

# Red Hat Enterprise Linux 5

## Configuration Example

### - NFS Over GFS

Configuring NFS over GFS in a Red Hat Cluster



# Red Hat Enterprise Linux 5 Configuration Example - NFS Over GFS

## Configuring NFS over GFS in a Red Hat Cluster

### Edition 3

Copyright © 2011 Red Hat Inc..

The text of and illustrations in this document are licensed by Red Hat under a Creative Commons Attribution–Share Alike 3.0 Unported license ("CC-BY-SA"). An explanation of CC-BY-SA is available at <http://creativecommons.org/licenses/by-sa/3.0/>. In accordance with CC-BY-SA, if you distribute this document or an adaptation of it, you must provide the URL for the original version.

Red Hat, as the licensor of this document, waives the right to enforce, and agrees not to assert, Section 4d of CC-BY-SA to the fullest extent permitted by applicable law.

Red Hat, Red Hat Enterprise Linux, the Shadowman logo, JBoss, MetaMatrix, Fedora, the Infinity Logo, and RHCE are trademarks of Red Hat, Inc., registered in the United States and other countries.

Linux® is the registered trademark of Linus Torvalds in the United States and other countries.

Java® is a registered trademark of Oracle and/or its affiliates.

XFS® is a trademark of Silicon Graphics International Corp. or its subsidiaries in the United States and/or other countries.

MySQL® is a registered trademark of MySQL AB in the United States, the European Union and other countries.

All other trademarks are the property of their respective owners.

1801 Varsity Drive  
Raleigh, NC 27606-2072 USA  
Phone: +1 919 754 3700  
Phone: 888 733 4281  
Fax: +1 919 754 3701

This book describes a procedure for configuring NFS over GFS in a 5-node Red Hat Cluster using Conga.

---

---

<b>Introduction</b>	<b>v</b>
1. About This Guide .....	v
2. Audience .....	v
3. Software Versions .....	v
4. Related Documentation .....	v
5. Feedback .....	vi
6. Document Conventions .....	vi
6.1. Typographic Conventions .....	vi
6.2. Pull-quote Conventions .....	viii
6.3. Notes and Warnings .....	viii
<b>1. NFS over GFS in a Red Hat Cluster</b>	<b>1</b>
<b>2. Prerequisite Configuration</b>	<b>3</b>
<b>3. Components to Configure</b>	<b>5</b>
<b>4. Configuring the Cluster Resources</b>	<b>7</b>
4.1. Configuring an IP Address Resource .....	7
4.2. Configuring a GFS Resource .....	7
4.3. Configuring an NFS Export Resource .....	8
4.4. Configuring NFS Client Resources .....	8
<b>5. Configuring an NFS Cluster Service</b>	<b>11</b>
5.1. Add a Service to the Cluster .....	11
5.2. Adding an IP Address Resource to an NFS Service .....	12
5.3. Adding a GFS Resource to an NFS Service .....	12
5.4. Adding an NFS Export Resource to an NFS Service .....	12
5.5. Adding NFS Client Resources to an NFS Service .....	13
<b>6. Testing the NFS Cluster Service</b>	<b>15</b>
<b>7. Troubleshooting</b>	<b>17</b>
<b>8. The Cluster Configuration File</b>	<b>19</b>
<b>9. Configuration Considerations</b>	<b>23</b>
9.1. Locking Considerations .....	23
9.2. Additional Configuration Considerations .....	23
<b>A. Revision History</b>	<b>25</b>
<b>Index</b>	<b>27</b>



---

# Introduction

## 1. About This Guide

This book describes a procedure for configuring NFS over GFS in a 5-node Red Hat Cluster using Conga.

## 2. Audience

This book is intended to be used by system administrators managing systems running the Linux operating system. It requires familiarity with Red Hat Enterprise Linux 5, Red Hat Cluster Suite, GFS file system administration, and basic knowledge of NFS administration.

## 3. Software Versions

Table 1. Software Versions

Software	Description
RHEL5	refers to RHEL5 and higher
GFS	refers to GFS for RHEL5 and higher

## 4. Related Documentation

For more information about using Red Hat Enterprise Linux, refer to the following resources:

- *Red Hat Enterprise Linux Installation Guide* — Provides information regarding installation of Red Hat Enterprise Linux 5.
- *Red Hat Enterprise Linux Deployment Guide* — Provides information regarding the deployment, configuration and administration of Red Hat Enterprise Linux 5.

For more information about Red Hat Cluster Suite for Red Hat Enterprise Linux 5, refer to the following resources:

- *Red Hat Cluster Suite Overview* — Provides a high level overview of the Red Hat Cluster Suite.
- *Configuring and Managing a Red Hat Cluster* — Provides information about installing, configuring and managing Red Hat Cluster components.
- *LVM Administrator's Guide: Configuration and Administration* — Provides a description of the Logical Volume Manager (LVM), including information on running LVM in a clustered environment.
- *Global File System: Configuration and Administration* — Provides information about installing, configuring, and maintaining Red Hat GFS (Red Hat Global File System).
- *Global File System 2: Configuration and Administration* — Provides information about installing, configuring, and maintaining Red Hat GFS2 (Red Hat Global File System 2).
- *Using Device-Mapper Multipath* — Provides information about using the Device-Mapper Multipath feature of Red Hat Enterprise Linux 5.
- *Using GNBD with Global File System* — Provides an overview on using Global Network Block Device (GNBD) with Red Hat GFS.

## Introduction

---

- *Linux Virtual Server Administration* — Provides information on configuring high-performance systems and services with the Linux Virtual Server (LVS).
- *Red Hat Cluster Suite Release Notes* — Provides information about the current release of Red Hat Cluster Suite.

Red Hat Cluster Suite documentation and other Red Hat documents are available in HTML, PDF, and RPM versions on the Red Hat Enterprise Linux Documentation CD and online at <http://www.redhat.com/docs/>.

## 5. Feedback

If you spot a typo, or if you have thought of a way to make this manual better, we would love to hear from you. Please submit a report in Bugzilla (<http://bugzilla.redhat.com/bugzilla/>) against the component **Documentation-cluster**.

Be sure to mention the manual's identifier:

```
Bugzilla component: Documentation-cluster
Book identifier: Cluster_Config_Example_NFS_GFS(EN)-5 (2011-07-21T15:20)
```

By mentioning this manual's identifier, we know exactly which version of the guide you have.

If you have a suggestion for improving the documentation, try to be as specific as possible. If you have found an error, please include the section number and some of the surrounding text so we can find it easily.

## 6. Document Conventions

This manual uses several conventions to highlight certain words and phrases and draw attention to specific pieces of information.

In PDF and paper editions, this manual uses typefaces drawn from the *Liberation Fonts*<sup>1</sup> set. The Liberation Fonts set is also used in HTML editions if the set is installed on your system. If not, alternative but equivalent typefaces are displayed. Note: Red Hat Enterprise Linux 5 and later includes the Liberation Fonts set by default.

### 6.1. Typographic Conventions

Four typographic conventions are used to call attention to specific words and phrases. These conventions, and the circumstances they apply to, are as follows.

#### **Mono-spaced Bold**

Used to highlight system input, including shell commands, file names and paths. Also used to highlight keycaps and key combinations. For example:

To see the contents of the file **my\_next\_bestselling\_novel** in your current working directory, enter the **cat my\_next\_bestselling\_novel** command at the shell prompt and press **Enter** to execute the command.

---

<sup>1</sup> <https://fedorahosted.org/liberation-fonts/>

The above includes a file name, a shell command and a keycap, all presented in mono-spaced bold and all distinguishable thanks to context.

Key combinations can be distinguished from keycaps by the hyphen connecting each part of a key combination. For example:

Press **Enter** to execute the command.

Press **Ctrl+Alt+F2** to switch to the first virtual terminal. Press **Ctrl+Alt+F1** to return to your X-Windows session.

The first paragraph highlights the particular keycap to press. The second highlights two key combinations (each a set of three keycaps with each set pressed simultaneously).

If source code is discussed, class names, methods, functions, variable names and returned values mentioned within a paragraph will be presented as above, in **mono-spaced bold**. For example:

File-related classes include **filesystem** for file systems, **file** for files, and **dir** for directories. Each class has its own associated set of permissions.

### Proportional Bold

This denotes words or phrases encountered on a system, including application names; dialog box text; labeled buttons; check-box and radio button labels; menu titles and sub-menu titles. For example:

Choose **System** → **Preferences** → **Mouse** from the main menu bar to launch **Mouse Preferences**. In the **Buttons** tab, click the **Left-handed mouse** check box and click **Close** to switch the primary mouse button from the left to the right (making the mouse suitable for use in the left hand).

To insert a special character into a **gedit** file, choose **Applications** → **Accessories** → **Character Map** from the main menu bar. Next, choose **Search** → **Find...** from the **Character Map** menu bar, type the name of the character in the **Search** field and click **Next**. The character you sought will be highlighted in the **Character Table**. Double-click this highlighted character to place it in the **Text to copy** field and then click the **Copy** button. Now switch back to your document and choose **Edit** → **Paste** from the **gedit** menu bar.

The above text includes application names; system-wide menu names and items; application-specific menu names; and buttons and text found within a GUI interface, all presented in proportional bold and all distinguishable by context.

### ***Mono-spaced Bold Italic*** or ***Proportional Bold Italic***

Whether mono-spaced bold or proportional bold, the addition of italics indicates replaceable or variable text. Italics denotes text you do not input literally or displayed text that changes depending on circumstance. For example:

To connect to a remote machine using ssh, type **ssh *username@domain.name*** at a shell prompt. If the remote machine is **example.com** and your username on that machine is john, type **ssh *john@example.com***.

The **mount -o remount *file-system*** command remounts the named file system. For example, to remount the **/home** file system, the command is **mount -o remount */home***.

To see the version of a currently installed package, use the **rpm -q *package*** command. It will return a result as follows: ***package-version-release***.

## Introduction

---

Note the words in bold italics above — *username*, *domain.name*, *file-system*, *package*, *version* and *release*. Each word is a placeholder, either for text you enter when issuing a command or for text displayed by the system.

Aside from standard usage for presenting the title of a work, italics denotes the first use of a new and important term. For example:

Publican is a *DocBook* publishing system.

## 6.2. Pull-quote Conventions

Terminal output and source code listings are set off visually from the surrounding text.

Output sent to a terminal is set in **mono-spaced roman** and presented thus:

```
books      Desktop  documentation  drafts  mss    photos  stuff  svn
books_tests Desktop1  downloads      images  notes  scripts svgs
```

Source-code listings are also set in **mono-spaced roman** but add syntax highlighting as follows:

```
package org.jboss.book.jca.ex1;

import javax.naming.InitialContext;

public class ExClient
{
    public static void main(String args[])
        throws Exception
    {
        InitialContext iniCtx = new InitialContext();
        Object          ref    = iniCtx.lookup("EchoBean");
        EchoHome        home   = (EchoHome) ref;
        Echo             echo   = home.create();

        System.out.println("Created Echo");

        System.out.println("Echo.echo('Hello') = " + echo.echo("Hello"));
    }
}
```

## 6.3. Notes and Warnings

Finally, we use three visual styles to draw attention to information that might otherwise be overlooked.



### Note

Notes are tips, shortcuts or alternative approaches to the task at hand. Ignoring a note should have no negative consequences, but you might miss out on a trick that makes your life easier.



### Important

Important boxes detail things that are easily missed: configuration changes that only apply to the current session, or services that need restarting before an update will apply. Ignoring a box labeled 'Important' will not cause data loss but may cause irritation and frustration.



**Warning**

Warnings should not be ignored. Ignoring warnings will most likely cause data loss.



# NFS over GFS in a Red Hat Cluster

This document provides a procedure to configure an NFS service in a Red Hat cluster using the Conga configuration tool. The configured cluster will have the following characteristics:

- There are 5 nodes in the cluster.
- The NFS service runs over a GFS file system.
- There are five NFS clients.

Note that this configuration is not a "high capacity" configuration in the sense that more than one server is providing NFS service. In this configuration, the floating IP moves about as needed, but only one server is active at a time.

*Figure 1.1, "NFS over GFS in a 5-Node Cluster"* shows the NFS over GFS configuration that this procedure yields.

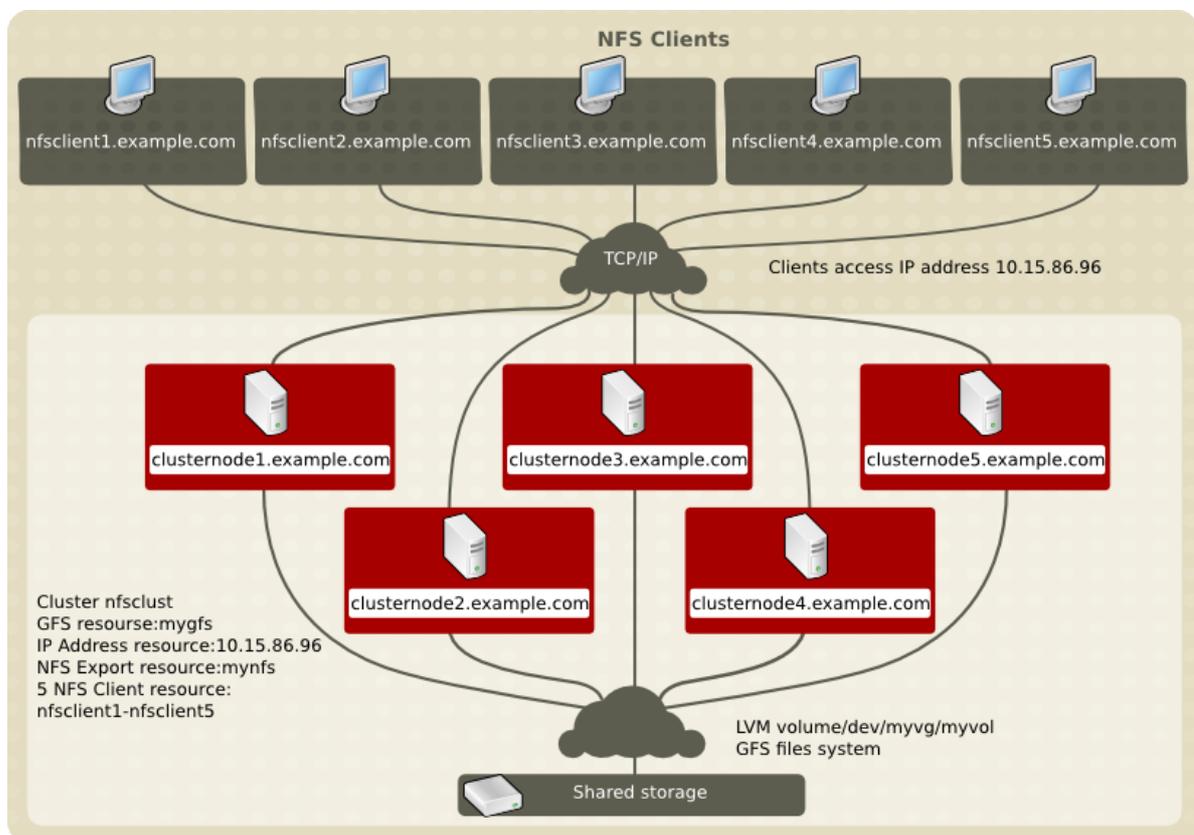


Figure 1.1. NFS over GFS in a 5-Node Cluster

In this configuration, Cluster Suite assigns which server will run the NFS service. If that server goes down, the NFS service will automatically fail over to another server in the cluster. The client will not be aware of any loss of service.

The NFS resource that this configuration defines will be NFS Version 3 by default. If you need to restrict what NFS protocol your system provides to its clients, you can do this at NFS startup; this is not part of cluster configuration.

This remainder of this document is organized as follows:

- [Chapter 2, Prerequisite Configuration](#) describes the prerequisite configuration components that have been set up before the procedure documented in this manual begins.

- [Chapter 3, Components to Configure](#) summarizes the cluster resources that this procedure configures.
- [Chapter 4, Configuring the Cluster Resources](#) provides the procedures for configuring the cluster resources needed for an NFS service.
- [Chapter 5, Configuring an NFS Cluster Service](#) provides the procedure for configuring an NFS service in a Red Hat Cluster Suite.
- [Chapter 6, Testing the NFS Cluster Service](#) provides a procedure to check that the NFS service is working and that it will continue to work as expected if one of the nodes goes down.
- [Chapter 7, Troubleshooting](#) provides some guidelines to follow when your configuration does not behave as expected.
- [Chapter 8, The Cluster Configuration File](#) shows the cluster configuration file as it appears before configuring the NFS service and after configuration the NFS service in a Red Hat Cluster Suite.
- [Chapter 9, Configuration Considerations](#) summarizes some general concerns to consider when configuring an NFS service over a GFS file system in a Red Hat Cluster Suite.

# Prerequisite Configuration

This document describes a procedure for configuring cluster resources and a cluster service. Before beginning this procedure, you must configure a cluster with a fencing device configured for each node in the cluster. In addition, you must configure a logical volume with a GFS filesystem that can be shared among the nodes in the cluster.

[Table 2.1, “Configuration Prerequisites”](#) summarizes the prerequisite configuration components that have been set up before this procedure begins.

Table 2.1. Configuration Prerequisites

Component	Name	Comment
cluster	nfsclust	five-node cluster
cluster node	clusternode1.example.com	node in cluster nfsclust configured with a fencing device of nfs-apc
cluster node	clusternode2.example.com	node in cluster nfsclust configured with a fencing device of nfs-apc
cluster node	clusternode3.example.com	node in cluster nfsclust configured with a fencing device of nfs-apc
cluster node	clusternode4.example.com	node in cluster nfsclust configured with a fencing device of nfs-apc
cluster node	clusternode5.example.com	node in cluster nfsclust configured with a fencing device of nfs-apc
LVM volume	/dev/myvg/myvol	The LVM device on which the GFS file system is created
GFS file system		The GFS file system to export by means of NFS, built on LVM volume <b>/dev/myvg/myvol</b> , mounted at <b>/mnt/gfs</b> , and shared among the members of cluster <b>nfsclust</b>
IP address	10.15.86.96	The IP address for the NFS service
NFS Client	nfsclient1.example.com	System that will access the NFS service
NFS Client	nfsclient2.example.com	System that will access the NFS service
NFS Client	nfsclient3.example.com	System that will access the NFS service
NFS Client	nfsclient4.example.com	System that will access the NFS service
NFS Client	nfsclient5.example.com	System that will access the NFS service

No failover domain has been defined for this cluster. The cluster software determines on which node to run the NFS service.

For information on configuring a cluster, see *Configuring and Managing a Red Hat Cluster*. For information on configuring an LVM volume, see *LVM Administrator's Guide: Configuration and Administration*.



# Components to Configure

This procedure documents the configuration of cluster resources and an NFS cluster service that will be named **nfssvc**.

[Table 3.1, “Cluster Resources to Configure”](#) summarizes the cluster resources that this procedure configures.

**Table 3.1. Cluster Resources to Configure**

Resource Type	Resource Name	Description
IP Address	10.15.86.96	The IP address for the NFS service
GFS	mygfs	The GFS file system that will be exported through the NFS service
NFS Export	mynfs	
NFS Client	nfsc1ient1	The NFS client system <b>nfsc1ient1.example.com</b>
NFS Client	nfsc1ient2	The NFS client system <b>nfsc1ient2.example.com</b>
NFS Client	nfsc1ient3	The NFS client system <b>nfsc1ient3.example.com</b>
NFS Client	nfsc1ient4	The NFS client system <b>nfsc1ient4.example.com</b>
NFS Client	nfsc1ient5	The NFS client system <b>nfsc1ient5.example.com</b>

The procedure for configuring these resources and the specific parameters to configure are described in [Chapter 4, Configuring the Cluster Resources](#).

You configure a Cluster service by adding a service and defining the composition of the service, which consists of the resources that the service requires. In this procedure, you will create an NFS cluster service named **nfssvc**. [Table 3.2, “Parameters to Configure for NFS Cluster Service nfssvc”](#) summarizes the resource configuration of **nfssvc**. The names of the resources are those that you assign when you define them, as noted in [Table 3.1, “Cluster Resources to Configure”](#)

**Table 3.2. Parameters to Configure for NFS Cluster Service nfssvc**

Resource	Name	Comment
IP Address Resource		
GFS Resource	mygfs	
NFS Export Resource	mynfs	NFS Export resource mynfs is a child of GFS resource mygfs.
NFS Client Resource	nfsc1ient1	NFS Client resource nfsc1ient1 is a child of NFS Export resource mynfs.
NFS Client Resource	nfsc1ient2	NFS Client resource nfsc1ient2 is a child of NFS Export resource mynfs.
NFS Client Resource	nfsc1ient3	NFS Client resource nfsc1ient3 is a child of NFS Export resource mynfs.
NFS Client Resource	nfsc1ient4	NFS Client resource nfsc1ient4 is a child of NFS Export resource mynfs.
NFS Client Resource	nfsc1ient5	NFS Client resource nfsc1ient5 is a child of NFS Export resource mynfs.

The procedure for configuring the NFS service is described in [Chapter 5, Configuring an NFS Cluster Service](#).



# Configuring the Cluster Resources

This chapter provides the procedures for configuring the cluster resources needed for an NFS service. For the procedure for configuring the NFS service itself, see [Chapter 5, Configuring an NFS Cluster Service](#).

Before you can configure an NFS service for a GFS file system in a cluster, you must add the following resources to your cluster:

- The IP address for the NFS service, as described in [Section 4.1, “Configuring an IP Address Resource”](#).
- The GFS file system, as described in [Section 4.2, “Configuring a GFS Resource”](#).
- The NFS export, as described in [Section 4.3, “Configuring an NFS Export Resource”](#).
- The NFS clients, as described in [Section 4.4, “Configuring NFS Client Resources”](#).

To add a resource to your cluster using **Conga**, perform the following procedure:

1. As an administrator of **luci** Select the **cluster** tab.
2. From the **Choose a cluster to administer** screen, select the cluster to which you will add resources. In this example, that is the cluster with the name **nfsclust**.
3. At the menu for cluster **nfsclust** (below the **clusters** menu), click **Resources**. This causes the display of menu items for resource configuration: **Add a Resource** and **Configure a Resource**.
4. Click **Add a Resource**. Clicking **Add a Resource** causes the **Add a Resource** page to be displayed.

The following sections provide instructions for adding the resources you need for an NFS service.

## 4.1. Configuring an IP Address Resource

Use the following procedure to add the IP Address resource 10.15.86.96 to cluster **nfsclust**.

1. At the **Add a Resource** page for cluster **nfsclust**, click the drop-down box under **Select a Resource Type** and select **IP Address**
2. For **IP Address**, enter 10.15.86.96.
3. Leave the **Monitor Link** checkbox selected to enable link status monitoring of the IP address resource.
4. Click **Submit**. Clicking **Submit** displays a verification page. Verifying that you want to add this resource displays a progress page followed by the display of **Resources** page, which displays the resources that have been configured for the cluster.

## 4.2. Configuring a GFS Resource

Use the following procedure to add the GFS file system resource **mygfs** to cluster **nfsclust**.

1. At the **Add a Resource** page for cluster **nfsclust**, click the drop-down box under **Select a Resource Type** and select **GFS file system**
2. For **Name**, enter **mygfs**.

3. For **Mount point**, enter `/mnt/gfs`. This is the path to which the GFS file system is mounted.
4. For **Device**, enter `/dev/myvg/myvol`. This is the LVM logical volume on which the GFS file system was created.
5. The **Options** field specifies the mount options for the GFS file system. For this example, we are mounting the file system with the `rw` (read-write) and `locallocks` option.
6. Leave the **File System ID** field blank. Leaving the field blank causes a file system ID to be assigned automatically after you click **Submit** at the **File System Resource Configuration** dialog box.
7. Leave the **Force Unmount** checkbox unchecked. **Force Unmount** kills all processes using the mount point to free up the mount when it tries to unmount. With GFS resources, the mount point is *not* unmounted at service tear-down *unless* this box is checked.
8. Click **Submit** and accept the verification screen.

### 4.3. Configuring an NFS Export Resource

Use the following procedure to add NFS export resource `mynfs` to cluster `nfsclust`.

1. At the **Add a Resource** page for cluster `nfsclust`, click the drop-down box under **Select a Resource Type** and select **NFS Export**
2. For **Name**, enter `mynfs`.
3. Click **Submit** and accept the verification screen.

The NFS Export resource that this configuration defines will be NFS Version 3 by default. If you need to restrict what NFS protocol your system provides to its clients, you can do this at NFS startup; this is not part of cluster configuration.

### 4.4. Configuring NFS Client Resources

This example procedure configures five NFS client resources for cluster `nfsclust`. The procedure for configuring the first two clients only is laid out explicitly.

Use the following procedure to add NFS client resource `nfsclient1` to cluster `nfsclust`.

1. At the **Add a Resource** page for cluster `nfsclust`, click the drop-down box under **Select a Resource Type** and select **NFS client**
2. For **Name**, enter `nfsclient1`.
3. For **Target**, enter `nfsclient1.example.com`. This is the first NFS client system.
4. This **Options** field specifies additional client access rights. Specify `rw` (read-write) in this field. For more information, refer to the General Options section of the `exports(5)` man page.
5. Check the **Allow Recover** checkbox. This indicates that if someone removes the export from the export list, the system will recover the export inline without taking down the NFS service.
6. Click **Submit** and accept the verification screen.

Use the following procedure to add NFS client resource `nfsclient2` to cluster `nfsclust`.

1. At the **Add a Resource** page for cluster **nfsclust**, click the drop-down box under **Select a Resource Type** and select **NFS client**
2. For **Name**, enter **nfscient2**.
3. For **Target**, enter **nfscient2.example.com**. This is the second NFS client system.
4. Leave the **Options** field blank.
5. Check the **Allow Recover** checkbox.
6. Click **Submit** and accept the verification screen.

Use the same procedure to configure the remaining three NFS client resources, using **nfscient3**, **nfscient4**, and **nfscient5** as the names of the resources and using **nfscient3.example.com**, **nfscient4.example.com**, and **nfscient5.example.com** as the targets.



# Configuring an NFS Cluster Service

This chapter provides the procedure for configuring an NFS service in a Red Hat Cluster Suite.

In order to configure an NFS service for a GFS file system in a cluster, you need to perform the following steps:

- Add a service to the cluster and provide a name for the service, as described in [Section 5.1, “Add a Service to the Cluster”](#).
- Add an IP address resource to service, as described in [Section 5.2, “Adding an IP Address Resource to an NFS Service”](#).
- Add a GFS resource to the service, as described in [Section 5.3, “Adding a GFS Resource to an NFS Service”](#).
- Add an NFS export resource to the service, as described in [Section 5.4, “Adding an NFS Export Resource to an NFS Service”](#).
- Add the NFS client resources to the services, as described in [Section 5.5, “Adding NFS Client Resources to an NFS Service”](#).

Some Red Hat Cluster resources must be configured as parent or child resources in a service. A resource that is configured as a parent resource is started before the children of that resource are started. All resources that are configured as a child resources of a parent must stop cleanly before a parent resource may be stopped. For a resource that is configured as a parent resource to be considered in good health, all the resources configured as its child resources must be in good health.

When you configure the NFS export resource for an NFS service, you must configure it as a child of the GFS resource. You must configure each NFS client resource for an NFS service as a child of the NFS export resource.

## 5.1. Add a Service to the Cluster

To add a service to your cluster using **Conga**, perform the following procedure:

1. As an administrator of **luci** Select the **cluster** tab.
2. From the **Choose a cluster to administer** screen, select the cluster to which you will add resources. In this example, that is the cluster with the name **nfsclust**.
3. At the menu for cluster **nfsclust** (below the **clusters** menu), click **Services**. This causes the display of menu items for service configuration: **Add a Service** and **Configure a Service**.
4. Click **Add a Service**. Clicking **Add a Service** causes the **Add a Service** page to be displayed.
5. For **Name**, enter **nfssvc**.
6. Leave the checkbox labeled **Automatically start this service** checked, which is the default setting. When the checkbox is checked, the service is started automatically when a cluster is started and running. If the checkbox is *not* checked, the service must be started manually any time the cluster comes up from the stopped state.
7. Leave the **Run Exclusive** checkbox unchecked. The **Run Exclusive** checkbox sets a policy wherein the service only runs on nodes that have *no other* services running on them. Since an NFS service consumes few resources, two services could run together on the same node without contention for resources and you do not need to check this.

8. For **Failover Domain**, leave the drop-down box default value of **None**. In this configuration, all of the nodes in the cluster may be used for failover.
9. For **Recovery Policy**, the drop-down box displays **Select a recovery policy**. Click the drop-down box and select **relocate**. This policy indicates that the system should relocate the service before restarting; it should not restart the node where the service is currently located.
10. Add the NFS service resources to this resource, as described in the following sections.
11. After you have added the NFS resources to the service, click **Submit**. The system prompts you to verify that you want to create this service. Clicking **OK** causes a progress page to be displayed followed by the display of **Services** page for the cluster. That page displays the services that have been configured for the cluster.

### 5.2. Adding an IP Address Resource to an NFS Service

Use the following procedure to add an IP Address resource to the NFS cluster service **nfssvc**.

1. At the **Add a Service** page for cluster **nfsc1ust**, click **Add a resource to this service**. Clicking **Add a resource to this service** causes the display of two drop-down boxes: **Add a new local resource** and **Use an existing global resource**.

For this example, we will use global resources, which are resources that were previously added as global resources. Adding a new local resource would add a resource that is available *only* to this service.

2. In the drop-down box underneath the **Use an existing global resource** display, click on the **Select a resource name** display. This displays the resources that have been defined for this cluster.
3. Select **10.15.86.96 (IP Address)**. This returns you to the **Add a Service** page with the IP Address resource displayed.

Leave the **Monitor link** checkbox selected, which is the default value. This enables link status monitoring of the IP address resource.

### 5.3. Adding a GFS Resource to an NFS Service

Use the following procedure to add a GFS resource to the NFS cluster service **nfssvc**.

1. At the **Add a Service** page for cluster **nfsc1ust**, click **Add a resource to this service**.
2. In the drop-down box underneath the **Use an existing global resource** display, click on the **Select a resource name** display.
3. Select **mygfs (GFS)**. This returns you to the **Add a Service** page with the GFS resource displayed, with the parameters that you defined in [Section 4.2, "Configuring a GFS Resource"](#) displayed.

### 5.4. Adding an NFS Export Resource to an NFS Service

Configure the NFS Export resource as a child of the GFS resource by following this procedure:

1. At the **Add a Service** page for cluster **nfsc1ust**, below the **GFS Resource Configuration** display, click **Add a child**. This causes the display of two drop-down boxes: **Add a new local resource** and **Use an existing global resource**.

2. In the drop-down box underneath the **Use an existing global resource** display, click on the **Select a resource name** display.
3. Select **my nfs (NFS Export)**. This returns you to the **Add a Service** page with the NFS Export resource displayed.

## 5.5. Adding NFS Client Resources to an NFS Service

Configure the NFS Client resources as children of the NFS export resource by following this procedure for each NFS client:

1. At the **Add a Service** page for cluster **nfsc1ust**, below the **NFS Export Resource Configuration** display, click **Add a child**. This causes the display of two drop-down boxes: **Add a new local resource** and **Use an existing global resource**.
2. Click on the **Select a resource name** display in the drop-down box underneath the **Use an existing global resource** display.
3. Select **nfsc1ent1 (NFS Client)**. This returns you to the **Add a Service** page with the NFS client resource displayed with the parameters you defined in [Section 4.4, "Configuring NFS Client Resources"](#).

Follow the same procedure to add a second, third, fourth, and fifth NFS client resource, selecting **nfsc1ent2 (NFS Client)**, **nfsc1ent3 (NFS Client)**, **nfsc1ent4 (NFS Client)**, and **nfsc1ent5 (NFS Client)** as the resources to add.

After you have added the NFS client resources to the service, you can click **Submit**. The system prompts you to verify that you want to create this service. Clicking **OK** causes a progress page to be displayed followed by the display of **Services** page for the cluster. That page displays the services that have been configured for the cluster.



## Testing the NFS Cluster Service

After you have configured the NFS service, you can check to be sure that the NFS service is working and that it will continue to work as expected if one of the nodes goes down. The following procedure tests an NFS mount on a client, fences the node on which the NFS service is running, and then checks to be sure that the NFS client can still access the file system.

1. If the GFS file system in the **nfsclust** cluster is currently empty, populate the file system with test data.
2. Log in to one of the client systems you defined as a target.
3. Mount the NFS file system on the client system, and check to see if the data on that file system is available.
4. On the **Luci** server, select **Nodes** from the menu for **nfsclust**. This displays the nodes in **nfsclust** and indicates which node is running the **nfssvc** service.
5. The drop-down box for each node displays **Choose a task**. For the node on which the **nfssvc** service is running, select **Fence this node**.
6. Refresh the screen. The **nfssvc** service should now be running in a different node.
7. On the client system, check whether the file system you mounted is still available. Even though the NFS service is now running on a different node in the cluster, the client system should detect no difference.
8. Restore the system to its previous state:
  - Unmount the file system from the client system.
  - Delete any test data you created in the GFS file system.
  - Click on **Choose a task** in the drop-down box for the node which you fenced and select **Reboot this node**.



### Note

For advanced troubleshooting once a system has gone into production, it is sometimes necessary to set up the **netconsole** and **kdump** services on a system. You may find it useful to implement and test these tools before a system goes into production, to help in troubleshooting down the line.



# Troubleshooting

If you find that you are seeing error messages when you try to configure your system, or if after configuration your system does not behave as expected, you can perform the following checks and examine the following areas.

- Connect to one of the nodes in the cluster and execute the **clustat(8)** command. This command runs a utility that displays the status of the cluster. It shows membership information, quorum view, and the state of all configured user services.

The following example shows the output of the **clustat(8)** command.

```
[root@clusternode4 ~]# clustat
Cluster Status for nfsclust @ Wed Dec  3 12:37:22 2008
Member Status: Quorate

Member Name                ID  Status
-----
clusternode5.example.com   1  Online, rgmanager
clusternode4.example.com   2  Online, Local, rgmanager
clusternode3.example.com   3  Online, rgmanager
clusternode2.example.com   4  Online, rgmanager
clusternode1.example.com   5  Online, rgmanager

Service Name                Owner (Last)                State
-----
service:nfssvc              clusternode2.example.com    starting
```

In this example, **clusternode4** is the local node since it is the host from which the command was run. If **rgmanager** did not appear in the **Status** category, it could indicate that cluster services are not running on the node.

- Connect to one of the nodes in the cluster and execute the **group\_tool(8)** command. This command provides information that you may find helpful in debugging your system. The following example shows the output of the **group\_tool(8)** command.

```
[root@clusternode1 ~]# group_tool
type      level name      id      state
fence     0      default  00010005 none
[1 2 3 4 5]
dlm       1      clvmd    00020005 none
[1 2 3 4 5]
dlm       1      rgmanager 00030005 none
[3 4 5]
dlm       1      mygfs    007f0005 none
[5]
gfs       2      mygfs    007e0005 none
[5]
```

The state of the group should be **none**. The numbers in the brackets are the node ID numbers of the cluster nodes in the group. The **clustat** shows which node IDs are associated with which nodes. If you do not see a node number in the group, it is not a member of that group. For example, if a node ID is not in dlm/rgmanager group, it is not using the rgmanager dlm lock space (and probably is not running rgmanager).

The level of a group indicates the recovery ordering. 0 is recovered first, 1 is recovered second, and so forth.

- Connect to one of the nodes in the cluster and execute the **cman\_tool nodes -f** command. This command provides information about the cluster nodes that you may want to look at. The following example shows the output of the **cman\_tool nodes -f** command.

```
[root@clusternode1 ~]# cman_tool nodes -f
Node Sts  Inc   Joined                Name
  1   M    752   2008-10-27 11:17:15 clusternode5.example.com
  2   M    752   2008-10-27 11:17:15 clusternode4.example.com
  3   M    760   2008-12-03 11:28:44 clusternode3.example.com
  4   M    756   2008-12-03 11:28:26 clusternode2.example.com
  5   M    744   2008-10-27 11:17:15 clusternode1.example.com
```

The **Sts** heading indicates the status of a node. A status of M indicates the node is a member of the cluster. A status of X indicates that the node is dead. The **Inc** heading indicates the incarnation number of a node, which is for debugging purposes only.

- Check whether the **cluster.conf** is identical in each node of the cluster. If you configure your system with Conga, as in the example provided in this document, these files should be identical, but one of the files may have accidentally been deleted or altered.
- In addition to using Conga to fence a node in order to test whether failover is working properly as described in [Chapter 6, Testing the NFS Cluster Service](#), you could disconnect the ethernet connection between cluster members. You might try disconnecting one, two, or three nodes, for example. This could help isolate where the problem is.
- If you are having trouble mounting or modifying an NFS volume, check whether the cause is one of the following:
  - The network between server and client is down.
  - The storage devices are not connected to the system.
  - More than half of the nodes in the cluster have crashed, rendering the cluster inquorate. This stops the cluster.
  - The GFS file system is not mounted on the cluster nodes.
  - The GFS file system is not writable.
  - The IP address you defined in the **cluster.conf** is not bounded to the correct interface / NIC (sometimes the **ip.sh** script does not perform as expected).
- Execute a **showmount -e** command on the node running the cluster service. If it shows up the right 5 exports, check your firewall configuration for all necessary ports for using NFS.
- If SELinux is currently in **enforcing** mode on your system, check your **/var/log/audit.log** file for any relevant messages. If you are using NFS to serve home directories, check whether the correct SELinux boolean value for **nfs\_home\_dirs** has been set to 1; this is required if you want to use NFS-based home directories on a client that is running SELinux. If you do not set this value on, you can mount the directories as root but cannot use them as home directories for your users.
- Check the **/var/log/messages** file for error messages from the NFS daemon.
- If you see the expected results locally at the cluster nodes and between the cluster nodes but not at the defined clients, check the firewall configuration at the clients.

# The Cluster Configuration File

Configuring a cluster with Conga modifies the cluster configuration file. This chapter shows the cluster configuration file before and after the procedures documented in [Chapter 4, Configuring the Cluster Resources](#) and [Chapter 5, Configuring an NFS Cluster Service](#) were performed.

Before the cluster resources and service were configured, the cluster.conf file appeared as follows.

```
<?xml version="1.0"?>
<cluster alias="nfsclust" config_version="1" name="nfsclust">
  <fence_daemon post_fail_delay="0" post_join_delay="3"/>
  <clusternodes>
    <clusternode name="clusternode1.example.com" nodeid="1" votes="1">
      <fence>
        <method name="apc-nfs">
          <device name="apc1" switch="3" port="1"/>
        </method>
      </fence>
    </clusternode>

    <clusternode name="clusternode2.example.com" nodeid="2" votes="1">
      <fence>
        <method name="apc-nfs">
          <device name="apc1" switch="3" port="2"/>
        </method>
      </fence>
    </clusternode>

    <clusternode name="clusternode3.example.com" nodeid="3" votes="1">
      <fence>
        <method name="apc-nfs">
          <device name="apc1" switch="3" port="3"/>
        </method>
      </fence>
    </clusternode>

    <clusternode name="clusternode4.example.com" nodeid="4" votes="1">
      <fence>
        <method name="apc-nfs">
          <device name="apc1" switch="3" port="4"/>
        </method>
      </fence>
    </clusternode>

    <clusternode name="clusternode5.example.com" nodeid="5" votes="1">
      <fence>
        <method name="apc-nfs">
          <device name="apc1" switch="3" port="5"/>
        </method>
      </fence>
    </clusternode>
  </clusternodes>
  <cman/>
  <fencedevices>
    <fencedevice name="apc1" agent="fence_apc" ipaddr="link-apc" login="apc"
      passwd="apc"/>
  </fencedevices>
</rm/>
</cluster>
```

After the cluster resources and service were configured, the cluster.conf file appeared as follows.



```

<?xml version="1.0"?>
<cluster alias="nfsclust" config_version="10" name="nfsclust">
  <fence_daemon clean_start="0" post_fail_delay="0" post_join_delay="3"/>
  <clusternodes>
    <clusternode name="clusternode1.example.com" nodeid="1" votes="1">
      <fence>
        <method name="apc-nfs">
          <device name="apc1" switch="3" port="1"/>
        </method>
      </fence>
    </clusternode>
    <clusternode name="clusternode2.example.com" nodeid="2" votes="1">
      <fence>
        <method name="apc-nfs">
          <device name="apc1" switch="3" port="2"/>
        </method>
      </fence>
    </clusternode>
    <clusternode name="clusternode3.example.com" nodeid="3" votes="1">
      <fence>
        <method name="apc-nfs">
          <device name="apc1" switch="3" port="3"/>
        </method>
      </fence>
    </clusternode>
    <clusternode name="clusternode4.example.com" nodeid="4" votes="1">
      <fence>
        <method name="apc-nfs">
          <device name="apc1" switch="3" port="4"/>
        </method>
      </fence>
    </clusternode>
    <clusternode name="clusternode5.example.com" nodeid="5" votes="1">
      <fence>
        <method name="apc-nfs">
          <device name="apc1" switch="3" port="5"/>
        </method>
      </fence>
    </clusternode>
  </clusternodes>
  <cman/>
  <fencedevices>
    <fencedevice name="apc1" agent="fence_apc" ipaddr="link-apc" login="apc"
      passwd="apc"/>
  </fencedevices>
  <rm>
    <failoverdomains/>
    <resources>
      <ip address="10.15.86.96" monitor_link="1"/>
      <clusterfs device="/dev/myvg/myvol" force_unmount="0"
        fsid="39669" fstype="gfs" mountpoint="/mnt/gfs"
        name="mygfs" options="rw,localflocks"/>
      <nfsexport name="my nfs"/>
      <nfsclient allow_recover="1" name="nfsclient1" options="rw"
        target="nfsclient1.example.com"/>
      <nfsclient allow_recover="1" name="nfsclient2" options="rw"
        target="nfsclient2.example.com"/>
      <nfsclient allow_recover="1" name="nfsclient3" options="rw"
        target="nfsclient3.example.com"/>
      <nfsclient allow_recover="1" name="nfsclient4" options="rw"
        target="nfsclient4.example.com"/>
      <nfsclient allow_recover="1" name="nfsclient5" options="rw"
        target="nfsclient5.example.com"/>
    </resources>
    <service autostart="1" exclusive="0" name="nfssvc" recovery="relocate">
      <ip ref="10.15.86.96"/>
      <clusterfs ref="mygfs">

```

---

```
        <nfsexport ref="my nfs">
            <nfsclient ref="nfsclient1"/>
            <nfsclient ref="nfsclient2"/>
            <nfsclient ref="nfsclient3"/>
            <nfsclient ref="nfsclient4"/>
            <nfsclient ref="nfsclient5"/>
        </nfsexport>
    </clusterfs>
</service>
</rm>
</cluster>
```



# Configuration Considerations

This chapter describes the caveats you should take into account when configuring an NFS service over a GFS or GFS2 file system.

## 9.1. Locking Considerations



### Warning

When a GFS or GFS2 filesystem is exported via NFS then you must mount the filesystem with the **locallocks** option. The intended effect of this is to allow the NFS server to manage locks on the GFS or GFS2 filesystem without the extra overhead of passing through the GFS and GFS2 locking layers.

For more information on the **locallocks** mount option and when it may be required, see the *Global File System* and *Global File System 2* manuals.

## 9.2. Additional Configuration Considerations

In addition to the locking considerations, you should take the following into account when configuring an NFS service over a GFS or GFS2 file system.

- Red Hat supports only Red Hat Cluster Suite configurations using NFSv3 with locking in an active/passive configuration with the following characteristics:
  - The backend file system is a GFS or GFS2 file system running on a 2 to 16 node cluster.
  - An NFSv3 server is defined as a service exporting the entire GFS/GFS2 file system from a single cluster node at a time.
  - The NFS server can fail over from one cluster node to another (active/passive configuration).
  - No access to the GFS/GFS2 file system is allowed *except* through the NFS server. This includes both local GFS/GFS2 file system access as well as access through Samba or Clustered Samba.
  - The GFS or GFS2 file system must be mounted with the **locallocks** option.
  - There is no NFS quota support on the system.

This configuration provides HA for the file system and reduces system downtime since a failed node does not result in the requirement to execute the **fsck** command when failing the NFS server from one node to another.

- The **fsid=** NFS option is mandatory for NFS exports of GFS/GFS2.
- There is currently an issue with failover and failback when using NFSv3 over GFS with TCP when the following scenario comes into play:
  1. Client A mounts from server 1.
  2. The system administrator moves NFS service from server 1 to server 2.
  3. The client resumes I/O operations.

4. The system administrator moves NFS service from server 2 to server 1.

In this situation, the NFS service on server 1 does not get shut down because this would render other NFS services inoperable.

Should this situation arise, you should move all NFS services off of server 1 and run the **service nfs restart**. After this you can safely migrate your NFS services back to server 1.

- If problems arise with your cluster (for example, the cluster becomes inquorate and fencing is not successful), the clustered logical volumes and the GFS/GFS2 file system will be frozen and no access is possible until the cluster is quorate. You should consider this possibility when determining whether a simple failover solution such as the one defined in this procedure is the most appropriate for your system.

---

# Appendix A. Revision History

**Revision 1.0-2 Thu Jul 21 2011**

**Steven Levine** [slevine@redhat.com](mailto:slevine@redhat.com)

Resolves: #676406

Clarifies information about necessity for mounting GFS file systems with the localflocks option.

**Revision 1.0-1 Tue Aug 3 2010**

**Steven Levine** [slevine@redhat.com](mailto:slevine@redhat.com)

Resolves: #562251

Clarifies support issues in "Configuration Considerations" chapter.

Resolves: #614526

Adds note about issues of Samba support to "Configuration Considerations" chapter.

**Revision 1.0 Thu Jan 29 2009**

**Steven Levine** [slevine@redhat.com](mailto:slevine@redhat.com)

First edition



---

# Index

## A

Allow Recover checkbox, 8

## C

child resource  
    configuration, 12, 13  
    definition, 11  
clustat command, 17  
cluster nodes, prerequisite, 3  
cluster resources, 5  
cluster service  
    adding GFS resource, 12  
    adding IP address resource, 12  
    adding NFS client resource, 13  
    adding NFS export resource, 12  
    adding to cluster, 11  
    composition, 5  
cluster, prerequisite, 3  
cluster.conf file, 19  
cman\_tool command, 17

## F

failover domain, 3, 11  
feedback, vi, vi  
Force Unmount checkbox, 7

## G

getenforce command, 17  
GFS file system  
    adding resource to cluster, 7  
    options, 7  
    prerequisite, 3  
    resource, 5  
GFS resource  
    adding to cluster service, 12  
group\_tool command, 17

## I

IP address  
    adding resource to cluster, 7  
    prerequisite, 3  
    resource, 5  
IP address resource  
    adding to cluster service, 12  
IP, floating, 1

## K

kdump service, 15

## L

LVM volume, prerequisite, 3

## M

Monitor Link checkbox, 7, 12

## N

netconsole service, 15  
NFS  
    failover and failback issues, 23  
    in a cluster, 23  
    locks, 23  
    startup, 1  
    Version 3, 1  
NFS client  
    adding resource to cluster, 8  
    options, 8  
    resource, 5  
NFS client resource  
    adding to cluster service, 13  
NFS client systems, prerequisite, 3  
NFS export  
    adding resource to cluster, 8  
    resource, 5  
NFS export resource  
    adding to cluster service, 12  
NFS service  
    configuration overview, 11  
    testing, 15

## P

parent resource  
    definition, 11  
POSIX locks, 23  
Prerequisite Configuration Components, 3

## R

recovery policy, 11  
resource  
    adding to cluster, 7  
    GFS file system, 7  
    IP address, 7  
    NFS client, 8  
    NFS export, 8

## S

SELinux enforcing mode, 17  
SELinux nfs\_home\_dirs value, 17  
service  
    adding to cluster, 11  
    automatic start, 11  
showmount command, 17

**T**

testing

    NFS service, 15