flashgrid

FlashGrid® Storage Fabric

version 25.05

Deployment Guide

for on-premises or private cloud deployments

rev. 25.05-2025.06.16

Table of Contents

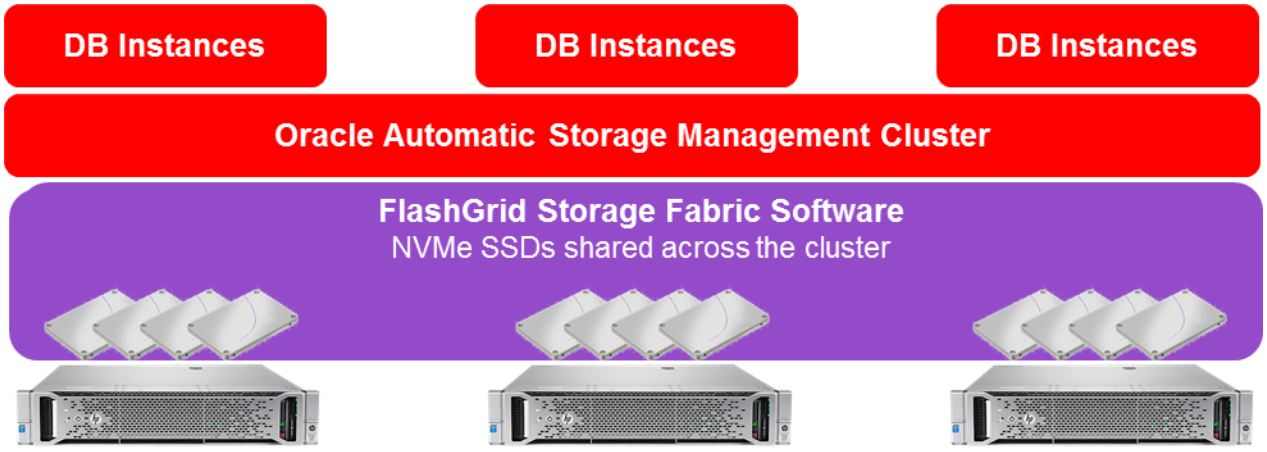
1 Introduction

This document is intended for system and database administrators who deploy and maintain storage for Oracle RAC based on FlashGrid Storage Fabric. The FlashGrid Storage Fabric software enables open and flexible architecture for Oracle RAC that is equally efficient in both small and large clusters.

For information specific to cloud deployments please refer to FlashGrid Help Center.

FlashGrid Storage Fabric highlights

- Primary shared storage based on standard NVMe PCIe SSDs, SAS SSDs, or locally attached virtual disks
- Physical storage located inside the database nodes (converged nodes) or in separate storage nodes
- VMs or standard x86 servers used as database and storage nodes
- FlashGrid Storage Fabric software manages SSD devices and connectivity, integrates with Oracle ASM
- Oracle ASM manages data, volumes, mirroring, snapshots
- 2-way or 3-way mirroring of data across different nodes
- Ethernet (physical or virtual) for network connectivity
- FlashGrid Read-Local Technology minimizes network overhead by serving reads from locally attached disks



2 Compatibility

- Oracle Linux 7, 8, or 9; Red Hat Enterprise Linux 7, 8, or 9
- Oracle Grid Infrastructure 19c. It's recommended to install with the latest Release Update.
- Oracle Database 19c, 18c, 12.2.0.1, 12.1.0.2, or 11.2.0.4. The latest Release Update / Patch Set Update is strongly recommended.

3 Overview of the Process of Deploying an Oracle RAC Cluster on FlashGrid Storage Fabric

The following steps outline the process of deploying FlashGrid. More detailed information for each step is provided in the subsequent sections of the guide.

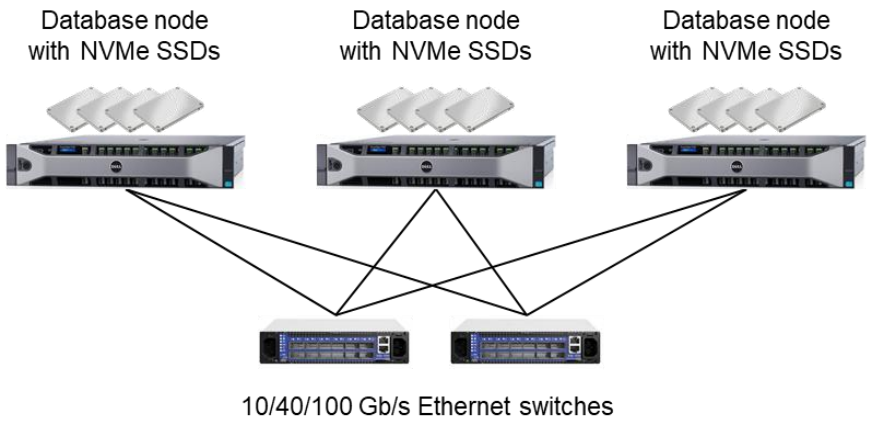
1. Get familiar with the FlashGrid architecture and determine the following:
    - Storage inside the database nodes (hyper-converged architecture) or in separate storage nodes
    - The number of database nodes and storage nodes
    - The number and type (Normal Redundancy or High Redundancy) of ASM disk groups
    - Placement of Grid Infrastructure files (GRID disk group) – on a FlashGrid disk group or on an external storage
    - The number of quorum disks required and their location
2. Prepare each node of the cluster
    a. Install and configure OS
    b. Install the FlashGrid software
    c. Configure network
    d. If required, create LVM volumes for quorum disks and/or disks for GRID disk group
    e. If required, configure external storage for quorum disks and/or disks for GRID disk group
3. Configure the FlashGrid cluster
4. If GRID disk group uses FlashGrid disks then
    a. Install Grid Infrastructure software in Software-Only mode
    b. Apply the latest Release Update to GI home on all nodes
    c. Configure Grid Infrastructure cluster
    d. During GI configuration create the GRID disk group using temporary GRID disks
    e. After GI installation replace the temporary disks with permanent GRID disks using flashgrid-fix-grid-dg-ca tool
5. Create ASM disk group(s) for data
6. Create database(s) or ACFS file system(s) on the disk group(s)

Note: Grid Infrastructure can be installed before configuring FlashGrid if an external storage is used for the GRID disk group.

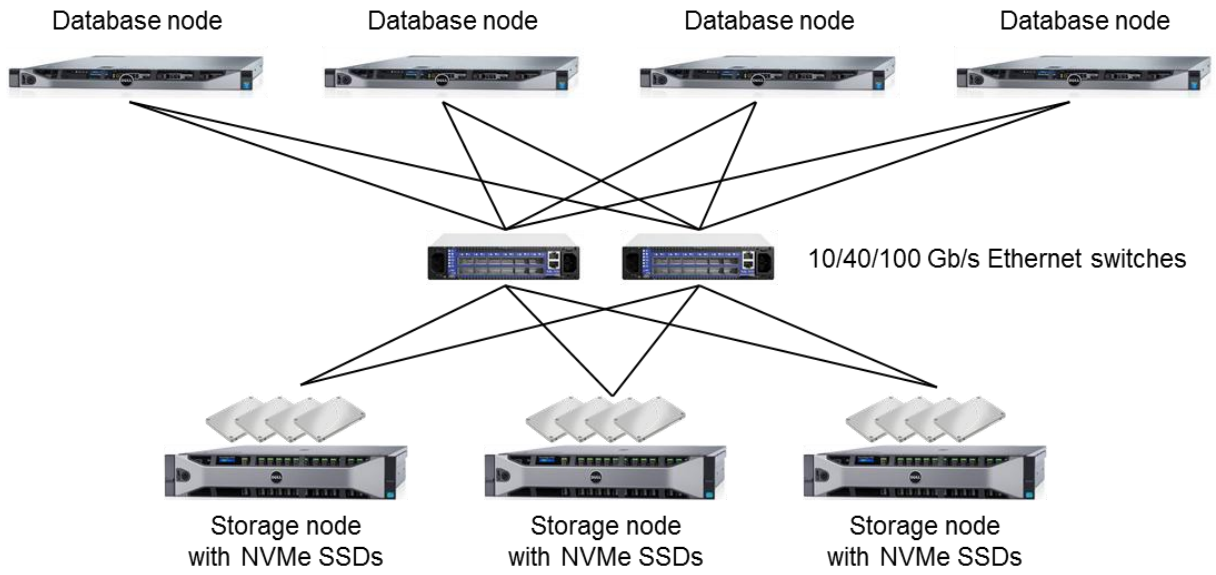
4 FlashGrid Storage Fabric Architecture

4.1 Hyper-converged architecture or separate storage nodes

With 2-node or 3-node clusters a converged configuration with storage located inside the database nodes is usually optimal. The following picture shows an example of such configuration with three database nodes.

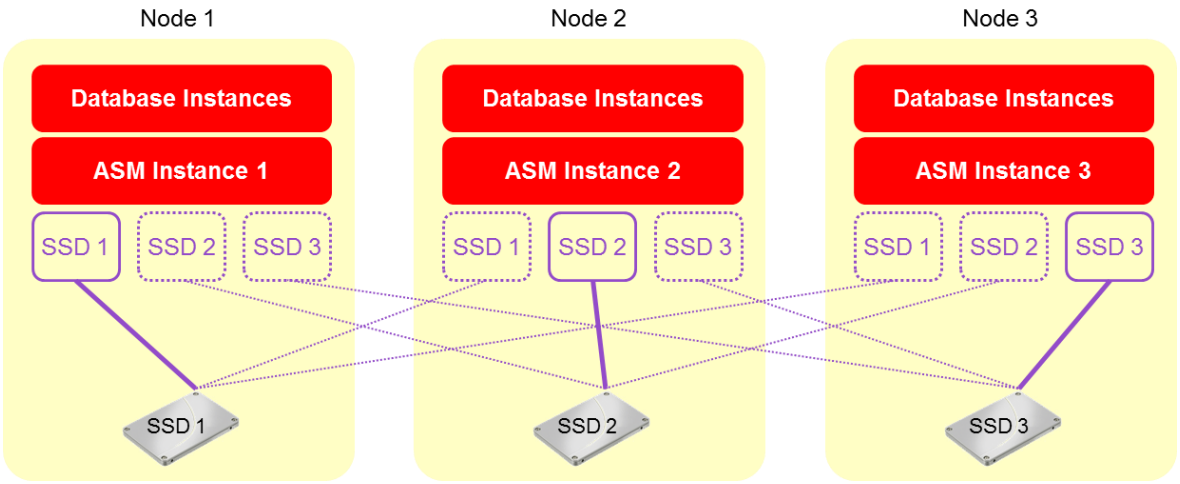


Placing storage in separate dedicated storage servers may be preferred in clusters with 4+ database nodes or if the database nodes do not have enough room for SSDs, for example, with blades or 1U database servers.



4.2 Shared Access

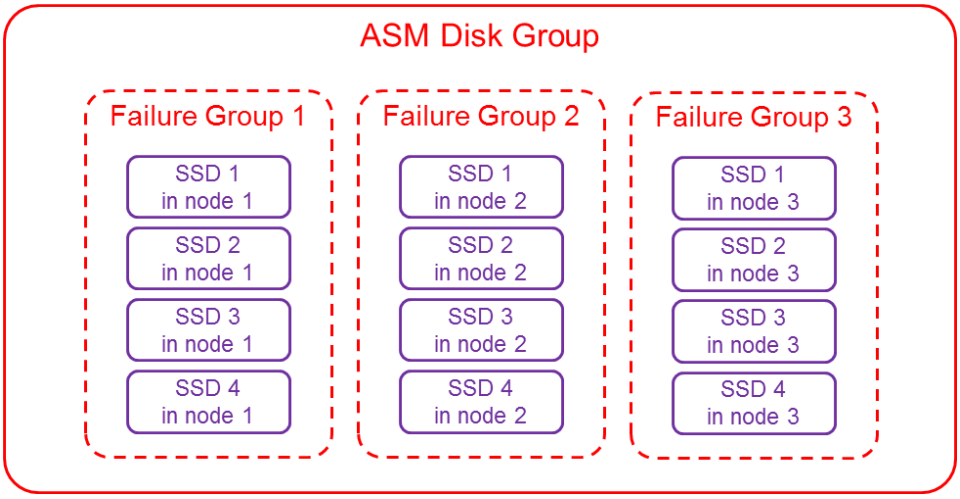
With the help of FlashGrid software each ASM instance can access each of the disks in the cluster. Each disk is visible in the OS as `/dev/flashgrid/nodename.diskname` device where *nodename* is the name of the node where the SSD is physically located.



4.3 Data Mirroring

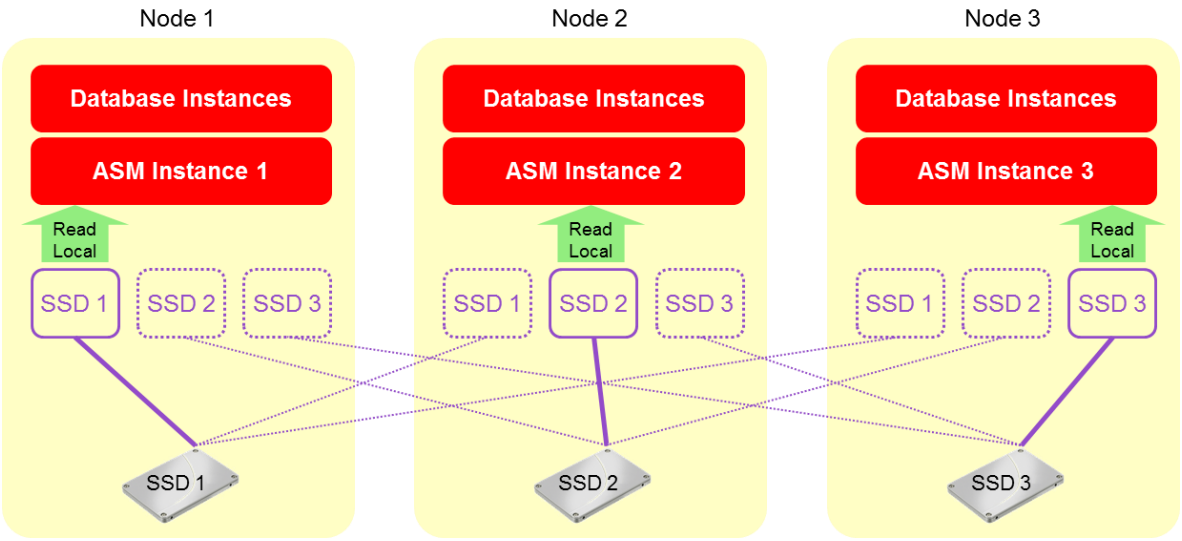
The FlashGrid architecture leverages Oracle ASM's existing capabilities for mirroring data. In Normal Redundancy mode each block of data has two mirrored copies. In High Redundancy mode each block of data has three mirrored copies. Each ASM disk group is divided into failure groups – one failure group per node. Each disk is configured to be a part of a failure group that corresponds to the node where the disk is physically located. ASM ensures that mirrored copies of each data block are placed in different failure groups.

In Normal Redundancy mode the cluster can withstand loss of one (converged or storage) node without interruption of service. In High Redundancy mode the cluster can withstand loss of two (converged or storage) nodes without interruption of service.



4.4 FlashGrid Read-Local™ Technology

In hyper-converged clusters the read traffic can be served from local SSDs at the speed of the PCIe bus instead of travelling over the network. In 2-node clusters with 2-way mirroring or 3-node clusters with 3-way mirroring 100% of the read traffic is served locally because each node has a full copy of all data. Because of the reduced network traffic the write operations are faster too. As a result, even 10 GbE network fabric can be sufficient for achieving outstanding performance in such clusters for both data warehouse and OLTP workloads. For example, a 3-node cluster with four NVMe SSDs per node can provide 30 GB/s of read bandwidth, even on a 10 GbE network.



4.5 Strict Read-Local Mode

ASM does not allow reads from disks that are resynchronizing data (SYNCING state) after being offline. As a result, if database is running on a node whose local disks are in SYNCING state, all reads will be performed remotely over the network. In cloud based or extended distance clusters that have relatively low network bandwidth this may result in lower performance of the database instance on a node that has just rebooted and is still resynchronizing its data.

Strict Read-Local mode prevents such performance asymmetry between nodes. When the *Strict Read-Local* mode is enabled, a database instance start will be delayed until its local disks complete resynchronization.

Use the following commands to enable, disable, and show status of Strict Read-Local mode:

```
flashgrid-cluster strict-read-local-enable
flashgrid-cluster strict-read-local-disable
flashgrid-cluster strict-read-local-show
```

Note that enabling Strict Read-Local mode changes the setting only for existing databases. Re-running the enable command is required after creating new database(s).

Note that in order to unmount a disk group while Strict Read-Local mode is enabled, `srvctl stop diskgroup` command with `-force` option must be used. Example:

```
srvctl stop diskgroup -diskgroup DGNAME -node rac1,rac2 -force
```

4.6 Using External Storage in Combination with FlashGrid

Any type of external storage can be used in combination with FlashGrid storage including FC, FCoE, iSCSI, or NFS. The external storage can be used for storing data that does not require the tier-0 performance of the FlashGrid storage, e.g. Grid Infrastructure files, ASM quorum disks, Grid/Database home, backups, or archive logs. Separate ASM disk groups must be configured for any external storage.

4.7 Quorum Disks

In certain disk group configurations one or two additional quorum disks may be required depending on the number of nodes in the cluster. The quorum disks may be required even in disk groups that do not store Voting files. ASM uses quorum disks to store additional copies of metadata that can be used for arbitration in certain failure scenarios.

One quorum disks requires 1 GiB of space. The quorum disks generate very small amount of storage traffic and can be stored on any type of external shared storage. The quorum disks may be stored on storage that does not provide high availability or redundancy. However, the storage for the quorum disks must be external to the FlashGrid nodes used for data storage.

Options for quorum disk placement:

- Virtual disks on quorum node VMs
- LVM volumes on database nodes (for clusters with separate storage nodes)
- External iSCSI storage
- NFS

The following table shows how many quorum disks are required for a disk group depending on the disk group redundancy level and the number of converged or storage nodes in the cluster.

	2 nodes	3 nodes	4 nodes	5+ nodes
Normal Redundancy	1 quorum disk	not needed	not needed	not needed
High Redundancy	N/A	2 quorum disks	1 quorum disk	not needed

4.8 Selecting Location for Grid Infrastructure Files (GRID disk group)

Two main options are available for the GRID disk group that contains Voting files and OCR:

- **FlashGrid Storage Fabric managed disks.** This option allows reduced dependency on any external storage. In most cases, it is recommended to use LVM volumes on the system boot drives of the converged or storage nodes as disks for the GRID disk group. The use of LVM volumes eliminates the need for separate physical SSDs dedicated to the GRID disk group, thus making more SSD slots available for data.
- **An external storage: iSCSI/FC/FCoE SAN or NFS.** This option may be preferred if high-availability external storage is readily available and unification of GRID disk group management process with other non-FlashGrid systems is desirable. 5 GB capacity is sufficient in most cases with moderate performance requirements. The external storage must provide high availability.

4.9 Dependencies between FlashGrid and Oracle Services

The FlashGrid rpm installer creates a *systemd* dependency on *flashgrid_wait* service that delays OHAS/CRS start until all storage nodes in the FlashGrid cluster are online and all FlashGrid devices are connected. This dependency allows avoiding a situation where CRS tries to mount an ASM disk group before all storage devices are available. You can use `flashgrid-node stop-waiting` command to override this dependency and allow CRS to start while some FlashGrid devices are still not ready.

Note that if you try to start CRS manually while it is being blocked by the *systemd* dependency then subsequent attempts to start it may fail even after the dependency is cleared. If this happens, use `systemctl restart oracle-ohasd` command to start CRS.

4.10 Persistent Names and Permissions for ASM Disks

FlashGrid Storage Fabric software provides persistent naming, sets device permissions, and configures multipathing for ASM disks managed by FlashGrid. There is no need to use ASMlib or UDEV rules for regular ASM disks managed by FlashGrid, including external iSCSI disks configured in FlashGrid. However, quorum disks located on external storage not managed by FlashGrid require persistent names, permissions, and multipathing configured separately outside of FlashGrid.

5 Installing and Configuring OS

5.1 Enabling Performance mode in system BIOS

It is recommended to configure *Performance* mode for CPUs in the system BIOS. With default settings a CPU can get into a power-saving mode that causes undesirable latencies.

5.2 Reserving space for LVM volumes during OS installation

In most cases it is recommended to reserve 300 GiB of space on the system drive during OS installation. This reserved space can be used later for creating LVM volumes for quorum disks or for the GRID disk group. The space can be reserved in the default LVM volume group, as a separate partition, or as a separate hardware RAID volume. In case of a virtual node, an additional virtual disk can be used instead.

5.3 Disabling swap

Swapping may happen when system is low on memory even when there is some physical memory still available. Swapping will cause system being unresponsive and potentially causing time outs at various levels including storage. While swapping may prevent or delay running out of memory, the results of swapping are likely to be worse than the results of running out of memory. Even if PGA aggregate limit/target parameters are carefully configured, FlashGrid recommends completely disabling swap to prevent failures resulting from potential mistakes in memory configuration.

5.4 Installing *kernel-devel* package

Installing *kernel-devel* package corresponding to the active kernel version is required for operation of *flashgrid-node-monitor* service.

```
# yum install kernel-devel-`uname -r`
```

5.5 Setting Performance mode for CPU governor

It is recommended to set *Performance* mode for the CPU governor service. This will guarantee that the CPUs are always running at their maximum frequency and will help reduce latencies.

```
# cpupower frequency-set --governor performance
```

5.6 Synchronizing clocks with NTP

System clocks must be within 30 seconds between all nodes in the cluster. Configuring CHRONYD or NTPD service is recommended before configuring FlashGrid cluster. Configure CHRONYD or NTPD service according to the Oracle Grid Infrastructure requirements.

5.7 Configuring ASM device owner user and group

Before configuring FlashGrid Storage Fabric make sure that owner (e.g. 'oracle' or 'grid') and group (e.g. 'asmadmin') for Grid Infrastructure are configured on all nodes where ASM will be installed. FlashGrid Storage Fabric will use these user and group in the disk device permissions.

5.8 Preparing a user account for cluster configuration operations

A user account used for configuring a FlashGrid Storage Fabric cluster must meet the following requirements:

- key based (passwordless) SSH access between all nodes of the cluster, including quorum nodes (required)
- sudo privileges for running any command without entering a password (required)
- *wheel* or device owner group (e.g. *asmadmin*) membership (recommended)

Creating a user named *fg* that meets the criteria above is recommended.

Key based SSH access can be configured using '`ssh-keygen`' command followed by '`ssh-copy-id username@nodename`' command for each node.

Example of configuring key based access to *fg* on three nodes (run as fg on each node):

```
$ ssh-keygen -t rsa
$ for i in node1 node2 node3; do ssh-copy-id -i ~/.ssh/id_rsa.pub fg@$i; done
```

6 Installing FlashGrid Storage Fabric Software

The FlashGrid Storage Fabric software is provided as an RPM package. Additionally, installation of FlashGrid Patch for iscsid, OS configuration, Diagnostics, and Health Checker RPMs is required. All RPMs must be installed on every node in a FlashGrid cluster including converged nodes, storage nodes, database nodes, and quorum server nodes.

To install the FlashGrid Storage Fabric software, complete the following steps on each node

1. Request latest versions of the following RPMs from FlashGrid support:
   - FlashGrid Python: ***flashgrid-python***
   - FlashGrid Patch for iscsid: ***flashgrid-iscsid-fix***
   - FlashGrid OS configuration: ***flashgrid-os-conf***
   - FlashGrid OS on premises configuration: ***flashgrid-os-conf-onprem***
   - FlashGrid Storage Fabric: ***flashgrid-sf***
   - FlashGrid Diagnostics: ***flashgrid-diags***
   - FlashGrid Health Checker: ***flashgrid-health-check***
2. Use YUM to install the downloaded RPMs and their dependencies:

```
# yum install flashgrid-python-*.rpm
# yum install flashgrid-iscsid-fix-*.rpm
# yum install flashgrid-os-conf-*.rpm
# yum install flashgrid-sf-*.rpm
# yum install flashgrid-diags-*.rpm
# yum install flashgrid-health-check-*.rpm
```

7 Installing License File

The license file specifies a type of the FlashGrid license and the FlashGrid support plan for the cluster. If the license file is not installed or is invalid or expired then the *flashgrid-cluster* command will be showing a warning and alert emails will be sent periodically.

If you have a license file available then place it as `/etc/flashgrid.license` on each node of the cluster. If you do not have a license file available then please contact sales@flashgrid.io

8 Configuring Storage Network

8.1 Sharing Network between FlashGrid and Oracle Private Network

Use of separate NICs for Oracle private network and for FlashGrid Storage Fabric is recommended in most cases. FlashGrid supports sharing network interfaces with Oracle private network. However, in case of RAC deployments the performance impact on Cache Fusion operation must be carefully assessed.

8.2 Storage network configuration requirements

- Equal number of network interfaces must be configured on all converged/database/storage nodes of a cluster
  - Quorum server nodes can have a different number of network interfaces, typically only one.
  - Exceptions are possible, but require additional configuration steps. Contact FlashGrid for assistance.
- Network interfaces must have the same names on all converged/database/storage nodes of a cluster
  - Quorum server nodes can have network interfaces with different names.
  - Exceptions are possible, but require additional configuration steps. Contact FlashGrid for assistance.
- Each network interface must have a static IPv4 address
- Network interfaces with the same name (e.g. 'em2') must belong to the same IPv4 subnet.
- Routing between the subnets must be disabled – do not configure gateways.
  - If a Quorum server node has to be on a separate subnet then additional configuration steps are required. Contact FlashGrid for assistance.
- IPv4 multicast must be enabled within each subnet
- Use of jumbo frames (MTU=9000) is recommended. Before enabling jumbo frames on NICs, need to verify that switches also are configured to support jumbo frames.
- The following ports must be open between all nodes of a cluster:
  - TCP 3260
  - TCP 5557
  - TCP 8753
  - TCP 22 (or a custom SSH port)

8.3 Configuring storage network redundancy

Two main options are recommended for ensuring network redundancy:

- Storage NICs connected to two (or more) separate network switches.
  NICs connected to each switch must be configured with separate subnet and VLAN. No routing must be enabled between those subnets.

- Storage NICs connected to two (or more) stacked network switches with LACP aggregation.
  In this scenario network link redundancy and load balancing is performed at the switch level. A single virtual network interface per node is configured in FlashGrid.

8.4 Configuring network across two geographically separated sites

For geographically separated sites one non-routable VLAN per network interface must be configured. Each VLAN must span all sites and include a separate set of physical switches to ensure network connection redundancy. Extra care must be taken to ensure IP multicast is enabled within each VLAN across all sites.

If enabling multicast between the two sites is problematic then FlashGrid® Cloud Area Network™ software can be configured to provide the multicast capability. Contact your FlashGrid technical representative for more information.

8.5 Configuring NICs

It is recommended to create NIC configuration files *ifcfg-<device>* in `/etc/sysconfig/network-scripts/` directory without using Network Manager. Use MAC address to assign persistent device names and corresponding static IP addresses.

Example of good NIC device names: *pub*, *priv1*, *priv2*, *storage1*, *storage2*

Example of a manually created configuration file:

```
HWADDR=XX:XX:XX:XX:XX:XX
DEVICE=storageN
IPADDR=192.168.N.Node#
PREFIX=24
BOOTPROTO=none
DEFROUTE=no
IPV4_FAILURE_FATAL=yes
IPV6INIT=no
ONBOOT=yes
NM_CONTROLLED=no
TYPE=Ethernet
MTU=9000
```

8.6 Configuring Enhanced Network Capabilities with FlashGrid® Cloud Area Network™

In some on-premises environments existing network capabilities may be insufficient for running FlashGrid Storage Fabric or Oracle RAC. In such cases it is possible to use FlashGrid Cloud Area Network (CLAN) software to provide the missing network capabilities on top of the existing physical network. Examples of when FlashGrid CLAN software may be required:

- Quorum server in a VM connected to the database servers via public network
- An extended distance cluster with limited network capabilities between sites (e.g. shared network links or lack of multicast).

If the FlashGrid CLAN capabilities are required then follow instructions below in this section for installing and configuring it. If the existing network capabilities are sufficient and FlashGrid CLAN is not required then skip the rest of this section.

8.6.1 Installing FlashGrid CLAN software

The FlashGrid CLAN software is provided as a single RPM package. The RPM must be installed on every node in a FlashGrid CLAN cluster including database nodes, storage nodes, quorum nodes, and CLAN client nodes.

To install the FlashGrid CLAN software, complete the following steps on each node

1. Request the latest version of FlashGrid CLAN software RPM from FlashGrid support.
2. Use YUM to install the downloaded RPMs and their dependencies:

   ```
   # yum install flashgrid-clan-*.rpm
   ```

8.6.2 Creating CLAN configuration file

This subsection describes creating a configuration file for connecting quorum server(s) to the database servers. If you need CLAN for an extended distance cluster then contact FlashGrid support for assistance with creating a configuration file.

Quorum server hosts or VMs are typically connected to database servers via low-speed public network. To avoid routing storage traffic via the public network, the corresponding public NICs should not be included in the FlashGrid Storage Fabric configuration. Instead, point-to-point virtual links can be created between the database nodes and the quorum server using the CLAN software.

Below is an example of a CLAN configuration file for a 2-node RAC cluster with one quorum node connected to the database nodes via a public network. The highlighted public IP addresses and public NIC device names must be customized. In most cases there is no need to change node names, role names, or other parameters.

```
[clan]

ssh_user = 'fg'

nodes = {
         'rac1': {'address': '10.10.10.11', 'id': 1, 'role': 'rac1'},
         'rac2': {'address': '10.10.10.12', 'id': 2, 'role': 'rac2'},
         'racq': {'address': '10.10.10.13', 'id': 3, 'role': 'racq'}
        }

vifs = {
        'rac1-quorum': {'max_bw': '100%', 'min_bw': '0%', 'net_id': 201, 'prio': 1},
        'rac2-quorum': {'max_bw': '100%', 'min_bw': '0%', 'net_id': 202, 'prio': 1},
             'quorum': {'max_bw': '100%', 'min_bw': '0%', 'net_id': 255, 'prio': 1},
       }

roles = {
         'racq': {'direct': {'max_bw': '100%', 'min_bw': '0%', 'prio': 2},
                    'interface': 'eth0',
                    'root_bw': None,
                    'upload_cfg': True,
                    'vifs': {
                             'rac1-quorum': {'peers': ['rac1']},
                             'rac2-quorum': {'peers': ['rac2']}
                            }
                  },
         'rac1': {'direct': {'max_bw': '100%', 'min_bw': '0%', 'prio': 2},
                   'interface': 'eth0',
                   'root_bw': None,
                   'upload_cfg': True,
                   'vifs': {'quorum': {'net_id': 201, 'peers': ['racq']}}
                  },
         'rac2': {'direct': {'max_bw': '100%', 'min_bw': '0%', 'prio': 2},
                   'interface': 'eth0',
                   'root_bw': None,
                   'upload_cfg': True,
                   'vifs': {'quorum': {'net_id': 202, 'peers': ['racq']}}
                  },
       }
```

The above configuration will create a virtual NIC named *quorum* on each of the database nodes and virtual NICs named *rac1-quorum* and *rac2-quorum* on the quorum server. IP addresses from the 192.168.x.x range will be assigned to the virtual NICs. Later when configuring FlashGrid Storage Fabric, these virtual NICs must be selected along with the primary storage NICs.

8.6.3 Blacklisting CLAN NICs in NetworkManager configuration

In /etc/NetworkManager/NetworkManager.conf file add the following lines:

```
[keyfile]
unmanaged-devices=interface-name:*quorum
```

8.6.4 Enabling FlashGrid CLAN service

To enable FlashGrid CLAN service

1. Place the configuration file as `/etc/flashgrid-clan.cfg` on one of the database nodes
2. On the same database node, as user fg deploy the configuration to all nodes of the CLAN cluster using force option (this requires user fg with sudo rights and passwordless SSH access to all of the nodes)

   ```
   $ sudo flashgrid-clan-cfg deploy-config -f
   ```

3. On all nodes of the CLAN cluster, enable and start flashgrid-clan services and verify their statuses

   ```
   # systemctl enable flashgrid-clan; systemctl start flashgrid-clan; systemctl status flashgrid-clan
   ```

   ```
   # systemctl enable flashgrid-clan-wait; systemctl start flashgrid-clan-wait; systemctl status flashgrid-clan-wait
   ```

4. Verify that the required virtual network interfaces were created

   ```
   # ip link
   ```

9 Creating LVM Volumes for Quorum and GRID Disks

9.1 Choosing an LVM volume group

An LVM volume group with sufficient amount of free space is required if you need to create LVM volumes for quorum or GRID disks.

To check available LVM volume groups or to create a new volume group

1. Install LVM2 rpm if it is not installed

   ```
   # yum install lvm2
   ```

2. Check available volume group(s) and the amount of free space. For quorum disks you will need 1 GiB multiplied by the number of disk groups. For GRID disks you need 5 GiB.

   ```
   # vgdisplay
   ```

3. If no volume group or no free space is available, create a new volume group on any unused disk or partition

   ```
   # pvcreate /dev/<disk>
   # vgcreate <vgname> /dev/<disk>
   ```

9.2 Creating LVM volumes for quorum disks

If you are planning to have quorum disks located on database nodes or on quorum servers then you need to create one 1 GiB LVM volume for each ASM disk group on each of the database nodes or quorum servers. The quorum disk volumes must have ‘quorum’ in either volume group name or logical volume name. Such volumes will be automatically shared by FlashGrid without additional configuration. The quorum disk volumes can be added before or after configuring the FlashGrid cluster.

To create LVM volumes for quorum disks

On each database node or quorum server create one volume for each disk group. Include word ‘quorum’ in the volume name unless volume group name already includes it.

Example for three disk groups:

```
# lvcreate <vgname> --size 1G --name quorum1
# lvcreate <vgname> --size 1G --name quorum2
# lvcreate <vgname> --size 1G --name quorum3
```

9.3 Creating LVM volumes for GRID disk group

If the GRID disk group is placed on FlashGrid disks then you need to create LVM volumes on each of the converged or storage nodes. The GRID disk volumes must have ‘grid’ in either volume group name or logical volume name. Such volumes will be automatically shared by FlashGrid without additional configuration. The grid disk volumes can be added before or after configuring the FlashGrid cluster, but they must be available before installing Grid Infrastructure.

To create LVM volumes for GRID disks

On each converged or storage node create one GRID disk volume and two temporary GRID disk volumes:

```
# lvcreate <vgname> --size 5G --name grid
# lvcreate <vgname> --size 5G --name gridtmp1
# lvcreate <vgname> --size 5G --name gridtmp2
```

10 Configuring FlashGrid Storage Fabric

Configuration of a FlashGrid Storage Fabric cluster is stored in `/etc/flashgrid.cfg` files on each node of the cluster. In most cases there is no need to edit the configuration files manually. FlashGrid Configuration Assistant tool provides an easy way for creating the configuration files.

It is important to remember that creating, changing, or erasing FlashGrid configuration does not change the contents of any ASM disks including ASM metadata. FlashGrid cluster configuration is stored separately from ASM disk group configuration and only determines how disks are shared between the nodes. However, extra care must be taken when changing FlashGrid cluster configuration while any ASM disk group is mounted. Accidentally removing access to any disk that is a member of a mounted disk group may lead to degraded data redundancy or to the disk group being dismounted by ASM.

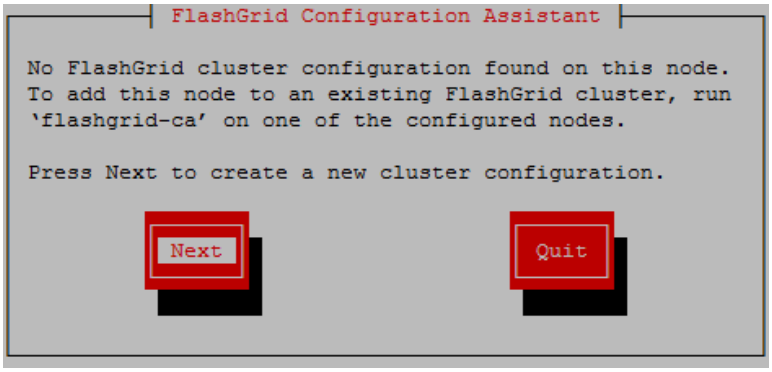
Before configuring a FlashGrid cluster, verify the following prerequisites

- Time is synchronized between all nodes that will be included in the cluster
- Owner (e.g. 'oracle' or 'grid') and group (e.g. 'asmadmin') for Grid Infrastructure are configured on all nodes where ASM will be installed.
- Current user account meets the requirements for performing cluster-wide operations. Using fg user account is recommended.
- If Grid Infrastructure is already installed then the CRS services must be stopped.

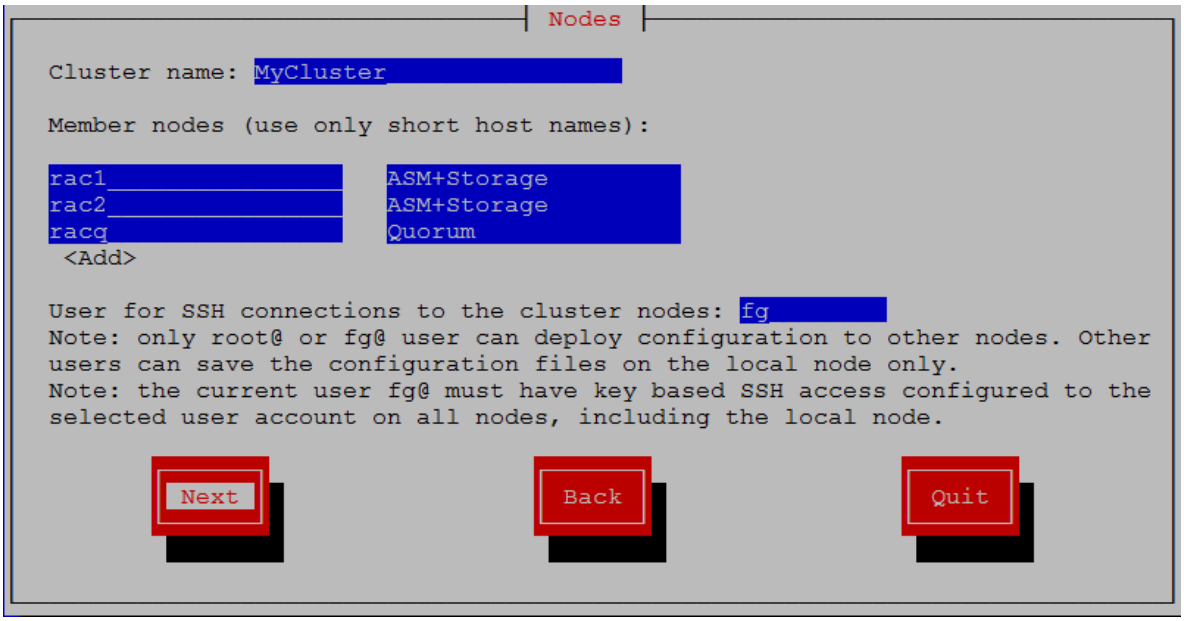
To configure a FlashGrid cluster

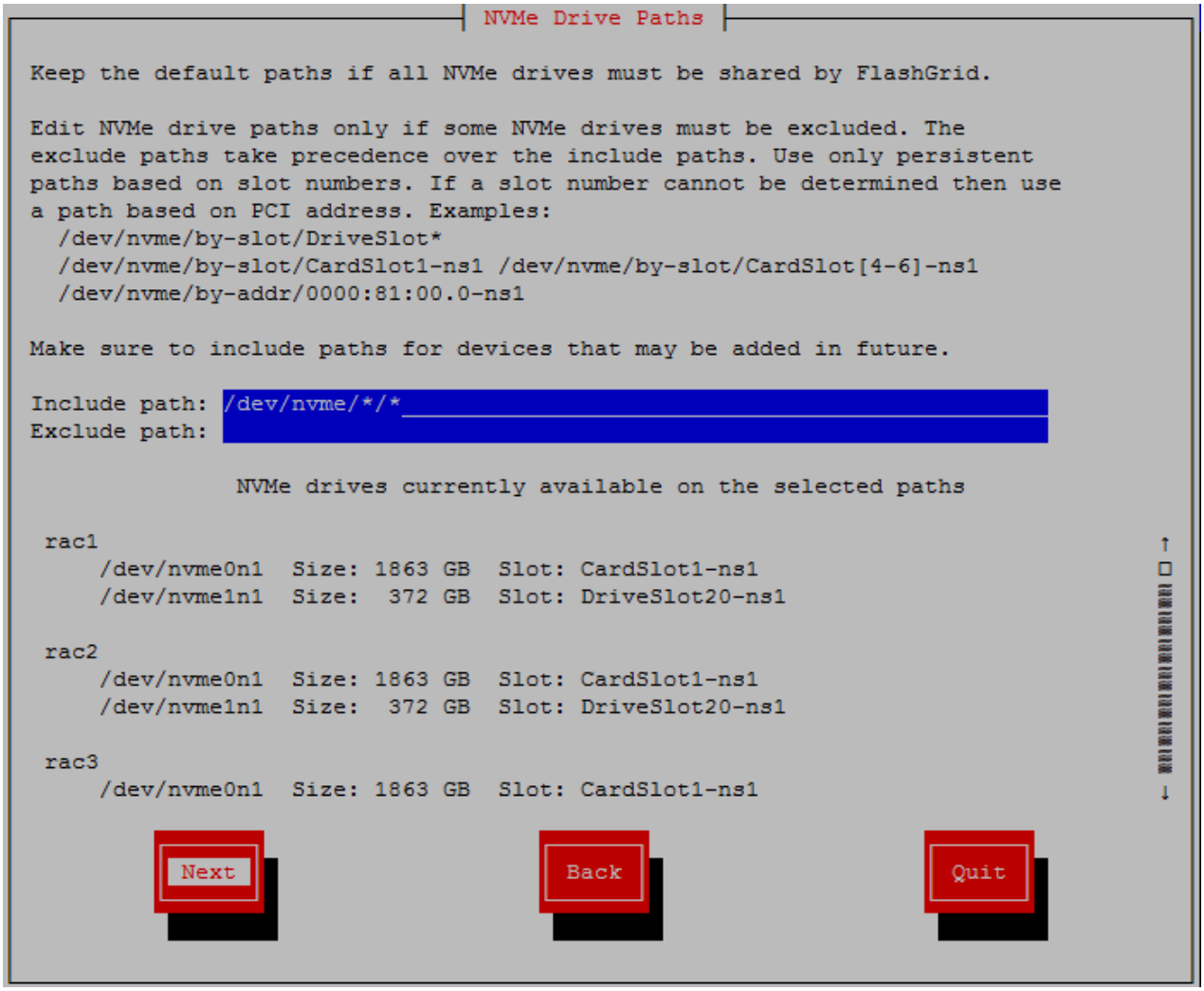
5. As user *fg*, run `flashgrid-ca` on any node of the cluster
6. Complete all steps of the FlashGrid Configuration Assistant following the instructions on each screen.
7. Run '`flashgrid-cluster`' command to check status of all nodes in the cluster and network connectivity between the nodes.
8. If any of the nodes shows *Warning* or *Critical* status then on that node run '`flashgrid-node`' command to find the source of the problem. Note that a *Warning* state is expected on those nodes that have the *ASM node* role and no Grid Infrastructure installed.
9. On each node run '`flashgrid-node test-alerts`' to check that email alerts work.

FlashGrid Configuration Assistant: Create new configuration

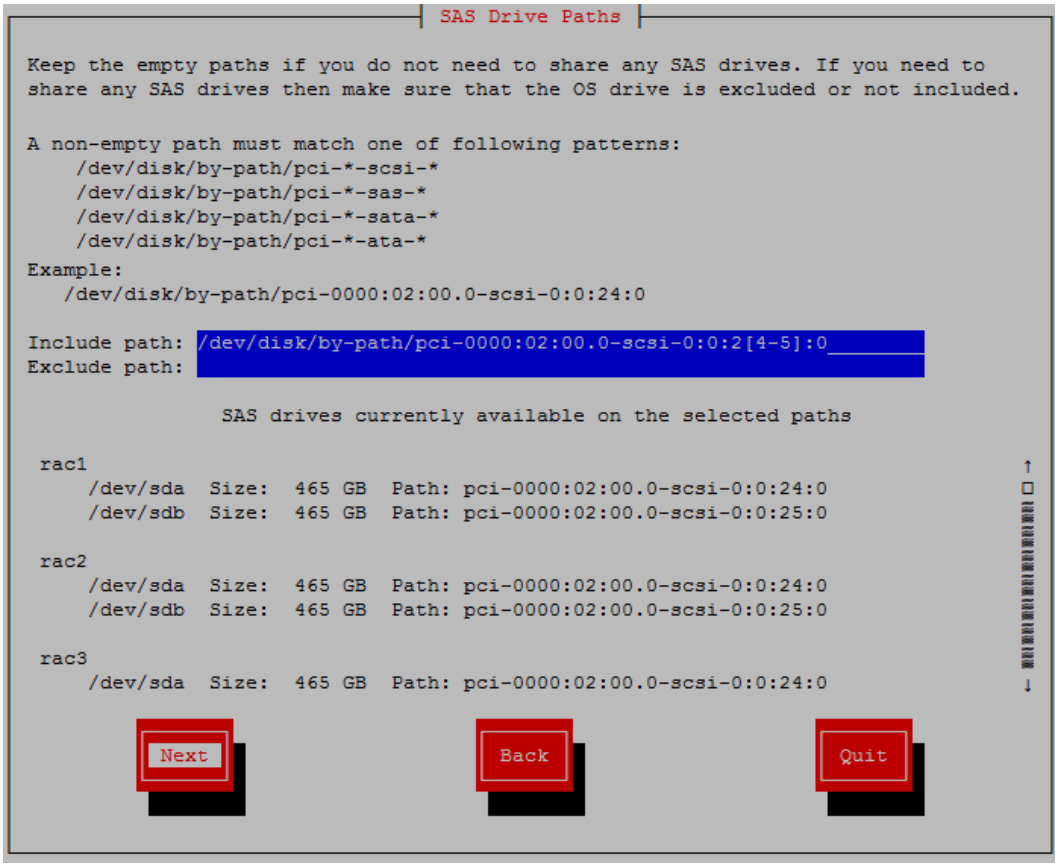


FlashGrid Configuration Assistant: Configuring cluster name and nodes

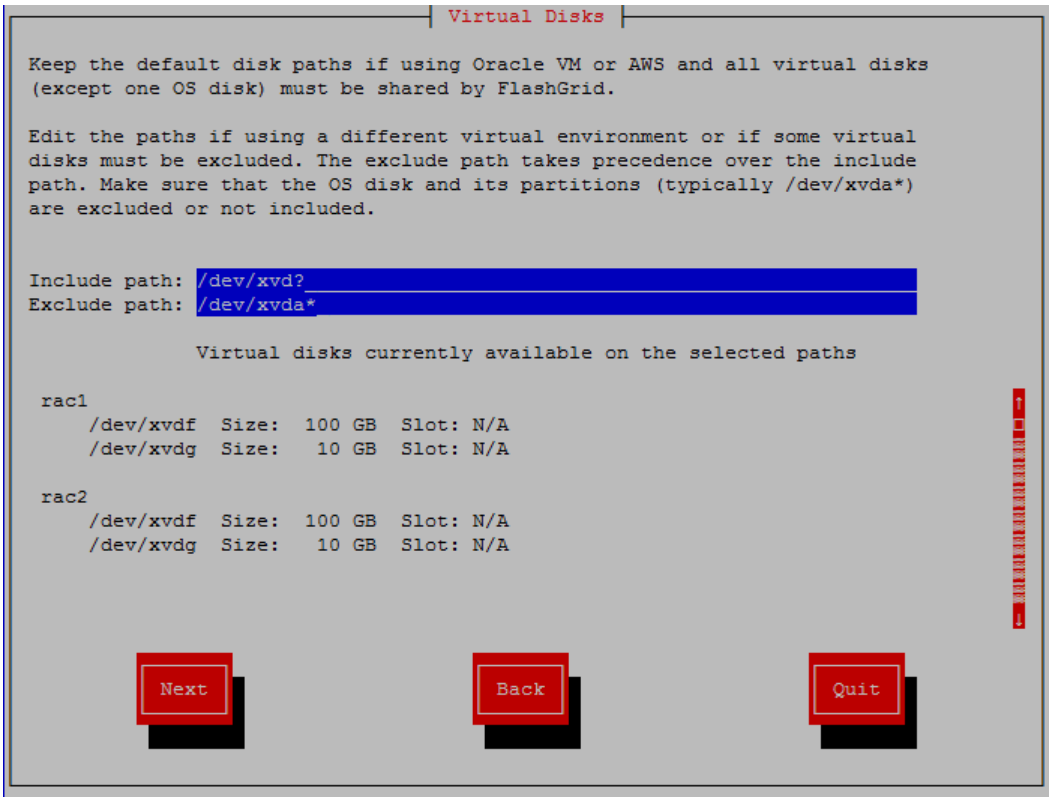



NVMe Drive Paths
Keep the default paths if all NVMe drives must be shared by FlashGrid.
Edit NVMe drive paths only if some NVMe drives must be excluded. The
exclude paths take precedence over the include paths. Use only persistent
paths based on slot numbers. If a slot number cannot be determined then use
a path based on PCI address. Examples:
/dev/nvme/by-slot/DriveSlot*
/dev/nvme/by-slot/CardSlot1-ns1 /dev/nvme/by-slot/CardSlot[4-6]-ns1
/dev/nvme/by-addr/0000:81:00.0-ns1
Make sure to include paths for devices that may be added in future.
Include path: /dev/nvme/*/*
Exclude path:
NVMe drives currently available on the selected paths
rac1
/dev/nvme0n1 Size: 1863 GB Slot: CardSlot1-ns1
/dev/nvme1n1 Size: 372 GB Slot: DriveSlot20-ns1
rac2
/dev/nvme0n1 Size: 1863 GB Slot: CardSlot1-ns1
/dev/nvme1n1 Size: 372 GB Slot: DriveSlot20-ns1
rac3
/dev/nvme0n1 Size: 1863 GB Slot: CardSlot1-ns1
Next
Back
Quit

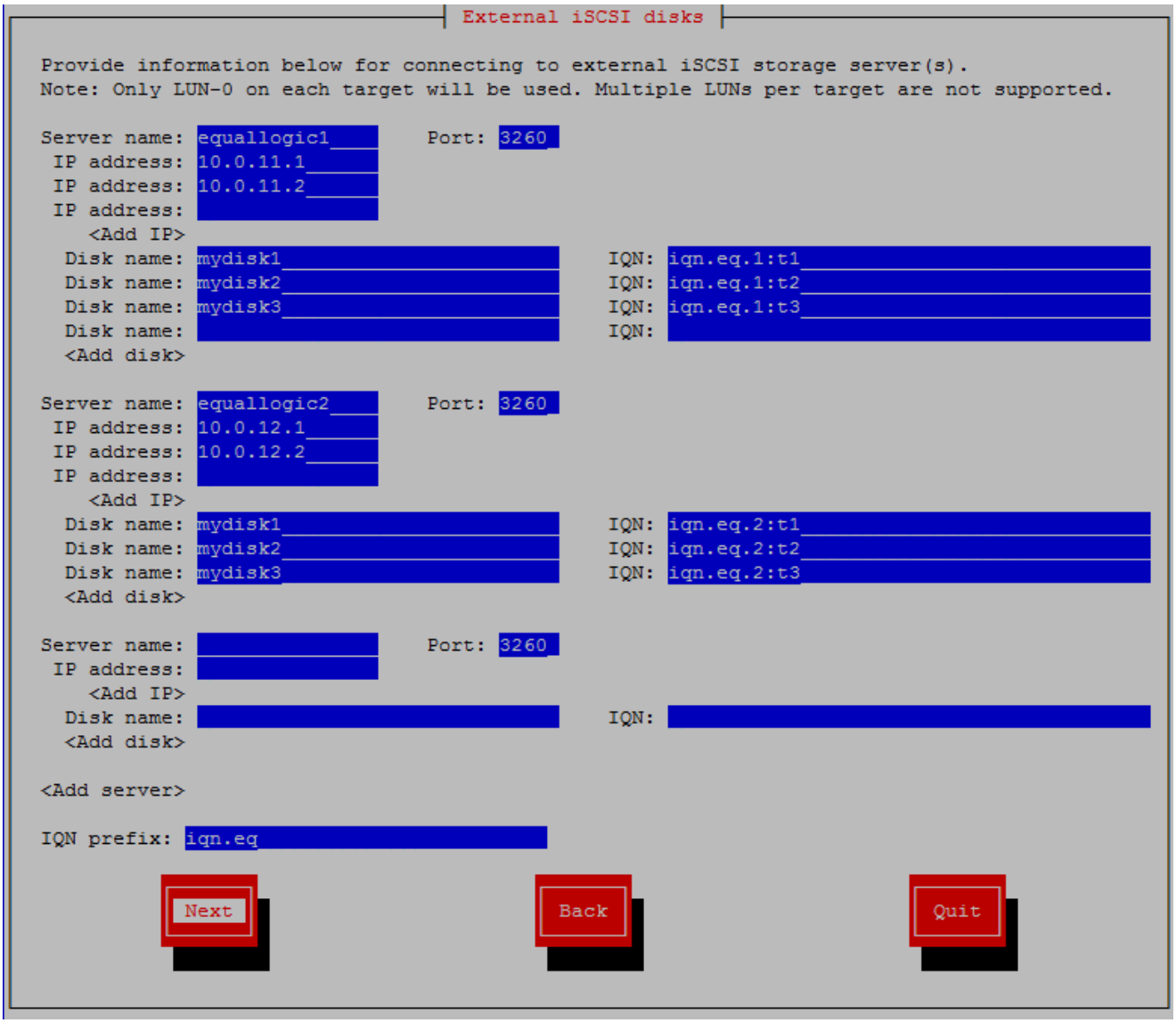
FlashGrid Configuration Assistant: Configuring SAS SSDs



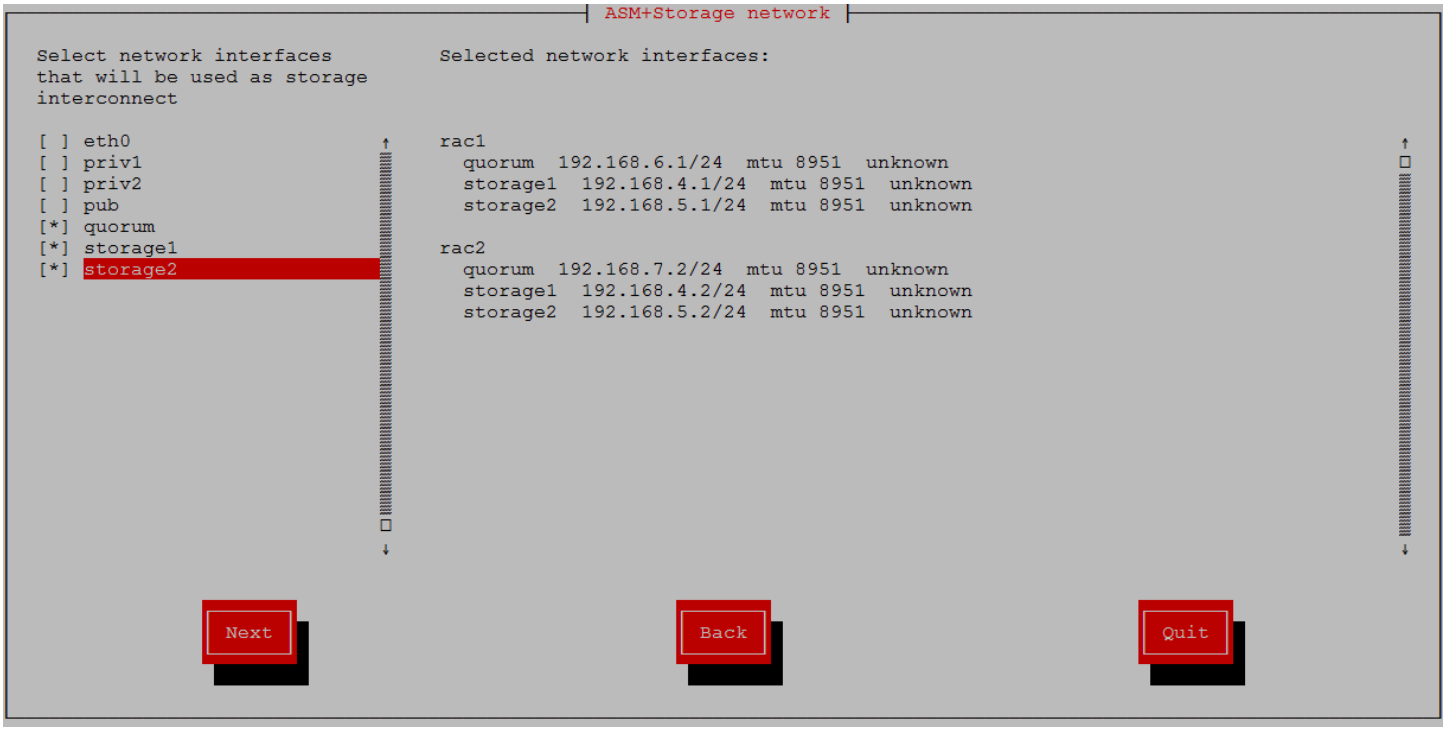
FlashGrid Configuration Assistant: Configuring Virtual Disks



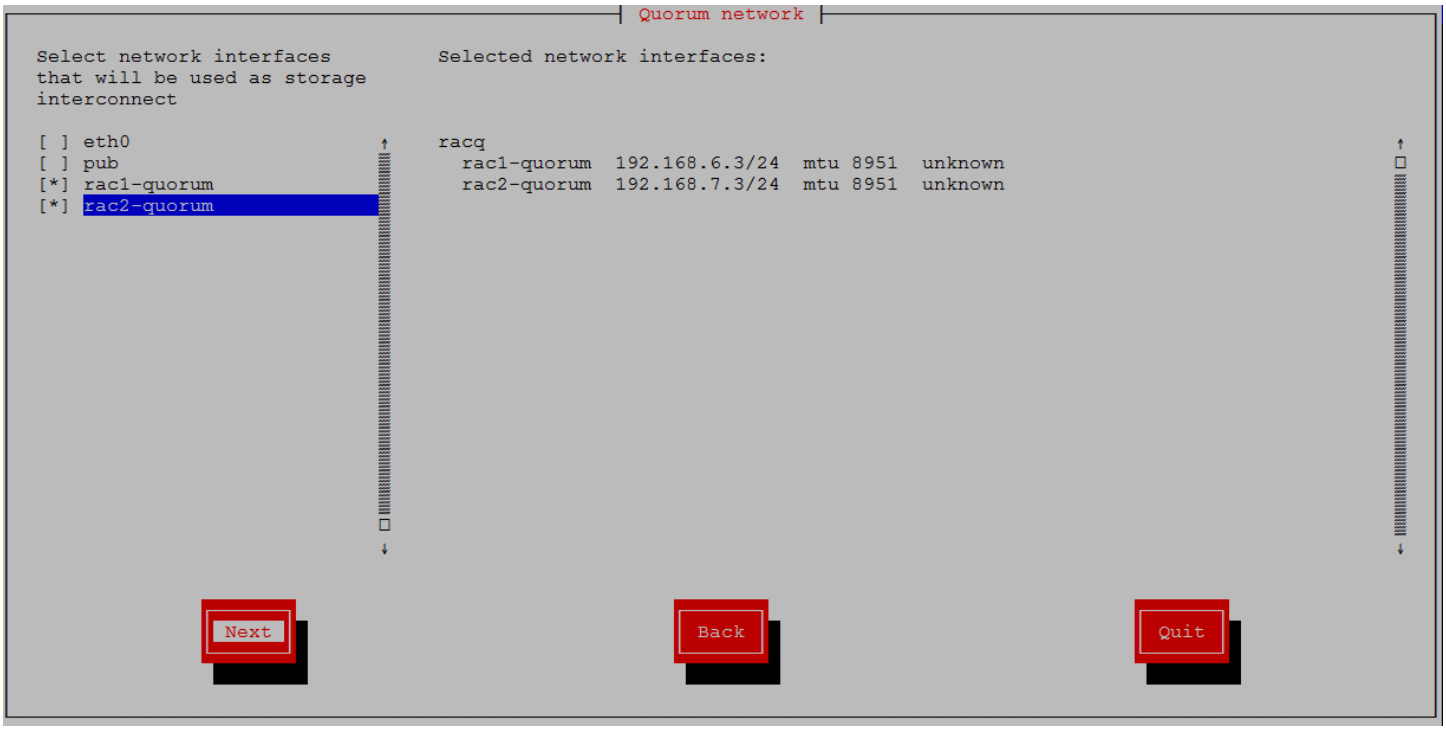
FlashGrid Configuration Assistant: Configuring External iSCSI Disks



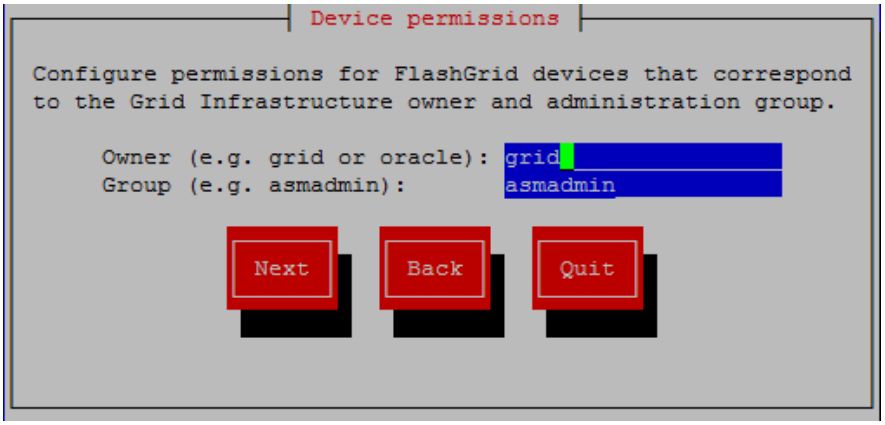
FlashGrid Configuration Assistant: Configuring NICs



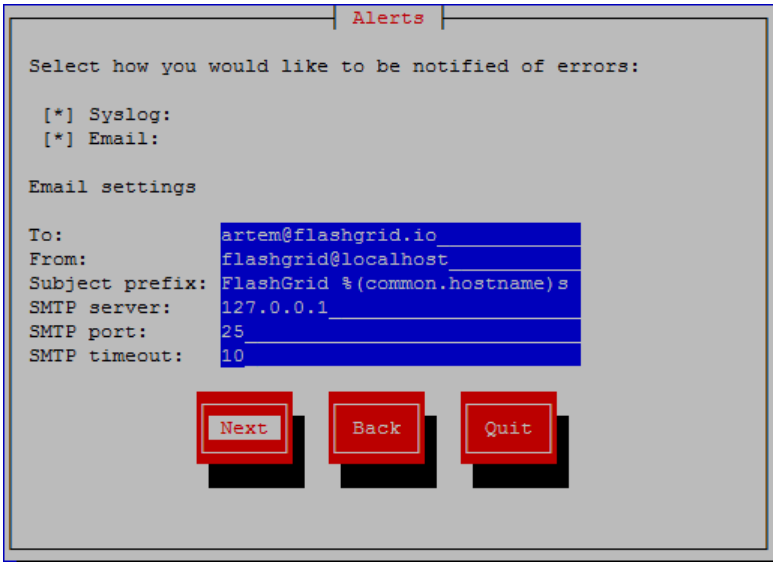
FlashGrid Configuration Assistant: Configuring NICs on Quorum Node(s)



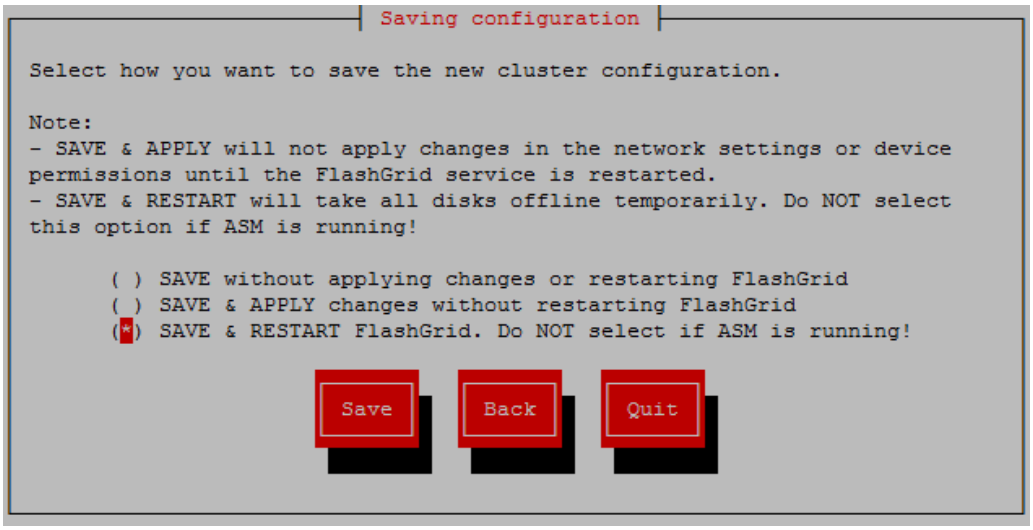
FlashGrid Configuration Assistant: Configuring device permissions



FlashGrid Configuration Assistant: Configuring Alerts



FlashGrid Configuration Assistant: Saving configuration as root



Example of a cluster status summary after initial configuration

```
[root@rac1 ~]# flashgrid-cluster
FlashGrid 18.6.29.67827 #49014476d930d22175fb859187a4b7d3a518d622
License: Active, Perpetual
Licensee: XYZ Corp
Support plan: 24x7
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
FlashGrid running: OK
Clocks check: OK
Configuration check: OK
Network check: OK

Querying nodes: rac02, rac01, quorum ...

Cluster Name: MyCluster
Cluster status: Good
-----------------------------------------------------------------
Node    Status   ASM_Node  Storage_Node  Quorum_Node  Failgroup
-----------------------------------------------------------------
rac01   Good     Yes       Yes           No           RAC01
rac02   Good     Yes       Yes           No           RAC02
quorum  Good     No        No            Yes          QUORUM
-----------------------------------------------------------------
```

Example of a node status summary after initial configuration

```
[root@rac1 ~]# flashgrid-node
FlashGrid 18.6.29.67827 #49014476d930d22175fb859187a4b7d3a518d622
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
rac1 node status: Good

Local NICs:
-----------------------------------------------------------
Address        Status  Iface       HW          Speed   MTU
-----------------------------------------------------------
192.168.100.1  Good    storage1    ethernet    100000  9000
192.168.101.1  Good    storage2    ethernet    100000  9000
192.168.201.1  Good    quorum      ethernet    N/A     1451
-----------------------------------------------------------

Local Drives:
-------------------------------------------------------------------------------------------------------------------------------
DriveName              Status  SizeGiB  Slot         WritesUsed  ASMName                ASMSizeGiB  DiskGroup  ASMStatus
-------------------------------------------------------------------------------------------------------------------------------
rac1.ft441500852p0egn  Good    1863     CardSlot1    0%          RAC1$FT441500852P0EGN  N/A         N/A        N/A
rac1.s1j0nyaf901288    Good    372      DriveSlot20  3%          RAC1$S1J0NYAF901288    N/A         N/A        N/A
rac1.ol7-grid          Good    10       N/A          N/A         RAC1$OL7_GRID          N/A         N/A        N/A
rac1.ol7-gridtmp       Good    5        N/A          N/A         RAC1$OL7_GRIDTMP       N/A         N/A        N/A
-------------------------------------------------------------------------------------------------------------------------------

Remote Drives:
-------------------------------------------------------------------------------------------------------------
DriveName              Status  SizeGiB  CfgPaths  ActPaths  ASMName                ASMSizeGiB  DiskGroup  ASMStatus
-------------------------------------------------------------------------------------------------------------
quorum.ol7-quorum1     Good    0        1         1         QUORUM$OL7_QUORUM1     N/A         N/A        N/A
quorum.ol7-quorum2     Good    0        1         1         QUORUM$OL7_QUORUM2     N/A         N/A        N/A
quorum.ol7-quorum3     Good    0        1         1         QUORUM$OL7_QUORUM3     N/A         N/A        N/A
rac2.ft516000fx2p0egn  Good    1863     2         2         RAC2$FT516000FX2P0EGN  N/A         N/A        N/A
rac2.s1j0nyaf901300    Good    372      2         2         RAC2$S1J0NYAF901300    N/A         N/A        N/A
rac2.ol7-grid          Good    10       2         2         RAC2$OL7_GRID          N/A         N/A        N/A
rac2.ol7-gridtmp       Good    5        2         2         RAC2$OL7_GRIDTMP       N/A         N/A        N/A
-------------------------------------------------------------------------------------------------------------
```

11 Installing Grid Infrastructure with GRID Disk Group on FlashGrid Disks

In most cases it is preferred to use FlashGrid disks for the GRID disk group (voting files, OCR). The Grid Infrastructure installer does not allow configuring a disk group on FlashGrid disks with custom disk names. Therefore, additional steps are required when placing GRID disk group on FlashGrid disks.

To create GRID disk group on FlashGrid disks

1. Make sure that you have LVM volumes created for use as GRID disks

2. During Grid Infrastructure configuration disable Grid Infrastructure Management Repository (GIMR).

3. During Grid Infrastructure configuration configure a disk group for GI files using the temporary GRID disk volumes:
   - Select *Normal* or *High* redundancy level for the disk group (do not select *External*)
   - Add `/dev/flashgrid/*` to the disk discovery string
   - Include all *gridtmp1* and *gridtmp2* disks from all nodes

4. Before running **root.sh** script on each node, clear page cache to avoid a bug in AMDU utility:

   ```
   # echo 1 > /proc/sys/vm/drop_caches
   ```

5. Immediately after Grid Infrastructure installation is complete, replace the temporary disks in the disk group with permanent GRID disks using the assistant tool `flashgrid-fix-grid-dg-ca`

   Note: Until the disk group created by GI installer is fixed, FlashGrid tools will be showing errors for the disk group and its disks.

12 Enabling *flashgrid-node-monitor* service

The flashgrid-node-monitor service is part of the FlashGrid Diagnostics package. It monitors disk I/O, network, and system clock and logs any detected for abnormalities. The logs can help with troubleshooting potential errors in the operation of the cluster.

To enable the flashgrid-node-monitor service, on each node of the cluster including quorum nodes

1. Create `/etc/flashgrid_debug_env` file with GI base home path. Example:

   ```
   PATH_GRID_BASE=/u01/app/grid
   ```

2. Enable and start the service:

   ```
   # systemctl enable flashgrid-node-monitor
   # systemctl start flashgrid-node-monitor
   ```

3. Verify that the service started successfully:

   ```
   # systemctl status flashgrid-node-monitor
   ```

13 Creating ASM Disk Groups

To add new hot-plug SSDs in a running cluster

1. Plug in new SSDs
2. Use `flashgrid-cluster drives` command to determine FlashGrid names of the SSDs, e.g. rac2.newserialnumber
3. Run `flashgrid-dg` to create a new disk group with the new SSDs or add the new SSDs to the ASM disk group. **flashgrid-dg** is a non-interactive CLI tool. Example:

```
$ flashgrid-dg add-disks -G FLASHDG -d /dev/flashgrid/rac1.newserialnumber1
/dev/flashgrid/rac2.newserialnumber2
```

Disk Group Compatibility

When you create a disk group, you need to specify the disk group compatibility attribute settings. This section discusses two compatibility attributes COMPATIBLE.ASM and COMPATIBLE.RDBMS.

These attributes determine the availability of certain ASM features described in **Oracle ASM features enabled by disk group compatibility attribute settings**. Once attributes are set, they **cannot be reverted** to a lower value, and can only be advanced.

ASM Compatibility

The COMPATIBLE.ASM attribute must be advanced before advancing other disk group compatibility attributes and its value must be greater than or equal to the value of other disk group compatibility attributes (**Reference**). Set this attribute to the version of a (Grid Infrastructure) GI stack, such as: **19.0**, **18.0**, **12.2**, etc.

RDBMS Compatibility

You should set RDBMS Compatibility attribute to match the database version(s) in use. If using one database version, then set the attribute to that version. If using two or more different database versions, then set the attribute to the lowest database version. Note that the version must be 11.2 or higher. Additional information is available in **Oracle documentation**.

If the disk group will be used for ACFS only then set RDBMS Compatibility to match ASM Compatibility.

13.1 Creating a disk group using flashgrid-dg CLI tool

Connect as the grid user to any database node to run the flashgrid-dg command. The following command creates a new NORMAL redundancy disk group with 2 disks and one quorum disk:

```
flashgrid-dg create --name FLASHDG \
                    --normal \
                    --asm-compat 19.0 \
                    --db-compat 19.0 \
                    --disks /dev/flashgrid/rac1.newserialnumber1
/dev/flashgrid/rac2.newserialnumber2 \
                    --quorum-disks /dev/flashgrid/quorum.newserialnumber3 \
                    --disk-repair-time 24000h \
                    --failgroup-repair-time 24000h \
                    --au-size 4M
```

Note: *Set disk-repair-time and failgroup-repair-time to 24000h, as shown in the example above. This will prevent ASM from dropping disks unnecessarily during transient disk or node failures.*

Note: *The disk group created in such a way will be mounted on the node where that command was run. Make sure you log in to the remaining database nodes and mount the disk group manually:*

```
$ asmcmd mount FLASHDG
```

You can get help with flashgrid-dg options using `$ flashgrid-dg create -h`

13.2 Verification

Please make sure that the created disk group is mounted on all nodes. For that, run the **flashgrid-cluster** command:

- **Mounted** should be **AllNodes**
- **Status** should be **Good**

```
[grid@rac1 ~]$ flashgrid-cluster
...
---------------------------------------------------------------------------------------------------------
GroupName  Status   Mounted   Type    TotalMiB  FreeMiB  OfflineDisks  LostDisks  Resync  ReadLocal  Vote
---------------------------------------------------------------------------------------------------------
FLASHDG    Good     AllNodes  NORMAL  8192      7880     0             0          No      Enabled    None
```

Example of a cluster status summary after configuring one disk group

```
[root@rac1 ~]# flashgrid-cluster
FlashGrid 18.6.29.67827 #49014476d930d22175fb859187a4b7d3a518d622
License: Active, Perpetual
Licensee: XYZ Corp
Support plan: 24x7
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
Cluster verification: OK
Querying nodes: rac02, rac01, quorum ...

Cluster Name: MyCluster
---------------------------------------------------------------
Node    Status   ASM_Node  Storage_Node  Quorum_Node  Failgroup
---------------------------------------------------------------
rac01   Good     Yes       Yes           No           RAC01
rac02   Good     Yes       Yes           No           RAC02
quorum  Good     No        No            Yes          QUORUM
---------------------------------------------------------------
---------------------------------------------------------------------------------------------------------
GroupName  Status  Mounted   Type    TotalMiB  FreeMiB  OfflineDisks  LostDisks  Resync  ReadLocal  Vote
---------------------------------------------------------------------------------------------------------
FLASHDG    Good    AllNodes  NORMAL  1142984   200243   0             0          No      Enabled    None
GRID       Good    AllNodes  NORMAL  20480     17786    0             0          No      Enabled    3/3
---------------------------------------------------------------------------------------------------------
```

Example of a node status summary after configuring one disk group

```
[root@rac1 ~]# flashgrid-node
FlashGrid 18.6.29.67827 #49014476d930d22175fb859187a4b7d3a518d622
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
rac1 node status: Good

Local NICs:
-----------------------------------------------------------
Address        Status  Iface       HW          Speed   MTU
-----------------------------------------------------------
192.168.100.1  Good    storage1    ethernet    100000  9000
192.168.101.1  Good    storage2    ethernet    100000  9000
192.168.201.1  Good    quorum      ethernet    N/A     1451
-----------------------------------------------------------
```

```
Local Drives:
-------------------------------------------------------------------------------------------------------------------
DriveName               Status  SizeGiB  Slot         WritesUsed  ASMName                ASMSizeGiB  DiskGroup  ASMStatus
-------------------------------------------------------------------------------------------------------------------
rac1.ft441500852p0egn   Good    1863     CardSlot1    0%          RAC1$FT441500852P0EGN  N/A         N/A        N/A
rac1.s1j0nyaf901288     Good    372      DriveSlot20  3%          RAC1$S1J0NYAF901288    372         FLASHDG    ONLINE
rac1.ol7-grid           Good    10       N/A          N/A         RAC1$OL7_GRID          10          GRID       ONLINE
rac1.ol7-gridtmp        Good    5        N/A          N/A         RAC1$OL7_GRIDTMP       N/A         N/A        N/A
-------------------------------------------------------------------------------------------------------------------

Remote Drives:
-------------------------------------------------------------------------------------------------------------
DriveName               Status  SizeGiB  CfgPaths  ActPaths  ASMName                ASMSizeGiB  DiskGroup  ASMStatus
-------------------------------------------------------------------------------------------------------------
quorum.ol7-quorum1      Good    0        1         1         QUORUM$OL7_QUORUM1     0           GRID       ONLINE
quorum.ol7-quorum2      Good    0        1         1         QUORUM$OL7_QUORUM2     0           FLASHDG    ONLINE
quorum.ol7-quorum3      Good    0        1         1         QUORUM$OL7_QUORUM3     N/A         N/A        N/A
rac2.ft516000fx2p0egn   Good    1863     2         2         RAC2$FT516000FX2P0EGN  N/A         N/A        N/A
rac2.s1j0nyaf901300     Good    372      2         2         RAC2$S1J0NYAF901300    372         FLASHDG    ONLINE
rac2.ol7-grid           Good    10       2         2         RAC2$OL7_GRID          10          GRID       ONLINE
rac2.ol7-gridtmp        Good    5        2         2         RAC2$OL7_GRIDTMP       N/A         N/A        N/A
-------------------------------------------------------------------------------------------------------------
```

14 Configuring Database Memory Settings

Larger SGA and PGA allocations can help with achieving higher database performance. However, it is critical to ensure that the settings are configured correctly to avoid running out of memory. Running out of memory will result in processes being killed with unpredictable results to the system stability. A typical mistake leading to low available memory condition is having the PGA size parameters set too high or too many HugePages configured.

On systems with 60 GiB or larger physical memory FlashGrid recommends allocating 80% of the total memory for use by the database(s). The remaining 20% must be reserved for the OS, Grid Infrastructure, and FlashGrid software. On systems with less than 60 GiB of physical memory, 12 GiB must be reserved for the OS, Grid Infrastructure, and FlashGrid software.

The optimal ratio of SGA and PGA is different for different types of databases. However, the sum of SGA and PGA allocations for all databases must not exceed the total memory *DatabaseMemory* value as calculated above.

If configuring HugePages then the amount of memory allocated as HugePages must match the SGA target (or sum of all SGA targets for multiple databases) plus 2 GiB for GIMR. Note that PGA and other software cannot use HugePages. Allocating too many HugePages may result in running out of memory.

Example of memory allocations calculation for a 40% PGA / 60% SGA ratio for a single database:

1) Calculate total database memory (for all databases)
   If Total_Memory_GiB >= 60 GiB: *Database_Memory_GiB = 0.8 x Total_Memory_GiB*
   If Total_Memory_GiB < 60 GiB: *Database_Memory_GiB = Total_Memory_GiB - 12 GiB*

2) Calculate PGA size parameters
   PGA_AGGREGATE_LIMIT_GiB = round(0.4 x Database_Memory_GiB)
   PGA_AGGREGATE_TARGET_GiB = round(0.5 x PGA_AGGREGATE_LIMIT_GiB)

   Note: In database version 11.2.0.4 explicitly setting *PGA_AGGREGATE_LIMIT* parameter is not supported. It is calculated automatically from *PGA_AGGREGATE_TARGET.*

3) Calculate SGA max size
   SGA_MAX_SIZE_GiB = round(0.6 x Database_Memory_GiB)

4) Calculate number of huge pages
   Number_HugePages = (SGA_MAX_SIZE_GiB + 2) x 512

In case of multiple databases sharing the same nodes, the ***sum of all*** PGA aggregate limit/target parameters must be equal to or lower than the values calculated using the formulas above.

15 Monitoring Cluster Health

The following methods of monitoring cluster health are available:

- *flashgrid-health-check* utility checks multiple items including database configuration, storage, OS kernel, config file modifications, errors in the logs, and other items that may affect health of the cluster or could help with troubleshooting. It is recommended for manual checks only.
- *flashgrid-cluster* utility displays status of the storage subsystem (FlashGrid Storage Fabric and ASM) and its main components. The utility can be used in monitoring scripts. It returns a non-zero value if status of the cluster is *Warning* or *Critical*.
- Alerts about failures are recorded in system log and can be analyzed by 3rd-party tools.
- Email alerts can be sent to one or several email addresses.
- ASM disk group monitoring and alerting via Oracle Enterprise Manager.

To test email alerts

10. On all nodes (including quorum node) run

    ```
    $ flashgrid-node test-alerts
    ```

11. Check that test alert emails were received from all cluster nodes at each of the configured email addresses.

To modify the list of email alert recipients

As user *fg* on any database node run

```
$ flashgrid-cluster set-email-alerts name1@host1 name2@host2 ...
```

Note that by default the *From* address is set to *flashgrid@localhost.localdomain*. This will ensure that delivery failure notifications are sent to root's mailbox on the originating node, which can help with troubleshooting delivery issues. It is recommended to add this address to the whitelist of senders on the receiving email server and in the email clients.

16 Before Going Live

Before switching the cluster to live use:

1. Verify health of the cluster: `$ sudo flashgrid-health-check`
2. Confirm that email alerts are configured and delivered: `$ flashgrid-node test-alerts`
3. Upload diags to FlashGrid support: `$ sudo flashgrid-diags upload-all`
4. Stop the cluster and back up all cluster nodes: https://support.flashgrid.io/hc/en-us/articles/4404887221655-Shutting-down-entire-cluster
5. Start the cluster and do final check of the cluster health: `$ sudo flashgrid-health-check`

17 Troubleshooting

The following troubleshooting steps are recommended in case of any issues with FlashGrid cluster configuration or operation:

1. Check status of all FlashGrid nodes, network, and disk groups by running '`flashgrid-cluster`' on any node
2. For any ASM disk group that has a *Warning* or *Critical* status, run the following command to check the list nodes where the disk group is mounted and the list of disks with their status:

   ```
   $ flashgrid-dg show -G <DGNAME>
   ```

3. On any node that has a *Warning*, *Critical*, or *Inaccessible* status:
   a. Check whether the FlashGrid service is active:
      `# systemctl status flashgrid`
   b. Check status of NICs, local disks, and remote disks:
      `# flashgrid-node`
   c. Check that the configuration has no errors:
      `# flashgrid-node verify-config`
4. If network verification fails then run '`flashgrid-cluster verify`' to get more detailed information
5. Check FlashGrid log files on the affected nodes. The log files are located in `/opt/flashgrid/log`

18 Additional Documentation

FlashGrid Storage Fabric CLI Reference Guide: https://support.flashgrid.io/hc/en-us/articles/1500011214681

FlashGrid Cloud Area Network CLI Reference Guide: https://support.flashgrid.io/hc/en-us/articles/1500011214661

FlashGrid Storage Fabric Release Notes: https://support.flashgrid.io/hc/en-us/articles/1500011214781

FlashGrid Diagnostics Release Notes: https://support.flashgrid.io/hc/en-us/articles/1500011175642

19 Contacting FlashGrid Technical Support

For help with troubleshooting an issue on an existing FlashGrid cluster please use Technical Support Request form located at https://www.flashgrid.io/support/

To expedite troubleshooting please also collect and upload diagnostic data to the secure storage used by FlashGrid support by running the following command:

```
# sudo flashgrid-diags upload-all
```

For reporting *emergency* type of issues that require immediate attention please also use the 24x7 telephone hotline: +1-650-641-2421 ext 7. Please note that use of the 24x7 hotline is reserved for emergency situations only.