Technical Report

Sun Cluster with NetApp MetroCluster—Disaster Recovery Solution

Karthikeyan Nagalingam and Suresh Vundru, NetApp

October 2009 | TR-3639

TABLE OF CONTENTS

1 INTRODUCTION

In today's economy, continuous access to information is a critical business need. Enterprises must continue their services through a disaster and need to be able to pick up where they left off before the disaster as quickly and seamlessly as possible. Revenue cannot be generated unless the business is up and running. The NetApp® storage solution is a comprehensive offering that addresses all causes of application downtime—minimizing operator errors, rapidly recovering from operator and application errors, minimizing planned downtime, maximizing system uptime, and recovering from disaster.

NetApp MetroCluster and Sun® Cluster deliver exceptional application and data availability with architectural simplicity and fast data recovery.

1.1 PURPOSE

This technical report documents the installation, configuration, and best practice guidelines for Sun Cluster and NetApp MetroCluster. This document should be treated as a starting reference point. It uses minimum requirements for deploying the solution.

1.2 INTENDED AUDIENCE

This technical report is intended for field personnel, storage architects, and administrators who are responsible for designing and deploying MetroCluster high-availability and disaster-recovery configurations.

This technical report assumes that readers are familiar with Sun Cluster software, operation of NetApp storage systems, and operation of the Solaris™ operating system and that they have general knowledge of networking and MetroCluster concepts.

2 OVERVIEW

2.1 METROCLUSTER

NetApp MetroCluster is an integrated high-availability and disaster-recovery solution that can reduce system complexity and simplify management while ensuring greater return on investment. MetroCluster uses clustered server technology to replicate data synchronously between sites located miles apart, eliminating data loss in case of a disruption. The simple and powerful recovery process minimizes downtime, with little or no user action required.

Integrated Solution Minimizes Costs: With MetroCluster you can simultaneously protect critical data and increase system availability. The MetroCluster solution reduces costs by combining high-availability clustering with powerful site failover capability.

Rapid, Automatic Recovery in Case of a Disruption: To eliminate downtime and ensure uninterrupted data availability, MetroCluster offers automatic recovery for any single component failure. One-button recovery for site disasters means that your business is up and running in minutes, even after a catastrophic site failure.

Near-Zero Data Loss Ensures Business Continuance: MetroCluster uses synchronous replication to a secondary data center located within 100KM. In the event of a site disaster, the mirror copy is instantly available, preserving productivity with near-zero data loss.

2.2 METROCLUSTER TYPES

MetroCluster configurations consist of a pair of active-active storage controllers configured with mirrored aggregates and extended distance capabilities to create a high-availability disaster-recovery solution. The primary benefits include:

- Faster high-availability and disaster-recovery protection
- Minimal risk of lost data, easier disaster recovery, and reduced system downtime
- Quicker recovery when a disaster occurs
- Minimal disruption to users and client applications

In a MetroCluster configuration, each disk shelf on a storage controller has a mirror shelf on its partner controller. This includes the shelf that contains the root volume. Not all shelves on the local node need to be mirrored (other than the shelf that contains the root volume). However, any data contained on a disk shelf that is not mirrored on its partner will obviously not be available in a failover situation.

STRETCH METROCLUSTER

Stretch MetroCluster, sometimes referred to as non-switched, is simply an active-active configuration that can extend up to 500m, depending on speed and cable type (see Figure 1). It also includes synchronous mirroring (SyncMirror®) and the ability to perform a site failover with a single command.

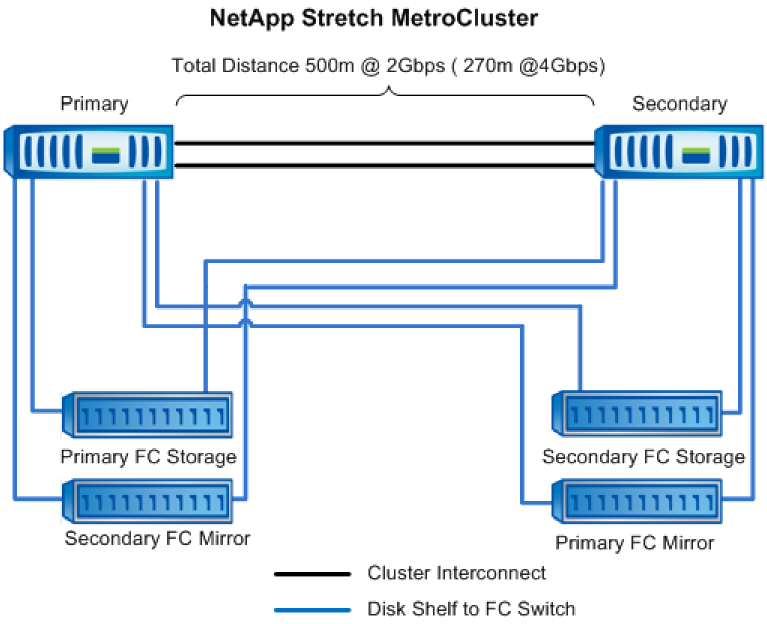


Figure 1) Stretch MetroCluster (non-switched).

FABRIC METROCLUSTER

Fabric MetroCluster also referred to as switched, uses fibre channel (FC) switches in a fabric to achieve greater distances (up to 100km). NetApp recommends using a minimum of four FC switches in a dual fabric configuration, with a separate cluster interconnect card between primary and disaster recovery locations.

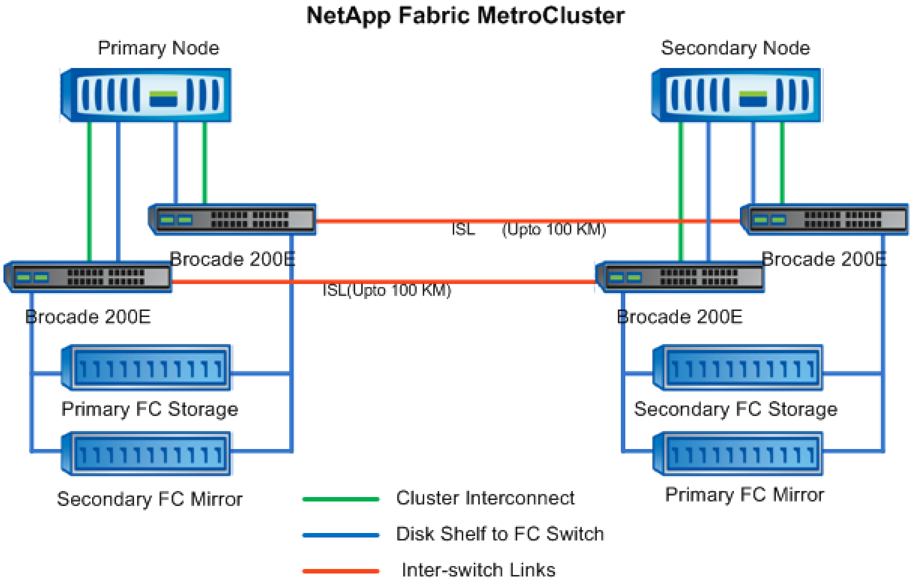


Figure 2) Fabric MetroCluster (switched).

Mirroring

NetApp SyncMirror, an integral part of MetroCluster, combines the disk-mirroring protection of RAID1 with NetApp industry-leading RAID4 and RAID-DP™ technology. In the event of an outage—whether it's due to a disk problem, cable break, or host bus adapter (HBA) failure—SyncMirror can instantly access the mirrored data without any operator intervention or disruption to client applications. SyncMirror maintains a strict physical separation between the two copies of your mirrored data. Each of these copies is referred to as a *plex*. Each controller's data has its "mirror" at the other site.

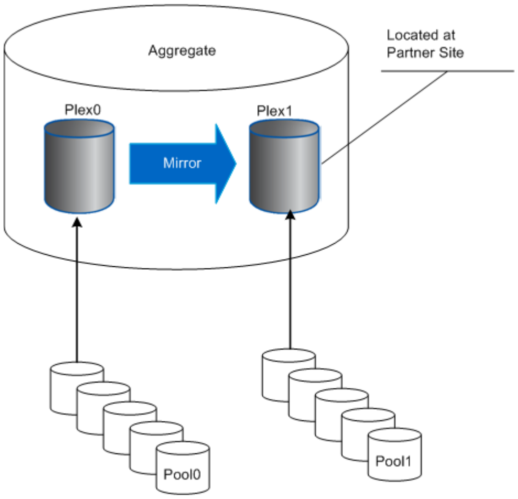


Figure 3) SyncMirror pools and plexes.

When SyncMirror is licensed and hardware ownership is used (see Disk Ownership), spare disks are split into two pools—Pool0 and Pool1. Each plex of a mirror uses disks from a separate pool. Disks assigned to

Pool0 must be on separate loops from disks assigned to Pool1. When software ownership is used, disks are explicitly assigned by the administrator, as described in Disk Ownership.

To maximize availability, Pool0 and Pool1 disks must be on separate loops and use separate HBAs, cables, and shelves.

Before enabling the SyncMirror license, verify that disks for each pool are located on the appropriate loops, and that the loops are fault isolated from each other.

Disk Ownership In a MetroCluster where disk shelves at either side are mirrored and therefore are accessible by either controller, disk ownership comes into play. There are two methods of establishing disk ownership: hardware and software. Hardware-based ownership is the default for all platforms except V-Series, FAS6000 series, and FAS3040/3070. Software disk ownership capability became available in Data ONTAP® 6.3.1. A brief description of each method follows. For more detail, see section 2.2.2, "Installation and Configuration"; or see the Data ONTAP documentation**.**

Hardware disk ownership establishes which controller owns which disks by how the shelves are connected. For more information, see the *System Configuration Guide:* http://now.netapp.com/NOW/knowledge/docs/hardware/NetApp/syscfg/

SWITCHED FABRIC CONFIGURATION

The NetApp Fabric MetroCluster (switched) configuration shown in Figure 4 uses four Brocade 200E FC switches in a dual fabric configuration to connect two active-active controllers. These switches can be combined with other 4Gbps Brocade FC switches or FC director.

A Fabric MetroCluster configuration contains two FC fabrics, each spanning the primary and remote sites. An FC fabric consists of an FC switch on the primary controller connected to a switch on the remote controller (see Figure 4). The two FC switches are connected to each other through interswitch link (ISL) cables.

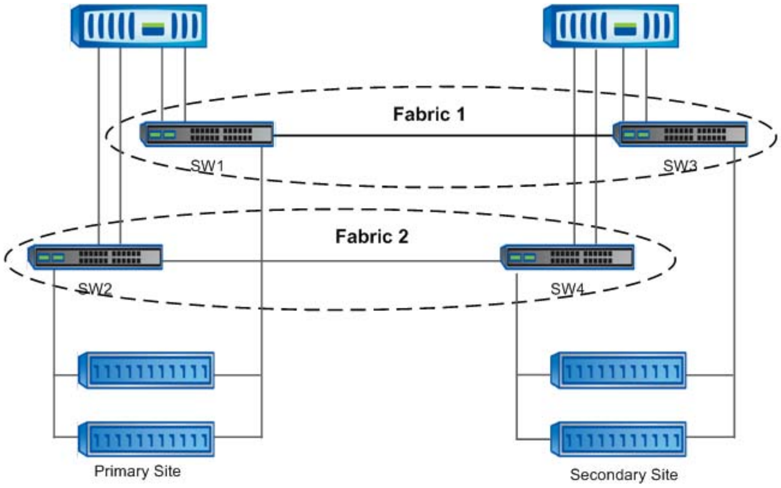


Figure 4) Fabric MetroCluster switch outline.

Figure 4 shows a Fabric MetroCluster configuration in which the first fabric is created by connecting FC switch SW1 at the primary site to FC switch SW3 at the remote site by using an ISL. The second fabric is created by connecting FC switch SW2 at the primary site to FC switch SW4 at the remote site by using an ISL. Note that redundancy for each fabric could also be obtained by using two ISLs to connect two switches in the same fabric. Two FC fabrics are recommended as best practice for fabric redundancy. The loss of a switch in a fabric or the loss of a fabric does not affect the availability of the Fabric MetroCluster.

See the MetroCluster Compatibility Matrix at http://now.netapp.com/NOW/knowledge/docs/olio/guides/metrocluster_compatibility/.

INSTALLATION AND CONFIGURATION

For detailed information about the procedure for setting up the primary and remote sites, see the *MetroCluster Design and Implementation Guide* at http://www.netapp.com/library/tr/3548.pdf.

2.3 SUN CLUSTER

Solaris Cluster can provide a multi-site disaster recovery solution that manages the availability of application services and data across local, regional, and widely dispersed data centers.

Sun Microsystems has released the Solaris Cluster source code through the HA Clusters community on OpenSolaris. The first contribution includes the application modules, called agents. A Sun Cluster agent is a k-sh script, a C-program, or a binary that manages the availability of an application. The agent starts, stops, and monitors the health of the application and takes corrective action to regain application availability upon failure. Applications do not need to be modified to benefit from the enhanced availability offered by the Sun Cluster agent.

The Solaris Cluster group, as well as other third-party software vendors, has created several agents for popular applications such as Java™ Enterprise System applications, Oracle, Siebel, SAP, Sybase, Sun middleware applications, and many others.

The Solaris Cluster software is a framework that extends the high-availability features of Solaris. It includes Solaris Cluster, developer tools, and support for commercial and open-source applications through the use of agents. Solaris Cluster software provides application and service failover for up to 16 data center nodes; this integrated software provides high availability and disaster recovery for clusters.

Heartbeat Mechanism The servers (nodes) in a cluster communicate through private interconnects, which carry important data as well as a cluster "heartbeat." This heartbeat lets each server know the health of any other servers within the cluster, ensuring that each server is "alive." If one of the servers ceases its heartbeat and goes offline, the rest of the devices in the cluster isolate the server and fail over any application or data from the failing node to a working node. This failover process is done quickly and is transparent to the users of the system. By exploiting the redundancy in the cluster, Solaris Cluster ensures the highest levels of availability.

NETAPP STORAGE WITH SUN CLUSTER ON SAN

You can access the LUN by using NetApp host utilities and incorporate the logical drives into Sun Cluster by using Sun Volume Manager or Veritas® Volume Manager.

Host utilities both simplify the products and provide more options for the host stacks. A single FCP Solaris host utility can contain software tools and documentation packages to support different host configurations.

The host utilities provide the tools and information necessary to enable a Solaris host to connect to your storage systems. The host utilities consist of the following elements:

- The SAN Toolkit, which contains:
  - Configuration tools that help you configure host, HBA, and system files and create persistent bindings.
    You can configure the Solaris host by using the tools provided in the FCP Solaris host utilities, HBA provided tools, or a combination of the two.
  - The `sanlun` utility, which helps you manage LUNs and the host HBA.
  - Diagnostic scripts that provide diagnostic information about components in your configuration. Customer support may ask for output from these scripts.
- Documentation that describes how to configure the Solaris host, including this setup guide, release notes, and a quick command reference guide.

For detailed procedures about Sun Volume Manager and Veritas Volume Manager, see http://docs.sun.com/app/docs/doc/819-0420.

Storage Area Network (SAN)

With Sun Certification for NetApp Fibre Channel storage, Solaris OS customers using NetApp storage powered by Data ONTAP can double their storage utilization, automatically manage fine-grained data, and reduce storage management costs while reaping the high-availability benefits associated with Sun Cluster software

NETAPP STORAGE WITH SUN CLUSTER ON NAS

When you install the support package (provided by NetApp) on Sun Cluster hosts running the NetApp application, it performs the following two tasks:

- In the event of a host failure, the support package prevents the failed host from writing data to the storage systems. It fences off the failed host from the storage systems.
- The support package also lets the Sun Cluster hosts use a LUN on the storage systems as a quorum device. Having the quorum capabilities ensures that multiple independent clusters do not exist if a host cluster failure occurs. The Sun Cluster uses a SCSI-3 persistent reservation mechanism to maintain membership information about the cluster on the LUN.

The Sun Cluster system attempts to prevent data corruption and ensure data integrity by the use of a quorum device and I/O fencing. The quorum device is used when a cluster becomes partitioned into separate sets of nodes, to establish which set of nodes constitutes the new cluster. In other words, it prevents multiple independent clusters from existing in case of a cluster failure. Once the quorum device has helped to determine which cluster is active, it fences off I/O from any node that does not belong to the cluster. This I/O fencing mechanism ensures that any failed nodes do not have access to the shared data.

NetApp uses the storage system’s LUN as a quorum device and iSCSI as the transport between Sun Cluster hosts and the NetApp storage systems.

The software supplied by NetApp consists of two tools:

- The first part accomplishes the fencing of the cluster nodes from the data exported by the storage system. This is shipped in the form of a binary called NTAPfence.
- The second part enables the usage of the storage system’s LUN as the quorum device. This is shipped in the form of a binary called NTAPquorum.

NTAPfence and NTAPquorum are supplied as part of the Solaris NTAPclnas package, which is available on the NOW™ (NetApp on the Web) site: http://now.netapp.com.

3 DEPLOYMENT SETUP

The following configuration example is used to illustrate concepts throughout the rest of this document.

The two physical sites are named SUN4500-WHQL01 and SUN4500-WHQL02. The storage controllers are separated by a distance of 25.2 km using single mode fiber spools. This solution uses NetApp MetroCluster on the back end for storage availability and a two-node Sun Cluster on the front end for application availability. It is also possible to have more than two nodes, depending on the application’s requirement in one cluster. The application used in this test configuration example is Oracle Database. The nodes run Solaris, Sun Cluster software, and Sun Cluster Support Package for NetApp. The servers access their storage via the iSCSI protocol for quorum and the NFS protocol for Oracle® database. In the “normal” situation, one server is active on the primary site and accesses the storage at that site. The DR site is passive and is able to access the storage on the primary site. The general layout of the components used in this sample configuration is shown in Figure 7.

The following notes are specific to this solution; we discovered these points during the test:

- Provide a separate system for the setup and avoid allowing other applications to access the system for NFS share.
- Data ONTAP version 7.2.x or later is required.
- If you are using Solaris 8, use up-to-date patches; or, at minimum, install the 108987-16, 109147-28, 117000-05, 108993-39, 110934-21, 111111-04, 108434-18, and 108435-18 patches for the Sun Cluster setup.
- The systems require a password to log in.
- For vFiler™ support of Sun Cluster, Data ONTAP will have the necessary support in 2008. Contact NetApp technical support to request an early version.

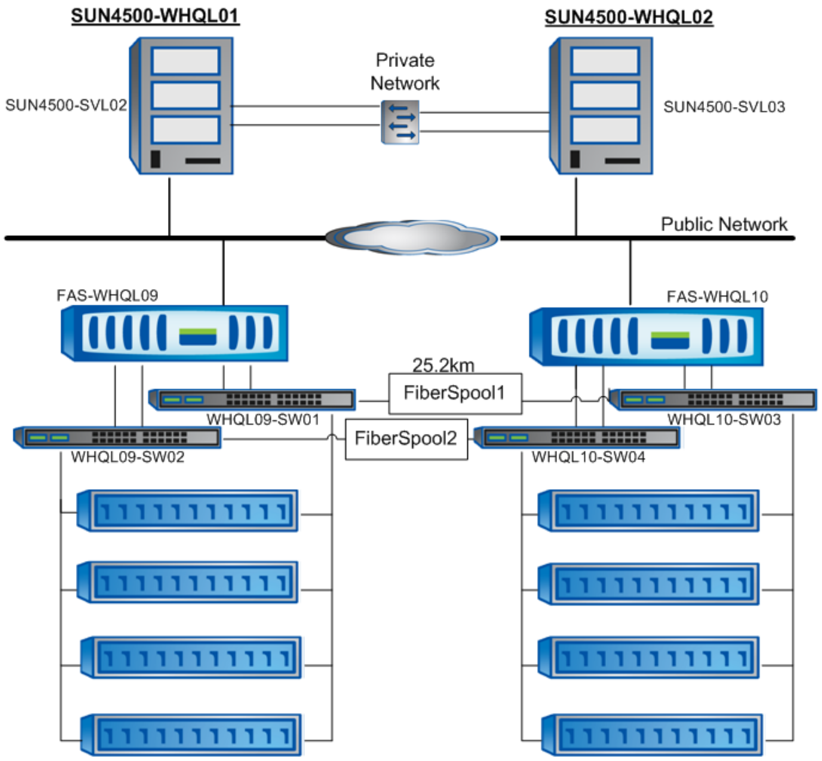


Figure 5) Cluster setup.

4 DEPLOYMENT PLANNING

4.1 DATA GATHERING

To deploy a successful MetroCluster with Sun Cluster configuration, you should gather the following information and items before beginning the deployment:

- Distance between the primary and remote sites

  This information is necessary to determine which type of MetroCluster is appropriate, or even whether either version is appropriate. In calculating the effective distance, consider such factors as cable type, speed, and number of patch panels. Although NetApp recommends that dedicated dark fiber be used for MetroCluster, WDM devices are supported. For information about supported devices, see the Brocade Compatibility Guide at www.brocade.com.

- NetApp and Fibre Channel switch licenses
- Hostnames and IP addresses for each of the nodes and the Fibre Channel switches
- Brocade switch licenses
- Cables

4.2 DISTANCE CONSIDERATIONS

Stretch MetroCluster can support a maximum of 500 meters between nodes at a speed of 2Gbps. Through the use of Fibre Channel switches, Fabric MetroCluster can extend the distance even further. This distance depends on the FC switch that is being used. Currently, a Brocade 4Gbps FC switch and qualified SPFs typically support 100km, and even farther distance at lower speed. For exact distances supported, refer to the FC switch specification. This extended distance capability gives customers greater flexibility in the physical location of the active-active controllers while maintaining the high-availability benefits of active-active controller configuration

This section describes a number of factors that affect the overall effective distance possible between the MetroCluster nodes:

- Physical distance
- Number of connections
- Desired speed
- Cable type

As stated earlier, the Stretch MetroCluster configuration can extend to a maximum of 500m (2Gbps). This distance is reduced by speed, cable type, and number of connections. A Fabric MetroCluster can extend out to 100km. This distance is affected by the same factors. At a distance of 100km, latency is around 1ms. Greater distances obviously result in greater latencies (500km = 5ms), which may be unacceptable to an application.

For more information about cable types and cabinets, see the *MetroCluster Design and Implementation Guide* at http://www.netapp.com/library/tr/3548.pdf .

5 CLUSTER CONFIGURATION CHECKER

The Cluster Configuration Checker is a Perl script that detects errors in the configuration of a pair of active-active NetApp controllers. It can be run as a command from a UNIX® shell or Windows® prompt, but it also doubles as a CGI script that can be executed by a UNIX Web server. The script uses rsh or ssh to communicate with the controllers you're checking, so you must have the appropriate permissions for rsh to run on both controllers in the cluster pair. This script detects and reports the following:

- Services licensed that are not identical on the partner (some services may be unavailable on takeover)
- Options settings that are not identical on the partner (some options may be changed on takeover)
- Network interfaces that are configured incorrectly (clients disconnect during takeover)
- FCP cfmode settings that are not identical on controllers that have FCP licensed
- Checks `/etc/rc` on each controller to see that all interfaces have a failover set

This script is available for download from the NOW site. NetApp recommends running this script as part of the implementation process.

If the controllers being implemented were part of an active-active configuration, then the configurations are probably compatible. The safe practice is to run the Cluster Configuration Checker.

6 RECOMMENDATIONS

NetApp strongly recommends that you complete the installation planning worksheets before beginning the installation. A little time spent up front will expedite the installation process. Section 7 gives an example of a configuration with specific equipment used. The following setup points are recommended:

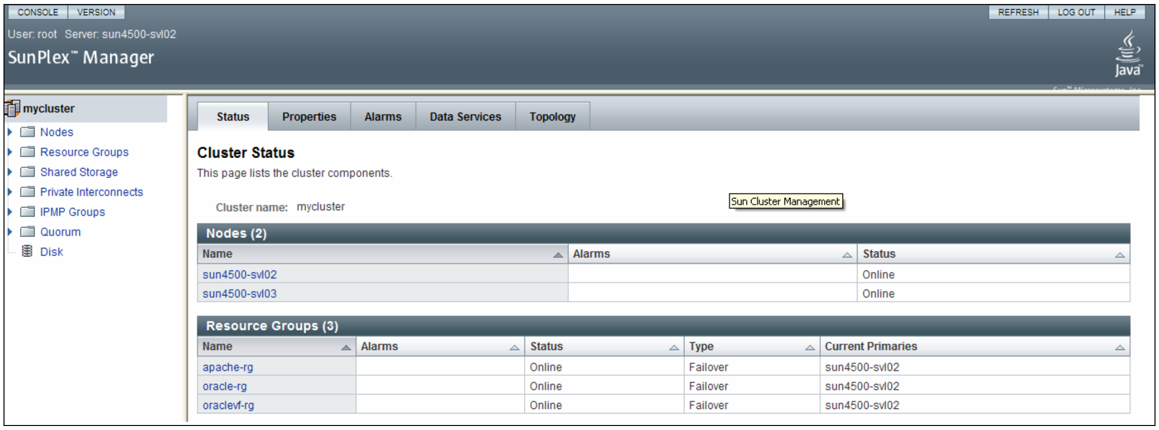
- Separate networks for front end and back end
- Redundancy network between hosts and storage systems
- Multipath for LUNs to access the storage systems

7 MANAGING THE DR SETUP

You can manage the Sun Cluster by using SunPlex™ Manager from Sun Microsystems. Also, you can configure and manage MetroCluster and the NetApp storage system by using NetApp FilerView®.

7.1 SUN CLUSTER – SUNPLEX MANAGER

- SunPlex Manager is a browser-based GUI for Sun Cluster.
- It is used to monitor and manage cluster configurations.
- You can access SunPlex Manager by using https://<clusternode name>:6789:
  1. Log in as root/xxxxxx.
  2. Click SunPlex Manager.
- You can perform all cluster manageability functions by using SunPlex Manager.



User: root Server: sun4500-svl02

Cluster Status

This page lists the cluster components.

Cluster name: mycluster

Nodes (2)

Name	Alarms	Status
sun4500-svl02		Online
sun4500-svl03		Online

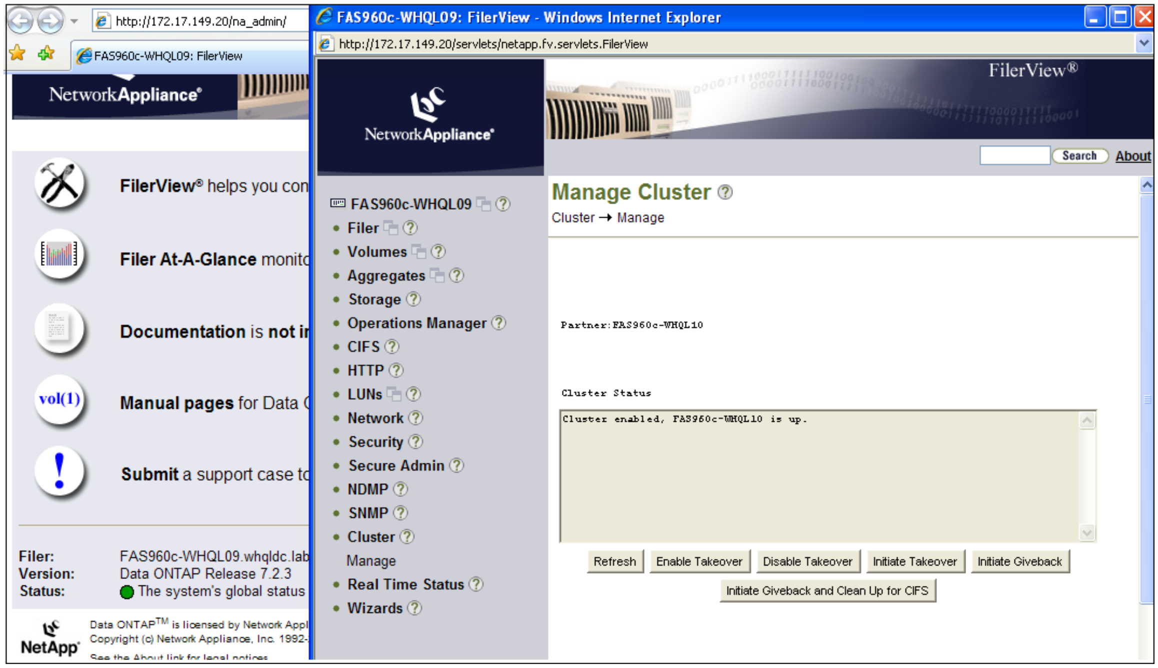
Resource Groups (3)

Name	Alarms	Status	Type	Current Primaries
apache-rg		Online	Failover	sun4500-svl02
oracle-rg		Online	Failover	sun4500-svl02
oraclevf-rg		Online	Failover	sun4500-svl02

Figure 6) SunPlex Manager.

7.2 NETAPP STORAGE SYSTEM – FILERVIEW

- FilerView is a browser-based GUI for NetApp storage systems.
- It is used to monitor and manage NetApp storage systems.
- You can access FilerView by using https://<filername name>/na_admin.
  1. Log in as root/xxxxxx.
  2. Click FilerView.
- You can perform all the storage system management functions by using FilerView.



FilerView® helps you con

Filer At-A-Glance monito

Documentation is **not i**

Manual pages for Data

Submit a support case to

Filer: FAS960c-WHQL09.whqldc.lab
Version: Data ONTAP Release 7.2.3
Status: ● The system's global status

Data ONTAP™ is licensed by Network Appl
Copyright (c) Network Appliance, Inc. 1992-
See the About link for legal notices

Manage Cluster

Cluster → Manage

```
Partner:FAS960c-WHQL10

Cluster Status

Cluster enabled, FAS960c-WHQL10 is up.
```

8 EXAMPLE SETUP

The example setup explained the procedure to configure NetApp MetroCluster with Sun Cluster for Oracle Database. For the detailed procedure, test scenarios, and FAQ, see the appendix.

9 REFERENCES

MetroCluster Documentation

TR 3517: MetroCluster Upgrade Planning Guide

Active-Active Configuration Guide

Data Protection Online Backup and Recovery Guide (Chapter 8, SyncMirror)—MetroCluster Compatibility Matrix

TR 3548: MetroCluster Design and Implementation Guide

TR 3412: MetroCluster Upgrade Planning Guide

TR 3614: Implementing Oracle Database 11*g* Running with Direct NFS Client on NetApp MetroCluster

Brocade Switch Documentation: http://now.netapp.com/NOW/knowledge/docs/brocade/relbroc30_40/

Sun Cluster_Documentation

TR 3570: Sun Cluster with NetApp Storage (NFS) for High Availability

Sun Cluster Data Service for Oracle Guide for Solaris OS: http://docs.sun.com/app/docs/doc/819-0694

10 APPENDIX

This appendix explains the detailed step-by-step procedure to configure NetApp MetroCluster with Sun Cluster for Oracle Database:

1. Materials list
2. NetApp storage system
3. Switch configuration
4. Sun Cluster
5. NetApp and Sun Cluster for Oracle
6. Test scenarios and FAQ

10.1 MATERIALS LIST

Hardware	Vendor	Name	Version	Description
Storage	NetApp	FAS960AA (2)	N/A	Storage Controller
Hosts	SUN	One Solaris Node on each site 1 X sun 4500 (8x400MHz/7GB RAM)	N/A	Host Server for Sun Cluster
Front-End Network	Cisco	4948 (4)	IOS 12.1	48 Port Ethernet Switch
Back-End SAN (MetroCluster)	Brocade	200E (4)	5.1.0	16 Port FC Switch
Software	**Vendor**	**Name**	**Version**	**Description**
Storage	NetApp	SyncMirror	7.2.x or later	Replication
	NetApp	Data ONTAP	7.2.x or later	Operating system
	NetApp	Cluster_Remote	7.2.x or later	Failover
Hosts	Sun	Solaris Sparc 64-bit with quad port NIC	8,9 or 10	Operating system. Two NICs for cluster interconnect and another two NICs for public network and IP multipathing
	Oracle	Oracle Database	9i,10g or 11gR1	Database
	Sun	Sun Cluster	3.1 or later	Solaris Cluster software
	NetApp	Support Package	1.2	NetApp support package to provide the fencing and quorum for Sun Cluster

10.2 NETAPP STORAGE SYSTEM

FAS STORAGE CONTROLLER

The controller and back-end Fibre Channel switches were configured by using the instructions in the *Data ONTAP 7.3.1 Active/Active Configuration Guide* and the current firmware levels and other notes found on the NOW site.

- Data ONTAP 7.2.x or later
- Brocade firmware 5.3.0

Two FAS960 series controllers (each with four DS14mk2-HA shelves full of 66GB 15k drives), connected with the VI-MC interconnect, and four Brocade 200E switches were used in this test. The controllers were named FAS-WHQL09 and FAS-WHQL10, and the switches were named WHQL09-SW01, WHQL09-SW02, WHQL10-SW03, and WHQL10-SW04.

SLOT ASSIGNMENTS

The controllers were configured identically in terms of hardware with the following cards and slot assignments.

Slot #	Card	Purpose
1	X3300A: Remote management card	Remote monitoring and management
5	X2050A: Dual optical Fibre Channel for mirroring	Disk connection
6	X1922A: VI-MetroCluster	Cluster interconnect
7	X3140A: NVRAM4	NVRAM card
8	X2050A: Dual optical Fibre Channel for mirroring	Disk connection
11	X2050A: Dual optical Fibre Channel for target interconnect	Target card

NETWORK SETTINGS

FAS-WHQL09

Interface	IP Address	Purpose
E0	172.17.149.20, Partner e0	LAN

FAS-WHQL10

Interface	IP Address	Purpose
E0	172.17.149.25, Partner e0	LAN

AGGREGATE LAYOUT

Controller	Aggregate Name	Options	# of Disks	Purpose
FAS-WHQL09	aggr1	nosnap=off, raidtype=raid_dp, raidsize=28, ignore_inconsistent=off, snapmirrored=off, resyncsnaptime=60, fs_size_fixed=off, snapshot_autodelete=on, lost_write_protect=on	14	Sun Cluster quorum and Oracle Database for primary site
FAS-WHQL10	ora_linux	nosnap=off, raidtype=raid_dp, raidsize=16, ignore_inconsistent=off, snapmirrored=off, resyncsnaptime=60, fs_size_fixed=off, snapshot_autodelete=on, lost_write_protect=on	14	Oracle Database, Sun Cluster 3.1 software, and patches are stored for the demo setup, which is not considered for Sun Cluster demo on the secondary site

VOLUME LAYOUT

Controller	Volume Name	Options	Size (GB)	Purpose
FAS-WHQL09	Vol0	root, diskroot, nosnap=off, raidtype=raid_dp, raidsize=16, ignore_inconsistent=off, snapmirrored=off, resyncsnaptime=60, fs_size_fixed=off, snapshot_autodelete=on, lost_write_protect=on	43	Root volume
FAS-WHQL09	Sun Cluster	nosnap=off, nosnapdir=off, minra=off, no_atime_update=off, nvfail=off, ignore_inconsistent=off, snapmirrored=off, create_ucode=off, convert_ucode=off, maxdirsize=62914, schedsnapname=ordinal, fs_size_fixed=off, guarantee=volume, svo_enable=off, svo_checksum=off, svo_allow_rman=off, svo_reject_errors=off, no_i2p=off, fractional_reserve=100, extent=off, try_first=volume_grow	48	For Sun Cluster quorum, but for the quorum only 1GB is needed
FAS-WHQL09	Data	nosnap=off, nosnapdir=off, minra=off, no_atime_update=off, nvfail=off, ignore_inconsistent=off, snapmirrored=off, create_ucode=off, convert_ucode=off, maxdirsize=62914, schedsnapname=ordinal, fs_size_fixed=off, guarantee=volume, svo_enable=off, svo_checksum=off, svo_allow_rman=off, svo_reject_errors=off, no_i2p=off, fractional_reserve=100, extent=off, try_first=volume_grow	58	Oracle Database accessed through NFS protocol

FAS-WHQL10	Data	nosnap=off, nosnapdir=off, minra=off, no_atime_update=off, nvfail=off, ignore_inconsistent=off, snapmirrored=off, create_ucode=off, convert_ucode=off, maxdirsize=62914,schedsnapname=ordinal, fs_size_fixed=off, guarantee=volume, svo_enable=off, svo_checksum=off, svo_allow_rman=off, svo_reject_errors=off, no_i2p=off, fractional_reserve=100, extent=off, try_first=volume_grow	47	Oracle Database, Sun Cluster 3.1 software, and patches are stored for the demo setup, which is not considered for Sun Cluster demo on the secondary site

SUNCLUSTER LUN AND NFS FOLDER

Controller	Volume Name	Options in Volume Level	Protocol		Size (GB)	Purpose
FAS-WHQL09	Sun Cluster	root, diskroot, nosnap=off, raidtype=raid_dp, raidsize=16, ignore_inconsistent=off, snapmirrored=off, resyncsnaptime=60, fs_size_fixed=off, snapshot_autodelete=on, lost_write_protect=on	ISCSI		43	Volume for quorum partition
FAS-WHQL09	Data	nosnap=off, nosnapdir=off, minra=off, no_atime_update=off, nvfail=off, ignore_inconsistent=off, snapmirrored=off, create_ucode=off, convert_ucode=off, maxdirsize=62914, schedsnapname=ordinal, fs_size_fixed=off, guarantee=volume, svo_enable=off, svo_checksum=off, svo_allow_rman=off, svo_reject_errors=off, no_i2p=off, fractional_reserve=100, extent=off, try_first=volume_grow	NFS	NFS exported options rw,bg,hard, forcedirectio,nointr, rsize=32768, wsize=32768, proto=tcp, vers=3	58	Oracle Database accessed through NFS protocol

FAS-WHQL10	Data	nosnap=off, nosnapdir=off, minra=off, no_atime_update=off, nvfail=off, ignore_inconsistent=off, snapmirrored=off, create_ucode=off, convert_ucode=off, maxdirsize=62914, schedsnapname=ordinal, fs_size_fixed=off, guarantee=volume, svo_enable=off, svo_checksum=off, svo_allow_rman=off, svo_reject_errors=off, no_i2p=off, fractional_reserve=100, extent=off, try_first=volume_grow	NFS	NFS exported options rw,bg,hard,forcedirectio,nointr,rsize=32768,wsize=32768,proto=tcp,vers=3	47	Oracle Database, Sun Cluster 3.1 software, and patches are stored for the demo setup, which is not considered for Sun Cluster demo on the secondary site

10.3 SWITCH CONFIGURATION

For the solution to function properly, the back-end FC switches in a MetroCluster environment must be set up in a specific manner. This section details the switch and port connections, which should be implemented exactly as documented.

WHQL09_SW01
IP Address: 172.17.149.235
Domain ID: 1

Port	Bank/Pool	Connected with	Purpose
0	1/0	FAS-WHQL09, 5a	FAS-WHQL09 FC HBA
1	1/0	FAS-WHQL09, 8a	FAS-WHQL09 FC HBA
2	1/0		
3	1/0		
4	1/1		
5	1/1	FAS-WHQL10, Pool1 Shelf 3B	
6	1/1		
7	1/1		
8	2/0		
9	2/0	FAS-WHQL10, Pool0, Shelf 1B	
10	2/0		
11	2/0		
12	2/1	FAS-WHQL09, FCVI, 6a	Cluster Interconnect
13	2/1	WHQL10_SW03, Port 5	ISL
14	2/1		
15	2/1		

WHQL09_SW02
IP Address: 172.17.149.236
Domain ID: 2

Port	Bank/Pool	Connected with	Purpose
0	1/0	FAS-WHQL09, 5a	DISK HBA for Bank 2 Shelves
1	1/0	FAS-WHQL09, 8a	DISK HBA for Bank 2 Shelves
2	1/0		
3	1/0		

4	1/1		
5	1/1	FAS-WHQL10, Pool1 Shelf 3A	
6	1/1		
7	1/1	FAS-WHQL09 FCVI, 6b	Cluster Interconnect
8	2/0		
9	2/0	FAS-WHQL09, Pool0, Shelf 1A	
10	2/0		
11	2/0		
12	2/1		
13	2/1	WHQL10_SW04, Port4	ISL
14	2/1		
15	2/1		

WHQL10_SW03
IP Address: 172.17.149.237
Domain ID: 3

Port	Bank/Pool	Connected with	Purpose
0	1/0	FAS-WHQL10, Pool1 Shelf 3B	
1	1/0		
2	1/0		
3	1/0	FAS-WHQL09, FCVI, 6a	Cluster Interconnect
4	1/1		
5	1/1	WHQL09_SW01, Port 13	ISL
6	1/1		
7	1/1		
8	2/0	FAS-WHQL10, 5a	Disk HBA Bank 2 Shelves
9	2/0	FAS-WHQL10, 8a	Disk HBA Bank 2 Shelves
10	2/0		
11	2/0		
12	2/1	FAS-WHQL10, Pool 0, Shelf 1B	
13	2/1		
14	2/1		
15	2/1		

WHQL10_SW04
IP Address: 172.17.149.238
Domain ID: 4

Port	Bank/Pool	Connected with	Purpose
0	1/0	FAS-WHQL09, Pool 1, Shelf 3A	
1	1/0		
2	1/0		
3	1/0		
4	1/1	WHQL09_SW02, Port 13	ISL
5	1/1		
6	1/1		
7	1/1		
8	2/0	FAS-WHQL10, 5b	Disk HBA Bank 2 Shelves
9	2/0	FAS-WHQL10, 8b	Disk HBA Bank 2 Shelves
10	2/0		
11	2/0		
12	2/1	FAS-WHQL10, Pool 0, Shelf 1A	
13	2/1	FAS-WHQL10 FCVI, 6b	Cluster Interconnect
14	2/1		
15	2/1		

10.4 SUN CLUSTER

HARDWARE

See section 10.1, "Materials List."

HOST CONFIGURATION

The Sun Cluster setup spans across sites with two nodes as members. The hosts in the Sun Cluster are named SUN4500-SVL02 and SUN4500-SVL03.

SOFTWARE CONFIGURATION

The hosts in the Sun Cluster are installed according to vendor-supplied procedures with:

- Sun Solaris 8, 9 or 10 SPARC 64 bit
- Solaris up-to-date patches; for latest patches see http://sunsolve.sun.com
- Oracle Database 9i, 10g or 11gR1
- Sun Cluster 3.2 U2
- NetApp support package

NETWORK CONFIGURATION

The following table details the network settings for various Oracle hosts.

Hostname	IP Address	Purpose
Sun4500-svl02	172.17.148.235	LAN
Sun4500-svl02	172.16.1.2	Heartbeat
Sun4500-svl02	172.16.0.130	Heartbeat
Oracle-ip	172.17.148.242	Virtual IP for Oracle
Sun4500-svl03	172.17.148.240	LAN
Sun4500-svl03	172.16.1.1	Heartbeat
Sun4500-svl0b3	172.16.0.129	Heartbeat

PLANNING SUN CLUSTER HA FOR ORACLE

Configuration Requirements

- **Oracle application files**: These files include Oracle binaries, configuration files, and parameter files. You can install these files on the local file system, on the highly available local file system, or on the cluster file system. In our setup, we kept the files on the local file system.
- **Database-related files**: These files include the control file, redo logs, and data files. You must install these files on the highly available local file system or on the cluster file system either as raw devices or as regular files. In our setup we kept these files in the NFS folder, which we accessed from the NetApp storage system.

Configuration Planning Questions

The answers to the following questions are based on our setup. You will provide the answers based on your setup.

Question: What resource groups will you use for network addresses and application resources and the dependencies between them?

Answer: Network Address: 172.17.148.242
Application resource: oracle-rg
Dependencies: oracle-ip
oracle-server-l
oracle-listener-l

Question: What is the logical hostname (for failover services) or shared address (for scalable services) for clients that will access the data service?

Answer: Logical hostname for failover service: oracle-ip

Question: Where will the system configuration files reside?

Answer: On the local file system.

Question: Where will the Oracle Database files reside?

Answer: In the NFS folder on the NetApp storage system.

10.5 NETAPP STORAGE AND SUN CLUSTER FOR ORACLE

LUN AS A QUORUM DEVICE

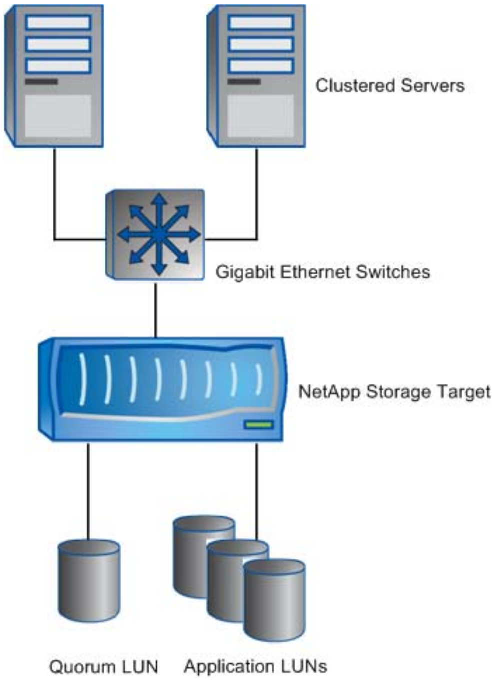


Figure 7) iSCSI- based cluster.

1. The quorum device that we used for Sun Cluster is accessed through iSCSI protocol from the NetApp storage system.
2. Configure the storage systems and nodes to use NTP for syncing the time.
3. Make sure that the host’s entry in the storage systems and cluster nodes are identical. For example:

```
bash-2.03# cat /etc/hosts
127.0.0.1       localhost
172.17.148.235  sun4500-svl02   loghost
172.17.148.240  sun4500-svl03
172.17.149.20   filer1
172.17.149.25   filer2
172.17.148.241  apache-ip
172.17.148.242  oracle-ip
```

```
172.17.148.247  oraclevf-ip
172.17.149.23   FAS-WHQL09-e9a vfiler1
172.17.148.246  vapache-ip
```

4. Set the password for *root* in physical storage system and vFiler by using *passwd.*
5. Provide the HTTP setting on the storage system.

   Enable the HTTP access on the storage system by turning on `http.admin.access` for all hosts that require admin access:

   ```
   FAS-WHQL09> options httpd.admin
   httpd.admin.access           all
   httpd.admin.enable           on
   httpd.admin.hostsequiv.enable off
   httpd.admin.max_connections  512
   httpd.admin.ssl.enable       off
   httpd.admin.top-page.authentication on
   FAS-WHQL09> vfiler context vfiler1
   vfiler1@FAS-WHQL09> options httpd.admin
   httpd.admin.access           all
   httpd.admin.enable           on
   httpd.admin.hostsequiv.enable on
   httpd.admin.max_connections  512
   httpd.admin.top-page.authentication on
   ```

6. Check the iSCSI license.

   The LUN designed as a quorum device uses the iSCSI protocol. Check whether the iSCSI is licensed on the storage system that contains that LUN.

   At the storage prompt:

   ```
   FAS-WHQL09> iscsi status
   iscsi: iSCSI is not licensed.
   Now you need to add the license for iSCSI
   FAS-WHQL09> license add xxxxxx
   FAS-WHQL09> iscsi start
   FAS-WHQL09> iscsi status
   iSCSI service is running
   ```

7. Get the iSCSI node name (optional).

   You need the iSCSI node name of the storage system to configure the LUN for quorum. You can get the iSCSI node name by using the following command from the storage system:

   ```
   FAS-WHQL09> iscsi nodename
   iSCSI target nodename: iqn.1992-08.com.netapp:sn.50393364
   ```

   You can set the hostname as the suffix for the IQN. The following example replaces the hostname FAS-WHQL09 as the suffix for the IQN of the storage system:

   ```
   FAS-WHQL09> iscsi nodename iqn.1992-08.com.netapp: FAS-WHQL09
   FAS-WHQL09> iscsi nodename
   ```

```
iSCSI target nodename: iqn.1992-08.com.netapp: FAS-WHQL09
```

8. Create a LUN on a storage system.

   On the storage system, create an igroup for each system in the cluster:

   ```
   igroup create { -f | -i } [ -t <ostype> ] <initiator group> [ <node> ... ]
   ```

   Create a LUN by using the *lun create* command, which has the following format:

   ```
   lun  create –s <size> [-t <type>] [-o noreserve] <lun_path>
   ```

   If you don't want space reservations, use the *–o noreserve* option.

   Note: You can also create a LUN interactively by using the *lun setup* command

9. Requirements for the LUN used as a quorum device:

   The iSCSI group used for this LUN must be exclusively reserved for this purpose.

   Note: Only the iSCSI initiator node names of the cluster nodes can be in the group. This is because if someone else uses the LUN, the quorum information could be destroyed.

10. Steps to create a LUN as a quorum device.

    a. Log in to the storage system.

    b. Enable the password for the storage systems and vFiler, which is required for the Sun Cluster agent.

    c. Display a list of LUNs, using *lun show –m*:

    ```
    lun show –m
    FAS-WHQL09> lun show -m
    LUN path                      Mapped to          LUN ID  Protocol
    ------------------------------------------------------------------
    /vol/suncluster/quorum     clusterA               0     iSCSI
    ```

11. If no LUN is listed for your cluster, proceed with the following steps.

    a. Create an aggregate for the cluster by using the ***aggr create*** command.

    For example, create an aggregate *aggr1* with size 160GB from two disks of size 80GB. You can use the –d option to specify the disks.

    ```
    FAS-WHQL09> aggr create aggr1 2@80
    ```

    b. Create a volume for quorum.

    For example, create a volume of size 5GB only for quorum.

    ```
    FAS-WHQL09> vol create sun cluster aggr1 5g
    ```

    c. Create a LUN for the cluster by using *lun create.*

    For example, create a LUN named quorum of size 1GB with the OS type Solaris in the Sun Cluster volume.

    ```
    FAS-WHQL09> lun create –s 1g –t Solaris /vol/suncluster/quorum
    ```

    d. Confirm that the LUN exists by using lun show:

    ```
    FAS-WHQL09> lun show
    /vol/suncluster/quorum    1g (1073741824)  (r/w, online)
    ```

    e. Display the list of igroups:

    ```
    FAS-WHQL09> igroup show
    iqn.1991-05.com.microsoft:hpdl380-whql02.whqldc.lab.netapp.com
    (iSCSI) (ostype: windows):
    ```

```
iqn.1991-05.com.microsoft:hpdl380-whql02.whqldc.lab.netapp.com (not
logged in)

iqn.1991-05.com.microsoft:hpdl380-whql01.whqldc.lab.netapp.com (not
logged in)
```

f. If the igroup for the cluster is not listed, create the igroup for the cluster by using `igroup create`.

For example, create an igroup named *clusterA.*

```
FAS-WHQL09> igroup create –i –t Solaris clusterA

FAS-WHQL09> igroup show

clusterA (iSCSI) (ostype: Solaris):

iqn.1991-05.com.microsoft:hpdl380-whql02.whqldc.lab.netapp.com
(iSCSI) (ostype: windows):

iqn.1991-05.com.microsoft:hpdl380-whql02.whqldc.lab.netapp.com (not
logged in)

iqn.1991-05.com.microsoft:hpdl380-whql01.whqldc.lab.netapp.com (not
logged in)
```

g. Add the cluster nodes to the igroup by using `igroup add`:

For example, add the cluster nodes sun4500-svl02 and sun4500-svl03 to igroup clusterA.

```
FAS-WHQL09> igroup add clusterA iqn.1986-05.com.sun:cluster-iscsi-
quorum.sun4500-svl02

FAS-WHQL09> igroup add clusterA iqn.1986-05.com.sun:cluster-iscsi-
quorum.sun4500-svl03
```

You can find the cluster node name by using `scconf` from the cluster node:

```
bash-2.03# scconf -pvv | grep -i "cluster node name"

Cluster node name:          sun4500-svl03

Cluster node name:          sun4500-svl02
```

h. Check the igroups that have the cluster nodes:

For example, check that the sun4500-svl02 and sun4500-svl03 nodes are in included in the clusterA igroup.

```
FAS-WHQL09> igroup show

clusterA (iSCSI) (ostype: Solaris):

iqn.1986-05.com.sun:cluster-iscsi-quorum.sun4500-svl02 (not logged
in)

iqn.1986-05.com.sun:cluster-iscsi-quorum.sun4500-svl03 (not logged
in)

iqn.1991-05.com.microsoft:hpdl380-whql02.whqldc.lab.netapp.com
(iSCSI) (ostype: windows):

iqn.1991-05.com.microsoft:hpdl380-whql02.whqldc.lab.netapp.com (not
logged in)

 iqn.1991-05.com.microsoft:hpdl380-whql01.whqldc.lab.netapp.com (not
logged in)
```

i. Map the LUN with the igroup:

For example, map the quorum LUN to the clusterA igroup.

```
FAS-WHQL09> lun map /vol/suncluster/quorum clusterA

lun map: auto-assigned clusterA=0
```

j. Check the LUN mapped to the igroup.

For example, check the quorum LUN mapped to clusterA.

```
FAS-WHQL09> lun show -m

LUN path                   Mapped to          LUN ID  Protocol
---------------------------------------------------------------
/vol/suncluster/quorum     clusterA               0     iSCSI
```

k. Log out of the storage system.

SUN CLUSTER

Solaris Server Preparation

1. During installation, select Entire Distribution + OEM Software Group.
2. When you install the Solaris operating environment, make sure that the system disk partitions meet the minimum requirement, as shown in the following table. If necessary, change the partitions.

Device Name	File System	Size
c0t10d0s0	/	8GB
c0t10d0s1	Swap	1.5 X RAM
c0t10d0s3	/globaldevices	1GB

   If necessary, add the `/usr`, `/opt`, and `/var` partitions.

3. Check the hardware details by using following commands:

prtconf –v	System memory and reports information about peripheral devices
psrinfo –v	Displays information about processors
showrev –p	Reports which patches are installed
prtdiag –v	Displays system diagnostics information

4. In `/etc/default/login`, comment the `CONSOLE=/dev/console` parameter, which enables the telnet.
5. Enable *rsh* between cluster nodes and storage systems. In the storage systems, enable *rsh* for public IP and logical IP of cluster nodes. In nodes, enable *rsh* for storage systems and vfilers**.**
6. Check that the server has a minimum of two NICs for private connection, which will provide cluster interconnection between cluster members and unplump them before starting the Sun Cluster installation.
7. Keep the domain name the same across cluster members, and change the `/etc/domainname` file accordingly:

```
bash-2.03# cat /etc/defaultdomain

lab.netapp.com
```

8. Create a 1GB partition for `/globaldevices`, if one has not already been created:

```
bash-2.03# format

Searching for disks...done

AVAILABLE DISK SELECTIONS:

   0. c0t10d0 <SUN9.0G cyl 4924 alt 2 hd 27 sec 133>

    /pci@3,4000/SUNW,isptwo@3/sd@a,0

   1. c0t11d0 <SUN9.0G cyl 4924 alt 2 hd 27 sec 133>
```

```
    /pci@3,4000/SUNW,isptwo@3/sd@b,0
Specify disk (enter its number):
```

Select the disk number, either 0 or 1; select the free space partition from the partition table; create the partition with the size 1GB; and enter *label* to write the modification to partition table.

NetApp Sun Cluster Agent Setup

1. Support of fencing and quorum.:

   When you install the support package (Sun Cluster agent) on cluster nodes, it performs two tasks:

   - In the event of cluster node failure, the support package prevents the failed cluster node from writing data to the storage system. It fences off the failed host from the storage system.
   - Sun Cluster nodes use a LUN from the storage system as a quorum device. Having the quorum capabilities ensures that multiple independent clusters do not exist if cluster node failure occurs. The Sun Cluster uses a SCSI-3 persistent reservation mechanism to maintain membership information about the cluster on the LUN.

2. Support package components:

   The Sun Cluster system attempts to prevent data corruption and ensure data integrity by the use of quorum device and I/O fencing. The quorum device is used when a cluster becomes partitioned into separate sets of nodes, to establish which set of nodes constitutes the new cluster. In other words, it ensures that multiple independent clusters do not exist in case of a cluster failure. Once the quorum device has helped to determine which cluster is active, it fences off I/O from any node that does not belong to the cluster. This I/O fencing mechanism ensures that any failed nodes do not have access to the shared data.

   NetApp uses the storage system's LUN as a quorum device and iSCSI as the transport between Sun Cluster hosts and the NetApp storage systems.

   The software supplied by NetApp consists of two tools:

   - The first part accomplishes the fencing of the cluster nodes from the data exported by the storage system. This will be shipped in the form of a binary called NTAPfence.
   - The second part enables the usage of the storage system's LUN as the quorum device. This will be shipped in the form of a binary called NTAPquorum.

   NTAPfence and NTAPquorum will be supplied as part of the Solaris NTAPclnas package, which is available on NOW site at *http://now.netapp.com*. The package can be installed by using the Solaris **`pkgadd`** command:

   ```
   # pkgadd –d . NTAPclnas
   ```

   The binaries NTAPfence and NTAPquorum will be installed in the `/usr/sbin` directory. The package can be removed by using the Solaris **`pkgrm`** command:

   ```
   # pkgrm NTAPclnas
   ```

3. Support package solution:

   The high-availability configuration consists of two Solaris systems (Server A and Server B) running Sun Cluster connected via NFS through two Gigabit Ethernet switches to a NetApp storage system cluster.

   If one server (for example, Server B) fails, the other server (Server A) runs the support package. This package executes the NTAPfence tool, which rewrites the `/etc/exports` file on Storage system A and on Storage system B so that Server B cannot write to storage on either storage system.

4. Getting the support package:

   Download the NTAPclnas package, which contains two binaries, from NOW.

   - NTAPFence
   - NTAPQuorum

Do the following:

i. Log in to now.netapp.com, using netapp NOW username and password*.*

ii. In the Software Download section, click Download Software.

iii. In the Package for Sun Cluster and NetApp NFS *row*, select the Solaris platform and then click OK.

iv. In the Package for Sun Cluster and NetApp NFS 1.2 row, click View & Download.

v. In the Software Download Instructions section, click Continue.

vi. Accept the agreement.

vii. Download the package for your architecture—either Sparc or x86 and x64.

5. Installing the package:

   Install the NTAPclnas package in each cluster member:

   ```
   # pkgadd -d . NTAPclnas
   ```

6. Removing the package:

   To remove the support package the cluster member, enter:

   ```
   # pkgrm NTAPclnas
   ```

Planning the Sun Cluster Environment and Worksheet

Software Patches

Install the software patches required for Sun Cluster and Oracle.

IP Addresses

- One IP for public network.
- IP network multipathing groups—one primary IP address and one test IP address for each adapter in the group.

Logical Addresses

- One IP for checking the basic functionality of Sun Cluster.
- One IP for Oracle to check the physical storage system.
- One IP for Oracle to check the vFiler.

Sun Cluster Configurable Components

- Cluster Name
- Node Names
- Private Network
- Private Hostnames
- Cluster Interconnect
- Public Networks
- IP Network Multipathing Groups
- Quorum devices
- NFS Volume for Oracle

Worksheets

1. Example: Cluster and Node Names Worksheet

Component	Default	Actual
Cluster Name		Mycluster
Private Network Address	172.16.0.0	172.16.0.0
Private Network Mask	255.255.0.0	255.255.0.0
First-Installed Node Name		sun4500-svl02

Private Hostname	clusternode1-priv	sun4500-svl02-priv
Additional Node Name		sun4500-svl03
Private Hostname	clusternode2-priv	sun4500-svl03-priv
Additional Node Name		
Private Hostname		

2. Example: Cluster Interconnect Worksheet

Node Name	Adapter Name	Transport Type	Junction Name	Junction Type	Port Name
sun4500-svl02	qfe1	Dlpi	switch1	switch	1
Sun4500-svl02	qfe2	Dlpi	switch2	switch	1
Sun4500-svl03	qfe1	Dlpi	switch1	switch	2
Sun4500-svl03	qfe2	Dlpi	switch2	switch	2

3. Example: Public Network Worksheet

Component	Name
Node Name	sun4500-svl02
Primary Hostname	Sun4500-svl02
IP Network Multipathing Group	IPMP1
Adapter Name	qfe0
Backup Adapters (optional)	qfe3
Network Name	net-85
Secondary Hostname	Sun4500-svl03
IP Network Multipathing Group	IPMP1
Adapter Name	Qfe0
Backup Adapters (optional)	Qfe3
Network Name	net-86

4. Example: Local Devices Worksheet

Local Disk Name	Size
C0t0d0	8G
C0t1d0	8G

5. Example: NFS Client for the Cluster Member Worksheet (for Oracle)

Storage system	Exported Folder	Protocol	Options	Mount Point	On Boot
Filer1 [172.17.149.20]	/vol/data	NFS	rw,bg,hard,forcedirectio, nointr,rsize=32768, wsize=32768, proto=tcp,vers=3	/oradata	Yes
vFiler1 [172.17.149.23]	/vol/vfiler1 /qtree1	Nfs	rw,bg,hard,forcedirectio, nointr,rsize=32768, wsize=32768, proto=tcp,vers=3	/ora9vfiler	Yes

Note: Configure the IPMP for the public network, which is required for Sun Cluster 3.1 and optional for Sun Cluster 3.2.

Sun Cluster Installation

1. Check that the minimum patches are installed for Sun Cluster, such as 108987-16, 109147-28, 117000-05, 108993-39, 110934-21, 111111-04, 108434-18, and 108435-18.
2. Configure the systems and nodes time with the NTP server to sync the time.
3. Add the peer parameter in `/etc/inet/ntp.cluster`; for example:

   ```
   bash-2.03# cat /etc/inet/ntp.cluster | grep sun
   peer sun4500-svl03
   peer sun4500-svl02
   ```
4. For the graphical installation, enable the DISPLAY variable.
5. Log in as root user in one of the cluster nodes and start the installer without the GUI; for example:

   ```
   ./installer -nodisplay
   Agree the license
   select the language
   select the sun cluster 3.1
   select the oracle agent and other agents to install
   choose the configure later
   ```

   Repeat the procedure in the other cluster node. to here

Sun Cluster Configuration

1. Change the JAVA_HOME and JAVA variables to the newly installed j2se in `/etc/default/sccheck`:

   ```
   bash-2.03# cat /etc/default/sccheck  | grep -i java
   JAVA_HOME=/usr/j2se
   JAVA=${JAVA_HOME}/bin/java
   # Minimum acceptable java version
   #   for SC31.U3: java 1.4.x
   MIN_JAVA_MAJOR_VER=1
   MIN_JAVA_MINOR_VER=4
   ```

2. Make the same change in another cluster node and then reboot both nodes.
3. Check that the time of the cluster nodes and the systems are in sync.
4. Log in to one of the cluster nodes. Start the `scinstall` utility in interactive mode as root.

   In the following example, the user inputs are in bold:

   ```
   bash-2.03# scinstall
   Option: 1  [Install a cluster or cluster node ]
   Option: 1  [Install all nodes of a new cluster ]
   Do you want to continue (yes/no) [yes]?  yes [ This option is used to
   install and configure a new cluster.]
   Option: 1  [ Typical installation ]
   Node name:  sun4500-svl02
   Node name:  sun4500-svl03
   Node name (Control-D to finish):  ^D
   This is the complete list of nodes:
   sun4500-svl02
   sun4500-svl03
   Is it correct (yes/no) [yes]?  yes
   What is the name of the first cluster transport adapter (help) [qfe1]? qfe1
   What is the name of the second cluster transport adapter (help) [qfe2]?
   qfe2
   Do you want to disable automatic quorum device selection (yes/no) [no]? yes
   Is it okay to begin the installation (yes/no) [yes]?  yes
   Interrupt the installation for sccheck errors (yes/no) [no]?  no
   Configuring "sun4500-svl03" ... done
   Rebooting "sun4500-svl03" ... done
   Configuring "sun4500-svl02" ... done
   Rebooting "sun4500-svl02" ...
   ```

   When the *`scinstall`* utility finishes, the installation and configuration of the basic Sun Cluster software are complete. The cluster is now ready to configure Oracle to support high availability.
5. Register the storage system information.

   i. From any cluster node, add the device.

   - For Sun Cluster 3.2:

     ```
     # clnasdevice add -t netapp -p userid=root myfiler
     Please enter password
     -t netapp Enter netapp as the type of device you are adding.
     -p userid=root Enter the HTTP administrator login for the NAS device.
     myfiler - Enter the name of the NAS device you are adding.
     ```

     For example: **`#clnasdevice add –t netapp –p userid=root filer1`**

     You can use the same command for vFiler.
   - For Sun Cluster 3.1:

     ```
     # scnas -a -h myfiler -t netapp  -o userid=root
     Please enter password
     -a Add the device to cluster configuration.
     ```

```
-h myfiler Enter the name of the NAS device you are adding.
-o userid=root Enter the HTTP administrator login for the NAS device.
```

`Eg: # scnas -a -h filer1 -t netapp  -o userid=root`

You can use the same command for vFiler.

ii. Confirm that the device has been added to the cluster:

- For Sun Cluster 3.2:

  `# clnasdevice list`

- For Sun Cluster 3.1:

  `# scnas -p`

  For example:

```
bash-2.03# scnas -p
Filers of type "netapp":
    Filer name:                 filer1
        type:                   netapp
        password:               *******
        userid:                 root

    Filer name:                 172.17.149.20
        type:                   netapp
        password:               *******
        userid:                 root

    Filer name:                 vfiler1
        type:                   netapp
        password:               *******
        userid:                 root
```

iii. Using `scconf –aq` from the cluster node, add the device as quorum.

For example, set the quorum LUN on the storage system [filer1] as the quorum device. If the LUN_ID is 0 (zero), there is no need to explicitly specify `scconf`.

`# scconf -aq name=netapp,type=netapp_nas,filer=filer1,lun_id=0`

Or

add the device as quorum by using `scsetup`.

```
# scsetup
Option: 1 [ for quorum ]
Option: 1 [ Add a quorum ]
Is it okay to continue (yes/no) [yes]?  yes
What name do you want to use for this quorum device?  netapp
What is the name of the filer [netapp]?  filer1
What is the LUN id on the filer [0]?  0
Is it okay to proceed with the update (yes/no) [yes]?  yes
scconf -a -q name=netapp,type=netapp_nas,filer=filer1,lun_id=0
```

```
Command completed successfully.
```

iv. Check the quorum device by using `scstat –q`.

```
bash-2.03# scstat -q
-- Quorum Summary --
  Quorum votes possible:      3
  Quorum votes needed:        2
  Quorum votes present:       3

-- Quorum Votes by Node --

 Node Name           Present Possible Status
                    ---------           ------- -------- ------
  Node votes:       sun4500-svl03       1        1       Online
  Node votes:       sun4500-svl02       1        1       Online


-- Quorum Votes by Device --

                     Device Name         Present Possible Status
                     -----------         ------- -------- ------
   Device votes:     netapp              1        1       Online
```

NetApp Storage (NFS) Volume for Oracle

1. Create a new volume on NetApp storage.
   - Make sure that the Data ONTAP version is 7.2 or higher and that it is NFS licensed.
   - Create and export the volume for storing Oracle Database files on the storage.

     For example, create the volume *data* in *aggr1* aggregate with 58GB size:

     ```
     FAS-WHQL09> vol create data aggr1 58g
     ```

   Export the data volume with “rw and root” permission to sun4500-svl02 and sun4500-svl03 cluster nodes only using exportfs

   ```
   FAS-WHQL09> exportfs –p rw=sun4500-svl02:sun4500-svl03,root=sun4500-
   svl02:sun4500-svl03,anon=0 /vol/data
   ```

   Export the new volume

   ```
   FAS-WHQL09> exportfs –a
   ```

   Note: NetApp recommends using flexible volumes in your database environment. NetApp FlexVol® technology pools storage resources automatically and enables you to create multiple flexible volumes on a large pool of disks. This flexibility means that you can simplify operations, gain maximum spindle utilization and efficiency, and make changes quickly and seamlessly.

2. Add the exported directories from the NAS device to Sun Cluster.

   i. Add the directories from one for the cluster nodes.

   o For Sun Cluster 3.2:

```
# clnasdevice add-dir -d /export/dir1, /export/dir2 myfiler
```

-d /export/dir1, /export/dir2: Enter the directory or directories that you are adding.

myfiler: Enter the name of the NAS device containing the directories.

For example: **`clnasdevice add-dir -d /vol/data filer1`**

You can use the same command for vFiler.

- For Sun Cluster 3.1:

```
# scnasdir -a -h myfiler -d /vol/DB1 -d /vol/DB2
```

`-a` Add the directory or directories to the cluster configuration.

`-h myfiler` Enter the name of the NAS device whose directories you are adding.

`-d` Enter the directory to add. Use this option once for each directory that you are adding.

For example: `#scnasdir -a -h filer1 -d /vol/data`

You can use the same command for vFiler.

This value must match the name of one of the directories exported by the NAS device.

ii. Confirm that the device has been added to the cluster.

- For Sun Cluster 3.2:

```
# clnasdevice list
```

- For Sun Cluster 3.1:

```
# scnasdir –p
Eg: bash-2.03# scnasdir -p
Filers of type "netapp":
Filer name:            filer1
directories:           /vol/data
directories:           /vol/vfiler1
Filer name:            172.17.149.20
Filer name:             vfiler1
directories:           /vol/vfiler1/qtree1
```

iii. Create a folder to mount the exported folder.

```
#mkdir –p /path-to-mountpoint
```

For example: **`# mkdir  -p /oradata`**

```
# mkdir –p /ora9vfiler
```

iv. Make an entry in `/etc/vfstab` for the mount point.

v. Make an entry like the following for the mount point in `/etc/vfstab`:

```
172.17.149.20:/vol/data  /vol/data       /oradata               nfs    -
yes
rw,bg,hard,forcedirectio,nointr,rsize=32768,wsize=32768,proto=tcp,vers=3

172.17.149.23:/vol/vfiler1/qtree1  /vol/vfiler1/qtree1        /ora9vfiler
nfs    -  yes
rw,bg,hard,forcedirectio,nointr,rsize=32768,wsize=32768,proto=tcp,vers=3
```

Note: The `/vol/data` and `/vol/vfiler1/qtree1` volumes are mounted on the Sun Cluster nodes with proper mount options. In this situation, NetApp storage acts as an NFS server and Sun Cluster nodes act as NFS clients.

All cluster nodes mount the NFS file system from NetApp storage; if an active node fails, the other node detects the failure and starts the application processes on the standby node, thus maintaining the high availability. In this scenario, Sun Cluster HAStoragePlus is not used for an NFS mounted file system.

Installing and Configuring Sun Cluster for Oracle.

1. Install the Oracle software and create an Oracle Database and listener.

   Refer to the Oracle installation documentation.

2. Set up Oracle Database permissions for Sun Cluster.

   Enable access for the user and password to be used for fault monitoring.

   To use the Oracle authentication method for all of the supported Oracle releases, enter the following script at the sqlplus prompt:

   ```
   # sqlplus “/as sysdba”
   SQL> grant connect, resource to oravcs identified by xxxxxx;
   SQL> alter user oravcs default tablespace system quota 1m on system;
   SQL> grant select on v_$sysstat to oravcs;
   SQL> grant create session to oravcs;
   SQL> grant create table to oravcs;
   SQL> exit;
   ```

   To use the Solaris authentication method, grant permission for the database to use Solaris authentication.

   Note: The user for whom you enable Solaris authentication is the user who owns the files under the $ORACLE_HOME directory. The following code sample shows that the user oracle owns these files:

   ```
   # sqlplus “/as sysdba”
   SQL> create user ops$oracle identified by externallydefault tablespace
      system quota 1m on system;
   SQL> grant connect, resource to ops$oracle;
   SQL> grant select on v_$sysstat to ops$oracle;
   SQL> grant create session to ops$oracle;
   SQL> grant create table to ops$oracle;
   SQL> exit
   ```

3. Install the Sun Cluster HA packages for Oracle by using the scinstall utility.

   i. Load the Sun Java Enterprise System Accessory CD Volume 3 into the CD-ROM drive.

   ii. Run the `scinstall` utility with no options.

   This step starts the `scinstall` utility in interactive mode.

   iii. Select the menu option Add Support for New Data Service to This Cluster Node. The `scinstall` utility prompts you for additional information.

   iv. Provide the path to the Sun Java Enterprise System Accessory CD Volume 3.

   The utility refers to the CD as the “data services cd.”

   v. Specify the data service to install.

   The `scinstall` utility lists the data service that you selected and prompts you to confirm your choice.

   vi. Exit the scinstall utility.

   vii. Unload the CD from the drive

4. Register and configure Sun Cluster HA for Oracle.

i. Log in to the server as super user.

ii. Register the resource types for the data service:

For Oracle, you must register two resource types: SUNW.oracle_server and SUNW.oracle_listener:

```
# scrgadm –a –t SUNW.oracle_server
```

```
# scrgadm –a –t SUNW.oracle_listener
```

iii. Create a failover resource group to hold the network and application resources:

```
# scrgadm  -a –g resourcegroup-name –h node-list
```

```
-g resourcegroup-name – specify the name of the resource group
```

```
-h node-list – comma separated list of physical node names, this parameter
is optional
```

For example: **# scrgadm –a –g oracle-rg –h sun45-svl03, sun4500-svl02**

iv. Check the `/etc/inet/hosts` for network resource IP and hostname for resolve.

v. Add a network resource to the failover resource group:

```
# scrgadm -a -L -g resource-group -l logical-hostname [-n netiflist
```

```
-l logical-hostname – specify the network resource name
```

For example: **scrgadm –a –K –g oracle-rg –l oracle-ip**

There is no need for the `SUNW.HAStoragePlus` resource type for storage access, as it will be managed by the NetApp support package.

vi. Add the Oracle application resources to the resource group:

```
# scrgadm -a -j oracle-server-1 -g oracle-rg -t SUNW.oracle_server -x
ORACLE_HOME=/oracle/OraHome1 -x
Alert_log_file=/oradata/ora9i/admin/ora9db5/bdump/alert_ora9db5.log  -x
ORACLE_SID=ora9db5 -x Connect_string=oravcs/xxxxxx
```

```
# scrgadm -a -j oracle-listener-1 -g oracle-rg -t SUNW.oracle_listener -x
ORACLE_HOME=/oracle/OraHome1 -x LISTENER_NAME=LISTENER
```

vii. Bring the resource group online and enable fault monitoring:

```
# scswitch –Z –g resource-group
```

```
-Z  Enables the resource and monitor, moves the resource group
```

```
to the MANAGED state, and brings it online.
```

```
-g resource-group     Specifies the name of the resource group
```

For example: **scswitch –Z –g oracle-rg**

viii. Check the resources online by using `scstat –g`:

For example: **bash-2.03# scstat –g**

```
Resource Groups and Resources --

Group Name                    Resources

----------                    ---------

Resources: oracle-rg          oracle-ip oracle-server-l oracle-listener-l

Resources: oraclevf-rg        oraclevf-ip oraclevf-server-1 oraclevf-
                              listener-1

Resource Groups --
```

```
     Group Name          Node Name           State
     ----------          ---------           -----
Group: oracle-rg           sun4500-svl03       Offline
Group: oracle-rg           sun4500-svl02       Online

Group: oraclevf-rg         sun4500-svl03       Offline
Group: oraclevf-rg         sun4500-svl02       Online

Resources --
Resource Name       Node Name           State     Status Message
-------------       ---------           -----     --------


Resource: oracle-ip           sun4500-svl03       Offline   Offline
Resource: oracle-ip           sun4500-svl02       Online    Online -
LogicalHostname online.


Resource: oracle-server-l     sun4500-svl03       Offline   Offline
Resource: oracle-server-l     sun4500-svl02       Online    Online


Resource: oracle-listener-l   sun4500-svl03       Offline   Offline
Resource: oracle-listener-l   sun4500-svl02       Online    Online


Resource: oraclevf-ip         sun4500-svl03       Offline   Offline
Resource: oraclevf-ip         sun4500-svl02       Online    Online -
LogicalHostname online.


Resource: oraclevf-server-1   sun4500-svl03       Offline   Offline
Resource: oraclevf-server-1   sun4500-svl02       Online    Online


Resource: oraclevf-listener-1 sun4500-svl03       Offline   Offline
Resource: oraclevf-listener-1 sun4500-svl02       Online    Online
```

ix. Repeat the procedure for the vFiler Oracle instance.

5. Verify the Sun Cluster HA for Oracle installation.

   Perform the following verification tests to make sure that you have correctly installed Sun Cluster HA for Oracle.

   These sanity checks ensure that all of the nodes that run Sun Cluster HA for Oracle can start the Oracle instance and that the other nodes in the configuration can access the Oracle instance. Perform these sanity checks to isolate any problems in starting the Oracle software from Sun Cluster HA for Oracle.

i. Log in as `oracle` to the node that currently masters the Oracle resource group.

ii. Set the environment variables ORACLE_SID and ORACLE_HOME.

iii. Confirm that you can start the Oracle instance from this node.

iv. Confirm that you can connect to the Oracle instance.

v. Use the **`sqlplus`** command with the user/password variable that is defined in the `connect_string` property.

```
# sqlplus username/password@tns_service
```

vi. Shut down the Oracle instance.

The Sun Cluster software restarts the Oracle instance because the Oracle instance is under Sun Cluster control.

vii. Switch the resource group that contains the Oracle Database resource to another cluster member.

The following example shows how to complete this step:

```
# scswitch -z -g resource-group -h node
```

For example: `#scswitch -z -g oracle-rg  -h sun4500-svl03`

viii. Log in as `oracle` to the node (sun4500-svl03) that now contains the resource group.

ix. Repeat steps 3 and 4 to confirm interactions with the Oracle instance.

10.6 TEST SCENARIOS AND FAQ

This section answers some frequently asked queries on Sun Clusters and various test scenarios that were executed upon successful build of the solution discussed earlier in this document. The test scenarios include various component failures, including server hardware, network, storage system, and so on. Unless stated otherwise, the environment was reset to the normal running state before each test. The normal running state had all the Sun Cluster nodes operational and the ora9db5 database active on Site A (SUN4500-WHQL01) and ora9vf1 database active on Site B (SUN4500-WHQL02).

COMPLETE LOSS OF POWER TO DISK SHELF

No single point of failure should exist in the scenario. Therefore the loss of an entire disk shelf was tested. This test was accomplished by simply turning off both power supplies while a load was applied.

Task	Power off the FAS-WHQL09 shelf. Observe the results and then power it back on.
Expected/Observed Results	Relevant disks go offline, plex is broken, but service to clients (availability and performance) is unaffected. When power is returned to the shelf, the disks are detected and a resync of the plexes occurs without any manual action.
Sun Cluster Results	Oracle services to clients were not interrupted. No failure or failover of Oracle resources.

LOSS OF ONE LINK ON ONE DISK LOOP

No single point of failure should exist in this scenario. Therefore the loss of one disk loop was tested. This test was accomplished by removing a fiber patch lead from one of the shelves.

Task	Remove fiber entering FAS-WHQL09 Pool0, ESH A. Observe the results and then reconnect the fiber
Expected/Observed Results	A controller message displays that some disks are connected to only one switch, but service to clients (availability and performance)unaffected. When the fiber is reconnected, a controller message is displayed that disks are now connected to two switches.
Sun Cluster Results	Oracle services to clients were not interrupted. No failure or failover of Oracle resources.

LOSS OF A BROCADE SWITCH

No single point of failure should exist in this scenario. Therefore the loss of an entire Brocade switch was tested. This test was accomplished by simply removing the power cord from the switch while Oracle was running.

Task	Power off the WHQL10-SW4 Fibre Channel switch. Observe the results and then power it back on.
Expected/Observed Results	A controller message displays that some disks are connected to only one switch and that one of the cluster interconnects is down. but service to clients (availability and performance) is unaffected. When the power is restored and the switch completes its boot process, a controller message is displayed to indicate that the second cluster interconnect is again active.
Sun Cluster Results	Oracle services to clients were not interrupted. No failure or failover of Oracle resources.

LOSS OF ONE ISL

No single point of failure should exist in the solution. Therefore the loss of one of the interswitch links (ISLs) was tested. This test was accomplished by simply removing the fiber connection between two of the switches while a load was applied.

Task	Remove the fiber between WHQL09-SW1 and WHQL10-SW3.
Expected/Observed Results	A controller message is displayed that some disks are connected to only one switch and that one of the cluster interconnects is down, but service to clients (availability and performance) is not affected. When ISL is reconnected, a controller messages is displayed to indicate that the disks are now connected to two switches and that the second cluster interconnect is again active.
Sun Cluster Results	Oracle services to clients were not interrupted. No failure or failover of Oracle resources.

FAILURE OF A CONTROLLER

No single point of failure should exist in the solution. Therefore the loss of one of the controllers itself was tested.

Task	Power off the running FAS-WHQL09 controller.
Expected/Observed Results	As a result of the change of processing from one controller to the other, host interruption should be minimal if any, because the failover is masked by the “disk time out” value in Oracle, which is set to 200 seconds by default. Failure of database should not occur.
Sun Cluster Results	Oracle services to clients were not interrupted. No failure or failover of Oracle resources.

FAILBACK OF A CONTROLLER

As a follow up to the previous test, the data serving must be failed back to the previously failed controller to return to the normal operating state. This test was accomplished by issuing a command on the surviving controller to request that processing be returned to the previously failed controller.

Task	Power on FAS-WHQL09. Issue a `cf giveback` command on FAS-WHQL10 to cause the failback to occur.
Expected/Observed Results	As a result of the change of processing from one controller to the other, host interruption should be minimal if any, because the failover is masked by the “disk time out” value in Oracle, which is set to 200 seconds by default. Failure of database should not occur.
Sun Cluster Results	Oracle services to clients were not interrupted. No failure or failover of Oracle resources.

CRASH OF A NODE IN THE ACTIVE CLUSTER MEMBER

To test the availability of the Sun Cluster setup, we powered off one of the nodes in the Sun Cluster.

Task	Power off sun4500-svl02.
Expected/Observed Results	The host crashes, and the remaining cluster nodes realize what has happened and transfer the resources to one or more surviving nodes. Clients operations fail. The failover should happen to the passive node..
Sun Cluster Results	The Oracle resource groups failed over to the passive node; Oracle services to clients resumed.

FAILURE OF A LAN CONNECTION ON THE ACTIVE NODE

To test the availability of the Sun Cluster solution, we removed the LAN (public) interface from a node on the active cluster node, which has the Oracle resource group.

Task	Remove the LAN cable from the LAN (public) interface on node sun4500-svl02.
Expected/Observed Results	The cluster realizes that the interface resource has gone offline and relocates the cluster group on a surviving node. Oracle client operations fail. The Oracle client has to reconnect to the Oracle server to perform operations.
Sun Cluster Results	The Oracle resource groups failed over to the passive node. Oracle services to clients resumed.

FAILURE OF HEARTBEAT CONNECTION ON THE ACTIVE NODE

To test the availability of the Sun Cluster, we removed the link used for the cluster heartbeat connection on an active node.

Task	Move oracle-rg to sun4500-svl02; remove the cables from the heartbeat interface.
Expected/Observed Results	No failover occurs because there are multiple interfaces that can be used for heartbeat traffic.
Sun Cluster Results	Oracle services to clients were not affected. No failure or failover of Oracle resources.

LOSS OF ENTIRE SITE: DISASTER DECLARED

To test the availability of the overall solution, we simulated loss of an entire site.

Task	Test the failure of the FAS-WHQL09 site by interrupting the following components in this order, in rapid succession: **Step 1: Simulate failure.** i. Remove both ISLs. ii. Remove power to the active cluster node. iii. Remove power from FAS-WHQL09 and disk shelves. **Step 2: Recovery.** i. Declare the disaster and perform a takeover at the surviving site, Site B. Issue the following command on FAS-WHQL10: `FAS-WHQL10> cf forcetakeover -d` ii. Use the partner command on FAS-WHQL10 to access FAS-WHQL09 (now running on the same controller as FAS-WHQL10). iii. Start the cluster service on one of the nodes on FAS-WHQL10. iv. To speed further steps, take all resource groups offline and then set online only the disk resources. Ensure that all disks are online. v. To test disk connections, set group ownership for the ‘data’ node on FAS-WHQL10. Disks should be transferred and come online, while other resources

	should remain offline. Set the ownership to the node on FAS-WHQL10 for the Quorum cluster group; this also verifies that the node is fully functional. vi. Start the cluster service on the remaining node on FAS-WHQL10.
Expected/Observed Results	The cluster initially goes offline because Quorum disk access is not possible. When the FAS-WHQL10 **`takeover`** command is issued, the steps of connecting any disk resources on the FAS-WHQL09 controller should be completed, allowing the cluster resources to come online. Obviously there should be no loss of data or corruption.
Sun Cluster Results	Service to clients was affected when the cluster went offline. Oracle operations were successful after evicting the failed node.

RESTORE OF ENTIRE SITE/RECOVER FROM DISASTER

To test the availability of the overall solution, we simulated recovery after loss of an entire site.

Task	i. Reconnect the ISL between sites so that FAS960-WHQL10 can see the disk shelves from FAS960-WHQL09. After connection, the FAS960-WHQL10 Pool1 volumes automatically begin to resync. ii. Perform the following steps on the cluster node in site FAS960-WHQL09. iii. Individually power on the cluster node. Verify that the cluster services start correctly and that the node has become part of the cluster; then power up the next node, and so on. iv. Power on the controller (FAS-WHQL09). v. Set all resources in data and quorum offline, set only the disk resources online, and then move the resource group to the FAS960-WHQL09 node to verify connectivity. vi. When all cluster nodes in the cluster are online, turn on FAS960-WHQL09. Use the cf status command to verify that a giveback is possible, and use cf giveback to fail back. Once FAS960-WHQL09 is online, manually start the resync of volumes by using: **`vol mirror <good volname> -v <outdated volname>`** For example: `vol mirror data -v data(1)` For example: `vol mirror quorum -v quorum(1)`
Expected/Observed Results	On the cluster giveback to the Site A controller, the results should be similar to the normal giveback. There should be no downtime during the failback process.
Sun Cluster Results	Oracle services were unaffected; failed nodes booted successfully after restoring power on the nodes in Site A and were back online.

QUORUM DEVICE ADD PROBLEM

Problem: When the quorum device is added by using `scconf –a –q`, the error message "`Failed to add quorum device (netapp) _ cannot scrub the device, check the device configuration`" appears.

Solution: Check the NTAPClans support package installed in the cluster nodes.

Problem: When the quorum device is added by using using `scconf –a –q`, the error message "`scconf: Failed  to add quorum device (netapp) – invalid quorum, Error in controller LUN (or) igroup configuration`" appears.

Solution: Unmap the LUN, remove the quorum LUN and igroup, and recreate the igroup. Create the LUN and then remap the LUN to igroup.

Problem: The quorum igroup member in the controller always shows "`not logged in`" for the quorum device. Is the iSCSI configuration wrong?

Solution: The agent does not retain the session; if needed that time only, the agent checks the quorum status and logs out from the cluster node.

REMOVING THE QUORUM DEVICE FROM THE CLUSTER

To remove the quorum device from the cluster, enter:

```
# scconf –r –q installmode

# scconf –r –q name=netapp

# scstat –q
```

REBOOTING SOLARIS WITHOUT THE CLUSTER

To restart solaris without the cluster, do one of the following:

- Enter:a
  ```
  # reboot -- -x
  ```
- Boot from `ok` prompt as:
  ```
  ok boot -x
  ```

10 AKNOWLEDGEMENTS

The author would like to thank the following individuals for their contributions to the testing process and to this technical report:

Uday Shet—Senior Manager, TME, NetApp

Michelle Nguyen—Senior Product Manager, Enterprise Applications, NetApp

DISCLAIMER

NetApp provides no representations or warranties regarding the accuracy, reliability, or serviceability of any information or recommendations provided in this publication or with respect to any results that might be obtained by the use of the information or observance of any recommendations provided herein. The information in this document is distributed AS IS, and the use of this information or the implementation of any recommendations or techniques herein is a customer's responsibility and depends on the customer's ability to evaluate and integrate them into the customer's operational environment. This document and the information contained herein must be used solely in connection with the NetApp products discussed in this document.

Please send any errors, omissions, differences, new discoveries, and comments about this paper to nkarthik@netapp.com, suresh.vundru@netapp.com .

www.netapp.com