



Red Hat Enterprise Linux 5

全局网络块设备

在红帽全局文件系统中使用 GNBD
版 3

Landmann

在红帽全局文件系统中使用 GNBD 版 3

Landmann
rlandmann@redhat.com

法律通告

Copyright © 2009 Red Hat, Inc.

This document is licensed by Red Hat under the [Creative Commons Attribution-ShareAlike 3.0 Unported License](#). 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, 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.

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.

摘要

本书提供了为红帽企业版 Linux 5 在红帽 GFS 中使用全局网络块设备（GNBD）的概述。

目录

简介	3
1. 有关此指南	3
2. 读者	3
3. 软件版本	3
4. 相关文档	3
5. 反馈	3
6. 文档约定	4
6.1. 排版约定	4
6.2. 抬升式引用约定	5
6.3. 备注及警告	6
第 1 章 使用带 Red Hat GFS 的 GNBD	7
第 2 章 考虑在设备映射器多路径中使用 GNBD	8
2.1. Linux 页面缓存	8
2.2. 栅栏 GNBD 服务器节点	8
第 3 章 GNBD 驱动程序及命令用法	9
3.1. 从服务器中导出 GNBD	9
用法	9
示例	11
3.2. 在客户端中导入 GNBD	11
用法	11
示例	11
第 4 章 在 GNBD 服务器节点中运行 GFS	12
修订记录	13
索引	13
D	13
E	13
F	13
G	13
I	13
L	13
S	13

简介

1. 有关此指南

本书描述了如何在全局文件系统（GFS）中使用全局网络块设备（GNBD），其中包括有关设备映射器多路径、GNBD 驱动程序和命令用法以及在 GNBD 服务器节点中运行 GFS。

2. 读者

本书主要面向管理运行 Linux 操作系统的系统的管理员。它要求您熟悉红帽企业版 Linux 以及 GFS 文件系统管理。

3. 软件版本

表 1. 软件版本

软件	描述
RHEL5	指的是 RHEL5 和更高的版本
GFS	指的是 RHEL5 和更高的版本中的 GFS

4. 相关文档

有关使用红帽企业版 Linux 的详情请参考以下资源：

- ▶ 《红帽企业版 Linux 安装指南》 — 提供有关安装红帽企业版 Linux 5 的信息。
- ▶ 《红帽企业版 Linux 部署指南》 — 提供有关部署、配置和管理红帽企业版 Linux 5 的信息。

有关 Red Hat Enterprise Linux 5 中 Red Hat Cluster Suite 的详情请参考以下资源：

- ▶ 《Red Hat Cluster Suite 概述》 — 提供红帽群集套件的高级概述。
- ▶ 《配置和管理红帽群集》 — 提供有关安装、配置和管理红帽群集套件的信息。
- ▶ *Logical Volume Manager Administration* — Provides a description of the Logical Volume Manager (LVM), including information on running LVM in a clustered environment.
- ▶ 《全局文件系统：配置及管理》 — 提供有关安装、配置和维护红帽 GFS（红帽全局文件系统）的信息。
- ▶ 《全局文件系统 2：配置及管理》 — 提供有关安装、配置和维护红帽 GFS2（红帽全局文件系统 2）的信息。
- ▶ 《使用设备映射器多路径》 — 提供有关在 Red Hat Enterprise Linux 5 中使用设备映射器多路径的信息。
- ▶ 《Linux 虚拟服务器管理》 — 提供使用 Linux 虚拟服务器（LVS）配置高性能系统及服务的信息。
- ▶ 《红帽群集套件发行注记》 — 提供有关红帽群集套件当前发行本的信息。

在 Red Hat Enterprise Linux 文档 CD 以及在线文档 <http://www.redhat.com/docs/> 中都有 Red Hat Cluster Suite 文档以及其它 Red Hat 文档的 HTML、PDF 和 RPM 版本。

5. 反馈

我们很乐于了解您有关打印错误或者可使本手册更好工作的想法。请在 Bugzilla (<http://bugzilla.redhat.com/bugzilla/>) 中根据 **rh-cs** 提交您的报告。

Be sure to mention the manual's identifier:

```
Bugzilla component: Documentation-cluster
Book identifier: Global_Network_Block_Device(EN)-5 (2009-08-18T15:25)
```

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

如果您有任何可改进此文档的建议，请尽量说得具体一些。如果您发现任何错误，请将错误所在部分号码以及上下文包括在内，以便我们比较容易找到这些错误之处。

6. 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](#) 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。

6.1. 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。

按键组合与单独按键之间的区别是按键组合是使用加号将各个按键连在一起。例如：

按 `Enter` 执行该命令。

按 `Ctrl+Alt+F2` 切换到虚拟终端。

第一个示例突出的是要按的特定按键。第二个示例突出了按键组合：一组要同时按下的三个按键。

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` 文件中插入特殊字符，请在主菜单栏中选择「应用程序」 → 「附件」 → 「字符

映射表」。接下来选择从 **Character Map** 菜单中选择 **Search** → 「查找.....」，在「搜索」字段输入字符名称并点击「下一个」按钮。此时会在「字符映射表」中突出您搜索的字符。双击突出的字符将其放在「要复制的文本」字段中，然后点击「复制」按钮。现在返回您的文档，并选择 **gedit** 菜单中的「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。

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***。

请注意上述使用黑斜体的文字 -- *username*、*domain.name*、*file-system*、*package*、*version* 和 *release*。每个字都是一个站位符，可用作您执行命令时输入的文本，也可作为该系统显示的文本。

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。

6.2. 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
```

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


```

static int kvm_vm_ioctl_deassign_device(struct kvm *kvm,
                                       struct kvm_assigned_pci_dev *assigned_dev)
{
    int r = 0;
    struct kvm_assigned_dev_kernel *match;

    mutex_lock(&kvm->lock);

    match = kvm_find_assigned_dev(&kvm->arch.assigned_dev_head,
                                  assigned_dev->assigned_dev_id);
    if (!match) {
        printk(KERN_INFO "%s: device hasn't been assigned before, "
               "so cannot be deassigned\n", __func__);
        r = -EINVAL;
        goto out;
    }

    kvm_deassign_device(kvm, match);

    kvm_free_assigned_device(kvm, match);

out:
    mutex_unlock(&kvm->lock);
    return r;
}

```

6.3. 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。



注意

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。

第 1 章 使用带 Red Hat GFS 的 GNBD

GNBD（全局网络块设备）通过以太网 LAN 提供块级别存储访问。GNBD 组件在 GFS 节点中作为客户端运行，在 GNBD 服务器节点中作为服务器运行。GNBD 服务器节点从其存储（可以是直接附加的存储或者 SAN 存储）中将块等级存储导出到 GFS 节点。

[表 1.1 “GNBD 软件子系统组件”](#) summarizes the GNBD software subsystems components.

表 1.1. GNBD 软件子系统组件

软件子系统	组件	描述
GNBD	gnbd.ko	在客户端中部署 GNBD 设备驱动程序的内核模块。
	gnbd_export	在 GNBD 服务器中创建、导出和管理 GNBD 的命令。
	gnbd_import	在 GNBD 服务器中导入和管理 GNBD 的命令。
	gnbd_serv	允许节点通过网络导出本地存储的服务器守护进程。

您矿浆 GNBD 服务器配置为与设备映射器多路径一同工作。带设备映射器多路径的 GNBD 允许您配置多个 GNBD 服务器节点以便提供到存储设备的冗余路径。GNBD 服务器依次通过冗余 GNBD 为 GFS 节点显示多路径存储路径。当使用带设备映射器多路径的 GNBD 时，如果一个 GNBD 服务器节点不可用，则另一个 GNBD 服务器节点会提供可访问存储设备的 GFS 节点。

本文档提供如何使用带 Red Hat GFS 的 GNBD，它由以下章节组成：

- ▶ [第 2 章 考虑在设备映射器多路径中使用 GNBD](#), which describes some of the issues you should take into account when configuring multipathed GNBD server nodes
- ▶ [第 3 章 GNBD 驱动程序及命令用法](#), which describes the restrictions that apply when you are running GFS on a GNBD server node
- ▶ [第 4 章 在 GNBD 服务器节点中运行 GFS](#), which describes the user commands that configure GNBD

第 2 章 考虑在设备映射器多路径中使用 GNBD

带设备映射器多路径的 GNBD 允许您配置多个 GNBD 服务器节点（将 GNBD 导出到 GFS 节点的节点）来为存储设备提供冗余路径。GNBD 服务器节点依次通过冗余 GNBD 为 GFS 节点显示多存储路径。当使用带设备映射器多路径的 GNBD 时，如果某个 GNBD 服务器节点不可用，另一个 GNBD 服务器节点就提供可访问存储设备的 GFS 节点。

如果您正在使用带设备映射器多路径的 GNBD，您需要考虑以下几点：

- ▶ Linux page caching, as described in [第 2.1 节“Linux 页面缓存”](#).
- ▶ Fencing GNBD server nodes, as described in [第 2.2 节“栅栏 GNBD 服务器节点”](#).
- ▶ GNBD device names; export names for GNBD devices must be unique. Additionally, you must specify the **-u** or **-U** when using the **gnbd_export** command. Exporting GNBD devices is described in [第 3 章 GNBD 驱动程序及命令用法](#).

2.1. Linux 页面缓存

For GNBD with device-mapper multipath, do *not* specify Linux page caching (the **-c** option of the **gnbd_export** command). All GNBDs that are part of a logical volume must run with caching *disabled*. Data corruption occurs if the GNBDs are run with caching enabled. Refer to [第 3.1 节“从服务器中导出 GNBD”](#) for more information about using the **gnbd_export** command for GNBD with device-mapper multipath.

2.2. 栅栏 GNBD 服务器节点

必须使用防护法物理删除来自网络的节点对 GNBD 服务器节点进行防护。要物理删除 GNBD 服务器节点，您可以使用栅栏设备：以下除外：**fence_brocade** 防护代理、**fence_vixel** 防护代理、**fence_mcddata** 防护代理、**fence_sanbox2** 防护代理、**fence_scsi** 防护代理。另外，您不能使用 GNBD 防护设备（**fence_gnbd** 防护代理）来防护 GNB 服务器节点。有关为 GNB 的服务器节点配置防护的详情请参考《全局文件系统》手册。

第 3 章 GNBD 驱动程序及命令用法

全局网络块设备（GNBD）驱动程序允许节点作为一个 GNBD 通过网络导出其本地存储，以便网络中的其它节点可共享此存储。导入 GNBD 的客户端节点可象使用其它块设备一样使用它。在多客户端中导入 GNBD 可形成共享存储配置，通过该共享存储配置可使用 GFS。

GNBD 驱动程序是通过以下组件实施的。

- ▶ **gnbd_serv** — 部署 GNBD 服务器。它是一个用户空间守护进程，允许节点通过网络导出本地存储。
- ▶ **gnbd.ko** — 在 GNBD 客户端（使用 GNBD 设备的节点）中部署 GNBD 设备驱动程序。

可用来配置 GNBD 的两个命令：

- ▶ **gnbd_export** (for servers) — User program for creating, exporting, and managing GNBDs on a GNBD server.
- ▶ **gnbd_import** (用于客户端) — 在 GNBD 客户端中导入和管理 GNBD 的用户程序。

3.1. 从服务器中导出 GNBD

在某个节点可作为 GNBD 导出存储前必须在其中运行 **gnbd_serv** 守护进程。您可以按以下方法运行 **gnbd_serv** 启动 **gnbd_serv**：

```
#gnbd_serv
gnbd_serv: startup succeeded
```

确定要导出的本地存储后，可使用 **gnbd_export** 命令导出它。

备注

当您使用设备映射器多路径配置 GNBD 服务器时，您一定不能使用页面缓存。所有是本地卷一部分的 GNBD 必须在禁用缓存的情况下运行。默认情况是 **gnbd_export** 命令导出时缓存是关闭的。

注意

服务器不应该导入 GNBD 使用其作为客户端。如果服务器导出的设备是没有缓冲的，**gfs** 也可以使用基本设备。

用法

```
gnbd_export -d pathname -e gnbdname [-c][-u][-U]
```

pathname

指定要导出的存储设备。

gnbdname

为 GNBD 选择任意指定的名称。它可作为 GNBD 客户端中的设备名称。这个名字必须在网络中导出的所有 GNBD 中是唯一的。

-o

以只读方式导出该设备。

-c

启用缓冲。从导出的 GNBD 读取并利用 Linux 页面缓存。

默认情况下，`gnbd_export` 命令不启用缓存。

**备注**

当您使用设备映射器多路径配置 GNBD 服务器时，不要指定 `-c` 选项。所有是逻辑卷一部分的 GNBD 必须在禁用缓存的情况下运行。

**备注**

如果您已经使用了 GFS 5.2 或者早期的版本且不想要更改您的 GNBD 设置，您应该指定 `-c` 选项。在 GFS 发行本 5.2.1 之前的版本中，是默认为 `gnbd_export` 启用 Linux 缓存。如果没有指定 `-c` 选项，那么 GNBD 的性能会显著下降。同时，如果没有指定 `-c` 选项，导出的 GNBD 会出现超时，使用默认超时值（`-t` 选项）。有关 `gnbd_export` 命令的详情及其选项，请参考 `gnbd_export man page`。

-u uid

为导出的设备手动设定全局识别符。这个选项和 `-e` 一同使用。UID 是设备映射器多路径来决定哪些设备属于多路径映射。一个设备必须有一个用于多路径的 UID。但是，对大多数 SCSI 设备来说，默认获得 UID 的命令 `/usr/sbin/gnbd_get_uid` 将返回适当的值。

**备注**

UID 指向要导出的设备，而不是 GNBD 本身。只有在导出同一个基本设备时，两个 GNBD 设备的 UID 才应该是相等的。这就是说两个 GNBD 服务器连接到同一个物理设备。

**警告**

当 `-U command` 选项不工作时，这个选项应该只用于导出共享存储设备。这在 SCSI 设备中基本不会发生。如果两个 GNBD 设备没有导出同一基本设备，但给出相同的 UID，则会发生数据崩溃。

-U Command

获得 UID 的命令。UID 命令是一个为导出的设备获得统一识别符的 `gnbd_export` 命令。该 UID 是使用带 GNBD 的设备映射器多路径所必须的。该命令必须使用您要运行的任意可执行全路径。一个命令中可包含 `%M`、`%m` 或者 `%n` 换码序列。`%M` 可扩展为导出设备的主号码，`%m` 可扩展为导出设备的副号码，`%n` 可扩展为该设备的 `sysfs` 名称。如果没有给出命令，GNBD 将默认使用命令 `/usr/sbin/gnbd_get_uid`。这个命令对大多数 SCSI 设备都适用。

示例

这个示例是配置了 GNBD 多路径的 GNBD 服务器。它将设备 `/dev/sdc2` 导出为 `gamma`。默认禁用缓冲。

```
gnbd_export -d /dev/sdc2 -e gamma -U
```

这个示例是没有配置 GNBD 多路径的 GNBD 服务器。它将 `/dev/sdb2` 设备导出为 GNBD `delta`，并禁用缓冲。

```
gnbd_export -d /dev/sdb1 -e delta -c
```

这个示例将 `/dev/sdb2` 设备导出为 GNBD `delta`，并启用缓冲。

```
gnbd_export -d /dev/sdb2 -e delta -c
```

3.2. 在客户端中导入 GNBD

在节点可以导入 GNBD 前必须在其中载入 `gnbd.ko` 内核模块。当导入 GNBD 时，会在 `/dev/gnbd/` 中为其生成设备节点，并分配给它们导出时的名称。

用法

```
gnbd_import -i Server
```

Server

使用导入 GNBD 的主机名和 IP 地址指定 GNBD 服务器。所有从该服务器中导出的 GNBD 将被导入到运行此命令的客户端。

示例

这个示例导入所有来自名为 `nodeA` 服务器的 GNBD。

```
gnbd_import -i nodeA
```

第 4 章 在 GNBD 服务器节点中运行 GFS

您可以在 GNBD 服务器节点中运行 GFS，但有一定限制。另外，在 GNBD 服务器节点中运行 GFS 可降低性能。以下是在 GNBD 服务器节点中运行 GFS 时的限制。



重要

当在 GNBD 服务器节点中运行 GFS 时，您必须遵循以下限制，否则 GNBD 服务器节点会失败。

1. GNBD 服务器节点必须有到所有需要挂载 GFS 文件系统的所有设备的本地访问。GNBD 服务器节点一定不能导入 (`gnbd_import` 命令) GNBD 以外的设备来运行此文件系统。
2. GNBD 服务器必须在无换出模式下导出所有 GNBD，它必须导出所有原始设备，而不是逻辑卷设备。
3. GFS 必须在逻辑卷设备顶部运行，而不是在原始设备中运行。



备注

您需要在导出的 GNBD 中增加超时区间来适应降低的性能。需要增加的超时区间要视硬件情况而定。

修订记录

修订 3-7.400	2013-10-31	Rüdiger Landmann
Rebuild with publican 4.0.0		
修订 3-7	2012-07-18	Anthony Towns
Rebuild for Publican 3.0		
修订 1.0-0	Thu Jan 29 2009	

索引

D

device-mapper multipath, [考虑在设备映射器多路径中使用 GNBD](#)

- fencing GNBD server nodes, [栅栏 GNBD 服务器节点](#)
- Linux page caching, [Linux 页面缓存](#)

driver and command usage, [GNBD 驱动程序及命令用法](#)

- exporting from a server, [从服务器中导出 GNBD](#)
- importing on a client, [在客户端中导入 GNBD](#)

E

exporting from a server daemon, [从服务器中导出 GNBD](#)

F

feedback, [反馈](#)

fencing GNBD server nodes, [栅栏 GNBD 服务器节点](#)

G

GFS, using on a GNBD server node, [在 GNBD 服务器节点中运行 GFS](#)

GNBD, using with Red Hat GFS, [使用带 Red Hat GFS 的 GNBD](#)

gnbd.ko module, [GNBD 驱动程序及命令用法](#), [在客户端中导入 GNBD](#)

gnbd_export command, [GNBD 驱动程序及命令用法](#), [用法](#)

gnbd_import command, [GNBD 驱动程序及命令用法](#), [用法](#)

gnbd_serv daemon, [GNBD 驱动程序及命令用法](#), [从服务器中导出 GNBD](#)

I

importing on a client module, [在客户端中导入 GNBD](#)

L

Linux page caching, [Linux 页面缓存](#)

S

software subsystem components, [使用带 Red Hat GFS 的 GNBD](#)