CI for Oracle Real Application Clusters on Extended Distance clusters on four nodes using Hitachi Virtual Storage Platform G900. This environment addresses the high availability, performance, and scalability requirements for OLTP and OLAP workloads. Your solution implementation can be tailored to meet your specific needs.

Continuous application availability in traditional and cloud designs requires continuous storage. This solution uses the unique Hitachi Storage Virtualization Operating System (SVOS) and enterprise-class Hitachi Virtual Storage Platform G-series systems for the following:

- Global storage virtualization
- Distributed continuous storage
- Zero recovery time and point objectives (RTO/RPO)
- Simplified distributed system design and operations

Global storage virtualization provides “global active volumes.” These are storage volumes with the ability to have read and write copies of the same data in two systems at the same time. The active-active storage design enables production workloads on both systems in a local or metro cluster configuration while maintaining full data consistency and protection

Configuring Oracle Real Application Clusters on extended distance with global-active device allows you to create and maintain synchronous, remote copies of data volumes on Hitachi Virtual Storage Platform F or VSP G series storage.

Business Benefits

This reference architecture provides the following benefits:

- Continuous server I/O when an unplanned outage, such as disaster or hardware failure, prevents access to a data volume of the database
- Automated configuration for global-active devices and quick recovery of global-active device pairs in error or suspended state storage operations using a web-based HDID UI without knowledge of manual Hitachi HORCM configurations
- Automated Pause, Resume, 2 datacenter replication swap, dissociate, revert, teardown, and delete global-active devices using HDID for planned outages
- Easy to understand global-active device internal operations using HDID informative log messages. This helps to quickly identify problems and complete troubleshooting.

High Level Infrastructure

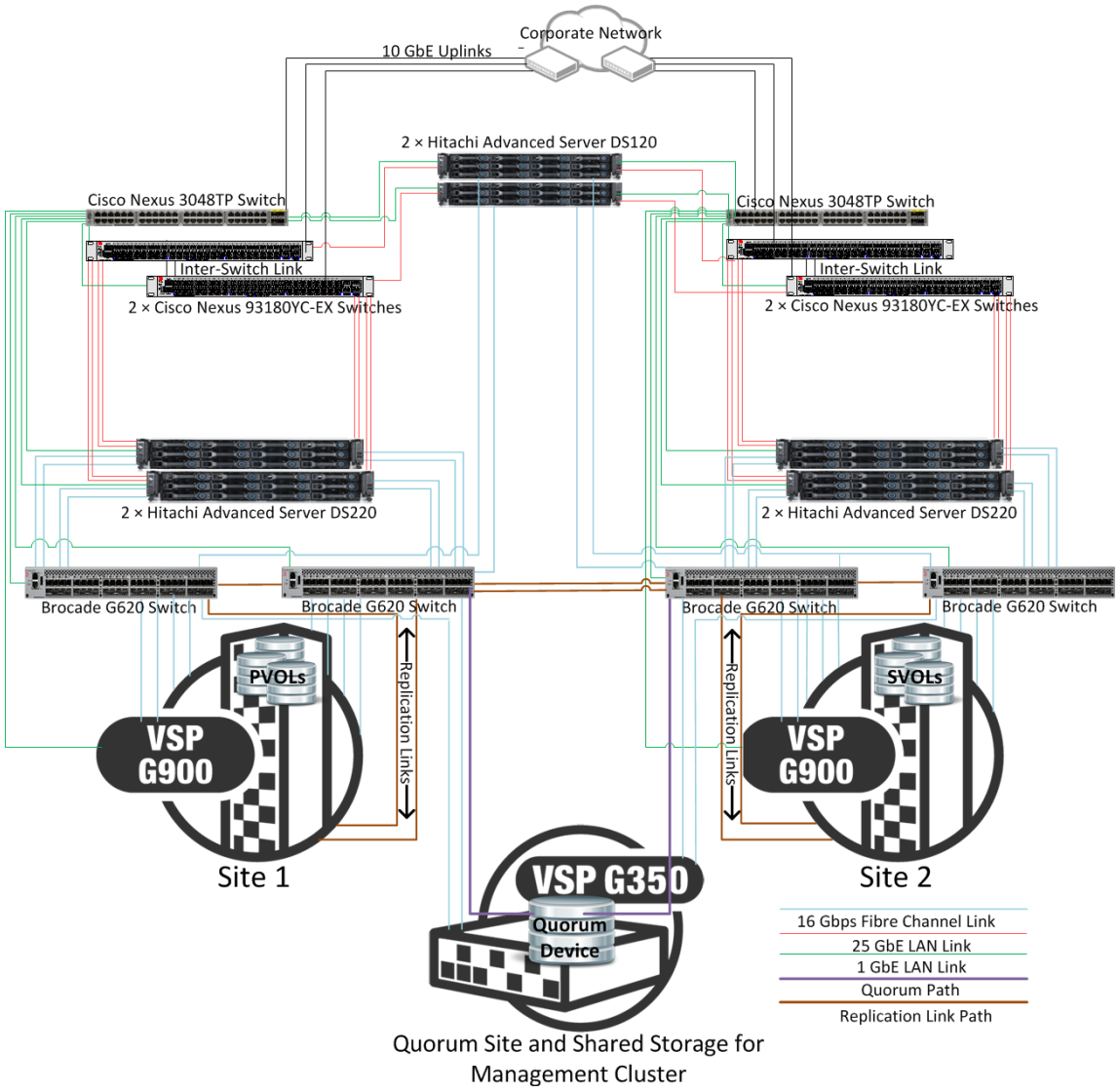
Figure 1 shows the high-level infrastructure for this solution.

The configuration of Virtual Storage Platform G900 and Hitachi Advanced Server DS220 have the following characteristics:

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

This high-level global-active device infrastructure is hosted in a single site environment. With a WAN, the physical configuration would be different.

Figure 1



To avoid any performance impact to the production database, Hitachi Vantara 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

The uplink speed to the corporate network depends on the customer environment and requirements. The Cisco Nexus 93180YC-EX switches can support uplink speeds of 25 GbE, 40 GbE, or 100 GbE if higher bandwidth is required.

Note — In the lab environment the management server setup was configured at Site 3. In the customer environment, the management server can be configured at Site 1 or Site 2.

Key Solution Components

The key solution components for this solution are listed in Table 1, Table 2, and Table 3.

TABLE 1. HARDWARE COMPONENTS

Hardware	Detailed Description	Firmware/Version	Quantity
Hitachi Virtual Storage Platform G900	■ Two Controllers ■ 16 × 16 Gbps Fibre Channel Ports ■ 8 × 12 Gbps Backend SAS Ports ■ 512 GB cache memory ■ 64 × 1.9 TB SSDs Plus 2 spares	88-03-24-60/00	2
Hitachi Virtual Storage Platform G350	■ Two Controllers ■ 4 × 16 Gbps Fibre Channel Ports ■ 128 GB cache memory ■ 4 × 6 TB 7.2 krpm SAS drives	88-03-24-60/00	1
Hitachi Advanced Server DS220 servers	■ 2 × Intel Xeon Gold 6140 CPU @ 2.30GHz ■ 768 GB (64 GB × 12) DIMM DDR4 Synchronous Registered (Buffered) 2666 MHz	BIOS: 3A10.H8 BMC: 4.23.06 CPLD:10	4
	■ 2 × Intel Corporation Ethernet Controller XXV710 for 25GbE SFP28	Driver: i40e Version: 2.3.2-k Firmware: 6.02 0x80003620 1.1747.0	
	■ 2 × Emulex LightPulse LPe31002-M6 2-Port 16 Gb Fibre Channel Adapter	Driver: 12.0.0.5 Boot: 11.4.204.34 Firmware: 11.4.204.34	

TABLE 1. HARDWARE COMPONENTS (CONTINUED)

Hardware	Detailed Description	Firmware/Version	Quantity
Hitachi Advanced Server DS120 server	■ 2 × Intel Xeon Sliver Processor 4110, 8-core, 2.1GHz, 85W ■ 8 × 32GB DDR4 R-DIMM 2666 Mhz (256GB total) ■ 1 × 64 GB SATADOM	BIOS: 3A10.H8 BMC: 4.23.06 CPLD:10	2
	■ 1 × Intel Corporation Ethernet Controller XXV710 for 25 GbE SFP28	Driver: i40en Version: 1.3.1 Firmware: 5.51 0x80002bca 1.1568.0	
	■ 1 × Emulex LightPulse LPe31002-M6 2-Port 16 Gb Fibre Channel Adapter	Driver: 11.1.0.6 Boot: 11.2.154.0 Firmware: 11.2.156.27	
Brocade Fibre Channel Switches	■ G620 ■ 48 port Fibre Channel switch ■ 16 Gbps SFPs ■ Brocade hot-pluggable SFP+, LC connector	Kernel: 2.6.34.6 Fabric OS: v8.2.0b	4
Cisco Nexus	■ C93180YC-EX ■ 48 × 10/25 GbE fiber ports ■ 6 × 40/100 Gbps Quad SFP (QSFP28) ports	BIOS: version 07.61 NXOS: version 7.0(3)I4(7)	4
Cisco Nexus	■ 3048TP ■ 1 GE 48-Port Gb Ethernet Switch	BIOS: version 4.0.0 NXOS: version 7.0(3)I4(7)	2

TABLE 2. SOFTWARE COMPONENTS FOR COMPUTE NODES

Software	Version	Function
Red Hat Enterprise Linux	RHEL 7.6 (Kernel Version - 3.10.0-957.21.2.el7.x86_64)	Operating System
Oracle 12c	12c Release 2 (12.2.0.1.0)	Database Software
Oracle Real Application Cluster	12c Release 2 (12.2.0.1.0)	Cluster Software
Oracle Grid Infrastructure	12c Release 2 (12.2.0.1.0)	Volume Management, File System Software, and Oracle Automatic Storage Management
Red Hat Enterprise Linux Device Mapper Multipath	-	Multipath Software

TABLE 3. SOFTWARE COMPONENTS FOR MANAGEMENT NODES

Software	Version	Function
VMware ESXi	Version 6.7.0 Build 10302608	ESXi for management nodes
VMware vCenter Server	Version 6.7.0 Build 10244745	Management cluster
Hitachi Storage Virtualization Operating System	SVOS RF 8.3.1	Global-active Device - Replication software
Hitachi Data Instance Director (HDID)	6.7.8	Data protection software
Hitachi Command Control Interface software (CCI)	01-49-03/01	Storage configuration and data management software
Hitachi Storage Advisor (HSA)	2.3	Storage orchestration software
Hitachi Infrastructure Analytics Advisor (HIAA)	4.2.0-01	Analytics Software
Manager for Hitachi adapters for Oracle Database	2.3.1	Hitachi adapters management Virtual appliance software
Hitachi Storage Adapter for Oracle Enterprise Manager	2.2.3	Storage management software
Hitachi Server Adapter for Oracle Enterprise Manager	2.2.3	Server management software
Oracle Enterprise Manager Cloud Control 13c	13c Release 2 (13.2.0.0)	OEM software

TABLE 3. SOFTWARE COMPONENTS FOR MANAGEMENT NODES (CONTINUED)

Software	Version	Function
Oracle Enterprise Manager Cloud Control 13c plug-ins	13c Release 2	Hitachi Storage and Server OEM plugins
Virtual SVP (vSVP)	Microcode dependent	Storage management software

Hitachi Virtual Storage Platform G and F Series Family

Use Hitachi Virtual Storage Platform F series family storage for a flash-powered cloud platform for your mission critical applications. This storage meets demanding performance and uptime business needs. Extremely scalable, its 4.8 million random read IOPS allows you to consolidate more applications for more cost savings.

This solution uses Virtual Storage Platform F900/G900, which supports Oracle Real Application Clusters.

Hitachi Storage Virtualization Operating System RF

SVOS RF is at the heart of the Virtual Storage Platform F series family. It provides storage virtualization, high availability, flash optimized performance, quality of service controls, and advanced data protection. This proven, mature software provides common features, management, and interoperability across the Hitachi portfolio. This means you can reduce migration efforts, consolidate assets, reclaim space, and extend life.

Global-active device enables you to create and maintain synchronous, remote copies of data volumes. A virtual storage machine is configured in the primary and secondary storage systems using the actual information of the primary storage system, and the global-active device primary and secondary volumes are assigned the same virtual LDEV number in the virtual storage machine. This enables the host to see the pair volumes as a single volume on a single storage system, and both volumes receive the same data from the host.

A quorum disk, which can be located in a third and external storage system or in an iSCSI-attached host server, is used to monitor the global-active device pair volumes. The quorum disk acts as a heartbeat for the global-active device pair, with both storage systems accessing the quorum disk to check on each other.

Hitachi Advanced Server DS220 Server

Hitachi Advanced Server DS220 is a general-purpose rackmount server designed for optimal performance and power efficiency. This allows owners to upgrade computing performance without overextending power consumption and offers non-latency support to virtualization environments that require the maximum memory capacity. Hitachi Advanced Server DS220 provides flexible I/O scalability for today’s diverse data center application requirements.

Hitachi Advanced Server DS120 Server

Hitachi Advanced Server DS120 provides flexible and scalable configurations for hyper-converged datacenters, provides computing performance, sophisticated power and thermal design to avoid unnecessary OPEX with quick deployment. For this solution two DS120 servers are used. The two DS120 servers are configured as a VMware vCenter cluster. Virtual machines on the cluster are used to host management applications. The management applications installed depend on customer needs and requirements. The following applications were installed in individual virtual machines in this architecture and would be installed in most cases.

- Hitachi Data Instance Director
- Hitachi Command Control Interface software (CCI)
- vCenter
- Oracle Enterprise Manager (OEM) 13c
- Oracle Adapter Manager
- Hitachi Storage Advisor (HSA)
- Hitachi Infrastructure Analytics Advisor / Hitachi Datacenter Analytics (HIAA/HDCA)
- HDCA Probe

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

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.

Device Mapper Multipathing

Device mapper multipathing (DM-Multipath) allows you to configure multiple I/O paths between server nodes and storage arrays into a single device.

These I/O paths are physical SAN connections that can include separate cables, switches, and controllers. Multipathing aggregates the I/O paths, creating a new device that consists of the aggregated paths.

Hitachi Data Instance Director

Hitachi Data Instance Director is a copy data management platform that simplifies creating and managing policy-based workflows that support business functions with controlled copies of data. Hitachi Data Instance Director provides business-defined data protection for organizations looking to modernize, simplify and unify their operational recovery, disaster recovery and long-term retention operations.

Hitachi Infrastructure Analytics Advisor

With Hitachi Infrastructure Analytics Advisor, you can define and monitor storage service level objectives (SLOs) for resource performance. You can identify and analyze historical performance trends to optimize storage system performance and plan for capacity growth.

Use Hitachi Infrastructure Analytics Advisor to register resources (storage systems, hosts, servers, and volumes), and set service-level thresholds. You are alerted to threshold violations and possible performance problems (bottlenecks). Using analytics tools, you find which resource has a problem and analyze its cause to help solve the problem. The Infrastructure Analytics Advisor ensures the performance of your storage environment based on real-time SLOs.

Hitachi Storage Advisor

Hitachi Storage Advisor is an infrastructure management solution that unifies storage management solutions such as storage provisioning, data protection, and storage management; simplifies the management of large scale data centers by providing smarter software services; and is extensible to provide better programmability and better control.

Hitachi Storage Adapter for Oracle Enterprise Manager

Hitachi Storage Adapter for Oracle Enterprise Manager presents an integrated, detailed view of the Hitachi storage supporting your Oracle databases. By gaining visibility into capacity, performance and configuration information, administrators can manage service levels more effectively, and ensure service level agreements (SLAs) are met to support business goals.

Hitachi Server Adapter for Oracle Enterprise Manager

Hitachi Server Adapter for Oracle Enterprise Manager is an Oracle Enterprise Manager plug-in that enables monitoring of Hitachi Advanced servers in Oracle Enterprise Manager.

For Hitachi Advanced servers, it provides visibility into the components, including their status, health, and attributes. In addition, the adapter supplies information about any Oracle database instances running on the servers. Both RAC and non-RAC databases are supported.

Oracle Database With the Real Application Clusters Option

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.

Oracle Enterprise Manager

Oracle Enterprise Manager provides a “single pane of glass” that allows you to manage on-premises and cloud-based IT using the same familiar interface you know and use on-premises every day. Oracle Enterprise Manager today is the nerve center of IT operations among thousands of enterprises. Millions of assets in Oracle’s SaaS and PaaS public cloud operations are managed by Enterprise Manager round the clock.

Enterprise Manager is the industry’s first complete cloud solution with Cloud Management. This includes self-service provisioning balanced against centralized, policy-based resource management, integrated chargeback and capacity planning, and complete visibility of the physical and virtual environments from applications to disk.

This solution uses Oracle Enterprise Manager Cloud Control, version 13c release 2. This allows you to use these cloud management features:

- Use the Database Cloud Self Service Portal
- Benefit from the Improved Service Catalog
- Perform Snap Cloning using “Test Master Snapshot”
- Take advantage of the Chargeback and Consolidation Planner plugins

For more information, see New Features in Oracle Enterprise Manager Cloud Control 13c

VMware Esxi

VMware ESXi is the next-generation hypervisor, providing a new foundation for virtual infrastructure. This innovative architecture operates independently from any general-purpose operating system, offering improved security, increased reliability, and simplified management.

vCenter Appliance

The vCenter Server Appliance is a preconfigured Linux virtual machine, which is optimized for running VMware vCenter Server and the associated services on Linux.

vCenter Server Appliance comes as an Open Virtualization Format (OVF) template. The appliance is imported to an ESXi host and configured through the web-based interface. It comes pre-installed with all the components needed to run a vCenter Server, including vCenter SSO (Single Sign-on), Inventory Service, vSphere Web Client, and the vCenter Server itself.

Brocade Switches

Brocade and Hitachi Vantara 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.

SAN switches are optional and direct connect is also possible under certain circumstances, but customers should check the support matrix to ensure support prior to implementation.

The solution uses the Brocade G620, 48 port Fibre Channel switch.

Cisco Switches

The Cisco Nexus Switch product line provides a series of solutions that can make it easier to connect and manage disparate data center resources with software-defined networking (SDN). Leveraging the Cisco Unified Fabric, which unifies storage, data and networking (Ethernet/IP) services, the Nexus Switches create an open, programmable network foundation built to support a virtualized data center environment.

The solution uses the following Cisco products:

- Nexus 93180YC-EX, 48-port 10/25 GbE switch
- Nexus 3048TP, 48-port 1 GbE Switch

Solution Design

This describes the reference architecture environment to implement Hitachi Unified Compute Platform CI for Oracle Real Application Clusters on Extended Distance clusters on four nodes using Hitachi Virtual Storage Platform. The environment used for testing and validation of this solution used Hitachi Virtual Storage Platform G900.

The infrastructure configuration includes the following:

- **Site 1**
  - **Oracle RAC Servers** — Two server nodes were configured in an Oracle Real Application Cluster.
  - **Storage System** — There are VVOLs mapped to each port that are presented to the server as LUNs.
  - **SAN Connections** — There are SAN connections to connect the Fibre Channel HBA ports to the storage through Brocade G620 switches.
- **Site 2**
  - **Oracle RAC Servers** — Two server nodes were configured in an Oracle Real Application Cluster.
  - **Storage System** — There are VVols mapped to each port that are presented to the server as LUNs.
  - **SAN Connection** — There are SAN connections to connect the Fibre Channel HBA ports to the storage through Brocade G620 switches.
- **Site 3**
  - **Quorum Site**
    - **Storage System** — The Hitachi Virtual Storage Platform G350 used as a quorum device had an LDEV mapped to two ports presented as an external volume at site 1 and site 2 to each Virtual Storage Platform G900 on the sites.

Note — Testing used a separate Hitachi Virtual Storage Platform G350 storage system for the quorum device. When implementing this, you may use any other supported storage system.

Testing used a quorum disk, located in a third storage system and used to monitor the global-active device pair volumes. Global-active device Active-Active configuration without a quorum disk is also supported with the latest SVOS version.

- **Management server cluster**
  - Install one Hitachi Data Instance Director master node on the virtual machine.
- **A proxy node virtual** machine managed and monitored global**-active device pair operations is required for** the **P-VOLs only.**
- **SAN Connection** — Each 16 Gb/sec Fibre Channel HBA port was connected to the storage front-end ports through a switched SAN fabric.

Storage Architecture

This describes the storage architecture for this solution.

Storage Configuration

The storage configuration takes into consideration Hitachi Vantara for Hitachi Virtual Storage Platform and Oracle recommended best practices for the design and deployment of database storage.

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

Figure 2

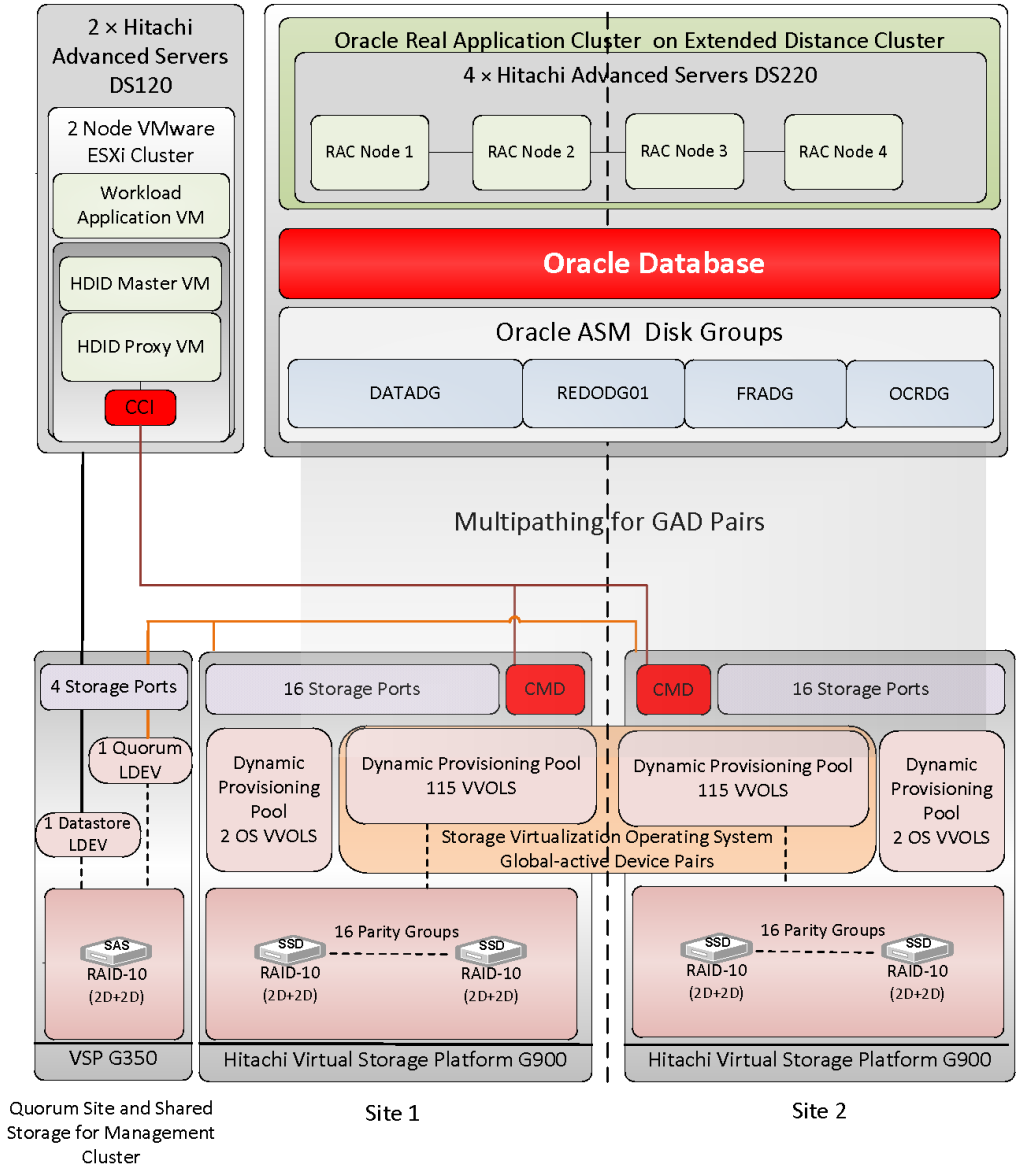


Table 4 shows the storage pool configuration used for this solution. In the current configuration OS and Oracle LDEVs are in different storage pools; however, users can create a single pool for OS and Oracle LDEVs.

TABLE 4. STORAGE POOL CONFIGURATION

Site	Site 1		Site 2	
Pool ID	Hitachi-UCP-CI-OS	Hitachi-UCP-CI-Oracle	Hitachi-UCP-CI-OS	Hitachi-UCP-CI-Oracle
Pool Type	Dynamic Provisioning	Dynamic Provisioning	Dynamic Provisioning	Dynamic Provisioning
RAID Group	1-1 – 1-1	1-2 – 1-17	1-1 – 1-1	1-2 – 1-17

TABLE 4. STORAGE POOL CONFIGURATION (CONTINUED)

RAID Level	RAID-10 (2D+2D)	RAID-10 (2D+2D)	RAID-10 (2D+2D)	RAID-10 (2D+2D)
Drive Type	1.9 TB SSDs	1.9 TB SSDs	1.9 TB SSDs	1.9 TB SSDs
Number of Drives	4	60	4	60
Number of Pool Volume LDEVs	1	64	1	64
Pool Volume LDEV size	880 GB	880 GB	880 GB	880 GB
Pool Capacity	880 GB	54.99 TB	880 GB	54.99 TB

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

TABLE 5. LOGICAL STORAGE CONFIGURATION

Site	Site 1		Site 2	
Pool ID	Hitachi-UCP-CI-OS	Hitachi-UCP-CI-Oracle	Hitachi-UCP-CI-OS	Hitachi-UCP-CI-Oracle
Number of VVOLs	2	115 PVOLs	2	115 SVOLs

TABLE 5. LOGICAL STORAGE CONFIGURATION (CONTINUED)

VVOL Size	2 × 200 GB	64 × 160 GB, 32 × 40 GB, 16 × 10 GB, 3 × 60 GB	2 × 200 GB	64 × 160 GB, 32 × 40 GB, 16 × 10 GB, 3 × 60 GB
Purpose	Operating System	▪ Oracle ▪ System ▪ Sysaux ▪ Undo ▪ Temp ▪ Redo Logs ▪ Parameter and Password file ▪ Oracle Cluster Registry and Voting Disk	Operating System	▪ Oracle ▪ System ▪ Sysaux ▪ Undo ▪ Temp ▪ Redo Logs ▪ Parameter and Password file ▪ Oracle Cluster Registry and Voting Disk
Storage Port	1A, 1B, 2A, 2B, 3A, 3B, 4A, 4B	1A, 2A, 3A, 4A, 5A, 6A, 7A, 8A, 1B, 2B, 3B, 4B, 5B, 6B, 7B, 8B, 1C,1D,2C,2D	1A, 1B, 2A, 2B, 3A, 3B, 4A, 4B	1A, 2A, 3A, 4A, 5A, 6A, 7A, 8A, 1B, 2B, 3B, 4B, 5B, 6B, 7B, 8B, 1C,1D,2C,2D

On Site 3 VSP G350 storage there is an additional RAID group consisting of four 6 TB 7.2 krpm SAS drives configured as RAID-10 (2D+2D).

This is used as shared storage for the management server cluster and for the quorum device. A single 6 TB LUN is mapped to four storage ports for the management server. 20 GB LDEV is used as a quorum device. 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 HDP pool, or it can use capacity on the HDP pool configured for the Oracle environment depending on customer requirements.

Database Layout

The database layout design uses recommended best practices from Hitachi Vantara for Hitachi Virtual Storage Platform G900 for small random I/O traffic, such as 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 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.

Oracle ASM Configuration

- **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 Data 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 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.
- **Database Block Size Settings** — Set the database block size to 8 KB.
- **ASM FILE SYSTEM I/O Settings** — Set the Oracle ASM I/O operations for database files as follows:
- FILESYSTEMIO_OPTIONS = setall

Table 6 shows the Oracle RAC Database Settings.

TABLE 6. ORACLE RAC DATABASE SETTINGS

Environment	*Value*
RAC configuration	Yes
ASM	Yes - Oracle RAC Database

Table 7 shows the Oracle Environment Parameters.

TABLE 7. ORACLE ENVIRONMENT PARAMETERS

Setting	*Value*
DB_CLOCK_SIZE	8 KB
SGA_TARGET	400 GB
PGA_AGGREGATE_TARGET	192 GB
DB_CACHE_SIZE	172 GB
DB_KEEP_CACHE_SIZE	96 GB
DB_RECYCLE_CACHE_SIZE	24 GB
INMEMORY_SIZE	48 GB
USE_LARGE_PAGES	TRUE

TABLE 7. ORACLE ENVIRONMENT PARAMETERS (CONTINUED)

Setting	*Value*
FILESYSTEMIO_OPTIONS	SETALL
DB_FILE_MULTIBLOCK_READ_COUNT	64
DISK_ASYNCH_IO	TRUE

Figure 3 shows the relationships between disk groups and replication pairs.

Figure 3

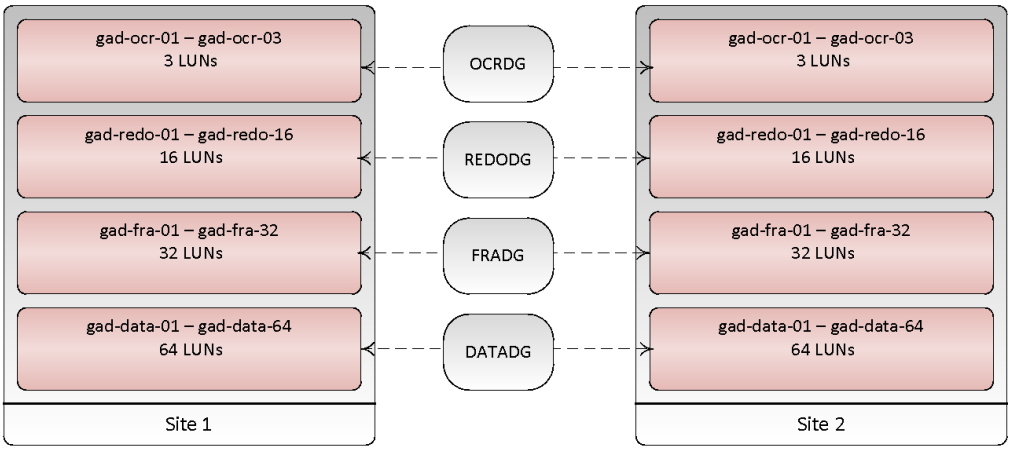


Table 8 shows the details of the disk mappings from the LUNs to the ASM disk groups for Oracle RAC Database tablespaces.

TABLE 8. LUNS AND ORACLE ASM DISK MAPPINGS FOR ORACLE DATABASE IN SITE 1 AND SITE 2

ASM Disk Group	*ASM Disk*	*DM-Multipath LUNs*	*LUN Details*	*Purpose*
NA	NA	/dev/mapper/mpatha	4 × 200 GB	OS and Oracle four node RAC Database
OCRDG	OCRDISK1 - OCRDISK3	/dev/mapper/mpathaa - /dev/mapper/mpathac	3 × 60 GB	Oracle Cluster Registry and Voting Disk
REDODG	REDODISK1 - REDODISK16	/dev/mapper/mpathad - /dev/mapper/mpathap /dev/mapper/mpathba - /dev/mapper/mpathbc	16 × 10 GB	Online REDO log group

TABLE 8. LUNS AND ORACLE ASM DISK MAPPINGS FOR ORACLE DATABASE IN SITE 1 AND SITE 2 (CONTINUED)

ASM Disk Group	*ASM Disk*	*DM-Multipath LUNs*	*LUN Details*	*Purpose*
FRADG	FRADISK1 - FRADISK32	/dev/mapper/mpathbd - /dev/mapper/mpathbp /dev/mapper/mpathcb - /dev/mapper/mpathcc	32 × 40 GB	Flash Recovery Area
DATADG	DATADISK1 - DATADISK64	/dev/mapper/mpathdd - /dev/mapper/mpathdp /dev/mapper/mpatheb - /dev/mapper/mpathep /dev/mapper/mpathfa - /dev/mapper/mpathfp /dev/mapper/mpathga - /dev/mapper/mpathgp	64 × 160 GB	Application Data

Server and Application Architecture

This reference architecture uses four Hitachi Advanced Server DS220 servers for a four-node Oracle RAC on an extended distance clusters configuration.

This provides the compute power for the Oracle RAC database to handle complex database queries and a large volume of transaction processing in parallel. Table 9 describes the details of the server configuration for this solution.

This reference architecture uses two Hitachi Advanced Server DS120 servers for *VMware ESXi* management server configuration.

Details of the VMware ESXi management servers are specified in Table 9.

TABLE 9. HITACHI ADVANCED SERVER DS220 AND DS120 SERVER SPECIFICATIONS

Hitachi Advanced Server	Server	Server Name	Role	CPU Core	RAM
Site 1 DS220	**Oracle Server1**	oracle-rac-01	Oracle RAC node 1	36	768 GB (64 GB × 12)
	Oracle Server2	oracle-rac-02	Oracle RAC node 2	36	768 GB (64 GB × 12)
Site 2 DS220	**Oracle Server3**	oracle-rac-03	Oracle RAC node 3	36	768 GB (64 GB × 12)
	Oracle Server4	oracle-rac-04	Oracle RAC node 4	36	768 GB (64 GB × 12)

TABLE 9. HITACHI ADVANCED SERVER DS220 AND DS120 SERVER SPECIFICATIONS (CONTINUED)

Hitachi Advanced Server	Server	Server Name	Role	CPU Core	RAM
Site 3 DS120	**Management server**	*VMware ESXi 1*	HDID Master VM HDID Proxy VM	16	256 GB (32 GB × 8)
		VMware ESXi 2	Workload application VM Hitachi Storage Advisor VM Hitachi Infrastructure Analytics Advisor VM Manager for Hitachi Adapters for Oracle Database VM Oracle Enterprise Manager Cloud Control 13c VM	16	256 GB (32 GB × 8)

SAN Architecture

Map the provisioned LDEVs to multiple ports on Hitachi Virtual Storage Platform G900 (VSP G900). These LDEV port assignments provide multiple paths to the storage system from the host for high availability.

- Site 1
- 16 SAN switch connections are being used for VSP G900 host ports.
- 16 SAN switch connections are being used for server HBA ports.
- Site 2
- 16 SAN switch connections are being used for VSP G900 host ports.
- 16 SAN switch connections are being used for server HBA ports.

Site 3:

- 4 SAN switch connections are being used for VSP G350 host ports.
- 4 SAN switch connections are being used for server HBA ports.

Table 10 shows details of the Fibre Channel switch connect configuration on the Hitachi Virtual Storage Platform G900 ports.

TABLE 10. SAN HBA CONNECTION CONFIGURATION BETWEEN DS220 AND VSP G900, DS120, AND VSP G350

Site	*Server*	*HBA Ports*	*Storage Host Group*	*Switch Zone*	*Connec-tion*	*Storage System*	*Port*	*Brocade G620 Switch*
Site 1	DS220 Server 1	HBA1_1	CN1_HBA1_1	CN1_HBA_1_1_ASE43_230_1A	Local	Site 1 VSP G900	1A	69
		HBA1_2	CN1_HBA1_2	CN1_HBA_1_2_ASE43_230_2A	Local	Site 1 VSP G900	2A	70
		HBA2_1	CN1_HBA2_1	CN1_HBA_2_1_ASE43_230_1B	Local	Site 1 VSP G900	1B	69
		HBA2_2	CN1_HBA2_2	CN1_HBA_2_2_ASE43_230_2B	Local	Site 1 VSP G900	2B	70
		HBA1_1	CN1_HBA1_1	CN1_HBA_1_1_ASE43_236_5A	Remote	Site 2 VSP G900	5A	69
		HBA1_2	CN1_HBA1_2	CN1_HBA_1_2_ASE43_236_6A	Remote	Site 2 VSP G900	6A	70
		HBA2_1	CN1_HBA2_1	CN1_HBA_2_1_ASE43_236_5B	Remote	Site 2 VSP G900	5B	69
		HBA2_2	CN1_HBA2_2	CN1_HBA_2_2_ASE43_236_6B	Remote	Site 2 VSP G900	6B	70

TABLE 10. SAN HBA CONNECTION CONFIGURATION BETWEEN DS220 AND VSP G900, DS120, AND VSP G350 (CONTINUED)

Site	*Server*	*HBA Ports*	*Storage Host Group*	*Switch Zone*	*Connec-tion*	*Storage System*	*Port*	*Brocade G620 Switch*
Site 1	DS220 Server 2	HBA1_1	CN2_HBA1_1	CN2_HBA1_1_ ASE_43_230_3A	Local	Site 1 VSP G900	3A	69
		HBA1_2	CN2_HBA1_2	CN2_HBA1_2_ ASE_43_230_4A	Local	Site 1 VSP G900	4A	70
		HBA2_1	CN2_HBA2_1	CN2_HBA2_1_ ASE_43_230_3B	Local	Site 1 VSP G900	3B	69
		HBA2_2	CN2_HBA2_2	CN2_HBA2_2_ ASE_43_230_4B	Local	Site 1 VSP G900	4B	70
		HBA1_1	CN2_HBA1_1	CN2_HBA1_1_ ASE_43_236_7A	Remote	Site 2 VSP G900	7A	69
		HBA1_2	CN2_HBA1_2	CN2_HBA1_2_ ASE_43_236_8A	Remote	Site 2 VSP G900	8A	70
		HBA2_1	CN2_HBA2_1	CN2_HBA2_1_ ASE_43_236_7B	Remote	Site 2 VSP G900	7B	69
		HBA2_2	CN2_HBA2_2	CN2_HBA2_2_ ASE_43_236_8B	Remote	Site 2 VSP G900	8B	70

TABLE 10. SAN HBA CONNECTION CONFIGURATION BETWEEN DS220 AND VSP G900, DS120, AND VSP G350 (CONTINUED)

Site	*Server*	*HBA Ports*	*Storage Host Group*	*Switch Zone*	*Connec-tion*	*Storage System*	*Port*	*Brocade G620 Switch*
Site 2	DS220 Server 3	HBA1_1	CN1_HBA1_1	CN1_HBA_1_1_ASE43_236_1A	Local	Site 2 VSP G900	1A	69
		HBA1_2	CN1_HBA1_2	CN1_HBA_1_2_ASE43_236_2A	Local	Site 2 VSP G900	2A	70
		HBA2_1	CN1_HBA2_1	CN1_HBA_2_1_ASE43_236_1B	Local	Site 2 VSP G900	1B	69
		HBA2_2	CN1_HBA2_2	CN1_HBA_2_2_ASE43_236_2B	Local	Site 2 VSP G900	2B	70
		HBA1_1	CN1_HBA1_1	CN1_HBA_1_1_ASE43_230_5A	Remote	Site 1 VSP G900	5A	69
		HBA1_2	CN1_HBA1_2	CN1_HBA_1_2_ASE43_230_6A	Remote	Site 1 VSP G900	6A	70
		HBA2_1	CN1_HBA2_1	CN1_HBA_2_1_ASE43_230_5B	Remote	Site 1 VSP G900	5B	69
		HBA2_2	CN1_HBA2_2	CN1_HBA_2_2_ASE43_230_6B	Remote	Site 1 VSP G900	6B	70

TABLE 10. SAN HBA CONNECTION CONFIGURATION BETWEEN DS220 AND VSP G900, DS120, AND VSP G350 (CONTINUED)

Site	*Server*	*HBA Ports*	*Storage Host Group*	*Switch Zone*	*Connec-tion*	*Storage System*	*Port*	*Brocade G620 Switch*
Site 2	DS220 Server 4	HBA1_1	CN2_HBA1_1	CN2_HBA1_1_ ASE_43_236_3A	Local	Site 2 VSP G900	3A	69
		HBA1_2	CN2_HBA1_2	CN2_HBA1_2_ ASE_43_236_4A	Local	Site 2 VSP G900	4A	70
		HBA2_1	CN2_HBA2_1	CN2_HBA2_1_ ASE_43_236_3B	Local	Site 2 VSP G900	3B	69
		HBA2_2	CN2_HBA2_2	CN2_HBA2_2_ ASE_43_236_4B	Local	Site 2 VSP G900	4B	70
		HBA1_1	CN2_HBA1_1	CN2_HBA1_1_ ASE_43_230_7A	Remote	Site 1 VSP G900	7A	69
		HBA1_2	CN2_HBA1_2	CN2_HBA1_2_ ASE_43_230_8A	Remote	Site 1 VSP G900	8A	70
		HBA2_1	CN2_HBA2_1	CN2_HBA2_1_ ASE_43_230_7B	Remote	Site 1 VSP G900	7B	69
		HBA2_2	CN2_HBA2_2	CN2_HBA2_2_ ASE_43_230_8B	Remote	Site 1 VSP G900	8B	70
Site 3	DS120 Server 1	HBA1_1	MN1_HBA1_1	MN1_HBA1_1_ ASE_43_240_1C	Local	Site 3 VSP G350	1C	67
		HBA1_2	MN1_HBA1_2	MN1_HBA1_2_ ASE_43_240_2C	Local	Site 3 VSP G350	2C	68
	DS120 Server 2	HBA2_1	MN1_HBA1_1	MN1_HBA1_1_ ASE_43_240_1D	Local	Site 3 VSP G350	1D	67
		HBA2_2	MN1_HBA1_2	MN1_HBA1_2_ ASE_43_240_2D	Local	Site 3 VSP G350	2D	68

SAN Switch Architecture between two VSP G900 Storage Systems.

TABLE 11. SAN SWITCH ARCHITECTURE BETWEEN VSP G900 STORAGE SYSTEMS

Storage System	*Storage Ports*	*Switch Zone*	*Storage System*	*Storage Port*	*Purpose*
Site 1 VSP G900	7B	ASE_43_230_7B_ ASE_43_236_7B	Site 2 VSP G900	7B	Replication link Remote Connection
	8B	ASE_43_230_7B_ ASE_43_236_8B		8B	Replication link Remote Connection

SAN Switch Architecture between Site 1, Site 2 G900 storage systems and Site 3 VSP G350 storage system.

TABLE 12. SWITCH ARCHITECTURE BETWEEN VSP G350 STORAGE SYSTEMS

Storage System	*Storage Ports*	*Switch Zone*	*Storage System*	*Storage Port*	*Purpose*
Site 3 VSP G350	4A	ASE_43_240_4A_ ASE_43_230_4A	Site 1 VSP G900	4A	Quorum Connection
	5A	ASE_43_240_5A_ ASE_43_230_5A	Site 1 VSP G900	5A	Quorum Connection
	6A	ASE_43_240_6A_ ASE_43_236_6A	Site 2 VSP G900	6A	Quorum Connection
	7A	ASE_43_240_7A_ ASE_43_236_7A	Site 2 VSP G900	7A	Quorum Connection

SAN Switch Architecture Between two G900 Storage Systems and ESXi cluster.

TABLE 13. SAN SWITCH ARCHITECTURE BETWEEN VSP G900 STORAGE SYSTEMS AND ESXI CLUSTER

Site	*Server*	*HBA Ports*	*Storage Host Group*	*Switch Zone*	*Storage System*	*Purpose*
Site 1	ESXi Cluster	HBA1_1	MN1_HBA1_1	MN1_HBA1_1_ ASE_43_230_8A	Site 1 VSP G900	Command device
Site 2	ESXi Cluster	HBA1_1	MN2_HBA1_1	MN2_HBA1_1_ ASE_43_236_8A	Site 2 VSP G900	Command device

Note — In a production environment, it is recommended to use separate storage ports for the management servers and quorum disks to avoid impact on the database performance. Shared storage ports can be used; however, port utilization should be monitored to avoid performance issues on 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 Network** — This network provides client connections to the applications and Oracle Real Application Clusters.
- BMC/management network — The Baseboard Management Controller (BMC) provides remote management capabilities including console redirection, logging, and power control.

Hitachi Vantara recommends using pairs of 25 Gbps 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.

Table 14 shows the network configuration, and Table 15 shows the virtual IP address and SCAN name configuration used when testing the environment. Your values may be different.

When creating NIC Bonding pairs, ports should be used on different cards to avoid single point of failure (SPoF). It is recommended that BMC connections go to a separate switch on the management network.

TABLE 14. NETWORK CONFIGURATION

Server	*NIC Ports*	*VLAN/ Subnet*	*NIC BOND*	*IP Address*	*Network*	*Bandwidth (Gbps)*	*Cisco Nexus 93180YC-EX Switch*	
							Switch Number	*Port*
DS220 Server1	NIC - 0	208	Bond0	192.168.208.xx	Private	25	1	41
	NIC - 2					25	2	
	NIC - 1	242	Bond1	192.168.242.xx	Public Oracle	25	1	42
	NIC - 3					25	2	
	BMC-Dedicated NIC	244	-	192.168.244.xx	Public Management	1	-	
DS220 Server2	NIC - 0	208	Bond0	192.168.208.xx	Private	25	1	43
	NIC - 2					25	2	
	NIC - 1	242	Bond1	192.168.242.xx	Public Oracle	25	1	44
	NIC - 3					25	2	
	BMC-Dedicated NIC	244	-	192.168.244.xx	Public Management	1	-	
DS220 Server3	NIC - 0	208	Bond0	192.168.208.xx	Private	25	1	45
	NIC - 2					25	2	
	NIC - 1	242	Bond1	192.168.242.xx	Public Oracle	25	1	46
	NIC - 3					25	2	
	BMC-Dedicated NIC	244	-	192.168.244.xx	Public Management	1	-	

TABLE 14. NETWORK CONFIGURATION (CONTINUED)

Server	*NIC Ports*	*VLAN/ Subnet*	*NIC BOND*	*IP Address*	*Network*	*Bandwidth (Gbps)*	*Cisco Nexus 93180YC-EX Switch*	
							Switch Number	*Port*
DS220 Server4	NIC - 0	208	Bond0	192.168.208.xx	Private	25	1	47
	NIC - 2					25	2	
	NIC - 1	242	Bond1	192.168.242.xx	Public Oracle	25	1	48
	NIC - 3					25	2	
	BMC-Dedicated NIC	244	-	192.168.244.xx	Public Management	1	-	
DS120 management server1	NIC - 0	242	-	192.168.242.xx	Public	25	1	49
	BMC-Dedicated NIC	244	-	192.168.244.xx	Public Management	1	-	
DS120 management server2	NIC - 0	242	-	192.168.242.xx	Public	25	1	50
	BMC-Dedicated NIC	244	-	192.168.244.xx	Public Management	1		

TABLE 15. VIRTUAL IP AND SCAN NAME CONFIGURATION

Server	Virtual IP	Scan Name - hitachi-cluster-scan
Database Server 1 (DS220 1)	192.168.242.xx	192.168.242.xx 192.168.242.xx 192.168.242.xx
Database Server 1 (DS220 2)	192.168.242.xx	
Database Server 3 (DS220 3)	192.168.242.xx	
Database Server 4 (DS220 4)	192.168.242.xx	

Table 16 lists virtual machine configuration running on the management server cluster.

TABLE 16. MANAGEMENT SERVER VIRTUAL MACHINES CONFIGURATION

Virtual Machine	vCPU	Virtual Memory	Disk capacity	IP Address	OS
HDID	2	8 GB	300 GB	192.168.242.xx	Windows Server 2012 R2 Standard
SwinchBench	2	8 GB	100 GB	192.168.242.xx	Windows Server 2012 R2 Standard
vCenter	2	10 GB	300 GB	192.168.242.xx	VMware Photon Linux 1.0
OEM	16	32 GB	200 GB	192.168.242.xx	RHEL 7.6
Oracle Adapters	2	6 GB	40 - 50 GB	192.168.242.xx	OL 7.3
Oracle VM Manager	2	10 GB	100 GB	192.168.242.xx	OL 7.3/7.4
HAS	4	16 GB	100 GB	192.168.242.xx	CentOS 7.2
HIAA/HDCA	4	32 GB	800 GB	192.168.242.xx	RHEL 7.3
HDCA Probe	4	10 GB	110 GB	192.168.242.xx	RHEL 7.3
vSVP - Storage Virtual Platform G900	2	32	120 GB	192.168.167.xx	Microsoft Windows 7 (64 bit)

Global-active Device Setup Pre-configuration

Before setup of global-active device configuration using HDID, manual pre-configuration steps must be performed on Site 1 and Site 2 VSP G900 storage. Manual pre-configuration includes creation of pools, VVols, host groups, zone configuration, multipathing configuration and the addition of the quorum disk from the quorum site on both of the VSP G900.

Table 17 shows the manual pre-configuration needed on Site 1 and Site 2 VSP G900 storage before setting up global-active device using HDID.

TABLE 17. MANUAL PRE-CONFIGURATION ON VSP G900

Site	*Site 1*		*Site 2*	
Pool Name	Hitachi-UCP-CI-OS	Hitachi-UCP-CI-Oracle	Hitachi-UCP-CI-OS	Hitachi-UCP-CI-Oracle
VVOLs	2 for OS	115 PVOLs for Oracle database	2 for OS	- (global-active device SVOLs will be created using HDID)
Host Groups	16 Host Groups		16 Host Groups	

TABLE 17. MANUAL PRE-CONFIGURATION ON VSP G900 (CONTINUED)

Zone configuration	16 Zones	16 Zones
Quorum disk	Map external volume, and specify quorum disk ID	Map external volume and specify quorum disk ID
Multipathing	4 owner paths per server Oracle database PVOLs	-
	(4 non-owner paths will be added after global-active device setup, total 8 paths per server Oracle database PVOLs) Please refer to Table 10, “SAN HBA Connection Configuration Between DS220 and VSP G900, DS120, and VSP G350,” on page 20 to get details of owner (local) and non-owner (remote) paths.	(4 owner and 4 non-owner paths will be added after global-active device setup, total 8 paths per server Oracle database SVOLs) Please refer to Table 10, “SAN HBA Connection Configuration Between DS220 and VSP G900, DS120, and VSP G350,” on page 20 to get details of owner (local) and non-owner (remote) paths.

Solution Implementation

Deploy the Solution

Implementing this solution requires doing the following high-level procedures:

1. Install the Hitachi Data Instance Director Master Node
2. Create the Hitachi Data Instance Director Nodes
3. Define the Hitachi Data Instance Director Policy
4. Define Hitachi Data Instance Director Data Flow

Your checklist might vary based on your environment.

Deploy and Configure Hitachi Data Instance Director

This includes steps to deploy and configure Hitachi Data Instance. To deploy Hitachi Data Instance Director in this solution, do these procedures.

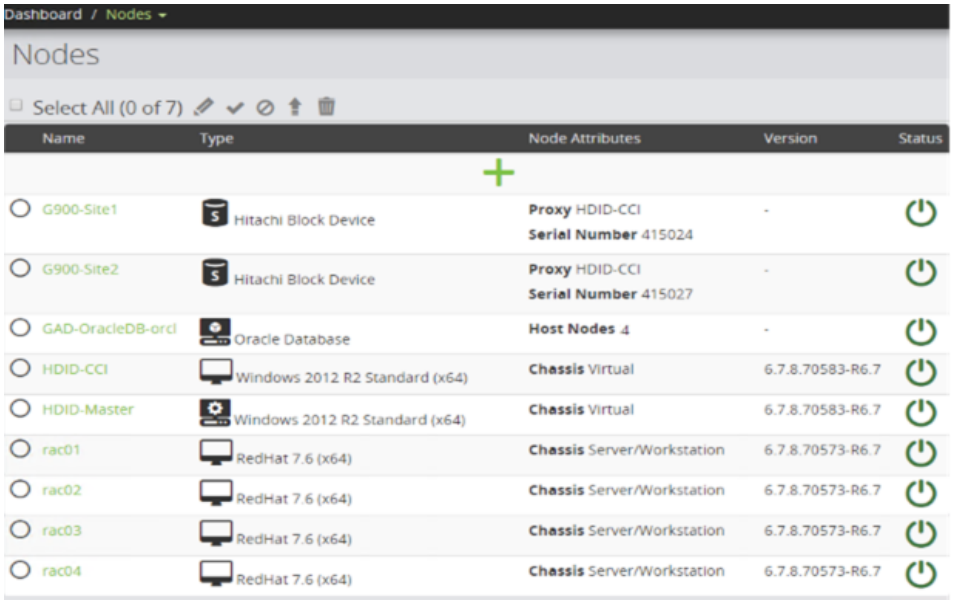
Install the Hitachi Data Instance Director Master Node

The HDID Installation Guide, User Guide, and other documentation are in the documentation folder and the HDID installable is located in the Linux/Windows folder of the HDID ISO image. Download the latest version of the media kit to get the HDID ISO image. Please visit https://knowledge.hitachivantara.com/ to get access to this content or contact your local Hitachi Vantara representative.

Create the Hitachi Data Instance Director Nodes

Login to HDID web console using https://HDID-Master URL with administrator@master user and local administrator password and add nodes into the HDID master. Figure 4 shows details of the nodes added.

Figure 4



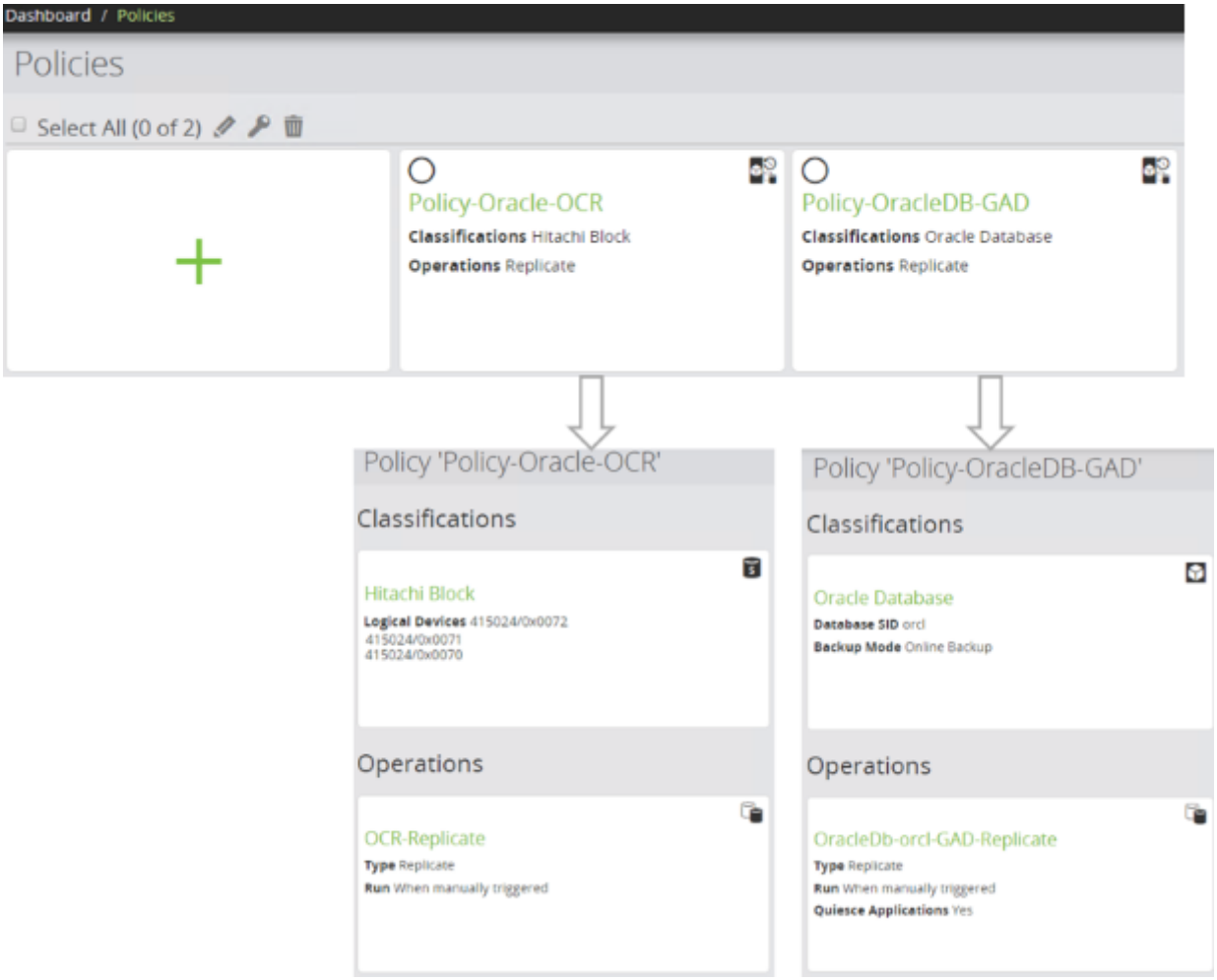
Define the Hitachi Data Instance Director Policy

A policy defines data classifications and data operations.

1. Policy-OracleDB-GAD was created to global-active device replicate Oracle database VVols. While creating a policy select the appropriate Oracle database from the added Oracle RAC nodes. In this case rac01, rac02, rac03, rac04 are the added Oracle RAC nodes and orcl is the Oracle database SID.
2. Policy-Oracle-OCR was created to global-active device replicate OCR VVols. Users need to specify the OCR VVOLs in the Serial/LDEV_ID format. HDID does not global-active device replicate OCR VVols as a part of Oracle database VVols.

Figure 5 shows the complete HDID policy details.

Figure 5


Dashboard / Policies
Policies
Select All (0 of 2)
Policy-Oracle-OCR
Classifications Hitachi Block
Operations Replicate
Policy-OracleDB-GAD
Classifications Oracle Database
Operations Replicate
Policy 'Policy-Oracle-OCR'
Classifications
Hitachi Block
Logical Devices 415024/0x0072
415024/0x0071
415024/0x0070
Operations
OCR-Replicate
Type Replicate
Run When manually triggered
Policy 'Policy-OracleDB-GAD'
Classifications
Oracle Database
Database SID orcl
Backup Mode Online Backup
Operations
OracleDb-orcl-GAD-Replicate
Type Replicate
Run When manually triggered
Quiesce Applications Yes

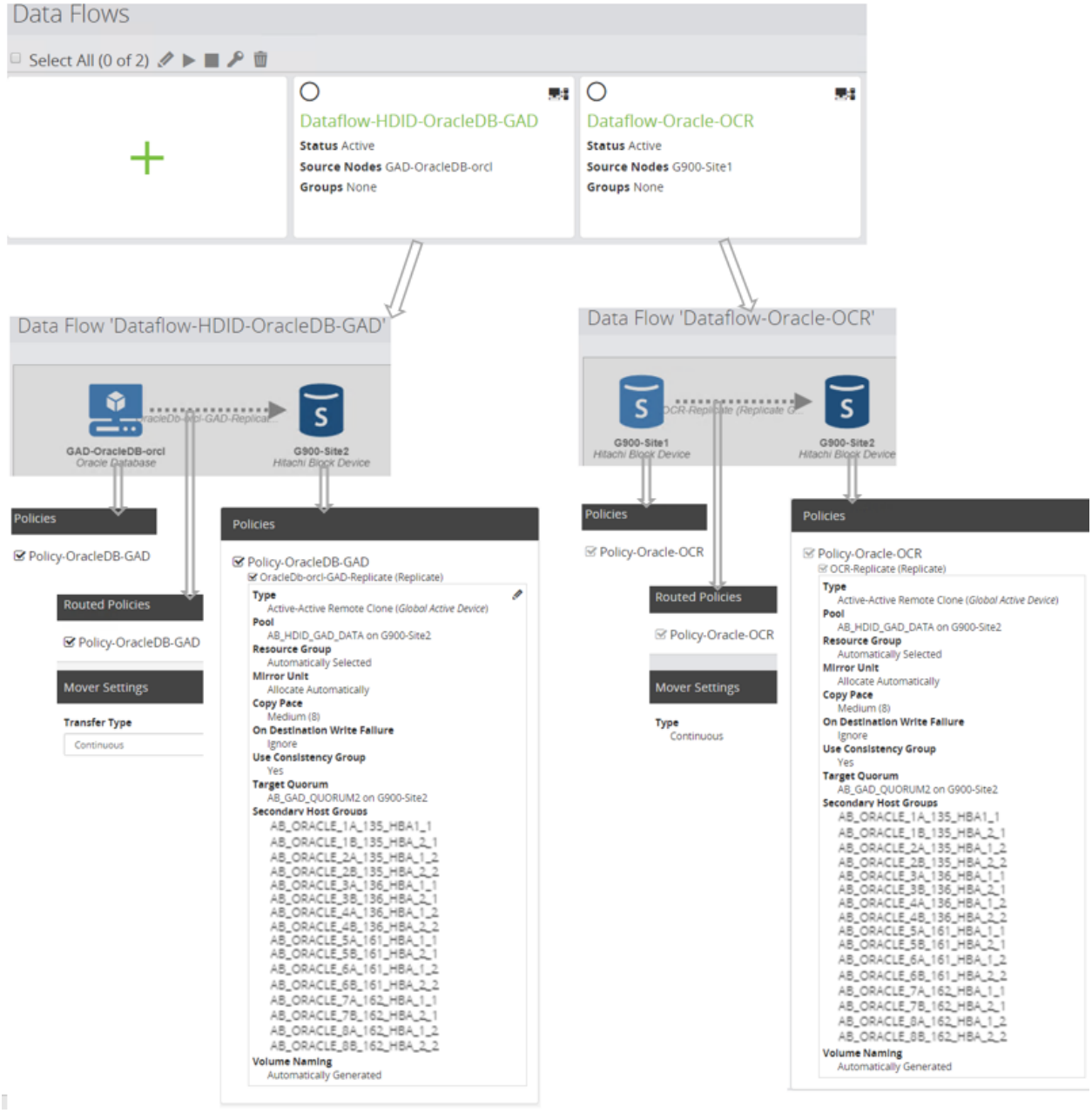
Define Hitachi Data Instance Director Data Flow

1. Dataflow-HDID-OracleDB-GAD:
   (1) Dataflow-HDID-OracleDB-GAD was created to global-active device replicate Oracle database VVols.
   (2) Select node GAD-OracleDB-orcl and assign the Policy-OracleDB-GAD policy.
   (3) Select replication mover type to **Continuous**.
   (4) For the G900-Site2 'Select Creation Mode' as Configure New Replication and replication type as 'Active-Active Remote Clone' on the next page and click on **Next**.
   (5) Select the target pool where global-active device replicated VVols will be created.
   (6) Select the Quorum disk from the available quorum disks. In this environment AB_GAD_QUORUM2 was used as the quorum disk and click on **Next**.
   (7) Select resource group as Automatically Selected and click on **Next**.
   (8) Select secondary host groups where VVols will be mapped and click on **Next**.
   (9) Match Origin option is selected for secondary LDEVs naming options.
2. Dataflow-Oracle-OCR:
   (1) Dataflow-Oracle-OCR was created to global-active device replicate OCR VVols.
   (2) Select G900-Site1 and assign the Policy-Oracle-OCR policy.
   (3) Select replication mover type to **Continuous**.
   (4) For the G900-Site2 configure **Select Creation Mode** as **Configure New Replication** and replication type as **Active-Active Remote Clone** on the next page and click on **Next**.
   (5) Select the target pool where global-active device replicated VVols will be created.
   (6) Select the Quorum disk from the available quorum disks and click on **Next**.

   In this environment AB_GAD_QUORUM2 was used as the quorum disk
   (7) Select resource group as Automatically Selected and click on **Next**.
   (8) Select secondary host groups where VVOLs will be mapped and click on **Next**.
   (9) The **Match Origin** option is selected for secondary LDEVs naming options.

Figure 6 shows the complete HDID data flow details.

Figure 6



Solution Execution

Execution of this solution consists of the following procedures:

- Perform global-active device replication for the Oracle database and OCR disks to the secondary VSP G900 storage using HDID.
- Recover Oracle Database After Storage Replication Link Failure between site 1 and site 2 storage systems.
- Perform Storage Replication operations using HDID.

Perform Global-active Device Replication for the Oracle Database and OCR Disks to the Secondary VSP G900 Storage

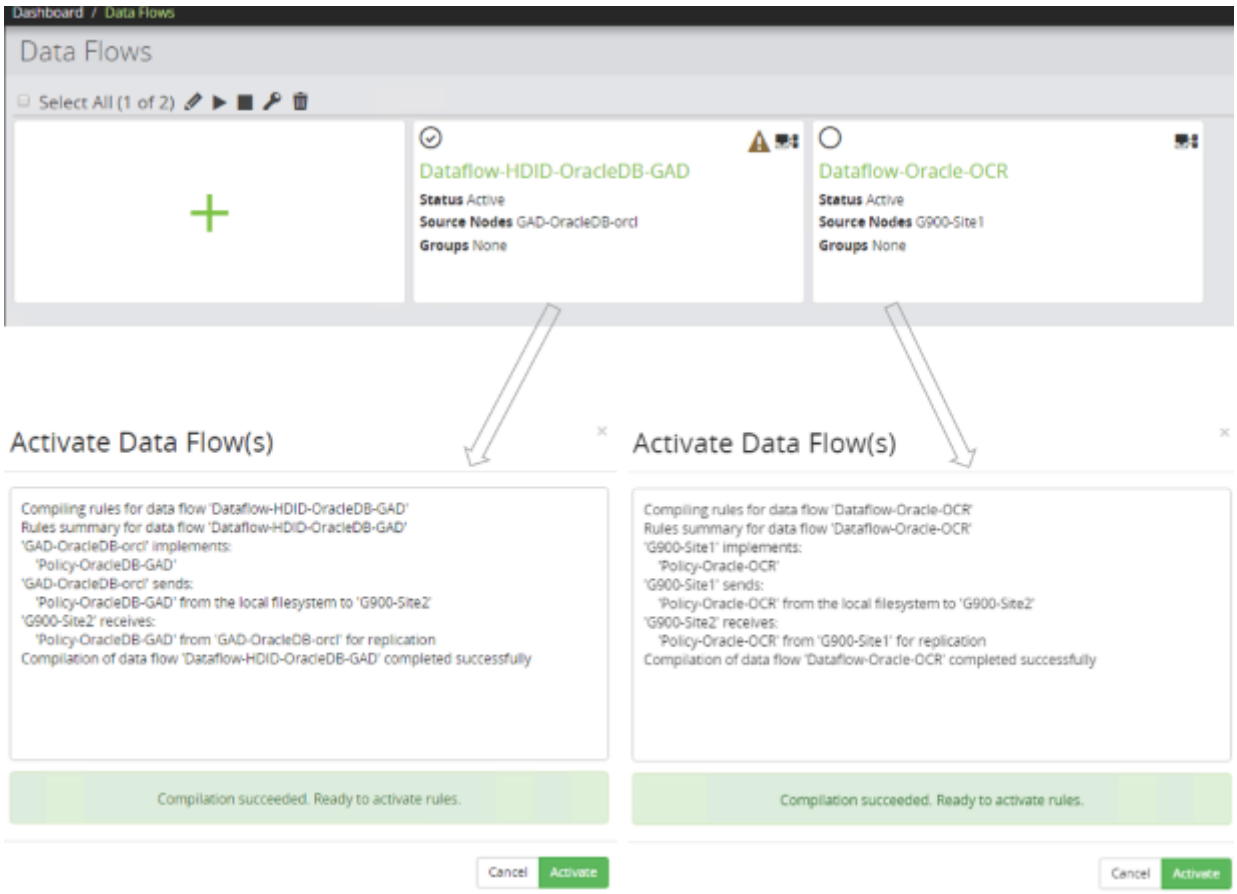
This is how to perform global-active device replication for the Oracle database and OCR disks to the secondary storage using Hitachi Data Instance Director.

Activate Hitachi Data Instance Director Data Flow

To execute the HDID data flow, **Activate** Hitachi Data Instance Director Data flow.

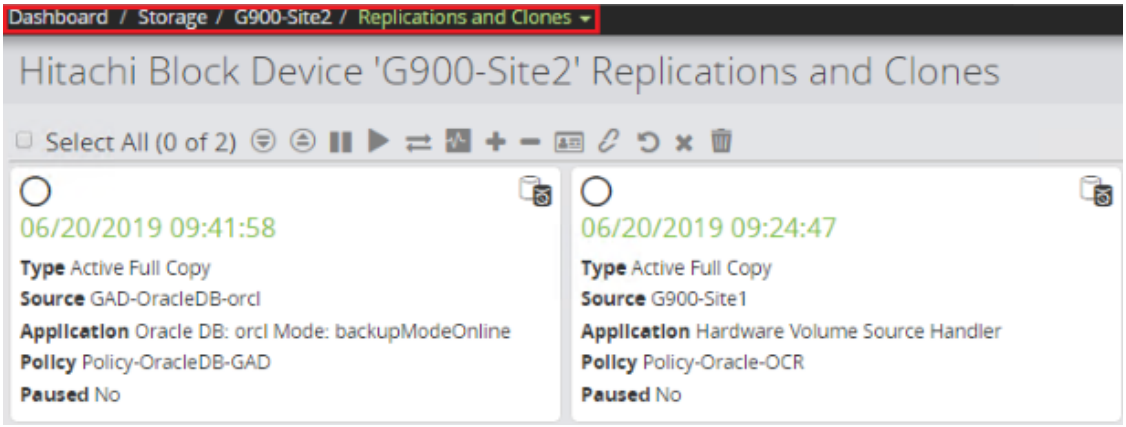
1. Select the appropriate Data Flow and click on the ▶ **(Activate)** button. The **Activate Data Flow(s)** dialog box displays with data flow compilation details.
2. Then click on the **Activate** button to execute the data flow.

Figure 7



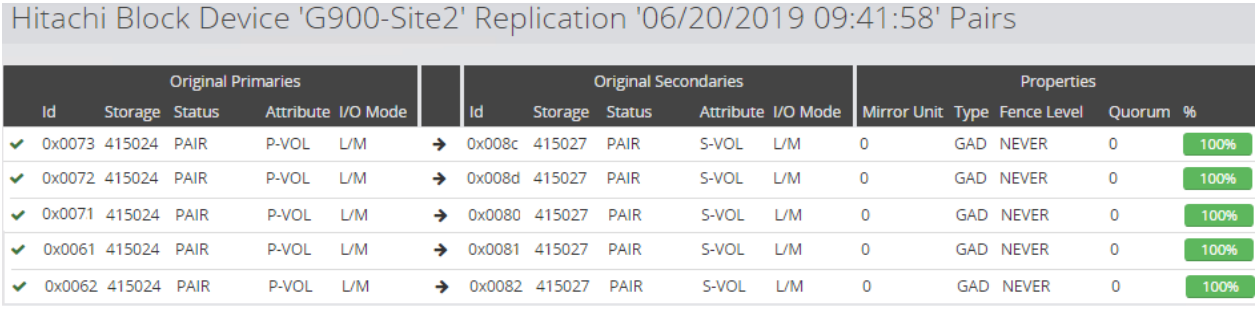
3. On the **Monitor** menu the user can monitor the HDID data flow progress operation.
4. After the HDID dataflow activation, source Oracle and OCR PVOLs will be tiered to the secondary VSP G900 storage.
5. Users can see global-active device pairs using HDID. Click on the 'Dashboard → Storage → G900-Site2 → Replication and Clones' to see the replications on the Site 2 VSP G900 storage.

Figure 8



6. Click on any of the replication pairs to see the global-active device pairing and progress details.

Figure 9



Original Primaries						Original Secondaries					Properties				
Id	Storage	Status	Attribute	I/O Mode		Id	Storage	Status	Attribute	I/O Mode	Mirror Unit	Type	Fence Level	Quorum	%
0x0073	415024	PAIR	P-VOL	L/M	→	0x008c	415027	PAIR	S-VOL	L/M	0	GAD	NEVER	0	100%
0x0072	415024	PAIR	P-VOL	L/M	→	0x008d	415027	PAIR	S-VOL	L/M	0	GAD	NEVER	0	100%
0x0071	415024	PAIR	P-VOL	L/M	→	0x0080	415027	PAIR	S-VOL	L/M	0	GAD	NEVER	0	100%
0x0061	415024	PAIR	P-VOL	L/M	→	0x0081	415027	PAIR	S-VOL	L/M	0	GAD	NEVER	0	100%
0x0062	415024	PAIR	P-VOL	L/M	→	0x0082	415027	PAIR	S-VOL	L/M	0	GAD	NEVER	0	100%

Recover Oracle Database After Storage Replication Link Failure Between Site 1 and Site 2 Storage Systems

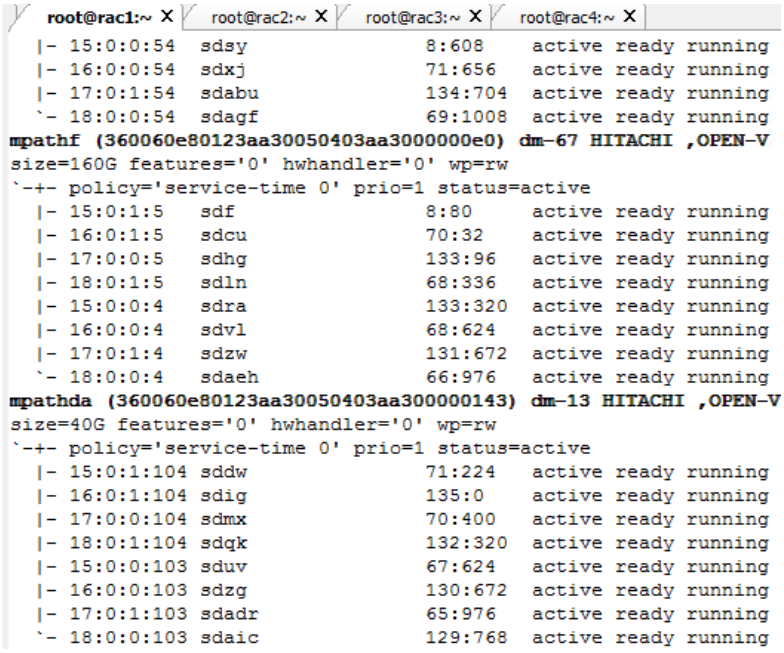
Objective for Use Case: Recover from storage replication link failure between site 1 and site 2 storage systems.

This was the procedure to evaluate the Use Case.

1. System Status Checks

(1) Verified that all paths were 'active ready running' for all Oracle RAC nodes.

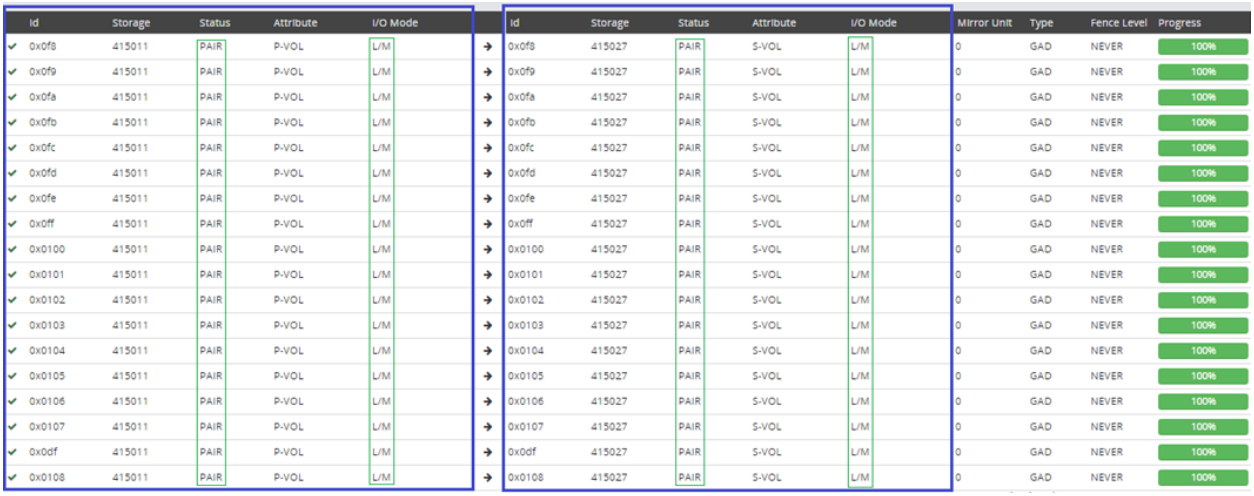
Figure 10



```
root@rac1:~ X | root@rac2:~ X | root@rac3:~ X | root@rac4:~ X
  |- 15:0:0:54  sdsy                8:608    active ready running
  |- 16:0:0:54  sdxj                71:656   active ready running
  |- 17:0:1:54  sdabu               134:704  active ready running
  `- 18:0:0:54  sdagf               69:1008  active ready running
mpathf (360060e80123aa30050403aa3000000e0) dm-67 HITACHI ,OPEN-V
size=160G features='0' hwhandler='0' wp=rw
`-+- policy='service-time 0' prio=1 status=active
  |- 15:0:1:5   sdf                 8:80     active ready running
  |- 16:0:1:5   sdcu                70:32    active ready running
  |- 17:0:0:5   sdhg                133:96   active ready running
  |- 18:0:1:5   sdln                68:336   active ready running
  |- 15:0:0:4   sdra                133:320  active ready running
  |- 16:0:0:4   sdvl                68:624   active ready running
  |- 17:0:1:4   sdzw                131:672  active ready running
  `- 18:0:0:4   sdaeh               66:976   active ready running
mpathda (360060e80123aa30050403aa300000143) dm-13 HITACHI ,OPEN-V
size=40G features='0' hwhandler='0' wp=rw
`-+- policy='service-time 0' prio=1 status=active
  |- 15:0:1:104 sddw                71:224   active ready running
  |- 16:0:1:104 sdig                135:0    active ready running
  |- 17:0:0:104 sdmx                70:400   active ready running
  |- 18:0:1:104 sdqk                132:320  active ready running
  |- 15:0:0:103 sduv                67:624   active ready running
  |- 16:0:0:103 sdzg                130:672  active ready running
  |- 17:0:1:103 sdadr               65:976   active ready running
  `- 18:0:0:103 sdaic               129:768  active ready running
```

(2) Verified that all the VVOLs pairs were in PAIR status using **HDID**.

Figure 11



Id	Storage	Status	Attribute	I/O Mode	Id	Storage	Status	Attribute	I/O Mode	Mirror Unit	Type	Fence Level	Progress
0x0f8	415011	PAIR	P-VOL	L/M	0x0f8	415027	PAIR	S-VOL	L/M	0	GAD	NEVER	100%
0x0f9	415011	PAIR	P-VOL	L/M	0x0f9	415027	PAIR	S-VOL	L/M	0	GAD	NEVER	100%
0x0fa	415011	PAIR	P-VOL	L/M	0x0fa	415027	PAIR	S-VOL	L/M	0	GAD	NEVER	100%
0x0fb	415011	PAIR	P-VOL	L/M	0x0fb	415027	PAIR	S-VOL	L/M	0	GAD	NEVER	100%
0x0fc	415011	PAIR	P-VOL	L/M	0x0fc	415027	PAIR	S-VOL	L/M	0	GAD	NEVER	100%
0x0fd	415011	PAIR	P-VOL	L/M	0x0fd	415027	PAIR	S-VOL	L/M	0	GAD	NEVER	100%
0x0fe	415011	PAIR	P-VOL	L/M	0x0fe	415027	PAIR	S-VOL	L/M	0	GAD	NEVER	100%
0x0ff	415011	PAIR	P-VOL	L/M	0x0ff	415027	PAIR	S-VOL	L/M	0	GAD	NEVER	100%
0x0100	415011	PAIR	P-VOL	L/M	0x0100	415027	PAIR	S-VOL	L/M	0	GAD	NEVER	100%
0x0101	415011	PAIR	P-VOL	L/M	0x0101	415027	PAIR	S-VOL	L/M	0	GAD	NEVER	100%
0x0102	415011	PAIR	P-VOL	L/M	0x0102	415027	PAIR	S-VOL	L/M	0	GAD	NEVER	100%
0x0103	415011	PAIR	P-VOL	L/M	0x0103	415027	PAIR	S-VOL	L/M	0	GAD	NEVER	100%
0x0104	415011	PAIR	P-VOL	L/M	0x0104	415027	PAIR	S-VOL	L/M	0	GAD	NEVER	100%
0x0105	415011	PAIR	P-VOL	L/M	0x0105	415027	PAIR	S-VOL	L/M	0	GAD	NEVER	100%
0x0106	415011	PAIR	P-VOL	L/M	0x0106	415027	PAIR	S-VOL	L/M	0	GAD	NEVER	100%
0x0107	415011	PAIR	P-VOL	L/M	0x0107	415027	PAIR	S-VOL	L/M	0	GAD	NEVER	100%
0x0df	415011	PAIR	P-VOL	L/M	0x0df	415027	PAIR	S-VOL	L/M	0	GAD	NEVER	100%
0x0108	415011	PAIR	P-VOL	L/M	0x0108	415027	PAIR	S-VOL	L/M	0	GAD	NEVER	100%

(3) Verified that database resource was open and stable.

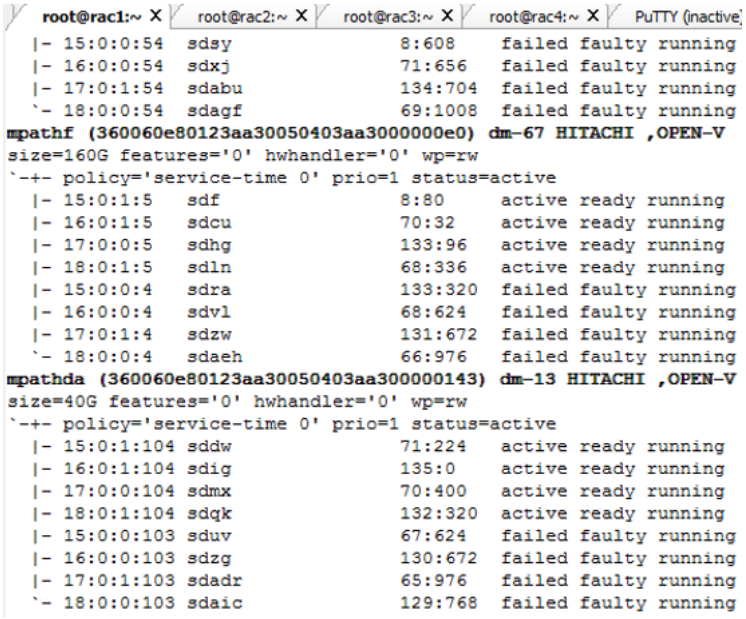
(4) Started Swingbench workload with the number of users configured as 20 on all the RAC nodes.

2. **Simulate Failure:** Disabled all ports used for remote connection (replication links) at site 1.

3. **Behavior after Failure**

(1) At site 1 Oracle RAC hosts, observed that the path status was 'failed faulty running' for non-owner paths and the status is online for owner paths. At site 2 Oracle RAC hosts, observed that the path status was offline for owner paths and the status was online for non-owner paths.

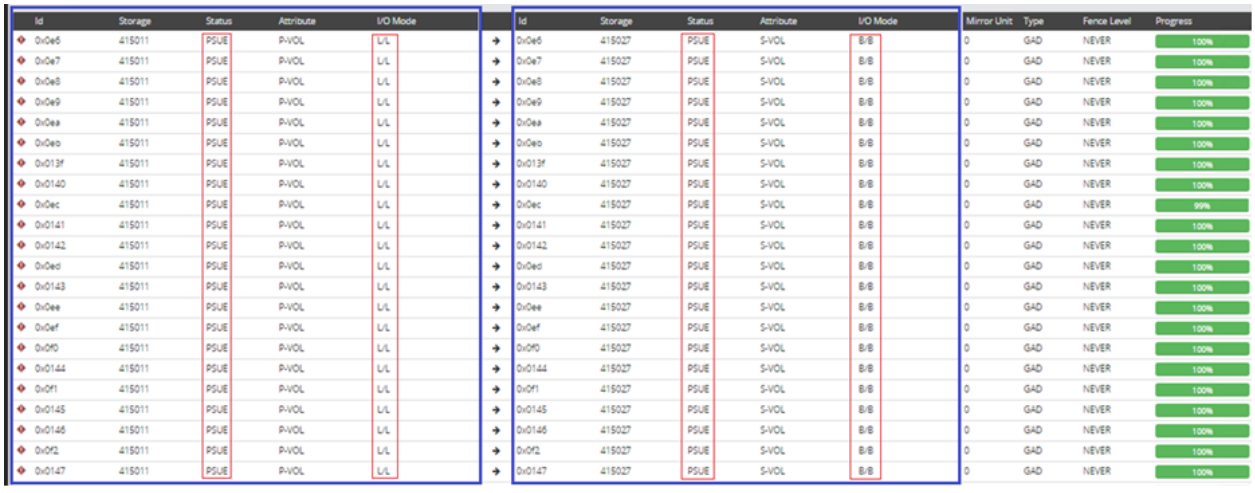
Figure 12



(2) Global-active device pair status after remote replication failure at site 1.

(3) Figure 13 shows that by using **HDID**, users can see VVols pair status in 'PSUE' state after a replication link failure between two storage sites.

Figure 13



(4) Verified that all the database instances were in an online state.

(5) Verified that number of Swingbench user connections was 20.

(6) Checked the database for errors. There were no errors in the alert logs.

4. Recovery Procedure using HDID

(1) At site 1, enabled the Fibre Channel switch ports used for remote connections. Users need to resolve hardware issues before recovery replication using HDID.

(2) Click on Dashboard → Monitor → Dataflow-HDID-OracleDB-GAD.

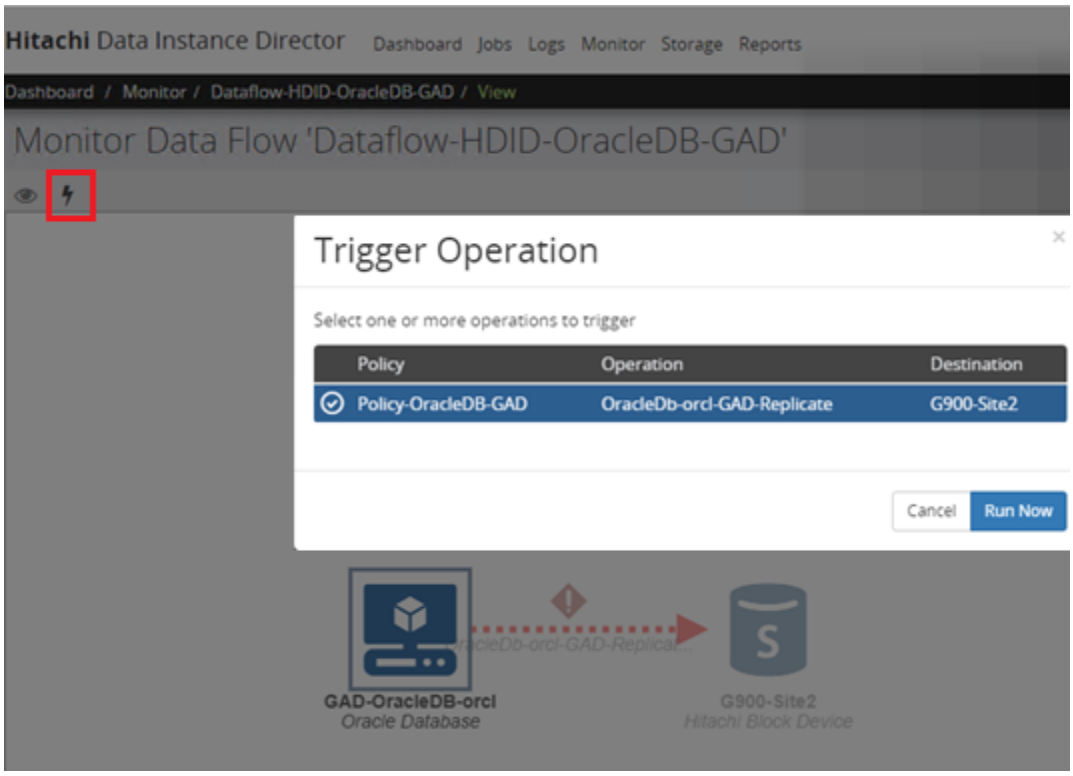
(3) Select GAD-OracleDB-orcl node.

(4) Click on **Trigger.**

(5) Select the policy **Policy-OracleDB-GAD** on the next screen and click **Run Now** to trigger the replication.

This will bring global-active device replication into **PAIR** state from the **PSUE** state. Figure 14 shows how to start the trigger operation on the **Monitor** screen.

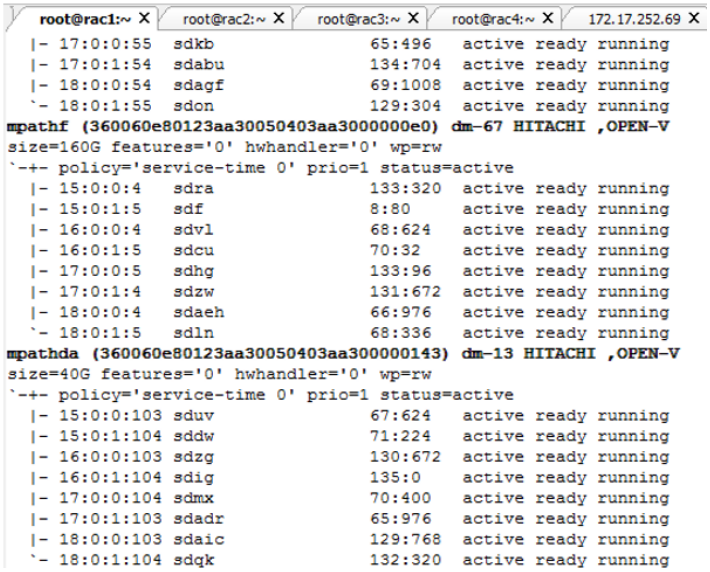
Figure 14



Note — For global-active device pairs in the PSUE error state or PSUS suspend state, users need to resolve the hardware issue first in the storage side and perform a 'Trigger Operation' using the HDID monitor screen which brings global-active device replication into the PAIR state from the PSUE error state or the PSUS suspend state.

(1) Observed path status on all the Oracle RAC hosts were 'active ready running'.

Figure 15

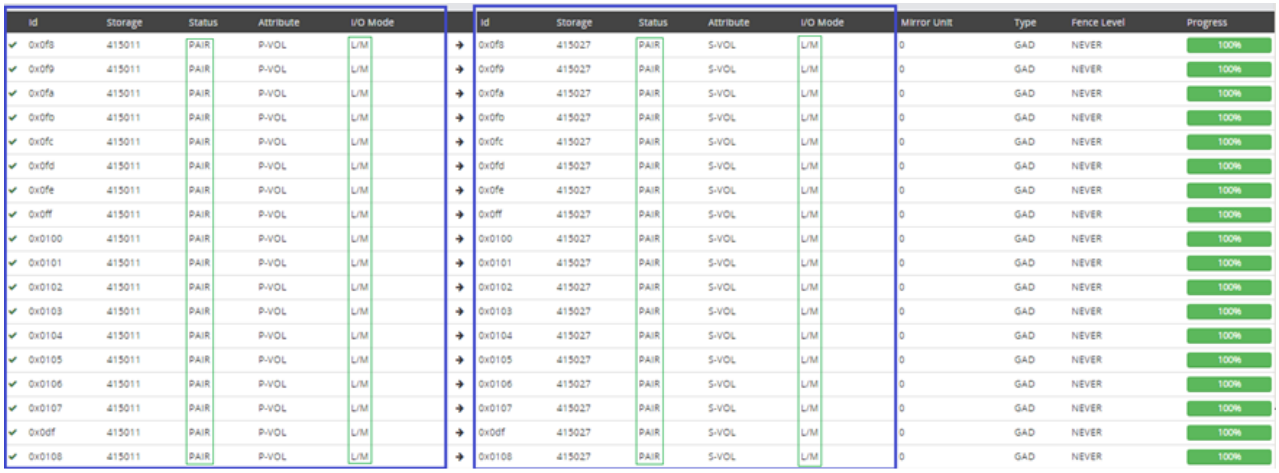


(2) Observed that all the instances were online.

(3) Verified that Swingbench was sending I/O to all the instances without any errors. The graphical user interface and Swingbench output logs showed no errors.

(4) Verified that all the VVol pairs were in PAIR status using **HDID**.

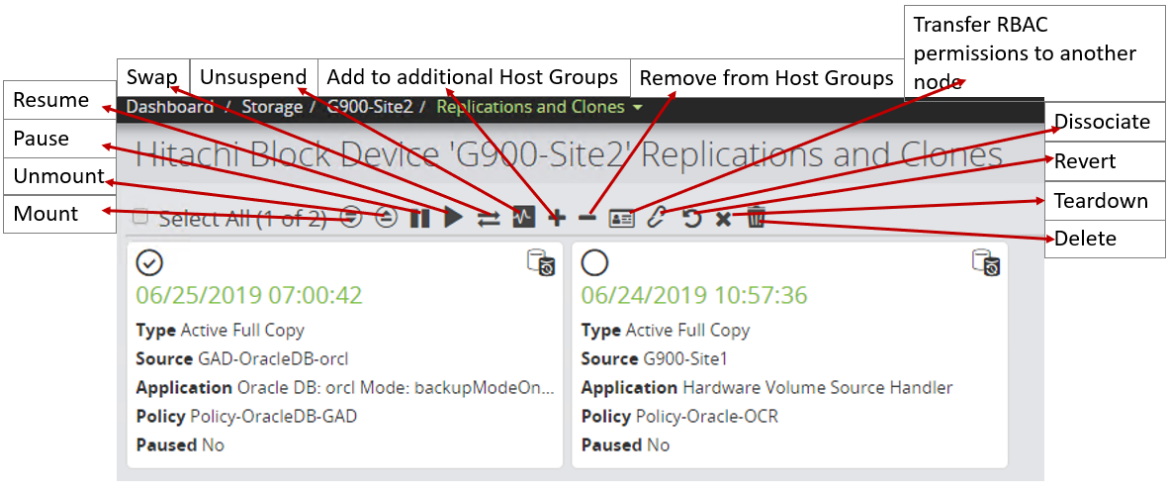
Figure 16



Perform Storage Replication Operations using HDID

HDID provides different options to manage the replications stored on a Block Storage node. Figure 17 shows all the options available with HDID to manage the replication.

Figure 17



This list provides details of options used for the Replications stored on a Block Storage node.

- **Mount:** Used to mount replication to operating system and add to host groups
- **Unmount:** Used to unmount replication from the operating system and delete from host groups
- **Pause:** Pauses the Replication**.** If the replication is live, then it can be paused.
- **Resume:** Resumes a paused Replication
- **Swap: Swaps** direction of GlobalActive Device, TrueCopy, and Universal Replicator
- **Unsuspend:** If a Swap operation cannot be completed due to a P-VOL or data link fault between the primary and secondary device, then the replication will enter the SSWS state (suspended for swapping) indicating that the swap is not yet complete. Unsuspend enables the replication process to be re-established once the cause has been rectified
- **Add to additional Host groups: This** enables LDEVs to be added to host groups in addition to the default HDIDProvisionedHostGroup used by HDID
- **Remove from Host Groups: This** enables LDEVs to be removed from host groups, including the default HDIDProvisionedHostGroup used by HDID
- **Transfer RBAC permissions to another node:** Allows RBAC ownership to be transferred from the current node to another node
- **Dissociate:** Dissociates a replication that was previously adopted by HDID. Removes the selected replication(s) from HDID including state information such as direction and mount location. The replication remains active on the hardware device(s).
- **Revert:** Reverts the replication to perform Oracle recovery operation
- **Teardown:** Tears down a replication using HDID **removes the volume pairings on the array**.
- **Delete:** Deletes the replication record from HDID. The replication is also removed from the block storage device.

Hitachi Block-based 2 datacenter Replication Swapping (Takeover/Takeback) Using HDID

HDID allows users to swap the direction of GlobalActive Device, TrueCopy, and Universal Replicator. When a replication is swapped, the S-VOL takes over the role of the primary volume and the P-VOL takes over the role of the secondary volume. A swapped replication can, of course, be swapped back to its normal state with the P-VOL as the primary and S-VOL as the secondary. A swap operation is typically performed either because maintenance is required, an application failure has occurred, a storage device has failed, or a disaster has occured at the primary site. Failover to the secondary site is therefore necessary.

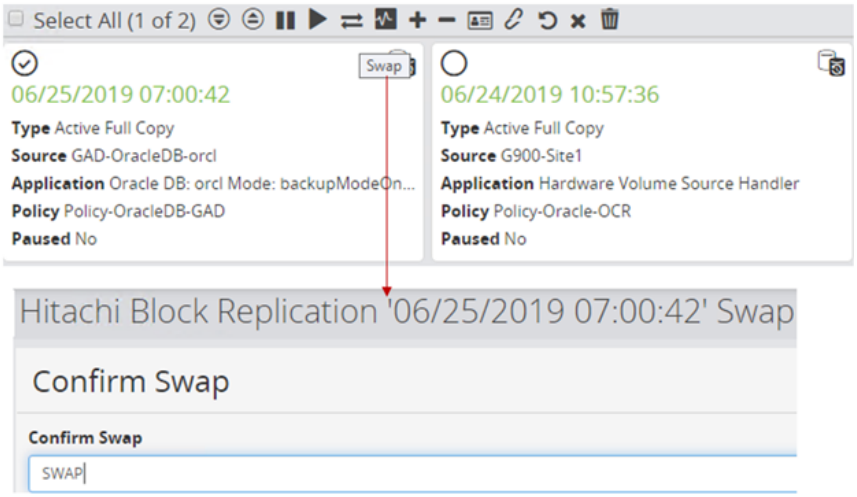
For active-active replications (global-active device):

- A Swap operation may be performed to move array processing load from the primary to the secondary device. If both P-VOL and S-VOL are operable and the link between the two sites is available, the secondary array will assume the higher processing load.
- If the replication cannot be established because the pair has entered an error or suspended state, then once the problem is resolved, the site with the most recent data must be used to re-establish the replication. Because the replication is active-active and cross-path set-ups are possible, depending on the nature of the fault, the P-VOL or S-VOL could contain the most recent data:
  - If the P-VOL contains the most recent data, no swap is required:
    i. If necessary, unsuspend and resume the replication.
    ii. Resynchronize the replication (via manual trigger or data flow reactivation).
  - If the S-VOL contains the most recent data:
    iii. Swap the replication to copy the data from the SVOL to the PVOL.
    iv. Swap the replication again to restore the original direction. This is optional, but highly recommended.
- The swap operation will result in the both P-VOL and the S-VOL remaining writable.

To perform 2 datacenter replication global-active device swap operation

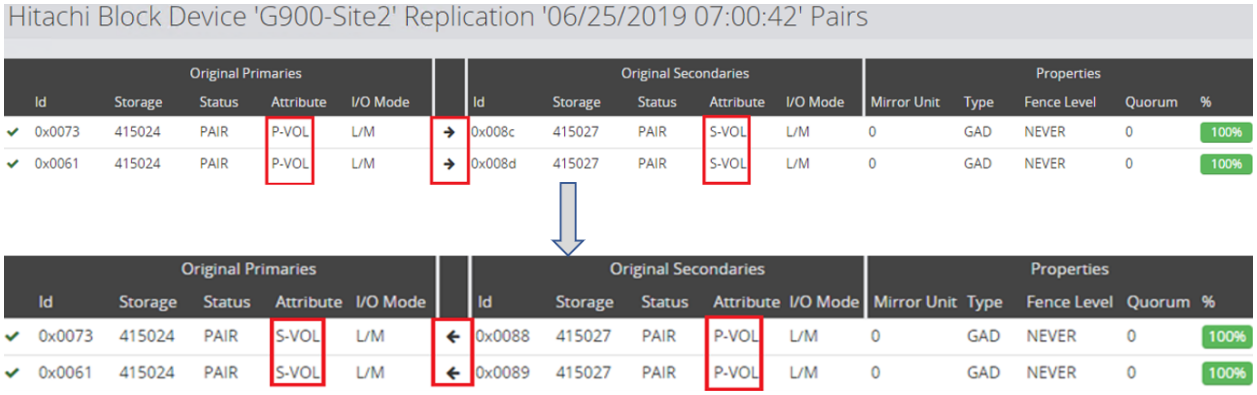
1. Click on the Dashboard → Storage → G900-Site2 → Replication and Clones.
2. Select the required replication to swap.
3. Click on the **Swap option.**
4. On the next screen type `SWAP` to confirm the swap operation.

Figure 18



This figure shows the results of the global-active device replication swap operation using HDID. The S-VOL takes over the role of the primary volume and the P-VOL takes over the role of the secondary volume.

Figure 19



Benefits of using HDID Versus Using Manual Commands for Global-active Device Setup and Configuration

Table 18 shows the benefits of using HDID versus using manual commands for global-active device setup and configuration.

TABLE 18. COMPARISON OF HDID VS MANUAL COMMANDS FOR GLOBAL-ACTIVE DEVICE SETUP

Global-active Device Setup and Configuration	Using Manual Commands	Using HDID
Web-Based UI for Management Flexibility	No	Yes
Quick and Easy automated global-active device setup and recovery of global-active device replication in error or suspended state storage operations	No	Yes

TABLE 18. COMPARISON OF HDID VS MANUAL COMMANDS FOR GLOBAL-ACTIVE DEVICE SETUP

One Click global-active device pair Pause, Resume, Swap, Dissociate, Revert, Teardown, Delete, Re-setup global-active device pair replication option - Less administration	No	Yes
Automated global-active device HORCM configuration and setup. No need of HORCM storage command knowledge to setup and manage global-active device	No	Yes
Informative log messages and monitoring which explains entire process of the global-active device pairing, Pause, Resume, Swap, Teardown, Delete operation	No	Yes
Option to select secondary site storage Pool, Quorum disk and mapping multiple host groups on a single screen	No	Yes

Engineering Validation

This summarizes the key observations from the test results for the Hitachi Unified Compute Platform CI architecture for Oracle Real Application Clusters on Extended Distance clusters in a two-site environment using HDID and global-active device in Hitachi Virtual Storage Platform.

Oracle RAC deployment with Hitachi Virtual Storage Platform G900 and Hitachi Advanced Server DS220.

Test Methodology

The test results were demonstrated using the Swingbench tool.

Swingbench

The workload generation application was Swingbench. Swingbench is a free load generator (and benchmark tool) designed to stress test an Oracle database. Swingbench consists of a load generator, a coordinator, and a cluster overview. The software enables a load to be generated and the transactions/response times to be charted.

Swingbench can be used to demonstrate and test technologies such as Real Application Clusters, Online table rebuilds, Standby databases, Online backup and recovery etc. Please refer to the Swingbench documentation for more information about Swingbench.

Workload Configuration

Testing ran simulated and synthetic workloads using Swingbench. This simulated the workloads for Hitachi Virtual Storage Platform G900 with Storage Virtualization Operating System to test the global-active device.

Test Results

Use Case	Services Impacted After Failure	Services Affected During Recovery	Total Service Downtime
Recover Oracle Database After Storage Replication Link Failure between Site 1 and Site 2 storage systems using HDID	No	No	Zero Time
Recover Oracle Database After Site 2 Storage Failure using HDID	No	No	Zero Time
Recover Oracle Database After Quorum Site Failure using HDID	No	No	Zero Time
Recover Oracle Database After Path Failure Between Servers and Local Storage System using HDID	No	No	Zero Time
Perform global-active device 2 datacenter swap operations from primary to secondary and secondary to primary devices using HDID	No	No	Zero Time
Perform global-active device pair Pause, Resume, Teardown, and Delete operations using HDID	No	No	Zero Time

Hitachi Vantara

Corporate Headquarters
2535 Augustine Drive
Santa Clara, CA 95054 USA
HitachiVantara.com | community.HitachiVantara.com

Contact Information
Phone: 1-800-446-0744
Sales: 1-858-225-2095
HitachiVantara.com/contact

© Hitachi Vantara Corporation, 2019. All rights reserved. HITACHI is a trademark or registered trademark of Hitachi, Ltd. Microsoft and Windows are trademarks or registered trademarks of Microsoft Corporation.All other trademarks, service marks, and company names are properties of their respective owners

Notice: This document is for informational purposes only, and does not set forth any warranty, expressed or implied, concerning any equipment or service offered or to be offered by Hitachi Vantara Corporation.

MK-SL-119-02, August 2019