

红帽企业版 Linux 6

全局文件系统 2

红帽全局文件系统 2



红帽企业版 Linux 6 全局文件系统 2

红帽全局文件系统 2

版 7

Copyright © 2011 Red Hat, Inc. and others.

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

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

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

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

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

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

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

All other trademarks are the property of their respective owners.

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

本书提供了有关为红帽企业版 Linux 6 安装、配置和维护红帽 GFS2（红帽全局文件系统 2）的信息。

简介	v
1. 读者	v
2. 相关文档	v
3. 我们需要反馈意见!	v
4. 文档约定	vi
4.1. 排版约定	vi
4.2. 抬升式引用约定	vii
4.3. 备注及警告	vii
1. GFS2 概述	1
1.1. 新的和更改的功能	2
1.1.1. 红帽企业版 Linux 6.0 新的和更改的功能	2
1.1.2. 红帽企业版 Linux 6.1 新的和更改的功能	2
1.2. 设置 GFS2 前	2
1.3. GFS 和 GFS2 之间的差别	3
1.3.1. GFS2 命令名称	3
1.3.2. GFS 和 GFS2 之间的其它不同之处	4
1.3.3. GFS2 性能改进	5
1.4. GFS2 节点锁定	6
1.4.1. 使用 GFS2 执行调节	6
1.4.2. 使用 GFS2 锁定转储排除 GFS2 性能故障	7
2. 开始	11
2.1. 前提任务	11
2.2. 初始设定任务	11
2.3. 部署 GFS2 集群	12
3. 管理 GFS2	13
3.1. 生成文件系统	13
3.2. 挂载文件系统	16
3.3. 卸载文件系统	18
3.4. 挂载 GFS2 文件系统时的具体注意事项	19
3.5. GFS2 配额管理	19
3.5.1. 配置磁盘配额	20
3.5.2. 管理磁盘配额	23
3.5.3. 保持配额准确	23
3.5.4. 使用 quotasync 命令同步配额	23
3.5.5. 参考	25
3.6. 增大的文件系统	25
3.7. 在文件系统中添加日志	26
3.8. 数据日志	28
3.9. 配置 atime 更新	29
3.9.1. 使用 relatime 挂载	29
3.9.2. 使用 noatime 挂载	29
3.10. 在文件系统中挂起一个动作	30
3.11. 修复文件系统	31
3.12. 绑定挂载以及上下文关联路径名	32
3.13. 绑定挂载和文件系统挂载顺序	33
3.14. GFS2 撤回功能	35
4. 诊断并修正 GFS2 文件系统的问题	37
4.1. GFS2 文件系统出现性能缓慢	37
4.2. 通过 GFS2 设定 NFS	37
4.3. GFS2 文件系统挂起并需要在节点中重启	37
4.4. GFS2 文件系统挂起并需要重启所有节点	38
4.5. GFS2 文件系统不挂载新添加的集群节点	38

4.6. 空格代表在空文件系统中使用	38
A. 使用 gfs2_quota 命令执行 GFS2 配额管理	39
A.1. 使用 gfs2_quota 命令设定配额	39
A.2. 使用 gfs2_quota 命令显示配额限制和用量	40
A.3. 使用 gfs2_quota 命令同步配额	41
A.4. 启用/禁用配额强制	42
A.5. 启用配额计数	43
B. 将文件系统从 GFS 转换为 GFS2	45
C. 修订记录	47
索引	49

简介

本书提供有关配置和维护红帽 GFS2（红帽全局文件系统 2）的信息，GFS2 包含在弹性存储附加组件中。

1. 读者

本书主要面向熟悉以下活动的 Linux 系统管理员：

- 包括内核配置在内的 Linux 系统管理流程
- 安装和配置共享存储网络，比如光纤 SAN

2. 相关文档

有关使用红帽企业版 Linux 的详细信息请参考以下资源：

- 《安装指南》 — 记录有关安装红帽企业版 Linux 6 的信息。
- 《部署指南》 — 记录有关部署、配置及管理红帽企业版 Linux 6 的信息。
- 《存储管理指南》 — 提供如何有效管理红帽企业版 Linux 6 中的存储设备和文件系统的说明。

有关 Red Hat Enterprise Linux 6 中高可用性添加组件和弹性存储添加组件的详细信息请参考以下资源：

- 《高可用性附加组件概述》 — 提供高可用性附加组件的高级概述。
- 《集群管理》 — 提供有关安装、配置和管理高可用性附加组件的信息。
- 《逻辑卷管理器管理》 — 提供逻辑卷管理器（LVM）的描述，其中包括在集群环境中运行 LVM 的信息。
- 《设备映射器多路径》 — 提供有关使用红帽企业版 Linux 设备映射器多路径功能的信息。
- 《负载均衡管理》 — 提供使用负载均衡附加组件配置高性能系统和服务的信息，这是一组整合的软件组件，可在一组真实服务器间平衡 IP 负载时提供 Linux 虚拟服务器（LVS）。
- 《发行注记》 — 提供有关红帽产品当前发行本信息。

在红帽企业版 Linux 文档光盘和在线文档 <http://www.redhat.com/docs/> 中都含有高可用性附加组件文档和其它红帽文档的 HTML、PDF 和 RPM 版本。

3. 我们需要反馈意见！

如果您在这本手册中发现任何印刷错误，或者您对本手册有改进意见，我们非常乐于倾听！请在 Bugzilla <http://bugzilla.redhat.com/> 中根据产品 Red Hat Enterprise Linux 6 和组件 doc-Global_File_System_2 提交报告。当提交 bug 报告时，请确定包含了手册识别符：rh-gfs2(EN)-6 (2010-11-19T15:15)。

如果您有针对文档的建议，请尽量具体描述。如果您发现任何错误，请附带章节号以及上下文，以便我们查找。

4. 文档约定

本手册使用几个约定来突出某些用词和短语以及信息的某些片段。

在 PDF 版本以及纸版中，本手册使用在 [Liberation 字体](https://fedorahosted.org/liberation-fonts/)¹ 套件中选出的字体。如果您在您的系统中安装了 Liberation 字体套件，它还可用于 HTML 版本。如果没有安装，则会显示可替换的类似字体。请注意：红帽企业 Linux 5 以及其后的版本默认包含 Liberation 字体套件。

4.1. 排版约定

我们使用四种排版约定突出特定用词和短语。这些约定及其使用环境如下。

单行粗体

用来突出系统输入，其中包括 shell 命令、文件名以及路径。还可用来突出按键以及组合键。例如：

要看到文件您当前工作目录中文件 `my_next_bestselling_novel` 的内容，请在 shell 提示符后输入 `cat my_next_bestselling_novel` 命令并按 `Enter` 键执行该命令。

以上内容包括一个文件名，一个 shell 命令以及一个按键，它们都以固定粗体形式出现，且全部与上下文有所区别。

组合键可通过使用连字符连接组合键的每个部分来与按键区别。例如：

按 `Enter` 执行该命令。

按 `Ctrl+Alt+F2` 切换到第一个虚拟终端。`Ctrl+Alt+F1` 返回您的 X-Windows 会话。

第一段突出的是要按的特定按键。第二段突出了两个按键组合（每个组合都要同时按）。下。

如果讨论的是源码、等级名称、方法、功能、变量名称以及在段落中提到的返回的数值，那么都会以上述形式出现，即固定粗体。例如：

与文件相关的等级包括用于文件系统的 `filesystem`、用于文件的 `file` 以及用于目录的 `dir`。每个等级都有其自身相关的权限。

比例粗体

这是指在系统中遇到的文字或者短语，其中包括应用程序名称、对话框文本、标记的按钮、复选框以及单选按钮标签、菜单标题以及子菜单标题。例如：

在主菜单条中选择「系统」→「首选项」→「鼠标」启动 鼠标首选项。在「按钮」标签中点击「惯用左手鼠标」复选框并点击 关闭切换到主鼠标按钮从左向右（让鼠标适合左手使用）。

要在 `gedit` 文件中插入一个特殊字符，请在主菜单中选择「应用程序」→「附件」→「字符映射表」。下一步在 字符映射表菜单条中选择「搜索」→「查找」，在「搜索」字段输入字符名称并点击 下一个 按钮。您输入的字符会在「字符表」中突出出来。双击这个突出的字符将其放入「要复制的文本」字段，然后点击 复制 按钮。现在切换回您的文档并在 `gedit` 菜单条中选择「编辑」→「粘贴」。

以上文本包括应用程序名称、系统范围菜单名称及项目、应用程序特定菜单名称以及按钮和 GUI 界面中的文本，所有都以比例粗体出现并与上下文区别。

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

固定粗斜体 或者 比例粗斜体

无论固定粗体或者比例粗体，附加的斜体表示是可替换或者变量文本。斜体表示那些不直接输入的文本或者那些根据环境改变的文本。例如：

要使用 `ssh` 连接到远程机器，请在 `shell` 提示符后输入 `ssh username@domain.name`。如果远程机器是 `example.com` 且您在该其机器中的用户名为 `john`，请输入 `ssh john@example.com`。

`mount -o remount file-system` 命令会重新挂载命名的文件系统。例如：要重新挂载 `/home` 文件系统，则命令为 `mount -o remount /home`。

要查看目前安装的软件包版本，请使用 `rpm -q package` 命令。它会返回以下结果：
: package-version-release。

请注意以上文字中的粗斜体字 — `username`、`domain.name`、`file-system`、`package`、`version` 和 `release`。无论您输入文本或者运行一个命令，还是该系统显示的文本，每个字都是一个占位符。

不考虑工作中显示标题的标准用法，斜体表示第一次使用某个新且重要的用语。例如：

`Publican` 是一个 `DocBook` 发布系统。

4.2. 抬升式引用约定

终端输出和源代码列表要与周围文本明显分开。

将发送到终端的输出设定为 `Mono-spaced Roman` 并显示为：

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

源代码列表也设为 `Mono-spaced Roman`，但添加下面突出的语法：

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

import javax.naming.InitialContext;

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

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

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

4.3. 备注及警告

最后，我们使用三种视觉形式来突出那些可能被忽视的信息。



备注

备注是对手头任务的提示、捷径或者备选解决方法。忽略提示不会造成负面后果，但您可能会错过一个更省事的诀窍。



重要

重要框中的内容是那些容易错过的事情：配置更改只可用于当前会话，或者在应用更新前要重启的服务。忽略‘重要’框中的内容不会造成数据丢失但可能会让您抓狂。



警告

警告是不应被忽略的。忽略警告信息很可能导致数据丢失。

GFS2 概述

红帽 GFS2 文件系统包含在弹性存储附加组件中，是固有文件系统，直接和 Linux 内核文件系统界面（VFS 层）相接。当作为集群文件系统使用时，GFS2 采用分布式元数据和多个日志（multiple journal）。红帽只支持将 GFS2 文件系统作为在高可用性附加组件中的部署使用。

注意

虽然 GFS2 文件系统可部署为独立的系统或者作为集群配置的一部分，但在红帽企业版 Linux 6 发行本中，红帽不支持将 GFS2 作为单节点文件系统使用。红帽支持很多高性能单节点文件系统，这些系统是为单节点优化的，因此一般比集群文件系统的消耗要小。红帽建议在只需要将单一节点挂载到文件系统的情况下，首先使用这些文件系统，而不是 GFS2。

红帽将继续支持单一节点 GFS2 文件系统挂载集群文件系统快照（例如：用于备份）。

注意

红帽不支持使用 GFS2 部署超过 16 个节点的集群文件系统。

GFS2 是基于 64 位构架，理论上可提供 8 EB 文件系统。但是，目前支持的 64 位硬件的最大 GFS2 文件系统为 100 TB，为 32 位硬件支持的最大 GFS2 文件系统为 16 TB。如果您的系统要求更大的 GFS2 文件系统，请联络您的红帽服务代表。

当决定您的文件系统大小时，您应该考虑您的恢复需求。在一个很大的文件系统中运行 `fsck.gfs2` 命令需要很长时间且消耗大量内存。另外，在磁盘或者磁盘子系统失败事件中，恢复时间受您的备份介质速度的限制。有关 `fsck.gfs2` 所需内存量的详情请参考 [第 3.11 节 “修复文件系统”](#)。

当在集群中进行配置时，可使用高可用性附加组件配置和管理工具对红帽 GFS2 节点进行配置和管理。然后红帽 GFS2 会在集群的 GFS2 节点间提供数据共享，即在 GFS2 节点间单一、一致的文件系统名称查看。这样就允许不同节点中的进程可以共享 GFS2 文件，其方法与同一节点中的进程可共享本地文件系统中的文件相同，没有明显的差别。有关高可用性附加组件的详情请参考《配置和管理红帽集群》。

当在 LVM 之外使用 GFS2 文件系统时，红帽只支持在 CLVM 逻辑卷中创建的 GFS2 文件系统。CLVM 包含在弹性存储附加组件中，这是在集群范围内部署 LVM，由在集群中管理 LVM 逻辑卷的 CLVM 守护进程 `clvmd` 启用。该守护进程可使 LVM2 跨集群管理逻辑卷，允许集群中的所有节点共享该逻辑卷。有关 LVM 卷管理器的详情请参考《LVM 管理员指南》。

`gfs2.ko` 内核模块部署 GFS2 文件系统，并在 GFS2 集群节点中载入。

备注

当您为 GFS2 文件系统配置为集群文件系统时，您必须确定该集群中的所有节点都可访问共享存储。在不对称集群配置中，不支持这样的情况，即有些节点可访问共享存储而另一些不可以。这并不需要所有节点都确实挂载到 GFS2 文件系统中。

本章提供一些基本的简略资料作为参考以便帮助您了解 GFS2，其中包括以下部分：

- 第 1.1 节 “新的和更改的功能”
- 第 1.2 节 “设置 GFS2 前”
- 第 1.3 节 “GFS 和 GFS2 之间的差别”
- 第 1.4 节 “GFS2 节点锁定”

1.1. 新的和更改的功能

本小节列出了红帽企业版 Linux 6 初始发行本以及后续版本中所包含的 GFS2 文件系统以及 GFS2 文档的新功能和更改的功能。

1.1.1. 红帽企业版 Linux 6.0 新的和更改的功能

红帽企业版 Linux 6.0 包括以下文档和功能更新和更改。

- 在红帽企业版 Linux 6 发行本中，红帽不支持将 GFS2 作为单节点文件系统使用。
- 在红帽企业版 Linux 6 发行本中，GFS2 文件系统中的 `gfs2_convert` 命令较之在 GFS 中有所提高。有关这个命令的详情请参考 [附录 B, 将文件系统从 GFS 转换为 GFS2](#)。
- 红帽企业版 Linux 6 支持 `discard`、`nodiscard`、`barrier`、`nobarrier`、`quota_quantum`、`statfs_quantum` 和 `statfs_percent` 挂载选项。有关挂载 GFS2 文件系统的详情请参考 [第 3.2 节 “挂载文件系统”](#)。
- 红帽企业版 Linux 6 版本的这个文档包含一个新的部分 [第 1.4 节 “GFS2 节点锁定”](#)。这部分描述了一些 GFS2 文件系统的内置功能。

1.1.2. 红帽企业版 Linux 6.1 新的和更改的功能

红帽企业版 Linux 6.1 包括以下文档和功能更新和更改。

- 从红帽企业版 Linux 6.1 发行本开始，GFS2 支持标准 Linux 配额工具。GFS2 配额管理文档请参考 [第 3.5 节 “GFS2 配额管理”](#)。

在红帽企业版 Linux 之前的发行本中，GFS2 需要使用 `gfs2_quota` 命令管理配额。现在在 [附录 A, 使用 gfs2_quota 命令执行 GFS2 配额管理](#) 中记录了 `gfs2_quota` 命令。

- 这个文档现在包含一个新章节，[第 4 章 诊断并修正 GFS2 文件系统的问题](#)。
- 对整篇文档进行了技术修正和澄清。

1.2. 设置 GFS2 前

在您安装和设定 GFS2 前，请记录您的 GFS2 文件系统的以下重要功能：

GFS2 节点

决定集群中的哪个节点将挂载 GFS2 文件系统。

文件系统数

决定最初创建多少个 GFS2 文件系统（以后可以添加更多的文件系统）。

文件系统名称

为每个文件系统取一个唯一名称。该名称必须与集群中的所有 lock_dlm 文件系统不同。每个文件系统名称都要求使用参数变量格式。例如：在本书中的一些示例过程中使用 mydata1 和 mydata2。

日志

决定 GFS2 文件系统的日志数。每个挂载到 GFS2 文件系统的节点都需要一个日志。GFS2 允许您以后动态添加日志作为挂载文件系统的附加服务器。有关在 GFS2 文件系统中添加日志的详情请参考第 3.7 节“在文件系统中添加日志”。

存储设备和分区

决定在文件系统中用来创建逻辑卷（使用 CLVM）的存储设备和分区。



备注

当同时在一个目录中进行多个生成和删除操作时，您可能会看到 GFS2 的性能下降。如果这导致您系统性能出现问题，您应该尽量将某个节点中的文件生成和删除动作定位到具体目录。

1.3. GFS 和 GFS2 之间的差别

这部分列出了相比 GFS，GFS2 提供的改进和更改。

从 GFS 迁移到 GFS2 需要您使用 gfs2_convert 程序将 GFS 文件系统转换为 GFS2。有关 gfs2_convert 程序的详情请参考附录 B，将文件系统从 GFS 转换为 GFS2。

1.3.1. GFS2 命令名称

通常 GFS2 和 GFS 的功能是一致的，只是在文件系统命令的名称中要将 GFS 改为 GFS2。表 1.1 “GFS 和 GFS2 命令”中显示了功能等同的 GFS 和 GFS2 命令。

表 1.1. GFS 和 GFS2 命令

GFS 命令	GFS2 命令	描述
mount	mount	挂载文件系统。系统可以确定文件系统的类型是 GFS 还是 GFS2。有关 GFS2 挂载选项的详情请参考 gfs2_mount(8) man page。
umount	umount	卸载文件系统
fscck gfs_fscck	fscck fscck.gfs2	检查并修复卸载的文件系统。
gfs_grow	gfs2_grow	增大挂载的文件系统。
gfs_jadd	gfs2_jadd	在某个挂载的文件系统中添加日志
gfs_mkfs mkfs -t gfs	mkfs.gfs2 mkfs -t gfs2	在存储设备中创建文件系统。
gfs_quota	gfs2_quota	在挂载的文件系统中管理配额。
gfs_tool	gfs2_tool	配置、调试文件系统或者收集有关文件系统的信息。
gfs_edit	gfs2_edit	显示、输出或者编辑文件系统内部结构。gfs2_edit 命令可用于 GFS 文件系统，也可用于 GFS2 文件系统。
gfs_tool setflag jdata/ inherit_jdata	chattr +j (preferred)	在文件或者目录中启用日志功能。

GFS 命令	GFS2 命令	描述
setfacl/getfacl	setfacl/getfacl	为文件或者目录设置或者获得文件访问控制。
setfattr/ getfattr	setfattr/ getfattr	设置或者获得文件的扩展属性。

GFS2 文件系统命令支持选项的完整列表请参考那些命令的 man page。

1.3.2. GFS 和 GFS2 之间的其它不同之处

这部分总结了在 第 1.3.1 节 “GFS2 命令名称” 中没有论述的 GFS 和 GFS2 管理中的其它不同之处。

上下文关联路径名

GFS2 文件系统不提供对上下文关联路径名的支持，上下文关联路径名允许您生成指向各种目的地文件或者目录的符号链接。在 GFS2 中，您可以使用 mount 命令的 bind 选项来实现这个功能。有关在 GFS2 中管理路径名的详情请参考 第 3.12 节 “绑定挂载以及上下文关联路径名”。

gfs2.ko 模块

使用 GFS 文件系统的内核模块是 gfs.ko。使用 GFS2 文件系统的内核模块是 gfs2.ko。

在 GFS2 中启用配额强制

在 GFS2 文件系统中，默认情况是禁用配额强制，且必须单独启用。有关启用和禁用配额强制的详情请参考 第 3.5 节 “GFS2 配额管理”。

数据日志

GFS2 文件系统支持使用 chattr 命令设定和清除某个文件或者目录的 j 标签。在某个文件中设定 +j 标签会启用该文件的数据日志记录。在某个目录中设定 +j 标签意味着“继承 jdata”，即对之后所有在那个目录中生成的文件或者目录进行日志记录操作。使用 chattr 命令是在某个文件中启用和禁用数据日志记录的首选方法。

动态添加日志

在 GFS2 文件系统中，日志是存在于文件系统之外的内嵌元数据，这就需要在添加日志前扩展包含该文件系统的逻辑卷大小。在 GFS2 文件系统中，日志是纯文本文件（虽然是隐藏的）。这意味着对于 GFS2 文件系统来说，只要在该文件系统中可放置附加日志的空间，即可动态添加日志将其作为附加服务器挂载到文件系统中。有关在 GFS2 文件系统中添加日志的详情请参考 第 3.7 节 “在文件系统中添加日志”。

删除 atime_quantum 参数

GFS2 文件系统不支持 atime_quantum 可调节参数，GFS 文件系统可使用该参数指定 atime 更新的频率。在 GFS2 中支持 relatime 和 noatime 挂载选项。建议您使用 relatime 挂载选项获得与在 GFS 中使用 atime_quantum 参数时得到的类似行为。

mount 命令的 data= 选项

在挂载 GFS2 文件系统时，您可以指定 mount 命令的 data=ordered 或者 data=writeback 选项。当设定 data=ordered 时，传输修改的用户数据会在传输被提交到磁盘前被冲入磁盘。这样可以让用户无法在崩溃后的文件中看到未初始化的块。设定 data=writeback 时，用户数据会在磁盘有数据后的任何时候被写入磁盘。这不会提供 ordered 模式可提供的一致性保障，但对某些工作负载来说可稍微加快一些速度。默认设置是 ordered 模式。

gfs2_tool 命令

gfs2_tool 为 GFS2 支持的选项组与 gfs_tool 命令为 GFS 支持的选项组不同:

- gfs2_tool 命令支持 journals 参数, 它可输出当前配置日志有关信息, 其中包括文件系统包含的日志数。
- gfs2_tool 命令不支持 counters 标签, 而 gfs_tool 命令可使用该标签显示 GFS 统计。
- gfs2_tool 命令不支持 inherit_jdata 标签。要将某个命令标记为 "inherit jdata", 您可以在目录中设定 jdata 标签, 或者使用 chattr 命令在目录中设定 +j 标签。使用 chattr 命令是在文件中启用和禁用数据日志记录的首选方法。

gfs2_edit 命令

gfs2_edit 命令为 GFS2 支持的选项组与 gfs_edit 命令为 GFS 支持的选项组不同。有关该命令每个版本具体支持的选项, 请参考 gfs2_edit 和 gfs_edit man page。

1.3.3. GFS2 性能改进

GFS2 文件系统的很多功能与 GFS 文件系统在界面上没有什么不同, 但文件系统性能会有提高。

GFS2 文件系统在以下方面提供改进的文件系统性能:

- 在大量使用单一目录时有较好的性能。
- 更快的同步 I/O 操作
- 更快的缓存读取 (无锁定消耗)
- 对预先分配的文件有更快的直接 I/O (提供合理的较大 I/O 值, 比如 4M 大的块)
- 普遍更快的 I/O 操作
- 执行 df 命令的速度更快, 因为 statfs 调用的速度更快。
- 与 GFS 相比, 我们改进了 atime 模式以减少 atime 生成的写入 I/O 操作数量。

GFS2 文件系统在以下方面提供更广泛和主流的支持:

- GFS2 是上游内核 (整合到 2.6.19) 的一部分。
- GFS2 支持以下特性:
 - SELinux 扩展的属性。
 - lsattr() 和 chattr() 属性是通过标准 ioctl() 调用设置的。
 - 纳秒时间戳

GFS2 为文件系统的内在效率提供以下改进。

- GFS2 使用更少的内核内存
- GFS2 需要非元数据生成数
分配 GFS2 元数据不需要读取。多日志中的元数据块副本是由从锁定释放前的日志中调用的块管理的。
- GFS2 的日志管理程序更为简单, 它对未链接的内节点或者配额修改一无所知。

- `gfs2_grow` 和 `gfs2_jadd` 命令使用锁定来防止多个事件同时运行。
- 为类似 `creat()` 和 `mkdir()` 的调用简化了 ACL 编码。
- 在没有重新挂载日志的情况下恢复未连接的内节点以及配额和 `statfs` 的更改。

1.4. GFS2 节点锁定

要获得最佳 GFS2 文件系统性能，则需要理解其操作的基本原理。单节点文件系统与缓存一同使用，其目的是在频繁使用请求的数据时可消除磁盘访问延迟。Linux 页面缓存（以及缓冲缓存）提供这个缓存功能。

使用 GFS2，每个节点都可有其自身的页面缓存，该缓存中包含 on-disk 数据的一部分。GFS2 使用 `glocks`（发音为 gee-locks）锁定机制维护节点间缓存的完整性。`glock` 子系统提供缓存管理功能，该功能使用分布式锁管理器（DLM）部署作为基础沟通层。

`glocks` 在每个内节点中为缓存提供保护，因此在每个内节点中都有一个锁定用来控制缓冲层。如果为那个 `glock` 赋予共享模式（DLM 锁定模式：PR），那么那个 `glock` 保护下的数据可同时被一个或者多个节点缓存，这样多有节点就都有到该数据的本地访问。

如果为 `glock` 赋予专用模式（DLM 锁定模式：EX），那么只有一个节点可缓存那个 `glock` 保护的数据。所有修改数据的操作（例如 `write` 系统调用）都使用这个模式。

如果另一个节点请求 `glock`，但无法立刻获得，那么 DLM 会向该节点发送一条信息，或者向目前使用 `glock` 并妨碍新的请求的节点发送信息，要求它们释放其锁定。释放 `glock`（大多数文件系统操作标准）需要很长时间。释放共享 `glock` 只需要使该缓存无效，相对缓冲的数据来说速度较快。

释放专用 `glock` 需要 `log flush`，并向磁盘写回所有更改的数据，之后要使每个共享的 `glock` 失效。

单一节点文件系统与 GFS2 之间的区别在于单一节点文件系统只有一个缓存，而 GFS2 在每个节点中都有独立的缓存。在这两种情况下，对缓冲数据访问的延迟程度类似，但如果另一个节点之前缓冲了同样的数据，GFS2 对非缓冲数据访问的延迟要大得多。



注意

根据采用 GFS2 缓冲的方法，在出现以下任意情况之一时都会获得最佳性能：

- 在所有节点中都使用只读方式使用内节点
- 只在单一节点中写入或者修改内节点

请注意：在创建和删除文件的过程中插入和删除目录条目也算在内，因为要写入到目录内节点。

也可能不遵守这个规则，但并不常发生。过度忽略这个规则会对性能有严重影响。

如果您在 GFS2 中使用 `read/write` 映射 `mmap()` 某个文件，但只读取它，那么这只计为读取。而在 GFS 中会将其计为写入，因此 GFS2 使用 `mmap()` I/O 时更灵活。

如果您没有设定 `noatime mount` 参数，那么读取也会导致写入来更新文件时间戳。文件建议所有 GFS2 用户应该使用 `noatime` 挂载，除非对 `atime` 有具体要求。

1.4.1. 使用 GFS2 执行调节

通常可以更改出问题的程序保存其数据的方法，这样可获得可观的性能优势。

典型的问题程序例子有电子邮件服务器。通常会制定包含每个用户的文件的 spool 目录 (mbox)，或者为每个用户创建一个目录，其中有包含所有信息的文件 (maildir)。当有来自 IMAP 的请求时，理想的情况是为每个用户赋予一个对特定节点的亲和性。那么将会使用那个节点中的缓存处理他们查看和删除电子邮件信息的请求。显然，如果那个节点失败，那么可在不同的节点中重启该会话。

当邮件通过 SMTP 送达时，那么还是将独立节点设定为默认将某个用户的邮件传递到特定节点。如果没有启动默认节点，那么会将该信息由接收节点直接保存到用户的邮件 spool 中。同样，这个设计的目的在于在通常情况下只在一个节点中保留缓冲文件的具体设置，但在节点失败的情况下允许直接访问。

这个设置可使用最佳 GFS2 页面缓存，还可让程序 (imap 或者 smtp) 清楚地了解失败情况。

备份通常是另一个令人纠结的领域。如果可能最好直接从节点备份每个节点的工作组件，这样可缓存具体的内节点组。如果您有可定期运行的备份脚本，且与在 GFS2 中运行的应用程序反应时间完全一致，那么很有可能集群无法最有效地使用页面缓存。

显然，如果您 (幸运地) 可以停止该程序以便执行备份，那么这就不是个问题。否则，如果只在一个节点中执行备份，那么当它完成后，文件系统的大部分将被缓存在该节点中，这样就对以后从其它节点访问有一定的性能损失。使用下面的命令在备份结束后放弃一些备份节点中的 VFS 页面缓存可将此影响缓解到一定程度：

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

但这并不是一个好的解决方案，如同要保证在每个节点中的工作组要么是共享的，大多数为集群内只读，或者是从单一节点的大量访问。

1.4.2. 使用 GFS2 锁定转储排除 GFS2 性能故障

如果由于浪费 GFS2 缓冲而造成集群性能问题，您会看到大量和增长的 I/O 等待时间。您可以使用 GFS2 的锁定转储信息确定造成此问题的原因。

GFS2 锁定转储信息可从 debugfs 文件中获得，您可以根据路径名找到该文件，假设 debugfs 是挂载在 /sys/kernel/debug/ 中：

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

文件的内容是一些行。每行使用 G: 开始，代表一个 glock。接下来的行使用一个空格缩进，代表文件中最新与 glock 关联的信息项目。

当程序出现问题时，使用 debugfs 文件的最佳方法是使用 cat 命令获得该文件内容副本（如果您的 RAM 较大，且有很多缓冲的内节点，则需要较长时间），以后再查看得到的数据。



窍门

复制两份 debugfs 文件会很有帮助，两个副本的间隔可在几秒甚至一分钟左右。比较与同一 glock 号关联的两个追踪信息，您可以了解负载是否在增长（就是说只是速度慢了）或者它在哪里卡住了（这通常是由 bug 造成，您应该立即向红帽支持提交报告）。

debugfs 文件中以 H: (拥有者) 开始的行代表赋予锁定请求或者正在等待赋予锁定请求。Holders 行中的标签字段 f: 表示是哪一个：‘w’代表等待的请求；‘H’代表授予的请求。有大量等待的 glocks 可能正处于争论阶段。

表 1.2 “Glock 标签” 显示不同 glock 标签的含义，表 1.3 “Glock 拥有者标签” 显示不同 glock 拥有者标签的含义。

表 1.2. Glock 标签

标签	名称	含义
d	等待降级	递延（远程）降级请求
D	降级	降级请求（本地或者远程）
f	清除日志	释放这个 glock 前需要提交该日志
F	冻结	回复忽略的远程节点 -- 恢复进行中。
i	使进程无效	这个 glock 下无效页面的进程中
I	启动	设定何时将 DLM 锁定与这个 glock 关联
l	锁定的	这个 glock 处于更改状态中
p	降级中	该 glock 正在响应降级请求
r	回复等待	从远程节点中接收的回复正在等待过程中
y	脏数据	释放这个 glock 前要刷新到磁盘中的数据

表 1.3. Glock 拥有者标签

标签	名称	含义
a	Async	不等待 glock 结果（以后轮询结果）
A	任意	接受任意兼容锁定模式
c	没有缓存	取消锁定时立即降级 DLM 锁定
e	没有过期	忽略随后的锁定取消请求
E	准确	必须有准确的锁定模式
F	第一	设定赋予这个锁定的第一个拥有者
H	拥有者	表示赋予请求的锁定
p	优先权	在队列头入队的拥有者
t	尝试	"try" 锁定
T	Try lcb	发送回呼的 "try" 锁定
W	等待	等待请求完成的设置

确定造成问题的 glock 后，下一步是要找到关联的内节点。glock 号 (G: 行中的 n:) 指的就是这个，其格式为 type/number，如果 type 是 2，那么 glock 就是一个内节点 glock，且 number 就是内节点号。要追踪内节点，您可以运行 `find -inum number`，其中 number 是将 glock 文件中的十六进制格式转换为十进制格式的内节点号。



警告

如果您在有锁定冲突的文件系统中运行 `find`，事情可能会变得更糟糕。当您查找冲突的内节点时，最好在运行 `find` 前停止该程序。

表 1.4 “Glock 类型” 显示不同 glock 类型含义。

表 1.4. Glock 类型

类型号	锁定类型	使用
1	Trans	事务锁定
2	Inode	内节点元数据和数据
3	Rgrp	员组群元数据
4	Meta	超级块
5	Iopen	内节点的最近探测
6	Flock	flock(2) 系统调用
8	Quota	配额操作
9	Journal	日志互斥

如果识别的 `glock` 是不同的类型，那么最可能是类型 3: (资源组)。如果您在常规负载时看到大量进程正在等待其它 `glock` 类型，请向红帽支持提交报告。

如果您看到在资源组锁定中有大量等待的请求，那么可能有很多原因。其中之一是在文件系统中相对于资源组有大量的内节点。另一个原因就是该文件系统可能接近饱和（按平均计算，需要较长的搜索）。在这两种情况下可通过添加更多存储以及使用 `gfs2_grow` 命令扩大该文件系统进行改善。

开始

本章论述了初始设定 GFS2 的步骤，其中包括以下部分：

- 第 2.1 节 “前提任务”
- 第 2.2 节 “初始设定任务”
- 第 2.3 节 “部署 GFS2 集群”

2.1. 前提任务

在设置红帽 GFS2 前，请确定您记录了 GFS2 节点的关键特征（请参考第 1.2 节 “设置 GFS2 前”）。同时还请确定 GFS2 节点的时钟是同步的。建议您使用红帽企业版 Linux 发行本提供的网络时间协议（NTP）软件。



备注

GFS2 节点中的系统时钟间的差别必须在几分钟之内以便防止内节点时间戳更新。不必要的内节点时间戳更新会严重影响集群的性能。

2.2. 初始设定任务

初始 GFS2 设定包含以下任务：

1. 设定逻辑卷。
2. 生成 GFS2 文件系统。
3. 挂载文件系统。

开始按照以下步骤设定 GFS2。

1. 使用 LVM 为每个红帽 GFS2 文件系统生成逻辑卷。



备注

您可以使用红帽集群套件中的 `init.d` 脚本自动激活和失活逻辑卷。有关 `init.d` 脚本的详情请参考《配置和管理红帽集群》。

2. 在第一步生成的逻辑卷中创建 GFS2 文件系统。为每个文件系统选择唯一的名称。有关创建 GFS2 文件系统的详情请参考第 3.1 节 “生成文件系统”。

您可以使用以下格式之一创建一个集群 GFS2 文件系统：

```
mkfs.gfs2 -p lock_d1m -t ClusterName:FSName -j NumberJournals BlockDevice
```

```
mkfs -t gfs2 -p lock_d1m -t LockTableName -j NumberJournals BlockDevice
```

有关创建 GFS2 文件系统的详情请参考 [第 3.1 节 “生成文件系统”](#)。

3. 在每个节点中挂载 GFS2 文件系统。有关挂载 GFS2 文件系统的详情请参考 [第 3.2 节 “挂载文件系统”](#)。

命令用法:

```
mount BlockDevice MountPoint
```

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

-o acl 选项允许操作文件 ACL。如果挂载某个文件系统是没有使用 -o acl 挂载选项，用户可以查看 ACL（使用 getfacl 命令），但不可以设定它们（使用 setfacl 命令）。



备注

您可以使用红帽高可用性附加组件中的 `init.d` 脚本自动挂载和卸载 GFS2 文件系统。

2.3. 部署 GFS2 集群

部署集群文件系统不是简单的替换单节点部署。我们建议您预留约 8-12 周的时间测试新的安装，以便测试该系统并保证其在要求的性能水平正常工作。在此期间可解决所有性能和功能问题，且所有疑问都可提交红帽支持团队。我们还建议考虑部署集群的客户在部署前请红帽支持团队审核其配置，这样可以避免之后可能存在的支持问题。

管理 GFS2

本章论述了管理 GFS2 的任务和命令，由以下部分组成：

- 第 3.1 节 “生成文件系统”
- 第 3.2 节 “挂载文件系统”
- 第 3.3 节 “卸载文件系统”
- 第 3.5 节 “GFS2 配额管理”
- 第 3.6 节 “增大的文件系统”
- 第 3.7 节 “在文件系统中添加日志”
- 第 3.8 节 “数据日志”
- 第 3.9 节 “配置 atime 更新”
- 第 3.10 节 “在文件系统中挂起一个动作”
- 第 3.11 节 “修复文件系统”
- 第 3.12 节 “绑定挂载以及上下文关联路径名”
- 第 3.13 节 “绑定挂载和文件系统挂载顺序”
- 第 3.14 节 “GFS2 撤回功能”

3.1. 生成文件系统

您可使用 `mkfs.gfs2` 命令创建 GFS2 文件系统。您还可以使用指定了 `-t gfs2` 选项的 `mkfs` 命令。文件系统是在活跃的 LVM 卷中创建的。运行 `mkfs.gfs2` 命令时需要以下信息：

- 锁定协议/模块名称（集群的锁定协议为 `lock_dlm`）
- 集群名称（当作为集群配置的一部分运行时）
- 日志数目（每个可能挂载文件系统的节点都需要一个日志）

当创建 GFS2 文件系统时，您可以直接使用 `mkfs.gfs2`，或者使用带 `-t` 参数的 `mkfs` 命令，并使用 `gfs2` 文件系统选项将文件系统类型指定为 `gfs2`。



备注

当您使用 `mkfs.gfs2` 命令创建 GFS2 文件系统时，您不能缩小该文件系统。但您可以使用 `gfs2_grow` 命令增大现有文件系统的大小，如第 3.6 节 “增大的文件系统” 所述。

用法

当创建集群的 GFS2 文件系统时，您可以使用以下任意格式之一：

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

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

您可以使用以下任意格式之一创建本地 GFS2 文件系统:



注意

在红帽企业版 Linux 6 发行本中，红帽不支持将 GFS2 作为单节点文件系统使用。

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

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



警告

请确定您非常熟悉 LockProtoName 和 LockTableName 参数的使用。不正确的 LockProtoName 和 LockTableName 参数使用可能导致文件系统或者锁定空间崩溃。

LockProtoName

指定要使用的锁定协议名称，集群的锁定协议为 lock_dlm。

LockTableName

这个参数是用来指定集群配置中的 GFS2 文件系统。它有两部分，用冒号隔开（没有空格）如下：
ClusterName:FSName

- ClusterName，用来创建 GFS2 文件系统的集群名称。
- FSName，文件系统名称，长度可在 1-16 个字符之间。该名称必须与集群中所有 lock_dlm 文件系统以及每个本地节点中的所有文件系统（lock_dlm 和 lock_nolock）不同。

Number

指定由 mkfs.gfs2 命令生成的日志数目。每个要挂载文件系统的节点都需要一个日志。对于 GFS2 文件系统来说，以后可以添加更多的日志而不会增大文件系统，如 [第 3.7 节 “在文件系统中添加日志”](#) 所述。

BlockDevice

指定逻辑卷或者物理卷。

示例

在这些示例中，lock_dlm 是文件系统使用的锁定协议，因为这是一个集群的文件系统。集群名称为 alpha，文件系统名为 mydata1。文件系统包含八个日志，日志是在 /dev/vg01/lvo10 中生成的。

```
mkfs.gfs2 -p lock_dlm -t alpha:mydata1 -j 8 /dev/vg01/lvo10
```

```
mkfs -t gfs2 -p lock_dlm -t alpha:mydata1 -j 8 /dev/vg01/lvo10
```

在这些示例中，生成了第二个 lock_dlm 文件系统，它可用于集群 alpha。文件系统名为 mydata2。文件系统包含八个日志，日志是在 /dev/vg01/lvo11 中生成的。

```
mkfs.gfs2 -p lock_dlm -t alpha:mydata2 -j 8 /dev/vg01/lvo11
```

```
mkfs -t gfs2 -p lock_dlm -t alpha:mydata2 -j 8 /dev/vg01/lvo11
```

完整选项

表 3.1 “命令选项: mkfs.gfs2” 描述 mkfs.gfs2 命令选项（标签和参数）。

表 3.1. 命令选项: mkfs.gfs2

标签	参数	描述
-c	Megabytes	以 Megabytes 为单位设定每个日志配额更改文件的初始大小。
-D		启用调试输出。
-h		帮助，显示可用选项。
-J	MegaBytes	以 MB 为单位指定日志大小。默认日志大小为 128MB，最小为 8MB。较大的日志可提高性能，但会比较小的日志占用更多的内存。
-j	Number	指定由 mkfs.gfs2 命令生成的日志数目。挂载文件系统的每个节点都需要一个日志。如果没有指定该选项，则会生成一个日志。对于 GFS2 文件系统，您可以后添加附加日志而不会增大文件系统。
-O		防止 mkfs.gfs2 命令在写入文件系统前进行确认。
-p	LockProtoName	指定要使用的锁定协议名称，可识别的锁定协议包括： lock_dlm — 标准锁定模式，用于集群的文件系统。 lock_nolock — 当 GFS2 作为本地文件系统作用时使用（只有一个节点）。
-q		静默，什么都不显示。
-r	MegaBytes	以 MB 为单位指定源组群大小，最小源组群值为 32MB，最大源组群值为 2048MB。在大型文件系统中源组群越大性能越高。如果没有指定这个信息，mkfs.gfs2 会根据文件系统大小选择源组群大小：中等大小的文件系统的源组群为 256MB，大一点的文件系统会有较大的源组群以获得更好的性能。
-t	LockTableName	在您使用 lock_dlm 协议时用来指定锁定表格字段的唯一识别程序，lock_nolock 协议不使用这个参数。这个参数有两个部分，用冒号隔开（没有空格）如下： ClusterName:FSName。 ClusterName 是用来创建 GFS2 文件系统的集群名称，只有集群成员有使用此文件系统的权限。集群名称

标签	参数	描述
		可使用 Cluster Configuration Tool 在文件 <code>/etc/cluster/cluster.conf</code> 中设定，并在红帽集群套件集群管理 GUI 的 Cluster Status Tool 中显示。 FSName，文件系统名称，长度可在 1-16 个字符之间，且必须不同于集群中的其它文件系统名。
-u	MegaBytes	指定每个日志未链接标签文件的初始大小。
-V		显示命令版本信息。

3.2. 挂载文件系统

在您挂载 GFS2 文件系统前，该文件系统必须存在（请参考第 3.1 节“生成文件系统”），文件系统所在的卷必须是活跃的，且必须启动了集群和锁定系统支持（请参考《配置和管理红帽集群》）。达到这些要求后，您就可以将这个 GFS2 文件系统挂载到任意 Linux 文件系统中。

注意

在没有启动集群管理器（cman）时尝试挂载 GFS2 文件系统会产生以下出错信息：

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

要控制文件 ACL，您必须使用 `-o acl` 挂载选项挂载文件系统。如果挂载文件系统时没有使用 `-o acl` 选项，用户可以查看 ACL（使用 `getfacl`），但不能对其进行设置（使用 `setfacl`）。

用法

不使用 ACL 控制挂载

```
mount BlockDevice MountPoint
```

使用 ACL 控制挂载

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

`-o acl`

允许控制文件 ACL 的具体 GFS2 选项。

BlockDevice

指定 GFS2 文件系统所在的块设备。

MountPoint

指定要挂载 GFS2 文件系统的目录。

示例

在这个示例中，位于 `/dev/vg01/lvo10` 的 GFS2 文件系统被挂载到 `/mygfs2` 目录中。

```
mount /dev/vg01/lvo10 /mygfs2
```

完整用法

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

`-o option` 参数包含 GFS2 具体选项（请参考表 3.2 “GFS2 特定挂载选项”）或者可接受的标准 Linux `mount -o` 选项，或者两者之和。多个 `option` 参数可使用逗号分开，没有空格。



备注

`mount` 命令是 Linux 系统命令。除了使用这部分论述的 GFS2 具体选项，您还可以使用其它标准 `mount` 命令选项（例如：`-r`）。有关其它 Linux `mount` 命令选项请参考 `mount man page`。

表 3.2 “GFS2 特定挂载选项” 描述在挂载时可传递给 GFS2 的 GFS2 特定 `-o option` 选项值。



注意

这个表格包含了只用于本地文件系统选项的描述。但请注意：在红帽企业版 Linux 6 发行本中红帽不支持将 GFS2 作为单节点文件系统使用。红帽将继续在挂载集群文件系统快照时支持单节点 GFS2 文件系统（例如：用于备份）。

表 3.2. GFS2 特定挂载选项

选项	描述
<code>acl</code>	允许控制文件 ACL。如果挂载文件系统时没有使用 <code>acl</code> 挂载选项，那么用户可以查看 ACL（使用 <code>getfacl</code> ），但不能对其进行设置（使用 <code>setfacl</code> ）。
<code>data=[ordered writeback]</code>	当设置 <code>data=ordered</code> 时，事务中修改的用户数据会将该事物递交到磁盘前被冲入磁盘。这样可以让用户无法在崩溃后的文件中看到未初始化的块。设定 <code>data=writeback</code> 时，用户数据会在磁盘有数据后的任何时候被写入磁盘。这样无法提供 <code>ordered</code> 模式可提供的一致性保障，但可稍微提高一些工作负载的速度。默认设置为 <code>ordered</code> 模式。
<code>ignore_local_fs</code> 警告：在共享 GFS2 文件系统时不应该使用这个选项。	强制 GFS2 将文件系统视为多主机文件系统。默认情况是使用 <code>lock_nolock</code> 自动打开 <code>locallocks</code> 标签。

选项	描述
localflocks 警告：在共享 GFS2 文件系统时不应该使用这个选项。	告知 GFS2 让 VFS（虚拟文件系统）层完成所有 flock 和 fcntl 操作。lock_nolock 可自动打开 localflocks 标签。
lockproto=LockModuleName	允许用户指定文件系统要使用的锁定协议。如果没有指定 LockModuleName，则会从文件系统的超级块中读取锁定协议。
locktable=LockTableName	允许用户指定文件系统要使用的锁定表。
quota=[off/account/on]	为文件系统打开或者关闭配额。在 account 状态中配置配额可让文件系统正确维护对每个 UID/GID 使用统计，忽略限制和警告值。默认值为 off。
errors=panic withdraw	当指定 errors=panic 时，文件系统错误将导致内核 panic。默认的行为与指定 errors=withdraw 一致，即将该系统从文件系统中退出，且直到重启前都无法访问。在有些情况下，该系统仍保持运行。有关 GFS2 退出功能请参考第 3.14 节“GFS2 撤回功能”。
discard/nodiscard	导致 GFS2 为释放的块生成 "discard" I/O 请求。合适的硬件可使用这些请求部署自动精简配置以及类似的方案。
barrier/nobarrier	当清洗日志时 GFS2 会发送 I/O 屏障。默认值为 on。如果基础设备不支持 I/O 屏障则会自动将这个选项改为 off。强烈建议您随时在 GFS2 中使用 I/O 屏障，除非块设备设计为不使用，这样它就不会丢失其写入缓存内容（例如：如果它在 UPS 中或者没有写入缓存）。
quota_quantum=secs	在将更改的配额信息写入配额文件前将其保存在某个节点的秒数。这是设定此参数的首选方法。该数值是一个大于 0 的整数。默认为 60 秒。设定为较短的间隔会让配额信息更快地更新，且更不可能让某些人超过其配额。较长的间隔可让文件系统操作更迅速有效地包括配额。
statfs_quantum=secs	设置 statfs 慢速版本的优选方法是将 statfs_quantum 设定为 0。默认值为 30 秒，该值设定了将 statfs 与主 statfs 文件同步前的最大时间段。可将该值调整为更迅速但不准确的 statfs 值，也可将其设定为更慢但更准确的值。当将该选项设定为 0 时，statfs 将总是报告真实值。
statfs_percent=value	提供在没有超时前，将 statfs 信息与主 statfs 文件同步前该信息更改的最大比例值。如果将 statfs_quantum 设定为 0，那么会忽略这个设置。

3.3. 卸载文件系统

可使用与卸载 Linux 文件系统相同的方法卸载 GFS2 文件系统 — 即使用 umount 命令。



备注

`umount` 是 Linux 系统命令。有关此命令的详情请参考 Linux `umount` 命令 man page。

用法

```
umount MountPoint
```

MountPoint

指定当前挂载 GFS2 文件系统的目录。

3.4. 挂载 GFS2 文件系统时的具体注意事项

该系统不会了解那些在系统关闭的过程中卸载的，通过手动挂载而不是使用 `fstab` 文件中的条目自动挂载的 GFS2 文件系统。因此，GFS2 脚本不会卸载 GFS2 文件系统。在运行 GFS2 关闭脚本后，标准关闭进程会杀死所有保留的用户进程，包括集群基础结构，并尝试卸载该文件系统。没有集群基础架构这个卸载会失败，且该系统会停滞。

要防止卸载 GFS2 文件系统时的系统停滞，您应该进行以下操作之一：

- 总是使用 `fstab` 文件中的条目挂载 GFS2 文件系统。
- 如果已经手动使用 `mount` 命令挂载了 GFS2 文件系统，请确定在重启或者关闭该系统前手动使用 `umount` 命令卸载该文件系统。

如果在这些情况下关闭系统的过程中，卸载文件系统时该文件系统停滞，请执行硬件重启。这样到不会丢失数据，因为在关闭进程的初期已经同步了该文件系统。

3.5. GFS2 配额管理

文件系统配额是用来限制某个用户或者组群使用的文件系统空间。在设置前对用户或者组群没有配额限制。当使用 `quota=on` 或者 `quota=account` 选项时，GFS2 会不断跟踪每个用户或者组群使用的空间，即使没有设定限制也是如此。GFS2 以互动的方式更新配额信息，因此系统崩溃并不需要重建配额用量。

为防止性能下降，GFS2 节点只会定时为配额文件更新同步。“fuzzy”配额核算可允许用户或者组群稍微超过设定的限制。为最小化这种情况，GFS2 会在接近“hard”配额限制时动态缩短同步周期。



注意

从红帽企业版 Linux 6.1 发行本开始，GFS2 支持标准 Linux 配额工具。要使用这个工具，您需要安装 quota RPM。这是在 GFS2 中管理配额的首选方法，且应该在所有使用配额新部署的 GFS2 中使用。本小节记录了如何使用这些工具管理 GFS2 配额。

在红帽企业版 Linux 之前的发行本中，GFS2 文件系统中的 gfs2_quota 命令管理配额。有关 gfs2_quota 命令的详情请参考 [附录 A, 使用 gfs2_quota 命令执行 GFS2 配额管理](#)。

3.5.1. 配置磁盘配额

请使用以下步骤实施磁盘配额：

1. 设置配额的强制或者计数模式。
2. 使用当前块使用信息初始化配额数据库文件。
3. 分配配额策略。（在计数模式中不强制这些策略。）

在以下小节中会详细讨论这些步骤的具体内容。

3.5.1.1. 将配额设定为强制或者计数模式

在 GFS2 文件系统中，默认禁用配额。要为文件系统启用配额，请在挂载文件系统时指定 quota=on 选项。

有可能在没有强制限制和警告值的情况下为每个用户和组群跟踪磁盘用量并维护配额核算。要做到这一点，请使用 quota=account 选项挂载文件系统。

用法

要挂载启用配额的文件系统，请在挂载文件系统时使用 quota=on 选项。

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

要在挂载文件系统时即使没有强制配额限制也要使用配额计数维护，则请在挂载文件系统时指定 quota=account 选项。

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

要在挂载文件系统时禁用配额，请使用 quota=off 选项挂载文件系统。这是默认设置。

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

quota={on|off|account}

on - 指定挂载文件系统时启用配额。

off - 指定挂载文件系统时禁用配额。

account - 即使在没有强制配额限制的情况下，也指定根据文件系统维护用户和组群的用量统计。

BlockDevice

指定 GFS2 文件系统所在的块设备。

MountPoint

指定要挂载 GFS2 文件系统的目录。

示例

在这个示例中，`/dev/vg01/lvo10` 中的 GFS2 文件系统被挂载到 `/mygfs2` 目录中并启用了配额。

```
mount -o quota=on /dev/vg01/lvo10 /mygfs2
```

在这个示例中，`/dev/vg01/lvo10` 中的 GFS2 文件系统被挂载到 `/mygfs2` 目录中并启用了配额计数，但没有强制。

```
mount -o quota=account /dev/vg01/lvo10 /mygfs2
```

3.5.1.2. 创建配额数据库文件

挂载了每个启用了配额的文件系统后，该系统就可以使用磁盘配额。但是该系统本身还不支持配额。下一步就是要运行 `quotacheck` 命令。

`quotacheck` 命令检查启用了配额的文件系统，并在每个文件系统中构建当前磁盘用量表。该表格可用于更新磁盘用量的操作系统副本。另外还会更新该文件系统的磁盘配额文件。

要在该文件系统中创建配额文件，请使用 `quotacheck` 命令的 `-u` 和 `-g` 选项。必须为用户和组群指定这两个选项方可进行初始化。如果为 `/home` 文件系统启用配额，则请在 `/home` 目录中生成该文件：

```
quotacheck -ug /home
```

3.5.1.3. 为每个用户分配配额

最后一步是使用 `edquota` 命令分配磁盘配额。请注意：如果您使用计数模式挂载文件系统（即指定 `quota=account` 选项），则不强制使用配额。

请在 shell 提示符后成为 root 用户，并执行以下命令为用户配置配额：

```
edquota username
```

为每个需要配额的用户执行这个步骤。例如：如果在 `/etc/fstab` 中为 `/home` 分区（在下面的示例中为 `/dev/Vo1Group00/LogVo102`）启用配额，并执行 `edquota testuser` 命令，则在系统默认的编辑器中会显示以下内容：

```
Disk quotas for user testuser (uid 501):
```

Filesystem	blocks	soft	hard	inodes	soft	hard
/dev/VolGroup00/LogVo102	440436	0	0			



备注

edquota 使用由 EDITOR 环境变量定义的文本编辑器。要更改编辑器，请在 `~/.bash_profile` 文件中将 EDITOR 环境变量设定为到您选择的编辑器的完整路径。

第一列是启用了配额的文件系统的名称。第二列显示目前该用户使用的块数。下两列是为该用户在该文件系统中设定的软限制和硬限制。

软块限制定义可使用的最大磁盘空间量。

硬块限制是用户或者组群可以使用的绝对最大磁盘空间量。达到这个上限后就再没有可以使用的磁盘空间了。

GFS2 文件系统不为内节点维护配额，因此这些列不适用于 GFS2 文件系统，为空白。

如果有任何值为 0，就是没有设定那个限制。您可以使用文本编辑器更改所需限制。例如：

```
Disk quotas for user testuser (uid 501):
Filesystem      blocks  soft  hard  inodes  soft  hard
/dev/VolGroup00/LogVo102 440436 500000 550000
```

请使用这个命令确认是否为用户设定了配额：

```
quota testuser
```

3.5.1.4. 为每个组群分配配额

还可以根据组群分配配额。请注意：如果您使用计数模式挂载您的文件系统（指定 `account=on` 选项），则不强制配额。

请使用以下命令为 `devel` 组群设定组群配额（设定组群配额前就存在该组群）：

```
edquota -g devel
```

这个命令在文本编辑器中显示该组群的现有配额：

```
Disk quotas for group devel (gid 505):
Filesystem      blocks  soft  hard  inodes  soft  hard
/dev/VolGroup00/LogVo102 440400 0 0
```

GFS2 文件系统不为内节点维护配额，因此这些列不适用于 GFS2 文件系统，为空白。您可以修改这些限制，然后保存文件。

请使用以下命令确认是否设定了组群配额：

```
quota -g devel
```

3.5.2. 管理磁盘配额

如果使用配额，则需要对其进行维护 — 大多数是查看是否超过了配额，并确定配额是准确的。

当然，如果用户不断超过其配额，或者持续达到其软限制，系统管理员可根据用户类型以及影响其工作的磁盘空间选择不同的处理方法。管理员可帮助用户决定如何使用较少的磁盘空间，或者增加该用户的磁盘配额。

您可以运行 `repquota` 程序创建磁盘用量报告。例如：命令 `repquota /home` 可有这样的输出：

```
*** Report for user quotas on device /dev/mapper/Vo1Group00-LogVo102
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
```

请使用以下命令查看所有启用了配额的文件系统的磁盘用量报告（选项 `-a`）：

```
repquota -a
```

虽然这个报告很好理解，但有几点需要注意。每个用户后显示的 `--` 可让您迅速确定是否超过块限制。如果超过块软限制，在输出结果的第一个 `-` 位置会出现 `+`。第二个 `-` 表示内节点限制，但 GFS2 文件系统不支持内节点限制，因此那个字符会保留为 `-`。GFS2 文件系统不支持宽限期，因此 `grace` 一列将为空白。

备注：不考虑基础文件系统，NFS 都不支持 `repquota` 命令。

3.5.3. 保持配额准确

如果您已经禁用配额运行文件系统一段时间并要再次启用配额，您应该运行 `quotacheck` 命令创建、检查并修复配额文件。另外，如果您认为配额文件不准确，您可以运行 `quotacheck` 命令，因为有可能在系统崩溃时没有完全卸载文件系统。

有关 `quotacheck` 命令的详情请查看 `quotacheck man page`。



注意

当文件系统在所有节点中相对停滞的时候运行 `quotacheck`，因为磁盘活性会消耗计算的配额值。

3.5.4. 使用 `quotasync` 命令同步配额

GFS2 在其磁盘自身内部文件中保存所有配额信息。GFS2 节点不需要在每次文件系统写入时更新这个配额文件，它会每 60 秒更新一次这个配额文件。这在避免节点间在写入配额文件时发生冲突是很有必要的，这种冲突会降低节点性能。

如果某个用户或者组群的配额使用接近了配额限制，GFS2 会动态减小配额文件更新的时间周期以防止超过配额限制。通常配额同步的时间周期为一个可调整的参数 `quota_quantum`，您可以使用 `gfs2_tool` 命令修改它。默认情况下，时间周期为 60 秒。同时，必须在每个节点以及每次挂载文件系统时设定 `quota_quantum` 参数。（卸载后不会保留修改的 `quota_quantum` 参数。）

您可以使用 `quotasync` 命令在 GFS2 执行自动更新时将某个节点的配额信息与磁盘配额文件同步。

用法

同步配额信息

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

u

同步该用户的配额文件。

g

同步该组群的配额文件

a

同步所有目前启用配额并支持同步的文件系统。缺少 `-a` 时，应是定文件系统挂载点。

mntpnt

指定要执行动作的 GFS2 文件系统。

调整同步时间间隔

```
gfs2_tool settune MountPoint quota_quantum Seconds
```

MountPoint

指定要执行动作的 GFS2 文件系统。

Seconds

指定 GFS2 对常规配额文件进行同步化的新时间周期。数值越小，竞争越激烈，同时还会降低性能。

示例

这个示例与为文件系统 `/mnt/mygfs2` 在磁盘配额文件中运行的节点同步所有缓冲的脏配额。

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

这个示例将单一节点中的文件系统 `/mnt/mygfs2` 的常规配额文件更新的默认时间周期改为一小时（3600 秒）。

```
gfs2_tool settune /mnt/mygfs2 quota_quantum 3600
```

3.5.5. 参考

有关磁盘配额的详情请参考以下命令的 man page:

- quotacheck
- edquota
- repquota
- quota

3.6. 增大的文件系统

`gfs2_grow` 是在其所在文件系统设备被扩展后用来扩展 GFS2 文件系统的。在现有 GFS2 文件系统中运行 `gfs2_grow` 命令则会填满目前文件系统终点和新初始化的 GFS2 文件系统扩展设备终点之间的所有剩余空间。当完成填充工作后，会为文件系统更新源索引。集群中的所有节点则可以使用以添加的额外存储空间。

`gfs2_grow` 必须在挂载的文件系统中运行，但只需要在集群的一个节点中运行。其它节点可感觉到扩展的发生，并可自动使用新的空间。



备注

您使用 `mkfs.gfs2` 命令创建 GFS2 文件系统后，您就无法缩小该文件系统的大小。

用法

```
gfs2_grow MountPoint
```

MountPoint

指定要执行动作的 GFS2 文件系统。

注释

在运行 `gfs2_grow` 命令前请您:

- 备份文件系统中的重要数据。
- 运行 `df MountPoint` 命令确定要进行扩展的文件系统的容量。
- 使用 LVM 扩展基础集群容量。有关管理 LVM 容量的信息请参考 LVM 管理员指南。

运行 `gfs2_grow` 命令后，请运行 `df` 命令查看文件系统中新的可用空间。

示例

在这个示例中扩展了 `/mygfs2fs` 目录中的文件系统。

```
[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.
```

完整用法

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

MountPoint

指定要挂载 GFS2 文件系统的目录。

Device

指定文件系统的设备节点。

表 3.3 “扩展文件系统是可用的 GFS2 具体选项” 描述在扩展 GFS2 文件系统时所要使用的 GFS2 具体选项。

表 3.3. 扩展文件系统是可用的 GFS2 具体选项

选项	描述
-h	帮助，显示简短用法信息。
-q	静默，降低详细等级。
-r MegaBytes	指定新资源组群大小，默认值为 256MB。
-T	测试。完成所有计算，但不要向磁盘中写入数据，也不要扩展文件系统。
-V	显示命令版本信息。

3.7. 在文件系统中添加日志

gfs2_jadd 命令可用来在 GFS2 文件系统中添加日志。您可以在任意点动态在 GFS2 文件系统中添加日志，且不需要扩展基础逻辑卷。gfs2_jadd 必须在挂载的文件系统中运行，但只需要在集群的一个节点中运行。其它节点可感觉到扩展的发生。



备注

如果某个 GFS2 文件系统已满，则 gfs2_jadd 将会失败，即使将逻辑卷包含的文件系统扩展到超过该文件系统的大小也是如此。这是因为在 GFS2 文件系统中，日志是纯文本文件，而不是嵌入的元数据，因此只是增大基础逻辑卷大小不会为日志提供空间。

在向 GFS 文件系统中添加日志前，您可以使用 `gfs2_tool` 命令的 `journals` 选项找出 GFS2 文件系统目前含有多少日志。以下示例显示挂载在 `/mnt/gfs2` 的文件系统中的日志数目和大小。

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

用法

```
gfs2_jadd -j Number MountPoint
```

Number

指定要添加的新日志数目。

MountPoint

指定要挂载 GFS2 文件系统的目录。

示例

这个示例中，在 `/mygfs2` 目录的文件系统中添加了一个日志。

```
gfs2_jadd -j1 /mygfs2
```

这个示例中，在 `/mygfs2` 目录的文件系统中添加了两个日志。

```
gfs2_jadd -j2 /mygfs2
```

完整用法

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

MountPoint

指定要挂载 GFS2 文件系统的目录。

Device

指定文件系统的设备节点。

表 3.4 “添加日志时可用的 GFS2 具体选项” 描述用来在 GFS2 文件系统中添加日志的 GFS2 具体选项。

表 3.4. 添加日志时可用的 GFS2 具体选项

标签	参数	描述
-h		帮助，显示简短用法信息。

标签	参数	描述
-J	MegaBytes	以 MB 为单位指定新日志的大小。默认日志大小为 128MB，最小值为 32MB。要在文件系统添加不同大小的日志，必须为每个不同大小的日志运行 <code>gfs2_jadd</code> 命令。指定的大小会不断下降，因此在生成文件系统时会指定多个日志区段。
-j	Number	用 <code>gfs2_jadd</code> 命令指定要添加的新日志数目，默认值为 1。
-q		静默，降低详细等级。
-V		显示命令版本信息。

3.8. 数据日志

通常 GFS2 只在其日志中写入元数据，接下来会由内核的周期性 `sync` 将文件内容写入磁盘，并被冲入系统缓冲。对某个文件的 `fsync()` 调用可导致将文件数据立刻写入磁盘，但磁盘报告所有数据已经安全写入后会返回调用。

数据日志可缩短非常小的文件的 `fsync()` 时间，因为文件数据在写入元数据外还要写入日志。随着文件的增大这个优势会明显降低。写入中等到大文件时打开数据日志会非常慢。

依赖 `fsync()` 同步文件数据的应用程序可能因使用数据日志而使性能有所提高。在被标记的目录及其所有子目录中生成的 GFS2 文件可自动启用数据日志。现有长度为 0 的文件也可以打开或者关闭其数据日志功能。

在将目录设定为 "inherit jdata" 的目录中启用数据日志，这个设定表示以后所有在那个目录中生成的文件和目录都进行日志操作。您可以使用 `chattr` 命令在文件中启用和禁用数据日志。

下面的命令在 `/mnt/gfs2/gfs2_dir/newfile` 文件中启用了数据日志，并查看是否正确设定了标签。

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

以下命令禁用了 `/mnt/gfs2/gfs2_dir/newfile` 文件中的数据日志，并查看是否正确设定了标签。

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

您还可以使用 `chattr` 命令在目录中设定 `j` 标签。当您为某个目录设定此标签时，以后在那个目录中创建的所有文件和目录也都会进行日志操作。下面的一组命令可在 `gfs2_dir` 目录中设定 `j` 标签，然后查看是否正确设定了该标签。此后，该命令会在 `/mnt/gfs2/gfs2_dir` 目录中生成一个名为 `newfile` 新文件，然后查看是否将为该文件设定了 `j` 标签。因为为该目录设定了 `j` 标签，那么应该也为 `newfile` 启用了日志操作。

```
[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
```

3.9. 配置 atime 更新

每个文件内节点和目录内节点都有三个与之相关的时间戳:

- ctime — 最后一次修改内节点状态的时间
- mtime — 最后一次修改文件（或者目录）数据的时间
- atime — 最后一次访问文件（或者目录）数据的时间

如果启用了 atime 更新，因为在 GFS2 和其它 Linux 文件系统中这是默认设置，那么每次在读取文件时都需要更新其内节点。

因为有些应用程序使用 atime 提供的信息，那些更新可能会需要大量不必要的写入流量和文件锁定流量。那个流量可能会降低性能，因此请尽量关闭或降低 atime 更新频率。

有两个可用来降低 atime 更新效果的方法:

- 使用 relatime (relative atime) 挂载，可在前一个 atime 更新比 mtime 或者 ctime 更新旧时更新 atime。
- 使用 noatime 挂载，可在那个文件系统中禁用 atime 更新。

3.9.1. 使用 relatime 挂载

当挂载文件系统时，可指定 relatime (relative atime) Linux 挂载选项。如果前一个 atime 更新比 mtime 或者 ctime 更新旧，这就可指定更新 atime。

用法

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

BlockDevice

指定 GFS2 文件系统所在的块设备。

MountPoint

指定要挂载 GFS2 文件系统的目录。

示例

在这个示例中，GFS2 文件系统位于 /dev/vg01/lvol0，并挂载到目录 /mygfs2。只有在前一个 atime 更新比 mtime 或者 ctime 更新旧时才会进行 atime 更新。

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

3.9.2. 使用 noatime 挂载

当挂载文件系统时，可指定 Linux 挂载选项 `noatime`，它可在那个文件系统中禁用 `atime` 更新。

用法

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

BlockDevice

指定 GFS2 文件系统所在的块设备。

MountPoint

指定要挂载 GFS2 文件系统的目录。

示例

在这个示例中，GFS2 文件系统位于 `/dev/vg01/lvo10`，并挂载到关闭了 `atime` 更新的 `/mygfs2` 目录中。

```
mount /dev/vg01/lvo10 /mygfs2 -o noatime
```

3.10. 在文件系统中挂起一个动作

您可以使用 `gfs2_tool freeze` 命令挂起对某个文件系统的写入动作。挂起写入动作允许使用基于硬件的设备快照捕获处于一致状态的文件系统。`gfs2_tool unfreeze` 命令可终止挂起。

用法

Start Suspension

```
gfs2_tool freeze MountPoint
```

End Suspension

```
gfs2_tool unfreeze MountPoint
```

MountPoint

指定文件系统。

示例

这个示例挂起了对文件系统 `/mygfs2` 的写入动作。

```
gfs2_tool freeze /mygfs2
```

这个示例终止了对文件系统 `/mygfs2` 写入动作的挂起。

```
gfs2_tool unfreeze /mygfs2
```

3.11. 修复文件系统

当挂载文件系统节点失败时，文件系统日志允许快速恢复。但如果存储设备断电或者断开物理连接，则会发生文件系统崩溃。（无法使用日志进行存储子系统失败修复。）当这种崩溃发生时，您可以使用 `fsck.gfs2` 命令修复 GFS2 文件系统。



警告

`fsck.gfs2` 命令必须只能在从所有节点中卸载的文件系统中运行。



备注

如果您以前有在 GFS 文件系统中使用 `gfs_fsck` 命令的经验，请注意 `fsck.gfs2` 命令在以下方面和之前发布的 `gfs_fsck` 有所不同：

- 在运行 `fsck.gfs2` 时按 `Ctrl+C` 会中断进程并显示提示信息，询问您是否要取消该命令，跳过剩余操作或者继续该进程。
- 您可以使用 `-v` 标签提高详细等级。添加第二个 `-v` 标签会再次提高等级。
- 您可以使用 `-q` 标签降低详细等级。添加第二个 `-q` 标签会再次降低等级。
- `-n` 会以只读方式打开某个文件系统并自动对所有查询回答 `no`。该选项提供了在不允许 `fsck.gfs2` 命令生效的前提下使用命令找出错误的方法。

有关其它命令选项详情请参考 `fsck.gfs2 man page`。

运行 `fsck.gfs2` 命令要求系统内存高于操作系统使用的内存而低于内核使用的内存。GFS2 文件系统中的每块内存本身需要大约 5 比特额外内存，或者 5/8 字节。因此要估算在您的文件系统中运行 `fsck.gfs2` 命令所需内存字节数，您可以确定包含的文件系统块的数量，然后乘以 5/8。

例如：要确定在大小为 16TB，块大小为 4K 的 GFS2 文件系统中运行 `fsck.gfs2` 命令大约需要多少内存，首先要确定该文件系统中包含多少内存块，可使用 16TB 除以 4K：

$$17592186044416 / 4096 = 4294967296$$

这个文件系统中包含 4294967296 个块，再乘以 5/8 就可确定需要多少内存：

$$4294967296 * 5/8 = 2684354560$$

该文件系统大约需要 2.6GB 可用内存方可运行 `fsck.gfs2` 命令。请注意：如果块大小为 1K，那么运行 `fsck.gfs2` 命令将需要四倍于以上值的内存，即大约 11GB。

用法

```
fscck.gfs2 -y BlockDevice
```

-y

-y 标签可使对所有问题的回答都为 yes。如果指定 -y 标签, fscck.gfs2 命令则不会在进行修改前提示您回答问题。

BlockDevice

指定 GFS2 文件系统所在的块设备。

示例

在这个示例中, 修复了位于块设备 /dev/testvol/testlv 中的 GFS2 文件系统。所有关于修复查询的回答都自动为 yes。

```
[root@dash-01 ~]# fscck.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
fscck.gfs2 complete
```

3.12. 绑定挂载以及上下文关联路径名

GFS2 文件系统不提供对上下文关联路径名 (CDPN) 的支持, CDPN 允许您生成指向不同目的地文件或目录的符号链接。在 GFS2 中使用 mount 命令的 bind 选项实现此功能。

mount 命令的 bind 选项允许您在不同位置重新挂载部分文件结构, 且使其在初始位置仍可用。该命令的格式如下:

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

执行此命令后, olddir 目录中的内容在两个位置可用: olddir 和 newdir。您还可以使用这个选项生成在两个位置可用的独立文件。

例如: 在执行以下命令后, /root/tmp 中的内容将和之前挂载的 /var/log 目录内容一致。

```
[root@mencryfa ~]# cd ~root
[root@mencryfa ~]# mkdir ./tmp
```

```
[root@menscryfa ~]# mount --bind /var/log /root/tmp
```

另外，您可以使用 `/etc/fstab` 文件中的条目在挂载时得到同样的结果。`/etc/fstab` 中的以下条目可使 `/root/tmp` 的内容和 `/var/log` 目录中的内容一致。

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

在您挂载文件系统后，您可以使用 `mount` 命令查看该文件系统是否被挂载了，如下示例所示：

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

对于支持上下文关联路径名的文件系统，您可以将 `/bin` 目录定义为上下文关联路径名，并根据系统构架将其解析为以下路径之一：

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

您可以通过生成空 `/bin` 目录得到同样的结果。然后使用脚本或者在 `/etc/fstab` 文件中的条目，将每个独立构架目录使用 `mount -bind` 命令挂载到 `/bin` 目录。例如：您可以使用以下命令作为脚本中的一行：

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

另外，您还可以使用以下行作为 `/etc/fstab` 文件的条目：

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

绑定挂载可为您提供比上下文关联路径名更大的灵活性，因为您可以使用此特性根据您定义的条件挂载不同的目录（比如文件系统的 `%f11` 值）。上下文关联路径名对其可处理的环境有更多的限制。请注意：您将需要根据条件（比如 `%f11`）编写您自己的挂载脚本。



警告

当您使用 `bind` 选项挂载文件系统，且起始文件系统以 `rw` 挂载时，新的文件系统也会被以 `rw` 形式挂载，即使您使用的是 `ro` 标签，`ro` 则被静默忽略了。在这种情况下，可能会在 `/proc/mounts` 目录中将新的文件系统标记为 `ro` 而引起误导。

3.13. 绑定挂载和文件系统挂载顺序

当您使用 `mount` 命令的 `bind` 选项时，您必须确定使用正确的顺序挂载该文件系统。在以下示例中，必须在 `/tmp` 目录中执行绑定挂载前挂载 `/var/log` 目录：

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

按以下方法决定文件系统挂载顺序:

- 通常文件系统挂载顺序由 `fstab` 文件中文件系统出现的顺序决定。例外情况包括使用 `_netdev` 标签挂载的文件系统或者有自身初始化脚本的文件系统。
- 有自身初始化脚本的文件系统在初始化进程后期挂载，即在挂载 `fstab` 文件中的文件系统之后挂载。
- 使用 `_netdev` 标签挂载的文件系统会在该系统中启用网络时挂载。

如果您的配置需要创建绑定挂载以便挂载 GFS2 文件系统，您可以命令 `fstab` 文件进行如下操作:

1. 挂载绑定挂载所需本地文件系统。
2. 绑定挂载要挂载 GFS2 文件系统的目录。
3. 挂载 GFS2 文件系统。

如果您的配置需要您在 GFS2 文件系统中绑定挂载本地目录或者文件系统，在 `fstab` 文件中列出正确的文件系统顺序也不会正确挂载文件系统，因为在 GFS2 初始化脚本运行前不会挂载该 GFS2 文件系统。在这种情况下，您应该在初始化脚本中写入命令执行绑定挂载，这样在挂载 GFS2 文件系统前就不会发生绑定挂载。

以下脚本是自定义初始化脚本示例。这个脚本在 GFS2 文件系统的两个目录中执行绑定挂载。在这个示例中，在 `/mnt/gfs2a` 有一个 GFS2 挂载点，可在集群启动并运行初始化脚本时执行挂载。

在这个示例脚本中 `chkconfig` 状态值说明:

- 345 说明启动该脚本的运行等级
- 29 是启动优先性，在这个示例中表示该脚本会在启动时在 GFS2 初始化脚本之后运行，后者的启动优先性为 26
- 73 为停止优先性，在这个示例中表示该脚本会在关闭时在 GFS2 脚本之后停止，后者的停止优先性为 74

启动和停止值表示您可以通过执行 `service start` 和 `service stop` 命令手动执行所示动作。例如：如果该脚本名为 `fredwilma`，则您可以执行 `service fredwilma start`。

应将这个脚本放到 `/etc/init.d` 目录并拥有与该目录中其他脚本相同的权限。您可以执行 `chkconfig on` 命令将该脚本链接到所示的运行等级中。例如：如果该脚本名为 `fredwilma`，那么您可以执行 `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

```

3.14. GFS2 撤回功能

GFS2 撤回功能是集群中 GFS2 文件系统的数据完整功能。如果 GFS2 内核模块探测到 GFS2 文件系统中存在不一致性，并伴随 I/O 操作，则该文件系统对该集群来说就不可用。该 I/O 操作会停止，同时该系统会等待下一个出错的 I/O 操作，防止进一步的破坏。当出现这种情况时，您可以停止手动停止任意服务或者程序，然后重启并重新挂载 GFS2 文件系统以便重新执行日志操作。如果问题仍存在，您可以在集群的所有节点中卸载该文件系统并使用 `fsck.gfs2` 命令执行文件系统恢复。GFS2 撤回功能没有内核 panic 那么严重，但可造成使用另一个节点 fence 这个节点。

如果您使用启用了启动脚本的 `gfs2` 配置您的系统，且在 `/etc/fstab` 文件中包含 GFS2 文件系统，则会在重启时重新挂载该 GFS2 文件系统。如果因为文件系统崩溃造成 GFS2 文件系统撤回，则建议您在重新挂载该文件系统前运行 `fsck.gfs2` 命令。在这种情况下，要防止您的文件系统在引导时重新挂载，您可以执行以下步骤：

1. 使用以下命令暂时在受影响的节点中禁用启动脚本：

```
# chkconfig gfs2 off
```

2. 重启受影响的节点，启动集群软件。此时将不会挂载该 GFS2 文件系统。
3. 在集群的所有节点中卸载该文件系统。
4. 只在确定没有文件系统崩溃的一个节点的文件系统中运行 `fsck.gfs2`。
5. 运行以下命令在受影响的节点中重新启用该启动脚本：

```
# chkconfig gfs2 on
```

6. 在集群的所有节点中重新挂载 GFS2 文件系统。

可能造成 GFS2 收回的不一致性示例为错误的块计数。当 GFS 内核删除文件系统中的某个文件上，它会系统地删除与那个文件关联的所有数据和元数据。完成此操作后，它会检查块计数。如果该块计数不是 1（意思是只剩下磁盘内节点自己），则表明文件系统不一致，因为块计数与找到的块列表不匹配。

您可以使用 `-o errors=panic` 选项挂载文件系统覆盖 GFS2 收回功能。当指定这个选项时，所有导致系统收回的错误都将导致系统 panic。这样可停止节点的集群沟通，并 fence 该节点。

在内部，GFS2 通过让内核向 `gfs_controld` 守护进程发送请求撤回的信息启动撤回功能。`gfs_controld` 会运行 `dmsetup` 程序替换该文件系统中的设备映射器错误目标以防止进一步访问块设备。然后它会告诉内核操作完成。这是 GFS2 支持要求总是在 GFS2 使用 CLVM 设备的理由，否则就不可能插入设备映射器目标。

设备映射器错误目标的目的是确保将来所有的 I/O 操作都将有一个 I/O 出错信息会让该文件系统以较旧的方式卸载。结果是当出现撤回时，通常会在系统日志中看到来自设备映射器的 I/O 出错信息。

偶尔在 `dmsetup` 程序不可能根据要求插入错误目标则撤回会失败。如果撤回时缺少内存，且由于造成撤回的问题而无法重新使用内存会发生这种情况。

撤回并不总是意味着 GFS2 中有错误。有时撤回功能可由与基础块设备有关的设备 I/O 错误引发。强烈建议您检查日志来查看发生撤回是否是这样的原因。

诊断并修正 GFS2 文件系统的问题

本章提供常见 GFS2 问题以及如何处理这些问题的信息。

4.1. GFS2 文件系统出现性能缓慢

您会发现 GFS2 文件系统显示出比 EXT3 文件系统缓慢的性能。在某些情况下 GFS2 的性能可能受到一些因素的影响。本文档中有关于处理 GFS2 性能问题的信息。

4.2. 通过 GFS2 设定 NFS

由于 GFS2 锁定子系统增加的复杂性及其集群的本质，通过 GFS2 设定 NFS 要求您非常仔细、小心进行配置。本小节论述了您在通过 GFS2 文件系统配置 NFS 服务时应该注意的地方。



警告

如果是使用 NFS 导出 GFS2 文件系统，且 NFS 客户端应用程序使用 POSIX 锁定，那么您必须使用 `localflocks` 选项挂载该文件系统。这样做预期的效果是将每个服务器强制 POSIX 锁定为本地：例如，非集群，独立的。（如果 GFS2 尝试使用 NFS 在集群的多个节点中部署 POSIX 锁定会有很多问题。）对于在 NFS 客户端中运行的应用程序，如果在不同的服务器中挂载两个客户端，本地的 POSIX 锁定意味着两个客户端可同时拥有同一个锁定。如果所有客户端都使用一个服务器挂载 NFS，那么不同服务器分别赋予同一锁定的问题就不存在了。

除锁定注意事项外，您应该在通过 GFS2 文件系统配置 NFS 时考虑以下问题。

- 红帽只支持使用主动/被动锁定的 NFSv3 进行红帽高可用性附加组件配置，并具有以下特征：
 - 后端文件系统是在 2 -- 16 个节点集群中运行的 GFS2 文件系统。
 - NFSv3 服务器的定义是一次从单一集群节点中导出整个 GFS2 文件系统的服务。
 - 该 NFS 服务器可能在从一个集群节点到另一个节点时失败（主动/被动配置）。
 - 不允许到 GFS2 文件系统的访问，除非是通过 NFS 服务器。这包括本地 GFS2 文件系统访问以及通过 Samba 或者集群的 Samba 的访问。
 - 在该系统中不支持 NFS 配额。

这个配置为该文件系统提供 HA，并降低系统当机时间，因为失败的节点不会造成在从一个节点到另一个节点有 NFS 服务器失败时要执行 `fsck` 命令。

- GFS2 的 NFS 导出中 `fsid= NFS` 选项是必须的。
- 如果您的集群出现问题（例如：集群成为 `inquorate`，且 `fencing` 不成功），集群的逻辑卷和 GFS2 文件系统将会停滞，且不可能有访问直到集群成为 `quorate`。您在决定在这个过程中定义的简单故障切换解决方案是否为您系统的最佳解决方案时应该考虑这种可能性。

4.3. GFS2 文件系统挂起并需要在一个节点中重启

如果您的 GFS2 文件系统挂起，且不返回由此运行的命令，但重启一个返回常态的具体节点，这可能表示有锁定问题或者 bug。出现这种情况时您应该收集以下数据：

- 在每个节点中为该文件系统执行的 gfs2 锁定转储:

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

- 每个节点中为该文件系统执行的 DLM 锁定转储: 您可以使用以下 dlm_tool 命令收集这个信息:

```
dlm_tool lockdebug -sv lsname.
```

在这个命令中, lsname 是有问题的附加系统 DLM 使用的锁定空间名称。您可以在 group_tool 命令的输出结果中找到这个值。

- sysrq -t 命令的输出结果。
- /var/log/messages 文件的内容。

您收集到数据后, 可以在红帽支持生成一个 ticket, 并提供您收集的数据。

4.4. GFS2 文件系统挂起并需要重启所有节点

如果您的 GFS2 文件系统挂起且不会返回由此运行的命令, 则需要您在使用它之前重启该集群中的所有节点, 并检查以下问题。

- 您可能会有一个失败的 fence。GFS2 文件系统将停滞以保证在失败的 fence 事件中数据的完整性。检查信息日志查看挂起时是否有失败的 fence。请确定您正确配置了 fencing。
- GFS2 文件系统可能已经撤回。检查信息日志查看关键字 withdraw, 看看是否有来自 GFS2 表示已经撤回的文件系统的信息和 calltrace。撤回表示可能为系统崩溃、存储失败或者是一个 bug。卸载该文件系统, 更新 gfs2-utils 软件包并在该文件系统中执行 fsck 命令以便返回该服务。在红帽支持生成一个支持 ticket。通知它们您有 GFS2 撤回问题, 并提供有日志的 sosreport。

有关 GFS2 撤回功能的详情请参考 [第 3.14 节 “GFS2 撤回功能”](#)。

- 这个出错信息表示有锁定问题或者 bug。在这种情况下之一收集数据并在红帽支持生成一个支持 ticket, 如 [第 4.3 节 “GFS2 文件系统挂起并需要在节点中重启”](#) 所述。

4.5. GFS2 文件系统不挂载新添加的集群节点

如果您在集群中添加一个新节点, 且您发现您无法在那个节点中挂载 GFS2 文件系统, 那么相较尝试访问 GFS2 文件系统的节点, 您的 GFS2 文件系统日志可能较少。您必须在每个要挂载文件系统的 GFS2 主机中都有一个日志 (使用 spectator 挂载选项组挂载的 GFS2 文件系统除外)。您可以使用 gfs2_jadd 命令在 GFS2 文件系统中添加日志, 如 [第 3.7 节 “在文件系统中添加日志”](#) 所述。

4.6. 空格代表在空文件系统中使用

如果您有一个空 GFS2 文件系统, df 命令将显示使用的空间。这是因为 GFS2 文件系统日志消耗磁盘空间 (日志数乘以日志大小)。如果您创建有大量日志的 GFS2 文件系统或者指定大日志, 那么您将在执行 df 时看到空间已经被使用 (日志数乘以日志大小)。即使您没有指定较大的日志数或者大日志, 小 GFS2 文件系统 (1GB 或者更小) 将使用大量默认 GFS2 日志大小的空间。

附录 A. 使用 gfs2_quota 命令执行 GFS2 配额管理

从红帽企业版 Linux 601 发行本开始，GFS2 支持标准 Linux 配额设备。要使用这个功能，您需要安装 quota RPM。这是在 GFS2 中管理配额的首选方法，并应在所有使用配额的 GFS2 新部署中使用。有关使用标准 Linux 配额设备的详情请参考 [第 3.5 节 “GFS2 配额管理”](#)。

在红帽企业版 Linux 之前的版本中，GFS2 需要 gfs2_quota 命令来管理配额。这个附录记录了使用 gfs2_quota 命令进行 GFS2 文件系统配额管理的内容。

A.1. 使用 gfs2_quota 命令设定配额

每个用户 ID (UID) 或者组群 ID (GID) 都有两个配额设置：硬限制和软限制。

硬限制是可以使用的空间数量。该文件系统不会让用户或者组群使用超过该数量的磁盘空间。硬限制值为零表示没有限制。

软限制通常是一个小于硬限制的值。当达到软限制时该文件系统会通知用户或者组群以便软化其使用的空间量。软限制值为 0 表示没有强制的限制。

您可以使用 gfs2_quota 命令设定限制。只需要在 GFS2 挂载的单一节点中运行此命令即可。

默认情况下不会在 GFS2 文件系统设定配额强制。要启用配额计数，请在挂在 GFS2 文件系统时使用 mount 命令的 quota= 选项，如 [第 A.4 节 “启用/禁用配额强制”](#) 所述。

用法

Setting Quotas, Hard Limit

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

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

Setting Quotas, Warn Limit

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

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

User

限制或者警告的用户 ID。可以是密码文件中的用户名，也可以是 UID 号。

Group

限制或者警告的组群 ID。可以是组群文件中的组群名称或者 GID 号。

Size

指定要限制或者警告的新值。默认情况下该值以 MB 为单位。使用 -k、-s 和 -b 标签可分别将单位改为 kb、扇区和文件系统块。

MountPoint

指定要执行动作的 GFS2 文件系统。

示例

这个示例在文件系统 /mygfs 中将用户 Bert 的硬限制设定为 1024MB (1GB)。

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

这个示例在文件系统 /mygfs 中将组群 ID 21 的软限制设定为 50kb。

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

A.2. 使用 gfs2_quota 命令显示配额限制和用量

可使用 gfs2_quota get 命令为具体用户或者组群显示配额限制和当前用量。还可使用 gfs2_quota list 命令显示配额文件的整个内容，其中所有 ID 都是非 0 硬限制或者列出的值。

用法

Displaying Quota Limits for a User

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

Displaying Quota Limits for a Group

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

Displaying Entire Quota File

```
gfs2_quota list -f MountPoint
```

User

显示具体用户信息的用户 ID。可以使用密码文件中的用户名或者 UID 号。

Group

显示具体组群信息的组群 ID。可以是组群文件中的组群名称或者 GID 号。

MountPoint

指定要执行动作的 GFS2 文件系统。

命令输出

gfs2_quota 命令显示 GFS2 配额信息如下：

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

LimitSize、WarnSize 和 Value 数（值）默认使用 MB 为单位。在命令行中使用 `-k`、`-s` 或者 `-b` 标签可分别将单位更改为 kb、扇区或者文件系统块。

User

与该数据关联的用户名或者 ID。

Group

与该数据关联的组群名称或者 ID。

LimitSize

用户或者组群的硬限制。如果没有设定限制则该值为 0。

Value

用户或者组群实际使用的磁盘空间量。

注释

显示配额信息时，如果在 `gfs2_quota` 中添加 `-n` 选项，则该命令不会将 UID 或者 GID 解析为名称。

在命令行中添加 `-d` 选项会在显示值中将为 root UID 和 GID 分配的 GFS 隐藏文件空间排除在外。当您尝试将 `gfs2_quota` 中的数字与 `du` 命令的结果匹配时很有用。

示例

这个示例显示所有设定了限制或者使用文件系统 `/mygfs2` 所有空间的用户和组群的配额信息。

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

这个手里以扇区为单位为文件系统 `/mygfs2` 中的组群显示配额信息。

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

A.3. 使用 `gfs2_quota` 命令同步配额

GFS2 在其磁盘的内部文件中保存所有配额信息。GFS2 节点不会在每次写入系统文件时更新这个配额文件，默认情况下它每 60 秒更新一次配额文件。这样可避免节点在写入配额文件时出现竞争，这种竞争可导致性能下降。

随着用户或者组群接近其配额限制，GFS2 会动态减少更新配额文件的时间间隔以避免超过限额。配额同步的时间间隔 `quota_quantum` 是一个可调参数，您可使用 `gfs2_tool` 命令将其从默认的 60 秒更改为其它数值。另外，每次挂载文件系统时都必须在每个节点中都设置 `quota_quantum` 参数。（卸载后不会保留对 `quota_quantum` 参数的更改。）

您可以使用 `gfs2_quota sync` 命令会在 GFS2 执行自动更新时将某个节点中的配额信息与磁盘中的配额文件同步。

用法

Synchronizing Quota Information

```
gfs2_quota sync -f MountPoint
```

MountPoint

指定要执行动作的 GFS2 文件系统。

Tuning the Time Between Synchronizations

```
gfs2_tool settune MountPoint quota_quantum Seconds
```

MountPoint

指定要执行动作的 GFS2 文件系统。

Seconds

指定 GFS2 常规配额文件同步的新时间间隔。较小的值可增加竞争并降低性能。

示例

这个示例是与在文件系统 /mygfs2 中运行的节点同步配额信息。

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

这个示例将单一节点文件系统 /mygfs2 的常规配额文件更新的默认时间间隔改为一个小时（3600 秒）。

```
gfs2_tool settune /mygfs2 quota_quantum 3600
```

A.4. 启用/禁用配额强制

在 GFS2 文件系统中默认为禁用配额强制。要为文件系统启用配额强制，您可以指定 `quota=on` 选项挂载文件系统。

用法

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

要在挂载文件系统时禁用配额强制，请指定 `quota=off` 选项挂载文件系统。这是默认设置。

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

`-o quota={on|off}`

挂载文件系统时指定是启用还是禁用配额强制。

BlockDevice

指定 GFS2 文件系统所在的块设备。

MountPoint

指定应挂载 GFS2 文件系统的目录。

示例

在这个示例中，将 `/dev/vg01/lvo10` 中的 GFS2 文件系统挂载到 `/mygfs2` 目录并启用配额强制。

```
mount -o quota=on /dev/vg01/lvo10 /mygfs2
```

A.5. 启用配额计数

可在不强制的情况下为每个用户和组群追踪磁盘用量并维护配额计数。要做到这一点，需要在挂载文件系统时指定 `quota=account` 选项。

用法

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

`-o quota=account`

指定由该文件系统维护的用户和组群用量统计，即使没有强制配额限制。

BlockDevice

指定 GFS2 文件系统所在的块设备。

MountPoint

指定应挂载 GFS2 文件系统的目录。

示例

在这个示例中将 `/dev/vg01/lvo10` 中的 GFS2 文件系统挂载到 `/mygfs2` 目录并启用配额计数。

```
mount -o quota=account /dev/vg01/lvo10 /mygfs2
```


附录 B. 将文件系统从 GFS 转换为 GFS2

由于红帽企业版 Linux 6 不支持 GFS 文件系统，因此您必须使用 `gfs2_convert` 命令将现有 GFS 文件系统升级到 GFS2 文件系统。请注意：您必须在升级到红帽企业版 Linux 6 之前在红帽企业版 Linux 5 中执行这个转换过程。



警告

在转换 GFS 文件系统前，您必须备份您的文件系统，因为转换过程是不可逆的，且在转换过程中出现的错误可导致进程意外终结，从而使文件系统不可用。

在转换 GFS 文件系统前，您必须使用 `gfs_fck` 命令检查文件系统并修复所有错误。

如果由于停电或者其它问题导致 GFS 到 GFS2 的转换中断，请重启转换工具。在转换完成前不要在文件系统中执行 `fsck.gfs2` 命令。



上下文关联路径名

GFS2 文件系统不提供对上下文独立路径名 (CDPN) 的支持，CDPN 允许您生成指向不同目的文件或目录的符号链接。您可使用 `mount` 命令的 `bind` 选项获得 CDPN 在 GFS2 文件系统中的作用。

`gfs2_convert` 命令识别 CDPN，并使用有相同名称的空目录替换它们。要配置绑定挂载替换 CDPN，则需要您了解要替换的 CDPN 目标的完整路径。转换您的文件系统前，您可以使用 `find` 命令识别该链接。

以下命令列出指向 `hostname` CDPN 的符号链接：

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

同样，您可以为其它 CDPN (`mach`、`os`、`sys`、`uid`、`gid`、`jid`) 执行 `find` 命令。请注意：因为 CDPN 名称的格式可以是 `@hostname` 或者 `{hostname}`，所以您需要为每个变体运行 `find` 命令。

有关 GFS2 中绑定挂载和文本独立路径名的详情请参考 [第 3.12 节 “绑定挂载以及上下文关联路径名”](#)。

当您转换全部或者接近全部文件系统时，可能没有足够的空间放置 GFS2 文件系统数据结构。在这种情况下，所有日志都统一缩小为适应可用空间的大小。

采用以下步骤将 GFS 文件系统转换为 GFS2 文件系统。

1. 在红帽企业版 Linux 系统中，请备份您的现有 GFS 文件系统。
2. 将 GFS 文件系统从集群中的所有节点中卸载。
3. 在 GFS 文件系统中执行 `gfs_fck` 命令以确定没有文件系统崩溃。

4. 执行 `gfs2_convertgfsfilesystem`。该系统会显示警告信息并在将 `gfsfilesystem` 转换为 GFS2 之前进行确认。
5. 升级到红帽企业版 Linux 6。

以下示例是在块设备 `/dev/she11_vg/500g` 中将 GFS 文件系统转换为 GFS2 文件系统。

```
[root@she11-01 ~]# /root/cluster/gfs2/convert/gfs2_convert /dev/she11_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/she11_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/she11_vg/500g: filesystem converted successfully to gfs2.
```

附录 C. 修订记录

修订 2.0-1 Thu May 19 2011

Steven Levine slevine@redhat.com

红帽企业版 Linux 6.1 初始发行本

解决: #549838

编写红帽企业版 Linux 6.1 中标准 Linux 配额工具支持文档。

解决: #608750

澄清 GFS2 撤回功能。

解决: #660364

修正最大 GFS2 文件系统大小信息。

解决: #687874

在 GFS2 故障排除中添加新的一章。

解决: #664848

添加从 GFS 转换为 GFS2 之前查找上下文独立路径名称的信息。

修订 1.0-1 Wed Nov 15 2010

Steven Levine slevine@redhat.com

红帽企业版 Linux 6 初始发行本

索引

符号

- 修复文件系统, 31
- 初始任务
 - 设定, 初始阶段, 11
- 前提任务
 - 配置, 初始阶段, 11
- 前言 (见 简介)
- 功能, 新的和更改的, 2
- 卸载命令, 18
- 卸载文件系统, 18, 19
- 卸载时系统停滞, 19
- 卸载, 系统停滞, 19
- 反馈
 - 此手册的联络信息, v
- 在文件系统中暂停动作, 30
- 在文件系统中添加日志, 26
- 增大文件系统, 25
- 性能调节, 6
- 挂载命令, 16
- 挂载文件系统, 16, 19
- 挂载表格, 17
- 撤回功能, GFS2, 35
- 数据日志, 28
- 文件系统
 - atime, 配置更新, 29
 - 使用 noatime 挂载, 29
 - 使用 relatime 挂载, 29
 - 上下文相关的路径名 (CDPN), 32
 - 修复, 31
 - 卸载, 18, 19
 - 增大, 25
 - 挂载, 16, 19
 - 挂载顺序, 33
 - 数据日志, 28
 - 暂停动作, 30
 - 添加日志, 26
 - 生成, 13
 - 绑定挂载, 32
 - 配额管理, 19, 20,
 - 同步配额, 23, 41
 - 启用/禁用配额强制, 42
 - 启用配额计数, 43
 - 显示配额限制, 40
 - 设定配额, 39
- 最大值, GFS2 文件系统,
- 概述,
 - 功能, 新的和更改的, 2
 - 配置, 之前, 2
- 添加日志的 GFS2 具体选项表, 27
- 生成文件系统, 13
- 用于扩展文件系统表的 GFS2 具体选项, 26

- 磁盘配额
 - 为每个用户分配, 21
 - 为每个组群分配, 22
 - 启用, 20
 - quotacheck 运行, 21
 - 创建配额文件, 21
 - 硬限制, 22
 - 管理, 23
 - quotacheck 命令, 用来检查, 23
 - 报告, 23
 - 软限制, 22
 - 附加资源, 25
- 简介,
 - 读者, v
- 管理 GFS2,
- 绑定挂载, 32
 - 挂载顺序, 33
- 节点锁定, 6
- 表格
 - mkfs.gfs2 命令选项, 15
 - 挂载选项, 17
 - 添加日志的 GFS2 具体选项, 27
 - 用于扩展文件系统的 GFS2 具体选项, 26
- 设置, 初始阶段
 - 初始任务, 11
- 读者, v
- 调节, 性能, 6
- 路径名, 上下文关联 (CDPN), 32
- 配置, 之前, 2
- 配置, 初始阶段,
 - 前提任务, 11
- 配额管理, 19, 20,
 - 同步配额, 23, 41
 - 启用/禁用配额强制, 42
 - 启用配额计数, 43
 - 显示配额限制, 40
 - 设定配额, 39

A

- acl 挂载选项, 16
- atime, 配置更新, 29
 - 使用 noatime 挂载, 29
 - 使用 relatime 挂载, 29

D

- debugfs 文件, 7

F

- fsck.gfs2 命令, 31

G

- GFS2
 - atime, 配置更新, 29

- 使用 noatime 挂载, 29
- 使用 relatime 挂载, 29

撤回功能, 35

管理,

配额管理, 19, 20,

- 同步配额, 23, 41

- 启用/禁用配额强制, 42

- 启用配额计数, 43

- 显示配额限制, 40

- 设定配额, 39

GFS2 文件系统最大值,

gfs2_grow 命令, 25

gfs2_jadd 命令, 26

gfs2_quota command,

glock 拥有者标签, 8

glock 标签, 8

glock 类型, 8

M

mkfs 命令, 13

mkfs.gfs2 命令选项表, 15

Q

quota= mount option, 39

quotacheck , 21

quotacheck 命令

- 检查配额准确, 23

quota_quantum tunable parameter, 24, 41