



# **Red Hat Enterprise Linux 6**

## **Sistema de archivos global 2 (GFS2)**

Sistema de archivos global 2 (GFS2) de Red Hat

Edición 7



# Red Hat Enterprise Linux 6 Sistema de archivos global 2 (GFS2)

---

Sistema de archivos global 2 (GFS2) de Red Hat  
Edición 7

## Legal Notice

Copyright © 2014 Red Hat, Inc. and others.

This document is licensed by Red Hat under the [Creative Commons Attribution-ShareAlike 3.0 Unported License](https://creativecommons.org/licenses/by-sa/3.0/). If you distribute this document, or a modified version of it, you must provide attribution to Red Hat, Inc. and provide a link to the original. If the document is modified, all Red Hat trademarks must be removed.

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, OpenShift, 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.

Node.js ® is an official trademark of Joyent. Red Hat Software Collections is not formally related to or endorsed by the official Joyent Node.js open source or commercial project.

The OpenStack ® Word Mark and OpenStack logo are either registered trademarks/service marks or trademarks/service marks of the OpenStack Foundation, in the United States and other countries and are used with the OpenStack Foundation's permission. We are not affiliated with, endorsed or sponsored by the OpenStack Foundation, or the OpenStack community.

All other trademarks are the property of their respective owners.

## Resumen

Este libro proporciona información sobre configuración y mantenimiento de GFS2 (Sistema de archivos global 2) de Red Hat para Red Hat Enterprise Linux 6.

## Table of Contents

<b>INTRODUCCIÓN</b> .....	<b>5</b>
1. AUDIENCIA	5
2. DOCUMENTACIÓN RELACIONADA	5
3. ¡NECESITAMOS SUS COMENTARIOS!	6
<b>CAPÍTULO 1. VISIÓN GENERAL DE GFS2</b> .....	<b>7</b>
1.1. FUNCIONALIDADES NUEVAS Y CAMBIADAS	8
1.1.1. Funcionalidades nuevas y cambiadas de Red Hat Enterprise Linux 6.0	8
1.1.2. Funcionalidades nuevas y cambiadas para Red Hat Enterprise Linux 6.1	8
1.1.3. Funcionalidades nuevas y cambiadas para Red Hat Enterprise Linux 6.2	9
1.1.4. Funcionalidades nuevas y cambiadas para Red Hat Enterprise Linux 6.3	9
1.1.5. Funcionalidades nuevas y cambiadas para Red Hat Enterprise Linux 6.4	9
1.1.6. Funcionalidades nuevas y cambiadas para Red Hat Enterprise Linux 6.6	10
1.2. ANTES DE CONFIGURAR GFS2	10
1.3. INSTALACIÓN DE GFS2	10
1.4. DIFERENCIAS ENTRE GFS Y GFS2	11
1.4.1. Nombres de los comandos de GFS2	11
1.4.2. Diferencias adicionales entre GFS y GFS2	12
Nombre de rutas dependientes del contexto	12
Módulo gfs2.ko	13
Cómo activar la activación de la cuota en GFS2	13
Diario de datos	13
Cómo añadir diarios dinámicamente	13
Parámetro atime_quantum eliminado	13
La opción data= del comando mount	13
El comando gfs2_tool	14
El comando gfs2_edit	14
1.4.3. Mejoras de rendimiento en GFS2	14
<b>CAPÍTULO 2. CONFIGURACIÓN DE GFS2 CONSIDERACIONES OPERACIONALES</b> .....	<b>16</b>
2.1. CONSIDERACIONES DE FORMATEO	16
2.1.1. Tamaño del sistema de archivos: Entre más pequeño, mejor	16
2.1.2. Tamaño de bloque: El predeterminado es el preferido (4K)	16
2.1.3. Número de diarios: uno por cada nodo que monte	17
2.1.4. Tamaño de diario: El predeterminado suele ser óptimo (128MB)	17
2.1.5. Tamaño y número de grupos de recursos	17
2.2. FRAGMENTACIÓN DE SISTEMA DE ARCHIVOS	18
2.3. PROBLEMAS DE ASIGNACIÓN DE BLOQUES	18
2.3.1. Abandonar espacio libre en el sistema de archivos	18
2.3.2. Procure que cada nodo asigne sus propios archivos	19
2.3.3. Preasignar, si es posible	19
2.4. CONSIDERACIONES DE CLÚSTER	19
2.5. CONSIDERACIONES DE USO	19
2.5.1. Opciones de montaje: noatime y nodiratime	19
2.5.2. Opciones de ajuste DLM: Aumento de tamaño de tablas DLM	20
2.5.3. Opciones de ajuste VFS: Investigación y experimentación	20
2.5.4. SELinux: Evitar SELinux en GFS2	20
2.5.5. Configuración de NFS en GFS2	20
2.5.6. Archivo Samba (SMB o Windows) sirviendo en GFS2	22
2.6. COPIAS DE SEGURIDAD DE SISTEMA DE ARCHIVOS	22
2.7. CONSIDERACIONES DE HARDWARE	22
2.8. PROBLEMAS DE RENDIMIENTO: REVISE EL PORTAL DEL CLIENTE DE RED HAT	23

2.9. BLOQUEO DE NODO GFS2	23
2.9.1. Problemas con Bloqueo de Posix	24
2.9.2. Ajuste de rendimiento con GFS2	24
2.9.3. Resolución de problemas de rendimiento de GFS2 con el volcado de cerrojo GFS2	25
<b>CAPÍTULO 3. CÓMO INICIAR</b>	<b>29</b>
3.1. TAREAS PREVIAS	29
3.2. TAREAS DE CONFIGURACIÓN INICIAL	29
<b>CAPÍTULO 4. ADMINISTRACIÓN DE GFS2</b>	<b>31</b>
4.1. CÓMO CREAR UN SISTEMA DE ARCHIVOS	31
Uso	32
Ejemplos	33
Opciones completas	33
4.2. CÓMO MONTAR UN SISTEMA DE ARCHIVOS	35
Uso	36
Ejemplo	36
Uso completo	36
4.3. CÓMO DESMONTAR UN SISTEMA DE ARCHIVOS	39
Uso	39
4.4. CONSIDERACIONES ESPECIALES EN EL MONTAJE DE LOS SISTEMAS DE ARCHIVOS GFS2	40
4.5. ADMINISTRACIÓN DE CUOTAS EN GFS2	40
4.5.1. Configuración de cuotas de disco	40
4.5.1.1. Cómo establecer cuotas en modo impositivo o de contabilidad	41
Uso	41
Ejemplos	41
4.5.1.2. Cómo crear archivos de base de datos de cuotas	42
4.5.1.3. Cómo asignar cuotas por usuario	42
4.5.1.4. Cómo asignar cuotas por grupo	43
4.5.2. Cómo administrar cuotas de disco	44
4.5.3. Cómo mantener cuotas exactas	44
4.5.4. Sincronización de cuotas con el comando quotasync	45
Uso	45
Ejemplos	46
4.5.5. Referencias	46
4.6. CÓMO EXPANDIR UN SISTEMA DE ARCHIVOS	46
Uso	46
Comentarios	47
Ejemplos	47
Uso completo	47
4.7. CÓMO AÑADIR DIARIOS A UN SISTEMA DE ARCHIVOS	48
Uso	48
Ejemplos	48
Uso completo	49
4.8. DIARIO DE DATOS	49
4.9. CÓMO CONFIGURAR LAS ACTUALIZACIONES ATIME	50
4.9.1. Montar con relatime	51
Uso	51
Ejemplo	51
4.9.2. Montar con noatime	51
Uso	51
Ejemplo	52
4.10. CÓMO SUSPENDER LA ACTIVIDAD EN UN SISTEMA DE ARCHIVOS	52

Uso	52
Ejemplos	52
4.11. CÓMO REPARAR UN SISTEMA DE ARCHIVOS	53
Uso	54
Ejemplo	54
4.12. NOMBRES DE RUTAS DEPENDIENTES DEL CONTEXTO Y MONTAJES ENLAZADOS	55
4.13. ORDEN DE MONTAJE DE ENLACE Y MONTAJE DE SISTEMA DE ARCHIVOS	56
4.14. LA FUNCIÓN DE RETIRO DE GFS2	58
<b>CAPÍTULO 5. DIAGNÓSTICO Y CORRECCIÓN DE PROBLEMAS CON SISTEMAS DE ARCHIVOS GFS2</b>	<b>60</b>
5.1. EL SISTEMA DE ARCHIVOS GFS2 PRESENTA UN BAJO RENDIMIENTO	60
5.2. EL SISTEMA DE ARCHIVOS GFS2 SE CUELGA Y REQUIERE REARRANQUE DE UN NODO	60
5.3. EL SISTEMA DE ARCHIVOS GFS2 SE CUELGA Y DEBE REARRANCAR TODOS LOS NODOS	60
5.4. EL SISTEMA DE ARCHIVOS GFS2 NO SE MONTA EN UN NODO DE CLÚSTER RECIÉN AÑADIDO	61
5.5. ESPACIO INDICADO COMO USADO EN SISTEMA DE ARCHIVOS VACÍO	61
<b>CAPÍTULO 6. CÓMO CONFIGURAR EL SISTEMA DE ARCHIVOS GFS2 EN UN CLÚSTER PACEMAKER</b>	<b>62</b>
<b>APÉNDICE A. ADMINISTRACIÓN DE CUOTAS GFS2 CON EL COMANDO GFS2_QUOTA</b>	<b>64</b>
A.1. CÓMO ESTABLECER CUOTAS CON EL COMANDO GFS2_QUOTA	64
Uso	64
Ejemplos	65
A.2. DESPLIEGUE DE LÍMITES DE CUOTAS Y USO CON EL COMANDO GFS2_QUOTA	65
Uso	65
Salida de comando	66
Comentarios	66
Ejemplos	66
A.3. SINCRONIZACIÓN DE CUOTAS CON EL COMANDO GFS2_QUOTA	67
Uso	67
Ejemplos	67
A.4. CÓMO REALIZAR ACTIVACIÓN Y DESACTIVACIÓN DE CUOTAS	68
Uso	68
Ejemplos	68
A.5. ACTIVACIÓN DE CONTABILIDAD DE CUOTAS	68
Uso	68
Ejemplo	69
<b>APÉNDICE B. CÓMO CONVERTIR EL SISTEMA DE ARCHIVOS DE GFS A GFS2</b>	<b>70</b>
B.1. CONVERSIÓN DE NOMBRES DE RUTAS DEPENDIENTES DE CONTEXTO	70
B.2. PROCEDIMIENTO DE CONVERSIÓN DE GFS A GFS2	71
<b>APÉNDICE C. PUNTOS DE TRAZADO GFS2 Y ARCHIVO GLOCKS DEBUGFS</b>	<b>72</b>
C.1. TIPOS DE PUNTOS DE TRAZADO GFS2	72
C.2. PUNTOS DE TRAZADO	72
C.3. GLOCKS	73
C.4. INTERFAZ DE GLOCK DEBUGFS	75
C.5. PORTADORES DE GLOCK	78
C.6. PUNTOS DE TRAZADO GLOCK	79
C.7. PUNTOS DE TRAZADO DE BMAP	79
C.8. PUNTOS DE TRAZADO DE REGISTRO	80
C.9. ESTADÍSTICAS DE GLOCK	80
C.10. REFERENCIAS	81
<b>APÉNDICE D. HISTORIA DE REVISIONES</b>	<b>82</b>

**ÍNDICE** ..... **84**

# INTRODUCCIÓN

Este libro proporciona información sobre configuración y mantenimiento de GFS2 de Red Hat (Red Hat Global File System 2/Sistema de archivos global 2 de Red Hat), el cual se incluye en la adición de almacenaje resistente.

## 1. AUDIENCIA

Este libro está destinado principalmente a los administradores de sistemas de Linux con conocimientos en:

- Procedimientos de administración de sistemas de linux, incluyendo la configuración del Kernel
- Instalación y configuración de redes de almacenaje compartido, tales como canales de fibra SAN

## 2. DOCUMENTACIÓN RELACIONADA

Para más información sobre el uso de Red Hat Enterprise Linux, consulte los siguientes recursos:

- *Guía de instalación* — Proporciona información sobre la instalación de Red Hat Enterprise Linux 6.
- *Guía de implementación* — Proporciona información importante sobre la implementación, configuración y administración de Red Hat Enterprise Linux 6.
- *Guía de administración de almacenamiento* — Proporciona información para administrar efectivamente los dispositivos de almacenamiento y sistemas de archivos en Red Hat Enterprise Linux 6.

Para obtener más información sobre las Adiciones de Alta disponibilidad y las Adiciones de almacenaje resistente para Red Hat Enterprise Linux 6, consulte los siguientes recursos:

- *Vista General de adiciones de Alta disponibilidad*— Proporciona una visión general de alto nivel de las adiciones de Alta disponibilidad de Red Hat.
- *Administración de clúster*— Proporciona información sobre instalación, configuración y administración de adiciones de Alta disponibilidad.
- *Administración del Gestor de volúmenes lógicos*— Proporciona una descripción del Gestor de Volúmenes Lógicos (LVM) e incluye información sobre la ejecución de LVM en un entorno de clúster.
- *DM Multipath* — Proporciona información sobre el uso del Mapeador de dispositivo Multipath de Red Hat Enterprise Linux.
- *Administración del equilibrador de cargas*— Proporciona información sobre configuración de sistemas de alto rendimiento y servicio con la adición de equilibrador de cargas, un conjunto de componentes de software integrados que proporcionan Servidores Virtuales de Linux (LVS) para equilibrio de carga IP a través de un conjunto de servidores reales.

- *Notas de lanzamiento* — Proporciona información sobre el lanzamiento actual de los productos de Red Hat.

La documentación de Red Hat Cluster Suite y otros documentos de Red Hat están disponibles en versiones HTML, PDF, y RPM en el CD de documentación de Red Hat Enterprise Linux y en línea en <https://access.redhat.com/site/documentation/>.

### 3. ¡NECESITAMOS SUS COMENTARIOS!

Si encuentra algún error tipográfico en esta guía o si tiene alguna sugerencia para mejorarla, nos gustaría saberlo. Por favor complete un reporte en Bugzilla: <http://bugzilla.redhat.com/> con el nombre del producto **Red Hat Enterprise Linux 6** y el componente **doc-Global\_File\_System\_2**. Al enviar el informe de errores, asegúrese de mencionar el identificador de la guía:

rh-gfs2(EN)-6 (2014-10-8T15:15)

En caso de que tenga una sugerencia para mejorar la documentación, trate de ser lo más específico posible. Si encontró algún error, por favor incluya el número de la sección y parte del contexto para que pueda localizarse con facilidad.

## CAPÍTULO 1. VISION GENERAL DE GFS2

El sistema de archivos GFS2 de Red Hat está incluido en la adición de almacenaje resistente. Es un sistema de archivos nativo que interactúa directamente con la interfaz del sistema de archivos del kernel de Linux (capa de VFS). Un sistema de archivos GFS2 emplea metadatos distribuidos y varios diarios. Red Hat soporta el uso de sistemas de archivos GFS2 únicamente como se implementan en la adición de Alta disponibilidad.



### NOTA

Aunque un sistema de archivos GFS2 puede ser implementado en un sistema autónomo o como parte de una configuración de clúster, para el lanzamiento de Red Hat Enterprise Linux 6 Red Hat no soporta el uso de GFS2 como un sistema de archivos de nodo único. Red Hat soporta un número de sistemas de archivos de nodo único de alto rendimiento que se optimizan para nodo único y por lo tanto, suelen tener una carga inferior que un sistema de archivos de clúster. Red Hat recomienda el uso de estos sistemas en lugar de GFS2 en casos en que solamente un nodo único necesite montar el sistema de archivos.

Red Hat continuará el soporte de sistemas de archivos GFS2 de nodo único para montar instantáneas de sistemas de archivos de clúster (por ejemplo, para propósitos de respaldo).



### NOTA

Red Hat no soporta el uso de GFS2 para implementaciones de sistema de archivos de clúster mayores de 16 nodos.

Un GFS2 se basa en una arquitectura de 64 bits, la cual en teoría, aloja un sistema de archivos de 8 EB. Sin embargo, el tamaño máximo actual soportado de un sistema de archivos GFS2 es de 100 TB. El tamaño máximo soportado de un sistema de archivos para hardware de 32 bits es de 16 TB. Si su sistema requiere sistemas de archivos GFS2 más grandes, contacte a su representante de servicio técnico de Red Hat.

Al determinar el tamaño de su sistema de archivos, debe considerar sus necesidades de recuperación. La ejecución del comando **fsck.gfs2** en un sistema de archivos muy grande puede tardar mucho y consumir mucha memoria. Además, en el evento de que un disco o subsistema de disco falle, el tiempo de recuperación está limitado por la velocidad de los medios de respaldo. Para obtener información sobre la cantidad de memoria que el comando **fsck.gfs2** requiere, consulte la [Sección 4.11, “Cómo reparar un sistema de archivos”](#).

Cuando GFS2 de Red Hat está configurado en un clúster, los nodos de GFS2 de Red Hat pueden ser configurados y administrados con la configuración de adición de Alta disponibilidad y las herramientas de administración. GFS2 de Red Hat proporciona datos compartidos entre los nodos GFS2 de un clúster de Red Hat, con una única vista consistente del espacio de nombres del sistema de archivos entre todos los nodos de GFS2. Esto le permite a los procesos en nodos diferentes compartir archivos GFS2 del mismo modo que los procesos en el mismo nodo pueden compartir archivos en un sistema de archivos local, sin ninguna diferencia discernible. Para obtener más información sobre adición de Alta disponibilidad, consulte el documento *Configuración y administración de Red Hat clúster*.

Si bien un sistema de archivos de GFS2 puede utilizarse fuera de LVM, Red Hat admite únicamente sistemas de archivos GFS2 que se crean en un volumen lógico de CLVM. CLVM se incluye en la adición de almacenamiento resistente. Es una aplicación de todo el clúster de LVM, habilitado por el demonio CLVM **c1vmd**, el cual administra los volúmenes lógicos LVM en un clúster. El demonio hace

posible el uso de LVM2 para administrar volúmenes lógicos en un clúster, lo que permite que todos los nodos del clúster compartan los volúmenes lógicos. Para obtener información sobre el administrador de volúmenes LVM, consulte *Gestión del Administrador de volúmenes lógicos*.

El módulo de kernel **gfs2.ko** implementa el sistema de archivos GFS2 y se carga en los nodos de clúster GFS2.



## NOTA

Al configurar un sistema de archivos GFS2 como un sistema de archivos de clúster, debe asegurarse de que todos los nodos en el clúster tengan acceso al almacenamiento compartido. Las configuraciones de clúster asimétricas en las cuales algunos nodos tienen acceso al almacenamiento compartido y otros no, no están soportadas. En realidad, no se requiere que todos los nodos monten el sistema de archivos GFS2.

Este capítulo proporciona información básica que lo ayudará a entender GFS2. Contiene las siguientes secciones:

- [Sección 1.1, “Funcionalidades nuevas y cambiadas”](#)
- [Sección 1.2, “Antes de configurar GFS2”](#)
- [Sección 1.4, “Diferencias entre GFS y GFS2”](#)
- [Sección 1.3, “Instalación de GFS2”](#)
- [Sección 2.9, “Bloqueo de nodo GFS2”](#)

## 1.1. FUNCIONALIDADES NUEVAS Y CAMBIADAS

Esta sección lista las funcionalidades nuevas y cambiadas del sistema de archivos GFS2 y la documentación que se incluye con el lanzamiento inicial y subsiguientes de Red Hat Enterprise Linux 6.

### 1.1.1. Funcionalidades nuevas y cambiadas de Red Hat Enterprise Linux 6.0

Red Hat Enterprise Linux 6.0 incluye la siguiente documentación, actualizaciones y cambios.

- En el lanzamiento de Red Hat Enterprise Linux 6, Red Hat no soporta el uso de GFS2 como un sistema de archivos de un nodo único.
- En el lanzamiento de Red Hat Enterprise Linux 6, el comando **gfs2\_convert** para actualizar de un sistema de archivos GFS a GFS2 ha sido mejorado. Para obtener más información sobre este comando, consulte [Apéndice B, \*Cómo convertir el sistema de archivos de GFS a GFS2\*](#).
- El lanzamiento de Red Hat Enterprise Linux 6 soporta las opciones de montaje **discard**, **nodiscard**, **barrier**, **nobarrier**, **quota\_quantum**, **statfs\_quantum**, y **statfs\_percent**. Para obtener más información sobre montaje de un sistema de archivos GFS2, consulte la [Sección 4.2, “Cómo montar un sistema de archivos”](#).
- La versión de Red Hat Enterprise Linux 6 de este documento contiene una nueva sección, la [Sección 2.9, “Bloqueo de nodo GFS2”](#). Esta sección describe algunos de los aspectos internos del sistema de archivos GFS2.

### 1.1.2. Funcionalidades nuevas y cambiadas para Red Hat Enterprise Linux 6.1

Red Hat Enterprise Linux 6.1 incluye la siguiente documentación, actualizaciones y cambios.

- A partir del lanzamiento de Red Hat Enterprise Linux 6.1, GFS2 soporta los servicios estándar de cuotas de Linux. La administración de cuotas GFS2 se documenta en la [Sección 4.5, “Administración de cuotas en GFS2”](#).

En lanzamientos anteriores de Red Hat Enterprise Linux, GFS2 requería el comando **gfs2\_quota** para administrar cuotas. La documentación para **gfs2\_quota** ahora se provee en el [Apéndice A, Administración de cuotas GFS2 con el comando gfs2\\_quota](#).

- Ahora este documento contiene un nuevo capítulo, [Capítulo 5, Diagnóstico y corrección de problemas con sistemas de archivos GFS2](#).
- Se han hecho correcciones técnicas menores y aclaraciones a lo largo del documento.

### 1.1.3. Funcionalidades nuevas y cambiadas para Red Hat Enterprise Linux 6.2

Red Hat Enterprise Linux 6.2 incluye la siguiente documentación y presenta actualizaciones y cambios.

- A partir del lanzamiento de Red Hat Enterprise Linux 6.2, GFS2 soporta el comando **tunegfs2**, el cual reemplaza algunas de las funcionalidades del comando **gfs2\_tool**. Para obtener más información, consulte la página de manual **tunegfs2**.

Las secciones a continuación han sido actualizadas para proporcionar procedimientos administrativos que no requieren el uso del comando **gfs2\_tool**:

- La [Sección 4.5.4, “Sincronización de cuotas con el comando quotasync”](#) y la [Sección A.3, “Sincronización de cuotas con el comando gfs2\\_quota”](#) ahora describen cómo cambiar el parámetro **quota\_quantum** de su valor predeterminado de 60 segundos mediante la opción de montaje **quota\_quantum=**.
- La [Sección 4.10, “Cómo suspender la actividad en un sistema de archivos”](#) ahora describe cómo suspender la actividad de escritura para una sistema de archivos mediante el comando **dmsetup suspend**.
- Este documento incluye el nuevo apéndice, [Apéndice C, Puntos de trazado GFS2 y archivo glocks debugfs](#). Este apéndice describe la interfaz **debugfs** y los puntos de trazado de GFS2. Se dirige a usuarios avanzados que estén familiarizados con el interior del sistema de archivos y que deseen aprender más sobre el diseño de GFS2 y la depuración de problemas específicos de GFS2.

### 1.1.4. Funcionalidades nuevas y cambiadas para Red Hat Enterprise Linux 6.3

Para el lanzamiento de Red Hat Enterprise Linux 6.3, este documento contiene un nuevo capítulo, el [Capítulo 2, Configuración de GFS2 consideraciones operacionales](#). Este capítulo proporciona las recomendaciones para mantener un sistema de archivos GFS2.

Además, se han hecho pequeñas aclaraciones y correcciones a lo largo del documento.

### 1.1.5. Funcionalidades nuevas y cambiadas para Red Hat Enterprise Linux 6.4

Para el lanzamiento de Red Hat Enterprise Linux 6.4, el [Capítulo 2, Configuración de GFS2 consideraciones operacionales](#) se ha actualizado con pequeñas aclaraciones.

## 1.1.6. Funcionalidades nuevas y cambiadas para Red Hat Enterprise Linux 6.6

Para el lanzamiento de Red Hat Enterprise Linux 6.6, este documento contiene un nuevo capítulo, [Capítulo 6, Cómo configurar el sistema de archivos GFS2 en un clúster Pacemaker](#). Este capítulo proporciona un resumen de los pasos requeridos para configurar un clúster Pacemaker que incluya un sistema de archivos GFS2.

Además, se han hecho pequeñas aclaraciones y correcciones a lo largo del documento.

## 1.2. ANTES DE CONFIGURAR GFS2

Antes de instalar y configurar GFS2, anote las siguientes características claves de su archivo de sistemas GFS2:

### Nodos GFS2

Determine cuáles nodos en el clúster montarán el sistema de archivos GFS2.

### Número de sistemas de archivos

Determine cuántos sistemas de archivos GFS2 deben ser creados inicialmente (más sistemas de archivos pueden ser añadidos posteriormente).

### Nombre del sistema de archivos

Determine un nombre único para cada sistema de archivos. El nombre debe ser único para todos los sistemas de archivos `lock_dlm` en el clúster. Cada nombre de sistema de archivo se requiere en la forma de una variable de parámetro. Por ejemplo, este libro utiliza los nombres de archivos `mydata1` y `mydata2` en algunos procedimientos de ejemplo.

### Diarios

Determine el número de diarios para su sistema de archivos GFS2. Se requiere un diario por cada nodo que monte el sistema de archivos GFS2. GFS2 le permite añadir dinámicamente diarios en un punto posterior de que los servidores adicionales monten un sistema de archivos. Para obtener información sobre cómo añadir diarios al sistema de archivos GFS2, consulte la [Sección 4.7, “Cómo añadir diarios a un sistema de archivos”](#).

### Particiones y dispositivos de almacenaje

Determine los dispositivos de almacenaje y las particiones que serán usadas para crear volúmenes lógicos (a través de CLVM) en los sistemas de archivos.



### NOTA

Se pueden presentar problemas de rendimiento con GFS2 cuando muchas operaciones de creación o borrado se emiten desde más de un nodo en el mismo directorio al mismo tiempo. Si esto ocurre, debe localizar la creación y borrado de archivo por un nodo a directorios específicos de ese nodo tanto como sea posible.

Para otras recomendaciones sobre cómo crear, usar y mantener el sistema de archivos GFS2, consulte el [Capítulo 2, Configuración de GFS2 consideraciones operacionales](#).

## 1.3. INSTALACIÓN DE GFS2

Además de los paquetes requeridos para Red Hat High Availability Add-On, debe instalar el paquete **gfs2-utils** para GFS2 y el paquete **lvm2-cluster** para el Gestor de volumen lógico en clúster (CLVM). Los paquetes **lvm2-cluster** y **gfs2-utils** son parte del canal ResilientStorage, el cual debe estar activado antes de instalar los paquetes.

Utilice el siguiente comando **yum install** para instalar los paquetes de software de alta disponibilidad de Red Hat:

```
# yum install rgmanager lvm2-cluster gfs2-utils
```

Para información general sobre Red Hat High Availability Add-On y administración de clúster, consulte el manual *Administración de clúster*.

## 1.4. DIFERENCIAS ENTRE GFS Y GFS2

Esta sección lista los cambios y mejoras que GFS2 ofrece en comparación con GFS.

Para migrar de GFS a GFS2 debe convertir su sistema de archivos GFS con la herramienta **gfs2\_convert**. Para obtener información sobre la herramienta **gfs2\_convert**, consulte [Apéndice B, Cómo convertir el sistema de archivos de GFS a GFS2](#).

### 1.4.1. Nombres de los comandos de GFS2

En general, la funcionalidad de GFS2 es idéntica a la de GFS. Los nombres de los comandos del sistema de archivos, sin embargo, especifican GFS2 en lugar de GFS. La [Tabla 1.1, “Comandos de GFS y GFS2”](#) muestra el GFS equivalente, los comandos de GFS2 y la funcionalidad.

**Tabla 1.1. Comandos de GFS y GFS2**

Comandos de GFS	Comandos de GFS2	Descripción				
<b>mount</b>	<b>mount</b>	Monta un sistema de archivos. El sistema puede determinar si el sistema de archivos es GFS o GFS2. Para obtener información sobre las opciones de montaje de GFS2 vea la página de manual de <code>gfs2_mount(8)</code> .				
<b>umount</b>	<b>umount</b>	Desmonta un sistema de archivos.				
<table border="1"> <tr> <td><b>fsck</b></td> </tr> <tr> <td><b>gfs_fsck</b></td> </tr> </table>	<b>fsck</b>	<b>gfs_fsck</b>	<table border="1"> <tr> <td><b>fsck</b></td> </tr> <tr> <td><b>fsck.gfs2</b></td> </tr> </table>	<b>fsck</b>	<b>fsck.gfs2</b>	Revisa y repara un sistema de archivos sin montar.
<b>fsck</b>						
<b>gfs_fsck</b>						
<b>fsck</b>						
<b>fsck.gfs2</b>						
<b>gfs_grow</b>	<b>gfs2_grow</b>	Expande un sistema de archivos montado.				
<b>gfs_jadd</b>	<b>gfs2_jadd</b>	Añadir un diario a un sistema de archivos montado				

Comandos de GFS	Comandos de GFS2	Descripción
<b>gfs_mkfs</b> <b>mkfs -t gfs</b>	<b>mkfs.gfs2</b> <b>mkfs -t gfs2</b>	Crea un sistema de archivos en un dispositivo de almacenaje.
<b>gfs_quota</b>	<b>gfs2_quota</b>	Manejo de cuotas en un sistema de archivos montado. A partir del lanzamiento de Red Hat Enterprise Linux 6.1, GFS2 soporta los servicios estándar de cuotas de Linux. Para obtener más información sobre el manejo de cuotas en GFS2, consulte la <a href="#">Sección 4.5, “Administración de cuotas en GFS2”</a> .
<b>gfs_tool</b>	<b>tunegfs2</b> Parámetros de montaje <b>dmsetup suspend</b>	Configure, ajuste o reúna información sobre el sistema de archivos. El comando <b>tunegfs2</b> tiene soporte a partir del lanzamiento de Red Hat Enterprise Linux 6.2. También está el comando <b>gfs2_tool</b> .
<b>gfs_edit</b>	<b>gfs2_edit</b>	Muestra, imprime o edita las estructuras internas de un sistema de archivos. El comando <b>gfs2_edit</b> puede ser usado por los sistemas de archivos GFS y GFS2.
<b>gfs_tool setflag jdata/inherited_jdata</b>	<b>chattr +j</b> (preferido)	Activa el diario en un archivo o directorio.
<b>setfacl/getfacl</b>	<b>setfacl/getfacl</b>	Establece u obtiene lista de control de acceso de archivo para un archivo o directorio.
<b>setfattr/getfattr</b>	<b>setfattr/getfattr</b>	Establece u obtiene atributos extendidos de un archivo.

Para obtener una lista completa de las opciones soportadas por cada comando del sistema de archivos GFS2, vea las páginas de manual para estos comandos.

### 1.4.2. Diferencias adicionales entre GFS y GFS2

Esta sección resume las diferencias adicionales en administración de GFS y GFS2 que no están descritas en la [Sección 1.4.1, “Nombres de los comandos de GFS2”](#).

#### Nombre de rutas dependientes del contexto

Los sistemas de archivos GFS2 no proporcionan soporte para los nombres de rutas dependientes del contexto, los cuales le permiten crear enlaces simbólicos que apunten a directorios o archivos de destino variable. Para obtener esta funcionalidad en GFS2, puede usar la opción **bind** del comando **mount**. Para obtener más información sobre montajes de enlaces y nombres de ruta dependiente del contexto, consulte la [Sección 4.12, “Nombres de rutas dependientes del contexto y montajes enlazados”](#).

## Módulo gfs2.ko

El módulo de kernel que implementa el sistema de archivos GFS es **gfs.ko**. El de GFS2 es **gfs2.ko**.

## Cómo activar la activación de la cuota en GFS2

En sistemas de archivos GFS2, la imposición de la cuota está desactivada por defecto y debe ser activada explícitamente. Para obtener información sobre cómo activar o desactivar la aplicación de la cuota, consulte la [Sección 4.5, “Administración de cuotas en GFS2”](#).

## Diario de datos

El sistema de archivos GFS2 soporta el uso del comando **chattr** para establecer y limpiar la opción **j** en un archivo o directorio. Establecer la opción **+j** en un archivo permite diario de datos en ese archivo. Establecer el indicador **+j** en un directorio significa que "inherit jdata", lo que indica que todos los archivos y directorios creados posteriormente en ese directorio están en diario. El uso del comando **chattr** es la forma preferida de activar y desactivar el diario de datos en un archivo.

## Cómo añadir diarios dinámicamente

En los sistemas de archivos GFS2, los diarios son metadatos incorporados que existen fuera del sistema de archivos, lo que hace necesario extender el tamaño de volúmenes lógicos que contienen el sistema de archivos antes de añadir diarios. En sistemas de archivos GFS2, los diarios son archivos planos (aunque escondidos). Esto significa que para los sistemas de archivos GFS2, los diarios pueden añadirse de forma dinámica cuando servidores adicionales montan un sistema de archivos, siempre y cuando se tenga suficiente espacio en el sistema de archivos para diarios adicionales. Para obtener más información sobre cómo añadir diarios a los sistemas de archivos GFS2, consulte la [Sección 4.7, “Cómo añadir diarios a un sistema de archivos”](#).

## Parámetro atime\_quantum eliminado

El sistema de archivos GFS2 no soporta el parámetro modificador **atime\_quantum**, el cual puede ser utilizado por el sistema de archivos GFS para especificar la frecuencia de actualizaciones **atime**. En su lugar, GFS2 soporta la opciones de montaje **relatime** y **noatime**. La opción de montaje **relatime** se recomienda para lograr una conducta similar para establecer el parámetro **atime\_quantum** en GFS.

## La opción data= del comando mount

Cuando se montan sistemas de archivos GFS2 se puede especificar la opción **data=ordered** o **data=writeback** de **mount**. Cuando se establece **data=ordered**, los datos del usuario modificados por una transacción son volcados al disco antes de que la transacción sea enviada al disco. Esto evita que el usuario vea, después de una falla, bloques sin inicializar en un archivo. Cuando se establece **data=writeback**, los datos del usuario son escritos en disco en cualquier momento después de ser ensuciado. Esto no ofrece las misma garantías que el modo **ordered**, pero puede ser un poco más rápido en algunas cargas de trabajo. El modo predeterminado es **ordered**.

## El comando `gfs2_tool`

`gfs2_tool` soporta una serie de opciones para GFS2 diferentes a las que el comando `gfs_tool` soporta para GFS:

- El comando `gfs2_tool` soporta un parámetro `journals` que imprime información sobre los diarios configurados actualmente, incluyendo cuántos diarios contiene el sistema de archivos.
- El comando `gfs2_tool` no soporta la opción `counters`, la cual utiliza el comando `gfs_tool` para mostrar estadísticas de GFS.
- El comando `gfs2_tool` no soporta la opción `inherit_jdata`. Para marcar un directorio como "inherit jdata", puede establecer la opción `jdata` en el directorio o puede utilizar el comando `chattr` para establecer la opción `+j` en el directorio. El comando `chattr` es el preferido para activar o desactivar el diario de datos en un archivo.



### NOTA

A partir del lanzamiento de Red Hat Enterprise Linux 6.2, GFS2 soporta el comando `tunegfs2`, el cual reemplaza algunas de las funcionalidades del comando `gfs2_tool`. Para obtener más información, consulte la página de manual (8) `tunegfs2`. Las funciones `settune` y `gettune` del comando `gfs2_tool` han sido reemplazadas por la opción de línea de comandos `mount`, la cual permite establecerlas mediante el archivo `fstab` cuando se requiera.

## El comando `gfs2_edit`

El comando `gfs2_edit` soporta un conjunto de opciones para GFS2 diferente a las que el comando `gfs_edit` soporta para GFS. Para obtener información sobre las opciones específicas que cada versión del comando soporta, consulte `gfs2_edit` y las páginas de manual `gfs_edit`.

### 1.4.3. Mejoras de rendimiento en GFS2

Hay muchas funcionalidades del sistema de archivos GFS2 que no son diferentes en la interfaz de usuario en comparación con el sistema de archivos GFS, pero que mejoran el rendimiento del sistema de archivos.

El sistema de archivos GFS2 ofrece un rendimiento de sistema de archivos mejorado así:

- Mejor rendimiento para uso pesado en un directorio único.
- Operaciones de sincronización de E/S más rápidas
- Lecturas de cache más rápidas (sin gasto de bloqueo)
- Aumento en la velocidad de E/S directa con archivos asignados con anterioridad (si el tamaño de E/S es razonablemente grande, tal como bloques de 4M)
- Más velocidad en las operaciones de E/S en general
- Una ejecución más ágil del comando `df`, debido a llamadas más rápidas de `statfs`.
- El modo `atime` para reducir el número de operaciones de escritura E/S generadas por `atime` cuando se compara con GFS.

Los sistemas de archivos GFS2 proporcionan un soporte más amplio y más convencional así:

- GFS2 es parte del kernel de la línea de desarrollo principal (integrado en 2.6.19).
- GFS2 soporta las siguientes funcionalidades.
  - atributos de archivo extendidos (**xattr**)
  - Parámetros de atributos **lsattr()** y **chattr()** a través de las llamadas **ioctl()** estándar
  - marcas de tiempo en nano segundos

Un sistema de archivos GFS2 proporciona las siguientes mejoras en la eficacia interna del sistema de archivos.

- GFS2 utiliza menos memoria del kernel
- GFS2 no requiere números de generación de metadatos.

La asignación de los metadatos de GFS2 no requiere lecturas. Las copias de bloques de metadatos en varios diarios se administran mediante la remoción de bloques de un diario antes de liberar el cerrojo.

- GFS2 incluye un administrador de registro más simple que no tiene conocimiento sobre los cambios de cuotas o inodos sin enlazar.
- Los comandos **gfs2\_grow** y **gfs2\_jadd** utilizan bloqueo para evitar la ejecución simultánea de varias instancias.
- El código ACL ha sido simplificado para llamadas como **creat()** y **mkdir()**.
- Inodos sin enlace, cambios de cuota y cambios de **statfs** se recobran sin remontar el diario.

## CAPÍTULO 2. CONFIGURACIÓN DE GFS2 CONSIDERACIONES OPERACIONALES

El Sistema de archivos global 2 (GFS2) permite a varios computadores (“nodos”) en un clúster compartir de forma cooperativa el mismo almacenaje. Para realizar esta cooperación y mantener la consistencia de datos entre los nodos, los nodos emplean un esquema de bloqueo de todo el clúster para recursos de sistemas de archivos. Este esquema de bloqueo usa protocolos de comunicación tales como TCP e IP para intercambiar información de bloqueo.

Para mejorar el rendimiento, siga las recomendaciones descritas en este capítulo, incluidas las recomendaciones para crear, usar y mantener un sistema de archivos GFS2.



### IMPORTANTE

Asegúrese de que la implementación de la adición de alta disponibilidad de Red Hat satisfaga sus necesidades y pueda recibir soporte. Consulte a un representante autorizado de Red Hat para verificar su configuración antes de implementarla.

## 2.1. CONSIDERACIONES DE FORMATEO

Esta sección proporciona recomendaciones sobre cómo formatear su sistema de archivos GFS2 para optimizar el rendimiento.

### 2.1.1. Tamaño del sistema de archivos: Entre más pequeño, mejor

GFS2 está basado en una arquitectura de 64 bits, la cual en teoría, acomoda un sistema de archivos de 8 EB. Sin embargo, el tamaño máximo de un sistema de archivos GFS2 para un hardware de 64 bits es de 100 TB y el tamaño máximo soportado actual de un sistema de archivos GFS2 para un hardware de 32 bits es 16 TB.

Observe que a pesar de que los sistemas de archivos GFS2 sean posibles, no significa que sean recomendables. La regla con GFS2 es que entre más pequeño mejor: es mejor tener 10 sistemas de archivos de 1TB en lugar de un sistema de archivos de 10 TB.

Hay varias razones por las cuales se deben mantener sistemas de archivos GFS2 pequeños:

- Se requiere menos tiempo para hacer la copia de seguridad de cada sistema de archivo.
- Se requiere menos tiempo para chequear el sistema de archivos con el comando `fsck.gfs2`.
- Se requiere menos memoria para chequear el sistema de archivos con el comando `fsck.gfs2`.

Además, menos grupos de recursos para mantener significa mejor rendimiento.

Claro está que si desea crear un sistema de archivos GFS2 demasiado pequeño, podría quedarse sin espacio, y esto tiene sus propias consecuencias. Debe considerar sus razones para decidir el tamaño.

### 2.1.2. Tamaño de bloque: El predeterminado es el preferido (4K)

A partir del lanzamiento de Red Hat Enterprise Linux 6, el comando `mkfs.gfs2` intenta estimar un tamaño de bloque óptimo basado en la topología del dispositivo. En general, los bloques de 4K se prefieren porque 4K es el tamaño de página (memoria) predeterminada para Linux. A diferencia de otros

sistemas de archivos, GFS2 realiza la mayoría de sus operaciones mediante búferes de kernel de 4k. Si su tamaño de bloque es de 4K, el kernel tiene menos trabajo para manipular los búferes.

Se recomienda usar el tamaño de bloque predeterminado, el cual produce el máximo rendimiento. Solo necesita usar un tamaño de bloque diferente si requiere almacenamiento eficiente de varios archivos muy pequeños.

### 2.1.3. Número de diarios: uno por cada nodo que monte

GFS2 requiere un diario para cada nodo en el clúster que necesite montar el sistema de archivos. Por ejemplo, si tiene un clúster de 16 nodos, pero solo necesita montar el sistema de archivos desde dos nodos, necesitará únicamente dos diarios. Si necesita montar un tercer nodo, siempre podrá añadir un diario con el comando `gfs2_jadd`. Con GFS2, usted puede añadir diarios sobre la marcha.

### 2.1.4. Tamaño de diario: El predeterminado suele ser óptimo (128MB)

Cuando ejecute el comando `mkfs.gfs2` para crear un sistema de archivos GFS2, especifique el tamaño de los diarios. Si no lo desea, se predeterminará a 128MB, el cual deberá ser óptimo para la mayoría de las aplicaciones.

Algunos administradores de sistemas podrían pensar que 128MB es excesivo y verse tentados a reducir el tamaño del diario a un mínimo de 8MB o al conservador de 32MB. Aunque esto podría funcionar, puede afectar gravemente el rendimiento. Al igual que muchos sistemas de diarios, cada vez que GFS2 escribe metadatos, los metadatos se envían al diario antes de ponerlo en su sitio. Esto garantiza que si el sistema se daña o pierde energía, usted recuperará todos los metadatos cuando el diario se reproduzca en tiempo de montaje. Sin embargo, no se necesita mucha actividad de sistema de archivos para llenar un diario de 8MB y, cuando el diario esté lleno, el rendimiento disminuirá porque GFS2 tiene que esperar a que se escriba al almacenamiento.

Por lo general, se recomienda usar el tamaño predeterminado de 128 MB. Si su sistema de archivos es muy pequeño (por ejemplo de 5 GB), un diario de 128 MB no será práctico. Si tiene un sistema de archivos mayor y si puede disponer de más espacio, el uso de diarios de 256 MB podría mejorar el rendimiento.

### 2.1.5. Tamaño y número de grupos de recursos

Cuando un sistema de archivos GFS2 es creado con el comando `mkfs.gfs2`, divide el almacenamiento en dos pedazos uniformes conocidos como grupos de recursos. Intenta estimar un tamaño de grupo de recursos óptimo (que va de 32 MB a 2 GB). Puede sobrescribir el predeterminado con la opción `-r` del comando `mkfs.gfs2`.

El tamaño óptimo del grupo de recursos depende de cómo va a utilizar el sistema de archivos. Considere qué tan lleno estará o si será severamente fragmentado.

Experimente con diferentes tamaños de grupos de recursos para ver qué resulta en rendimiento óptimo. Se recomienda experimentar con un clúster de prueba antes de implementarlo en plena producción.

Si su sistema de archivos tiene demasiados grupos de recursos (cada uno de los cuales es demasiado pequeño), las asignaciones de bloques perderán demasiado tiempo en la búsqueda de decenas de miles (o cientos de miles) de grupos de recursos de un bloque libre. Entre más lleno esté su sistema de recursos, más recursos de grupos serán buscados y cada uno de ellos requerirá un cerrojo en todo el clúster. Esto conducirá a un rendimiento bajo.

No obstante, si su sistema de archivos tiene muy pocos grupos de recursos (cada uno de los cuales es demasiado grande), la asignación de bloques podría pelear más a menudo por el mismo cerrojo de

grupos de recursos, lo cual también impactará el rendimiento. Por ejemplo, si tiene un archivo de 10 GB que está repartido entre cinco grupos de recursos de 2 GB, los nodos en su clúster se luchará más a menudo por esos cinco grupos de recursos que si el mismo sistema de archivos estuviera repartido entre 320 grupos de recursos de 32 MB. El problema se exagera si su sistema de archivos está casi lleno porque cada asignación de bloques podría tener que mirar a través de varios grupos de recursos antes de que encuentre uno con un bloque libre. GFS2 intenta mitigar este problema en dos formas:

- Primero, cuando un grupo de recursos está completamente lleno, este recuerda y trata de evitar chequearlo en futuras asignaciones (antes de que un bloque se libere de él). Si usted nunca borra archivos, la competencia será menos grave. No obstante, si su aplicación está constantemente borrando bloques y asignando nuevos bloques en un sistema de archivos que está en su mayoría lleno, la competencia será muy grande y esto tendrá un grave impacto de rendimiento.
- Segundo, cuando se añaden nuevos bloques a un archivo existente, GFS2 intentará agrupar los nuevos bloques en el mismo grupo de recursos que el archivo. Esto se realiza para incrementar el rendimiento: en un disco giratorio, las búsquedas toman menos tiempo cuando están físicamente juntos.

El peor escenario es cuando hay un directorio central en el cual todos los nodos crean archivos, ya que todos los nodos lucharán constantemente por bloquear el mismo grupo de recursos.

## 2.2. FRAGMENTACIÓN DE SISTEMA DE ARCHIVOS

Red Hat Enterprise Linux 6.4 introduce mejoras a la administración de fragmentación de archivos en GFS2. Con Red Hat Enterprise Linux 6.4, la escritura simultánea resulta en menor fragmentación de archivos y por lo tanto, en mejor rendimiento para estas cargas.

Aunque no hay una herramienta de desfragmentación para GFS2 en Red Hat Enterprise Linux, usted puede desfragmentar archivos individuales al identificarlos con la herramienta `filefrag`, copiándolos en los archivos temporales, y renombrándolos para reemplazar los originales. (Este procedimiento puede también hacerse en versiones anteriores a Red Hat Enterprise Linux 6.4 siempre y cuando la escritura se haga en forma secuencial).

## 2.3. PROBLEMAS DE ASIGNACIÓN DE BLOQUES

Esta sección provee un resumen de los aspectos relacionados con la asignación de bloques en sistemas de archivos GFS2. Aunque las aplicaciones solo escriben datos, por lo general no se preocupan en cómo o dónde se asigna el bloque, un poco de conocimiento acerca de cómo funciona la asignación de bloques puede ayudar a optimizar el rendimiento.

### 2.3.1. Abandonar espacio libre en el sistema de archivos

Cuando un sistema de archivos GFS2 está casi lleno, el asignador de bloques comienza a tener dificultades para encontrar espacio para los nuevos bloques que van a ser asignados. Como resultado, los bloques entregados por el asignador tienden a comprimirse hasta el final del grupo de recursos o en partes diminutas donde la fragmentación de archivos es mucho más probable. Esta fragmentación puede causar problemas de rendimiento. Además, cuando un GFS2 está casi lleno, el asignador de bloques de GFS2 gasta más tiempo buscando a través de múltiples grupos de recursos y que añade competencia de cerrojo que no estaría necesariamente en un sistema de archivos, que tenga bastante espacio libre. Esto también puede causar problemas de rendimiento.

Por estas razones, se recomienda ejecutar un sistema de archivos que esté más del 85 por ciento lleno, aunque esta cantidad puede variar según la carga de trabajo.

### 2.3.2. Procure que cada nodo asigne sus propios archivos

Debido a la forma como funciona el Gestor de bloqueo distribuido (DLM), habrá más competencia de cerrojo si todos los archivos son asignados por un nodo y otros nodos necesitan añadir bloques a dichos archivos.

En GFS (versión 1), el Gestor de cerrojo central administraba todos los cerrojos cuyo trabajo era controlar el bloqueo en el clúster. Este Gran gestor de cerrojo unificado (GULM) era problemático porque era un solo punto de falla. El esquema de cerrojo de remplazo de GFS2, DLM, extiende los cerrojos a través del clúster. Si cualquier nodo en el clúster se cae, sus cerrojos serán recuperados por otros nodos.

Con DLM, el primer nodo para bloquear un recurso (tal como un archivo) se convierte en el “maestro de cerrojo” para ese bloque. Otros nodos pueden bloquear ese recurso, pero primero deben pedir permiso al maestro de cerrojo. Cada nodo sabe qué cerrojos le corresponden al maestro de cerrojo y cada nodo sabe a qué nodo le ha prestado el cerrojo. Bloquear un cerrojo en el nodo maestro es mucho más rápido que bloquear uno en otro nodo que tenga que detenerse y pedir permiso al maestro del cerrojo.

Como en muchos sistemas de archivos, el asignador GFS2 intenta mantener bloques en el mismo archivo cerca el uno del otro para reducir el movimiento de las cabezas de disco y aumentar el rendimiento. Un nodo que asigne los bloques a un archivo probablemente necesitará usar y bloquear los mismos grupos para los nuevos bloques (a menos que todos los bloques en ese grupo de recursos estén en uso). El sistema de archivos se ejecuta más rápido si el cerrojo maestro para el grupo de recursos que contiene el archivo, asigna sus propios bloques de datos (es decir, es más rápido hacer que el nodo que primero abrió el archivo ejecute toda la escritura de los nuevos bloques).

### 2.3.3. Preasignar, si es posible

Si se preasignan los archivos, se evitarán las asignaciones de bloques y el sistema de archivos se ejecutará de forma más eficiente. Las versiones más recientes de GFS2 incluyen la llamada del sistema `fallocate(1)`, la cual puede usarse para preasignar bloques de datos.

## 2.4. CONSIDERACIONES DE CLÚSTER

Cuando determine el número de nodos que su sistema puede contener, note que hay un intercambio entre alta disponibilidad y rendimiento. Con un número mayor de nodos, se dificulta cada vez más hacer escala de cargas de trabajo. Por esta razón, Red Hat no acepta el uso de GFS2 para implementaciones de sistemas de archivos mayores de 16 nodos.

La implementación de un sistema de archivos no es un remplazo de “caída” para la implementación de un nodo único. Se recomienda dejar un periodo de 8 a 12 semanas en instalaciones nuevas para probar el sistema y asegurarse de que está funcionando en el nivel de rendimiento requerido. Durante este periodo cualquier problema de rendimiento o funcionalidad puede solucionarse y las consultas deben dirigirse al equipo de soporte técnico de Red Hat.

Recomendamos que los usuarios que consideren implementar clústeres hagan revisar su configuración por el equipo de soporte de Red Hat antes para evitar cualquier problema de soporte más adelante.

## 2.5. CONSIDERACIONES DE USO

Esta sección proporciona recomendaciones generales sobre el uso de GFS2.

### 2.5.1. Opciones de montaje: `noatime` y `nodiratime`

Por lo general se recomienda montar los archivos de sistemas GFS2 con los argumentos **noatime** y **nodiratime**. Esto permite a GFS2 emplear menos tiempo en actualizar nodos de disco para cada acceso.

### 2.5.2. Opciones de ajuste DLM: Aumento de tamaño de tablas DLM

DLM usa varias tablas para administrar, coordinar, y pasar información de bloques entre nodos en el clúster. El aumento del tamaño de las tablas DLM podría incrementar el rendimiento. En Red Hat Enterprise Linux 6.1 y posterior, se han aumentado los tamaños predeterminados de dichas tablas, aunque usted puede aumentarlas con los siguientes comandos:

```
echo 1024 > /sys/kernel/config/dlm/cluster/lkbtbl_size
echo 1024 > /sys/kernel/config/dlm/cluster/rsbtbl_size
echo 1024 > /sys/kernel/config/dlm/cluster/dirtbl_size
```

Estos comandos no son persistentes y no sobrevivirán al reinicio, por lo tanto, debe añadirlos a uno de los scripts de inicio y debe ejecutarlos antes de montar los a algún sistema de archivos GFS2, o de lo contrario, los cambios se ingorarán.

Para obtener más información sobre bloqueo de nodos For GFS2, consulte la [Sección 2.9, “Bloqueo de nodo GFS2”](#).

### 2.5.3. Opciones de ajuste VFS: Investigación y experimentación

Como todos los sistemas de archivos de Linux, GFS2 se sitúa en la parte superior de la capa llamada el sistema de archivos virtual (VFS). Ajuste la capa VFS para mejorar el rendimiento subyacente de GFS2 mediante el comando **sysctl**(8). Por ejemplo, se pueden ajustar los valores para **dirty\_background\_ratio** y **vfs\_cache\_pressure** según su situación. Para obtener los valores actuales, use los siguientes comandos:

```
sysctl -n vm.dirty_background_ratio
sysctl -n vm.vfs_cache_pressure
```

Los siguientes comandos ajustan los valores:

```
sysctl -w vm.dirty_background_ratio=20
sysctl -w vm.vfs_cache_pressure=500
```

Para cambiar permanentemente los valores de estos parámetros edite el archivo **/etc/sysctl.conf**.

Para hallar los valores óptimos para sus casos de uso, busque las opciones VFS y experimente en un clúster de prueba antes de implementar en toda la producción.

### 2.5.4. SELinux: Evitar SELinux en GFS2

La seguridad mejorada de Linux (SELinux) es altamente recomendable por razones de seguridad en la mayoría de situaciones, pero no tiene soporte si se utiliza con GFS2. SELinux guarda información mediante atributos extendidos acerca de cada objeto de sistema de archivos. Es posible leer, escribir o mantener estos atributos, pero reduce considerablemente la velocidad del GFS2. Debe apagar SELinux en sistemas de archivos GFS2.

### 2.5.5. Configuración de NFS en GFS2

Debido a la complejidad añadida del subsistema de bloqueo de GFS2 y a su naturaleza de agrupamiento, la configuración de NFS en GFS2 requiere muchas precauciones y una configuración cuidadosa. Esta sección describe los riesgos que puede correr en una cuenta al configurar un servicio NFS en un sistema de archivos GFS2.



### AVISO

Si el sistema de archivos GFS2 es exportado con NFS, y las aplicaciones de cliente NFS usan cerrojos POSIX, entonces deberá montar el sistema de archivos con la opción **locallocks**. El efecto que se pretende es forzar a los cerrojos de POSIX de cada servidor para que sean locales: es decir no agrupados, independientes uno del otro. (Se presentan una serie de problemas si GFS2 intenta implementar cerrojos de POSIX desde NFS a través de los nodos de un clúster). Para las aplicaciones que se ejecutan en clientes NFS, los cerrojos POSIX localizados significan que dos clientes pueden tener simultáneamente el mismo cerrojo si dos clientes montan desde servidores diferentes. Si todos los clientes montan NFS desde un servidor, entonces desaparecerá el problema de que servidores independientes otorguen los mismos cerrojos por separado. Si no está seguro de montar su sistema de archivos con la opción **locallocks**, no debería usarla; siempre es más seguro si los cerrojos operan en una base agrupada.

Además de las consideraciones de cerrojo, debería tener en cuenta lo siguiente al configurar un servicio NFS en un sistema de archivos GFS2.

- Red Hat soporta únicamente las configuraciones de alta disponibilidad adicionales mediante NFSv3 con cerrojo en configuración activa/pasiva con las siguientes características:
  - El Sistema de archivos de fondo es el sistema de archivos GFS2 que se ejecuta en un cluster de 2 a 16 nodos.
  - Un servidor NFSv3 se define como un servicio que exporta simultáneamente todo el sistema de archivos GFS2 desde un nodo de cluster único.
  - El servidor NFS puede conmutarse de un nodo de cluster a otro (configuración activa y pasiva).
  - No se permite el acceso al sistema de archivos GFS2 *excepto* a través del servidor NFS. Esto incluye tanto el acceso del sistema de archivos local como el acceso a través de Samba o Samba en cluster.
  - No hay soporte de cuota NFS en el sistema.

Esta configuración proporciona alta disponibilidad para el sistema de archivos y reduce el tiempo de inactividad porque un nodo fallido no resulta en el requerimiento para ejecutar el comando **fsck** cuando falla el servidor NFS de un nodo a otro.

- La opción de NFS **fsid=** es obligatoria para exportaciones NFS de GFS2.
- Si se presentan problemas con su clúster (por ejemplo, el clúster no tiene quórum y el cercado no es exitoso), los volúmenes lógicos en clúster y el sistema de archivos GFS2 se congelarán y no se podrá acceder hasta que el clúster tenga quórum. Considere esta posibilidad al

determinar si la solución de conmutación única tal como la definida en este procedimiento es la más apropiada para su sistema.

### 2.5.6. Archivo Samba (SMB o Windows) sirviendo en GFS2

A partir del lanzamiento de Red Hat Enterprise Linux 6.2, puede usar el archivo Samba (SMB o Windows) sirviendo desde un sistema de archivos GFS2 con CTDB, el cual permite configuraciones activa/activa. Para obtener información sobre configuración en clúster de Samba, consulte el documento de *Administración de clúster*.

El acceso simultáneo a datos en el recurso compartido de Samba desde fuera de Samba no recibe soporte. Actualmente, no se ofrece soporte para arriendos de clústeres GFS2, lo cual reduce la velocidad del servicio de archivos de Samba.

## 2.6. COPIAS DE SEGURIDAD DE SISTEMA DE ARCHIVOS

Es importante hacer con regularidad copias de seguridad de su sistema de archivos GFS2 en caso de emergencia, independiente del tamaño de su sistema de archivos. Muchos administradores de sistemas se sienten seguros porque están protegidos por RAID, multirrutas, espejos, instantáneas, y otras formas de redundancia, pero no hay nada lo suficientemente seguro.

Puede ser un problema crear una copia de seguridad ya que el proceso de copia de un nodo o de una serie de nodos suele involucrar la lectura de todo el sistema de archivos en secuencia. Si esto se realiza desde un nodo único, ese nodo retendrá toda la información en la memoria cache hasta que otros nodos en el clúster empiecen a solicitar cerrojos. Si ejecuta este tipo de programa de copia de seguridad mientras el clúster se está ejecutando impactará negativamente el rendimiento.

Al abandonar la memoria cache una vez se termine el copiado, se reducirá el tiempo que los nodos requieren para recobrar la propiedad de sus cerrojos de clústeres o de cache. No obstante, esto no es lo ideal, puesto que otros nodos habrán detenido los datos que estaban guardando en cache antes de que el proceso de copiado de seguridad comenzara. Puede abandonar las cache mediante el siguiente comando después de que el copiado de seguridad se complete:

```
echo -n 3 > /proc/sys/vm/drop_caches
```

Es más rápido si cada nodo en el clúster respalda sus propios archivos para que la tarea se divida entre los nodos. Podría hacerlo con un script que usa el comando **rsync** en directorios de nodos específicos.

La mejor forma de hacer una copia de seguridad de GFS2 es mediante la creación de una instantánea de hardware en la SAN, presente la instantánea a otro sistema y cópiela allí. El sistema de respaldo deberá montar la instantánea con **-o lockproto=lock\_nolock** ya que no estará en un clúster.

## 2.7. CONSIDERACIONES DE HARDWARE

Tenga en cuenta las siguientes consideraciones al implementar un sistema de archivos GFS2.

- Use opciones de almacenaje de alta calidad

GFS2 puede funcionar en opciones de almacenaje más baratas, tales como iSCSI o Canal de fibra por Ethernet (FCoE), pero obtendrá mejor rendimiento si compra almacenaje de alta calidad con mayor capacidad de cache. Red Hat realiza las pruebas de máxima calidad, sanidad y rendimiento en almacenamiento SAN con interconexión de Canal de fibra. Como regla general, siempre es mejor implementar algo que ya ha sido probado.

- Pruebe el equipo de red antes de la implementación

Un equipo de mayor calidad y agilidad hará que las comunicaciones de clúster y GFS2 se ejecuten con mayor rapidez y confiabilidad. No siempre el hardware más costoso es el mejor. Algunos de los interruptores de red más costosos tienen problemas al pasar los paquetes multidifusión, los cuales sirven para pasar los cerrojos (flocks) `fcntl`, mientras que interruptores de red más baratos son algunas veces más rápidos y confiables. Es una buena práctica ensayar el equipo antes de implementarlo en una producción total.

## 2.8. PROBLEMAS DE RENDIMIENTO: REVISE EL PORTAL DEL CLIENTE DE RED HAT

Para obtener información sobre las mejores prácticas para implementar y actualizar los clústeres de Red Hat Enterprise Linux mediante la adición de alta disponibilidad y el Sistema de archivos globales 2 de Red Hat (GFS2), consulte el artículo "Red Hat Enterprise Linux Cluster, High Availability, y GFS Deployment Best Practices" en Red Hat Customer Portal, <https://access.redhat.com/site/articles/40051>.

## 2.9. BLOQUEO DE NODO GFS2

Para obtener el mejor rendimiento de un sistema de archivos GFS2, es muy importante entender alguna teoría básica de su operación. Un sistema de archivos de nodo único se implementa junto a la memoria cache, con el propósito de eliminar latencia de acceso de disco cuando se utilicen datos solicitados con frecuencia. En la página cache de Linux (e históricamente la cache de buffer) se proporciona esta función de cache.

Con GFS2, cada nodo tiene su propia página cache que contiene alguna parte de datos en disco. GFS2 utiliza un mecanismo de bloqueo llamado *glocks* para mantener la integridad de la cache entre nodos. El subsistema de glock ofrece una función de manejo de cache que se implementa con el *administrador de cerrojo distribuido* (DLM) como la capa de comunicación subyacente.

Los glock proporcionan protección para la cache en una base por inodo, por lo tanto hay un cerrojo por inodo, el cual sirve para controlar la capa de memoria cache. Si ese glock se otorga en modo compartido (modo de cerrojo DLM: PR) entonces los datos bajo ese glock pueden ser puestos en memoria cache tras uno o más nodos al mismo tiempo, para que todos los nodos puedan tener acceso local a datos.

Si se le otorga modo exclusivo al glock (modo de cerrojo DLM: EX) entonces solamente un nodo único puede poner en memoria cache los datos bajo ese glock. Este modo es utilizado por todas las operaciones que modifican los datos (tal como la llamada a sistema `write`).

Si otro nodo solicita un glock que no puede ser otorgado inmediatamente, entonces el DLM envía un mensaje al nodo o nodos que actualmente guardan los glock bloqueando la nueva solicitud para pedirles que coloquen los cerrojos. Colocar los glock (por las normas de la mayoría de operaciones del sistema de archivos) puede ser un proceso largo. Colocar un glock compartido requiere que la cache sea invalidada, lo cual es relativamente rápido y proporcional a la cantidad de datos en memoria cache.

Colocar un glock exclusivo requiere un vaciado de registro, y volver a escribir al disco los datos cambiados, seguido de la invalidación como el glock compartido.

La diferencia entre un sistema de archivos de nodo único y GFS2, es que el sistema de archivos de nodo único tiene una cache única y GFS2 tiene una cache independiente para cada nodo. En ambos casos, la latencia de acceso a los datos en memoria cache es de un orden de magnitud similar, pero la latencia para acceder datos que no están en memoria cache es mucho mayor en GFS2 si otro nodo ha puesto anteriormente en memoria cache esos mismos datos.



## NOTA

Debido a la forma en la que la puesta en memoria cache de GFS2 se implementa, el mejor rendimiento se obtiene cuando algo de esto tiene lugar:

- Un inodo se utiliza en modo de solo lectura a través de todos los nodos.
- Un inodo se escribe o modifica desde un solo nodo únicamente.

Observe que insertar o retirar entradas de un directorio durante la creación y borrado de un archivo, cuenta como escribir en el inodo del directorio.

Es posible romper esta regla si pocas veces se rompe. Si ignora esta regla con mucha frecuencia, resultará en una pena severa en el rendimiento.

Si ejecuta `mmap()` de un archivo en GFS2 con una asignación de lectura/escritura, pero solamente lee desde él, solo cuenta como de lectura. En GFS, sin embargo, este archivo cuenta como de escritura, por lo tanto GFS2 es mucho más escalable con E/S de `mmap()`.

Si no ha establecido el parámetro `noatime mount`, entonces las lecturas también resultarán en escrituras para actualizar las marcas de tiempo del archivo. Se recomienda que todos los usuarios de GFS2 monten con `noatime` a menos que tengan un requerimiento específico para `atime`.

### 2.9.1. Problemas con Bloqueo de Posix

Al utilizar el bloqueo de Posix, debe tener en cuenta lo siguiente:

- El uso de Flocks generará un procesamiento más rápido que el uso de cerrojos Posix.
- Los programas que usan cerrojos Posix en GFS2 deben evitar el uso de la función `GETLK` ya que en un entorno de clúster, el ID de proceso puede ser para un nodo diferente en el clúster.

### 2.9.2. Ajuste de rendimiento con GFS2

Usualmente es posible alterar la forma en que una aplicación problemática almacena sus datos para obtener una ventaja considerable en el rendimiento.

Un ejemplo típico de una aplicación problemática es un servidor de correo-e. A menudo se establece con un directorio de cola que contiene los archivos para cada usuario (`mbx`), o con un directorio para cada usuario que contenga un archivo para cada mensaje (`maildir`). Cuando llegan solicitudes sobre IMAP, el mecanismo ideal es dar a cada usuario una afinidad a un nodo particular. De este modo sus solicitudes para ver y eliminar mensajes de correo electrónico tenderán a ser servidas desde otra memoria cache en ese nodo único. Obviamente, si ese nodo falla, la sesión se reiniciará en un nodo diferente.

Cuando un correo llega a través de SMTP, los nodos individuales pueden configurarse como para pasar algún correo de usuario a un nodo particular predeterminado. Si el nodo predeterminado no está activado, entonces el mensaje puede ser guardado por el nodo que lo recibe. De nuevo, este diseño es para conjuntos de archivos particulares en memoria cache en solo un nodo en el caso normal, pero permite acceso directo en caso de fallar un nodo.

Esta configuración usa la memoria cache de página de GFS2 y también hace transparente las fallas a la aplicación ya sea `imap` o `smtp`.

Las copias de seguridad son otra área engañosa. Es posible y preferible hacer copias del conjunto de

trabajo de cada nodo de forma directa desde el nodo en el que se está guardando en cache ese conjunto particular de inodos. Si tiene un script de respaldo que se ejecuta en un momento determinado y que parece coincidir con un punto máximo en tiempo de respuesta de una aplicación ejecutándose en GFS2, entonces es muy probable que el cluster no esté haciendo el uso más eficiente de la cache de página.

Obviamente, si está en la posición envidiable de poder detener la aplicación para realizar una copia de seguridad, entonces no habrá problema. Por otra parte, si la copia de seguridad se ejecuta solamente desde un nodo, entonces después de haber completado una gran porción del sistema de archivos será puesta en memoria en ese nodo, con una pena de rendimiento para accesos posteriores desde otros nodos. Esto se puede mitigar hasta cierto punto, colocando la cache de página VFS en el nodo de respaldo después de que la copia de seguridad haya completado con el siguiente comando:

```
echo -n 3 >/proc/sys/vm/drop_caches
```

Sin embargo, esa no es una buena solución, es mejor asegurarse de que el conjunto de trabajo en cada nodo sea compartido, principalmente leído solo a través del cluster o accedido ampliamente desde un nodo único.

### 2.9.3. Resolución de problemas de rendimiento de GFS2 con el volcado de cerrojo GFS2

Si su rendimiento de cluster está sufriendo debido a un uso ineficiente de guardado de memoria en cache de GFS2, verá que los tiempos de espera aumentan y se prolongan.

Esta sección ofrece una visión general del vaciado de bloques GFS2. Para obtener una descripción completa del vaciado de bloque GFS2, consulte el [Apéndice C, Puntos de trazado GFS2 y archivo glocks debugfs](#).

La información del volcado de cerrojo GFS2 puede obtenerse del archivo **debugfs**, el cual se encuentra en el siguiente nombre de ruta, suponiendo que **debugfs** esta montado en **/sys/kernel/debug/**:

```
/sys/kernel/debug/gfs2/fsname/glocks
```

El contenido del archivo es una serie de líneas. Cada línea que inicia con G: representa un glock, y las siguientes líneas, con un solo espacio, representan un elemento de información que relaciona el glock inmediatamente delante de ellas en el archivo.

La mejor manera de usar el archivo **debugfs** es usar el comando **cat** para tomar una copia del contenido completo del archivo (podría tomarse bastante tiempo si tiene una gran cantidad de RAM y un montón de inodos en memoria) mientras la aplicación está teniendo problemas y luego mirar los datos resultantes en una fecha posterior.



#### NOTA

Puede ser útil hacer dos copias del archivo **debugfs**, una, unos segundos o incluso un minuto o dos después de la otra. Al comparar la información del portador en los dos trazados que se relacionan con el mismo número de glock, se sabe si la carga de trabajo está haciendo algún progreso o si se ha trancado (lo cual es siempre un error y debe reportarse inmediatamente al soporte técnico de Red Hat).

Las líneas en el archivo **debugfs** que inician con H de 'holders' (portadores), representan solicitudes de cerrojo otorgadas o en espera de ser otorgadas. Los campos de indicadores en la línea de portadores f:

muestra Which (cuál) : El indicador 'W' se refiere a la solicitud en espera, el indicador 'H' se refiere a la solicitud otorgada. Es probable que los glocks que tienen una gran cantidad de solicitudes en espera sean los que estén experimentando contención particular.

Tabla 2.1, “Indicadores Glock” shows the meanings of the different glock flags and Tabla 2.2, “Indicadores de portador de Glock” shows the meanings of the different glock holder flags in the order that they appear in the glock dumps.

**Tabla 2.1. Indicadores Glock**

Indicador	Nombre	Significado
por	Bloqueo	Es válido cuando el indicador se establece a bloqueado e indica si la operación que ha sido solicitada desde el DLM puede bloquear. Este indicador es retirado para degradación y para los cerrojos 'try'. El propósito de este indicador es permitir reunir estadísticas de tiempo de respuesta independiente del tiempo tomado por otros nodos para degradar cerrojos.
d	Pending demote	Una solicitud para degradar diferida o remota
D	Demote	Una solicitud para degradar (local o remota)
f	Log flush	El registro necesita ser enviado antes de liberar este Glock
F	Frozen	Las respuestas de los nodos remotos se ignoran- la recuperación está en curso. Este indicador no se relaciona con la congelación del sistema de archivos, el cual usa un mecanismo diferente, pero se utiliza únicamente en recuperación.
i	Invalidate in progress	En el curso de invalidar páginas en este Glock
I	Initial	Establecer cuándo un bloqueo de DLM se asocia con este Glock
l	Locked	El Glock está en el proceso de cambio de estado
L	LRU	Se establece cuando el glock está en la lista LRU
o	Objeto	Se establece cuando el glock se asocia con un objeto (es decir, un inodo para tipo de glocks 2 y grupo de recursos para tipo de glocks3 )
p	Demote in progress	El Glock está en el proceso de responder a una solicitud de degradación
s	En cola	Se establece cuando un portador es puesto en cola para un glock y se retira cuando el glock se mantiene pero no quedan portadores remanentes. Se utiliza como parte de el algoritmo que calcula el mínimo de tiempo de retención para un glock.

Indicador	Nombre	Significado
r	Reply pending	La respuesta recibida del nodo remoto está pendiente de procesamiento
y	Dirty	Es necesario vaciar los datos al disco antes de lanzar este Glock

Tabla 2.2. Indicadores de portador de Glock

Indicador	Nombre	Significado
a	Async	No espera ningún resultado de Glock (emitirá un resultado más adelante)
A	Any	Se acepta cualquier modo de bloqueo compatible
c	No cache	Cuando se desbloquea, degrada de inmediato el bloque DML
e	No expire	Ignora las siguientes solicitudes de cancelación de bloqueo
E	exact	Debe tener un modo de bloqueo exacto
F	First	Cuando el portador es el primero en concedérsele este cerrojo
H	Holder	Indica que se ha otorgado el cerrojo solicitado
p	Priority	Pone a Holder a la cabeza de la cola
t	Try	Un cerrojo de "ensayo"
T	Try 1CB	Un cerrojo de "ensayo" que envía una llamada
W	Wait	Establecer mientras espera que la solicitud se complete

Al haber identificado el glock que está causando problemas, el siguiente paso es averiguar con qué inodo se relaciona. El número de glock (n: en la G: línea) lo indica. Es de la forma *tipo/número* y si el *tipo* es 2, entonces el glock es un glock de inodo y el *número* es un número de inodo. Para rastrear el inodo, puede entonces ejecutar **find -inum número** donde *número* es el número del inodo convertido desde el formato hex en el archivo glocks en decimales.



#### NOTA

Si ejecuta **find** en un sistema de archivos cuando experimente una contención de cerrojo, muy probable que el problema empeore. Se recomienda parar la aplicación antes de ejecutar **find** cuando esté buscando inodos enfrentados.

La [Tabla 2.3, "Tipos de glock"](#) muestra los significados de los diferentes tipos de glock.

**Tabla 2.3. Tipos de glock**

Número	Tipo de cerrojo	Uso
1	Trans	Cerrojo de transacción
2	Inodo	Metadatos y datos de inodos
3	Rgrp	Metadatos de grupo de recursos
4	Meta	El superbloque
5	lopen	El último nodo que se detectó
6	Flock	<b>flock(2)</b> syscall
8	Cuota	Operaciones de cuota
9	Diario	Diario mutex

Si el glock que ha sido identificado era de un tipo diferente, entonces es muy probable que sea del tipo 3: (grupo de recursos). Si usted ve una cantidad significativa de procesos que están esperando a otros tipos de glock bajo cargas normales, entonces por favor repórtelo al soporte técnico de Red Hat.

Si ve una cantidad de solicitudes esperando en cola en un cerrojo de grupo de recursos, puede haber varias explicaciones. Una es que hay un gran número de nodos comparados al número de grupo de recursos en el sistema de archivos. Otra es que el sistema de archivos puede estar casi lleno (requiriendo en promedio, búsquedas más prolongadas para bloques libres). La situación en ambos casos puede mejorarse añadiendo más almacenamiento y usando el comando **gfs2\_grow** para expandir el sistema de archivos.

## CAPÍTULO 3. CÓMO INICIAR

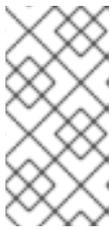
Este capítulo describe los procedimientos a seguir para la configuración de GFS2 y contiene las siguientes secciones:

- [Sección 3.1, “Tareas previas”](#)
- [Sección 3.2, “Tareas de configuración inicial”](#)

### 3.1. TAREAS PREVIAS

Debe completar las siguientes tareas antes de configurar GFS2 de Red Hat:

- Asegúrese de observar las características clave de la configuración de los nodos de GFS2 (consulte la [Sección 1.2, “Antes de configurar GFS2”](#)).
- Asegúrese de que los relojes en los nodos de GFS2 estén sincronizados. Se recomienda usar el software de Protocolo de tiempo de red (NTP) distribuido con su Red Hat Enterprise Linux.



#### NOTA

Los relojes del sistema en los nodos de GFS2 deben tener una diferencia de pocos minutos entre ellos para evitar actualizaciones innecesarias de marca de tiempo del inodo. La actualización innecesaria de marca de tiempo impacta de forma severa el rendimiento del clúster.

- Para usar GFS2 en un entorno de agrupamiento, configure su sistema para usar un Volumen lógico en clúster (CLVM), un grupo de extensiones de agrupamiento para el Gestor de volumen lógico, LVM. Para usar CLVM, el software de Red Hat Cluster Suite, incluido el demonio `clvmd` deben estar en ejecución. Para obtener información sobre el uso de CLVM, consulte *Administración del gestor de volumen lógico*. Para obtener información sobre cómo instalar y administrar Red Hat Cluster Suite, consulte *Administración de clúster*.

### 3.2. TAREAS DE CONFIGURACIÓN INICIAL

La configuración inicial de GFS2 consta de las siguientes tareas:

1. Configuración de volúmenes lógicos.
2. Creación del sistema de archivos GFS2.
3. Montaje del sistema de archivos.

Siga los siguientes pasos para establecer GFS2 inicialmente.

1. Con LVM, cree un volumen lógico para cada sistema de archivos GFS2 de Red Hat.



#### NOTA

Puede usar los scripts en `init.d` incluidos con Red Hat Cluster Suite para automatizar la activación y desactivación de los volúmenes lógicos. Para obtener más información sobre los scripts `init.d`, consulte *Configuración y administración de Red Hat Cluster*.

2. Cree un sistema de archivos GFS2 en los volúmenes lógicos creados en el paso 1. Escoja un nombre único para cada sistema de archivos. Para obtener más información sobre cómo crear sistemas de archivos GFS2, consulte la [Sección 4.1, “Cómo crear un sistema de archivos”](#).

Utilice cualquiera de los siguientes formatos para crear un sistema de archivos GFS2 en clúster:

```
mkfs.gfs2 -p lock_dlm -t ClusterName:FSName -j NumberJournals  
BlockDevice
```

```
mkfs -t gfs2 -p lock_dlm -t LockTableName -j NumberJournals  
BlockDevice
```

Para obtener más información sobre cómo crear un sistema de archivos GFS2, consulte la [Sección 4.1, “Cómo crear un sistema de archivos”](#).

3. En cada nodo, monte un sistema de archivos GFS2. Para obtener más información sobre cómo montar un sistema de archivos GFS2, consulte la [Sección 4.2, “Cómo montar un sistema de archivos”](#).

Uso del comando:

```
mount BlockDevice MountPoint
```

```
mount -o acl BlockDevice MountPoint
```

La opción de montaje **-o acl** permite la manipulación de las ACL en los archivos. Si un sistema de archivos es montado sin la opción **-o acl**, los usuarios pueden ver las ACL (con **getfacl**), pero no las pueden establecer (con **setfacl**).



#### NOTA

Use los scripts **init.d** incluidos en la adición de Alta disponibilidad de Red Hat para automatizar el montaje y desmontaje de sistemas de archivos GFS2.

## CAPÍTULO 4. ADMINISTRACIÓN DE GFS2

Este capítulo describe las tareas y comandos necesarios para administrar GFS2. Está conformado por las siguientes secciones:

- [Sección 4.1, “Cómo crear un sistema de archivos”](#)
- [Sección 4.2, “Cómo montar un sistema de archivos”](#)
- [Sección 4.3, “Cómo desmontar un sistema de archivos”](#)
- [Sección 4.5, “Administración de cuotas en GFS2”](#)
- [Sección 4.6, “Cómo expandir un sistema de archivos”](#)
- [Sección 4.7, “Cómo añadir diarios a un sistema de archivos”](#)
- [Sección 4.8, “Diario de datos”](#)
- [Sección 4.9, “Cómo configurar las actualizaciones `atime`”](#)
- [Sección 4.10, “Cómo suspender la actividad en un sistema de archivos”](#)
- [Sección 4.11, “Cómo reparar un sistema de archivos”](#)
- [Sección 4.12, “Nombres de rutas dependientes del contexto y montajes enlazados”](#)
- [Sección 4.13, “Orden de montaje de enlace y montaje de sistema de archivos”](#)
- [Sección 4.14, “La función de retiro de GFS2”](#)

### 4.1. CÓMO CREAR UN SISTEMA DE ARCHIVOS

El sistema de archivos GFS2 se crea con el comando `mkfs.gfs2`. También se puede utilizar el comando `mkfs` con la opción `-t gfs2`. Un sistema de archivos se crea en un volumen LVM activado. La siguiente información se requiere para ejecutar el comando `mkfs.gfs2`:

- El nombre del módulo/protocolo de cerrojo (el protocolo de cerrojo para un clúster es `lock_dlm`)
- Nombre de clúster (cuando se ejecuta como parte de una configuración de clúster)
- Número de diarios (un diario es requerido por cada nodo que monte el sistema de archivos)

Al crear un sistema de archivos GFS2, puede utilizar directamente el comando `mkfs.gfs2` o `mkfs` con el parámetro `-t` especificando un sistema de archivos de tipo `gfs2`, seguido de las opciones de sistema de archivos `gfs2`.



#### NOTA

Una vez haya creado un sistema de archivos GFS2 con el comando `mkfs.gfs2`, no podrá disminuir el tamaño del sistema de archivos. Sin embargo, puede aumentar el tamaño de un sistema de archivos existente con el comando `gfs2_grow`, tal como se describe en la [Sección 4.6, “Cómo expandir un sistema de archivos”](#).

## Uso

Al crear un sistema de archivos GFS2 en clúster, puede utilizar cualquiera de los siguientes formatos:

```
mkfs.gfs2 -p LockProtoName -t LockTableName -j NumberJournals BlockDevice
```

```
mkfs -t gfs2 -p LockProtoName -t LockTableName -j NumberJournals
BlockDevice
```

Al crear un sistema de archivos local GFS2, puede utilizar cualquiera de los siguientes formatos:



### NOTA

En el lanzamiento de Red Hat Enterprise Linux 6, Red Hat no soporta el uso del GFS2 como un sistema de archivos con un nodo único.

```
mkfs.gfs2 -p LockProtoName -j NumberJournals BlockDevice
```

```
mkfs -t gfs2 -p LockProtoName -j NumberJournals BlockDevice
```



### AVISO

Asegúrese de familiarizarse muy bien con el uso de los parámetros ***LockProtoName*** y ***LockTableName***. El uso incorrecto de los parámetros ***LockProtoName*** y ***LockTableName*** puede dañar el sistema de archivos o el espacio de cerrojos.

### ***LockProtoName***

Especifica el nombre del protocolo de cerrojo a usar. El protocolo de cerrojo para un clúster es ***lock\_dlm***.

### ***LockTableName***

Este parámetro se especifica para el sistema de archivos GFS2 en una configuración de clúster. Tiene dos partes separadas por dos puntos (sin espacios) así: ***clústerName:FSName***

- ***ClusterName***, el nombre del clúster para el cual el sistema de archivos GFS2 se ha creado.
- ***FSName*** es el nombre del sistema de archivos y puede tener entre 1 y 16 caracteres. El nombre debe ser único para todos los sistemas de archivos ***lock\_dlm*** en el clúster y para todos los sistemas de archivos (***lock\_dlm*** y ***lock\_nolock***) en cada nodo local.

### ***Number***

Especifica el número de diarios que serán creados por el comando ***mkfs.gfs2***. Se requiere un diario por cada nodo que monta el sistema de archivos. Para sistemas de archivos GFS2, se pueden

añadir más diarios posteriormente sin necesidad de expandir el sistema de archivos, así como se describe en la [Sección 4.7, “Cómo añadir diarios a un sistema de archivos”](#).

### **BlockDevice**

Especifica un volumen lógico o físico.

### **Ejemplos**

En este ejemplo, **lock\_dlm** es el protocolo de cerrojo que utiliza el sistema de archivos, ya que es un sistema de archivos en clúster. El nombre de clúster es **alpha** y el nombre del sistema de archivos es **mydata1**. El sistema de archivos contiene ocho diarios y es creado en **/dev/vg01/lvol0**.

```
mkfs.gfs2 -p lock_dlm -t alpha:mydata1 -j 8 /dev/vg01/lvol0
```

```
mkfs -t gfs2 -p lock_dlm -t alpha:mydata1 -j 8 /dev/vg01/lvol0
```

En este ejemplo, se crea un segundo sistema de archivos **lock\_dlm**, el cual se puede utilizar en el clúster **alpha**. El nombre del sistema de archivos es **mydata2**. El sistema de archivos contiene ocho diarios y se crea en **/dev/vg01/lvol1**.

```
mkfs.gfs2 -p lock_dlm -t alpha:mydata2 -j 8 /dev/vg01/lvol1
```

```
mkfs -t gfs2 -p lock_dlm -t alpha:mydata2 -j 8 /dev/vg01/lvol1
```

### **Opciones completas**

[Tabla 4.1, “Opciones del comando: mkfs.gfs2”](#) describe las opciones de comando **mkfs.gfs2** (indicadores y parámetros).

**Tabla 4.1. Opciones del comando: mkfs.gfs2**

Opción	Parámetro	Descripción
<b>-c</b>	<b><i>Megabytes</i></b>	Establece el tamaño inicial de cada archivo de cambio de cuota del diario en <b><i>Megabytes</i></b> .
<b>-D</b>		Activa los mensajes de salida de depurado.
<b>-h</b>		Ayuda. Muestra las opciones disponibles.
<b>-J</b>	<b><i>MegaBytes</i></b>	Especifica el tamaño del diario en megabytes. El tamaño predeterminado es 128 megabytes. El tamaño mínimo es 8 megabytes. Los diarios grandes mejoran el rendimiento pero utilizan más memoria que los diarios pequeños.

Opción	Parámetro	Descripción
<b>-j</b>	<b>Number</b>	Especifica el número de diarios que serán creados por el comando <b>mkfs.gfs2</b> . Un diario es requerido por cada nodo que monta el sistema de archivos. Si esta opción no se especifica, un diario será creado. En GFS2 se pueden crear diarios adicionales posteriormente sin necesidad de expandir el sistema de archivos.
<b>-o</b>		Previene que el comando <b>mkfs.gfs2</b> solicite la confirmación antes de escribir el sistema de archivos.
<b>-p</b>	<b>LockProtoName</b>	<div style="border: 1px solid #ccc; padding: 10px;"> <p>Especifica el nombre del protocolo de cerrojo a utilizar. Entre los protocolos de cerrojo reconocidos están:</p> <p><b>lock_dlm</b> — El módulo de bloqueo estándar, requerido por un sistema de archivos de clúster.</p> <p><b>lock_nolock</b> — Usado cuando GFS2 actúa como sistema de archivos local (en un solo nodo).</p> </div>
<b>-q</b>		En silencio. No muestra mensajes de salida.
<b>-r</b>	<b>MegaBytes</b>	Especifica el tamaño de los grupos de recursos en megabytes. El tamaño mínimo del grupo de recursos es 32 MB. El tamaño máximo es 2048 MB. Un grupo de recursos grande puede aumentar el rendimiento en sistemas de archivos muy grandes. Si no se especifica, <b>mkfs.gfs2</b> escoge el tamaño del grupo de recursos dependiendo del tamaño del sistema de archivos: el tamaño promedio del sistema de archivos tendrá grupos de recursos de 256 MB y sistemas de archivos más grandes tendrán grupos de recursos (RG) más grandes para mejorar el rendimiento.

Opción	Parámetro	Descripción
-t	<b>LockTableName</b>	<p>Un identificador único que especifica el campo de tabla de cerrojo se utiliza el protocolo <b>lock_dlm</b>; el protocolo <b>lock_nolock</b> no utiliza este parámetro.</p> <p>Este parámetro tiene dos partes separadas por dos puntos (sin espacio) de la siguiente manera: <b>ClusterName : FSName.</b></p> <p><b>ClusterName</b> es el nombre del clúster de Red Hat para el cual el sistema de archivos GFS2 ha sido creado; solamente los miembros de este clúster pueden utilizar este sistema de archivos. El nombre del clúster se establece en el archivo <b>/etc/cluster/cluster.conf</b> a través de la <b>Herramienta de configuración de Clúster</b> y se muestra en <b>Herramienta de estatus de Cluster</b> en la interfaz gráfica de administración del clúster Red Hat Cluster Suite.</p> <p><b>FSName</b> es el nombre del sistema de archivos, puede tener entre 1 y 16 caracteres. El nombre debe ser único entre todos los sistemas de archivos en el clúster.</p>
-u	<b>MegaBytes</b>	Especifica el tamaño inicial de cada archivo de etiquetas sin enlazar de diario.
-v		Muestra la información de la versión del comando.

## 4.2. CÓMO MONTAR UN SISTEMA DE ARCHIVOS

Antes de poder montar un sistema de archivos GFS2, el sistema de archivos debe existir (consulte la [Sección 4.1, “Cómo crear un sistema de archivos”](#)), el volumen donde el sistema de archivos existe debe estar activado y los sistemas de bloqueo y agrupamiento deben estar funcionando (consulte el documento *Configuración y administración de Red Hat Cluster*). Una vez que estos requerimientos hayan sido cumplidos, se puede montar el sistema de archivos GFS2 del mismo modo que se monta cualquier otro sistema de archivos de Linux.



## NOTA

El intento de montar un sistema de archivos GFS2 cuando el administrador de clúster (**cman**) no ha sido iniciado, produce el siguiente mensaje de error:

```
[root@gfs-a24c-01 ~]# mount -t gfs2 -o noatime
/dev/mapper/mpathap1 /mnt
gfs_control join connect error: Connection refused
error mounting lockproto lock_dlm
```

Para manipular las ACL de un archivo se debe montar el sistemas de archivos con la opción **-o acl**. Si un sistema de archivos es montado sin la opción **-o acl**, los usuarios pueden ver las ACL (con **getfacl**) pero no pueden establecerlas (con **setfacl**).

## Uso

### Montaje sin manipulación de ACL

```
mount BlockDevice MountPoint
```

### Montaje con manipulación de ACL

```
mount -o acl BlockDevice MountPoint
```

### **-o acl**

Opciones específicas de GFS2 que permiten la manipulación de las ACL de archivos.

### **BlockDevice**

Especifica los dispositivos de bloque en donde residen los sistemas de archivo GFS2.

### **MountPoint**

Especifica el directorio en donde debe montarse el sistema de archivos GFS2.

## Ejemplo

En este ejemplo, el sistema de archivos GFS2 en **/dev/vg01/lvol10** es montado en el directorio **/mygfs2**.

```
mount /dev/vg01/lvol10 /mygfs2
```

## Uso completo

```
mount BlockDevice MountPoint -o option
```

El argumento **-o opción** consta de opciones específicas de GFS2 (consulte la [Tabla 4.2, “Opciones de montaje específicas de GFS2”](#)) u opciones estándar aceptables de **mount -o** de Linux o una combinación de ambas. Varios parámetros de **option** están separados por una coma y sin espacios.

**NOTA**

El comando **mount** es un comando del sistema Linux. Además de poder usar las opciones específicas de GFS2 descritas en esta sección, puede utilizar otras opciones estándar del comando **mount** (por ejemplo **-r**). Para obtener información sobre otras opciones de comando **mount** de Linux, consulte la página de manual (man) de **mount**.

Tabla 4.2, “Opciones de montaje específicas de GFS2” describe los valores GFS2 **-o option** que pueden ser pasados a GFS2 en el momento de montaje.

**NOTA**

Esta tabla incluye descripciones de opciones que se utilizan con sistemas de archivos locales únicamente. Observe, sin embargo que para el lanzamiento de Red Hat Enterprise Linux 6, Red Hat no soporta el uso de GFS2 como un sistema de archivos de nodo único. Red Hat continuará ofreciendo soporte a los sistemas de archivos GFS2 para montar instantáneas de sistemas de sistemas de archivos (por ejemplo, para hacer copias de seguridad).

**Tabla 4.2. Opciones de montaje específicas de GFS2**

Opción	Descripción
<b>acl</b>	Permite la manipulación de las ACL de archivos. Si un sistema de archivos se monta sin la opción de montaje <b>acl</b> , los usuarios pueden ver las ACL (con <b>getfacl</b> ) pero no pueden establecerlas (con <b>setfacl</b> ).
<b>data=[ordered writeback]</b>	Cuando se establece <b>data=ordered</b> , los datos de usuario modificados en una transacción se volcan al disco antes de que la transacción sea enviada al disco. Esto debe evitar que el usuario vea bloques sin inicializar en un archivo después de una falla. Cuando se establece el modo <b>data=writeback</b> , los datos de usuario se escriben al disco en cualquier momento después de que es ensuciado; este modo no ofrece las mismas garantías de consistencia que el modo <b>ordered</b> , pero debe ser un poco más rápido en algunas cargas de trabajo. El valor predeterminado es el modo <b>ordered</b> .
<b>ignore_local_fs</b>  <b>Precaución:</b> Esta opción <i>no</i> debe ser utilizada cuando el sistema de archivos GFS2 es compartido.	Hace que GFS2 trate al sistema de archivos como un sistema de archivos multihost. Cuando se utiliza <b>lock_nolock</b> automáticamente se activa el indicador <b>localflocks</b> .

Opción	Descripción
<p><b>localflocks</b></p> <p><b>Precaución:</b> esta opción no debe ser usada cuando GFS2 es un sistema de archivos compartido.</p>	<p>Le avisa a GFS2 que debe dejar a VFS (Virtual File System) realizar todos los llamados flock y fcntl. La opción <b>localflocks</b> es activada automáticamente por <b>lock_nolock</b>.</p>
<b>lockproto=LockModuleName</b>	<p>Le permite al usuario especificar el protocolo de cerrojo a utilizar con el sistema de archivos. Si no se especifica <b>LockModuleName</b>, el nombre del protocolo de cerrojo se lee desde el superbloque del sistema de archivos.</p>
<b>locktable=LockTableName</b>	<p>Le permite al usuario especificar la tabla de bloqueo a utilizar con el sistema de archivos.</p>
<b>quota=[off/account/on]</b>	<p>Activa o desactiva las cuotas del sistema de archivos. Al establecer las cuotas para que estén en el estado <b>account</b> causa que las estadísticas de uso por UID/GID sean correctamente mantenidas por el sistema de archivos; los valores límite y de advertencia son ignorados. El valor predeterminado es <b>off</b>.</p>
<b>errors=panic withdraw</b>	<p>Cuando se especifica <b>errors=panic</b>, los errores de sistema causarán pánico de kernel. La conducta predeterminada, que es igual a especificar <b>errors=withdraw</b>, es para que el sistema se retire del sistema de archivos y lo convierta en inaccesible hasta el próximo reinicio; en algunos casos el sistema puede permanecer en ejecución. Para obtener información sobre la función de retiro de GFS2, consulte la <a href="#">Sección 4.14, "La función de retiro de GFS2"</a>.</p>
<b>discard/nodiscard</b>	<p>Hace que GFS2 genere solicitudes de E/S de "descarte" para bloques que han sido liberados. Estos bloques pueden ser utilizados por hardware apropiado para implementar esquemas de aprovisionamiento fino y similares.</p>
<b>barrier/nobarrier</b>	<p>Hace que GFS2 envíe barreras de E/S al vaciar el diario. El valor predeterminado es <b>encendido</b>. Esta opción automáticamente pasa a <b>apagado</b> si el dispositivo no soporta barreras de E/S. Se recomienda el uso de barreras de E/S con GFS2 en todo momento a menos que el dispositivo de bloque esté diseñado para que no pierda su contenido de escritura en cache (por ejemplo, si no es un UPS o si no tiene una memoria cache de escritura).</p>

Opción	Descripción
<b>quota_quantum=secs</b>	Establece el número de segundos para que se pueda establecer un cambio en la información de cuota en un nodo antes de que se escriba en el archivo de cuotas. Esa es la forma preferida de establecer este parámetro. El valor es un número entero de segundos mayor que cero. El valor predeterminado es 60 segundos. Configuraciones más cortas dan como resultado actualizaciones más rápidas de la información de cuota perezosa y es menos probable que alguien supere a su cuota. Configuraciones más prolongadas crean operaciones de sistemas de archivos que implican cuotas más rápidas y eficientes.
<b>statfs_quantum=secs</b>	Establecer <b>statfs_quantum</b> a 0 es la forma preferida de establecer la versión lenta de <b>statfs</b> . El valor predeterminado es 30 segundos que establece el tiempo máximo antes de que los cambios de <b>statfs</b> se sincronicen con el archivo maestro de <b>statfs</b> . Esto puede ajustarse para permitir valores <b>statfs</b> más rápidos y menos exactos o valores más lentos y menos exactos. Cuando esta opción se establece a 0, <b>statfs</b> siempre reportará los valores verdaderos.
<b>statfs_percent=valor</b>	Proporciona un vínculo en el cambio del porcentaje máximo en la información de <b>statfs</b> en la base local antes de que se vuelva a sincronizar al archivo maestro de <b>statfs</b> , incluso si el tiempo no ha expirado. Si la configuración de <b>statfs_quantum</b> es 0, entonces la configuración se ignora.

### 4.3. CÓMO DESMONTAR UN SISTEMA DE ARCHIVOS

El sistema de archivos GFS2 puede ser desmontado de la misma forma que cualquier otro sistema de archivos de Linux — con el comando **umount**.



#### NOTA

El comando **umount** es un comando del sistema Linux. Se puede encontrar información sobre este comando en las páginas de manual (man) de **umount**.

#### Uso

```
umount MountPoint
```

#### *MountPoint*

Especifica el directorio en donde el sistema de archivos GFS2 está montado actualmente.

## 4.4. CONSIDERACIONES ESPECIALES EN EL MONTAJE DE LOS SISTEMAS DE ARCHIVOS GFS2

El sistema de archivos GFS2 que ha sido montado en forma manual en lugar de automática a través de una entrada en el archivo **fstab**, no será reconocido por el sistema cuando el sistema de archivos sea desmontado en el momento de apagado del sistema. Como resultado, el script GFS2 no desmontará el sistema de archivos GFS2. Después de que el script de apagado de GFS2 se ejecute, el proceso estándar de apagado mata todos los procesos del usuario restantes, incluida la infraestructura de clúster, e intenta desmontar el sistema de archivos. Este desmontaje fallará sin la infraestructura de clúster y el sistema se colgará.

Para evitar que el sistema se cuelgue durante el desmontaje de los sistemas de archivos GFS2:

- Use siempre una entrada en el archivo **fstab** para montar el sistema de archivos GFS2.
- Si el sistema de archivos GFS2 ha sido montado en forma manual con el comando **mount**, asegúrese de desmontar el sistema de archivos con el comando **umount** antes de reiniciar o apagar el sistema.

Si el sistema de archivos se bloquea mientras está siendo desmontado durante el apagado del sistema en estas condiciones, realice un reinicio de hardware. Es poco probable que los datos se pierdan, puesto que el sistema de archivos se sincroniza antes en el proceso de apagado.

## 4.5. ADMINISTRACIÓN DE CUOTAS EN GFS2

Las cuotas de los sistemas de archivos se utilizan para limitar la cantidad de espacio del sistema de archivos que puede utilizar un usuario o un grupo. Ningún usuario o grupo tiene una cuota límite hasta que no se establezca. Cuando el sistema de archivos GFS2 se monta con la opción **quota=on** o **quota=account** guarda el registro del espacio utilizado por cada usuario y grupo aunque no hay límites establecidos. GFS2 actualiza la información de la cuota mediante transacciones para que en caso de una falla del sistema, los usos de cuotas no se tengan que reconstruir.

Para evitar una reducción de rendimiento, un nodo GFS2 solo sincroniza periódicamente las actualizaciones al archivo de cuotas. La contabilidad de cuotas "difusas" permite que usuarios y grupos se excedan levemente del límite. Para minimizar este efecto, GFS2 reduce en forma dinámica el periodo de sincronización cuando se alcanza el límite de cuota "duro".



### NOTA

A partir del lanzamiento de Red Hat Enterprise Linux 6.1, GFS2 soporta los servicios estándar de cuota de Linux. Para hacer uso de ellos necesita instalar el RPM de **quota**. Esta es la forma preferida para administrar cuotas en GFS2 y se debe utilizar para todas las implementaciones de GFS2 que usan cuotas. Esta sección documenta la administración de cuotas GFS2 mediante dichos servicios.

Para lanzamientos anteriores de Red Hat Enterprise Linux, GFS2 requería el comando **gfs2\_quota** para administrar cuotas. Si desea obtener más información sobre el uso del comando **gfs2\_quota**, consulte la [Apéndice A, Administración de cuotas GFS2 con el comando gfs2\\_quota](#).

### 4.5.1. Configuración de cuotas de disco

Para implementar cuotas de disco, siga los siguientes pasos:

1. Establezca las cuotas en modo impositivo o de contabilidad.
2. Inicie el archivo de base de datos de cuotas con información de uso de bloques actual.
3. Asigne políticas de cuotas. (En modo de contabilidad, estas políticas no se aplican.)

Cada uno de estos pasos se describe en detalle en las siguientes secciones.

#### 4.5.1.1. Cómo establecer cuotas en modo impositivo o de contabilidad

En los sistemas de archivos GFS2, las cuotas se inhabilitan automáticamente. Para habilitar las cuotas para un sistema de archivos, monte el sistema de archivos con la opción **quota=on** especificada.

Es posible mantener registro del uso de disco y mantener una contabilidad de cuotas para cada usuario sin tener que establecer los valores de advertencia y límite. Para ello, monte el sistema de archivos con la opción **quota=account** especificada.

##### Uso

Para montar un sistema de archivos con cuotas activadas, monte el sistema de archivos con la opción **quota=on** especificada.

```
mount -o quota=on BlockDevice MountPoint
```

Para montar un sistema de archivos con contabilidad de cuotas, a pesar de que los límites de cuotas no se han aplicado, monte el sistema de archivos con la opción **quota=account** especificada.

```
mount -o quota=account BlockDevice MountPoint
```

Para montar un sistema de archivos con cuotas desactivadas, monte el sistema de archivos con la opción **quota=off** especificada. Este es el valor predeterminado.

```
mount -o quota=off BlockDevice MountPoint
```

##### **quota={on|off|account}**

**on** - Especifica que las cuotas se activan o desactivan cuando se monta un sistema de archivos.

**off** - Especifica que las cuotas se desactivan cuando se monta el sistema de archivos.

**account** - Especifica que el uso de estadísticas de usuario y grupo es mantenido por el sistema de archivos, a pesar de que los límites de cuotas no están establecidos.

##### **BlockDevice**

Especifica los dispositivos de bloque en donde residen los sistemas de archivo GFS2.

##### **MountPoint**

Especifica el directorio en donde debe montarse el sistema de archivos GFS2.

##### Ejemplos

En este ejemplo, el sistema de archivos GFS2 en `/dev/vg01/lvol10` está montado en el directorio `/mygfs2` con cuotas activas.

```
mount -o quota=on /dev/vg01/lvol10 /mygfs2
```

En este ejemplo, el sistema de archivos GFS2 en `/dev/vg01/lvol10` está montado en el directorio `/mygfs2` con la contabilidad de cuotas activa mantenidas, pero no se imponen.

```
mount -o quota=account /dev/vg01/lvol10 /mygfs2
```

#### 4.5.1.2. Cómo crear archivos de base de datos de cuotas

Después de montar cada sistema de archivos con cuotas habilitadas, el sistema puede funcionar con cuotas de disco. Sin embargo, el sistema de archivos por sí mismo, no está listo para soportar cuotas. El siguiente paso es ejecutar el comando **quotacheck**.

El comando **quotacheck** examina los sistemas de archivos con cuotas activadas y genera una tabla del uso actual de disco por sistema de archivos. La tabla se utiliza para actualizar la copia del sistema operativo de uso de disco. Además, los archivos de cuotas de discos del sistema de archivos son actualizados.

Para crear los archivos de cuotas en el sistema de archivos, use las opciones **-u** y **-g** del comando **quotacheck**; ambas opciones deben especificarse para que las cuotas de usuario y grupo se inicien. Por ejemplo, si las cuotas están habilitadas para el sistema de archivos `/home`, cree los archivos en el directorio `/home`:

```
quotacheck -ug /home
```

#### 4.5.1.3. Cómo asignar cuotas por usuario

El último paso es asignar las cuotas de disco con el comando **edquota**. Observe que si ha montado el sistema de archivos en modo de contabilidad (con la opción **quota=account** especificada), las cuotas no se aplican.

Para configurar la cuota por usuario, como root en el indicador de shell prompt, ejecute el comando:

```
edquota nombre_de_usuario
```

Realice este paso para cada usuario que necesite una cuota. Por ejemplo, si habilita una cuota en `/etc/fstab` para la partición `/home` (`/dev/VolGroup00/LogVol102` en el siguiente ejemplo) y ejecuta el comando **edquota testuser**, en el editor configurado como el predeterminado para el sistema, se muestra lo siguiente:

```
Disk quotas for user testuser (uid 501):
Filesystem          blocks      soft      hard      inodes      soft
hard
/dev/VolGroup00/LogVol102  440436      0          0
```



## NOTA

El editor de texto definido por la variable de entorno **EDITOR** es utilizado por **edquota**. Para cambiar el editor, establezca la variable de entorno **EDITOR** en su archivo `~/ .bash_profile` para la ruta completa del editor de su escogencia.

La primera columna es el nombre del sistema de archivos con una cuota. La segunda columna muestra el número de bloques que el usuario está usando en el momento. Las siguientes dos columnas se utilizan para establecer los límites blandos y duros para el usuario en el sistema de archivos.

El límite de bloque blando define la cantidad máxima de espacio de disco que se puede utilizar.

El límite de bloque duro es la cantidad máxima absoluta de espacio de disco que puede usar un usuario o grupo. Cuando se alcanza este límite, no se puede usar más espacio de disco.

El sistema de archivos GFS2 no mantiene cuotas para inodos, por lo tanto estas columnas no se aplican al sistema de archivos GFS2 y quedarán en blanco.

Si alguno de los valores se establece a 0, ese límite no se establece. En el editor de texto, cambie los límites deseados. Por ejemplo:

```
Disk quotas for user testuser (uid 501):
Filesystem          blocks      soft      hard      inodes     soft
hard
/dev/VolGroup00/LogVol102  440436    500000    550000
```

Para verificar si la cuota para el usuario ha sido establecida, use el comando:

```
quota testuser
```

### 4.5.1.4. Cómo asignar cuotas por grupo

Las cuotas también se pueden asignar en función de cada grupo. Observe que si ha montado su sistema de archivos en modo de contabilidad (con la opción **account=on** especificada), las cuotas no se aplican.

Para establecer una cuota de grupos para el grupo **devel** (el grupo debe existir antes de establecer la cuota de grupos), use el siguiente comando:

```
edquota -g devel
```

Este comando despliega cuotas para el grupo en el editor de texto:

```
Disk quotas for group devel (gid 505):
Filesystem          blocks      soft      hard      inodes     soft      hard
/dev/VolGroup00/LogVol102  440400        0        0
```

El sistema de archivos GFS2 no mantiene cuotas para inodos, por lo tanto estas columnas no se aplican a sistemas de archivos GFS2 y quedarán en blanco. Modifique los límites y guarde el archivo.

Para verificar si la cuota de grupo ha sido establecida, use el siguiente comando:

```
quota -g devel
```

## 4.5.2. Cómo administrar cuotas de disco

Si las cuotas se implementan, deberán recibir mantenimiento — sobre todo para observar si las cuotas se exceden y asegurarse de que sean exactas.

Obviamente, si los usuarios exceden repetidamente sus cuotas o alcanzan sus límites blandos, el administrador de sistemas tendrá pocas alternativas dependiendo del tipo de usuarios y de cómo impacta la cantidad de espacio de disco en su trabajo. El administrador puede ayudar al usuario a determinar cómo usar menos espacio de disco o a incrementar el disco de cuota de usuario.

Puede crear un reporte de uso de disco al ejecutar la herramienta **repquota**. Por ejemplo, el comando **repquota /home** produce esta salida:

```
*** Report for user quotas on device /dev/mapper/VolGroup00-LogVol02
Block grace time: 7days; Inode grace time: 7days
  Block limits  File limits
User  used soft hard grace used soft hard grace
-----
root  --    36      0      0          4      0      0
kristin  --   540      0      0         125      0      0
testuser  -- 440400 500000 550000    37418      0      0
```

Para ver el reporte del uso de disco para todos los sistemas de archivos con cuota activada, (opción **-a**), use el comando:

```
repquota -a
```

Aunque el informe es fácil de leer, algunos puntos deben explicarse. **- -** que se muestra después de cada usuario es una forma rápida de determinar si los límites de bloques se han excedido. Si un límite blando de bloque se excede, aparecerá en la salida un **+** en lugar del primer **-**. El segundo **-** indica el límite de inodo, pero los sistemas de archivos GFS2 no soportan límites de inodo, por lo tanto ese caracter permanecerá como **-**. Los sistemas de archivos GFS2 no soportan un periodo de gracia, por lo tanto la columna **grace** permanecerá en blanco.

Observe que el comando **repquota** no recibe soporte en NFS, sea cual fuere el sistema de archivos subyacente.

## 4.5.3. Cómo mantener cuotas exactas

Si habilita las cuotas en su sistema de archivos después de un periodo de tiempo cuando ha estado ejecutando con cuotas inhabilitadas, ejecute el comando **quotacheck** para crear, revisar y reparar los archivos de cuotas. Además, ejecute **quotacheck** si considera que sus archivos de cuotas no son exactos, como puede ocurrir cuando el sistema de archivos no se monta correctamente después de un daño del sistema.

Para obtener más información sobre el comando **quotacheck**, consulte la página de manual **quotacheck**.



### NOTA

Ejecuta **quotacheck** cuando el sistema de archivos está relativamente inactivo en todos los nodos porque la actividad de discos puede afectar los valores de cuotas computados.

#### 4.5.4. Sincronización de cuotas con el comando `quotasync`

GFS2 almacena toda la información de cuotas en su propio archivo interno en disco. Un nodo de GFS2 no actualiza este archivo de cuotas para cada escritura de sistema de archivos, en su lugar, lo actualiza automáticamente cada 60 segundos. Es necesario evitar conflictos entre nodos que escriben al archivo de cuotas, lo cual reduciría el rendimiento.

Cuando un usuario o grupo se acerca a su límite de cuota, GFS2 reduce de forma dinámica el tiempo entre las actualizaciones del archivo de cuotas para evitar que se exceda el límite. El tiempo normal entre la sincronización de las cuotas es el parámetro ajustable `quota_quantum`. Puede cambiarlo de su valor predeterminado de 60 mediante la opción de montaje `quota_quantum=`, como se describe en la [Tabla 4.2, “Opciones de montaje específicas de GFS2”](#). El parámetro `quota_quantum` debe establecerse en cada nodo y cada vez que se monte el archivo. Los cambios al parámetro `quota_quantum` no persisten a través de desmontajes. Debe actualizar el valor `quota_quantum` con el comando `mount -o remount`.

Puede usar el comando `quotasync` para sincronizar la información de cuotas desde un nodo al archivo de cuotas en disco entre las actualizaciones automáticas realizadas por GFS2.

### Uso

#### Cómo sincronizar información de cuotas

```
quotasync [-ug] -a|mntpnt...
```

#### **u**

Sincroniza los archivos de cuotas de usuario.

#### **g**

Sincroniza los archivos de cuotas de grupos

#### **a**

Sincroniza todos los sistemas de archivos que actualmente tienen cuotas habilitadas y soporte de sincronización. Cuando `-a` está ausente, debe especificarse un punto de montaje de sistema de archivos.

#### **mntpnt**

Especifica el sistema de archivos GFS2 en el cual se aplicará la acción.

#### Cómo ajustar el tiempo entre sincronizaciones

```
mount -o quota_quantum=secs,remount BlockDevice MountPoint
```

#### **MountPoint**

Especifica el sistema de archivos GFS2 en el cual se aplicará la acción.

#### **secs**

Especifica el nuevo tiempo de sincronización del archivo de cuota por GFS2. Valores inferiores pueden incrementar los conflictos y reducir el rendimiento.

## Ejemplos

Este ejemplo sincroniza todas las cuotas sucias almacenadas en cache desde el nodo en que se ejecuta en el archivo `dd` cuotas ondisk para el sistema de archivos `/mnt/mygfs2`.

```
# quotasync -ug /mnt/mygfs2
```

Este ejemplo cambia los periodos predeterminados de las actualizaciones regulares del archivo de cuotas a una hora (3600 segundos) para el sistema de archivos `/mnt/mygfs2` al volver a montar ese sistema de archivos en un volumen lógico `/dev/volgroup/logical_volume`.

```
# mount -o quota_quantum=3600,remount /dev/volgroup/logical_volume
/mnt/mygfs2
```

### 4.5.5. Referencias

Para obtener más información sobre cuotas de disco, consulte los siguientes comandos en las páginas de manual `man`:

- `quotacheck`
- `edquota`
- `repquota`
- `quota`

## 4.6. CÓMO EXPANDIR UN SISTEMA DE ARCHIVOS

El comando `gfs2_grow` sirve para expandir un sistema de archivos GFS2 después de que el dispositivo donde el sistema de archivos reside ha sido expandido. Al ejecutar `gfs2_grow` en un sistema de archivos GFS2, se llena todo el espacio libre entre el final actual del sistema de archivos y el final del dispositivo, con una extensión recién inicializada del sistema de archivos GFS2. Cuando se completa esta operación, el índice de recursos para el sistema de archivos se actualiza. Todos los nodos en el clúster pueden utilizar el espacio de almacenamiento adicional que ha sido añadido.

El comando `gfs2_grow` debe ser ejecutado en un sistema de archivos montado, pero solo debe ser ejecutado en un nodo en el clúster. Todos los otros nodos verán la expansión y automáticamente empezarán a utilizar el nuevo espacio.



### NOTA

Una vez creado el sistema de archivos GFS2 con el comando `mkfs.gfs2`, no podrá disminuir el tamaño del sistema de archivos.

### Uso

```
gfs2_grow MountPoint
```

#### *MountPoint*

Especifica el sistema de archivos GFS2 en el cual se aplicará la acción.

## Comentarios

Antes de ejecutar el comando **gfs2\_grow**:

- Haga una copia de seguridad de los datos importantes en el sistema de archivos.
- Con el comando **df *MountPoint***, determine el volumen que está siendo usado por el sistema de archivos que va a ser expandido.
- Expanda el volumen del clúster subyacente con LVM. Para obtener más información sobre cómo administrar volúmenes LVM, vea *Administración del gestor de volumen lógico*.

Una vez ejecutado el comando **gfs2\_grow**, ejecute **df** para verificar si el nuevo espacio está disponible en el sistema de archivos.

## Ejemplos

En este ejemplo, el sistema de archivos en el directorio **/mygfs2fs** es expandido.

```
[root@dash-01 ~]# gfs2_grow /mygfs2fs
FS: Mount Point: /mygfs2fs
FS: Device:      /dev/mapper/gfs2testvg-gfs2testlv
FS: Size:       524288 (0x80000)
FS: RG size:    65533 (0xffffd)
DEV: Size:      655360 (0xa0000)
The file system grew by 512MB.
gfs2_grow complete.
```

## Uso completo

```
gfs2_grow [Options] {MountPoint | Device} [MountPoint | Device]
```

### *MountPoint*

Especifica el directorio en donde el sistema de archivos GFS2 está montado.

### *Device*

Especifica el nodo del dispositivo del sistema de archivos.

Tabla 4.3, “Opciones específicas de GFS2 disponibles mientras se expande un sistema de archivos” describe las opciones específicas de GFS2 que pueden utilizarse durante la expansión del sistema de archivos GFS2.

**Tabla 4.3. Opciones específicas de GFS2 disponibles mientras se expande un sistema de archivos**

Opción	Descripción
<b>-h</b>	Ayuda. Muestra un mensaje de uso corto.
<b>-q</b>	Silencioso. Baja el nivel de verbosidad.

Opción	Descripción
<b>-r MegaBytes</b>	Especifica el tamaño del nuevo grupo de recursos. El tamaño predeterminado es 256MB.
<b>-T</b>	Test. Hace todos los cálculos, pero no escribe ningún dato al disco y no expande el sistema de archivos.
<b>-V</b>	Muestra la información de la versión del comando.

## 4.7. CÓMO AÑADIR DIARIOS A UN SISTEMA DE ARCHIVOS

El comando **gfs2\_jadd** se utiliza para añadir diarios al sistema de archivos GFS2. Puede añadir diarios al sistema de archivos GFS2 de forma dinámica en cualquier momento, sin necesidad de expandir el volumen lógico subyacente. El comando **gfs2\_jadd** debe ser ejecutado en sistemas de archivos montados, pero debe ejecutarse en un solo nodo del clúster. Los demás nodos notarán que la expansión ha ocurrido.



### NOTA

Si un sistema de archivos GFS2 está lleno, el comando **gfs2\_jadd** fallará, incluso si el volumen lógico que contiene el sistema de archivos ha sido extendido y es más grande que el sistema de archivos. Esto se debe a que en un sistema de archivos GFS2, los diarios son archivos plenos y no metadatos incorporados, por lo tanto, la sola extensión del volumen lógico subyacente no proporcionará espacio para los diarios.

Antes de añadir diarios a un sistema de archivos GFS, puede utilizar la opción **journals** de **gfs2\_tool** para saber cuántos diarios tiene actualmente el sistema de archivos GFS2. El siguiente ejemplo muestra el número y el tamaño de los diarios en el sistema de archivos montado en **/mnt/gfs2**.

```
[root@roth-01 ../cluster/gfs2]# gfs2_tool journals /mnt/gfs2
journal2 - 128MB
journal1 - 128MB
journal0 - 128MB
3 journal(s) found.
```

### Uso

```
gfs2_jadd -j Number MountPoint
```

#### *Number*

Especifica el número de diarios a añadir.

#### *MountPoint*

Especifica el directorio en donde el sistema de archivos GFS2 está montado.

### Ejemplos

En este ejemplo, se añade un diario al sistema de archivos en el directorio `/mygfs2`

```
gfs2_jadd -j1 /mygfs2
```

En este ejemplo, se añaden dos diarios al sistema de archivos en el directorio `/mygfs2`

```
gfs2_jadd -j2 /mygfs2
```

## Uso completo

```
gfs2_jadd [Options] {MountPoint | Device} [MountPoint | Device]
```

### *MountPoint*

Especifica el directorio en donde el sistema de archivos GFS2 está montado.

### *Device*

Especifica el nodo del dispositivo del sistema de archivos.

La [Tabla 4.4, “Opciones específicas de GFS2 disponibles al añadir diarios”](#) describe las opciones específicas de GFS2 que se pueden utilizar para adicionar diarios a un sistema de archivos de GFS2.

**Tabla 4.4. Opciones específicas de GFS2 disponibles al añadir diarios**

Opción	Parámetro	Descripción
<b>-h</b>		Ayuda. Muestra un mensaje corto sobre el uso del comando.
<b>-J</b>	<i>MegaBytes</i>	Especifica el tamaño del nuevo diario en megabytes. El tamaño predeterminado es 128 MB. El tamaño mínimo permitido es de 32 MB. Para añadir diarios de tamaños diferentes al sistema de archivos, el comando <b>gfs2_jadd</b> debe ser ejecutado por cada tamaño diferente de diario. El tamaño especificado se redondea a la cifra inferior para que sea un múltiplo del tamaño de segmentos de diario que fue especificado cuando el sistema de archivos fue creado.
<b>-j</b>	<i>Number</i>	Especifica el número de diarios nuevos que serán añadidos con el comando <b>gfs2_jadd</b> . El valor predeterminado es 1.
<b>-q</b>		Silencioso. Baja el nivel de verbosidad.
<b>-v</b>		Muestra la información de la versión del comando.

## 4.8. DIARIO DE DATOS

Generalmente, GFS2 escribe solo metadatos en su diario. Luego, el contenido de los archivos se escribe al disco mediante la sincronización periódica del kernel que vuelca los búferes del sistema de archivos. Una llamada **fsync()** en un archivo, hace que los datos del archivo se escriban al disco de

forma inmediata. La llamada retorna cuando el disco informa que todos los datos han sido escritos de forma exitosa.

El diario de datos puede resultar en un tiempo de **fsync()** reducido, para archivos muy pequeños, porque los datos del archivo se escriben al diario además de los metadatos. Esta ventaja se reduce rápidamente cuando el tamaño del archivo aumenta. La escritura a un medio y a archivos más grandes será mucho más lenta con diarios de datos encendidos.

Las aplicaciones que dependen de **fsync()** para sincronizar los datos de archivo pueden tener un rendimiento mejorado con el uso de diarios de datos. Los diarios de datos pueden activarse automáticamente para cualquier archivo GFS2 creado en un directorio con la opción correcta (y todos sus subdirectorios). Los archivos existentes de longitud cero pueden también tener diarios de datos encendidos o apagados.

Activar el diario de datos en un directorio establece el directorio a "inherit jdata", lo cual indica que todos los archivos y directorios subsiguientes creados en ese directorio usarán el diario. Puede activar o desactivar el diario de datos en un archivo con el comando **chattr**.

Los siguientes comandos activan el diario de datos en el archivo **/mnt/gfs2/gfs2\_dir/newfile** y luego verifican si el indicador se ha establecido correctamente.

```
[root@roth-01 ~]# chattr +j /mnt/gfs2/gfs2_dir/newfile
[root@roth-01 ~]# lsattr /mnt/gfs2/gfs2_dir
-----j--- /mnt/gfs2/gfs2_dir/newfile
```

Los siguientes comandos desactivan el diario de datos en el archivo **/mnt/gfs2/gfs2\_dir/newfile** y luego verifican si el indicador se ha establecido correctamente.

```
[root@roth-01 ~]# chattr -j /mnt/gfs2/gfs2_dir/newfile
[root@roth-01 ~]# lsattr /mnt/gfs2/gfs2_dir
----- /mnt/gfs2/gfs2_dir/newfile
```

También puede utilizar el comando **chattr** para establecer el indicador **j** en un directorio. Cuando se establece este indicador para un directorio, todos los archivos y directorios creados posteriormente en ese directorio son puestos en el diario. La siguiente serie de comandos establece el indicador **j** en el directorio **gfs2\_dir**, luego verifica si el indicador se ha establecido correctamente. Después de esto, los comandos crean un archivo nuevo llamado **newfile** en el directorio **/mnt/gfs2/gfs2\_dir** y verifican si el indicador **j** ha sido establecido para el archivo. Debido a que el indicador **j** se establece para el directorio, **newfile** también deberá tener activado el diario.

```
[root@roth-01 ~]# chattr -j /mnt/gfs2/gfs2_dir
[root@roth-01 ~]# lsattr /mnt/gfs2
-----j--- /mnt/gfs2/gfs2_dir
[root@roth-01 ~]# touch /mnt/gfs2/gfs2_dir/newfile
[root@roth-01 ~]# lsattr /mnt/gfs2/gfs2_dir
-----j--- /mnt/gfs2/gfs2_dir/newfile
```

## 4.9. CÓMO CONFIGURAR LAS ACTUALIZACIONES ATIME

Cada inodo de archivo y directorio tiene tres marcas de tiempo asociadas:

- **ctime** — La última vez en que el estado del inodo cambió

- **mtime** — La última vez en que los datos del archivo (o directorio) fueron modificados
- **atime** — La última vez en que los datos del archivo (o directorio) fueron accedidos

Si las actualizaciones **atime** están activadas como lo están en GFS2 y otros sistemas de archivos de Linux, cada vez que el archivo es leído, su inodo necesita ser actualizado.

Ya que pocas aplicaciones utilizan la información proporcionada por **atime**, estas actualizaciones requieren una cantidad innecesaria de tráfico de escritura y tráfico de bloqueo de archivos. Este tráfico puede reducir el rendimiento; por lo tanto, es preferible apagar o reducir la frecuencia de actualizaciones **atime**.

Hay dos métodos para reducir los efectos de las actualizaciones de **atime**:

- Montar con **relatime** (relative atime), el cual actualiza **atime** si la actualización **atime** previa es menos reciente que **mtime** o **ctime**.
- Montar con **noatime**, el cual desactiva actualizaciones **atime** en el sistema de archivos.

#### 4.9.1. Montar con relatime

Se puede especificar la opción de montaje **relatime** (relative atime) de Linux cuando se monta el sistema de archivos. Esta opción especifica que **atime** se actualiza si la actualización de **atime** previa es menos reciente que **mtime** o **ctime**.

##### Uso

```
mount BlockDevice MountPoint -o relatime
```

##### **BlockDevice**

Especifica los dispositivos de bloque en donde residen los sistemas de archivo GFS2.

##### **MountPoint**

Especifica el directorio en donde debe montarse el sistema de archivos GFS2.

##### Ejemplo

En este ejemplo, el sistema de archivos GFS2 reside en **/dev/vg01/lvol0** y se monta en el directorio **/mygfs2**. Las actualizaciones **atime** solamente tienen lugar si la actualización **atime** previa es menos reciente que **mtime** o **ctime**.

```
mount /dev/vg01/lvol0 /mygfs2 -o relatime
```

#### 4.9.2. Montar con noatime

La opción de montaje de Linux **noatime**, puede ser especificada al momento de montar el sistema de archivos, el cual desactiva las actualizaciones de **atime** en ese sistema de archivos.

##### Uso

```
mount BlockDevice MountPoint -o noatime
```

### ***BlockDevice***

Especifica los dispositivos de bloque en donde residen los sistemas de archivo GFS2.

### ***MountPoint***

Especifica el directorio en donde debe montarse el sistema de archivos GFS2.

## **Ejemplo**

En este ejemplo, el sistema de archivos GFS2 reside en `/dev/vg01/lvol0` y es montado en el directorio `/mygfs2` con las actualizaciones `atime` apagadas.

```
mount /dev/vg01/lvol0 /mygfs2 -o noatime
```

## **4.10. CÓMO SUSPENDER LA ACTIVIDAD EN UN SISTEMA DE ARCHIVOS**

Puede suspender la actividad de escritura a un sistema de archivos con el comando `dmsetup suspend`. La suspensión de escritura permite que las instantáneas de dispositivos basadas en hardware sirvan para capturar el sistema de archivos en un estado consistente. El comando `dmsetup resume` finaliza la suspensión.

### **Uso**

#### **Iniciar suspensión**

```
dmsetup suspend MountPoint
```

#### **Terminar suspensión**

```
dmsetup resume MountPoint
```

### ***MountPoint***

Especifica el sistema de archivos.

## **Ejemplos**

Este ejemplo suspende la escritura al sistema de archivos `/mygfs2`.

```
# dmsetup suspend /mygfs2
```

Este ejemplo finaliza la suspensión de la escritura al sistema de archivos `/mygfs2`.

```
# dmsetup resume /mygfs2
```

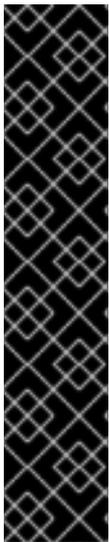
## 4.11. CÓMO REPARAR UN SISTEMA DE ARCHIVOS

Cuando los nodos fallan con el sistema de archivos montado, el diario del sistema de archivos permite una recuperación rápida. Sin embargo, si un dispositivo de almacenamiento pierde el poder o es físicamente desconectado, el sistema de archivos puede corromperse (los diarios no pueden ser usados para recuperarse de las fallas del subsistema de almacenaje). Cuando ocurre este tipo de corrupción, puede recuperar el sistema de archivos GFS2 con el comando **fsck.gfs2**.



### IMPORTANTE

El comando **fsck.gfs2** debe ser ejecutado solo en sistemas de archivos que están desmontados de todos los nodos.



### IMPORTANTE

No debe revisar el sistema de archivos GFS2 en el momento de arranque con el comando **fsck.gfs2**. El comando **fsck.gfs2** no puede determinar en tiempo de arranque si el sistema de archivos es montado por otro nodo en el clúster. Ejecute el comando **fsck.gfs2** de forma manual solo después de que el sistema arranque.

Para asegurarse de que el comando **fsck.gfs2** no se ejecute en un sistema de archivos GFS2 en el momento de arranque, modifique el archivo **/etc/fstab** para que las dos últimas columnas para un punto de montaje de un sistema de archivos GFS2 muestren "0 0" en lugar de "1 1" (o cualquier otro número), como en el siguiente ejemplo:

```
/dev/VG12/lv_svr_home /svr_home gfs2
defaults,noatime,nodiratime,noquota 0 0
```



### NOTA

Si tiene experiencia previa con el uso del comando **gfs\_fsck** en los sistemas de archivos GFS, observe que el comando **fsck.gfs2** difiere de algunas versiones anteriores de **gfs\_fsck** así:

- Al presionar **Ctrl+C** mientras ejecuta **fsck.gfs2** se interrumpe el procesamiento y muestra un indicador que le solicita si desea abortar el comando, ignorar el resto del paso actual o continuar el procesamiento.
- Puede incrementar el nivel de verbosidad con la opción **-v**. Si se añade una segunda opción **-v**, se incrementará aún más el nivel de verbosidad.
- Puede reducir el nivel de verbosidad con la opción **-q**. Si añade una segunda opción **-q** el nivel de verbosidad se reducirá de nuevo.
- La opción **-n** abre un sistema de archivos para sólo lectura y responde automáticamente **no** a cualquier pregunta. Esta opción ofrece una manera de observar los errores sin permitir que el comando **fsck.gfs2** surta efecto.

Consulte la página de manual **fsck.gfs2** para obtener más información acerca de otras opciones de comando.

La ejecución del comando **fsck.gfs2** requiere que la memoria del sistema esté por encima o más allá

de la memoria utilizada para el sistema operativo y el kernel. Cada bloque de memoria en el sistema de archivos GFS2 requiere aproximadamente cinco bits de memoria adicional, o 5/8 de un byte. Por lo tanto, para estimar cuántos bytes de memoria necesitará para ejecutar el comando **fsck.gfs2** en su sistema de archivos, determine cuántos bloques contiene el sistema de archivos y multiplíquelo por 5/8.

Por ejemplo, para determinar aproximadamente cuánta memoria se requiere para ejecutar el comando **fsck.gfs2** en un sistema de archivos GFS2 que es de 16TB con un tamaño de bloque de 4K, divida 16TB por 4K para determinar cuántos bloques de memoria contiene el sistema de archivos:

```
17592186044416 / 4096 = 4294967296
```

Puesto que el sistema de archivos contiene 4294967296 bloques, multiplique ese número por 5/8 para determinar cuántos bytes de memoria se requieren:

```
4294967296 * 5/8 = 2684354560
```

Este sistema de archivos aproximadamente requiere 2.6 GB de memoria libre para ejecutar el comando **fsck.gfs2**. Observe que si el tamaño del bloque fuera de 1K, la ejecución del comando **fsck.gfs2** requeriría cuatro veces la memoria o aproximadamente 11 GB.

## Uso

```
fsck.gfs2 -y BlockDevice
```

### -y

La opción **-y** hace que todas las preguntas sean respondidas afirmativamente con **yes**. Con la opción **-y** especificada, el comando **fsck.gfs2** no le preguntará nada antes de realizar los cambios.

### **BlockDevice**

Especifica los dispositivos de bloque en donde residen los sistemas de archivo GFS2.

## Ejemplo

En este ejemplo, el sistema de archivos GFS2 que reside en el dispositivo de bloque **/dev/testvol/testlv** es reparado. Todas las preguntas para reparar se responden automáticamente con **yes**.

```
[root@dash-01 ~]# fsck.gfs2 -y /dev/testvg/testlv
Initializing fsck
Validating Resource Group index.
Level 1 RG check.
(level 1 passed)
Clearing journals (this may take a while)...
Journals cleared.
Starting pass1
Pass1 complete
Starting pass1b
Pass1b complete
Starting pass1c
Pass1c complete
```

```
Starting pass2
Pass2 complete
Starting pass3
Pass3 complete
Starting pass4
Pass4 complete
Starting pass5
Pass5 complete
Writing changes to disk
fsck.gfs2 complete
```

## 4.12. NOMBRES DE RUTAS DEPENDIENTES DEL CONTEXTO Y MONTAJES ENLAZADOS

Los sistemas de archivos GFS2 no ofrecen soporte para Nombres de rutas dependientes del contexto (CDPN), los cuales permiten crear enlaces simbólicos que apuntan a directorios o archivos de destinos variables. Para obtener esta funcionalidad en GFS2, puede utilizar la opción **bind** del comando **mount**.

La opción **bind** del comando **mount** le permite remontar parte de la jerarquía de archivos en una ubicación diferente, mientras que permanece también disponible en la ubicación original. El formato de este comando es el siguiente:

```
mount --bind olddir newdir
```

Después de ejecutar este comando, el contenido del directorio **olddir** estará disponible en dos ubicaciones **olddir** y **newdir**. Puede también usar esta opción para que un archivo esté disponible en dos lugares diferentes:

Por ejemplo, después de ejecutar los siguientes comandos, el contenido de **/root/tmp** tendrá el mismo contenido que el directorio **/var/log** anteriormente montado.

```
[root@mencryfa ~]# cd ~root
[root@mencryfa ~]# mkdir ./tmp
[root@mencryfa ~]# mount --bind /var/log /root/tmp
```

Alternativamente, puede utilizar una entrada en el archivo **/etc/fstab** para lograr el mismo resultado durante el tiempo de montajes. La siguiente entrada de **/etc/fstab** hará que el contenido de **/root/tmp** sea igual que el contenido de **/var/log**.

```
/var/log                /root/tmp                none    bind
0 0
```

Después de montar el sistema de archivos, puede utilizar el comando **mount** para ver si el sistema de archivos ha sido montado, como en el siguiente ejemplo:

```
[root@mencryfa ~]# mount | grep /tmp
/var/log on /root/tmp type none (rw,bind)
```

Con un sistema de archivos que soporta el nombre de rutas dependiente del contexto, usted podría tener que definir el directorio **/bin** como un nombre de ruta dependiente del contexto que resuelve en una de las siguientes rutas, dependiendo de la arquitectura del sistema.

```
/usr/i386-bin
/usr/x86_64-bin
/usr/ppc64-bin
```

Esta misma funcionalidad se puede lograr al crear un directorio vacío **/bin**. Luego, mediante un script o una entrada en el archivo **/etc/fstab**, monte cada uno de los directorios de arquitectura individuales en el directorio **/bin** con el comando **mount -bind**. Por ejemplo, puede utilizar el siguiente comando como una línea en un script:

```
mount --bind /usr/i386-bin /bin
```

Alternativamente, puede utilizar la siguiente entrada en el archivo **/etc/fstab**.

```
/usr/i386-bin          /bin          none          bind          0 0
```

Un montaje enlazado puede proporcionar mayor flexibilidad que un nombre de ruta dependiente del contexto, ya que puede utilizar esta funcionalidad para montar diferentes directorios con cualquier criterio (como el valor de **%fill** para el sistema de archivos). Los nombres de rutas dependientes del contexto son más limitados en lo que pueden abarcar. Observe, sin embargo, que tendrá que escribir sus propios scripts para realizar el montaje según el criterio, como por ejemplo, el valor de **%fill**.



#### AVISO

Cuando monta un sistema de archivos con la opción **bind** y el sistema de archivo original fue montado con **rw**, el nuevo sistema de archivos será montado como **rw** incluso si usted pasa la opción **ro**; la opción **ro** es ignorada silenciosamente. En este caso, el nuevo sistema de archivos podría ser marcado como **ro** en el directorio **/proc/mounts**, lo cual es confuso.

## 4.13. ORDEN DE MONTAJE DE ENLACE Y MONTAJE DE SISTEMA DE ARCHIVOS

Cuando utilice la opción **bind** del comando **mount**, usted debe asegurarse de que los sistemas de archivos se monten en el orden correcto. En el siguiente ejemplo, el directorio **/var/log** debe estar montado antes de ejecutar el montaje enlazado en el directorio **/tmp**:

```
# mount --bind /var/log /tmp
```

El orden del montaje del sistema de archivos se determina así:

- En general, el orden del montaje del sistema de archivos está determinado por el orden en el que aparecen en el archivo **fstab**. Las excepciones a este orden son sistemas de archivos montados con el indicador **\_netdev** o sistemas de archivos que tienen sus propios scripts **init**.
- Un sistema de archivos con su propio script **init** se monta más adelante en el proceso de inicialización, después del sistema de archivos en el archivo **fstab**.

- Los sistemas de archivos montados con el indicador **\_netdev** se montan cuando la red ha sido activada en el sistema.

Si su configuración requiere que usted cree un montaje enlazado para montar el sistema de archivos GFS2, puede ordenar su archivo **fstab** así:

1. Monte los sistemas de archivos locales que se requieran para el montaje de enlace.
2. Vincule el montaje del directorio en el cual monta el sistema de archivos GFS2.
3. Monte el sistema de archivos GFS2.

Cuando su configuración requiere que usted monte en enlace un directorio local o sistema de archivos en un sistema de archivos GFS2, si lista los sistemas de archivos en el orden correcto en el archivo **fstab** no montará el sistema de archivos correctamente, puesto que GFS2 solamente se montará hasta que se ejecute el script **init** de GFS2. En este caso, debe escribir un script **init** para ejecutar el montaje de enlace para que el montaje de enlace se realice después de que se monte el sistema de archivos GFS2.

El siguiente es un ejemplo de un script personalizado de **init**. El script realiza un montaje de enlace de dos directorios encima de dos directorios de sistemas de archivos GFS2. En este ejemplo, hay un punto de montaje de GFS2 en **/mnt/gfs2a**, el cual se monta cuando el script **init** de GFS2 se ejecuta, después del inicio de clúster.

En este script de ejemplo, los valores de la declaración **chkconfig** indican lo siguiente:

- 345 indica los niveles de ejecución en los que el script se iniciará
- 29 es la prioridad de inicio, la cual indica que el script se ejecutará en el inicio después del script **init** GFS2, el cual tiene una prioridad de inicio de 26.
- 73 es la prioridad de parada, lo cual indica que en este caso el script se detendrá durante el apagado antes del script GFS2, el cual tiene una prioridad de parada de 74

Los valores de inicio y de parada indican que puede realizar manualmente la acción indicada al ejecutar el comando **service start** y un comando **service stop**. Por ejemplo, si el script se llama **fredwilma**, ejecute **service fredwilma start**.

Este script debe ponerse en el directorio **/etc/init.d** con los mismos permisos como los otros scripts en ese directorio. Puede entonces ejecutar un comando **chkconfig on** para enlazar el script a los niveles de ejecución indicados. Por ejemplo, si el script se llama **fredwilma**, entonces puede ejecutar **chkconfig fredwilma on**.

```
#!/bin/bash
#
# chkconfig: 345 29 73
# description: mount/unmount my custom bind mounts onto a gfs2
# subdirectory
#
### BEGIN INIT INFO
# Provides:
### END INIT INFO

. /etc/init.d/functions
```

```

case "$1" in
  start)
    # In this example, fred and wilma want their home directories
    # bind-mounted over the gfs2 directory /mnt/gfs2a, which has
    # been mounted as /mnt/gfs2a
    mkdir -p /mnt/gfs2a/home/fred &> /dev/null
    mkdir -p /mnt/gfs2a/home/wilma &> /dev/null
    /bin/mount --bind /mnt/gfs2a/home/fred /home/fred
    /bin/mount --bind /mnt/gfs2a/home/wilma /home/wilma
    ;;

  stop)
    /bin/umount /mnt/gfs2a/home/fred
    /bin/umount /mnt/gfs2a/home/wilma
    ;;

  status)
    ;;

  restart)
    $0 stop
    $0 start
    ;;

  reload)
    $0 start
    ;;

  *)
    echo $"Usage: $0 {start|stop|restart|reload|status}"
    exit 1
esac

exit 0

```

## 4.14. LA FUNCIÓN DE RETIRO DE GFS2

La función de *retiro* de GFS2 es una funcionalidad de integridad de datos de sistemas de archivos de GFS2 en un clúster. Si el módulo de kernel de GFS2 detecta una inconsistencia en un sistema de archivos GFS2, debido a una operación de E/S, el sistema de archivos no estará disponible para el clúster. La operación de E/S se detiene y el sistema esperará otras operaciones de E/S para detenerse con un error, evitando más daños. Cuando esto ocurra, puede detener de forma manual, cualquier otro servicio o aplicación, tras de lo cual puede reiniciar o remontar el sistema de archivos GFS2 para reproducir los diarios. Si el problema persiste, desmonte el sistema de archivos de todos los nodos en el clúster y recupere el sistema de archivos con el comando **fsck.gfs2**. La función de retiro de GFS es menos severa que la de un pánico de kernel, lo cual haría que otro nodo cerque el nodo.

Si su sistema está configurado con el script de inicio de **gfs2** habilitado y el sistema de archivos GFS2 se incluye en el archivo **/etc/fstab**, el sistema de archivos GFS2 volverá a montarse durante el reinicio. Si el sistema de archivos GFS2 se retiró debido a que percibió un daño en el sistema de archivos, se recomienda ejecutar el comando **fsck.gfs2** antes de volver a montar el sistema de archivos. En este caso, para evitar que su sistema de archivos se vuelva a montar en el momento de arranque, puede realizar lo siguiente:

1. Temporalmente desactive el script de inicio en el nodo afectado mediante el siguiente comando:

```
# chkconfig gfs2 off
```

2. Reinicie el nodo afectado, al iniciar el software de clúster. El sistema de archivos GFS2 no será montado.
3. Desmunte el sistema de archivos de cada nodo en el clúster.
4. Ejecute **fsck.gfs2** en el sistema de archivos desde un nodo solamente para asegurarse de que no hay sistemas de archivos corruptos.
5. Rehabilita el script de inicio en el nodo afectado ejecutando el siguiente comando:

```
# chkconfig gfs2 on
```

6. Vuelva a montar el sistema de archivos GFS2 de todos los nodos en el cluster.

Un ejemplo de una inconsistencia que puede generar un retiro de GFS2 es un conteo de bloques incorrecto. Cuando el kernel de GFS borra un archivo de un sistema de archivos, sistemáticamente elimina todos los datos y bloques de metadatos con ese archivo. Cuando se hace esto, revisa el conteo de bloques. Si el conteo de bloques no es uno (lo que significa que lo resta es el inodo de disco mismo), eso indica una inconsistencia en un sistema de archivos puesto que el conteo de bloques no coincidió con la lista de bloques encontrada.

Puede sobrescribir la función de retiro de GFS2 al montar el sistema de archivos con la opción **-o errors=panic** especificada. Cuando se especifica esta opción, los errores que normalmente harían que el sistema de archivos se retire provocarían pánico en el sistema. Esta opción detiene comunicaciones de clúster de nodo, las cuales cercan al nodo con vallas.

Internamente, la función de retiro de GFS2 funciona al hacer que el kernel envíe un mensaje al demonio **gfs\_control** solicitando el retiro. El demonio **gfs\_control** ejecuta el programa **dmsetup** para colocar el destino de error de mapeador de dispositivo bajo el sistema de archivos previniendo el acceso al dispositivo de bloque. Luego, le dice al kernel que ha completado. Esta es la razón por la cual se requiere que el soporte de GFS2 siempre utilice el dispositivo CLVM bajo GFS2, puesto que de lo contrario no es posible insertar un destino de mapeador de dispositivo.

El propósito del destino de error de un mapeador de dispositivo es el de garantizar que todas las futuras operaciones de E/S resulten en un error de E/S que permitan el desmontaje ordenado del sistema de archivos. Como resultado, cuando se presente el retiro, es normal ver una cantidad de errores de E/S del mapeador de dispositivos reportado en los registros de sistema.

En ocasiones, el retiro puede fallar si no es posible para el programa **dmsetup** insertar el destino de error como se solicitó. Esto puede suceder si hay una escasez de memoria en el momento del retiro y no se puede reclamar debido al problema que activó el retiro en primer lugar.

Un retiro no siempre significa que hay un error en GFS2. Algunas veces la función de retiro puede activarse por errores de E/S del dispositivo relativos al dispositivo de bloque. Es altamente recomendable revisar los registros para ver si ese es el caso si se presenta un retiro.

## CAPÍTULO 5. DIAGNÓSTICO Y CORRECCIÓN DE PROBLEMAS CON SISTEMAS DE ARCHIVOS GFS2

Este capítulo proporciona información sobre algunos problemas comunes de GFS2 y cómo resolverlos.

### 5.1. EL SISTEMA DE ARCHIVOS GFS2 PRESENTA UN BAJO RENDIMIENTO

Es posible que su sistema de archivos GFS2 presente un rendimiento más bajo que un sistema de archivos ext3. El rendimiento GFS2 puede afectarse por un número de influencias y en algunos casos de uso. La información que aborda estos problemas de rendimiento de GFS2 se encuentra a través de todo el documento.

### 5.2. EL SISTEMA DE ARCHIVOS GFS2 SE CUELGA Y REQUIERE REARRANQUE DE UN NODO

Si su sistema de archivos GFS2 se cuelga y no retorna comandos, pero al reiniciar un nodo específico retorna al sistema normal, esto puede ser indicativo de un problema de bloqueo o algún error. Si esto llega a ocurrir, reúna la siguiente información:

- El vaciado de gfs2 para el sistema de archivos en cada nodo:

```
cat /sys/kernel/debug/gfs2/fsname/glocks >glocks.fsname.nodename
```

- El vaciado de cierre de DLM para el sistema de archivos en cada nodo: Puede obtener esta información con **d1m\_tool**:

```
d1m_tool lockdebug -sv lname.
```

En este comando, *lname* está el nombre de lockspace utilizado por DLM para el sistema de archivos en cuestión. Puede hallar este valor en la salida del comando **group\_tool**.

- La salida desde el comando **sysrq -t**.
- El contenido del archivo **/var/log/messages**.

Una vez que haya reunido la información, puede abrir un tiquete con Red Hat Support y proveer los datos que ha reunido.

### 5.3. EL SISTEMA DE ARCHIVOS GFS2 SE CUELGA Y DEBE REARRANCAR TODOS LOS NODOS

Si su sistema de archivos GFS2 se cuelga y no retorna comandos, requiriendo el reinicio de todos los nodos en el clúster antes de usarlo, revise los siguientes problemas.

- Puede que haya fallado una valla. Los sistemas de archivos GFS2 se congelarán para garantizar integridad de datos en el evento de que una valla falle. Revise los registros de mensajes para ver si hay vallas fallidas en el momento de colgarse. Asegúrese de que el cercado esté configurado correctamente.
- El sistema de archivos GFS2 puede haberse retirado. Busque en los registros de mensajes la palabra **withdraw** y revise si hay algún mensaje o trazado de llamadas de GFS2 indicando

que el sistema de archivos se ha caído. Un retiro es indicativo de daño de sistema de archivos, una falla o un error. Desmonte el sistema de archivos, actualice el paquete **gfs2-utils**, y ejecute el comando **fsck** en el sistema de archivos para retornarlo al servicio. Abra un tiquete de soporte con Asistencia técnica de Red Hat. Infórmeles que ha experimentado un retiro de GFS2 y proporcione los **sosreports** con registros.

Para obtener información sobre la función de retiro de GFS2, consulte la [Sección 4.14, “La función de retiro de GFS2”](#).

- Este error puede ser indicativo de un problema de bloqueo o error. Reúna información durante los sucesos y abra un tiquete de soporte con Asistencia técnica de Red Hat, como se describe en la [Sección 5.2, “El sistema de archivos GFS2 se cuelga y requiere re arranque de un nodo”](#).

## 5.4. EL SISTEMA DE ARCHIVOS GFS2 NO SE MONTA EN UN NODO DE CLÚSTER RECIÉN AÑADIDO

Si añadió un nuevo nodo al clúster y no puede montar su sistema de archivos GFS2 en ese nodo, puede que tenga menos diarios en el sistema de archivos GFS2 que nodos intentando acceder al sistema de archivos GFS2. Debe tener un diario por host de GFS2 que intente montar en el sistema de archivos (a excepción del sistema de archivos GFS2 montado con la opción de montaje **spectator**, puesto que ellos no requieren diarios). Añada diarios al sistema de archivos GFS2 con el comando **gfs2\_jadd**, como se describe en la [Sección 4.7, “Cómo añadir diarios a un sistema de archivos”](#).

## 5.5. ESPACIO INDICADO COMO USADO EN SISTEMA DE ARCHIVOS VACÍO

Si tiene un sistema de archivos GFS2 vacío, el comando **df** mostrará que se ha ocupado un espacio. Esto se debe a que los diarios de sistema de archivos GFS2 consumen el espacio (número de diarios \* tamaño del diario) en disco. Si creó un sistema de archivos GFS2 con una gran cantidad de diarios o si ha especificado un tamaño grande de diario, entonces verá en uso (número de diarios \* tamaño de diarios) al ejecutar el comando **df**. Incluso si no especificó un gran número de diarios o grandes diarios, sistemas de archivos GFS2 pequeños (en un rango de 1 GB o menor) se mostrará una gran cantidad de espacio que está en uso con el tamaño predeterminado de diario GFS2.

## CAPÍTULO 6. CÓMO CONFIGURAR EL SISTEMA DE ARCHIVOS GFS2 EN UN CLÚSTER PACEMAKER

El siguiente procedimiento es un resumen de los pasos requeridos para configurar un clúster Pacemaker:

Después de instalar el software de clúster y GFS2 y los paquetes LVM en clúster en cada nodo, inicie los servicios **cman**, **clvmd**, and **pacemaker** en cada nodo y cree el clúster Pacemaker. Debe configurar la valla para el clúster. Para obtener más información sobre cómo configurar un clúster Pacemaker, consulte *Cómo configurar Red Hat High Availability Add-On con Pacemaker*.

1. Establezca el parámetro de Pacemaker **no\_quorum\_policy** a **freeze**.



### NOTA

El valor de **no-quorum-policy** se predetermina a **stop**, indicando que cuando el cuórum se pierde, todos los recursos en la partición restante se detendrán inmediatamente. Por lo general, este valor predeterminado es el más seguro y la opción más óptima, pero a diferencia de la mayoría de los recursos, GFS2 requiere cuórum para funcionar. Cuando el cuórum se pierde, la aplicación que usa GFS2 se monta y no puede detenerse correctamente. Cualquier intento para detener estos recursos sin cuórum fallará, lo cual resulta en el cercado de todo el clúster cada vez que se pierda el cuórum.

Para afrontar esta situación, establezca el valor **no-quorum-policy=freeze** cuando GFS2 esté en uso. Esto significa que cuando el cuórum se pierde, la partición restante no hará nada hasta que se recupere el cuórum.

```
# pcs property set no-quorum-policy=freeze
```

2. Después de asegurarse de que el tipo que está viendo está configurado a 3 en el archivo **/etc/lvm/lvm.conf** para soportar el bloqueo en clúster, cree el LV en clúster y de formato al volumen con un sistema de archivos GFS2. Verifique si crea suficiente diarios para cada uno de los nodos en su clúster.

```
# pvcreate /dev/vdb
# vgcreate -Ay -cy cluster_vg /dev/vdb
# lvcreate -L5G -n cluster_lv cluster_vg
# mkfs.gfs2 -j2 -p lock_dlm -t rhel7-demo:gfs2-demo
/dev/cluster_vg/cluster_lv
```

3. Configure un recurso **clusterfs**.

No agregue el sistema de archivos al archivo **/etc/fstab** porque este será administrado como un recurso de clúster Pacemaker. Las opciones de montaje pueden especificarse como parte de la configuración de recursos con **options=options**. Ejecute el comando **pcs resource describe Filesystem** para obtener todas las opciones de configuración.

Este comando de creación de recursos de clúster especifica la opción de montaje **noatime**.

```
# pcs resource create clusterfs Filesystem
device="/dev/cluster_vg/cluster_lv" directory="/var/mountpoint"
fstype="gfs2" "options=noatime" op monitor interval=10s on-
```

```
fail=fence clone interleave=true
```

4. Verifique si el GFS2 se montó como esperaba

```
# mount |grep /mnt/gfs2-demo  
/dev/mapper/cluster_vg-cluster_lv on /mnt/gfs2-demo type gfs2  
(rw,noatime,seclabel)
```

5. (Opcional) Reinicie todos los nodos de clúster para verificar persistencia y recuperación de gfs2.

## APÉNDICE A. ADMINISTRACIÓN DE CUOTAS GFS2 CON EL COMANDO `GFS2_QUOTA`

A partir del lanzamiento de Red Hat Enterprise Linux 6.1, GFS2 soporta el estándar de servicios de cuotas de Linux. Para usarlo necesitará instalar el RPM `quota`. Esta es la forma preferida para administrar cuotas en GFS2 y se debe utilizar para todas las implementaciones de GFS2 que usan cuotas. Para obtener información sobre el uso del estándar de servicios de cuotas de Linux, consulte la [Sección 4.5, “Administración de cuotas en GFS2”](#).

Para lanzamientos anteriores de Red Hat Enterprise Linux, GFS2 requería el comando `gfs2_quota` para administrar cuotas. Este apéndice describe el uso del comando `gfs2_quota` para administrar las cuotas del sistema de archivos GFS2.

### A.1. CÓMO ESTABLECER CUOTAS CON EL COMANDO `GFS2_QUOTA`

La configuración de dos cuotas está disponible para cada ID (UID) de usuario ID (GID) de grupo: un *límite duro* y un *límite blando*.

Un límite blando es la cantidad de espacio que puede utilizarse. El sistema de archivos no permitirá al usuario o grupo usar más de esa cantidad de espacio de disco. Un valor de límite de *ceros* significa que no se impone ningún límite.

Un límite blando suele ser un valor menor que el dado por el límite duro. El sistema de archivos notificará al usuario o grupo cuando haya llegado al límite blando para advertirlos de la cantidad de espacio de disco que están usando. Un límite de advertencia blando de *ceros* significa que no se ha aplicado ningún límite.

Puede establecer límites mediante el comando `gfs2_quota`. El comando solamente necesita ser ejecutado en un nodo único donde GFS2 esté montado.

La aplicación de cuota no se establece de forma predeterminada en sistemas de archivos GFS2. Para permitir la contabilidad de cuotas, use la opción `quota=` del comando `mount` cuando monte el sistema de archivos GFS2, como se describe en la [Sección A.4, “Cómo realizar activación y desactivación de cuotas”](#).

#### Uso

##### Cómo establecer cuotas, límite duro

```
gfs2_quota limit -u User -l Size -f MountPoint
```

```
gfs2_quota limit -g Group -l Size -f MountPoint
```

##### Cómo configurar cuotas, límite de advertencia

```
gfs2_quota warn -u User -l Size -f MountPoint
```

```
gfs2_quota warn -g Group -l Size -f MountPoint
```

#### User

Un ID de usuario para limitar o advertir. Puede ser un nombre de usuario del archivo de contraseñas o el número de UID.

**Group**

Un ID de grupo para limitar o advertir. Puede ser un nombre de grupo o el número de GID.

**Size**

Especifica el nuevo valor para limitar o advertir. El valor predeterminado está en unidades de megabytes. Los indicadores **-k**, **-s** y **-b** cambian las unidades a Kilobytes, sectores y bloques de sistema de archivos, respectivamente.

**MountPoint**

Especifica el sistema de archivos GFS2 al cual se aplican las acciones.

**Ejemplos**

Este ejemplo establece el límite duro para usuario *Bert* a 1024 megabytes (1 gigabyte) en sistema de archivos */mygfs2*.

```
# gfs2_quota limit -u Bert -l 1024 -f /mygfs2
```

Este ejemplo establece el límite blando para ID de grupo de 21 a 50 kilobytes en sistema de archivos */mygfs2*.

```
# gfs2_quota warn -g 21 -l 50 -k -f /mygfs2
```

**A.2. DESPLIEGUE DE LÍMITES DE CUOTAS Y USO CON EL COMANDO GFS2\_QUOTA**

Los límites de cuotas y uso actual pueden desplegarse para un usuario o grupo específico mediante el comando **gfs2\_quota get**. El contenido total del archivo de cuotas puede desplegarse también mediante el comando **gfs2\_quota list** en cuyo caso, se listan todos los ID con un límite duro de no-cero, límite suave o valor.

**Uso****Despliegue de límite de cuotas para un usuario**

```
gfs2_quota get -u User -f MountPoint
```

**Despliegue de límites de cuotas para un grupo**

```
gfs2_quota get -g Group -f MountPoint
```

**Despliegue de todo el archivo de cuotas**

```
gfs2_quota list -f MountPoint
```

**User**

Un ID de usuario para desplegar información sobre usuario específico. Puede ser un nombre de usuario del archivo de contraseñas o el número de UID.

### **Group**

Un ID de grupo para desplegar información sobre un grupo específico. Puede ser un nombre de grupo del archivo de grupo o el nombre de GID.

### **MountPoint**

Especifica el sistema de archivos GFS2 al cual se aplican las acciones.

## **Salida de comando**

A continuación, se muestra Información de cuota GFS2 del comando **gfs2\_quota**:

```
user User: limit:LimitSize warn:WarnSize value:Value  
group Group: limit:LimitSize warn:WarnSize value:Value
```

Los números (valores) **LimitSize**, **WarnSize**, y **Value** están en unidades de megabytes de forma predeterminada. Al añadir los indicadores **-k**, **-s**, o **-b** a la línea de comandos cambian las unidades a kilobytes, sectores, o bloques de sistemas de archivos, respectivamente.

### **User**

Un nombre de usuario o ID al cual se asocian los datos.

### **Group**

Un nombre de grupo o ID al cual se asocian los datos.

### **LimitSize**

El límite duro establecido para el usuario o grupo. El valor es cero si no se ha establecido ningún límite.

### **Value**

La cantidad actual de espacio de disco utilizada por el usuario o grupo.

## **Comentarios**

Al desplegar información de cuotas, el comando **gfs2\_quota** no resuelve los UID y GID en nombres si se añade la opción **-n** a la línea de comandos.

Espacio asignado a los archivos ocultos de GFS2 que pueden dejarse fuera de los valores desplegados para el UUID y GID de root al añadir la opción **-d** para la línea de comandos. Esto es útil cuando se tratan de igualar los números de **gfs2\_quota** con los resultados del comando **du**.

## **Ejemplos**

Este ejemplo muestra información para todos los usuarios y grupos que tengan un límite establecido o que estén utilizando cualquier espacio de disco en un sistema de archivos **/mygfs2**.

```
# gfs2_quota list -f /mygfs2
```

Este ejemplo muestra información de cuotas en sectores para grupo **users** en el sistema de archivos **/mygfs2**.

```
# gfs2_quota get -g users -f /mygfs2 -s
```

### A.3. SINCRONIZACIÓN DE CUOTAS CON EL COMANDO GFS2\_QUOTA

GFS2 almacena toda la información de cuota en su archivo interno en disco. Un nodo GFS2 no actualiza este archivo de cuota para cada sistema escrito; en su lugar, de forma predeterminada, actualiza el archivo de cuota una vez cada 60 segundos. Esto se necesita para evitar la contención entre nodos que escriben al archivo de cuota, lo cual ocasionaría una reducción en rendimiento.

Cuando un usuario o grupo se acerca a su límite de cuota, GFS2 reduce de forma dinámica el tiempo entre las actualizaciones del archivo de cuotas para evitar que se exceda el límite. El tiempo normal entre la sincronización de las cuotas es el parámetro ajustable **quota\_quantum**. Puede cambiarlo de su valor predeterminado de 60 mediante la opción de montaje **quota\_quantum=**, como se describe en la [Tabla 4.2, “Opciones de montaje específicas de GFS2”](#). El parámetro **quota\_quantum** debe establecerse en cada nodo y cada vez que se monte el archivo. Los cambios al parámetro **quota\_quantum** no persisten a través de desmontajes. Debe actualizar el valor **quota\_quantum** con el comando **mount -o remount**.

Puede usar el comando **gfs2\_quota sync** para sincronizar la información de cuotas desde un nodo al archivo de cuotas en disco entre las actualizaciones automáticas realizadas por GFS2.

#### Uso

##### Información sobre sincronización de cuotas

```
gfs2_quota sync -f MountPoint
```

##### **MountPoint**

Especifica el sistema de archivos GFS2 al cual se aplican las acciones.

##### Cómo ajustar el tiempo entre sincronizaciones

```
mount -o quota_quantum=secs,remount BlockDevice MountPoint
```

##### **MountPoint**

Especifica el sistema de archivos GFS2 al cual se aplican las acciones.

##### **secs**

Especifica el periodo de tiempo entre sincronizaciones de archivos de cuota regulares por GFS2. Los valores inferiores pueden aumentar la contención y disminuir el rendimiento.

#### Ejemplos

Este ejemplo sincroniza la información de cuotas desde el nodo en el que se está ejecutando hasta el sistema de archivos **/mygfs2**.

```
# gfs2_quota sync -f /mygfs2
```

Este ejemplo cambia los periodos predeterminados de las actualizaciones regulares del archivo de cuotas a una hora (3600 segundos) para el sistema de archivos `/mnt/mygfs2` al volver a montar ese sistema de archivos en un volumen lógico `/dev/volgroup/logical_volume`.

```
# mount -o quota_quantum=3600,remount /dev/volgroup/logical_volume  
/mnt/mygfs2
```

## A.4. CÓMO REALIZAR ACTIVACIÓN Y DESACTIVACIÓN DE CUOTAS

En sistemas de archivos GFS2, la aplicación de cuotas está desactivada de forma predeterminada. Para activar la aplicación de cuotas para un sistema de archivos, monte el sistema de archivos con la opción **quota=on** especificada.

### Uso

```
mount -o quota=on BlockDevice MountPoint
```

Para montar un sistema de archivos con aplicación de cuotas desactivada, monte el sistema de archivos con la opción **quota=off** especificada. Esta es la configuración predeterminada.

```
mount -o quota=off BlockDevice MountPoint
```

### -o quota={on|off}

Especifica que la aplicación de cuotas está activada o desactivada cuando se monta el sistema de archivos.

### *BlockDevice*

Especifica el dispositivo de bloque donde reside el sistema de archivos GFS2.

### *MountPoint*

Especifica el directorio donde se monta el sistema de archivos GFS2.

## Ejemplos

En este ejemplo, el sistema de archivos GFS2 en `/dev/vg01/lvol0` está montado en el directorio `/mygfs2` con la aplicación de cuotas activada.

```
# mount -o quota=on /dev/vg01/lvol0 /mygfs2
```

## A.5. ACTIVACIÓN DE CONTABILIDAD DE CUOTAS

Es posible mantener el rastro del uso de disco y la contabilidad de cuotas por cada usuario y grupo al aplicar los valores de límite y advertencia. Para hacer esto, monte el sistema de archivos con la opción **quota=account** especificada.

### Uso

```
mount -o quota=account BlockDevice MountPoint
```

**-o quota=account**

Especifica que las estadísticas de uso de grupo y usuario se guarden en el sistema de archivos, aunque los límites de cuotas no se apliquen.

***BlockDevice***

Especifica el dispositivo de bloque donde reside el sistema de archivos GFS2.

***MountPoint***

Especifica el directorio donde se monta el sistema de archivos GFS2.

**Ejemplo**

En este ejemplo, el sistema de archivos GFS2 en **/dev/vg01/lvol0** está montado en el directorio **/mygfs2** con contabilidad de cuota activada.

```
# mount -o quota=account /dev/vg01/lvol0 /mygfs2
```

## APÉNDICE B. CÓMO CONVERTIR EL SISTEMA DE ARCHIVOS DE GFS A GFS2

El lanzamiento Red Hat Enterprise Linux 6 no soporta sistemas de archivos GFS, debe actualizar los sistemas de archivos GFS existentes a sistemas de archivos GFS2 con el comando **gfs2\_convert**. Observe que no debe realizar este procedimiento de conversión en un sistema de Red Hat Enterprise Linux 5 antes de actualizarlo a Red Hat Enterprise Linux 6.



### AVISO

Antes de convertir el sistema de archivos de GFS, debe hacer una copia de seguridad del sistema de archivos, ya que el proceso de conversión es irreversible y cualquier error durante la conversión puede hacer que el programa termine en forma abrupta y como consecuencia el sistema de archivos será inutilizable.

Antes de convertir un sistema de archivos de GFS, debe usar el comando **gfs\_fsck** para revisar el sistema de archivos y corregir los errores.

Si la conversión de un GFS a un GFS2 es interrumpida por una falla eléctrica o cualquier otro problema, reinicie la herramienta de conversión. No intente ejecutar el comando **fsck.gfs2** en el sistema de archivos hasta que la conversión esté completa.

Cuando se convierten sistemas de archivos completos o casi completos, es posible que no haya más espacio disponible para todas las estructuras de datos del sistema de archivos de GFS2. En tales casos, el tamaño de todos los diarios se reduce uniformemente para que quepa todo en el espacio disponible.

### B.1. CONVERSIÓN DE NOMBRES DE RUTAS DEPENDIENTES DE CONTEXTO

El sistema de archivos GFS2 no ofrece soporte para nombres de rutas dependientes de contexto (CDPN), el cual le permite crear enlaces simbólicos que apuntan a directorios o archivos de destino variables. Para obtener esta funcionalidad como sistemas de archivos CDPN en GFS2, utilice la opción **bind** del comando **mount**.

El comando **gfs2\_convert** identifica los CDPN y los reemplaza por directorios vacíos con el mismo nombre. Sin embargo, al configurar montajes de enlace para reemplazar los CDPN, necesita saber las rutas completas de los destinos de enlace de los CDPN que usted está reemplazando. Antes de convertir su sistema de archivos, utilice el comando **find** para identificar los enlaces.

El siguiente comando lista los enlaces simbólicos que apuntan a un CDPN de **hostname**:

```
[root@smoke-01 gfs]# find /mnt/gfs -lname @hostname
/mnt/gfs/log
```

Igualmente, usted puede ejecutar el comando **find** para otros CDPN (**mach**, **os**, **sys**, **uid**, **gid**, **jid**). Observe que como los nombres de CDPN pueden ser de la forma **@hostname** o **{hostname}**, necesitará ejecutar el comando **find** para cada variante.

Para obtener más información sobre montajes de enlace y nombres de ruta dependientes de contexto, consulte la [Sección 4.12, “Nombres de rutas dependientes del contexto y montajes enlazados”](#).

## B.2. PROCEDIMIENTO DE CONVERSIÓN DE GFS A GFS2

Utilice el siguiente procedimiento para convertir un sistema de archivos GFS a GFS2.

1. En un sistema de Red Hat Enterprise Linux, cree una copia de seguridad en su sistema de archivos GFS.
2. Desmonte todo el sistema de archivos GFS en todos los nodos del clúster.
3. Ejecute el comando **gfs\_fsck** en el sistema de archivos GFS para verificar que no haya corrupción del sistema de archivos.
4. Ejecute **gfs2\_convert gfsfilesystem**. El sistema desplegará advertencias y preguntas de confirmación antes de convertir **gfsfilesystem** a GFS2.
5. Actualice a Red Hat Enterprise Linux 6.

El siguiente ejemplo convierte un sistema de archivos GFS en el dispositivo de bloque **/dev/shell\_vg/500g** para un GFS2.

```
[root@shell-01 ~]# /root/cluster/gfs2/convert/gfs2_convert
/dev/shell_vg/500g
gfs2_convert version 2 (built May 10 2010 10:05:40)
Copyright (C) Red Hat, Inc. 2004-2006 All rights reserved.

Examining file system.....
This program will convert a gfs1 filesystem to a gfs2 filesystem.
WARNING: This can't be undone. It is strongly advised that you:

    1. Back up your entire filesystem first.
    2. Run gfs_fsck first to ensure filesystem integrity.
    3. Make sure the filesystem is NOT mounted from any node.
    4. Make sure you have the latest software versions.
Convert /dev/shell_vg/500g from GFS1 to GFS2? (y/n)y
Converting resource groups.....
Converting inodes.
24208 inodes from 1862 rgs converted.
Fixing file and directory information.
18 cdpn symlinks moved to empty directories.
Converting journals.
Converting journal space to rg space.
Writing journal #1...done.
Writing journal #2...done.
Writing journal #3...done.
Writing journal #4...done.
Building GFS2 file system structures.
Removing obsolete GFS1 file system structures.
Committing changes to disk.
/dev/shell_vg/500g: filesystem converted successfully to gfs2.
```

## APÉNDICE C. PUNTOS DE TRAZADO GFS2 Y ARCHIVO GLOCKS DEBUGFS

Este apéndice describe la interfaz de glock **debugfs** y los puntos de trazado de GFS2. Está dirigido a usuarios avanzados que están familiarizados con interioridades de sistemas de archivos y que desearían conocer más sobre el diseño de GFS2 y la forma de depuración de problemas específicos de GFS2.

### C.1. TIPOS DE PUNTOS DE TRAZADO GFS2

Actualmente hay tres tipos de trazado GFS2: puntos de trazado *glock*, puntos de trazado *bmap* y puntos de trazado *log*. Estos sirven para monitorizar un sistema de archivos GFS2 en ejecución y dar información adicional a la que pueda obtenerse con opciones de depuración con soporte en previos lanzamientos de Red Hat Enterprise Linux. Los puntos de trazado son muy útiles cuando se presenta un problema, tal como de rendimiento o de colgado, es reproducible y por lo tanto se puede obtener durante la operación problemática. En GFS2, los glocks son el mecanismo de control de cache y la clave para entender el rendimiento del núcleo de GFS2. Los puntos de trazado *bmap* (mapa de bloques) sirven para monitorizar asignaciones y bloquear mapeos (búsqueda de los bloques ya asignados en un árbol de metadatos en disco) a medida que ocurren y buscar problemas relacionados con el sitio de acceso. Los puntos de trazado del registro siguen el rastro de los datos que se escriben y publican del diario y pueden ofrecer información útil en la parte de GFS2.

Los puntos de trazado están diseñados para ser genéricos en lo posible. Esto significa que no se necesita cambiar la API durante el curso de Red Hat Enterprise Linux 6. Por otra parte, los usuarios de esta interfaz deben ser conscientes de que se trata de una interfaz de depuración y no parte del set normal de la API de Red Hat Enterprise Linux 6, y como tal Red Hat no garantiza que no se presenten cambios en la interfaz de puntos de trazado de GFS2.

Los puntos de trazado son una función genérica de Red Hat Enterprise Linux 6 y su alcance va más allá de GFS2. En particular, sirven para implementar la infraestructura de **blktrace** y los puntos de trazado **blktrace** pueden servir junto con los GFS2 para obtener un cuadro total del rendimiento del sistema. Debido al nivel en el que los puntos de trazado operan, pueden producir grandes volúmenes de datos en un periodo de tiempo muy corto. Están diseñados para poner un mínimo de carga en el sistema cuando están habilitados, pero es inevitable que tengan algún efecto. El filtraje de eventos mediante una variedad de medios puede ayudar a reducir el volumen de datos y a enfocarse en obtener solamente la información que es útil para entender una situación en particular.

### C.2. PUNTOS DE TRAZADO

Los puntos de trazado se encuentran en el directorio `/sys/kernel/debug/tracing/` si suponemos que **debugfs** está montado en el lugar estándar en el directorio `/sys/kernel/debug`. El subdirectorio **events** contiene todos los eventos de trazado que pueden especificarse y, siempre y cuando el módulo **gfs2** esté cargado, habrá un subdirectorio **gfs2** que contenga otros subdirectorios, uno para cada evento GFS2. El contenido del directorio `/sys/kernel/debug/tracing/events/gfs2` debe ser similar al siguiente:

```
[root@chywoon gfs2]# ls
enable          gfs2_bmap          gfs2_glock_queue    gfs2_log_flush
filter          gfs2_demote_rq     gfs2_glock_state_change gfs2_pin
gfs2_block_alloc gfs2_glock_put     gfs2_log_blocks      gfs2_promote
```

Para habilitar todos los puntos de trazado GFS2, ejecute el siguiente comando:

```
[root@chywoon gfs2]# echo -n 1
>/sys/kernel/debug/tracing/events/gfs2/enable
```

Para habilitar un punto de trazado específico, hay un archivo **enable** en cada uno de los subdirectorios de eventos individuales. También es cierto que el archivo **filter** puede servir para establecer un filtro de eventos o una serie de eventos. El significado de los eventos individuales se explica en más detalle a continuación.

La salida desde los puntos de trazado está disponible en formato o ASCII o en formato binario. Este apéndice no cubre actualmente la interfaz binaria. La interfaz ASCII está disponible en dos formas. Para listar el contenido actual del búfer de anillo, ejecute el siguiente comando:

```
[root@chywoon gfs2]# cat /sys/kernel/debug/tracing/trace
```

Esta interfaz es útil cuando está utilizando un largo proceso por un cierto periodo de tiempo y, después de algún evento, desea buscar la última información capturada en el búfer. Cuando se requiera toda la salida, puede utilizarse una interfaz alterna `/sys/kernel/debug/tracing/trace_pipe`. Los eventos se leen desde este archivo tal como se presentan; no hay historial disponible con esta interfaz. El formato de la salida es el mismo de ambas interfaces y se describe para cada uno de los eventos GFS2 en las secciones de este apéndice.

Una utilidad llamada **trace-cmd** está disponible para leer los datos de punto de trazado. Para obtener más información sobre esta herramienta, vaya al enlace en la [Sección C.10, “Referencias”](#). La herramienta **trace-cmd** puede servir de una forma similar para la herramienta **strace**, por ejemplo para ejecutar un comando mientras reúne la información de trazado de datos de varias fuentes.

### C.3. GLOCKS

El concepto más importante para entender el GFS2 y el que se establece al lado de otros sistemas de archivos, es el concepto de glocks. En términos de código de fuente, un glock es una estructura de datos que combina el DLM y la memoria cache en una sola máquina. Cada máquina tiene una relación 1:1 con un solo cerrojo DLM, y ofrece memoria cache para ese estado de cerrojo para que las operaciones repetitivas llevadas a cabo desde un solo nodo de un sistema de archivos no tenga que llamar varias veces al DLM, y por lo tanto, ayudan a evitar tráfico innecesario. Hay dos amplias categorías de glocks, los que guardan metadatos y los que no. Los glocks de inodo y los glocks de grupos de recursos ambos guardan en cache metadatos, otros tipos de glocks no guardan metadatos en cache. El glock de inodo también tiene que ver con el guardado en cache de datos, aparte de metadatos, y su lógica es la más compleja de todos los glocks.

**Tabla C.1. Modos Glock y modos de cerrojo DLM**

Modo Glock	Modo de cerrojo DLM	Notas
UN	IV/NL	Desbloqueado (No hay cerrojo DLM asociado con Glock o ningún cerrojo NL que dependa del indicador I)
SH	PR	Cerrojo compartido (protegido lectura)
EX	EX	Cerrojo exclusivo

Modo Glock	Modo de cerrojo DLM	Notas
DF	CW	Diferido (escritura simultánea) utilizado para E/S directa y congelación de sistema de archivos

Los Glocks permanecen en memoria hasta que se desbloqueen (a solicitud de algún nodo o a solicitud de la Máquina virtual) y si no hay otros usuarios locales. En ese momento se retiran de la tabla hash de glocks y se liberan. Cuando se crea un glock, el cerrojo DLM no se asocia con el glock de inmediato. El cerrojo DLM se asocia con el glock tras la primera solicitud al DLM, y si esta solicitud tiene éxito, entonces el indicador 'I' (initial) se establecerá en el glock. La [Tabla C.4, "Indicadores Glock"](#) muestra el significado de los diferentes indicadores. Una vez que el DLM haya sido asociado con el glock, el cerrojo DLM permanecerá por lo menos en el modo de cerrojo NL (Null) hasta que el glock deba ser liberado. La democión del cerrojo DLM del modo NL a desbloqueado es siempre la última operación en la vida de un glock.



### NOTA

Este aspecto particular de la conducta de cerrojo del DLM cambió desde Red Hat Enterprise Linux 5, el cual desbloquea algunas veces los cerrojos de DLM que se vinculan totalmente a los glocks y así, Red Hat Enterprise Linux 5 tiene un mecanismo diferente para garantizar que los bloques de valor de cerrojo (LVB) se preserven donde sea requerido. El nuevo esquema que utiliza Red Hat Enterprise Linux 6 se hizo posible debido a la fusión del módulo de cerrojo `lock_dlm` (no confundir con el DLM) dentro de GFS2.

Cada glock puede tener un número de "portadores" asociados, cada uno representa una solicitud de cerrojo de las capas más altas. Las llamadas del sistema relacionadas con GFS2 ponen y quitan de la cola a los portadores del glock para proteger la sección crítica de código.

La máquina de estados de glock se basa en 'workqueue'. Por razones de rendimiento, los 'tasklets' son preferibles, sin embargo, en la implementación actual, necesitamos enviar E/S desde ese contexto que prohíbe su uso.



### NOTA

Los 'Workqueues' tienen sus propios puntos de trazado que pueden utilizarse junto con los puntos de trazado de GFS2 si se desea.

La [Tabla C.2, "Modos Glock y tipos de datos"](#) muestra el estado en que pueden guardarse en cache cada uno de los modos de glock y si ese estado guardado en cache puede estar sucio. Esto aplica tanto a los cerrojos de grupo de recursos como al inodo, aunque no hay componente de datos para cerrojos de grupo de recursos, solo metadatos.

**Tabla C.2. Modos Glock y tipos de datos**

Modo Glock	Datos de cache	Metadatos de cache	Datos sucios	Metadatos sucios
UN	No	No	No	No

Modo Glock	Datos de cache	Metadatos de cache	Datos sucios	Metadatos sucios
SH	Yes	Yes	No	No
DF	No	Yes	No	No
EX	Yes	Yes	Yes	Yes

## C.4. INTERFAZ DE GLOCK DEBUGFS

La interfaz de glock **debugfs** permite la visualización del estado interno de los glocks y de los portadores y también incluye algunos detalles de resumen de los objetos que se bloquean en algunos casos. Cada línea del archivo puede comenzar por G: sin ninguna sangría (lo cual se refiere al glock) o por una letra diferente, con sangría con un solo espacio y se refiere a las estructuras asociadas con el glock inmediatamente por encima de este en el archivo (H: es un portador, I: un inodo, y R: un grupo de recursos) . Aquí vemos un ejemplo de cómo se podría ver el contenido:

```
G:  s:SH n:5/75320 f:I t:SH d:EX/0 a:0 r:3
   H:  s:SH f:EH e:0 p:4466 [postmark] gfs2_inode_lookup+0x14e/0x260 [gfs2]
G:  s:EX n:3/258028 f:yI t:EX d:EX/0 a:3 r:4
   H:  s:EX f:tH e:0 p:4466 [postmark] gfs2_inplace_reserve_i+0x177/0x780
   [gfs2]
   R:  n:258028 f:05 b:22256/22256 i:16800
G:  s:EX n:2/219916 f:yfI t:EX d:EX/0 a:0 r:3
   I:  n:75661/219916 t:8 f:0x10 d:0x00000000 s:7522/7522
G:  s:SH n:5/127205 f:I t:SH d:EX/0 a:0 r:3
   H:  s:SH f:EH e:0 p:4466 [postmark] gfs2_inode_lookup+0x14e/0x260 [gfs2]
G:  s:EX n:2/50382 f:yfI t:EX d:EX/0 a:0 r:2
G:  s:SH n:5/302519 f:I t:SH d:EX/0 a:0 r:3
   H:  s:SH f:EH e:0 p:4466 [postmark] gfs2_inode_lookup+0x14e/0x260 [gfs2]
G:  s:SH n:5/313874 f:I t:SH d:EX/0 a:0 r:3
   H:  s:SH f:EH e:0 p:4466 [postmark] gfs2_inode_lookup+0x14e/0x260 [gfs2]
G:  s:SH n:5/271916 f:I t:SH d:EX/0 a:0 r:3
   H:  s:SH f:EH e:0 p:4466 [postmark] gfs2_inode_lookup+0x14e/0x260 [gfs2]
G:  s:SH n:5/312732 f:I t:SH d:EX/0 a:0 r:3
   H:  s:SH f:EH e:0 p:4466 [postmark] gfs2_inode_lookup+0x14e/0x260 [gfs2]
```

El ejemplo anterior es una serie de extractos (de un archivo de 18 MB aproximadamente) generado por el comando `cat /sys/kernel/debug/gfs2/unity:myfs/glocks >my.lock` durante una ejecución del punto de referencia del matasellos en un único nodo de sistema de archivos GFS2. Los glocks en la figura han sido seleccionados para mostrar algunas de las funciones más interesantes de los vaciados de glock.

Los estados de glock pueden ser EX (exclusivo), DF (diferido), SH (compartido) o UN (desbloqueado). Estos estados corresponden directamente con extractos de modos de cerrojo DLM para 'UN' el cual puede representar un estado de cerrojo DLM null (nulo) o ese GFS2 no mantiene el cerrojo DLM (según el indicador I como se explica arriba). El s: campo del glock indica el estado actual del cerrojo y el mismo campo en el portador indica el modo solicitado. Si el cerrojo se otorga, el portador tendrá el bit H bit establecido en sus indicadores (f: campo). De lo contrario, deberá tener el bit W de esperar.

El campo n: (número) indica el número asociado con cada elemento. Para glocks, es el número del tipo seguido por el número de glock, entonces en el ejemplo anterior, el primer glock es n:5/75320; es decir,

un glock **iopen** se relaciona con el inodo 75320. En el caso del inodo y de los glocks **iopen**, el número de glock siempre es idéntico al número de bloque del disco del inodo.



## NOTA

Los números glock (n: campo) en el archivo de glocks `debugfs` están en hexadecimales, aunque en la salida de puntos de trazado aparece en decimales. Esto se debe a razones históricas; los números glock siempre se escribían en hexadecimales, pero se seleccionaron los decimales para los puntos de trazado a fin de que los números pudieran ser comparados fácilmente con la otra salida de punto de trazado (desde **blktrace** por ejemplo) y con salida desde **stat(1)**.

El listado completo de todos los indicadores para el portador y el glock se configuran en la [Tabla C.4, “Indicadores Glock”](#) y en la [Tabla C.5, “Indicadores de portadores de glock”](#). El contexto de bloques de valor de cerrojo no está disponible actualmente mediante la interfaz de glock **debugfs**.

La [Tabla C.3, “Tipos Glock”](#) muestra los significados de los diferentes tipos de Glock.

**Tabla C.3. Tipos Glock**

Número de tipo	Tipo de cerrojo	Uso
1	trans	Cerrojo de transacción
2	inodo	Metadatos y datos de inodo
3	rgrp	Metadatos de grupo de recursos
4	meta	El súperbloque
5	iopen	El último nodo que se detectó
6	Flock	<b>flock(2)</b> syscall
8	quota	Operaciones de cuota
9	Diario	Diario mutex

Uno de los indicadores de glock más importantes es el indicador I (Bloqueado). Este es el cerrojo de bits utilizado para arbitrar el acceso a un estado de glock cuando se realiza un cambio de estado. Se establece cuando la máquina de estados está a punto de enviar una petición de cerrojo a través del Gestor de cerrojo distribuido (DML), y solo se despeja cuando la operación se ha realizado. Algunas veces, esto puede significar que se ha enviado más de una solicitud de cerrojo con varias invalidaciones que se presentan entre los tiempos.

La [Tabla C.4, “Indicadores Glock”](#) muestra los significados de los diferentes indicadores Glock.

**Tabla C.4. Indicadores Glock**

Indicador	Nombre	Significado
d	Pending demote	Una solicitud para degradar diferida o remota
D	Demote	Una solicitud para degradar (local o remota)
f	Log flush	El registro necesita ser enviado antes de lanzar este Glock
F	Frozen	Responde desde los nodos remotos ignorados - recuperación en curso.
i	Invalidate in progress	En el proceso de invalidar páginas en este Glock
I	Initial	Definir cuando el cerrojo DLM esté asociado con este Glock
l	Locked	El Glock está en el proceso de cambiar estado
L	LRU	Se establece cuando el glock está en la lista LRU
o	Objeto	Se establece cuando el glock se asocia con un objeto (es decir, un inodo para tipo de glocks 2 y grupo de recursos para tipo de glocks3 )
p	Demote in progress	El Glock está en el proceso de responder a una solicitud de degradación
s	En cola	Se establece cuando un portador es puesto en cola para un glock y se retira cuando el glock se mantiene pero no quedan portadores remanentes. Se utiliza como parte de el algoritmo que calcula el mínimo de tiempo de retención para un glock.
r	Reply pending	Respuesta recibida del nodo remoto está a la espera del procesamiento
y	Dirty	Los datos necesitan vaciarse al disco antes de lanzar este glock

Cuando se recibe una llamada remota desde un nodo que desea obtener un cerrojo en un modo que entra en conflicto con el que se mantiene en un nodo local, luego uno u otro de los dos indicadores D (degradar) o d (pendiente para degradar) se establece. Con el fin de evitar las condiciones de hambre cuando hay contención en un cerrojo particular, a cada cerrojo se le asigna un tiempo de espera mínimo. El nodo que aún no tiene el cerrojo para el tiempo de espera mínimo puede retener ese cerrojo hasta que el intervalo de tiempo haya expirado.

Si el intervalo de tiempo ha expirado, entonces el indicador D (degradar) se enviará y el estado requerido será registrado. En ese caso, la próxima vez que no se otorguen cerrojos en la cola de portadores, el cerrojo será degradado. Si el tiempo de intervalo no ha expirado, entonces el indicador d (pendiente para degradar) se establecerá en su lugar. Esto también programa el la máquina de estados para limpiar d (pendiente para degradar) y establecer D (degradar) cuando el tiempo mínimo del portador mínimo ha expirado.

El indicador I (inicial) se establece cuando se ha asignado un cerrojo DLM al glock. Esto sucede cuando el glock se utiliza primero y cuando el indicador I permanece configurado hasta que el glock sea desbloqueado.

## C.5. PORTADORES DE GLOCK

La [Tabla C.5, "Indicadores de portadores de glock"](#) muestra el significado de los diferentes indicadores de portadores de glock.

**Tabla C.5. Indicadores de portadores de glock**

Indicador	Nombre	Significado
a	Async	No espera ningún resultado de Glock (emitirá un resultado más adelante)
A	Any	Acepta cualquier modo de cerrojo compatible
c	No cache	Cuando está desbloqueado, degrada al cerrojo DLM inmediatamente
e	No expire	Ignora solicitudes posteriores de cancelar el cerrojo
E	Exact	Debe tener el modo exacto de cerrojo
F	First	Cuando el portador es el primero en concedérsele este cerrojo
H	Holder	Indica que se ha otorgado el cerrojo solicitado
p	Priority	Pone al 'Portador' a la cabeza de la cola
t	Try	Un cerrojo de "ensayo"
T	Try 1CB	Un cerrojo de "ensayo" que envía una llamada
W	Wait	Se establece mientras espera que una solicitud se complete

Los indicadores de portadores son H (holder) y W (wait) como se mencionó anteriormente, ya que se establecen en solicitudes de cerrojo otorgadas y solicitudes de cerrojo en cola, respectivamente. El ordenamiento de los portadores en la lista es importante. Si hay portadores otorgados, siempre estarán en la cabeza de la cola, seguido por otros portadores en cola.

Si no hay portadores otorgados, entonces el primer portador de la lista será el primero en activar el siguiente cambio de estado. Debido a que las solicitudes para degradar son considerados de más prioridad que las solicitudes del sistema de archivos, puede ser que no siempre resulte en el cambio al estado solicitado.

El subsistema de glock soporta dos clases de cerrojos "try". Estos cerrojos son útiles porque permiten la toma de bloqueos del orden normal (con regreso y reintento apropiados) y porque pueden servir para ayudar a evitar recursos en uso por otros nodos. El cerrojo normal t (try) es básicamente lo que su

nombre indica; es el cerrojo de "ensayo" que no hace nada especial. Por otra parte, el cerrojo T (**try 1CB**), es idéntico al cerrojo t excepto que el DLM enviará una sola llamada a los portadores de cerrojos incompatibles. El cerrojo T (**try 1CB**) con los cerrojos **iopen**, sirve de árbitro entre nodos cuando el conteo de nodos **i\_nlink** es cero y, determina cuál de los nodos deberá cambiar la asignación del inodo. El glock **iopen** por lo general se lleva a cabo en el estado compartido, pero cuando el conteo **i\_nlink** se convierte en cero y **->delete\_inode()** es llamado, solicitará un cerrojo exclusivo con T (**try 1CB**). Si el cerrojo es otorgado, continuará cambiando la asignación del inodo. De lo contrario, hará que los nodos que impedían el cerrojo marque su(s) glock(s) con el indicador D (degradar), el cual se revisa en **->drop\_inode()** para garantizar que el cambio de asignación no se olvide.

Esto significa que los inodos que tienen un conteo de enlace cero, pero que aún están abiertos, serán cambiados de asignación por el nodo en el cual ocurre el **close()** final. También, al mismo tiempo que el conteo de enlace de inodo decrece a cero, el inodo se marca como si estuviera en el estado especial de conteo de enlace cero pero aún se utiliza en el bitmap de grupo de recursos. Esto funciona como un huérfano de la lista de sistema de archivos ext3 en que permite a cualquier lector posterior del bitmap saber que hay espacio potencial que podría ser reclamado, e intentar reclamarlo.

## C.6. PUNTOS DE TRAZADO GLOCK

Los puntos de trazado también se diseñan para poder confirmar la exactitud del control de la cache al combinarlos con la salida **blktrace** y con el conocimiento de la presentación en disco. Es entonces posible revisar si se ha emitido alguna E/S y completado en el cerrojo correcto y que no haya carreras en el momento.

El punto de trazado **gfs2\_glock\_state\_change** es el más importante. Hace el trazado de cada cambio de estado del glock desde la creación inicial correcta a través de la degradación final que termina con **gfs2\_glock\_put** y la NL final para transición desbloqueada. El indicador de glock I (Bloqueado) siempre se establece antes de que ocurra un cambio de estado y no será despejado sino hasta después de que termine. Nunca habrá portadores otorgados (indicador de portador de glock H) durante el cambio de estado. Si hay algún portador en cola, siempre estará en el estado W (espera). Cuando el cambio de estado se complete, entonces se podrán otorgar los portadores que es la operación final antes de que el indicador de glock I sea despejado.

El punto de trazado **gfs2\_demote\_rq** mantiene el rastro de las solicitudes 'degradar', tanto locales como remotas. Si suponemos que hay suficiente memoria en el nodo, las solicitudes locales de degradar pocas veces serán vistas y muy a menudo serán creadas por **umount** o al reclamar memoria ocasional. El número de solicitudes de degradar remotas es una medida de contención entre nodos para un inodo particular o grupo de recursos.

Cuando a un portador se le otorga un cerrojo, **gfs2\_promote** es llamado, esto sucede debido a que las etapas finales de un estado cambian o al solicitar un cerrojo que puede ser otorgado inmediatamente debido al estado de glock que ya está guardando en cache un cerrojo de un modo apropiado. Si es el primer portador otorgado para este glock, entonces el indicador f (primero) se establecerá en ese portador. Eso es utilizado actualmente solo por grupos de recursos.

## C.7. PUNTOS DE TRAZADO DE BMAP

La asignación de bloques es una tarea central para cualquier sistema de archivos. GFS2 usa un sistema de archivos tradicional basado en bitmap con dos bits por bloque. El principal propósito de los puntos de trazado en este subsistema es permitir la monitorización del tiempo tomado para asignar los bloques de mapas.

El punto de trazado **gfs2\_bmap** es llamado dos veces por cada operación **bmap**: una vez en el inicio para desplegar la solicitud de **bmap**, y otra al final para mostrar el resultado. Esto facilita la coincidencia

de solicitudes y resultados y la medida del tiempo tomado para mapear bloques en partes diferentes de sistemas de archivos, en diferentes desplazamientos de archivos, o incluso en diferentes archivos. También es posible ver el tamaño promedio cuando se devuelven en comparación con los que se solicitan.

Para mantener el rastro de los bloques asignados, **gfs2\_block\_alloc** no solo se llama para asignar, sino también para liberar bloques. Puesto que se hace referencia a las asignaciones según el inodo para el bloque destinado, esto puede servir para rastrear qué bloques físicos pertenecen a cuáles archivos en el sistema de archivos en vivo. Esto es en particular muy útil cuando se combina con **blktrace**, el cual muestra los patrones problemáticos de E/S que pueden referirse a los inodos relevantes mediante el mapa obtenido a través del punto de trazado.

## C.8. PUNTOS DE TRAZADO DE REGISTRO

Los puntos de trazado en este subsistema rastrean los bloques que se añaden o retiran del diario (**gfs2\_pin**), como también el tiempo tomado para enviar las transacciones al registro (**gfs2\_log\_flush**). Esto puede ser muy útil para depurar problemas de rendimiento de diarios.

El punto de trazado **gfs2\_log\_blocks** sigue el rastro de los bloques reservados en el registro, lo cual puede ayudar a ver si el registro es demasiado pequeño para la carga, por ejemplo:

El punto de trazado **gfs2\_aio\_flush** (Red Hat Enterprise Linux 6.2 y posterior) es similar al punto de trazado **gfs2\_log\_flush** porque ambos realizan el seguimiento del inicio y fin de las descargas de la lista AIL. La lista AIL contiene búferes que han pasado por el registro, pero que aún no han sido escritos en su lugar y se vacía con periodicidad para liberar más espacio de registro para ser usado por el sistema de archivos o cuando el proceso solicita un 'sync' o 'fsync'.

## C.9. ESTADÍSTICAS DE GLOCK

GFS2 mantiene estadísticas que pueden ayudar a rastrear lo que está sucediendo en el sistema de archivo. Le permite detectar los problemas de rendimiento.

GFS2 mantiene dos contadores:

- **dcount**, el cual cuenta el número de operaciones DLM solicitadas. Esto muestra cuántos datos entran en los cálculos de media o varianza..
- **qcount**, el cual cuenta el número del nivel de operaciones **syscall** solicitadas. Por lo general, **qcount** será igual o mayor que **dcount**.

Además, GFS2 mantiene 3 pares de media o de varianza. Los pares de media o varianza son estimaciones exponenciales regulares y el algoritmo utilizado es el que se utiliza para calcular veces de round trip en código de red. Los pares de media y varianza guardados en GFS2 no se escalan, pero están en unidades de nanosegundos enteros.

- **srtt/srttvar**: tiempo regular de round trip para operaciones de no-bloqueo
- **srttb/srttvarb**: tiempo regular de round trip para operaciones de bloqueo
- **irtt/irttvar**: tiempo de Inter-request (Por ejemplo, tiempo entre solicitudes DLM)

Una solicitud de no-bloqueo es aquella que se completa enseguida, sea cual sea el estado del cerrojo DLM en cuestión. Esto actualmente significa cualquier solicitud cuando: (a) el estado actual del cerrojo es exclusivo (b) el estado solicitado es nulo o no está bloqueado o (c) el indicador del "cerrojo try" está establecido. Una solicitud de bloqueo cubre las demás las solicitudes de bloqueo.

Los tiempos más prolongados son mejores para los IRTT, mientras que los tiempos más cortos son mejores para los RTT.

Las estadísticas se guardan en dos archivos **sysfs**:

- El archivo **glstats**. Este archivo es similar al archivo **glocks** a excepción de que este contiene estadísticas con un glock por línea. Los datos se inicializan de datos "por cpu" para dicho tipo de glock para el cual se crea el glock ( al lado de los contadores, los cuales se ponen en cero). Este archivo puede ser muy largo.
- El archivo **lkstats**. Contiene estadísticas "por cpu" para cada tipo de glock. Contiene una línea por estadística, en las cuales cada columna es un núcleo de CPU. Hay ocho líneas por tipo de glock, con tipos sucediéndose uno tras otro.

## C.10. REFERENCIAS

Para obtener más información sobre puntos de trazado y el archivo GFS2 **glocks**, consulte los siguientes recursos:

- Este apéndice ha sido adaptado parcialmente del documento entregado por Steve Whitehouse en el 'Linux Symposium 2009', el cual se encuentra en <http://git.kernel.org/?p=linux/kernel/git/torvalds/linux-2.6.git;a=blob;f=Documentation/filesystems/gfs2-glocks.txt;h=0494f78d87e40c225eb1dc1a1489acd891210761;hb=HEAD>.
- Para obtener información sobre reglas internas de cerrojo de glocks, consulte <http://git.kernel.org/?p=linux/kernel/git/torvalds/linux-2.6.git;a=blob;f=Documentation/filesystems/gfs2-glocks.txt;h=0494f78d87e40c225eb1dc1a1489acd891210761;hb=HEAD>.
- Para obtener información sobre el trazado de eventos, consulte <http://git.kernel.org/?p=linux/kernel/git/torvalds/linux-2.6.git;a=blob;f=Documentation/trace/events.txt;h=09bd8e9029892e4e1d48078de4d076e24eff3dd>.
- Para obtener información sobre la herramienta **trace-cmd**, consulte <http://lwn.net/Articles/341902/>.

## APÉNDICE D. HISTORIA DE REVISIONES

<b>Revisión 7.1-3.2</b> translated	<b>Tue Feb 3 2015</b>	<b>Gladys Guerrero-Lozano</b>
<b>Revisión 7.1-3.1</b> Los archivos de traducción sincronizados con fuentes XML 7.1-3	<b>Tue Feb 3 2015</b>	<b>Gladys Guerrero-Lozano</b>
<b>Revisión 7.1-3</b> Actualización para implementar sort_order en la página de inicio de RHEL 6.	<b>Tue Dec 16 2014</b>	<b>Steven Levine</b>
<b>Revisión 7.0-9</b> Versión para lanzamiento GA 6.6	<b>Wed Oct 8 2014</b>	<b>Steven Levine</b>
<b>Revisión 7.0-8</b> Versión para lanzamiento Beta 6.6	<b>Thu Aug 7 2014</b>	<b>Steven Levine</b>
<b>Revisión 7.0-4</b> Resuelve #1102591 Se agrega procedimiento para configurar GFS2 en un clúster Pacemaker	<b>Thu Jul 17 2014</b>	<b>Steven Levine</b>
<b>Revisión 7.0-3</b> Resuelve #1035119 Se actualiza la tabla de indicadores Glock y agrega una sección una sección sobre estadísticas de Glock	<b>Wed Jul 16 2014</b>	<b>Steven Levine</b>
<b>Revisión 7.0-1</b> Primer borrador para lanzamiento 6.6	<b>Thu Jun 5 2014</b>	<b>Steven Levine</b>
<b>Revisión 6.0-6</b> Versión para lanzamiento GA 6.5	<b>Wed Nov 13 2013</b>	<b>Steven Levine</b>
<b>Revisión 6.0-5</b> Versión para lanzamiento Beta 6.5	<b>Fri Sep 27 2013</b>	<b>Steven Levine</b>
<b>Revisión 6.0-3</b> Resuelve #960841 Se aclara la falta de soporte para SELinux con sistemas de archivos GFS2.	<b>Fri Sep 27 2013</b>	<b>Steven Levine</b>
<b>Revisión 6.0-1</b> Se agrega nota sobre Samba y GFS2	<b>Fri Sep 06 2013</b>	<b>Steven Levine</b>
<b>Revisión 5.0-7</b> Versión para lanzamiento de disponibilidad general 6.4	<b>Mon Feb 18 2013</b>	<b>Steven Levine</b>
<b>Revisión 5.0-5</b> Versión para lanzamiento beta 6.4	<b>Mon Nov 26 2012</b>	<b>Steven Levine</b>
<b>Revisión 5.0-4</b> Resuelve #860324 Actualización del capítulo sobre consideraciones operacionales de GFS2 con pequeñas aclaraciones.  Resuelve #807057 Se adiciona nota que recomienda la consulta a un representante autorizado de Red Hat para verificar su configuración antes de la implementación.	<b>Tue Nov 13 2012</b>	<b>Steven Levine</b>
<b>Revisión 5.0-1</b>	<b>Mon Oct 15 2012</b>	<b>Steven Levine</b>



# ÍNDICE

## A

Adición de diarios al sistema de archivos, [Cómo añadir diarios a un sistema de archivos](#)

Administración de cuota, [Administración de cuotas GFS2 con el comando gfs2\\_quota](#)  
sincronización de cuotas, [Sincronización de cuotas con el comando gfs2\\_quota](#)

administración de cuota

cómo establecer cuotas, [Cómo establecer cuotas con el comando gfs2\\_quota](#)

administración de cuotas, [Administración de cuotas en GFS2](#), [Cómo establecer cuotas en modo impositivo o de contabilidad](#)

Administración de cuotas

cómo aplicar activación y desactivación de cuotas, [Cómo realizar activación y desactivación de cuotas](#)

cómo desplegar límites de cuotas, [Despliegue de límites de cuotas y uso con el comando gfs2\\_quota](#)

Cómo habilitar la contabilidad de cuotas, [Activación de contabilidad de cuotas](#)

cómo sincronizar cuotas, [Sincronización de cuotas con el comando quotasync](#)

Administración de GFS2, [Administración de GFS2](#)

Ajuste de rendimiento, [Ajuste de rendimiento con GFS2](#)

archivo debugfs, [Resolución de problemas de rendimiento de GFS2 con el volcado de cerrojo GFS2](#)

atime, configuración de actualizaciones, [Cómo configurar las actualizaciones atime](#)

montaje con noatime , [Montar con noatime](#)

montaje con relatime , [Montar con relatime](#)

audiencia, [Audiencia](#)

## B

Bloqueo de nodos, [Bloqueo de nodo GFS2](#)

Bloqueo de Posix, [Problemas con Bloqueo de Posix](#)

## C

colgado de sistema al desmontar, [Consideraciones especiales en el montaje de los sistemas de archivos GFS2](#)

comando fsck.gfs2, [Cómo reparar un sistema de archivos](#)

comando gfs2\_jadd, [Cómo añadir diarios a un sistema de archivos](#)

Comando gfs2\_quota, [Administración de cuotas GFS2 con el comando gfs2\\_quota](#)

Comando mkfs, [Cómo crear un sistema de archivos](#)

comando mount, [Cómo montar un sistema de archivos](#)

comando umount, [Cómo desmontar un sistema de archivos](#)

Comentarios

información de contacto para este manual, [¡Necesitamos sus comentarios!](#)

Cómo extender un sistema de archivos, [Cómo expandir un sistema de archivos](#)

configuración inicial, [Cómo iniciar](#)

configuración, antes de, [Antes de configurar GFS2](#)

configuración, inicial

tareas previas, [Tareas previas](#), [Tareas de configuración inicial](#)

Consideraciones de configuración, [Configuración de GFS2 consideraciones operacionales](#)

Creación de un sistema de archivos, [Cómo crear un sistema de archivos](#)

Cuotas de disco

activación, [Configuración de cuotas de disco](#)

quotacheck, ejecución, [Cómo crear archivos de base de datos de cuotas](#)

administración del

comando quotacheck, para verificar, [Cómo mantener cuotas exactas](#)

asignación por grupo, [Cómo asignar cuotas por grupo](#)

asignación por usuario, [Cómo asignar cuotas por usuario](#)

límite duro, [Cómo asignar cuotas por usuario](#)

Cuotas de discos

activación

creación de archivos de cuotas, [Cómo crear archivos de base de datos de cuotas](#)

## D

debugfs, [Puntos de trazado GFS2 y archivo glocks debugfs](#)

desmontaje de un sistema de archivos, [Cómo desmontar un sistema de archivos](#),

[Consideraciones especiales en el montaje de los sistemas de archivos GFS2](#)

desmontar, colgado de sistema, [Consideraciones especiales en el montaje de los sistemas de archivos GFS2](#)

diario de datos, [Diario de datos](#)

Disco de cuotas

administración de, [Cómo administrar cuotas de disco](#)

reporte, [Cómo administrar cuotas de disco](#)

límite blando, [Cómo asignar cuotas por usuario](#)

recursos adicionales, [Referencias](#)

## F

función de retiro, GFS2, [La función de retiro de GFS2](#)

funcionalidades, nuevas y cambiadas, [Funcionalidades nuevas y cambiadas](#)

## G

GFS2

administración, [Administración de GFS2](#)

Administración de cuota, [Administración de cuotas en GFS2](#), [Cómo establecer cuotas en modo impositivo o de contabilidad](#)

administración de cuota, [Administración de cuotas GFS2 con el comando gfs2\\_quota](#)

administración de cuotas

cómo aplicar activación y desactivación de cuotas, [Cómo realizar activación y desactivación de cuotas](#)

cómo desplegar límites de cuota, [Despliegue de límites de cuotas y uso con el comando gfs2\\_quota](#)

cómo establecer cuotas, [Cómo establecer cuotas con el comando gfs2\\_quota](#)

cómo habilitar la contabilidad de cuotas, [Activación de contabilidad de cuotas](#)

cómo sincronizar cuotas, [Sincronización de cuotas con el comando gfs2\\_quota](#)

administrar cuotas

cómo sincronizar cuotas, [Sincronización de cuotas con el comando quotasync](#)

atime, cómo configurar actualizaciones, [Cómo configurar las actualizaciones atime](#)

montaje con noatime , [Montar con noatime](#)

montaje con relatime , [Montar con relatime](#)

Consideraciones de configuración, [Configuración de GFS2 consideraciones operacionales](#)

función de retiro, [La función de retiro de GFS2](#)

Operación, [Configuración de GFS2 consideraciones operacionales](#)

`gfs2_grow` command, [Cómo expandir un sistema de archivos](#)

`glock`, [Puntos de trazado GFS2 y archivo glocks debugfs](#)

I

Indicadores de portador de `glock`, [Resolución de problemas de rendimiento de GFS2 con el volcado de cerrojo GFS2](#)

Indicadores de portadores de `glock`, [Portadores de Glock](#)

Indicadores Glock, [Resolución de problemas de rendimiento de GFS2 con el volcado de cerrojo GFS2](#), [Interfaz de Glock debugfs](#)

Introducción, [Introducción](#)

audiencia, [Audiencia](#)

M

`mkfs.gfs2` tabla de opciones de comando, [Opciones completas](#)

Montaje de un sistema de archivos, [Cómo montar un sistema de archivos](#), [Consideraciones especiales en el montaje de los sistemas de archivos GFS2](#)

montaje enlazado

orden de montaje, [Orden de montaje de enlace y montaje de sistema de archivos](#)

montajes enlazados, [Nombres de rutas dependientes del contexto y montajes enlazados](#)

## N

nombres de ruta, dependientes del contexto (CDPN), [Nombres de rutas dependientes del contexto y montajes enlazados](#)

Nombres de rutas dependientes de contexto (CDPN)

Conversión de GFS a GFS2, [Conversión de nombres de rutas dependientes de contexto](#)

## O

opción acl mount, [Cómo montar un sistema de archivos](#)

Opciones específicas de GFS2 para añadir tablas de diario, [Uso completo](#)

Opciones específicas de GFS2 para extender tabla de sistemas de archivos, [Uso completo](#)

## P

Parámetro ajustable quota\_quantum, [Sincronización de cuotas con el comando gfs2\\_quota](#)

Prefacio (ver introducción)

Puntos de trazado, [Puntos de trazado GFS2 y archivo glocks debugfs](#)

## Q

quota= mount option, [Cómo establecer cuotas con el comando gfs2\\_quota](#)

quotacheck , [Cómo crear archivos de base de datos de cuotas](#)

quotacheck comando

cómo revisar exactitud de la cuota, [Cómo mantener cuotas exactas](#)

quota\_quantum tunable parameter, [Sincronización de cuotas con el comando quotasync](#)

## R

rendimiento, ajuste, [Ajuste de rendimiento con GFS2](#)

reparación de un sistema de archivos, [Cómo reparar un sistema de archivos](#)

## S

Sistema de archivos

adición de diarios, [Cómo añadir diarios a un sistema de archivos](#)

administración de cuota, [Administración de cuotas GFS2 con el comando gfs2\\_quota](#)

administración de cuotas, [Administración de cuotas en GFS2, Cómo establecer cuotas en modo impositivo o de contabilidad](#)

activación de contabilidad de cuotas, [Activación de contabilidad de cuotas](#)

cómo aplicar activación y desactivación de cuotas, [Cómo realizar activación y desactivación de cuotas](#)

cómo desplegar límites de cuotas, [Despliegue de límites de cuotas y uso con el comando gfs2\\_quota](#)

cómo establecer cuotas, [Cómo establecer cuotas con el comando gfs2\\_quota](#)

cómo sincronizar cuotas, [Sincronización de cuotas con el comando quotasync](#) , [Sincronización de cuotas con el comando gfs2\\_quota](#)

atime, cómo configurar actualizaciones, [Cómo configurar las actualizaciones atime](#)  
montaje con noatime , [Montar con noatime](#)

Creación, [Cómo crear un sistema de archivos](#)

diarios de datos, [Diario de datos](#)

extensión, [Cómo expandir un sistema de archivos](#)

Montajes enlazados, [Nombres de rutas dependientes del contexto y montajes enlazados](#)

nombres de rutas dependientes del contexto (CDPN), [Nombres de rutas dependientes del contexto y montajes enlazados](#)

orden de montaje, [Orden de montaje de enlace y montaje de sistema de archivos](#)

reparación, [Cómo reparar un sistema de archivos](#)

suspensión de actividad, [Cómo suspender la actividad en un sistema de archivos](#)

sistema de archivos

atime, cómo configurar actualizaciones

montaje con relatime , [Montar con relatime](#)

desmontaje, [Cómo desmontar un sistema de archivos](#), [Consideraciones especiales en el montaje de los sistemas de archivos GFS2](#)

montaje, [Cómo montar un sistema de archivos](#), [Consideraciones especiales en el montaje de los sistemas de archivos GFS2](#)

Suspensión de actividad en un sistema de archivos, [Cómo suspender la actividad en un sistema de archivos](#)

**T**

tabla de montaje, [Uso completo](#)

tablas

mkfs.gfs2 opciones de comando, [Opciones completas](#)

Opciones de montaje, [Uso completo](#)

Opciones específicas de GFS2 para añadir diarios, [Uso completo](#)

Opciones específicas de GFS2 para expandir sistemas de archivos, [Uso completo](#)

Tamaño máximo de sistema de archivos GFS2, [Visión general de GFS2](#)

tareas iniciales

configuración, inicial, [Tareas de configuración inicial](#)

tareas previas

configuración, inicial, [Tareas previas](#)

Tipos de glock, [Resolución de problemas de rendimiento de GFS2 con el volcado de cerrojo GFS2](#)

Tipos de Glock, [Interfaz de Glock debugfs](#)

**V**

Visión general, [Visión general de GFS2](#)

funcionalidades, nuevas y cambiadas, [Funcionalidades nuevas y cambiadas](#)

**Vista general**

antes de la configuración, [Antes de configurar GFS2](#)