FlashGrid® Storage Fabric
Version 19.06

Deployment and Maintenance Guide
for on-premises deployments

rev. 19.06-2019.08.17
# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1  Introduction</td>
<td>3</td>
</tr>
<tr>
<td>2  Compatibility</td>
<td>4</td>
</tr>
<tr>
<td>3  Overview of the Process of Deploying an Oracle RAC Cluster on FlashGrid Storage Fabric</td>
<td>4</td>
</tr>
<tr>
<td>4  FlashGrid Storage Fabric Architecture</td>
<td>5</td>
</tr>
<tr>
<td>4.1 Hyper-converged architecture or separate storage nodes</td>
<td>5</td>
</tr>
<tr>
<td>4.2 Shared Access</td>
<td>6</td>
</tr>
<tr>
<td>4.3 Data Mirroring</td>
<td>6</td>
</tr>
<tr>
<td>4.4 FlashGrid Read-Local™ Technology</td>
<td>7</td>
</tr>
<tr>
<td>4.5 Strict Read-Local Mode</td>
<td>7</td>
</tr>
<tr>
<td>4.6 Quorum Disks</td>
<td>8</td>
</tr>
<tr>
<td>4.7 Using External Storage in Combination with FlashGrid</td>
<td>8</td>
</tr>
<tr>
<td>4.8 Selecting Location for Grid Infrastructure Files (GRID disk group)</td>
<td>8</td>
</tr>
<tr>
<td>4.9 Dependencies between FlashGrid and Oracle Services</td>
<td>9</td>
</tr>
<tr>
<td>4.10 Persistent Names and Permissions for ASM Disks</td>
<td>9</td>
</tr>
<tr>
<td>5  Installing and Configuring OS</td>
<td>10</td>
</tr>
<tr>
<td>5.1 Enabling Performance mode in system BIOS</td>
<td>10</td>
</tr>
<tr>
<td>5.2 Reserving space for LVM volumes during OS installation</td>
<td>10</td>
</tr>
<tr>
<td>5.3 Disabling swap</td>
<td>10</td>
</tr>
<tr>
<td>5.4 Installing kernel-devel package</td>
<td>10</td>
</tr>
<tr>
<td>5.5 Setting Performance mode for CPU governor</td>
<td>10</td>
</tr>
<tr>
<td>5.6 Synchronizing clocks with NTP</td>
<td>10</td>
</tr>
<tr>
<td>5.7 Configuring ASM device owner user and group</td>
<td>10</td>
</tr>
<tr>
<td>5.8 Preparing a user account for cluster configuration operations</td>
<td>11</td>
</tr>
<tr>
<td>6  Installing FlashGrid Storage Fabric Software</td>
<td>12</td>
</tr>
<tr>
<td>7  Installing License File</td>
<td>12</td>
</tr>
<tr>
<td>8  Configuring Storage Network</td>
<td>13</td>
</tr>
<tr>
<td>8.1 Sharing Network between FlashGrid and Oracle Private Network</td>
<td>13</td>
</tr>
<tr>
<td>8.2 Common Network Configuration Requirements</td>
<td>13</td>
</tr>
<tr>
<td>8.3 Storage Network Topology</td>
<td>13</td>
</tr>
<tr>
<td>8.4 Configuring NICs</td>
<td>14</td>
</tr>
<tr>
<td>8.5 Configuring Enhanced Network Capabilities with FlashGrid® Cloud Area Network™</td>
<td>15</td>
</tr>
<tr>
<td>9  Creating LVM Volumes for Quorum and GRID Disks</td>
<td>17</td>
</tr>
<tr>
<td>9.1 Choosing an LVM volume group</td>
<td>17</td>
</tr>
<tr>
<td>9.2 Creating LVM volumes for quorum disks</td>
<td>17</td>
</tr>
<tr>
<td>9.3 Creating LVM volumes for GRID disk group</td>
<td>17</td>
</tr>
<tr>
<td>10 Configuring FlashGrid Storage Fabric</td>
<td>18</td>
</tr>
<tr>
<td>11 Installing Grid Infrastructure with GRID Disk Group on FlashGrid Disks</td>
<td>26</td>
</tr>
<tr>
<td>12 Enabling flashgrid-node-monitor service</td>
<td>26</td>
</tr>
<tr>
<td>13 Creating ASM Disk Groups</td>
<td>27</td>
</tr>
<tr>
<td>14 Configuring Database Memory Settings</td>
<td>32</td>
</tr>
<tr>
<td>15 Measuring Performance</td>
<td>33</td>
</tr>
<tr>
<td>16 Troubleshooting</td>
<td>34</td>
</tr>
<tr>
<td>17 Additional Documentation</td>
<td>34</td>
</tr>
<tr>
<td>18 Contacting FlashGrid Technical Support</td>
<td>34</td>
</tr>
</tbody>
</table>
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 the corresponding FlashGrid SkyCluster Deployment Guide.

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 local SSDs
2 Compatibility

- Oracle Linux 7 or Red Hat Enterprise Linux 7
  - Red Hat kernel version 3.10.0-957.el7.x86_64 or later
- Oracle Grid Infrastructure 19c (recommended), 18c, or 12.2.0.1 with July’19 or newer Release Update
- Oracle Database 19c, 18c, 12.2.0.1, 12.1.0.2, or 11.2.0.4 with July’19 or newer Release Update / Patch Set Update

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
   - Network type and topology

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 Patch Set 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 SSDs in the cluster. Each SSD is visible in the OS as /dev/flashgrid/nodename.drivename 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 makes sure that mirrored copies of a data block are placed on 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 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 disk 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:

- LVM volumes on dedicated quorum server(s), physical or VM
- LVM volumes on database nodes (for clusters with separate storage nodes)
- External iSCSI/FC/FCoE 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.

<table>
<thead>
<tr>
<th></th>
<th>2 nodes</th>
<th>3 nodes</th>
<th>4 nodes</th>
<th>5+ nodes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal Redundancy</td>
<td>1 quorum disk</td>
<td>not needed</td>
<td>not needed</td>
<td>not needed</td>
</tr>
<tr>
<td>High Redundancy</td>
<td>N/A</td>
<td>2 quorum disks</td>
<td>1 quorum disk</td>
<td>not needed</td>
</tr>
<tr>
<td>(with Voting Files)</td>
<td>required</td>
<td>required</td>
<td>recommended</td>
<td></td>
</tr>
<tr>
<td>High Redundancy</td>
<td>N/A</td>
<td>2 quorum disks</td>
<td>1 quorum disk</td>
<td>not needed</td>
</tr>
<tr>
<td>(no Voting Files)</td>
<td>recommended</td>
<td>recommended</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

4.7 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.

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

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

- **An external storage: FC/FCoE/iSCSI SAN or NFS.** This option provides simplified installation and more flexibility by storing Grid Infrastructure files separately from the high performance FlashGrid storage. 100 GB capacity is sufficient in most cases with moderate performance requirements. The external storage must provide high availability. Two or three separate storage systems can be combined in a Normal or High Redundancy disk group to achieve high availability.

- **FlashGrid Storage Fabric managed disks.** This option allows reduced dependency on any external storage. This option is preferred when there is no external storage available that can provide sufficient level of availability. 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.
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 cluster 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 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. asadmin) 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 a single RPM package. Additionally, installation of FlashGrid Diagnostics RPM 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 Storage Fabric: `flashgrid-sf`
   - FlashGrid Diagnostics: `flashgrid-diags`

2. Download `oracle-instantclient19.3-basic-19.3.0.0.0-1.x86_64.rpm` from
   https://www.oracle.com/database/technologies/instant-client/linux-x86-64-downloads.html#license-lightbox
   Note: Same version 19.3 of the instant client is used for Grid Infrastructure versions 19c, 18c, and 12.2.

3. Use YUM to install the downloaded RPMs and their dependencies:
   ```
   # yum install oracle-instantclient19.3-basic-19.3.0.0.0-1.x86_64.rpm
   # yum install flashgrid-sf-*.*.rpm
   # yum install flashgrid-diags-*.*.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

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 Common Network Configuration Requirements

In all storage network topologies the following requirements must be met:

- 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.
- Use of jumbo frames (MTU=9000) is recommended on Ethernet networks. If using switches then before enabling jumbo frames on NICs, need to verify that switches also are configured to support jumbo frames.
- 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
- The following ports must be open between all nodes of a cluster:
  - TCP 3260
  - TCP 5557
  - TCP 8753

8.3 Storage Network Topology

The following subsections describe typical storage network topologies. These topologies are for the storage network only and do not cover public network or Oracle private network.

8.3.1 Direct back-to-back network links

This topology allows implementing a high-speed storage network interconnect without the additional cost of the high-speed network switches.

In production environments at least two network links must be used for redundancy between each pair of nodes. Using three or more links is possible for extra bandwidth. There is no need for the storage network links between database-only nodes. There is no need for the storage network links between storage-only nodes.

Each network link must be configured as a separate IP subnet. FlashGrid software provides multi-path and load balancing across multiple links. NIC bonding is not required.

8.3.2 A single switch

This topology should not be used in a mission-critical environment. A failure of the single switch may cause the cluster down time.
Each node must be connected to one or more ports of the switch. Connecting to more than one port adds port and cable redundancy and increases the total network bandwidth available to each node.

8.3.3 Two or more separate switches, each node connected to all switches

Two separate switches are recommended for most configurations of 4 or more nodes per cluster. Three or more switches can be used for extra switch redundancy.

The switches should have no links between them. If links between the switches are required for purposes other than FlashGrid then the ports used for FlashGrid should be isolated in non-routable VLANs that do not include the inter-switch links.

Each node must be connected to all switches.

8.3.4 Two or more stacked Ethernet switches, each node connected to all switches

Stacked Ethernet switches may be configured with LACP link aggregation across the switches. 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.

8.3.5 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.

8.4 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=storage
IPADDR=192.168.100.Node#
PREFIX=24
BOOTPROTO=none
DEFROUTE=no
IPV4_FAILURE_FATAL=yes
IPV6INIT=no
ONBOOT=yes
NM_CONTROLLED=no
TYPE=Ethernet
MTU=9000
```
8.5 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.

8.5.1 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
   ```

8.5.2 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.

```plaintext
[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 = {
    'rac1': {'device': 'eth0'},
    'rac2': {'device': 'eth1'},
    'racq': {'device': 'eth2'}
}
```
The above configuration will create a virtual NIC named \textit{quorum} on each of the database nodes and virtual NICs named \textit{rac1-quorum} and \textit{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.

\textbf{8.5.3 Blacklisting CLAN NICs in NetworkManager configuration}

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

\begin{verbatim}
[keyfile] unmanaged-devices=interface-name:*quorum
\end{verbatim}

\textbf{8.5.4 Enabling FlashGrid CLAN service}

\textbf{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)

\begin{verbatim}
$ sudo flashgrid-clan-cfg deploy-config -f
\end{verbatim}

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

\begin{verbatim}
# systemctl enable flashgrid-clan; systemctl start flashgrid-clan; systemctl status flashgrid-clan
# systemctl enable flashgrid-clan-wait; systemctl start flashgrid-clan; systemctl status flashgrid-clan
\end{verbatim}

4. Verify that the required virtual network interfaces were created

\begin{verbatim}
# ip link
\end{verbatim}
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 150 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 100G --name grid
# lvcreate <vgname> --size 25G --name gridtmp1
# lvcreate <vgname> --size 25G --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 in 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

No FlashGrid cluster configuration found on this node. To add this node to an existing FlashGrid cluster, run `flashgrid-ca` on one of the configured nodes.

Press Next to create a new cluster configuration.

Next Quit

FlashGrid Configuration Assistant: Configuring cluster name and nodes

Cluster name: MyCluster

Member nodes (use only short host names):

- rac1 ASM+Storage
- rac2 ASM+Storage
- racq Quorum

<Add>

User for SSH connections to the cluster nodes: fg

Note: only root@ or fg@ user can deploy configuration to other nodes. Other users can save the configuration files on the local node only. Note: the current user fg@ must have key based SSH access configured to the selected user account on all nodes, including the local node.

Next Back Quit
FlashGrid Configuration Assistant: Configuring NVMe SSDs

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 FCI address. Examples:
/dev/nvme/by-slot/DriveSlot
/dev/nvme/by-sлот/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
FlashGrid Configuration Assistant: Configuring SAS SSDs

Keep the empty paths if you do not need to share any SAS drives. If you need to share any SAS drives then make sure that the OS drive is excluded or not included.

A non-empty path must match one of following patterns:
/dev/disk/by-path/poI-**-scsi-*
/dev/disk/by-path/poI-**-sas-*
/dev/disk/by-path/poI-**-ata-*

Example:
/dev/disk/by-path/poI-0000:02:00.0-scsi-0:0:24:0

Include path: /dev/disk/by-path/poI-0000:02:00.0-scsi-0:0:24:0
Exclude path: 

SAS drives currently available on the selected paths

rac1
/dev/sda Size: 465 GB Path: poI-0000:02:00.0-scsi-0:0:24:0
/dev/sdb Size: 465 GB Path: poI-0000:02:00.0-scsi-0:0:25:0

rac2
/dev/sda Size: 465 GB Path: poI-0000:02:00.0-scsi-0:0:24:0
/dev/sdb Size: 465 GB Path: poI-0000:02:00.0-scsi-0:0:25:0

rac3
/dev/sda Size: 465 GB Path: poI-0000:02:00.0-scsi-0:0:24:0

FlashGrid Configuration Assistant: Configuring Virtual Disks

Keep the default disk paths if using Oracle VM or AWS and all virtual disks (except one OS disk) must be shared by FlashGrid.

Edit the paths if using a different virtual environment or if some virtual disks must be excluded. The exclude path takes precedence over the include path. Make sure that the OS disk and its partitions (typically /dev/xvda*) are excluded or not included.

Include path: /dev/xvda*
Exclude path: 

Virtual disks currently available on the selected paths

rac1
/dev/xvdf Size: 100 GB Slot: N/A
/dev/xvdg Size: 10 GB Slot: N/A

rac2
/dev/xvdf Size: 100 GB Slot: N/A
/dev/xvdg Size: 10 GB Slot: N/A
### FlashGrid Configuration Assistant: Configuring External iSCSI Disks

Provide information below for connecting to external iSCSI storage server(s).

**Note:** Only LUN-0 on each target will be used. Multiple LUNs per target are not supported.

<table>
<thead>
<tr>
<th>Server Name</th>
<th>Port</th>
</tr>
</thead>
<tbody>
<tr>
<td>equallogic1</td>
<td>3260</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>IP Address</th>
<th>Disk Name</th>
<th>IQN</th>
</tr>
</thead>
<tbody>
<tr>
<td>10.0.11.1</td>
<td>mydisk1</td>
<td>iqn.eq.1:t1</td>
</tr>
<tr>
<td>10.0.11.2</td>
<td>mydisk2</td>
<td>iqn.eq.1:t2</td>
</tr>
<tr>
<td>10.0.11.2</td>
<td>mydisk3</td>
<td>iqn.eq.1:t3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Server Name</th>
<th>Port</th>
</tr>
</thead>
<tbody>
<tr>
<td>equallogic2</td>
<td>3260</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>IP Address</th>
<th>Disk Name</th>
<th>IQN</th>
</tr>
</thead>
<tbody>
<tr>
<td>10.0.12.1</td>
<td>mydisk1</td>
<td>iqn.eq.2:t1</td>
</tr>
<tr>
<td>10.0.12.2</td>
<td>mydisk2</td>
<td>iqn.eq.2:t2</td>
</tr>
<tr>
<td>10.0.12.2</td>
<td>mydisk3</td>
<td>iqn.eq.2:t3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Server Name</th>
<th>Port</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3260</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Disk Name</th>
<th>IQN</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Add server**

**IQN prefix:** iqn.eq

[Next] [Back] [Quit]
FlashGrid Configuration Assistant: Configuring device permissions

Configure permissions for FlashGrid devices that correspond to the Grid Infrastructure owner and administration group.

Owner (e.g. grid or oracle): grid
Group (e.g. asadmin): asadmin

Next Back Quit

FlashGrid Configuration Assistant: Configuring Alerts

Select how you would like to be notified of errors:
[*] Syslog:
[*] Email:

Email settings
To: artem@flashgrid.io
From: flashgrid@localhost
Subject prefix: FlashGrid *(common_hostname)*
SMTP server: 127.0.0.1
SMTP port: 25
SMTP timeout: 10

Next Back Quit

FlashGrid Configuration Assistant: Saving configuration as root

Select how you want to save the new cluster configuration.

Note:
- SAVE & APPLY will not apply changes in the network settings or device permissions until the FlashGrid service is restarted.
- SAVE & RESTART will take all disks offline temporarily. Do NOT select this option if ASM is running!

( ) SAVE without applying changes or restarting FlashGrid
( ) SAVE & APPLY changes without restarting FlashGrid
( ) SAVE & RESTART FlashGrid. Do NOT select if ASM is running!

Save Back Quit
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

<table>
<thead>
<tr>
<th>Node</th>
<th>Status</th>
<th>ASM_Node</th>
<th>Storage_Node</th>
<th>Quorum.Node</th>
<th>Failgroup</th>
</tr>
</thead>
<tbody>
<tr>
<td>rac01</td>
<td>Good</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>RAC01</td>
</tr>
<tr>
<td>rac02</td>
<td>Good</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>RAC02</td>
</tr>
<tr>
<td>quorum</td>
<td>Good</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>QUORUM</td>
</tr>
</tbody>
</table>

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:

<table>
<thead>
<tr>
<th>Address</th>
<th>Status</th>
<th>Iface</th>
<th>HW</th>
<th>Speed</th>
<th>MTU</th>
</tr>
</thead>
<tbody>
<tr>
<td>192.168.100.1</td>
<td>Good</td>
<td>storage1</td>
<td>ethernet</td>
<td>100000</td>
<td>9000</td>
</tr>
<tr>
<td>192.168.101.1</td>
<td>Good</td>
<td>storage2</td>
<td>ethernet</td>
<td>100000</td>
<td>9000</td>
</tr>
<tr>
<td>192.168.201.1</td>
<td>Good</td>
<td>quorum</td>
<td>ethernet</td>
<td>N/A</td>
<td>1451</td>
</tr>
</tbody>
</table>

Local Drives:

<table>
<thead>
<tr>
<th>DriveName</th>
<th>Status</th>
<th>SizeGiB</th>
<th>Slot</th>
<th>WriteUsed</th>
<th>ASMName</th>
<th>ASMSizeGiB</th>
<th>DiskGroup</th>
<th>ASMStatus</th>
</tr>
</thead>
<tbody>
<tr>
<td>rac1.ft441500852pSegn</td>
<td>Good</td>
<td>1863</td>
<td>CardSlot1</td>
<td>0%</td>
<td>RAC1FT441500852P0EGN</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>rac1.sl1j0nyaF901288</td>
<td>Good</td>
<td>372</td>
<td>DriveSlot20</td>
<td>3%</td>
<td>RAC1S1J0NYAF901288</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>rac1.ol7-grid</td>
<td>Good</td>
<td>10</td>
<td>N/A</td>
<td>N/A</td>
<td>RAC1OL7_Grid</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>rac1.ol7-gridtmp</td>
<td>Good</td>
<td>5</td>
<td>N/A</td>
<td>N/A</td>
<td>RAC1OL7_GRIDTMP</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
</tbody>
</table>

Remote Drives:

<table>
<thead>
<tr>
<th>DriveName</th>
<th>Status</th>
<th>SizeGiB</th>
<th>CfgPaths</th>
<th>ActPaths</th>
<th>ASMName</th>
<th>ASMSizeGiB</th>
<th>DiskGroup</th>
<th>ASMStatus</th>
</tr>
</thead>
<tbody>
<tr>
<td>quorum.ol7-quorum1</td>
<td>Good</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>QUORUM$OL7_QUORUM1</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>quorum.ol7-quorum2</td>
<td>Good</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>QUORUM$OL7_QUORUM2</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>quorum.ol7-quorum3</td>
<td>Good</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>QUORUM$OL7_QUORUM3</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>rac2.ft516000fx2pSegn</td>
<td>Good</td>
<td>1863</td>
<td>2</td>
<td>2</td>
<td>RAC2FT516000FX2P0EGN</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>rac2.sl1j0nyaF901300</td>
<td>Good</td>
<td>372</td>
<td>2</td>
<td>2</td>
<td>RAC2S1J0NYAF901300</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>rac2.ol7-grid</td>
<td>Good</td>
<td>10</td>
<td>2</td>
<td>2</td>
<td>RAC2OL7_Grid</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>rac2.ol7-gridtmp</td>
<td>Good</td>
<td>5</td>
<td>2</td>
<td>2</td>
<td>RAC2OL7_GRIDTMP</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
</tbody>
</table>
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, MGMTDB). 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 disabling Grid Infrastructure Management Repository (GIMR) is recommended (for GI 19c only).

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:
   
   ```bash
   # 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 and database base and home paths. Example:

   ```bash
   PATH_GRID_BASE=/u01/app/grid
   PATH_DB_BASE=/u01/app/oracle
   PATH_DB_HOME=/u01/app/oracle/product/12.2.0/dbhome_1
   ```

2. Enable and start the service:

   ```bash
   # systemct1 enable flashgrid-node-monitor
   # systemct1 start flashgrid-node-monitor
   ```

3. Verify that the service started successfully:

   ```bash
   # systemct1 status flashgrid-node-monitor
   ```
13 Creating ASM Disk Groups

**Note:** For configuring a disk group on FlashGrid storage for Grid Infrastructure files (voting files, OCR, MGMTDB) see section 11 for the GI installation steps.

Use of FlashGrid Disk Group Configuration Assistant is recommended for creating ASM disk groups on FlashGrid disks. This wizard-like tool helps avoid mistakes by automating the following tasks:

- Enforces correct disk group type and compatibility attributes
- Assigns each disk to a correct failgroup based on its physical location
- Checks disk number and size restrictions
- Helps determine correct number of quorum disks
- Offers optimal disk and failgroup repair time depending on the disk group configuration
- Configures preferred read settings

To create an ASM disk group on FlashGrid disks

1. If your cluster and disk group configuration requires one or more quorum disks (see section 0 Quorum disks) then create the required number of quorum disks. For quorum disks on external storage, except iSCSI disks configured in FlashGrid, add path to the disks to the ASM disk discovery string.

2. Make sure that ASM is running on all nodes that are configured to run ASM.

3. On any ASM node as the Grid Infrastructure owner user run `flashgrid-create-dg`

4. Complete all steps of the FlashGrid Disk Group Configuration Assistant following the instructions on each screen.

5. Run ‘flashgrid-cluster’ command to check status of the newly created disk group.

**FlashGrid Disk Group Configuration Assistant: Disk Group Attributes**

![Disk Group Parameters](image)

Name: FLASHDG
Type: High Redundancy (3-way mirroring)
ASM Compatibility: 12.1.0.2
Database Compatibility: 12.1

Note that setting Database Compatibility = 11.2.x.x.x will truncate disk sizes to 2 TiB max.

[Next] [Back] [Quit]
FlashGrid Disk Group Configuration Assistant: Selecting Disks

Select FlashGrid disks that you want to include in the new disk group FLASHG2

<-> rac1
[*] ft441500852p0egn 1863 GiB CardSlot1
<-> rac2
[*] ft816000fx2p0egn 1863 GiB CardSlot1
<-> rac3
[*] ft524500ch2p0egn 1863 GiB CardSlot1

Total capacity available to ASM: 5506 GiB | Capacity after mirroring: 2794 GiB

No errors

Next  Back  Quit

FlashGrid Disk Group Configuration Assistant: Selecting Quorum Disks

For the selected disk group configuration the minimum number of quorum disks is 1.

Quorum disks store copies of ASM metadata and must be located independently of the storage nodes. Possible locations are separate Quorum Servers, separate database nodes configured as Quorum Servers, or external iSCSI, FC, or NFS storage (quorum disks for disk groups with ACFS cannot reside on NFS). No user data is stored on the quorum disks. Recommended quorum disk size is 120 MiB.

<-> Other non-FlashGrid storage
[ ] /NFS DISKS/data1.quorumdisk1 100 MiB !ASM member!
[*] /NFS DISKS/data1.quorumdisk2 100 MiB

Next  Back  Reload  Quit
FlashGrid Disk Group Configuration Assistant: Setting Disk and FailGroup Repair Time

If needed, customize the disk and failgroup repair times. The disk repair time defines how long ASM waits for an offline disk to come back online before it drops the disk and starts rebalancing. The failgroup repair time is similar, but used only when an entire node (failgroup) goes offline.

These parameters are mainly for unexpected disk or node failure scenarios. For planned maintenance the parameters can be overridden in the manual offline command for a disk or a failgroup.

Disk repair time: 5.6h
Failgroup repair time: 240h

FlashGrid Disk Group Configuration Assistant: Allocation Unit Size

If needed, customize the Allocation Unit Size (AU_SIZE) for the disk group.

Allocation Unit Size: 4MB

FlashGrid Disk Group Configuration Assistant: Review Disk Group Summary

We are ready to create the new disk group.

Name: FLASHDG
Type: high
ASM Compatibility: 12.1.0.2
Database Compatibility: 12.1
Disks:
- rac1.disk4150000000001 1663 GiB
- rac2.disk8450000000001 1663 GiB
- rac3.disk8245000000001 1663 GiB
Quorum disks:
- /NFS_DISKS/flashdg.quorumdisk1
- /NFS_DISKS/flashdg.quorumdisk2
Total capacity available to ASM: 5999 GiB
Capacity after mirroring: 1663 GiB
Disk repair time: 3.6h
Failgroup repair time: 240h
Please review the SQL code for creating the disk group. Press Back if you need to make any changes.

```
CREATE DISKGROUP FLASHDG HIGH REDUNDANCY
  FAILGROUP rac1 DISK '/dev/flashgrid/rac1.ft415000852p0egn' NAME RAC1$FT415000852P0EHN SIZE 1663G
  FAILGROUP rac2 DISK '/dev/flashgrid/rac2.ft516000fx2p0egn' NAME RAC2$FT516000FX2P0EHN SIZE 1663G
  FAILGROUP rac3 DISK '/dev/flashgrid/rac3.ft5245000b2p0egn' NAME RAC3$FT5245000B2P0EHN SIZE 1663G
  QUORUM FAILGROUP quorum0 DISK '/NFS_DISKS/flashdg.quorumdisk1' NAME quorumdisk0
  QUORUM FAILGROUP quorum1 DISK '/NFS_DISKS/flashdg.quorumdisk2' NAME quorumdisk1
  ATTRIBUTE 'compatible.asm' = '12.1.0.2', 'compatible.rdzms' = '12.1';
  ALTER SYSTEM SET asm_preferred_read_failure_groups='FLASHDG.rac1' SID='+ASM1';
  ALTER SYSTEM SET asm_preferred_read_failure_groups='FLASHDG.rac2' SID='+ASM3';
  ALTER SYSTEM SET asm_preferred_read_failure_groups='FLASHDG.rac3' SID='+ASM2';
  ALTER DISKGROUP FLASHDG SET ATTRIBUTE 'failgroup_repair_time' = '240h';
  ALTER DISKGROUP FLASHDG SET ATTRIBUTE 'disk_repair_time' = '3.6h';
```

---

**FlashGrid Disk Group Configuration Assistant: Finish**

Do you want to create the new disk group now or just save the SQL code in a file?

- [ ] Create the new disk group and save the SQL code in a file
- [ ] Only save the SQL code in a file

Finish  Back  Quit
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

<table>
<thead>
<tr>
<th>Node</th>
<th>Status</th>
<th>ASM_Node</th>
<th>Storage_Node</th>
<th>Quorum_Node</th>
<th>Failgroup</th>
</tr>
</thead>
<tbody>
<tr>
<td>rac01</td>
<td>Good</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>RAC01</td>
</tr>
<tr>
<td>rac02</td>
<td>Good</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>RAC02</td>
</tr>
<tr>
<td>quorum</td>
<td>Good</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>QUORRM</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>GroupName</th>
<th>Status</th>
<th>Mounted</th>
<th>Type</th>
<th>TotalMiB</th>
<th>FreeMiB</th>
<th>OfflineDisks</th>
<th>LostDisks</th>
<th>Resync</th>
<th>ReadLocal</th>
<th>Vote</th>
</tr>
</thead>
<tbody>
<tr>
<td>FLASHDG</td>
<td>Good</td>
<td>AllNodes</td>
<td>NORMAL</td>
<td>1142984</td>
<td>200243</td>
<td>0</td>
<td>0</td>
<td>No</td>
<td>Enabled</td>
<td>None</td>
</tr>
<tr>
<td>GRID</td>
<td>Good</td>
<td>AllNodes</td>
<td>NORMAL</td>
<td>20480</td>
<td>17786</td>
<td>0</td>
<td>0</td>
<td>No</td>
<td>Enabled</td>
<td>3/3</td>
</tr>
</tbody>
</table>

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:

<table>
<thead>
<tr>
<th>Address</th>
<th>Status</th>
<th>Iface</th>
<th>HW</th>
<th>Speed</th>
<th>MTU</th>
</tr>
</thead>
<tbody>
<tr>
<td>192.168.100.1</td>
<td>Good</td>
<td>storage1</td>
<td>ethernet</td>
<td>100000</td>
<td>9000</td>
</tr>
<tr>
<td>192.168.101.1</td>
<td>Good</td>
<td>storage2</td>
<td>ethernet</td>
<td>100000</td>
<td>9000</td>
</tr>
<tr>
<td>192.168.201.1</td>
<td>Good</td>
<td>quorum</td>
<td>ethernet</td>
<td>N/A</td>
<td>1451</td>
</tr>
</tbody>
</table>

Local Drives:

<table>
<thead>
<tr>
<th>DriveName</th>
<th>Status</th>
<th>SizeGiB</th>
<th>Slot</th>
<th>WritesUsed</th>
<th>ASMName</th>
<th>ASMSizeGiB</th>
<th>DiskGroup</th>
<th>ASMStatus</th>
</tr>
</thead>
<tbody>
<tr>
<td>rac1.ft441500852pSegn</td>
<td>Good</td>
<td>1863</td>
<td>CardSlot1</td>
<td>0%</td>
<td>RAC1$FT441500852PSEGNN</td>
<td>N/A</td>
<td>FLASHDG</td>
<td>N/A</td>
</tr>
<tr>
<td>rac1.slj0nyaf901288</td>
<td>Good</td>
<td>372</td>
<td>DriveSlot20</td>
<td>3%</td>
<td>RAC1$SLJ0NYAFP901288</td>
<td>372</td>
<td>FLASHDG</td>
<td>ONLINE</td>
</tr>
<tr>
<td>rac1.ol7-grid</td>
<td>Good</td>
<td>10</td>
<td>N/A</td>
<td>N/A</td>
<td>RAC1$OL7_GRID</td>
<td>10</td>
<td>GRID</td>
<td>ONLINE</td>
</tr>
<tr>
<td>rac1.ol7-gridtmp</td>
<td>Good</td>
<td>5</td>
<td>N/A</td>
<td>N/A</td>
<td>RAC1$OL7_GRIDTMP</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
</tbody>
</table>

Remote Drives:

<table>
<thead>
<tr>
<th>DriveName</th>
<th>Status</th>
<th>SizeGiB</th>
<th>CfgPaths</th>
<th>ActPaths</th>
<th>ASMName</th>
<th>ASMSizeGiB</th>
<th>DiskGroup</th>
<th>ASMStatus</th>
</tr>
</thead>
<tbody>
<tr>
<td>quorum.ol7-quorum1</td>
<td>Good</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>QUORUM$OL7_QUORUM1</td>
<td>0</td>
<td>GRID</td>
<td>ONLINE</td>
</tr>
<tr>
<td>quorum.ol7-quorum2</td>
<td>Good</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>QUORUM$OL7_QUORUM2</td>
<td>0</td>
<td>FLASHDG</td>
<td>ONLINE</td>
</tr>
<tr>
<td>quorum.ol7-quorum3</td>
<td>Good</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>QUORUM$OL7_QUORUM3</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>rac2.ft516000fx2pSegn</td>
<td>Good</td>
<td>1863</td>
<td>2</td>
<td>2</td>
<td>RAC2$FT516000FX2PSEGNN</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>rac2.slj0nyaf901300</td>
<td>Good</td>
<td>372</td>
<td>2</td>
<td>2</td>
<td>RAC2$SLJ0NYAFP901300</td>
<td>372</td>
<td>FLASHDG</td>
<td>ONLINE</td>
</tr>
<tr>
<td>rac2.ol7-grid</td>
<td>Good</td>
<td>10</td>
<td>2</td>
<td>2</td>
<td>RAC2$OL7_GRID</td>
<td>10</td>
<td>GRID</td>
<td>ONLINE</td>
</tr>
<tr>
<td>rac2.ol7-gridtmp</td>
<td>Good</td>
<td>5</td>
<td>2</td>
<td>2</td>
<td>RAC2$OL7_GRIDTMP</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
</tbody>
</table>
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 swapping or 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 20% PGA / 80% 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(0.2 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(0.8 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.

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.
15 Measuring Performance

DBMS_RESOURCE_MANAGER.CALIBRATE_IO procedure provides an easy way for measuring storage performance including maximum bandwidth, random IOPS, and latency. The CALIBRATE_IO procedure generates I/O through the database stack on actual database files. The test is read-only and it is safe to run it on any existing database. It is also a good tool for directly comparing performance of two storage systems because the CALIBRATE_IO results do not depend on any non-storage factors, such as memory size or the number of CPU cores.

To measure storage performance with CALIBRATE_IO

1. Create or load a database on the corresponding ASM disk group
2. Make sure the total size of the database files is larger than 5 GB per disk. If needed, create an additional large table space / data file.
3. Customize the first parameter in the SQL code below with the number of disks corresponding to your storage setup. Keep the second parameter (max latency) with the minimum allowed value of 10 milliseconds.
4. Connect to the database with sqlplus and run the customized SQL code.
5. Wait for the CALIBRATE_IO to complete. This may take 10 to 30 minutes.

CALIBRATE_IO SQL code

```
SET SERVEROUTPUT ON
DECLARE
  lat INTEGER;
iops INTEGER;
mbps INTEGER;
BEGIN DBMS_RESOURCE_MANAGER.CALIBRATE_IO (12, 10, iops, mbps, lat);
  DBMS_OUTPUT.PUT_LINE ('max_iops = ' || iops);
  DBMS_OUTPUT.PUT_LINE ('latency = ' || lat);
  DBMS_OUTPUT.PUT_LINE ('max_mbps = ' || mbps);
end;
/
```

Example of running CALIBRATE_IO on a 3-node cluster with four NVMe SSDs per node

![Example of running CALIBRATE_IO on a 3-node cluster with four NVMe SSDs per node](image)
16 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

17 Additional Documentation

Maintenance Tasks: https://www.kb.flashgrid.io/maintenance/maintenance-on-prem

Troubleshooting Guide: https://www.kb.flashgrid.io/troubleshooting


FlashGrid Storage Fabric Release Notes: https://www.kb.flashgrid.io/release-notes/sf

FlashGrid Diagnostics Release Notes: https://www.kb.flashgrid.io/release-notes/diags

18 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.