SUSE Linux Enterprise High Availability Extension

11 www.novell.com

2009 年7 月31 日 High Availability 指南

High Availability 指南

目录

关于本指南

SUSE® Linux Enterprise High Availability Extension 是一个开放源代码群集技术的集成套件，使您能够实现高度可用的物理和虚拟 Linux 群集。为快速且有效地进行配置和管理，High Availability Extension 包括了一个图形用户界面 (GUI) 和一个命令行界面 (CLI)。

本指南适用于需要设置、配置和维护 High Availability (HA) 群集的管理员。它详细介绍了这两种界面（GUI 和 CLI），以帮助管理员选择与执行关键任务的需要匹配的相应工具。

本指南分为以下部分：

安装和设置
: 在开始安装和配置群集前，先熟悉群集的基本原理和体系结构，并概览关键功能和优点及在上一版本基础上的更改。了解必须满足的硬件和软件要求及执行以下步骤前要做的准备。使用 YaST 执行 HA 群集的安装和基本设置。

配置和管理
: 使用 GUI 或 `crm` 命令行界面添加、配置和管理资源。了解如何利用负载平衡和屏障。如果您考虑编写自己的资源代理或修改现有的资源代理，请了解一些有关如何创建不同类型的资源代理的背景信息。

储存和数据复制
: SUSE Linux Enterprise High Availability Extension 提供了群集感知的文件系统（Oracle 群集文件系统，OCFS2）和卷管理器（群集的逻辑卷管理器，cLVM）。High Availability Extension 还提供了 DRBD（分布式复制块设备）用于数据复制，您可以使用 DRBD 将高度可用的服务的数据从群集的活动节点镜像到备用节点。

查错和参考
: 管理群集需要执行一些查错操作。了解最常见的故障及解决方法。了解 High Availability Extension 提供的命令行工具的全面参考以便管理群集。还列出了有关群集资源和资源代理的最重要的信息。

本手册中的许多章节包含到附加文档资源的链接。 这包括系统上提供的附加文档以及因特网上提供的文档。

有关该产品可用文档的概述和最新文档更新，请参见 http://www.novell.com/documentation。

1 反馈

提供了多种反馈渠道：

- 要报告产品组件的 bug 或提交增强请求，请使用 https://bugzilla.novell.com/。如果您刚开始使用 Bugzilla，您会发现 *bug 书写常见问题* 可能很有用，该功能可以从 Novell Bugzilla 主页中找到。
- 我们希望听到您对本手册和本产品中包含的其他文档的意见和建议。请使用每页联机文档底部的用户意见功能并发表您的意见。

2 文档约定

以下是本手册中使用的版式约定：

- `/etc/passwd`：目录名称和文件名
- *`placeholder`*：将 *`placeholder`* 替换为实际值
- `PATH`：环境变量 PATH
- `ls`、`--help`：命令、选项和参数
- `user`：用户和组
- Alt、Alt + F1：按键或组合键；这些键以大写形式显示，如在键盘上一样
- *文件*、*文件* > *另存为*：菜单项，按钮
- 本段只与指定的体系结构有关。 箭头标记文本块的开始位置和结束位置。

  本段只与指定的体系结构有关。 箭头标记文本块的开始位置和结束位置。
- *跳舞的企鹅*（*企鹅*一章，↑其他手册）：这是对其他手册中的某章的参考。

部分 I. 安装和设置

1 概念概述

SUSE® Linux Enterprise High Availability Extension 是一个开放源代码群集技术的集成套件，使您能够实现高度可用的物理和虚拟 Linux 群集，并排除单一故障点。它确保了关键网络资源（包括数据、应用程序和服务）的高可用性和可管理性。因此，它有助于维持业务连续性、保护数据完整性及减少 Linux 关键任务工作负荷的计划外停机时间。

它提供了基本的监视、消息交换和群集资源管理功能，支持分别管理的群集资源的故障转移、故障回复和迁移（负载平衡）。High Availability Extension 作为 SUSE Linux Enterprise Server 11 的外接式附件提供。

1.1 产品功能

SUSE® Linux Enterprise High Availability Extension 有助于确保和管理网络资源的可用性。以下列表重点描述了某些重要功能：

支持多种群集方案

包括主动/主动和主动/被动（N+1、N+M、N 到 1、N 到 M）方案，及混合的物理和虚拟群集（允许虚拟服务器与物理服务器成为群集，以改善服务和资源的可用性）。

多节点主动群集，最多可包含 16 台 Linux 服务器。如果群集内的一台服务器发生故障，则群集内的任何其他服务器均可重启动此服务器上的资源（应用程序、服务、IP 地址和文件系统）。

灵活的解决方案

High Availability Extension 提供了 OpenAIS 消息交换和成员资格层以及 Pacemaker 群集资源管理器。使用 Pacemaker，管理员可以持续监视资源的状态，管理依赖性，并基于可配置程度很高的规则和策略自动停止和启动服务。High Availability Extension 允许您将群集调整为特定的应用程序和硬件基础结构，以适应您的组织。基于时间的配置使服务可以在特定时间自动迁移回已修复的节点。

储存和数据复制

使用 High Availability Extension 可以根据需要动态地指派和重指派服务器储存。它支持光纤通道或 iSCSI 储存区域网络 (SAN)。它还支持共享磁盘系统，但这不是必需的。SUSE Linux Enterprise High Availability Extension 还附带了群集感知的文件系统（Oracle 群集文件系统，OCFS2）和卷管理器（群集的逻辑卷管理器，cLVM）。High Availability Extension 还提供了 DRBD（分布式复制块设备）用于数据复制，您可以使用DRBD将高度可用的服务的数据从群集的活动节点镜像到备用节点。

支持虚拟环境

SUSE Linux Enterprise High Availability Extension 支持物理和虚拟 Linux 服务器的混合群集。SUSE Linux Enterprise Server 11 提供了 Xen，一款开放源代码的虚拟化超级管理程序。High Availability Extension 中的群集资源管理器可以识别、监视和管理运行于使用 Xen 创建的虚拟服务器中的服务，及运行在物理服务器中的服务。Guest 系统可作为服务由群集管理。

资源代理

SUSE Linux Enterprise High Availability Extension 包括了大量资源代理以管理资源，例如 Apache、IPv4、IPv6 等等。它还为通用的第三方应用程序（例如 IBM WebSphere Application Server）提供了资源代理。有关产品包含的开放群集框架 (OCF) 资源代理的列表，请参见第 18 章 *HA OCF 代理* [189]。最新的日期列表可从 www.novell.com/products/highavailability 获得。

用户友好的管理

为了方便配置和管理，High Availability Extension 提供了一个图形用户界面（如 YaST 和 Linux HA Management Client）和一个强大的统一命令行界面。这两个界面都提供了从一个位置即可有效监视和管理群集的管理方式。在以下章节中可了解如何操作。

1.2 本产品的优点

High Availability Extension 允许将最多 16 台 Linux 服务器配置为一个高度可用的群集（HA 群集），在群集中可以将资源动态地切换或移动到任何服务器上。可以将资源配置为自动迁移以防服务器故障，或手动移动资源以对硬件查错或平衡工作负荷。

High Availability Extension 通过商品组件提供了高度可用性。通过将应用程序和操作合并到群集中降低了成本。High Availability Extension 还允许集中管理整个群集并调整资源以满足变化的工作负荷要求（这样就手动地使群集“负载平衡”了）。允许群集的多个（两个以上）节点共享一个“热备份”也节约了成本。

一个同样重要的好处是潜在地减少了计划外服务中断及用于软件和硬件维护和升级的计划内中断。

实施群集的理由包括：

- 提高可用性
- 改善性能
- 降低操作成本
- 可伸缩性
- 灾难恢复
- 数据保护
- 服务器合并
- 储存合并

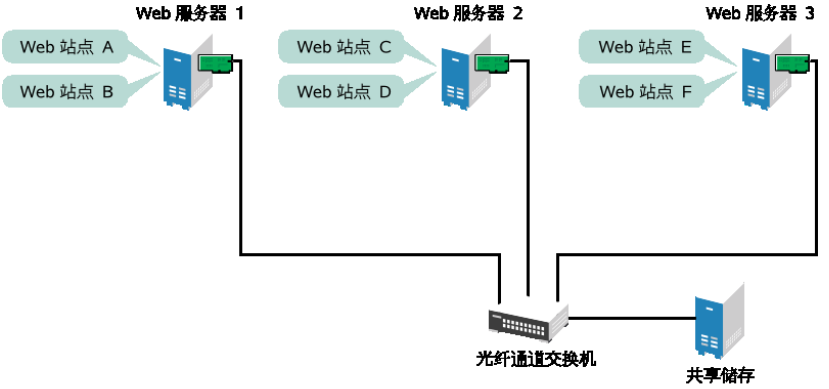
通过在共享磁盘子系统上实施 RAID 可获得共享磁盘容错。

以下方案说明了 High Availability Extension 提供的一些好处。

示例群集方案

假设您配置了一个包含三台服务器的群集，并在群集内的每台服务器上安装了 Web 服务器。群集内的每台服务器都主管两个 Web 站点。每个 Web 站点的全部数据、图形和 Web 页面内容都储存在一个连接到群集中每台服务器的共享磁盘子系统上。下图说明了该系统的结构。

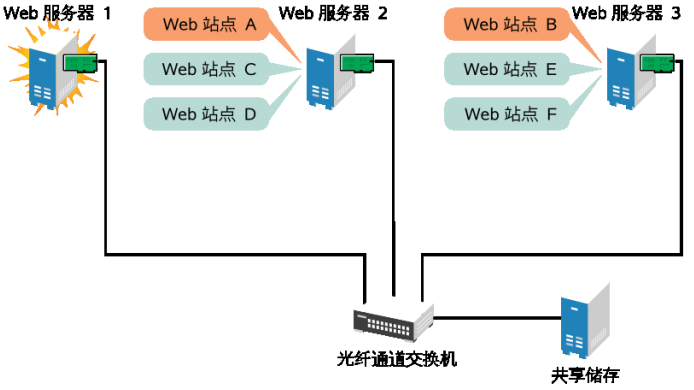
图 1.1　三台服务器的群集



在群集的正常工作状态下，每台服务器都与群集内的其它服务器持续通讯，并对所有已注册的资源进行定期巡回检测以检测故障。

假设 Web 服务器 1 出现硬件或软件故障，而依赖此 Web 服务器访问因特网、收发电子邮件和获取信息的用户失去了连接。下图说明了当万维网服务器 1 出现故障时，资源的移动情况。

图 1.2　三台服务器的群集（其中一台服务器出现故障后）



Web 站点 A 移至 Web 服务器 2，Web 站点 B 移至 Web 服务器 3。IP 地址和证书也移至 Web 服务器 2 和 Web 服务器 3。

在配置群集时，您决定了在出现故障的情况下，每台 Web 服务器上的 Web 站点将移至哪里。在上例中，您已配置将 Web 站点 A 移至 Web 服务器 2，将 Web 站点 B 移至 Web 服务器 3。这样，曾由 Web 服务器 1 处理的工作负荷继续存在且平均分配给剩余的群集成员。

当 Web 服务器 1 出现故障，High Availability Extension 软件

- 检测到故障，并与 STONITH 确认 Web 服务器 1 确实已出现故障
- 将以前安装在 Web 服务器 1 上的共享数据目录重新安装在 Web 服务器 2 和 Web 服务器 3 上。
- 在 Web 服务器 2 和 Web 服务器 3 上重启动以前运行于 Web 服务器 1 上的应用程序
- 将 IP 地址传送到 Web 服务器 2 和 Web 服务器 3

在此示例中，故障转移过程迅速完成，用户在几秒钟之内就可以重新访问 Web 站点信息，而且在多数情况下无需重新登录。

现在，假设 Web 服务器 1 的故障已解决，它已恢复到正常工作状态。Web 站点 A 和 Web 站点 B 可以自动故障回复（移回）至 Web 服务器 1，或者留在当前所在的服务器上。这取决于您是如何配置它们的资源的。将服务迁移回 Web 服务

器 1 将导致一段中断期，因此 High Availability Extension 也允许您将迁移推迟到某个将极少或不会造成服务中断的时段。这两种选择都各有优缺点。

High Availability Extension 也提供了资源迁移功能。可以根据系统管理的需要将应用程序、Web 站点等资源移动到群集中的其他服务器。

例如，您可以手动将 Web 站点 A 或 Web 站点 B 从 Web 服务器 1 移至群集内的其他任何一台服务器。对万维网服务器 1 进行升级或定期维护时，或者只是要提高万维网站点的性能或可访问性，都需要执行此操作。

1.3 群集配置

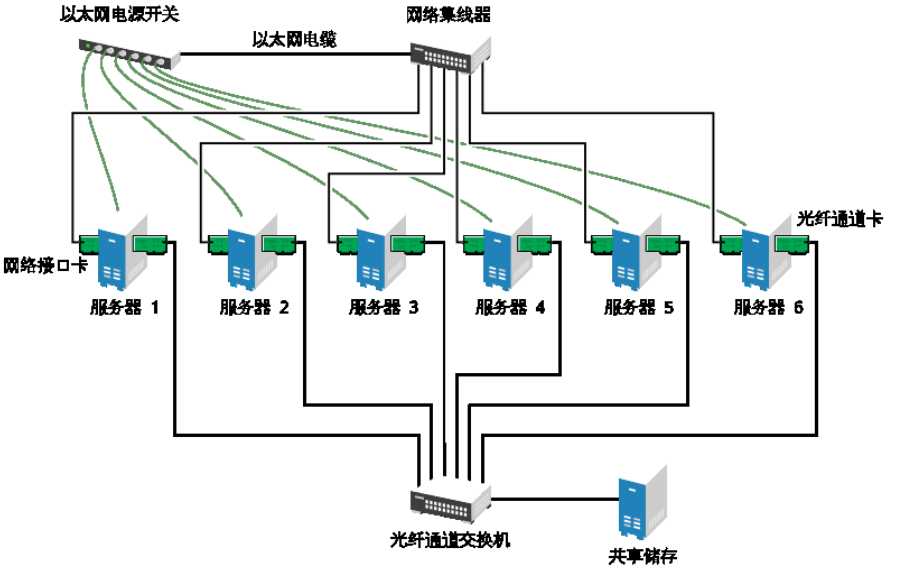
High Availability Extension 的群集配置可能包括或不包括共享磁盘子系统。共享磁盘子系统可通过高速光纤通道卡、电缆和交换机连接，也可配置为使用iSCSI。如果一台服务器出现故障，群集内的另一台指定服务器将自动装入以前安装在此有故障服务器上的共享磁盘目录。这样，网络用户就能继续访问共享磁盘子系统上的目录。

重要：带 cLVM 的共享磁盘子系统

使用带 cLVM 的共享磁盘子系统时，此子系统必须连接到群集中所有需要访问此子系统的服务器。

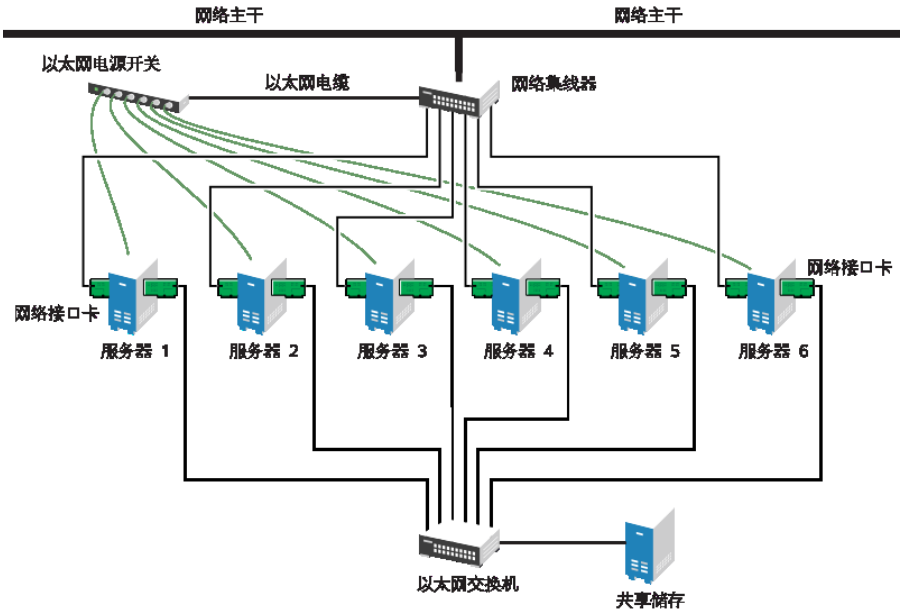
典型的资源包括数据、应用程序和服务。下图显示了一个典型的光纤通道群集配置的结构。

图 1.3　典型的光纤通道群集配置



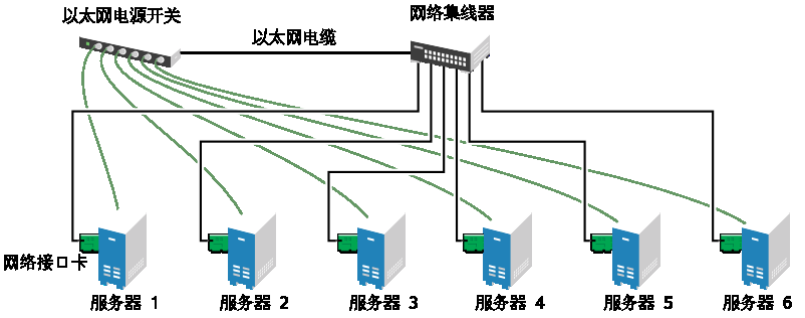
虽然光纤通道提供的性能最佳，但也可以将群集配置为使用 iSCSI。iSCSI 是除光纤通道外的另一种选择，可用于创建低成本的储存区域网络 (SAN)。下图显示了一个典型的 iSCSI 群集配置。

图 1.4　典型的 iSCSI 群集配置



虽然多数群集都包括共享磁盘子系统，但也可以创建不含共享磁盘子系统的群集。下图显示了一个不含共享磁盘子系统的群集。

图 1.5　典型的不含共享储存的群集配置



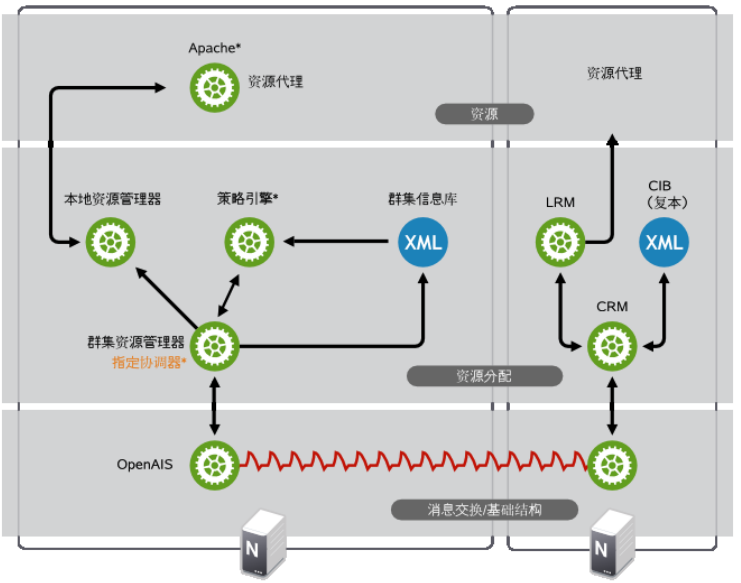
1.4 体系结构

本部分简要地概述了 High Availability Extension 体系结构。它提供了有关体系结构组件的信息，并描述了这些组件是如何协同工作的。

1.4.1 体系结构层

High Availability Extension 具有分层的体系结构。图 1.6 “体系结构” [11]说明了不同的层及其相关的组件。

图 1.6 体系结构



消息交换和基础结构层

主层或第一层是消息交换/基础结构层，也称为 OpenAIS 层。此层包含了发送含有“我在线”信号的消息及其他信息的组件。High Availability Extension 的程序就位于此消息交换/基础结构层。

资源分配层

下一层是资源分配层。此层最复杂，它包含以下组件：

群集资源管理器 (CRM)
: 在资源分配层中执行的每个操作都要经过群集资源管理器。如果资源分配层的其他组件（或更高层中的组件）需要通讯，则它们通过本地 CRM 进行。

  在每个节点上，CRM 维护群集信息库 (CIB) [12]，包含所有群集选项、节点、资源及其关系和当前状态的定义。如果选择群集中的 CRM 为指定协调程序 (DC)，则意味着它具有主 CIB。群集中的所有其他 CIB 是此主 CIB 的复本。对 CIB 的常规读写操作通过主 CIB 进行排序。DC 是群集中唯一可以决定需要在整个群集执行更改（例如节点屏障或资源移动）的实体。

群集信息库 (CIB)
: 群集信息库是整个群集配置和当前状态在内存中的 XML 表示。它包含所有群集选项、节点、资源、约束及其之间的关系的定义。CIB 还将更新同步到所有群集节点。群集中有一个主 CIB，由 DC 维护。所有其他节点包含一个 CIB 复本。

策略引擎 (PE)
: 每当指定协调程序需要进行整个群集的更改（对新 CIB 做出反应），策略引擎就会根据群集的当前状态和配置计算群集的下一个状态。PE 还生成一个转换图，包含用于达到下一个群集状态的（资源）操作和依赖性的列表。PE 在每个节点上都运行以加速 DC 故障转移。

本地资源管理器 (LRM)
: LRM 代表 CRM 调用本地资源代理（请参见“资源层”一节 [13]）。因此它可以执行启动/停止/监视操作并将结果报告给 CRM。它还隐藏资源代理支持的脚本标准（OCF、LSB、Heartbeat V1）之间的区别。LRM 是其本地节点上所有资源相关信息的权威来源。

资源层

最高层是资源层。资源层包括一个或多个资源代理 (RA)。资源代理是为启动、停止和监视某种服务（资源）而编写的程序，通常是壳层脚本。资源代理仅由 LRM 调用。第三方可将他们自己的代理放在文件系统中定义的位置，这样就为各自的软件提供了现成群集集成。

1.4.2 处理流程

SUSE Linux Enterprise High Availability Extension 使用 Pacemaker 作为 CRM。CRM 作为守护程序执行 (`crmd`)，它在每个群集节点上都有一个实例。Pacemaker 通过将某个 crmd 实例选为主实例，从而集中了所有的群集决策制定。如果选定的 crmd 过程（或它所在的节点）出现故障，则将建立一个新的过程。

在每个节点上保留了一个 CIB，它反映了群集的配置和群集中所有资源的当前状态。CIB 的内容会在整个群集的同步过程中自动保留下来。

群集中执行的许多操作都将导致整个群集的更改。这些操作包括添加或删除群集资源、更改资源约束等等。了解执行这样的操作时群集中会发生的状况是很重要的。

例如，假设您要添加一个群集 IP 地址资源。为此，您可以使用一种命令行工具或 GUI 修改 CIB。您不必在 DC 上执行此操作，可以使用群集中任何节点上的任何工具，此操作会被传送到 DC 上。然后 DC 将把此 CIB 更改复制到所有群集节点。

根据 CIB 中的信息，PE 便计算群集的理想状态及如何达到此状态，并将指令列表传递给 DC。DC 通过消息交换/基础结构层发出命令，其他节点上的 crmd 同级将收到此命令。每个 crmd 使用它的 LRM（作为 lrmd 实现）执行资源修改。lrmd 不是群集感知的，它直接与资源代理（脚本）交互。

所有同级节点将操作的结果报告给 DC。一旦 DC 做出所有必需操作已在群集中成功执行的结论，群集将返回至空闲状态并等待进一步事件。如果有操作未按计划执行，则会再次调用 PE，CIB 中将记录新信息。

在某些情况下，可能需要关闭节点以保护共享数据或完成资源恢复。为此，Pacemaker 附带了一个屏障子系统，stonithd。STONITH 是 “Shoot The Other Node In The Head（关闭其他节点）”的首字母缩写，通常通过一个远程电源开关实施。在 Pacemaker 中将 STONITH 设备构造成资源（并在 CIB 中配置）以便对

它们监视故障；然而，stonithd 负责了解 STONITH 拓扑，这样它的客户端只需请求屏障节点，余下的工作由它来完成。

1.5 新功能

SUSE Linux Enterprise Server 11 的群集堆栈已从 Heartbeat 更改为 OpenAIS。OpenAIS 实施行业标准 API，应用程序界面规范 (AIS)，由服务可用性论坛 (Service Availability Forum) 发布。SUSE Linux Enterprise Server 10 的群集资源管理器得以保留但有了显著增强，它已转换为 OpenAIS 且现在称为 Pacemaker。

有关 High Availability 组件从 SUSE® Linux Enterprise Server 10 SP2 到 SUSE Linux Enterprise High Availability Extension 11 的更改的更多细节，请参见以下部分。

1.5.1 新增功能

迁移阈值和故障超时
: 现在 High Availability Extension 附带了迁移阈值和故障超时的概念。可以对资源定义一个故障数量，达到此数量后资源将迁移到新节点。默认情况下，将不再允许节点运行出现故障的资源，直到管理员手动重设置资源的故障计数。但也可以通过设置资源的 `failure-timeout` 选项来使资源失效。

资源和操作默认值
: 现在可以设置资源选项和操作的全局默认值。

支持脱机配置更改
: 通常在详细更新配置前最好预览一下一系列更改的效果。现在，执行配置从而详细更改活动群集配置前，可以创建配置的“阴影”副本并用命令行界面进行编辑。

重用规则、选项和操作集
: 可以定义一次规则、实例属性、元属性和操作集，然后在多处参考。

对 CIB 中的某些操作使用 XPath 表达式
: 现在 CIB 接受基于 XPath 的 `create`、`modify`、`delete`操作。有关更多信息，请参见 `cibadmin` 帮助文本。

多维排列和排序约束
: 为创建一个排列资源集，以前可以定义一个资源组（无法总是准确地表达设计意图）或将每个关系定义为单独的约束，导致约束随着资源和组合的数量增长而激增。现在还可以使用排列约束的另一种形式，即定义 `resource_sets`。

从非群集的服务器连接到 CIB
: 如果服务器上安装了 Pacemaker，则即使服务器本身不是群集的一部分，也可以连接到群集。

在已知时间触发重现操作
: 默认情况下，重现操作是根据资源启动的时间来计划的，但这并不总令人满意。要指定操作应根据的日期/时间，请设置操作的间隔-起始时间。群集使用此时间计算正确的启动-延迟，这样操作将在起始时间 + （间隔 * N）时发生。

1.5.2 变更功能

资源和群集选项的命名约定
: 现在所有资源和群集选项都使用连字符 (-) 代替下划线 (_)。例如，`master_max` 元选项已重命名为 `master-max`。

重命名 `master_slave` 资源
: `master_slave` 资源已重命名为 `master`。主资源是一种特殊类型的克隆，可按两种模式之一运行。

属性的容器标记
: `attributes` 容器标记已删除。

先决条件的操作字段
: `pre-req` 操作字段已重命名为 `requires`。

操作间隔
: 所有操作都必须有间隔。对于启动/停止操作，间隔必须设置为 `0`。

排列和排序约束的属性
: 为了清晰起见，已重命名排列和排序约束的属性。

因故障而迁移的群集选项
: `resource-failure-stickiness` 群集选项已替换为 `migration-threshold` 群集选项。另请参见迁移阈值和故障超时 [14]。

命令行工具的自变量
: 已使命令行工具的自变量保持一致。另请参阅资源和群集选项的命名约定 [15]。

验证和分析 XML
: 群集配置是用 XML 编写的。现在，一种更强大的 RELAX-NG 纲要已取代文档类型定义 (DTD)，用于定义结构和内容的模式。`libxml2` 用作分析器。

`id` 字段
: 现在 `id` 字段是 XML ID，它有以下限制：

  - ID 不能包含冒号。
  - ID 不能以数字开始。
  - ID 必须是全局唯一的（不只是对标记唯一）。

参考其他对象
: 某些字段（例如那些参考资源的约束中的字段）是 IDREF。这意味着它们必须参考现有资源或对象，以使配置有效。无法删除在别处作为参考的对象。

1.5.3 删除功能

设置资源元选项
: 无法再将资源元选项设置为顶级属性。改为使用元属性。

设置全局默认值
: 不再从 `crm_config` 读取资源和操作默认值。

2 入门

在以下部分中，可了解系统要求及安装 High Availability Extension 前要做的准备。大致概览安装和设置群集的基本步骤。

2.1 硬件要求

以下列表指定了 SUSE® Linux Enterprise High Availability Extension 的群集的硬件要求。这些要求表示最低硬件配置。根据群集的用途，可能会需要其他硬件。

- 安装了第 2.2 节 “软件要求” [18]中指定的软件的 1 到 16 台 Linux 服务器。服务器的硬件（内存、磁盘空间等等）无需相同。
- 至少两个 TCP/IP 通讯媒体。群集节点使用多路广播进行通讯，所以网络设备必须支持多路广播。通讯媒体应支持 100 Mbit/s 或更高的数据传送速度。最好应绑定以太网通道。
- 可选：一个共享磁盘子系统，连接到群集中所有需要访问它的服务器。
- 一个 STONITH 机制。STONITH 是 “Shoot the other node in the head（关闭其他节点）”的首字母缩写。STONITH 设备是一个电源开关，用于群集重设置被认为出现故障的节点。重设置没有检测信号的节点是确保存在但出现故障的节点未执行数据破坏的唯一可靠方法。

  有关更多信息，请参考第 8 章 *屏障和 STONITH* [75]。

2.2 软件要求

请确保满足以下软件要求：

- 在所有将成为群集组成部分的节点上安装了包含全部可用联机更新的SUSE® Linux Enterprise Server 11。
- 在所有将成为群集组成部分的节点上安装了包含全部可用联机更新的 SUSE Linux Enterprise High Availability Extension 11。

2.3 共享磁盘系统需求

如果要使数据高度可用，建议为群集使用共享磁盘系统（储存区域网络或 SAN）。如果使用共享磁盘子系统，请确保符合以下要求：

- 根据制造商的说明正确设置共享磁盘系统并且共享磁盘系统可正确运行。
- 共享磁盘系统中包含的磁盘应配置为使用镜像或 RAID，来为共享磁盘系统增加容错性。建议使用基于硬件的 RAID。所有配置都不支持基于主机的软件 RAID。
- 如果准备对共享磁盘系统访问使用 iSCSI，则请确保正确配置了 iSCSI 启动器和目标。
- 使用 DRBD 实现在两台服务器间分发数据的镜像 RAID 系统时，请确保仅访问复制的设备。请与群集的剩余节点使用相同（绑定）的 NIC 来调整此处提供的冗余。

2.4 准备工作

请在安装前执行以下准备步骤：

- 通过在群集中的每台服务器上编辑 `/etc/hosts` 文件来配置主机名解析并使用静态主机信息。有关更多信息，请参见 `http://www.novell.com/documentation` 提供的 *SUSE Linux Enterprise Server 管理指南* 的*基本网络*一章的*配置主机名和 DNS* 部分。

群集的成员之间可按名称查找对方是基本的。如果名称不可用，则将无法进行群集内部通讯。

- 通过使群集节点与群集外的时间服务器同步来配置时间同步。有关更多信息，请参见 http://www.novell.com/documentation 提供的 *SUSE Linux Enterprise Server 管理指南* 的*与 NTP 同步时间*一章。

  群集节点将使用此时间服务器作为它们的时间同步源。

2.5 概述：安装和设置群集

准备工作完成后，以下步骤是用 SUSE® Linux Enterprise High Availability Extension 安装和设置群集所必需的：

1. 将 SUSE® Linux Enterprise Server 11 和 SUSE® Linux Enterprise High Availability Extension 11 作为外接式附件安装在 SUSE Linux Enterprise Server 之上。有关详细信息，请参见第 3.1 节 “安装 High Availability Extension” [21]。

2. 配置 OpenAIS。有关详细信息，请参见第 3.2 节 “初始群集设置” [22]。

3. 启动 OpenAIS 并监视群集状态。有关详细信息，请参见第 3.3 节 “使群集联机” [24]。

4. 用图形用户界面 (GUI) 或命令行添加和配置群集资源。有关详细信息，请参见第 4 章 *使用 GUI 配置群集资源* [29]或第 5 章 *通过命令行配置群集资源* [55]。

   为保护数据免于屏障和 STONITH 可能造成的破坏，请确保将 STONITH 设备配置为资源。有关详细信息，请参见第 8 章 *屏障和 STONITH* [75]。

   可能还需要在共享磁盘（储存区域网络，SAN）上创建文件系统（如果尚未存在）并将这些文件系统配置为群集资源（如果需要）。

   群集感知 (OCFS 2) 和非群集感知的文件系统都可以用 High Availability Extension 配置。如果需要，还可以通过 DRBD 使用数据复制。有关详细信息，请参见第 III 部分 “储存和数据复制” [99]。

3 用 YaST 进行安装和基本设置

有多种方式安装 High Availability 群集所需的软件：通过 `zypper` 用命令行安装，或通过 YaST 用图形用户界面安装。在所有将成为群集组成部分的节点上安装软件后，下一步是初始配置群集以使节点之间可以通讯。此步骤可以手动完成（通过编辑配置文件）或用 YaST 群集模块完成。

注意：安装软件包

High Availability 群集所需的软件包不会自动复制到群集节点。在所有将成为群集组成部分的节点上安装 SUSE® Linux Enterprise Server 11 和 SUSE® Linux Enterprise High Availability Extension 11。

3.1 安装 High Availability Extension

用 High Availability Extension 配置和管理群集所需的包在高可用性安装模式中。仅在将 SUSE® Linux Enterprise High Availability Extension 作为外接式附件安装后才提供此模式。有关如何安装外接式附件产品的信息，请参见 SUSE Linux Enterprise 11 *部署指南* 的*安装外接式附件产品*一章。

1. 启动 YaST 并选择*软件* > *软件管理*打开 YaST 包管理器。
2. 从*过滤器*列表中选择*模式*，然后在模式列表中激活*高可用性*模式。
3. 单击*接受*开始安装包。

3.2 初始群集设置

安装 HA 包后，可以用 YaST 配置初始群集设置。它包括节点间的通讯通道、安全性方面（如使用加密通讯和启动 OpenAIS 作为服务）。

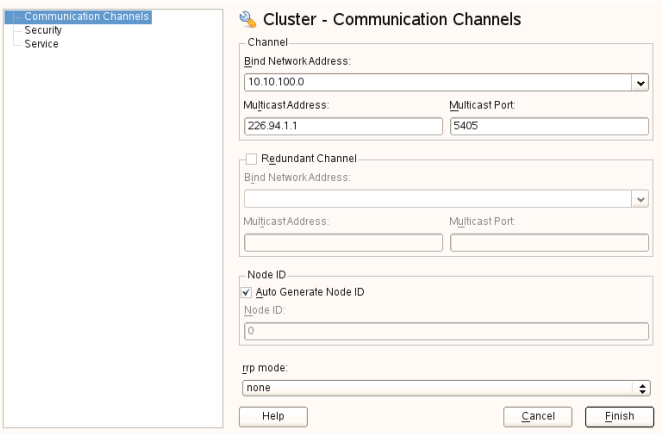
对于通讯通道，需要定义一个绑定网络地址 (`bindnetaddr`)、一个多路广播地址 (`mcastaddr`) 和一个多路广播端口 (`mcastport`)。`bindnetaddr` 是要绑定的网络地址。为方便在群集间共享配置文件，OpenAIS 使用网络界面网络掩码来屏蔽仅用于路由网络的地址位。`mcastaddr` 可以是 IPv4 或 IPv6 多路广播地址。`mcastport` 是为 `mcastaddr` 指定的 UDP 端口。

群集中的节点通过使用同一多路广播地址和同一端口号来相互了解。对于不同的群集，请使用不同的多路广播地址。

过程 3.1 *配置群集*

1. 启动 YaST 并选择*杂项* > *群集*或在命令行中运行 `yast2 cluster` 启动初始群集配置对话框。

2. 在*通讯通道*类别中，配置用于群集节点间通讯的通道。此信息会写入 `/etc/ais/openais.conf` 配置文件。

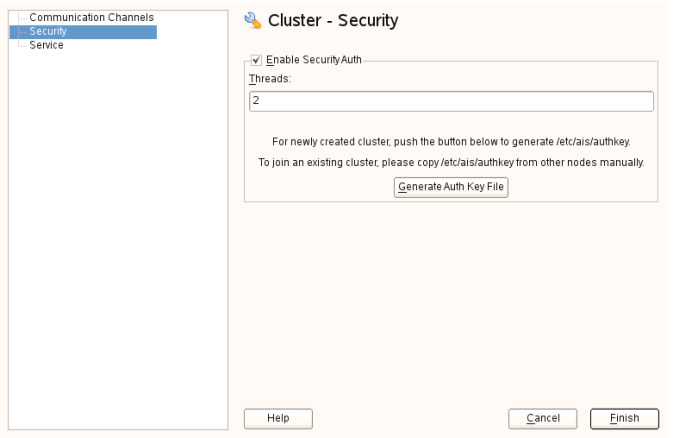


   定义用于所有群集节点的*绑定网络地址*、*多路广播地址*和*多路广播端口*。

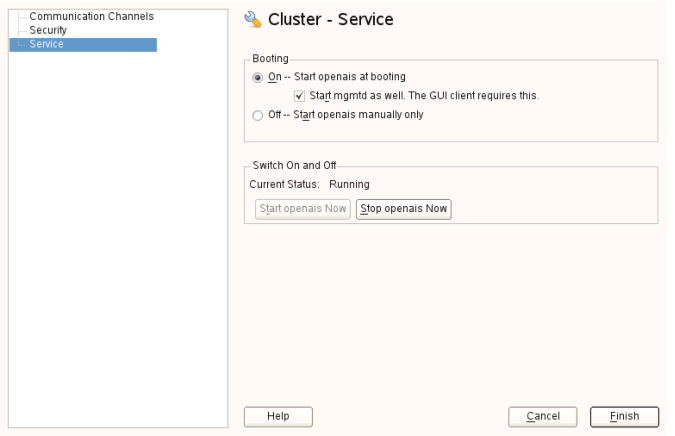
3. 为每个群集节点指定唯一的*节点 ID*。建议从 `1` 开始。

4 在*安全*类别中，定义群集的身份验证设置。如果激活*启用安全身份验证*，则会对群集节点间的通讯使用 HMAC/SHA1 身份验证。

此身份验证方法需要一个共享的机密，用于保护和鉴定消息。指定的身份验证密钥（密码）将用于群集中的所有节点。对于新创建的群集，请单击*生成身份验证密钥文件*创建身份验证密钥，它会写入 `/etc/ais/authkey`。



5 在*服务*类别中，选择是否要每次引导此群集服务器时都启动 OpenAIS。



如果选择关，则每次引导此群集服务器时必须手动启动OpenAIS。要手动启动 OpenAIS，请使用 `rcopenais start` 命令。

要立即启动 OpenAIS，请单击*立即启动 OpenAIS*。

6 如果已按照需要设置了所有选项，请单击*完成*。YaST 随后还会自动调整防火墙设置并打开用于多路广播的 UDP 端口。

7 初始配置完成后，需要将此配置传送到群集中的其他节点。执行此操作的最简单方法是将 `/etc/ais/openais.conf` 文件复制到群集中的其他节点。由于每个节点需要具有唯一的节点 ID，请确保复制文件后相应地调整节点 ID。

8 如果要使用加密通讯，请同时将 `/etc/ais/authkey` 复制到群集中的其他节点。

3.3 使群集联机

在基本配置后，可以使堆栈联机并检查状态。

1 在每个群集节点上运行以下命令以启动 OpenAIS：

```
rcopenais start
```

2 在任一节点上，用以下命令检查群集状态：

```
crm_mon
```

如果所有节点都联机，则输出应类似于如下内容：

```
============
Last updated: Thu Feb  5 18:30:33 2009
Current DC: d42 (d42)
Version: 1.0.1-node: b7ffe2729e3003ac8ff740bebc003cf237dfa854
3 Nodes configured.
0 Resources configured.
============

Node: d230 (d230): online
Node: d42 (d42): online
Node: e246 (e246): online
```

在基本配置完成且节点联机后，即可开始用 `crm` 命令行工具或图形用户界面配置群集资源。有关更多信息，请参见第 4 章 *使用 GUI 配置群集资源* [29]或第 5 章 *通过命令行配置群集资源* [55]。

部分 II. 配置和管理

4 使用 GUI 配置群集资源

HA 群集的主要目的是管理用户服务。用户服务的典型示例是 Apache Web 服务器或数据库。从用户角度来看，服务就是在客户的要求下执行某些操作。但对群集来说，服务则是可以启动或停止的资源 — 服务的本质与群集无关。

作为群集管理员，您需要在群集中为服务器上运行的每个资源或应用程序创建群集资源。群集资源可以包括 Web 站点、电子邮件服务器、数据库、文件系统、虚拟机和任何其他基于服务器的应用程序或在任意时间对用户都可用的服务。

要创建群集资源，请使用图形用户界面 (Linux HA Management Client) 或 `crm` 命令行实用程序。有关命令行方法，请参见第 5 章 *通过命令行配置群集资源* [55]。

本章先介绍 Linux HA Management Client，然后介绍配置群集时所需的若干个主题：创建资源、配置约束、指定故障转移节点和故障回复节点、配置资源监视、启动或删除资源、配置资源组或克隆资源，以及手动迁移资源。

配置群集资源的图形用户界面包括在 `pacemaker-pygui` 包中。

4.1 Linux HA Management Client

启动 Linux HA Management Client 时，需要连接到群集。

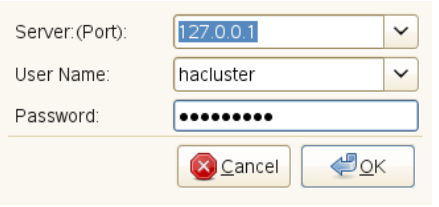
注意：hacluster 用户的密码

安装过程中会创建名为 `hacluster` 的 Linux 用户。在使用 Linux HA Management Client 之前，必须为 `hacluster` 用户设置密码。必须成为 `root` 才能执行此操作；在命令行输入 `passwd hacluster` 并为 `hacluster` 用户输入密码。

对要使用 Linux HA Management Client 连接的每个节点执行此操作。

要启动 Linux HA Management Client，请在命令行输入 `crm_gui`。要连接到群集，请选择 *Connection*（*连接*）> *登录*。默认情况下，*Server*（服务器）字段会显示本地主机的 IP 地址，*User Name*（用户名）字段会显示 `hacluster`。输入该用户的密码以继续。

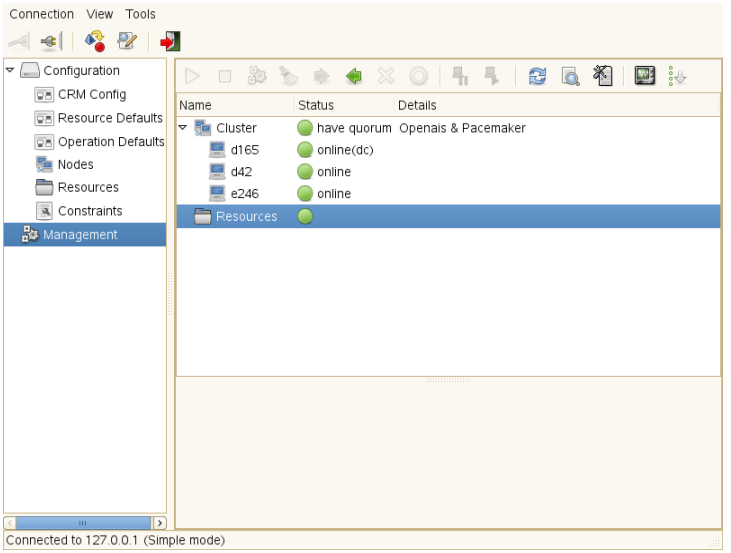
图 4.1 *连接群集*



如果是远程运行 Linux HA Management Client，请为 *Server*（服务器）字段输入群集节点的 IP 地址。在 *User Name*（用户名）字段中，您也可以使用任何其他属于 `haclient` 组的用户连接到群集。

连接后，系统会打开主窗口：

图 4.2　Linux HA Management Client - 主窗口



Linux HA Management Client 允许您添加和修改资源、约束及配置等。它还提供管理群集组件的功能，如启动、停止或迁移资源，清理资源，或将节点设置为待机。此外，还可以通过选择任意 *Configuration*（配置）子项并选择 *Show*（*显示*）> *XML Mode*（*XML 模式*），轻松地查看、编辑、导入和导出 CIB 的 XML 结构。

在下面可以找到一些如何使用 Linux HA Management Client 创建和管理群集资源的示例。

4.2 创建群集资源

可以创建以下类型的资源：

原始
: 原始资源是最基本的资源类型。

组
: 组包含一系列需要放置在一起、按顺序启动和以反序停止的资源。有关更多信息，请参考第 4.10 节 “配置群集资源组” [45]。

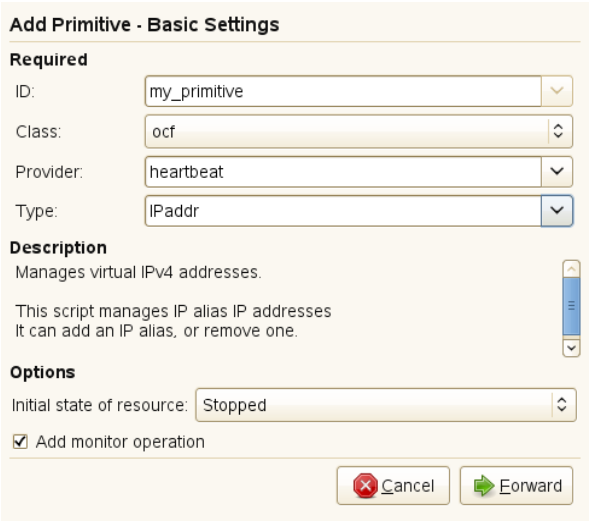
复制
克隆是可以在多个主机上处于活动状态的资源。如果各个资源代理支持，则任何资源均可克隆。有关更多信息，请参考第 4.11 节 “配置克隆资源” [50]。

主
主资源是一种特殊的克隆资源，主资源可以具有多种模式。主资源必须只能包含一个组或一个常规资源。

过程 4.1 *添加原始资源*

1. 按第 4.1 节 “Linux HA Management Client” [29]中所述，启动 Linux HA Management Client 并登录到群集。
2. 在左窗格中，选择*资源*并单击*添加* > *Primitive*（*原始*）。
3. 在下一个对话框中，为资源设置以下参数：
   - **3a** 为资源输入唯一的 `ID`。
   - **3b** 从 *Class*（类）列表中，选择要用于该资源的资源代理类：*heartbeat*、*lsb*、*ocf* 或 *stonith*。有关详细信息，参见第 17.1 节 “支持的资源代理类” [181]。
   - **3c** 如果选择了 *ocf* 作为类，请同时指定 OCF 资源代理的 *Provider*（提供程序）。OCF 规范允许多个供应商供应相同的资源代理。
   - **3d** 从 *Type*（类型）列表中，选择要使用的资源代理（例如 *IPaddr* 或 *Filesystem*）。该资源代理的简短描述显示在下方。

     Type（类型）列表中提供的选项取决于您选择的 *Class*（类）（对于 OCF 资源还取决于 *Provider*（提供程序）中选择的内容）。
   - **3e** 在 *Options*（选项）下面，设置 *Initial state of resource*（资源的初始状态）。
   - **3f** 如果希望群集监视资源状况是否仍然正常，请激活 *Add monitor operation*（添加监视操作）。



4 单击 *Forward*（前进）。下一个窗口将显示已为该资源定义的参数摘要。系统会列出该资源的所有必需的 *Instance Attributes*（实例属性）。若要设置相应的值，需要对实例属性进行编辑。可能还需要添加更多的属性，具体取决于您的部署和设置。有关如何操作的细节，请参考添加或修改元属性和实例属性 [34]。

5 如果所有参数都按您的需要进行了设置，请单击*应用*完成该资源的配置。配置对话框关闭，主窗口显示新添加的资源。

对于原始资源，您随时可以添加或修改下列参数：

元属性
: 元属性是可以为资源添加的选项。它们告诉 CRM 如何处理特定资源。有关可用的元属性、属性值和默认值的概览，请参考第 17.3 节 “资源选项” [184]。

实例属性
: 实例属性是特定资源类的参数，用于确定资源类的行为方式及其控制的服务实例。有关更多信息，请参考第 17.5 节 “实例属性” [186]。

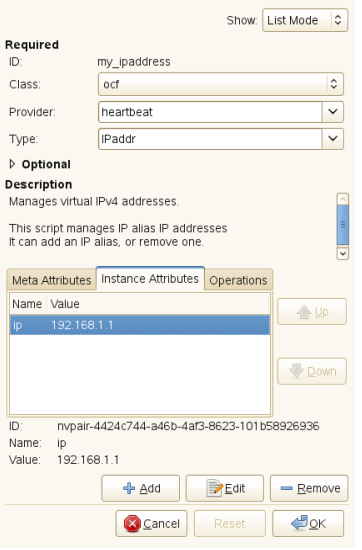
操作
: 可以为资源添加监视操作。监视操作指示群集确保资源状况依然正常。所有资源代理类都可以添加监视操作。您还可以设置特定参数，如为 `start` 或

stop 操作设置 timeout。有关更多信息，请参考第 4.7 节 “配置资源监视” [43]。

过程 4.2　添加或修改元属性和实例属性

1. 在 Linux HA Management Client 主窗口中，单击左窗格中的*资源*以查看已配置的群集资源。

2. 在右窗格中，选择要修改的资源并单击 *Edit*（编辑）（或双击资源）。下一个窗口将显示已为该资源定义的基本资源参数和 *Meta Attributes*（元属性）、*Instance Attributes*（实例属性）或 *Operations*（操作）。



3. 要添加新的元属性或实例属性，请选择相应的选项卡并单击*Add*（添加）。

4. 选择要添加的属性*名称*。将显示简短的*描述*。

5. 如果需要，请指定属性*值*。否则，使用该属性的默认值。

6. 单击 *OK*（确定）确认更改。新添加或新修改的属性便显示在选项卡上。

7. 如果所有参数都按您的需要进行了设置，请单击 *OK*（确定）完成该资源的配置。配置对话框关闭，主窗口显示已修改的资源。

提示：XML 源代码

通过 Linux HA Management Client 可以查看根据您为特定资源或所有资源定义的参数而生成的 XML。在资源配置对话框的右上角或在主窗口的 *Resources*（资源）视图下，选择*Show*（*显示*）> *XML Mode*（*XML 模式*）。

编辑器将显示 XML 代码，允许您 *Import*（导入）或 *Export*（导出）XML 元素或手动编辑 XML 代码。

4.3 创建 STONITH 资源

要配置屏障，需要配置一个或多个 STONITH 资源。

过程 4.3 *添加 STONITH 资源*

1. 按第 4.1 节 “Linux HA Management Client” [29]所述，启动 Linux HA Management Client 并登录到群集。
2. 在左窗格，选择 *Resources*（资源）并单击*Add*（*添加*）> *Primitive*（*原始*）。
3. 在下一个对话框中，为资源设置以下参数：
   - **3a** 为资源输入唯一的 `ID`。
   - **3b** 从 *Class*（类）列表，选择资源代理类 *stonith*。
   - **3c** 从 *Type*（类型）列表，选择 STONITH 插件以控制 STONITH 设备。该插件的简短描述显示在下方。
   - **3d** 在 *Options*（选项）下面，设置 *Initial state of resource*（资源的初始状态）。
   - **3e** 如果希望群集监视屏障设备，请激活 *Add monitor operation*（添加监视操作）。有关更多信息，请参考第 8.4 节 “监视屏障设备” [82]。
4. 单击 *Forward*（前进）。下一个窗口将显示已为该资源定义的参数摘要。系统会列出所选 STONITH 插件的所有必需的*实例属性*。若要设置相应的

值，需要对实例属性进行编辑。可能还需要添加更多的属性或监视操作，具体取决于您的部署和设置。有关如何操作的细节，请参考添加或修改元属性和实例属性 [34]和第 4.7 节 “配置资源监视” [43]。

5 如果所有参数都按您的需要进行了设置，请单击 *Apply*（应用）完成该资源的配置。配置对话框关闭，主窗口显示新添加的资源。

要完成屏障配置，请添加约束和/或使用克隆。有关详细信息，请参考第 8 章 *屏障和 STONITH* [75]。

4.4 配置资源约束

配置好所有资源只是完成了该作业的一部分。即便群集熟悉所有必需资源，它可能还无法进行正确处理。资源约束允许您指定在哪些群集节点上运行资源、以何种顺序装载资源，以及特定资源依赖于哪些其他资源。

提供三种不同的约束：

Resource Location（资源位置）
位置约束定义资源可以、不可以或首选在哪些节点上运行。

Resource Collocation（资源排列）
排列约束告诉群集资源可以或不可以在某个节点上一起运行。

Resource Order（资源顺序）
排序约束定义操作的顺序。

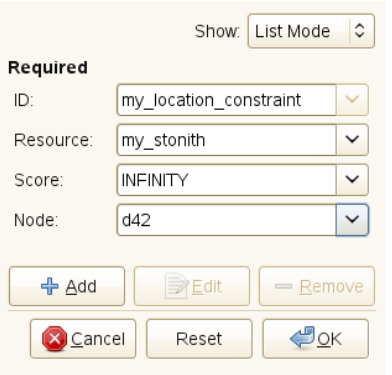
定义约束时，还需要指定分数。各种分数是群集工作方式的重要组成部分。其实，从迁移资源到决定在已降级群集中停止哪些资源的整个过程是通过以某种方式操纵分数来实现的。分数按每个资源来计算，资源分数为负的任何节点都无法运行该资源。在计算出资源分数后，群集选择分数最高的节点。`INFINITY`（无穷大）目前定义为 `1,000,000`。加减无穷大遵循以下 3 个基本规则：

- 任何值 + 无穷大 = 无穷大
- 任何值 - 无穷大 = -无穷大
- 无穷大 - 无穷大 = -无穷大

定义资源约束时，也可以指定每个约束的分数。分数表示您指派给此资源约束的值。分数较高的约束先应用，分数较低的约束后应用。通过使用不同的分数为既定资源创建更多位置约束，可以指定资源要故障转移至的目标节点的顺序。

过程 4.4 *添加或修改位置约束*

1. 按第 4.1 节 “Linux HA Management Client” [29]中所述，启动 Linux HA Management Client 并登录到群集。
2. 在 Linux HA Management Client 主窗口中，单击左窗格中的 *Constraints*（约束）以查看群集已配置的约束。
3. 在左窗格中，选择 *Constraints*（约束）并单击 *Add*（添加）。
4. 选择 *Resource Location*（资源位置）并单击 *OK*（确定）。
5. 为约束输入唯一的 *ID*。修改现有约束时，ID 已经定义并显示在配置对话框中。
6. 选择要定义约束的*资源*。列表中显示群集已配置的所有资源的 ID。
7. 设置约束的*分数*。正值表示资源可以在以下指定的*节点*上运行。负值表示资源不能在此节点上运行。值 `+/- INFINITY` 则由“可以”变为 `must`。
8. 选择约束的*节点*。



9 如果将*Node*（节点）和*Score*（分数）字段留为空白，也可以通过单击*Add*（*添加*）>*Rule*（*规则*）来添加规则。要添加有效期，只需单击*Add*（*添加*）>*Lifetime*（*有效期*）即可。

10 如果所有参数都按您的需要进行了设置，请单击*OK*（确定）完成约束配置。配置对话框关闭，主窗口显示新添加或新修改的约束。

过程 4.5 *添加或修改排列约束*

1 按第 4.1 节 “Linux HA Management Client” [29]中所述，启动 Linux HA Management Client 并登录到群集。

2 在 Linux HA Management Client 主窗口中，单击左窗格中的 *Constraints*（约束）以查看已为群集配置的约束。

3 在左窗格中，选择 *Constraints*（约束）并单击 *Add*（添加）。

4 选择 *Resource Collocation*（资源排列）并单击 *OK*（确定）。

5 为约束输入唯一的 *ID*。修改现有约束时，ID 已经定义并显示在配置对话框中。

6 选择要作为排列源的*资源*。列表中显示群集已配置的所有资源的 ID。

如果约束无法满足，群集可以决定根本不允许运行资源。

7 如果将 *Resource*（资源）和 *With Resource*（排列资源）字段都保留为空，也可以通过单击*Add*（*添加*）>*Resource Set*（*资源集*）来添加资源集。要添加有效期，只需单击 *Add*（*添加*）>*Lifetime*（*有效期*）即可。

8 在 *With Resource*（排列资源）中，定义排列目标。群集先决定将此资源放置在什么位置，再决定将此资源放置在*Resource*（资源）字段的什么位置。

9 定义 *Score*（分数）可确定两个资源的位置关系。正值表示两个资源应在相同的节点上运行。负值表示两个资源不应在相同的节点上运行。值 `+/- INFINITY` 则由 `should` 变为 `must`。该分数与其他系数结合使用可以决定放置节点的位置。

10 如果需要，请指定进一步参数，如 *Resource Role*（资源角色）。

根据选择的参数和选项，显示简短*描述*，解释您正在配置的排列约束的效果。

11 如果所有参数都按您的需要进行了设置，请单击 *OK*（确定）完成约束的配置。配置对话框关闭，主窗口显示新添加或新修改的约束。

过程 4.6 *添加或修改顺序约束*

1 按第 4.1 节 “Linux HA Management Client” [29]中所述，启动 Linux HA Management Client 并登录到群集。

2 在 Linux HA Management Client 主窗口中，单击左窗格中的 *Constraints*（约束）以查看群集已配置的约束。

3 在左窗格中，选择 *Constraints*（约束）并单击 *Add*（添加）。

4 选择 *Resource Order*（资源顺序）并单击 *OK*（确定）。

5 为约束输入唯一的 *ID*。修改现有约束时，ID 已经定义并显示在配置对话框中。

6 使用 *First*（首先）定义必须先于 *Then*（然后）资源启动的资源。

7 使用 *Then*（然后）定义必须后于 *First*（首先）资源启动的资源。

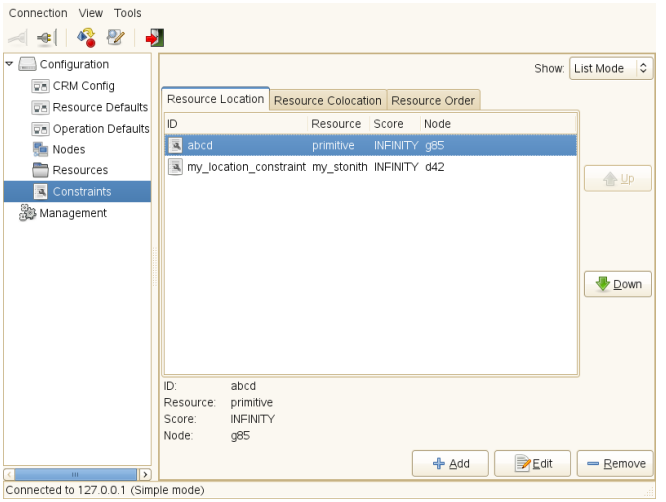
8 如果需要，定义进一步参数，例如，*Score*（分数，如果大于零，约束是强制的；否则，约束只是一种建议）或 *Symmetrical*（对称）（如果为 `true`，以相反的顺序停止资源。

根据选择的参数和选项，显示简短*描述*，解释您正在配置的顺序约束的效果。

9 如果所有参数都按您的需要进行了设置，请单击 *OK*（确定）完成约束的配置。配置对话框关闭，主窗口显示新添加或新修改的约束。

您可以访问和修改在 Linux HA Management Client 的 *Constraints*（约束）视图下配置的所有约束。

图 4.3　Linux HA Management Client - 约束



有关配置约束的更多信息以及顺序和排列基本概念的详细背景信息，请参考 http://clusterlabs.org/wiki/Documentation上提供的以下文档：

- *配置 1.0 说明*，*资源约束*章节
- *排列说明*
- *排序说明*

4.5 指定资源故障转移节点

资源在出现故障时会自动重启动。如果在当前节点上无法实现重启动，或如果在当前节点上发生 N 次故障，则资源会试图故障转移到其他节点。您可以多次定义资源的故障次数（`migration-threshold`），在该值之后资源会迁移到新节点。如果群集中存在两个以上的节点，特定资源故障转移的节点由 High Availability 软件选择。

如果您希望自己选择资源故障转移的节点，必须执行以下操作：

1 按添加或修改位置约束 [37]中所述，为该资源配置位置约束。

2 按添加或修改元属性和实例属性 [34]中所述，为该资源添加 `migration-threshold` 元属性，并输入迁移阈值的*值*。值应该是小于 INFINITY 的正值。

3 如果希望资源的故障计数自动失效，请按添加或修改元属性和实例属性 [34]中所述为该资源添加 `failure-timeout` 元属性，并输入故障超时的*值*。

4 如果希望为资源指定更多的首选故障转移节点，请创建更多的位置约束。

例如，假设您已经为 `r1` 资源配制了一个首选在 `node1` 节点上运行的位置约束。如果那里失败了，系统会检查 `migration-threshold` 并与故障计数进行比较。如果故障计数 >= migration-threshold，会将资源迁移到下一个自选节点。

默认情况下，一旦达到阈值，就只有在管理员手动重置资源的故障计数后（在修复故障原因后），才允许在该节点上运行有故障的资源。

但是，可以通过设置资源的 failure-timeout 选项使故障计数失效。因此，`migration-threshold=2` 和 `failure-timeout=60s` 的设置会导致资源在两次故障后迁移到新的节点，并且可能允许在一分钟后移回（取决于黏性和约束分数）。

迁移阈值概念有两个例外，在资源启动失败或停止失败时出现：启动故障会使故障计数设置为 INFINITY，因此总是导致立即迁移。停止故障会导致屏障（`stonith-enabled` 设置为 `true` 时，这是默认设置）。如果不定义 STONITH 资源（或 `stonith-enabled` 设置为 `false`），则该资源根本不会迁移。

要使用 Linux HA Management Client 清理资源计数，请选择左窗格中的 *Management*（管理），再选择右窗格中各自的资源，然后单击工具栏上的 *Cleanup Resource*（清理资源）。此操作对指定节点上的指定资源执行 `crm_resource -C` 和 `crm_failcount -D` 命令。有关详细信息，另请参见crm_resource(8) [157]和crm_failcount(8) [148]。

4.6 指定资源故障回复节点（资源黏性）

当原始节点恢复联机并位于群集中时，资源可能会故障回复到该节点。如果希望阻止资源故障回复到故障转移前运行的节点上，或如果希望指定其他的节点让资源进行故障回复，则必须更改资源黏性值。在创建资源时或在创建资源后，都可以指定指定资源黏性。

在指定资源黏性值时，请考虑以下情况：

值为 0：
这是默认选项。资源放置在系统中的最适合位置。这意味着当负载能力“较好”或较差的节点变得可用时才转移资源。此选项的作用几乎等同于自动故障回复，只是资源可能会转移到非之前活动的节点上。

值大于 0：
资源更愿意留在当前位置，但是如果有更合适的节点可用时会移动。值越高表示资源越愿意留在当前位置。

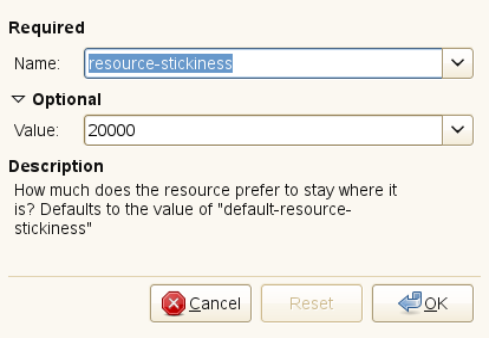
值小于 0：
资源更愿意移离当前位置。绝对值越高表示资源越愿意离开当前位置。

值为 INFINITY：
如果不是因节点不适合运行资源（节点关机、节点待机、达到 `migration-threshold` 或配置更改）而强制资源转移，资源总是留在当前位置。此选项的作用几乎等同于完全禁用自动故障回复。

值为 -INFINITY：
资源总是移离当前位置。

过程 4.7 *指定资源黏性*

1 按添加或修改元属性和实例属性 [34]中所述为资源添加 `resource-stickiness` 元属性。



2 在资源黏性*值*中，指定介于 -INFINITY 和 INFINITY 之间的值。

4.7 配置资源监视

虽然 High Availability Extension 可以检测节点故障，但也能够检测节点上的各个资源何时发生故障。如果希望确保资源运行，则必须为该资源配置资源监视。资源监视包括指定超时和/或启动延迟值以及间隔。间隔告诉 CRM 检查资源状态的频率。

过程 4.8 *添加或修改监视操作*

1 按第 4.1 节 “Linux HA Management Client” [29]中所述，启动 Linux HA Management Client 并登录到群集。

2 在 Linux HA Management Client 主窗口，单击左窗格中的 *Resources*（资源）以查看群集已配置的资源。

3 在右窗格中，选择要修改的资源并单击 *Edit*（编辑）。下一个窗口将显示为该资源定义的基本资源参数、元属性、实例属性和操作。

4 要添加新的监视操作，请选择各自选项卡并单击 *Add*（添加）。

要修改现有操作，请选择各自条目并单击 *Edit*（编辑）。

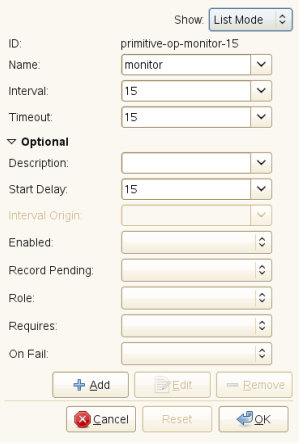
5 为监视操作输入唯一的 *ID*。修改现有监视操作时，ID 已经定义并显示在配置对话框中。

6 在 *Name*（名称）中，选择要执行的操作，例如 `monitor`、`start` 或 *stop*。

7 在 *Interval*（间隔）字段中，输入以秒表示的值。

8 在 *timeout*（超时）字段中，输入以秒表示的值。在指定的超时期间后，操作会被视为 `failed`。PE 会决定如何做或执行您在监视操作的 *On Fail*（失败时）字段中指定的操作。

9 如果需要，请设置可选参数，如 *On Fail*（失败时）（此操作失败时如何做？）。或 *Requires*（要求）（发生此操作前需要满足哪些条件？）。



10 如果所有参数都按您的需要进行了设置，请单击 *OK*（确定）完成该资源的配置。配置对话框关闭，主窗口显示修改后的资源。

如果不配置资源监视，则不会告知成功启动的资源故障，且群集始终显示资源状况正常。

如果资源监视检测到故障，会发生以下操作：

- 根据在 `/etc/ais/openais.conf` 的 `logging` 部分指定的配置，生成日志文件消息（默认情况下，写入系统日志，通常为 `/var/log/messages`）。

- 在 Linux HA Management Client 的 `crm_mon` 工具和 CIB 状态部分中反映故障。要在 Linux HA Management Client 中查看故障，请单击左窗格中的 *Management*（管理），然后在右窗格中，选择要查看其细节的资源。
- 群集启动有意义的恢复操作，包括停止资源修复故障状态和从本地或其他节点重启动资源。资源也可能根本不会重启动，具体取决于配置和群集状态。

4.8 启动新群集资源

注意：启动资源

使用 High Availability Extension 配置资源时，不应手动启动或停止相同的资源（在群集之外）。High Availability Extension 软件负责所有服务的启动或停止操作。

如果在创建时将资源的初始状态设置为 `stopped`（`target-role` 元属性的值为 `stopped`），则资源在创建后不会自动启动。要使用 Linux HA Management Client 启动新群集资源，请在左窗格中选择 *Management*（管理）。在右窗格中，右键单击资源并选择 *Start*（启动）（或从工具栏启动资源）。

4.9 删除群集资源

要使用 Linux HA Management Client 删除群集资源，请在左窗格中切换到 *Resources*（资源）视图，选择各自的资源并单击 *Remove*（删除）。

注意：删除引用的资源

如果群集资源的 ID 由任何约束引用，则无法删除该群集资源。如果您无法删除某个资源，请检查引用资源 ID 的位置，然后先从该约束删除资源。

4.10 配置群集资源组

某些群集资源与其他组件或资源相关，要求每个组件或资源以特定的顺序启动并在相同的服务器上运行。为了简化此配置，我们支持组的概念。

组具有以下属性：

启动和停止资源
: 资源以显示顺序启动，以相反顺序停止。

相关性
: 如果组中某个资源在某处无法运行，则该组中位于其之后的任何资源都不允许运行。

组内容
: 组可能仅包含一些原始群集资源。要引用组资源的子代，请使用子代的 ID 代替组的 ID。

限制
: 尽管在约束中可以引用组的子代，但通常倾向于使用组的名称。

黏性
: 黏性在组中可以累加。每个*活动*的组成员可以将其黏性值累加到组的总分中。因此，如果默认的`resource-stickiness`值为`100`，而组中有七个成员，其中五个成员是活动的，则组总分为 `500`，更喜欢其当前位置。

资源监控
: 要为组启用资源监视，必须为组中每个要监视的资源分别配置监视。

注意：空组

组必须包含至少一个资源，否则配置无效。

过程 4.9 *添加资源组*

1. 按第 4.1 节 “Linux HA Management Client” [29]中所述，启动 Linux HA Management Client 并登录到群集。
2. 在左窗格中，选择 *Resources（资源）*并单击 *Add（添加）* > *Group（组）*。
3. 为组输入唯一的 `ID`。
4. 在 *Options*（选项）下面，设置 *Initial state of resource*（资源的初始状态）并单击 *Forward*（前进）。

5 在下一个步骤中，可以添加原始资源作为组的子资源。它们的创建方式与添加原始资源 [32]中描述的步骤类似。

6 如果所有参数都按您的需要进行了设置，请单击 *Apply*（应用）完成原始资源的配置。

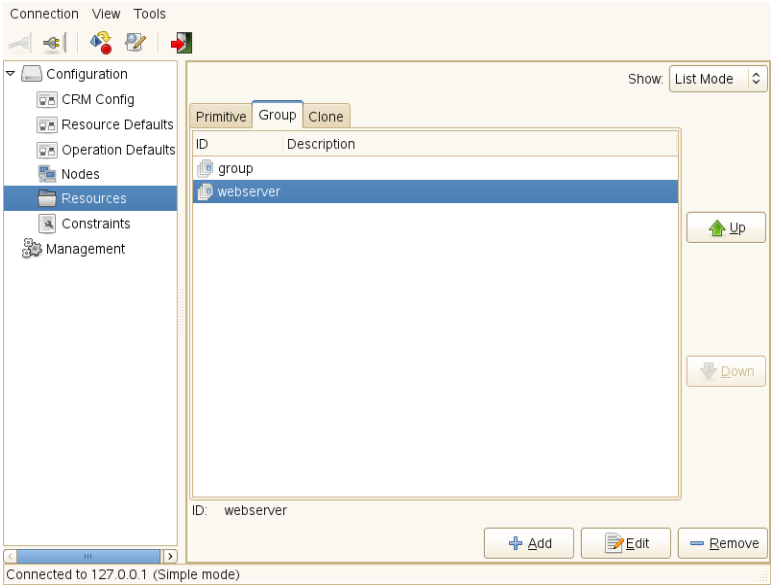
7 在下一个窗口中，可以通过再次选择 *Primitive*（原始）并单击 *OK*（确定）来继续为组添加子资源。

当不希望再向组中添加原始资源时，单击 *Cancel*（取消）。下一个窗口将显示您为该组定义的参数摘要。系统会列出组的 *Meta Attributes*（元属性）和 *Primitives*（原始）资源。资源在 *Primitive*（原始）选项卡上的位置代表资源在群集中的启动顺序。

8 由于资源在组中的顺序很重要，请使用 *Up*（向上）和 *Down*（向下）按钮对组中的 *Primitives*（原始）资源进行排序或重新排序。

9 如果所有参数都按您的需要进行了设置，请单击 *OK*（确定）完成该组的配置。配置对话框关闭，主窗口显示新创建或新修改的组。

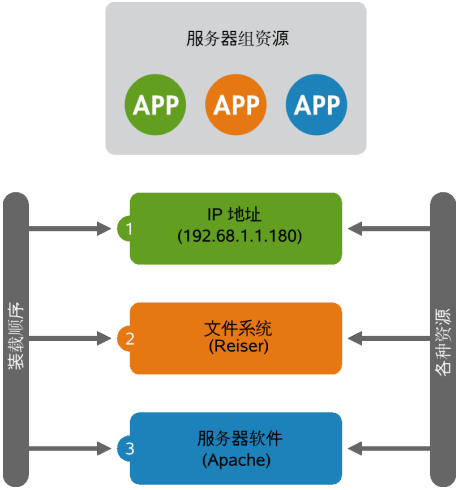
图 4.4 *Linux HA Management Client - 组*



例 4.1　Web 服务器的资源组

资源组示例是需要 IP 地址和文件系统的 Web 服务器。在这种情况下，每个组件都是一个会合并到群集资源组中的独立群集资源。资源组在一台或多台服务器上运行，如果软件或硬件有故障，故障转移至群集中的另一台服务器上，这与单个群集资源相同。

图 4.5　组资源



在添加资源组 [46]中，您已了解如何创建资源组。让我们假设您已经按上述说明创建了资源组。向现有组添加资源 [48]向您介绍如何修改组以匹配例 4.1 “Web 服务器的资源组” [48]。

过程 4.10　向现有组添加资源

1. 按第 4.1 节 “Linux HA Management Client” [29]中所述，启动 Linux HA Management Client 并登录到群集。
2. 在左窗格中，切换到 *Resources*（资源）视图；在右窗格中，选择要修改的组并单击 *Edit*（编辑）。下一个窗口将显示为该资源定义的基本组参数以及元属性和原始资源。
3. 单击 *Primitives*（原始）选项卡并单击 *Add*（添加）。
4. 在下一个对话框中，若要将 IP 地址添加为组的子资源，请设置以下参数：

4a 输入唯一 `ID`，例如 `my_ipaddress`。

4b 从 *Class*（类）列表中，选择 *ocf* 作为资源代理类。

4c 在 OCF 资源代理的 *Provider*（提供程序）中，选择 *heartbeat*。

4d 从 *Type*（类型）列表中，选择 *IPaddr* 作为资源代理。

4e 单击 *Forward*（前进）。

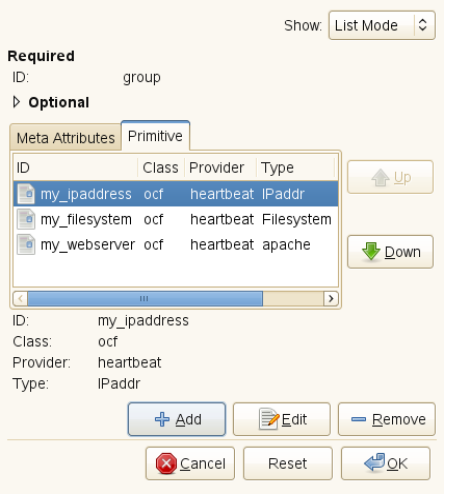
4f 在 *Instance Attribute*（实例属性）选项卡中，选择 *IP* 条目并单击 *Edit*（编辑）（或双击 *IP* 条目）。

4g 在 *Value*（值）中输入所需的 IP 地址，例如 `192.168.1.1`。

4h 单击 *OK*（确定）并单击 *Apply*（应用）。组配置对话框将显示新添加的原始资源。

5 再次单击 *Add*（添加）可添加下一个子资源（文件系统和 Web 服务器）。

6 设置每个子资源各自的参数，其过程类似于从步骤 4a [49]到步骤 4h [49]的步骤，直到为组配置完所有子资源为止。



由于已按子资源在群集中所需的启动顺序对子资源进行了配置，所以 *Primitives*（原始）选项卡上的顺序已是正确的。

7 如果需要更改组中的资源顺序，请使用 *Up*（向上）和 *Down*（向下）按钮对 *Primitive*（原始）选项卡上的资源进行重新排序。

8 要从组中删除资源，请在 *Primitive*（原始）选项卡上选择资源，并单击 *Remove*（删除）。

9 单击 *OK*（确定）完成该组的配置。配置对话框关闭，主窗口显示修改后的组。

4.11 配置克隆资源

您可能希望某些资源在群集的多个节点上同时运行。为此，必须将资源配置为克隆资源。可以配置为克隆资源的资源示例包括 STONITH 和群集文件系统（如 OCFS2）。如果受资源的资源代理支持，则可以克隆任何资源。克隆资源的配置甚至也有不同，具体取决于资源驻留的节点。

资源克隆有三种类型：

匿名克隆
: 这是最简单的克隆类型。这种克隆类型在所有位置上的运行方式都相同。因此，每台计算机上只能有一个匿名克隆实例是活动的。

全局唯一克隆
: 这些资源各不相同。一个节点上运行的克隆实例与另一个节点上运行的实例不同，同一个节点上运行的任何两个实例也不同。

状态克隆
: 这些资源的活动实例分为两种状态：主动和被动。有时也称为主要和辅助，或主和从。状态克隆可以是匿名克隆也可以是全局唯一克隆。

过程 4.11　添加或修改克隆

1 按第 4.1 节 “Linux HA Management Client” [29]中所述，启动 Linux HA Management Client 并登录到群集。

2 在左窗格中，选择 *Resources*（资源）并单击 *Add*（*添加*） > *Clone*（*克隆*）。

3 为克隆资源输入唯一的 `ID`。

4 在 *Options*（选项）下面，设置 *Initial state of resource*（资源的初始状态）。

5 为克隆资源激活要设置的各个选项，并单击 *Forward*（前进）。

6 在下一个步骤中，可以添加 *Primitive*（原始）或 *Group*（组）作为克隆资源的子资源。创建方式与添加原始资源 [32]或添加资源组 [46]中描述的过程类似。

7 如果所有参数都按您的需要进行了设置，请单击 *Apply*（应用）完成复制配置。

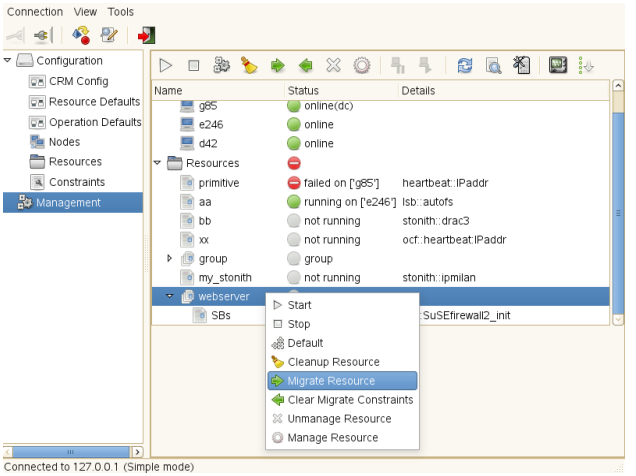
4.12 迁移群集资源

如第 4.5 节 “指定资源故障转移节点” [40]中所述，如果软件或硬件发生故障，群集会自动进行资源故障转移（迁移）—根据定义的特定参数（例如，迁移阈值或资源黏性）。除此之外，还可以手动将资源迁移到群集资源中的其他节点。

过程 4.12　手动迁移资源

1 按第 4.1 节 “Linux HA Management Client” [29]中所述，启动 Linux HA Management Client 并登录到群集。

2 在左窗格中切换到 *Management*（管理）视图，在右窗格中右键单击各个资源并选择 *Migrate Resource*（迁移资源）。



3 在新窗口的 *To Node*（目标节点）中，选择资源要移动到的节点。此操作将创建目标节点分数为 `INFINITY` 的位置约束。

4 如果只是希望临时迁移资源，请激活 *Duration*（持续时间）并输入资源迁移到新节点的时间范围。在持续时间到期后，资源*可以*移回到其原始位置，也可以留在当前位置（取决于资源黏性）。

5 如果资源无法迁移（若资源在当前节点上的黏性和约束总分数大于 `INFINITY`），请激活 *Force*（强制）选项。它通过创建当前位置规则和 `-INFINITY` 的分数强制资源移动。

注意

这可防止资源在此节点上运行，直到通过 *Clear Migrate Constraints*（清除迁移约束）删除约束或持续时间到期为止。

---

6 单击 *OK*（确定）确认迁移。

要允许资源重新移回，请切换到*Management*（管理），右键单击资源视图，并选择 *Clear Migrate Constraints*（清除迁移约束）。这使用 `crm_resource -U` 命令。资源*可以*移回到其原始位置，也可以留在当前位置（取决于资源黏性）。有关详细信息，请参考crm_resource(8) [157]；或*配置 1.0 说明*，*资源迁移*章节 (http://clusterlabs.org/wiki/Documentation)。

4.13 更多信息

http://clusterlabs.org/
Pacemaker 主页，随 High Availability Extension 提供的群集资源管理器。

http://linux-ha.org
高可用性 Linux 项目的主页。

http://clusterlabs.org/wiki/Documentation
CRM 命令行界面：`crm` 命令行工具介绍。

http://clusterlabs.org/wiki/Documentation
配置 1.0 说明：解释配置 Pacemaker 所涉及的概念。包含全面、详尽的参考信息。

5 通过命令行配置群集资源

就像在第 4 章 [29]中一样，必须为在群集中的服务器上运行的每个资源或应用程序创建群集资源。群集资源可包括 Web 站点、电子邮件服务器、数据库、文件系统、虚拟机和任何其他基于服务器的应用程序或在任意时间对用户都可用的服务。

您可以使用图形 HA Management Client 实用程序或 `crm` 命令行实用程序来创建资源。本章介绍几个 `crm` 实用程序。

5.1 命令行工具

安装后有一些可以管理群集的工具。通常您只需要 `crm` 命令。此命令包含几个子命令。运行 `crm help` 可以获取所有可用命令的概述。它包含一个带有嵌入示例的全面帮助系统。

`crm` 工具具有管理功能（子命令 `resources` 和 `node`），可用于配置 (`cib`, `configure`)。管理子命令可*立即*生效，但是配置需要最终提交。

5.2 调试配置更改

在将更改装载回群集之前，建议您使用`ptest` 查看更改。`ptest` 可显示由将要提交的更改引入的操作图表。您需要`graphiz` 包才能显示这些图表。以下示例是一个抄本，添加了监视操作：

```
# crm
crm(live)# configure
crm(live)configure# show fence-node2
primitive fence-node2 stonith:apcsmart \
        params hostlist="node2"
crm(live)configure# monitor fence-node2 120m:60s
crm(live)configure# show changed
primitive fence-node2 stonith:apcsmart \
        params hostlist="node2" \
        op monitor interval="120m" timeout="60s"
crm(live)configure# ptest
crm(live)configure# commit
```

5.3 创建群集资源

群集可使用三种类型的RA（资源代理）。首先，有旧式Heartbeat 1脚本。高可用性可使用 LSB 初始化脚本。最后，群集有其自己的一组 OCF (Open Cluster Framework) 代理。本文档重点介绍 LSB 脚本和 OCF 代理。

要创建群集资源，请使用 crm 工具。要向群集添加新的资源，常规步骤如下所示：

1. 打开壳层并成为 root。
2. 运行 crm 来打开 crm 的内部壳层。提示符变为 crm(live)#。
3. 配置原始 IP 地址：

   ```
   crm(live)# configure
   crm(live)configure# primitive myIP ocf:heartbeat:IPaddr \
        params ip=127.0.0.99 op monitor intervall=60s
   ```

   上一命令配置了名称为myIP的“原始资源”。您还需要类（在此为ocf）、提供程序(heartbeat)和类型(IPaddr)。此外，此原始资源还需要一些参数，如 IP 地址。必须将地址更改为您的设置。

4. 显示您所做的更改并进行复查：

   ```
   crm(live)configure# show
   ```

   要查看 XML 结构，请使用以下命令：

   ```
   crm(live)configure# show xml
   ```

5 提交更改使其生效：

```
crm(live)configure# commit
```

5.3.1 LSB 初始化脚本

通常可在目录 /etc/init.d 中找到所有 LSB 脚本。根据 http://www.linux-foundation.org/spec/refspecs/LSB_1.3.0/gLSB/gLSB/iniscrptact.html 中所述，它们必须已实现一些操作，至少包括 start、stop、restart、reload、force-reload 和 status。

这些服务的配置没有标准化。如果您打算将 LSB 脚本用于 High Availability，请确定您了解各个脚本是如何配置的。您通常可在 /usr/share/doc/packages/*包名*中的各个包的文档中找到有关这些内容的文档。

注意：请勿接触高可用性使用的服务

当某个服务由 High Availability 使用时，不可以通过其他方式接触该服务。这表示在引导、重引导或手动操作时不应启动或停止该资源。但是，如果要检查服务是否正确配置，可手动启动该服务，但请确定在 High Availability 接管前再次停止该服务。

在使用 LSB 资源之前，确定该资源的配置在所有群集节点上存在并且是相同的。该配置不由 High Availability 管理。您必须自己管理。

5.3.2 OCF 资源代理

所有 OCF 代理都位于 /usr/lib/ocf/resource.d/heartbeat/。这些是功能与 LSB 脚本功能相似的小型程序。但是，该配置始终使用环境变量来设置。要求所有 OCF 资源代理至少包含操作 start、stop、status、monitor 和 meta-data。meta-data 操作可检索有关如何配置代理的信息。例如，如果您想了解有关 IPaddr 代理的更多信息，可使用以下命令：

```
OCF_ROOT=/usr/lib/ocf /usr/lib/ocf/resource.d/heartbeat/IPaddr meta-data
```

输出是简单 XML 格式的详细信息。您可以使用 ra-api-1.dtd DTD 来验证该输出。此 XML 格式基本上包含三个部分，首先是一些通用的描述，然后是所有可用的参数，最后是该代理的可用操作。

此输出设计为机器可读，可读性不太好。出于此原因，crm 工具包含了 ra 命令，以便获取有关资源代理的不同信息：

```
# crm
crm(live)# ra
crm(live)ra#
```

命令 classes 可提供所有类和提供程序的列表：

```
crm(live)ra# classes
stonith
lsb
ocf / lvm2 ocfs2 heartbeat pacemaker
heartbeat
```

要获取有关某个类（和提供程序）的所有可用资源代理的概述，请使用 list 命令：

```
crm(live)ra# list ocf
AudibleAlarm       ClusterMon         Delay              Dummy
Filesystem         ICP                IPaddr             IPaddr2
IPsrcaddr          IPv6addr           LVM                LinuxSCSI
MailTo             ManageRAID         ManageVE           Pure-FTPd
Raid1              Route              SAPDatabase        SAPInstance
SendArp            ServeRAID          SphinxSearchDaemon Squid
...
```

有关资源代理的更多信息可使用 meta 命令来查看：

```
crm(live)ra# meta Filesystem ocf heartbeat
Filesystem resource agent (ocf:heartbeat:Filesystem)

Resource script for Filesystem. It manages a Filesystem on a shared storage
medium.

Parameters (* denotes required, [] the default):
...
```

按 Q 键可退出查看器。可在第 6 章 *设置简单的测试资源* [69]中查找配置示例。

5.3.3 NFS 服务器的示例配置

要设置 NFS 服务器，需要三种资源：文件系统资源、drbd 资源及一组 NFS 服务器和 IP 地址。以下子节介绍了如何执行该操作。

设置文件系统资源

filesystem资源已配置为OCF原始资源。它的任务是在发出启动和停止请求时将设备装入和卸载到某个目录。在这种情况下，该设备为 /dev/drbd0，用作安装点的目录为 /srv/failover。使用的文件系统为 xfs。

可使用 crm 壳层中的以下命令来配置文件系统资源：

```
crm(live)# configure
crm(live)configure# primitive filesystem_resource \
    ocf:heartbeat:Filesystem \
    params device=/dev/drbd0 directory=/srv/failover fstype=xfs
```

配置 drbd

在开始 drbd High Availability 配置之前，请手动设置 drbd 设备。这主要是在 /etc/drbd.conf 中配置 drbd 并使它同步。*存储管理指南* 中描述了配置 drbd 的详细步骤。现在，假定您配置了在两个群集节点上的设备 /dev/drbd0 中都可访问的资源 r0。

drbd 资源为 OCF 主从属资源。这在对 drbd RA 的元数据的描述中可找到。但是，更重要的是元数据的 actions 部分中包含操作 promote 和 demote。这些操作对于主从属资源是必需的，通常不提供给其他资源。

对于 High Availability，主从属资源可能在不同的节点上有多个主资源。甚至有可能主资源和从属资源在同一个节点上。因此，请以这样的方式配置此资源：只有一个主资源和一个从属资源，且各自在不同的节点上运行。可使用master资源的元属性来执行此操作。主从属资源是 High Availability 中的一种特殊的克隆资源。每个主资源和每个从属资源计为一个克隆。

使用 crm 壳层中的以下命令来配置主从属资源：

```
crm(live)# configure
crm(live)configure# primitive drbd_r0 ocf:heartbeat:drbd params
crm(live)configure# ms drbd_resource drbd_r0 \
   meta clone_max=2 clone_node_max=1 master_max=1 master_node_max=1 notify=true
crm(live)configure# commit
```

NFS 服务器和 IP 地址

要使 NFS 服务器在同一个 IP 地址上始终可用，可使用附加 IP 地址以及计算机用于正常操作的那些地址。除系统的 IP 地址以外，此地址随后会指派到活动的 NFS 服务器。

NFS 服务器和 NFS 服务器的 IP 地址在同一台计算机上应始终是活动的。在这种情况下，启动顺序就不是很重要了。它们甚至可以同时启动。这些是组资源的典型要求。

在启动 High Availability RA 配置之前，先使用 YaST 配置 NFS 服务器。不要让系统启动 NFS 服务器。只需要设置配置文件。如果要手动执行该操作，请参阅手册页导出 exports(5) (`man 5 exports`)。配置文件为 `/etc/exports`。将 NFS 服务器配置为 LSB 资源。

通过 High Availability RA 配置可完全配置 IP 地址。系统中不需要任何其他修改。IP 地址 RA 为 OCF RA。

```
crm(live)# configure
crm(live)configure# primitive nfs_resource lsb:nfsserver
crm(live)configure# primitive ip_resource ocf:heartbeat:IPaddr \
      params ip=10.10.0.1
crm(live)configure# group nfs_group nfs_resource ip_resource
crm(live)configure# commit
crm(live)configure# end
crm(live)# quit
```

5.4 创建 STONITH 资源

从 `crm` 透视图中来看，STONITH 设备只是另一种资源。要创建 STONITH 资源，请执行如下步骤：

1. 以系统管理员身份运行 `crm` 命令。提示符更改为 `crm(live)`。
2. 使用以下命令获取所有 STONITH 类型的列表：

```
crm(live)# ra list stonith
apcmaster               apcsmart                baytech
cyclades                drac3                   external/drac5
external/hmchttp        external/ibmrsa         external/ibmrsa-telnet
external/ipmi           external/kdumpcheck     external/rackpdu
external/riloe          external/sbd            external/ssh
external/vmware         external/xen0           external/xen0-ha
```

```
ibmhmc                  ipmilan                 meatware
null                    nw_rpc100s              rcd_serial
rps10                   ssh                     suicide
```

3 从以上列表中选择 STONITH 类型并查看可用的选项列表。可使用以下命令（按 Q 键可关闭查看器）：

```
crm(live)# ra meta external/ipmi stonith
IPMI STONITH external device (stonith:external/ipmi)

IPMI-based host reset

Parameters (* denotes required, [] the default):
...
```

4 使用类 `stonith` 创建 STONITH 资源，并输入您在步骤 3 中选择的类型以及各自的参数（如果需要），例如：

```
crm(live)# configure
crm(live)configure# primitive my-stonith stonith:external/ipmi \
   meta target-role=Stopped \
   operations my_stonith-operations \
     op monitor start-delay=15 timeout=15 hostlist='' \
                pduip='' community=''
```

5.5 配置资源约束

配置所有资源只是任务的一部分。即使群集了解所有需要的资源，它仍然不能正确处理它们。例如，尝试在 drbd 的从属节点上装入文件系统是没有意义的（实际上，这将由于 drbd 而失败）。要通知群集这些事项，需要定义约束。

在 High Availability 中，提供了三种类型的约束：

- 位置约束：定义资源可在哪个节点上运行（在 `crm` 壳层中使用命令 `location`）。
- 组合约束：告知群集哪些资源可以或不可以在一个节点上一起运行 (`colocation`)。
- 排序约束：定义操作的顺序 (`order`)。

5.5.1 位置约束

每个资源可多次添加此类约束。将对指定资源评估所有 rsc_location 约束。以下列出一个简单示例，该示例将在名称为 earth 的节点上运行 ID 为 fs1-loc 的资源的可能性增加到 100：

```
crm(live)configure# location fs1-loc fs1 100: earth
```

5.5.2 组合约束

colocation 命令用于定义哪些资源应在相同主机上运行，哪些资源应在不同主机上运行。通常情况下使用以下顺序：

```
crm(live)configure# order rsc1 rsc2
crm(live)configure# colocation rsc2 rsc1
```

只能设置 +INFINITY 或 -INFINITY 的分数来定义必须始终或决不能在同一节点上运行的资源。例如，要始终在同一个主机上运行 ID 为 filesystem_resource 和 nfs_group 的两个资源，可使用以下约束：

```
crm(live)configure# colocation nfs_on_filesystem inf: nfs_group
filesystem_resource
```

对于主从属配置，除在本地运行资源以外，还有必要了解当前节点是否为主节点。这可以通过附加 to_role 或 from_role 属性来检查。

5.5.3 排序约束

有时提供启动资源的顺序是必要的。例如，在设备可用于系统之前，您不能装入文件系统。使用排序约束可在另一个资源满足某个特殊条件之前或之后启动或停止某项服务，如已启动、已停止或已升级到主资源。可使用 crm 壳层中的以下命令来配置排序约束：

```
crm(live)configure# order nfs_after_filesystem mandatory: group_nfs
filesystem_resource
```

5.5.4 示例配置约束

如果没有附加约束，本章使用的示例将不能按预期工作。其中最基本的就是让所有资源在同一台计算机上作为 drbd 资源的主资源运行。另一个关键点就是在

任何其他资源启动之前，drbd 资源必须是主资源。在 drbd 不是主资源时尝试装入 drbd 设备只能失败。必须满足的约束如下所示：

- 文件系统必须始终与 drbd 资源的主资源位于相同的节点上。

```
crm(live)configure# colocation filesystem_on_master inf: \
    filesystem_resource drbd_resource:Master
```

- NFS 服务器及 IP 地址必须与文件系统位于相同的节点上。

```
crm(live)configure# colocation nfs_with_fs inf: \
   nfs_group filesystem_resource
```

- NFS 服务器及 IP 地址在装入文件系统后启动：

```
crm(live)configure# order nfs_second mandatory: \
   filesystem_resource nfs_group
```

- 必须在 drbd 资源升级为节点上的主资源之后将文件系统装入此节点上。

```
crm(live)configure# order drbd_first inf: \
    drbd_resource:promote filesystem_resource
```

5.6 指定资源故障转移节点

要确定资源故障转移，可使用元属性 migration-threshold。例如：

```
crm(live)configure# location r1-node1 r1 100: node1
```

通常 r1 会首选在 node1 上运行。如果失败，将检查 migration-threshold 并与将它与故障计数进行比较。如果故障计数 >= migration-threshold，则会将该资源迁移到具有下一个最佳自选设置的节点。

启动故障会根据 `start-failure-is-fatal` 选项将故障计数设置为 INFINITY。停止故障可导致屏障。如果没有定义 STONITH，则该资源将不会迁移。

5.7 指定资源故障回复节点（资源粘性）

根据故障计数迁移某个资源后，仅当管理员重设置故障计数或故障已过期（参见 failure-timeout 元属性）时，该资源才能故障回复。

```
crm resource failcount RSC delete NODE
```

5.8 配置资源监视

有两种方法可以监视资源：使用 op 关键字定义监视操作或使用 monitor 命令。以下示例配置了一个 Apache 资源并使用 op 关键字每 30 分钟对它进行监视一次：

```
crm(live)configure# primitive apache apache \
  params ... \
  op monitor interval=60s timeout=30s
```

同样也可以使用以下方式来实现：

```
crm(live)configure# primitive apache apache \
   params ...
crm(live)configure# monitor apache 60s:30s
```

5.9 启动新的群集资源

要启动新的群集资源，您需要相应的标识符。按如下所示继续：

1. 以系统管理员身份运行 crm 命令。提示符更改为 crm(live)。
2. 使用命令 status 搜索相应的资源。
3. 使用以下命令启动资源：

   ```
   crm(live)# resource start ID
   ```

5.10 删除群集资源

要删除群集资源，您需要相应的标识符。按如下所示继续：

1 以系统管理员身份运行 crm 命令。提示符更改为 crm(live)。

2 运行以下命令来获取您的资源列表：

```
crm(live)# resource status
```

例如，输出可以显示如下（其中 myIP 是您的资源的相应标识符）：

```
myIP    (ocf::IPaddr:heartbeat) ...
```

3 删除带有相应标识符的资源（也暗含 commit 命令）：

```
crm(live)# configure delete YOUR_ID
```

4 提交更改：

```
crm(live)# configure commit
```

5.11 配置群集资源组

群集的最常见元素之一就是需要位置集中、按顺序启动并以相反顺序停止的一组资源。为了简化此配置，我们支持组的概念。以下示例创建了两个原始资源（IP 地址和电子邮件资源）：

1 以系统管理员身份运行 crm 命令。提示符更改为 crm(live)。

2 配置这两个原始资源：

```
crm(live)# configure
crm(live)configure# primitive Public-IP ocf:IPaddr:heartbeat \
   params ip=1.2.3.4
crm(live)configure# primitive Email lsb:exim
```

3 为原始资源分组，它们相应的标识符的顺序应正确：

```
crm(live)configure# group shortcut Public-IP Email
```

5.12 配置克隆资源

最初将克隆构想成便于启动一个 IP 地址的 N 个实例并使它们分布在群集各处以保持负载平衡的一种方法。事实也证明它们在很多方面非常有用，包括与 DLM 集成、屏障子系统和 OCFS2。您可以克隆任何资源，只要资源代理支持。

存在以下类型的克隆资源：

匿名资源
: 匿名克隆是最简单的类型。这些资源在它们运行的每个位置的行为完全相同。正因如此，每台计算机只允许有一个活动的匿名克隆副本。

多状态资源
: 多状态资源是克隆的特殊形式。它们允许实例处于两个操作模式的任一个中。这些模式称为“主”和“从属”，但是它们可以表示您所需的任何含义。唯一的限制就是当启动一个实例时，该实例必须出现在从属状态中。

5.12.1 创建匿名克隆资源

要创建匿名克隆资源，首先要创建一个原始资源，然后使用 `clone` 命令来引用它。执行下列操作：

1. 以系统管理员身份运行 `crm` 命令。提示符更改为 `crm(live)`。
2. 配置原始资源，例如：

```
crm(live)# configure
crm(live)configure# primitive Apache lsb:apache
```

3. 克隆原始资源：

```
crm(live)configure# clone apache-clone Apache \
   meta globally-unique=false
```

5.12.2 创建有状态/多状态克隆资源

要创建有状态的克隆资源，首先要创建一个原始资源，然后再创建主从属资源。

1. 以系统管理员身份运行 `crm` 命令。提示符更改为 `crm(live)`。

2 配置原始资源。必要时更改时间间隔：

```
crm(live)# configure
crm(live)configure# primitive myRSC ocf:myCorp:myAppl \
  operations foo \
    op monitor interval=60 \
    op monitor interval=61 role=Master
```

3 创建主从属资源：

```
crm(live)configure# clone apache-clone Apache \
   meta globally-unique=false
```

5.13 迁移群集资源

尽管已将资源配置为在出现硬件或软件故障时自动故障转移（或迁移）到群集中的其他节点，您仍然可以使用 Linux HA Management Client 或命令行手动将资源迁移到群集中的其他节点。

1 以系统管理员身份运行 `crm` 命令。提示符更改为 `crm(live)`。

2 要将名为 `ipaddress1` 的资源迁移到名为 `node2` 的群集节点，请输入：

```
crm(live)# resource
crm(live)resource# migrate ipaddress1 node2
```

5.14 使用阴影配置进行测试

注意：仅适用于有经验的管理员

尽管概念很简单，仍然建议仅在您确实需要阴影配置或者对 High Availability 有所熟悉的情况下再使用它们。

阴影配置可用于测试不同的配置方案。如果您已创建几个阴影配置，可以逐个测试以查看更改的效果。

一般的流程显示如下：

1. 用户启动 `crm` 工具。

2. 切换到 `configure` 子命令：

```
crm(live)# configure
crm(live)configure#
```

3. 现在您可以做出更改。但是，当您确定这些配置有风险或者想以后应用它们时，可以将它们保存到新的阴影配置中：

```
crm(live)configure# cib new myNewConfig
INFO: myNewConfig shadow CIB created
crm(myNewConfig)configure# commit
```

4. 创建阴影配置后，即可进行更改。

5. 要切换回在线群集配置，可使用以下命令：

```
crm(myNewConfig)configure# cib use
crm(live)configure#
```

5.15 更多信息

http://linux-ha.org
高可用性 Linux 主页

http://www.clusterlabs.org/mediawiki/images/8/8d/Crm_cli.pdf
提供对 CRM CLI 工具的介绍

http://www.clusterlabs.org/mediawiki/images/f/fb/Configuration_Explained.pdf
说明 Pacemaker 配置

6 设置简单的测试资源

在按第 3 章 *用 YaST 进行安装和基本设置* [21]中所述安装和设置群集后，您学会了如何使用 GUI 或用命令行配置资源，本章提供配置简单资源（IP 地址）的基本示例。它演示了两种方法来完成资源配置：使用 Linux HA Management Client 或 `crm` 命令行工具。

在以下示例中，假设群集由至少两个节点组成。

6.1 使用 GUI 配置资源

创建样本群集资源并将它迁移到其他服务器可帮助您进行测试，以确保群集运行正确。要配置和迁移的简单资源就是 IP 地址。

过程 6.1 *创建 IP 地址群集资源*

1. 按第 4.1 节 “Linux HA Management Client” [29]中所述，启动 Linux HA Management Client 并登录到群集。
2. 在左侧窗格中切换到*资源*视图，在右侧窗格中，选择要修改的组并单击*编辑*。下一个窗口将显示基本组参数和元属性，以及该资源已定义的原始资源。
3. 单击*原始*选项卡，然后单击*添加*。
4. 在下一个对话框中，设置以下参数以添加一个IP地址作为该组的子资源：

   4a 输入唯一的 `ID`，例如，`myIP`。

4b 从*类*列表中，选择 *ocf* 作为资源代理类。

4c 对于您的 OCF 资源代理的*提供程序*，选择 *heartbeat*。

4d 从*类型*列表中，选择 *IPaddr* 作为资源代理。

4e 单击*前进*。

4f 在*实例属性*选项卡中，选择 *IP* 项并单击*编辑*（或双击 *IP* 项）。

4g 输入所需的 IP 地址作为*值*（例如，`10.10.0.1`）并单击*确定*。

4h *添加*新的实例属性，并将*名称*指定为 `nic`，将*值*指定为 `eth0`，然后单击*确定*。

名称和值取决于您的硬件配置和安装 High Availability Extension 软件期间选择的媒体配置。

5 如果您已根据需要设置好所有参数，请单击*确定*完成该资源的配置。配置对话框将关闭，主窗口将显示修改后的资源。

要使用 Linux HA Management Client 启动资源，请在左侧窗格中选择*管理*。在右侧窗格中，右键单击资源并选择*启动*（或从工具栏启动资源）。

要将 IP 地址资源迁移到其他节点 (`saturn`)，请按如下操作：

过程 6.2 *将资源迁移到其他节点*

1 切换到左侧窗格中的*管理*视图，然后右键单击右侧窗格中的 IP 地址资源，并选择*迁移资源*。

2 在新窗口中，从*目标节点*下拉列表中选择 `saturn` 以将选定的资源移到节点 `saturn` 上。

3 如果只想临时迁移资源，请激活*持续时间*并输入时间范围，在该时间段内资源应迁移到新的节点。

4 单击*确定*确认迁移。

6.2 手动配置资源

计算机提供的任何类型的服务都称为资源。资源能够由 High Availability 识别，当资源受 RA（资源代理）控制时，它们就是 LSB 脚本、OCF 脚本或旧式 Heartbeat 1 资源。所有资源都可以使用 `crm` 命令进行配置，或在 CIB（群集信息库）的 `resources` 部分配置为 XML。有关可用资源的概览，请查看第 18 章 *HA OCF 代理* [189]。

要将 IP 地址 `10.10.0.1` 作为资源添加到当前配置中，请使用 `crm` 命令：

过程 6.3 *创建 IP 地址群集资源*

1. 以系统管理员的身份运行 `crm` 命令。提示更改为 `crm(live)`。
2. 切换到 `configure` 子命令：

   ```
   crm(live)# configure
   ```

3. 创建 IP 地址资源：

   ```
   crm(live)configure# resource
   primitive myIP ocf:heartbeat:IPaddr params ip=10.10.0.1
   ```

注意

使用 High Availability 配置资源时，不应通过 `init` 初始化相同的资源。高可用性负责所有服务的启动或停止操作。

如果配置成功，新资源将显示在 `crm_mon` 中，它在群集的随机节点上启动。

要将资源迁移到其他节点，请执行以下操作：

过程 6.4 *将资源迁移到其他节点*

1. 启动壳层并成为 `root` 用户。
2. 将资源 `myip` 迁移到节点 `saturn`：

   ```
   crm resource migrate myIP saturn
   ```

7 添加或修改资源代理

应由群集管理的所有任务必须可作为资源使用。有两个应区分的主要组：资源代理和STONITH代理。对于这两个类别，您都可以添加自己的代理，根据需要扩展群集的功能。

7.1 STONITH 代理

群集有时会检测到某个节点行为不正常，需要将它删除。这称为*屏障*，通常使用 STONITH 资源来实现。所有 STONITH 资源都驻留在每个节点上的 `/usr/lib/stonith/plugins` 中。

警告：不支持 SSH 和 STONITH

由于无法了解 SSH 可能对其他系统问题如何做出反应。出于此原因，生产环境不支持 SSH 和 STONITH 代理。

要（从软件端）获取所有当前可用的 STONITH 设备列表，请使用 `stonith -L` 命令。

很抱歉，目前尚无有关写入 STONITH 代理的文档。如果要写入新的 STONITH 代理，请参考 `heartbeat-common` 包的源中提供的示例。

7.2 写入 OCF 资源代理

所有 OCF 资源代理都在 /usr/lib/ocf/resource.d/ 中提供，有关更多信息，请参见第 17.1 节 “支持的资源代理类” [181]。为了避免名称冲突，在创建新的资源代理时要创建不同的子目录。例如，如果您的一个资源组 kitchen 具有资源 coffee_machine，可将此资源添加到目录 /usr/lib/ocf/resource.d/kitchen/。要访问此资源代理，请执行命令 crm：

```
configure
primitive coffee_1 ocf:coffee_machine:kitchen ...
```

在实施您自己的 OCF 资源代理时，请提供该代理的一些操作。可在 http://www.linux-ha.org/OCFResourceAgent 中找到有关写入 OCF 资源代理的更多细节。有关 High Availability 2 的一些概念的特殊信息，可在第 1 章 *概念概述* [3]中查找。

8 屏障和 STONITH

屏障在HA（高可用性）计算机群集中是一个非常重要的概念。群集有时候会检测到某个节点功能不正常，需要将其删除。这就称为*屏障*，它通常通过STONITH资源完成。屏障可以定义为一种使 HA 群集具有已知状态的方法。

群集中的每个资源都附带一个状态，例如："资源 r1 在节点 1 上启动"。在 HA 群集中，这种状态暗示了"资源 r1 在除节点 1 的所有节点上都是停止的"，因为HA群集必须确保每个资源最多只能在一个节点上启动。每个节点都必须报告资源发生的每个更改。这样群集状态就是资源状态和节点状态的集合。

如果某些节点或资源的状态无法确定地建立（不论何种原因），就会采取屏障。即使当群集不了解某些节点上发生的状况时，屏障也可以确保这些节点没有运行任何重要资源。

8.1 屏障分类

有两类屏障：资源级别屏障和节点级别屏障。后者是本章的主题。

资源级别屏障
: 通过使用资源级别屏障，群集可以确保节点无法访问一个或多个资源。SAN就是一个典型的示例，屏障操作更改了 SAN 交换机上的规则从而拒绝节点的访问。

  通过使用要保护的资源所依赖的常规资源可以完成资源级别屏障。这种资源当然会拒绝在此节点上启动，所以依赖它的资源将不会在同一节点上运行。

节点级别屏障
: 节点级别屏障确保节点不会运行任何资源。通常用一种很简单但有点极端的方式来完成此种屏障：只需用电源开关重设置节点。这应该是必要的，因为节点可能根本不会作出响应。

8.2 节点级别屏障

在 SUSE® Linux Enterprise High Availability Extension 中，实现屏障的是 STONITH。它提供节点级别屏障。High Availability Extension 包括 `stonith` 命令行工具，一个能远程关闭群集中节点的可扩展界面。有关可用选项的概述，请运行 `stonith --help` 或参见 `stonith` 的手册页了解更多信息。

8.2.1 STONITH 设备

要使用节点级别屏障，首先需要有屏障设备。要获取 High Availability Extension 支持的 STONITH 设备的列表，请以 `root` 在任何节点上运行以下命令：

```
stonith -L
```

STONITH 设备可分为以下类别：

电源分配单元 (PDU)
: 电源分发单元是管理关键网络、服务器和数据中心设备的电源容量和功能的基本元素。它可以提供对已连接设备的远程负载监视和独立电源出口控制，以实现远程电源循环。

不间断电源 (UPS)
: 通过在公共用电不可用时从独立源供电，不间断电源为已连接设备提供紧急电源。

刀片电源控制设备
: 如果是在刀片组上运行群集，则刀片外壳中的电源控制设备就是提供屏障的唯一候选。当然，此设备必须能够管理单个刀片计算机。

无人值守设备
: 无人值守设备（IBM RSA、HP iLO、Dell DRAC）正变得越来越通用，将来甚至可能成为现成计算机的标准设备。然而，它们相比 UPS 设备有一点不足，因为它们与主机（群集节点）共享一个电源。如果节点没有电源，则设

想为控制节点的设备就等于没用。在那种情况下，CRM 屏障节点的尝试将是徒劳，而此情况将一直继续，因为所有其他资源操作将等待屏障/STONITH 操作成功完成。

测试设备
: 测试设备仅用于测试目的。它们通常对硬件更加友好。一旦群集进入生产阶段，它们必须替换为真实的屏障设备。

对 STONITH 设备的选择主要取决于您的预算和所用硬件的种类。

8.2.2 STONITH 实现

SUSE® Linux Enterprise High Availability Extension 的 STONITH 实现由两个组件组成：

stonithd
: stonithd 是一个守护程序，可以由本地进程或通过网络访问。它接受与屏障操作相应的命令：重设置、关闭电源和打开电源。它还可以检查屏障设备的状态。

  stonithd 守护程序在 CRM HA 群集中的每个节点上运行。在 DC 节点上运行的 stonithd 实例从 CRM 接收屏障请求。它会对请求作出响应，其他 stonithd 程序将执行所需的屏障操作。

STONITH 插件
: 对于每个受支持的屏障设备，都有一个能够控制此设备的 STONITH 插件。STONITH 插件是屏障设备的界面。所有 STONITH 插件都位于每个节点上的 `/usr/lib/stonith/plugins` 中。所有 STONITH 插件看上去都与 stonithd 一样，但显著区别在于反映了屏障设备的性质。

  某些插件支持多个设备。`ipmilan`（或 `external/ipmi`)就是一个典型的示例，它实施 IPMI 协议并可以控制任何支持此协议的设备。

8.3 STONITH 配置

要配置屏障，需要配置一个或多个 STONITH 资源，stonithd 守护程序不需要配置。所有配置都储存在 CIB 中。STONITH 资源就是 `stonith` 类的资源（请参见第 17.1 节 “支持的资源代理类” [181]）。STONITH 资源是 STONITH 插件在

CIB 中的代表。除了屏障操作，还可以启动、停止和镜像 STONITH 资源，就像任何其他资源一样。在这种情况下，启动或停止 STONITH 资源意味着启用或禁用 STONITH。启动和停止仅是管理操作，不会转换成对屏障设备自身的任何操作。但是，监视会转换成设备状态。

STONITH 资源可像任何其他资源一样进行配置。有关配置资源的更多信息，请参见第 4.3 节 “创建 STONITH 资源” [35]或第 5.4 节 “创建 STONITH 资源” [60]。

参数（属性）列表取决于相应的 STONITH 类型。要查看特定设备的参数列表，请使用 `stonith` 命令：

```
stonith -t stonith-device-type -n
```

例如，要查看 `ibmhmc` 设备类型的参数，请输入以下命令：

```
stonith -t ibmhmc -n
```

要获取设备的简短帮助文本，请使用 `-h` 选项：

```
stonith -t stonith-device-type -h
```

8.3.1 STONITH 资源配置示例

在以下部分中，可了解一些用 `crm` 命令行工具的语法编写的示例配置。要应用这些配置，请将示例放进文本文件（例如 `sample.txt`）并运行：

```
crm < sample.txt
```

有关使用 `crm` 命令行工具配置资源的更多信息，请参见第 5 章 *通过命令行配置群集资源* [55]。

警告：测试配置

以下一些示例仅用于演示和测试目的。请勿将任何测试配置示例用于真实的群集方案。

例 8.1 *测试配置*

```
configure
primitive st-null stonith:null \
params hostlist="node1 node2"
clone fencing st-null
commit
```

例 8.2 *测试配置*

备用配置：

```
configure
 primitive st-node1 stonith:null \
 params hostlist="node1"
 primitive st-node2 stonith:null \
 params hostlist="node2"
 location l-st-node1 st-node1 -inf: node1
 location l-st-node2 st-node2 -inf: node2
 commit
```

考虑到群集软件，此配置示例是完全正确的。与真实配置的唯一区别是没有发生屏障操作。

例 8.3 *测试配置*

以下 external/ssh 配置是一个更真实的示例，但仍然仅用于测试：

```
configure
 primitive st-ssh stonith:external/ssh \
 params hostlist="node1 node2"
 clone fencing st-ssh
 commit
```

此配置也可以重设置节点。此配置非常类似于针对空 STONITH 设备的第一个示例。在此示例中使用了克隆。这是 CRM/Pacemaker 的一个功能。克隆基本上是一种快捷方式：一个克隆的资源可以取代定义 n 个相同但名称不同的资源。到目前为止，克隆的最常用途是与 STONITH 资源一起使用（如果可以从所有节点访问 STONITH 设备）。

*例 **8.4** IBM RSA 无人值守设备的配置*

真实的设备配置没有太大区别，尽管某些设备可能要求更多属性。可以如下配置 IBM RSA 无人值守设备：

```
configure
primitive st-ibmrsa-1 stonith:external/ibmrsa-telnet \
params nodename=node1 ipaddr=192.168.0.101 \
userid=USERID passwd=PASSW0RD
primitive st-ibmrsa-2 stonith:external/ibmrsa-telnet \
params nodename=node2 ipaddr=192.168.0.102 \
userid=USERID passwd=PASSW0RD
location l-st-node1 st-ibmrsa-1 -inf: node1
location l-st-node2 st-ibmrsa-2 -inf: node2
commit
```

在此示例中使用了位置约束，原因如下：STONITH操作总有失败的可能性。因此，对作为操作执行者的节点的STONITH操作是不可靠的。如果重设置节点，则它将无法发送有关屏障操作结果的通知。唯一的方法是假设操作会成功并提前发送通知。但如果操作失败，则会遇到问题。所以stonithd拒绝按约定终止其主机。

例 8.5 *UPS 屏障设备的配置*

UPS 类型的屏障设备的配置类似于上述示例，细节就留给读者作为练习吧。所有 UPS 设备对屏障都使用相同的机制，但访问设备的方式是不同的。旧的 UPS 设备（被认为是专业的）只有一个串行端口，通常使用一根特殊的串行电缆以 1200 波特连接。许多新的 UPS 设备仍有一个串行端口，但通常它们还提供了一个 USB 接口或一个以太网接口。可以使用的连接类型取决于插件支持的连接。

例如，通过使用 `stonith -t` *`stonith 设备类型`* `-n` 命令比较 `apcmaster` 与 `apcsmart` 设备：

```
stonith -t apcmaster -h
```

返回以下信息：

```
STONITH Device: apcmaster - APC MasterSwitch (via telnet)
NOTE: The APC MasterSwitch accepts only one (telnet)
connection/session a time. When one session is active,
subsequent attempts to connect to the MasterSwitch will fail.
For more information see http://www.apc.com/
List of valid parameter names for apcmaster STONITH device:
ipaddr
login
 password
```

使用

```
stonith -t apcsmart -h
```

得到以下输出：

```
STONITH Device: apcsmart - APC Smart UPS
(via serial port - NOT USB!).
Works with higher-end APC UPSes, like
Back-UPS Pro, Smart-UPS, Matrix-UPS, etc.
(Smart-UPS may have to be >= Smart-UPS 700?).
See http://www.networkupstools.org/protocols/apcsmart.html
for protocol compatibility details.
For more information see http://www.apc.com/
List of valid parameter names for apcsmart STONITH device:
ttydev
hostlist
```

第一个插件支持带有一个网络端口的 APC UPS 和 telnet 协议。第二个插件使用 APC SMART 协议，而不使用许多不同的 APC UPS 产品系列都支持的串行线路。

8.3.2 约束与克隆

您在第 8.3.1 节 “STONITH 资源配置示例” [78]中了解了配置 STONITH 资源有多种方法：使用约束和/或克隆。选择哪种用于配置取决于多种因素（屏障设备的性质、设备管理的主机数、群集节点数）和最后但并非最不重要的一点，个人偏好。

简而言之：如果将克隆用于配置是安全的且如果它们减少了配置，则使用克隆的 STONITH 资源。

8.4 监视屏障设备

就像任何其他资源一样，STONITH 类代理也支持用于检查状态的监视操作。

注意：监视 STONITH 资源

强烈建议监视 STONITH 资源。定期但谨慎地进行监视。

屏障设备是 HA 群集不可缺少的组成部分，但越少需要使用它们越好。电源管理设备在通讯方面非常脆弱是众所周知的。如果在线路上有太多广播通讯，某些设备就会损坏。某些设备无法处理每分钟多于十个连接的情况。如果两个客户端同时尝试连接，某些设备就会陷入混乱。多数设备无法同时处理多个会话。

所以多数情况下，每几个小时检查一次屏障设备足够了。在这几个小时内需要执行屏障操作及电源开关出现故障的可能性通常很小。

有关如何配置监视操作的详细信息，对于 GUI 方法请参见添加或修改元属性和实例属性 [34]，对于命令行方法请参见第 5.8 节 “配置资源监视” [64]。

8.5 特殊的屏障设备

除了处理真实设备的插件，某些 STONITH 插件有点不寻常，值得特别关注。

`external/kdumpcheck`
: 有时候有一个内核转储是很重要的。此插件可用于检查转储是否在工作。如果在工作，则它将返回 true，就像节点已被屏障，因为它此时无法运行任何

资源，所以确是事实。kdumpcheck 通常与另一个真实的屏障设备一起使用。有关更多细节，请参见 /usr/share/doc/packages/heartbeat/stonith/README_kdumpcheck.txt。

external/sbd
这是一个自屏障设备。它对可以插入共享磁盘的所谓的“毒药”作出反应。当共享储存连接丢失时，它还使节点自终止。有关详细信息，请参见http://www.linux-ha.org/SBD_Fencing。

meatware
meatware 需要人为帮助才能运行。调用 meatware 时，它会记录一条 CRIT 严重性消息，显示在节点的控制台上。操作员随后需要确保节点已关闭并发出 meatclient(8) 命令。此命令会告诉 meatware 可以通知群集认为此节点已出现故障。有关更多信息，请参见 /usr/share/doc/packages/heartbeat/stonith/README.meatware。

null
这是一个用于各种测试方案的假设设备。它总是声明它已关闭某个节点，但其实未做任何操作。除非您了解您所执行的操作，否则请勿使用它。

suicide
这是一个仅有软件的设备，它可以使用 reboot 命令重引导它运行所处的节点。这需要节点的操作系统的操作，在某些情况下可能失败。因此，请尽量避免使用此设备（但它可用于一个节点的群集）。

suicide 和 null 是 “do not shoot my host（不关闭我的主机）”规则的唯一例外。

8.6 更多信息

/usr/share/doc/packages/heartbeat/stonith/
在已安装系统中，此目录包含许多 STONITH 插件和设备的自述文件。

http://linux-ha.org/STONITH
高可用性 Linux 项目的主页上有关 STONITH 的信息。

http://linux-ha.org/fencing
高可用性 Linux 项目的主页上有关屏障的信息。

http://linux-ha.org/ConfiguringStonithPlugins
高可用性 Linux 项目的主页上有关 STONITH 插件的信息。

http://linux-ha.org/CIB/Idioms
高可用性 Linux 项目的主页上有关 STONITH 的信息。

http://clusterlabs.org/wiki/Documentation，*Configuration 1.0 Explained*
说明了用于配置 Pacemaker 的概念。包含全面而非常详细的信息供参考。

http://techthoughts.typepad.com/managing_computers/2007/10/split-brain-quo.html
说明 HA 群集中节点分裂、仲裁人数和屏障的概念的文章。

使用 Linux Virtual Server 进行负载平衡

9

Linux Virtual Server (LVS) 的目标是提供一个基本框架，将网络连接定向到共享其工作负载的多台服务器。Linux Virtual Server 是服务器群集（一个或多个负载平衡器及多台用于运行服务的真实的服务器），对于外部客户端显示为一台大型的快速服务器。这种看上去像是单台服务器的服务器被称为*虚拟服务器*。Linux Virtual Server 作为一种高级负载平衡解决方案，可用于构建高度可缩放且高度可用的网络服务，如 Web、缓存、邮件、FTP、媒体和 VoIP 服务。

真实的服务器和负载平衡器可通过高速 LAN 或地理位置分散的 WAN 互相连接。负载平衡器可以将请求发送到不同的服务器，使群集的并行服务显示为单个 IP 地址上的虚拟服务，并且可以使用 IP 负载平衡技术或应用程序级别的负载平衡技术来发送请求。系统的可伸缩性通过在群集中透明地添加或去除节点来实现。高可用性通过检测节点或守护程序故障并相应地重配置系统来提供。

9.1 概念概述

LVS 由两个主要组件构成：

内核代码：ip_vs（或 IPVS）
: 运行已增补的 Linux 内核以包括 IPVS 代码的节点称为*控制器*。控制器上运行的 IPVS 代码是 LVS 的基本功能。

  客户端连接到控制器，控制器将包转发给真实的服务器。控制器是第 4 层路由器，具有能够使 LVS 工作的修改过的路由规则集（例如，连接未从控制器发起或终止，控制器就不会发送确认信息）。控制器与真实的服务器就是对客户端显示为一台服务器的虚拟服务器。

有不同的转发方法可确定控制器如何将包从客户端发送到真实的服务器。对于客户端请求的新连接，使用哪台真实的服务器是使用不同算法来实施的，这些算法可作为模块并可适应特定的需求。从客户端接收到连接请求时，控制器会根据*日程表*将一台真实的服务器指派给此客户端。调度程序是 IPVS 内核代码的组成部分，它决定哪台真实的服务器将获取下一个新连接。

用户空间控制器：ipvsadm

由 `ipvsadm` 包提供的 ipvsadm 是用于管理 Linux Virtual Server 的用户界面。例如，可以为处理的服务设置规则、处理故障转移或设置调度程序类型。

从命令行（或在 RC 文件中）使用 ipvsadm 可以设置：

- 控制器控制的服务/服务器（例如，http 转到所有真实的服务器，而 ftp 仅转到某台真实的服务器）
- 为每台真实的服务器指定的权重（当部分服务器快于其他服务器时，这很有用）
- 调度算法

还可以使用 ipvsadm 执行以下任务：

- 添加服务
- 关闭服务
- 删除服务

9.2 High Availability

要构造高度可用的 Linux Virtual Server 群集，可以使用软件的多种内置功能。通常，负载平衡器上会运行设备监视器守护程序，以定期检查服务器运行状况。如果服务器在指定时间内对服务访问请求或 ICMP ECHO_REQUEST 没有响应，则设备监视器将认为此服务器出现故障，并将其从负载平衡器的可用服务器列表中去除。这样就不会将新请求发送给这台出现故障的服务器了。如果设备监视器检测到发生故障的服务器已恢复并重新工作，则它会将此服务器重新添加到可用服务器列表中。因此，负载平衡器可以自动屏蔽服务守护程序或服务器的故障。

此外，管理员还可以使用系统工具添加新服务器以增加系统吞吐量，或去除服务器以进行系统维护，而无需关闭整个系统服务。

为防止负载平衡器成为整个系统的单一故障点，需要设置负载平衡器的备份（或多个备份）。两个检测信号守护程序分别运行在主平衡器和备份平衡器上。它们通过串行线路和/或网络接口定期互相发送“正常运行”检测信号讯息。如果备份平衡器的检测信号守护程序在指定时间内未收到来自主平衡器的检测信号讯息，它将接管虚拟 IP 地址以提供负载平衡服务。

发生故障的负载平衡器重新恢复后，有两种可能的结果：它可以自动成为备份负载平衡器，或活动的负载平衡器释放 VIP 地址，这样恢复的平衡器将接管 VIP 地址并重新成为主负载平衡器。主负载平衡器具有连接状态，即表示它知道将连接转发给哪台服务器。如果备份负载平衡器在接管时没有收到此连接信息，则客户端必须重新发送其请求后才能访问服务。为使负载平衡器故障转移对客户端应用程序透明，在 IPVS 中会进行连接同步：主 IPVS 负载平衡器通过 UDP 多路广播将连接信息同步到备份负载平衡器。当备份负载平衡器在主负载平衡器发生故障后接管它时，备份负载平衡器将具有大多数连接状态，这样几乎所有连接都可以通过备份负载平衡器继续访问服务。

9.3 更多信息

要了解更多有关 Linux Virtual Server 的信息，请参阅 http://www.linuxvirtualserver.org/ 上的项目主页。

10 网络设备联接

对于许多系统，需要实施符合典型以太网设备的标准数据安全性或可用性要求的网络连接。在这种情况下，可以将多种以太网设备聚合到单个联接设备。

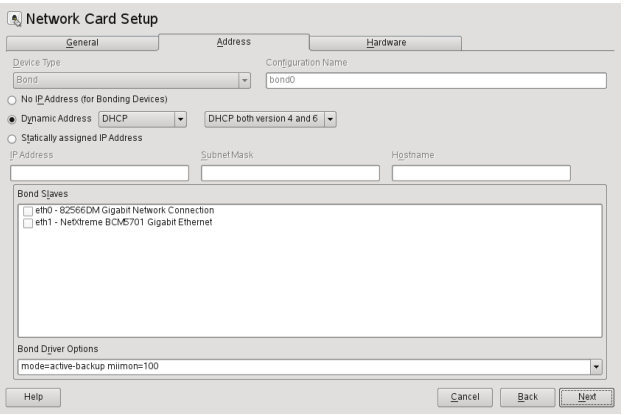
联接设备的配置通过联接模块选项来完成。其行为取决于联接设备的模式。默认情况下是 `mode=active-backup`，即如果活动从属设备发生故障，则其他从属设备将变成为活动设备。

使用 OpenAIS 时，不通过群集软件来管理联接设备。因此，必须在每个可能需要访问联接设备的群集节点上配置联接设备。

要配置联接设备，请使用以下过程：

1 运行 *YaST* > *网络设备* > *网络设置*。

2 使用*添加*并将*设备类型*更改为*联接*。按*下一步*继续。



3 选择如何为联接设备指派 IP 地址。有三种方法可供选择：

- 无 IP 地址
- 动态地址（使用 DHCP 或 Zeroconf）
- 静态指派的 IP 地址

使用最适合您环境的方法。如果 OpenAIS 管理虚拟 IP 地址，请选择*静态指派的 IP 地址*并在界面上指派一个基本 IP 地址。

4 选择需要包含在联接中的以太网设备，方法是激活相关*联接从属设备*前的复选框。

5 编辑*联接驱动程序选项*。可以使用以下模式：

balance-rr
: 提供负载平衡和容错。

active-backup
: 提供容错

balance-xor
: 提供负载平衡和容错。

broadcast
: 提供容错

802.3ad
提供动态链接集合（如果连接的交换机支持动态链接集合）。

balance-tlb
为外发的通讯量提供负载平衡。

balance-alb
为进来的和外发的通讯量提供负载平衡（如果使用的网络设备允许在使用中修改网络设备的硬件地址）。

6 确保将参数 `miimon=100` 添加到*联接驱动程序选项*。如果没有此参数，则不会定期检查数据完整性。

7 单击*下一步*，然后单击*确定*退出 YaST 以创建设备。

*Linux 以太网联接驱动程序操作指南*中详细解释了所有模式及许多其他选项，在安装 `kernel-source` 包后可以在 `/usr/src/linux/Documentation/networking/bonding.txt` 上找到这些内容。

将群集更新为 SUSE Linux Enterprise 11

11

如果现有群集基于 SUSE® Linux Enterprise Server 10 SP2，则可以将群集更新为在 SUSE® Linux Enterprise Server 11 上与 High Availability Extension 一起运行。为了进行迁移，所有群集节点都必须处于脱机状态，且必须将群集作为整体迁移—不支持混合的 SUSE Linux Enterprise Server 10/SUSE Linux Enterprise Server 11 群集。

为方便起见，SUSE® Linux Enterprise High Availability Extension 包括一个 `hb2openais.sh` 脚本，使用它可在从 Heartbeat 移动到 OpenAIS 群集堆栈的同时转换数据。脚本会分析储存在 `/etc/ha.d/ha.cf` 中的配置，并为 OpenAIS 群集堆栈生成新的配置文件。此外，它调整 CIB 以匹配 OpenAIS 约定、转换 OCFS2 文件系统以及用 cLVM 替换 EVMS。

要成功将群集从 SUSE Linux Enterprise Server 10 SP2 迁移到 SUSE Linux Enterprise Server 11，需要执行以下步骤：

1. 准备 SUSE Linux Enterprise Server 10 SP2 群集 [94]
2. 更新到 SUSE Linux Enterprise 11 [95]
3. 测试转换 [95]
4. 转换数据 [96]

成功完成转换后，可以重新使更新后的群集联机。

注意：更新后还原

更新到 SUSE Linux Enterprise Server 11 后，就*不*支持还原回 SUSE Linux Enterprise Server 10。

11.1 准备和备份

在将群集更新为最新的产品版本并相应地转换数据前，需要准备当前群集。

过程 11.1 *准备 SUSE Linux Enterprise Server 10 SP2 群集*

1. 登录到群集。
2. 查看检测信号配置文件 `/etc/ha.d/ha.cf` 并检查是否所有通讯媒体都支持多路广播。
3. 确保以下文件在所有节点上都是相同的：`/etc/ha.d/ha.cf` 和 `/var/lib/heartbeat/crm/cib.xml`。
4. 在每个节点上执行 `rcheartbeat stop` 以使所有节点都脱机。
5. 在更新到最新版本前，除了建议进行常规的系统备份，还建议备份以下文件，因为在更新到 SUSE Linux Enterprise Server 11 后运行转换脚本时需要使用这些文件：
   - `/var/lib/heartbeat/crm/cib.xml`
   - `/var/lib/heartbeat/hostcache`
   - `/etc/ha.d/ha.cf`
   - `/etc/logd.cf`

11.2 更新/安装

准备好群集并备份完文件后，就可以开始将群集节点更新到最新的产品版本。如果不运行更新，也可以在群集节点上进行全新的 SUSE Linux Enterprise 11 安装。

过程 11.2 *更新到 SUSE Linux Enterprise 11*

1. 在所有群集节点上执行从 SUSE Linux Enterprise Server 10 SP2 到 SUSE Linux Enterprise Server 11 的更新。有关如何更新产品的信息，请参阅 SUSE Linux Enterprise Server 11 *部署指南*，*更新 SUSE Linux Enterprise* 一章。

   如果不进行更新，也可以在所有群集节点上全新安装 SUSE Linux Enterprise Server 11。

2. 在所有群集节点上，将 SUSE Linux Enterprise High Availability Extension 11 安装为 SUSE Linux Enterprise Server 之上的外接式附件。有关详细信息，请参见第 3.1 节 “安装 High Availability Extension” [21]。

11.3 数据转换

安装完 SUSE Linux Enterprise Server 11 和 High Availability Extension 后，可以开始数据转换。High Availability Extension 附带的转换脚本已谨慎设置过，但是它无法在全自动模式下处理所有设置。它会其所作更改显示警报，但是需要您进行干预和决策。您需要详细地了解群集—因为由您来校验更改是否有意义。转换脚本位于 `/usr/lib/heartbeat`（如果使用 64 位系统，则位于 `/usr lib64/heartbeat`）。

注意：执行测试运行

要熟悉转换进程，我们强烈建议您首先测试一下转换（不作任何更改）。可以使用同一测试目录执行重复的测试运行，但是只需要复制一次文件。

过程 11.3 *测试转换*

1. 在某个节点上创建测试目录，并将备份文件复制到此测试目录：

```
$ mkdir /tmp/hb2openais-testdir
$ cp /etc/ha.d/ha.cf /tmp/hb2openais-testdir
```

```
$ cp /var/lib/heartbeat/hostcache /tmp/hb2openais-testdir
$ cp /etc/logd.cf /tmp/hb2openais-testdir
$ sudo cp /var/lib/heartbeat/crm/cib.xml /tmp/hb2openais-testdir
```

2 使用以下命令开始测试运行

```
$ /usr/lib/heartbeat/hb2openais.sh -T /tmp/hb2openais-testdir -U
```

如果使用 64 位系统，请使用以下命令：

```
$ /usr/lib64/heartbeat/hb2openais.sh -T /tmp/hb2openais-testdir -U
```

3 阅读并校验生成的 `openais.conf` 和 `cib-out.xml` 文件：

```
$ cd  /tmp/hb2openais-testdir
$ less openais.conf
$ crm_verify -V -x cib-out.xml
```

有关转换阶段的详细信息，请参阅安装的 High Availability Extension 中的 `/usr/share/doc/packages/pacemaker/README.hb2openais`。

过程 11.4 转换数据

执行测试运行并检查输出后，可以立即开始数据转换。只需在一个节点上运行转换。主群集配置 (CIB) 会自动复制到其他节点。需要复制的所有其他文件会由转换脚本自动进行复制。

1 确保所有节点上都在运行 sshd，并允许 `root` 访问所有节点，以便转换脚本将文件成功复制到其他群集节点。

2 High Availability Extension 附带了一个默认的 OpenAIS 配置文件。如果要防止在后面的步骤中重写此默认配置，请制作 `/etc/ais/openais.conf` 配置文件的副本。

3 以 `root` 身份启动转换脚本。如果使用 sudo，请使用 `-u` 选项指定特权用户：

```
$ /usr/lib/heartbeat/hb2openais.sh -u root
```

基于储存在 `/etc/ha.d/ha.cf` 中的配置，脚本将为 OpenAIS 群集堆栈生成新的配置文件，`/etc/ais/openais.conf`。由于从 Heartbeat 更改到 OpenAIS，它还将分析 CIB 配置并让您了解群集配置是否需要更改。在转换运行的节点上完成所有文件处理，并将文件处理复制到其他节点。

4 按照屏幕指导执行操作。

成功完成转换后，按照第 3.3 节 “使群集联机” [24]中所述启动新的群集堆栈。

升级进程完成后，就不支持还原回 SUSE Linux Enterprise Server 10。

11.4 更多信息

有关转换脚本和转换阶段的更多细节，请参阅安装的 High Availability Extension 中的 `/usr/share/doc/packages/pacemaker/README.hb2openais`。

部分 III. 储存和数据复制

12 Oracle Cluster File System 2

Oracle Cluster File System 2 (OCFS2) 是一个一般用途的日志文件系统，它完全集成到 Linux 2.6 和更高版本的内核中。OCFS2 可将应用程序二进制文件、数据文件和数据库储存到共享储存设备。群集中的所有节点对文件系统都有并行的读和写权限。通过克隆资源管理的用户空间控制守护程序提供了与 HA 堆栈（尤其是 OpenAIS 和 DLM）的集成。

12.1 功能和优点

在 SUSE Linux Enterprise Server 10 和更高版本中，OCFS2 可用于以下储存解决方案：

- 一般应用程序和工作负荷
- 在群集中储存的 XEN 图形

  XEN 虚拟计算机和虚拟服务器可以储存在由群集服务器安装的 OCFS2 卷上，以在服务器之间提供 XEN 虚拟机的快速、简便性。
- LAMP（Linux、Apache、MySQL 和 PHP | Perl | Python）堆栈

此外，它还可以与 OpenAIS 完全集成。

作为一个高性能、均衡的并行群集文件系统，OCFS2 支持以下功能：

- 群集中的所有节点都可以使用应用程序文件。用户只需在群集中的 OCFS2 上安装它一次。

- 所有节点可以通过标准的文件系统接口直接同时读写到存储器，方便地管理运行在群集中的应用程序。
- 通过分布式锁管理器（DLM）协调文件访问。

  DLM 控制适合大多数情况，但是如果应用程序的设计与 DLM 竞争来协调文件的访问，则可能会限制其可测量性。
- 所有后端储存上都可以使用储存备份功能。可以方便地创建共享应用程序文件的图形，它能够帮助提供有效的故障恢复。

OCFS2 还提供以下功能：

- 元数据缓存。
- 元数据日记。
- 跨节点的文件数据一致性。
- 支持最多 4 KB 的多个块大小（每个卷可有不同的块大小），最大卷大小为 16 TB。
- 支持最多 16 个群集节点。
- 对于数据库文件的异步和直接 I/O 支持，提高了数据库性能。

12.2 管理实用程序和命令

下表中描述了 OCFS2 实用程序。有关这些命令的语法信息，请参阅它们的手册页。

表 12.1 *OCFS2 实用程序*

OCFS2 实用程序	说明
debugfs.ocfs2	为了调试检查 OCFS 文件系统的状态
fsck.ocfs2	检查文件系统的错误并进行选择性的修改。

OCFS2 实用程序	说明
mkfs.ocfs2	在某个设备上创建 OCFS2 文件系统，通常是共享物理或逻辑磁盘上的某个分区。
mounted.ocfs2	检测并列出群集系统上所有的 OCFS2 卷。检测并列出已经安装了 OCFS2 设备的系统上的所有节点或列出所有的 OCFS2 设备。
tunefs.ocfs2	更改 OCFS2 文件系统参数，包括卷标、节点槽号、所有节点槽的日志大小和卷大小。

12.3 OCFS2 包

SUSE Linux Enterprise Server 11 HAE 中自动安装了 OCFS2 内核模块 (`ocfs2`)。要使用 OCFS2，请使用 YaST（如果需要，也可以使用命令行）在群集中的每个节点上安装 `ocfs2-tools` 及匹配的 `ocfs2-kmp-*` 包。

1. 以 `root` 用户身份或等价的用户身份登录，然后打开 YaST 控制中心。
2. 选择*软件* > *外接式附件产品*。
3. 选择*添加*使 SUSE Linux Enterprise High Availability Extension 可用。
4. 选中*运行软件管理器*，然后选择*过滤器* > *模式*。
5. 选择*高可用性*安装模式。
6. 单击*接受*并遵照屏幕指导操作。

12.4 创建 OCFS2 卷

按照本部分中的过程配置您的系统以使用 OCFS2 并创建 OCFS2 卷

12.4.1 前提条件

开始操作之前，请完成以下操作：

- 准备计划用于 OCFS2 卷的块设备。将这些设备留作可用空间。

  建议您在不同的 OCFS2 卷上存储应用程序文件和数据文件，但是，如果您的应用程序卷和数据卷有不同的装入要求，则不强制您这样做。

- 确定 `ocfs2-tools` 包已安装。使用 YaST 或命令行方法安装它们（如果它们不存在）。有关 YaST 的说明，请参阅第 12.3 节 “OCFS2 包” [103]。

12.4.2 配置 OCFS2 服务

您必须首先配置 OCFS2 服务，才能创建 OCFS2 卷。

为群集中的某个节点执行本部分描述的过程。

1 打开终端窗口并以 `root` 用户身份或相当的用户身份登录。

2 添加分布式锁管理器配置。

2a 启动 crm 壳层，从头创建新配置：

```
crm
cib new stack-glue
```

2b 创建 DLM 服务，并使它在群集中的所有计算机上运行。

```
configure
primitive dlm ocf:pacemaker:controld op monitor interval=120s
clone dlm-clone dlm meta globally-unique=false interleave=true
end
```

2c 在提交之前，请校验对群集所做的更改：

```
cib diff
configure verify
```

2d 将配置上载到群集并退出壳层：

```
cib commit stack-glue
quit
```

3 使用 *crm* 添加 O2CB 配置。

3a 启动 crm 壳层，从头创建新配置：

```
crm
cib new oracle-glue
```

3b 配置 Pacemaker 以在群集中的所有节点上启动 o2cb 服务。

```
configure
primitive o2cb ocf:ocfs2:o2cb op monitor interval=120s
clone o2cb-clone o2cb meta globally-unique=false interleave=true
```

3c 确保 Pacemaker 仅在同样包含已在运行的 dlm 服务的节点上启动 o2cb 服务：

```
colocation o2cb-with-dlm INFINITY: o2cb-clone dlm-clone
order start-o2cb-after-dlm mandatory: dlm-clone o2cb-clone
end
```

3d 将配置上载到群集并退出壳层：

```
cib commit stack-glue
quit
```

12.4.3 创建 OCFS2 卷

应该仅在群集中的某个节点上执行创建 OCFS2 文件系统并将新节点添加到群集。

1 打开终端窗口并以 `root` 用户身份或相当的用户身份登录。

2 检查群集是否与 crm_mon 命令联机。

3 使用以下方法之一创建和格式化卷：

- 使用 mkfs.ocfs2 实用程序。有关此命令语法的信息，请参阅 mkfs.ocfs2 手册页。

  要在支持最多 16 个群集节点的 /dev/sdb1 上创建新的 OCFS2 文件系统，请使用以下命令：

  ```
  mkfs.ocfs2 -N 16 /dev/sdb1
  ```

请参阅下表以获得建议的设置。

OCFS2 参数	描述和建议
卷标	卷的描述性名称能够在不同节点上安装卷时唯一标识它。 使用 tunefs.ocfs2 实用程序根据需要修改该卷标。
群集大小	群集大小是分配给文件以保存数据的最小空间单元。 选项是 4、8、16、32、64、128、256、512 和 1024 KB。格式化卷后不能再修改群集大小了。 Oracle 建议数据库卷的群集大小是 128 KB 或更大。Oracle 还建议 Oracle Home 的群集大小是 32 或 64 KB。
节点槽的号码	可以同时安装卷的最大节点数。OCFS2 会为每个节点创建单独的系统文件（如日记）。访问卷的节点可以是小尾端结构（如 x86 x86-64 和 ia64）和大尾端结构（如 ppc64 和 s390x)的组合。 特定于节点的文件作为本地文件。节点槽号附加到该本地文件。例如：journal:0000 属于任何槽号为 0 的节点。 当您创建卷时，要根据您希望同时安装卷的节点数，设置每个卷的最大节点槽号。使用 tunefs.ocfs2 实用程序根据需要增加节点槽号；该值不能减少。

OCFS2 参数	描述和建议
块大小	文件系统可寻址的最小空间单元创建卷时请指定块大小。 选项有 512 字节（不建议使用）、1 KB、2 KB 或 4 KB（对大多数卷建议使用）。格式化卷后不能再修改块大小了。

12.5 装入 OCFS2 卷

1 打开终端窗口并以 `root` 用户身份或相当的用户身份登录。

2 检查群集是否与 `crm_mon` 命令联机。

3 使用以下某个方法装入卷

警告：手动装入的 OCFS2 设备

如果您出于测试目的手动安装了 ocfs2 文件系统，则开始使用它之前，需要通过 OpenAIS 再次卸载该文件系统。

- 在 `ocfs2console` 中，在可用设备列表中选择一个设备，单击*安装*，指定目录装入点和装入选项（可选），然后单击*确定*。
- 使用 `mount` 命令从命令行装入卷。
- 使用群集管理器装入文件系统。可使用 ocf 资源 `Filesystem` 来完成此任务。有关细节，请参阅使用群集管理器安装文件系统 [108]。

成功安装后，`ocfs2console` 中的设备列表显示装入点和设备。

选项	描述
`datavolume`	确保 Oracle 进程打开具有 o_direct 标志的文件。

选项	描述
nointr	没有中断。确保 IO 未被信号中断。

要配置文件系统资源，请使用以下步骤：

过程 12.1 *使用群集管理器安装文件系统*

1. 启动 crm 壳层，从头创建新配置：

```
crm
cib new filesystem
```

2. 配置 Pacemaker 以在群集中的所有节点上安装文件系统。

```
configure
primitive fs ocf:heartbeat:Filesystem \
   params device="/dev/sdb1" directory="/mnt/shared" fstype="ocfs2" \
   op monitor interval=120s
clone fs-clone fs meta interleave="true" ordered="true"
```

3. 确保 Pacemaker 仅在同样包含已在运行的 o2cb 资源的节点上启动 o2cb 克隆资源：

```
colocation fs-with-o2cb INFINITY: fs-clone o2cb-clone
order start-fs-after-o2cb mandatory: o2cb-clone fs-clone
end
```

4. 将配置上载到群集并退出壳层：

```
cib commit filesystem
quit
```

12.6 其他信息

有关使用 OCFS2 的信息，请参阅 OCFS2 project at Oracle [http://oss.oracle.com/projects/ocfs2/]（Oracle 上的 OCFS2 项目）处的 *OCFS2 User Guide [http://oss.oracle.com/projects/ocfs2/documentation/]*（OCFS2 用户指南）。

13 群集 LVM

当管理群集上的共享储存时，所有节点必须收到有关对储存子系统所做更改的通知。Linux Volume Manager 2 (LVM2)（广泛应用于管理本地储存）的功能已扩展为支持透明管理整个群集中的卷组。可使用与本地储存相同的命令来管理群集卷组。

要实施 cLVM，共享储存设置（如 Fibre Channel 提供的 FCoE、SCSI、iSCSI SAN 或 DRBD）必须可用。

cLVM 在内部使用群集堆栈的 Distributed Lock Manager (DLM) 组件来协调对 LVM2 元数据的访问。DLM 会依次与堆栈的其他组件（如屏障）集成，因此共享储存的完整性将一直受到保护。

cLVM 不会协调对共享数据自身的访问；要执行此操作，必须在 cLVM 管理的储存的基础上配置 OCFS2 或其他群集感知应用程序。

13.1 cLVM 配置

要设置群集感知卷组，必须成功完成以下几项任务：

1 将 LVM2 的锁定类型更改为群集感知的。

编辑文件 `/etc/lvm/lvm.conf` 并找到以下行：

```
locking_type = 1
```

将锁定类型更改为 `3`，并将配置写入磁盘。将此配置复制到所有节点。

2 将 clvmd 资源作为克隆品包含在 Pacemaker 配置中，并让它依赖于 DLM 克隆资源。Crm 配置壳层中的典型片段将按如下显示：

```
primitive dlm ocf:pacemaker:controld
primitive clvm ocf:lvm2:clvmd \
        params daemon_timeout="30"
clone dlm-clone dlm \
        meta target-role="Started" interleave="true" ordered="true"
clone clvm-clone clvm \
        meta target-role="Started" interleave="true" ordered="true"
colocation colo-clvm inf: clvm-clone dlm-clone
order order-clvm inf: dlm-clone clvm-clone
...
```

在开始之前，确认这些资源已在您的群集中成功启动。您可以使用crm_mon 或 GUI 来检查正在运行的服务。

3 使用以下命令准备 LVM 的物理卷：

```
pvcreate <physical volume path>
```

4 创建群集感知卷组：

```
vgcreate --clustered y <volume group name> <physical volume path>
```

5 根据需要创建逻辑卷，例如：

```
lvcreate --name testlv -L 4G <volume group name>
```

6 要确保在群集范围内激活卷组，请按如下方式配置 LVM 资源：

```
primitive vg1 ocf:heartbeat:LVM \
        params volgrpname="<volume group name>"
clone vg1-clone vg1 \
        meta interleave="true" ordered="true"
colocation colo-vg1 inf: vg1-clone clvm-clone
order order-vg1 inf: clvm-clone vg1-clone
```

7 如果希望只在一个节点上激活卷组，可使用以下示例；在这种情况下，cLVM 将防止在多个节点上激活 VG 内的所有逻辑卷，作为非群集应用程序的附加保护措施：

```
primitive vg1 ocf:heartbeat:LVM \
```

```
        params volgrpname="<volume group name>" exclusive="yes"
colocation colo-vg1 inf: vg1 clvm-clone
order order-vg1 inf: clvm-clone vg1
```

8 现在VG内的逻辑卷可作为文件系统装入或原始用法提供。确保使用逻辑卷的服务必须具备适当的依赖性，以便在激活VG后对它们进行排列和排序。

完成这些配置步骤后，即可像在任何独立工作站中一样进行 LVM2 配置。

13.2 显式配置合格的 LVM2 设备

当一些设备看似共享同一个物理卷签名（例如对于多路径设备或 drbd 来说）时，建议显式配置 LVM2 扫描 PV 的设备。

例如，如果命令 `vgcreate` 使用的是物理设备而不是镜像块设备，DRBD 将会感到困惑，并可能导致 DRBD 的裂脑情况。

要停用 LVM2 的单个设备，请执行以下操作：

1 编辑文件 `/etc/lvm/lvm.conf` 并搜索以 `filter` 开头的行。

2 其中的模式作为正则表达式来处理。前面的“a”表示接受扫描的设备模式，前面的“r”表示拒绝遵守该设备模式的设备。

3 要删除名为 `/dev/sdb1` 的设备，请向过滤规则添加以下表达式：

```
"r|^/dev/sdb1$|"
```

完整的过滤行将显示如下：

```
filter = [ "r|^/dev/sdb1$|", "r|/dev/.*/by-path/.*|",
"r|/dev/.*/by-id/.*|", "a/.*/" ]
```

接受 DRBD 和 MPIO 设备但拒绝所有其他设备的过滤行将显示如下：

```
filter = [ "a|/dev/drbd.*|", "a|/dev/.*/by-id/dm-uuid-mpath-.*|",
"r/.*/" ]
```

4 编写配置文件并将它复制到所有群集节点。

13.3 更多信息

Pacemaker 邮件列表中提供了更多信息，地址为：http://www.clusterlabs.org/wiki/Help:Contents。

官方 cLVM 常见问题可在以下网址中找到：http://sources.redhat.com/cluster/wiki/FAQ/CLVM。

14 分布式复制块设备 (DRBD)

通过 DRBD，您可以为位于 IP 网络上两个不同站点的两个块设备创建镜像。当用于 OpenAIS 时，DRBD 支持分布式高可用性 Linux 群集。

重要

镜像之间的数据通讯是不加密的。为了保护数据交换，您应为该连接部署虚拟专用网 (VPN) 解决方案。

主设备的数据将以确保两个数据副本始终相同的方式复制到次设备。当使用群集感知文件系统（如 ocfs2）时，也可能将两个节点都作为主设备来运行。

默认情况下，DRBD 使用 TCP 端口 7788 在 DRBD 节点之间进行通讯。请确定您的防火墙不会阻止此端口上的通讯。

在 DRBD 设备上创建文件系统之前，必须先设置 DRBD 设备。只能通过 `/dev/drbd<n>` 设备（而非原始设备）操纵用户数据，因为 DRBD 使用原始设备的最后 128 MB 储存元数据。确保仅在 `/dev/drbd<n>` 设备上创建文件系统，而不在原始设备上创建。

例如，如果原始设备大小为 1024 MB，则 DRBD 设备仅有 896 MB 可用于数据，128 MB 隐藏并保留用于元数据。任何访问 896 MB 和 1024 MB 之间的空间的尝试都会失败，因为它不可用于用户数据。

14.1 安装 DRBD 服务

根据第 I 部分 “安装和设置” [1]中所述，要安装所需的 drbd 包，需要在联网的群集中的两台 SUSE Linux Enterprise Server 计算机上都安装 High Availability Extension 外接式附件产品。安装 High Availability Extension 还会安装 drbd 程序文件。

如果您不需要完整的群集堆栈而只想使用 drbd，也可以添加 High Availability Extension 外接式附件产品，并继续安装 drbd。这同样也会安装所需的内核模块。

14.2 配置 DRBD 服务

注意

以下过程使用服务器名节点 1 和节点 2，以及群集资源名称 r0。它将节点 1 设置为主节点。确保修改这些说明，以使用您自己的节点和文件名。

1. 启动 YaST 并选择配置模块*杂项* > *drbd*。
2. 在*启动配置* > *引导*中，选择*开启*始终在引导时启动 drbd。
3. 如果您需要配置多个复制资源，请选择*全局配置*。输入字段*次要计数*用于选择在无需重启动计算机的前提下可以配置的不同 drbd 资源数。
4. 资源的实际配置可在*资源配置*中完成。按*添加*可创建新资源。必须设置以下参数：

资源名称	资源的名称，通常称为 r0。
名称	各个节点的主机名。
地址:端口	各个节点的 IP 地址和端口号。
设备	保管各个节点上的复制数据的设备。此设备可用于创建文件系统和装入操作。

磁盘	在两个节点之间进行复制的设备。
元磁盘	可将*元磁盘*的值设为 `internal`，或指定一个由索引扩展的明确设备来保管 drbd 所需的元数据。 当使用 `internal` 时，复制设备的最后 128 MB 可用于储存元数据。 实际设备也可用于多个 drbd 资源。例如，如果第一个资源的*元磁盘*为 `/dev/sda6[0]`，则可以将 `/dev/sda6[1]` 用于第二个资源。但是，必须为此磁盘上的每个资源保留至少 128 MB 空间。

所有这些选项在 `/usr/share/doc/packages/drbd/drbd.conf` 文件的示例和 `drbd.conf(5)` 的手册页中均有说明。

5 将 `/etc/drbd.conf` 文件复制到二级服务器（节点 2）上的 `/etc/drbd.conf` 位置。

```
scp /etc/drbd.conf <node 2>:/etc
```

6 通过在每个节点上输入以下命令，初始化并启动两个系统上的 DRBD 服务：

```
drbdadm create-md r0
rcdrbd start
```

7 通过在 `node1` 上输入以下命令，将 `node1` 配置为主节点：

```
drbdsetup /dev/drbd0 primary --overwrite-data-of-peer
```

8 通过在每个节点上输入以下命令，检查 DRBD 服务状态：

```
rcdrbd status
```

继续之前，等待两个节点上的块设备完全同步。重复 `rcdrbd status` 命令以跟踪同步进度。

9 两个节点上的块设备都完全同步之后，使用诸如 reiserfs 的文件系统格式化主节点上的 DRBD 设备。可以使用任何 Linux 文件系统。例如，输入

```
mkfs.reiserfs -f /dev/drbd0
```

重要

请在命令中始终使用 /dev/drbd<n> 名称，而不是实际 /dev/disk 设备名。

14.3 测试 DRBD 服务

如果安装和配置过程和预期一样，则您就准备好运行DRBD功能的基本测试了。此测试还有助于了解该软件的工作原理。

1 在节点 1 上测试 DRBD 服务。

1a 打开终端控制台并作为 root 用户或具有相同权限的用户登录。

1b 通过输入以下命令，在节点 1 上创建安装点，例如 /srv/r0mount

```
mkdir -p /srv/r0mount
```

1c 通过输入以下命令，装入 drbd 设备

```
mount -o rw /dev/drbd0 /srv/r0mount
```

1d 通过输入以下命令，从主节点创建一个文件

```
touch /srv/r0mount/from_node1
```

2 在节点 2 上测试 DRBD 服务。

2a 打开终端控制台并作为 root 用户或具有相同权限的用户登录。

2b 通过在节点 1 上输入以下命令，卸下节点 1 上的磁盘：

```
umount /srv/r0mount
```

2c 通过在节点 1 上输入以下命令，降级节点 1 上的 DRBD 服务：

```
drbdadm secondary r0
```

2d 在节点 2 上，通过输入以下命令，将 DRBD 服务提升为主节点

```
drbdadm primary r0
```

2e 在节点 2 上，通过输入以下命令，检查节点 2 是否是主节点

```
rcdrbd status
```

2f 通过输入以下命令，在节点 2 上创建安装点，例如 `/srv/r0mount`

```
mkdir /srv/r0mount
```

2g 在节点 2 上，通过输入以下命令，装入 DRBD 设备

```
mount -o rw /dev/drbd0 /srv/r0mount
```

2h 通过输入以下命令，校验您在节点 1 上创建的文件可查看。

```
ls /srv/r0mount
```

`/srv/r0mount/from_node1` 文件应列出。

3 如果该服务在两个节点上都运行正常，则 DRBD 安装即已完成。

4 再次将节点 1 设置为主节点。

4a 通过在节点 2 上输入以下命令，卸下节点 2 上的磁盘：

```
umount /srv/r0mount
```

4b 通过在节点 2 上输入以下命令，降级节点 2 上的 DRBD 服务：

```
drbdadm secondary r0
```

4c 在节点 1 上，通过输入以下命令，将 DRBD 服务提升为主节点

```
drbdadm primary r0
```

4d 在节点 1 上，通过输入以下命令，检查节点 1 是否是主节点

```
rcdrbd status
```

5 要使服务在服务器有问题时自动启动并故障转移，可以使用 OpenAIS 将 DRBD 设置为高可用性服务。有关为 SUSE Linux Enterprise 11 安装和配置 OpenAIS 的信息，请参见第 II 部分 “配置和管理” [27]。

14.4 DRBD 查错

drbd 设置涉及很多不同的组件，这些不同的源可能产生问题。以下几节中介绍了一些常见问题，并提供了解决这些问题的提示。

14.4.1 配置

如果初始 `drbd` 安装未能和预期一样，则配置中可能出错了。

获取关于配置的信息：

1 打开终端控制台并作为 `root` 用户登录。

2 通过运行 `drbdadm`（带 `-d` 选项），测试配置文件。输入

```
drbdadm -d adjust r0
```

在 `adjust` 选项的干运行中，`drbdadm` 将 DRBD 资源的实际配置与您的 DRBD配置文件进行比较，但它不会执行这些调用。检查输出以确保您了解任何错误的根源。

3 如果 drbd.conf 文件中有任何错误，请先更正它们再继续。

4 如果分区和设置正确，请在不使用 -d 的情况下再次运行 drbdadm。输入

```
drbdadm adjust r0
```

这会将配置文件应用到 DRBD 资源。

14.4.2 主机名

对于 DRBD，主机名区分大小写，因此 Node0 和 node0 不是一个主机。

如果您具有几个网络设备，并希望使用一个专用的网络设备，则主机名将可能不会解析为已用的 IP 地址。在这种情况下，可使用参数 disable-ip-verification 使 DRBD 忽略此事实。

14.4.3 TCP 端口 7788

如果您的系统无法连接到同级，则本地防火墙可能有问题。默认情况下，DRBD 使用 TCP 端口 7788 访问另一个节点。确保在两个节点上该端口均可访问。

14.4.4 DRBD 设备在重引导后损坏

如果 DRBD 不知道哪个实际设备保管了最新数据，它就会变为节点分裂状态。在这种情况下，DRBD 子系统将分别成为次系统，并且互不相连。在这种情况下，会将以下消息写入 /var/log/messages：

```
Split-Brain detected, dropping connection!
```

要解决这种情况，必须选择一个应丢弃其修改的节点。登录到该节点，并运行以下命令：

```
drbdadm secondary r0
drbdadm -- --discard-my-data connect r0
```

在另一个节点上，运行以下命令：

```
drbdadm connect r0
```

14.5 其他信息

以下开放源代码资源可用于 DRBD：

- 该分发包中有 DRBD 的以下手册页：

  ```
  drbd(8)
  drbddisk(8)
  drbdsetup(8)
  drbdadm(8)
  drbd.conf(5)
  ```

- 可在 `/usr/share/doc/packages/drbd/drbd.conf` 中查找注释过的 DRBD 示例配置

- 项目主页：http://www.drbd.org。

- Linux Pacemaker 群集堆栈项目的 http://clusterlabs.org/wiki/DRBD_HowTo_1.0。

部分 IV. 查错和参考

15 查错

尤其是在开始使用 Heartbeat 时，可能会出现不易理解的奇怪问题。但是，有一些实用程序可以仔细地观察 Heartbeat 的内部进程。通过本章内容可获得有关如何解决您的问题的一些建议。

15.1 安装问题

如果您在安装包或使群集联机方面遇到困难，请根据以下列表进行操作。

是否安装了 HA 包?

配置和管理群集所需的包位于 High Availability Extension 提供的*高可用性*安装模式中。

根据第 3.1 节 “安装 High Availability Extension” [21]中所述，检查是否将 High Availability Extension 作为 SUSE Linux Enterprise Server 11 的外接式附件安装在每个群集节点上，以及每台计算机上是否安装了*高可用性*模式。

所有群集节点的初始配置是否相同?

根据第 3.2 节 “初始群集设置” [22]中所述，为了相互通讯，属于同一个群集的所有节点需要使用相同的 `bindnetaddr`、`mcastaddr` 和 `mcastport`。

检查在 `/etc/ais/openais.conf` 中配置的所有群集节点的通讯通道和选项是否相同。

如果您使用的是加密通讯，请检查所有群集节点上的 `/etc/ais/authkey` 文件是否可用。

防火墙是否允许通过 `mcastport` 进行通讯?

如果用于群集节点之间通讯的 mcastport 由防火墙阻止，这些节点将无法相互可见。根据第 3.1 节 “安装 High Availability Extension” [21]中所述，在使用 YaST 配置初始设置时，通常会自动调整防火墙设置。

要确保 mcastport 不被防火墙阻止，请检查每个节点上的 `/etc/sysconfig/SuSEfirewall2` 中的设置。或者，在每个群集节点上启动 YaST 防火墙模块。在单击*允许的服务* > *高级*后，将 mcastport 添加到允许的 *UDP 端口*列表中并确认更改。

是否在每个群集节点上启动了 OpenAIS?

使用 `/etc/init.d/openais status` 检查每个群集节点上的 OpenAIS 状态。如果 OpenAIS 未在运行，执行 `/etc/init.d/openais start` 命令来启动它。

15.2 “调试” HA 群集

如果您的群集中的某部分不工作，可首先尝试以下操作。它显示了资源操作历史记录（选项 `-o`）和不活动的资源 (`-r`)：

```
crm_mon -o -r
```

此显示每 10 秒钟刷新一次（您可以使用 Ctrl + C 取消它）。示例显示如下：

例 15.1　已停止的资源

```
Refresh in 10s...

============
Last updated: Mon Jan 19 08:56:14 2009
Current DC: d42 (d42)
3 Nodes configured.
3 Resources configured.
============

Online: [ d230 d42 ]
OFFLINE: [ clusternode-1 ]

Full list of resources:

Clone Set: o2cb-clone
         Stopped: [  o2cb:0 o2cb:1o2cb:2 ]
Clone Set: dlm-clone
         Stopped [ dlm:0 dlm:1 dlm:2 ]
mySecondIP      (ocf::heartbeat:IPaddr):        Stopped

Operations:
* Node d230:
   aa: migration-threshold=1000000
    + (5) probe: rc=0 (ok)
    + (37) stop: rc=0 (ok)
    + (38) start: rc=0 (ok)
    + (39) monitor: interval=15000ms rc=0 (ok)
* Node d42:
   aa: migration-threshold=1000000
    + (3) probe: rc=0 (ok)
    + (12) stop: rc=0 (ok)
```

首先使您的节点联机（参见第 15.3 节 [127]）。然后，检查您的资源和操作。

http://clusterlabs.org/wiki/Documentation 中的 *Configuration Explained* PDF 的*群集如何解释 OCF 返回代码？*一节中介绍了三种不同的恢复类型。

15.3 常见问题解答

我的群集状态是什么？

要检查您的群集的当前状态，请使用程序 `crm_mon`。这样可以显示当前 DC 以及当前节点已知的所有节点和资源。

我的群集的一些节点无法互相可见。

这可能有几个原因：

- 首先查看配置文件 /etc/ais/openais.conf，检查群集中的所有节点的多路广播地址是否相同（使用 mcastaddr 键搜索接口部分）。
- 检查您的防火墙设置。
- 检查您的交换机是否支持多路广播地址
- 另一个原因可能是您的节点之间的连接已中断。这通常是错误配置防火墙的结果。这也可能是*节点分裂*情况（其中群集已分区）的原因。

我想列出当前已知的资源。

使用命令 crm_resource -L 可以了解您的当前资源。

我配置了一个资源，但是它总是失败。

尝试手动运行资源代理。对于 LSB，只需运行*脚本名称* start 和 *脚本名称* stop。要检查 OCF 脚本，请首先设置所需的环境变量。例如，当测试 IPaddr OCF 脚本时，您必须通过设置一个变量名称前缀为 OCF_RESKEY_ 的环境变量来设置变量 ip 的值。对于此示例，可运行以下命令：

```
export OCF_RESKEY_ip=<your_ip_address>
/usr/lib/ocf/resource.d/heartbeat/IPaddr validate-all
/usr/lib/ocf/resource.d/heartbeat/IPaddr start
/usr/lib/ocf/resource.d/heartbeat/IPaddr stop
```

如果此操作不成功，很可能是您缺少某个必需变量或者只是输错了参数。

我刚刚收到一条失败消息。有可能收到更多信息吗？

您可以始终将 -V 参数添加到命令。如果您多次执行该操作，该调试输出会变得非常详细。

如何清理我的资源？

如果您知道您的资源 ID（可使用 crm_resource -L 获取），可使用 crm_resource -C -r *资源 ID* -H *主机*命令删除特定的资源。

我无法装入 ocfs2 设备。

检查 /var/log/message 是否存在以下行：

```
Jan 12 09:58:55 clusternode2 lrmd: [3487]: info: RA output:
(o2cb:1:start:stderr) 2009/01/12_09:58:55
  ERROR: Could not load ocfs2_stackglue
```

```
Jan 12 16:04:22 clusternode2 modprobe: FATAL: Module ocfs2_stackglue not
 found.
```

在这种情况下，将缺少内核模块`ocfs2_stackglue.ko`。请根据安装的内核安装包`ocfs2-kmp-default`、`ocfs2-kmp-pae`或`ocfs2-kmp-xen`。

15.4 更多信息

有关 Linux 和 Heartbeat 上的高可用性的更多信息（包括配置群集资源以及管理和自定义 Heartbeat 群集），请参见 http://clusterlabs.org/wiki/Documentation。

16 群集管理工具

High Availability Extension 附带了一组全面的工具，来帮助您通过命令行管理群集。本章主要介绍管理 CIB 中的群集配置和群集资源所需的工具。有关用于管理资源代理的其他命令行工具或用于对设置进行调试和查错的工具的内容包含在第 15 章 *查错* [125]中。

以下列表提供了一些与群集管理相关的任务，并简要介绍了完成这些任务所使用的工具：

监视群集的状态

`crm_mon` 命令可用于监视您的群集状态和配置。其输出包括节点数、uname、uuid、状态、群集中配置的资源和每个资源的当前状态。`crm_mon` 的输出可显示在控制台上或打印到 HTML 文件。当具有不包含状态部分的群集配置文件时，`crm_mon` 会按文件中所指定的方式创建节点和资源概览。有关对此工具的使用和命令语法的详细介绍，请参见crm_mon(8) [153]。

管理 CIB

`cibadmin` 命令是用于操作 Heartbeat CIB 的低级管理命令。它可用于转储、更新和修改全部或部分 CIB，删除整个 CIB 或执行其他 CIB 管理操作。有关对此工具的使用和命令语法的详细介绍，请参见cibadmin(8) [133]。

管理配置更改

`crm_diff` 命令可帮助您创建和应用 XML 增补程序。它对于观察群集配置的两个版本之间的更改或保存这些更改供日后使用cibadmin(8) [133]来应用它们非常有用。有关对此工具的使用和命令语法的详细介绍，请参见crm_diff(8) [145]。

操作 CIB 属性

您可以使用 crm_attribute 命令来查询和操作 CIB 中使用的节点属性和群集配置选项。有关对此工具的使用和命令语法的详细介绍，请参见 crm_attribute(8) [142]。

验证群集配置

crm_verify 命令可检查配置数据库 (CIB) 的一致性和其他问题。它可检查包含配置的文件或连接到运行中的群集。它可报告两类问题。虽然警告解决方法已经传达到管理员，但是必须修复错误 Heartbeat 才能正常工作。

crm_verify 可帮助创建新的或已修改的配置。您可以本地复制运行的群集中的 CIB，编辑它，使用 crm_verify 验证它，然后使用 cibadmin 使新配置生效。有关对此工具的使用和命令语法的详细介绍，请参见 crm_verify(8) [179]。

管理资源配置

crm_resource 命令可在群集上执行各种资源相关的操作。它可以修改已配置资源的定义，启动和停止资源，删除资源或在节点间迁移资源。有关对此工具的使用和命令语法的详细介绍，请参见crm_resource(8) [157]。

管理资源故障计数

crm_failcount 命令可查询指定节点上每个资源的故障计数。此工具还可用于重设置故障计数，同时允许资源在它多次失败的节点上再次运行。有关对此工具的使用和命令语法的详细介绍，请参阅crm_failcount(8) [148]。

管理节点的备用状态

crm_standby 命令可操作节点的备用属性。备用模式中的所有节点都不再具备主管资源的资格，并且必须移动那里的所有资源。备用模式对于执行维护任务（如内核更新）很有用。当备用属性应再次成为群集的完全活动的成员时，将它从节点中删除。有关对此工具的使用和命令语法的详细介绍，请参阅crm_standby(8) [176]。

cibadmin (8)

cibadmin — 提供对群集配置的直接访问

大纲

允许查询、修改、替换和删除配置或配置的某些部分。

```
cibadmin (--query|-Q) -[Vrwlsmfbp] [-i xml-object-id|-o
  xml-object-type] [-t t-flag-whatever] [-h hostname]

cibadmin (--create|-C) -[Vrwlsmfbp] [-X xml-string]
  [-x xml- filename] [-t t-flag-whatever] [-h hostname]

cibadmin (--replace-R)  -[Vrwlsmfbp]  [-i xml-object-id|
  -o xml-object-type]  [-X  xml-string] [-x xml-filename] [-t
  t-flag- whatever] [-h hostname]

cibadmin (--update|-U) -[Vrwlsmfbp]  [-i xml-object-id|
  -o xml-object-type]  [-X  xml-string] [-x xml-filename] [-t
  t-flag- whatever] [-h hostname]

cibadmin (--modify|-M) -[Vrwlsmfbp]  [-i xml-object-id|
           -o xml-object-type]  [-X  xml-string] [-x xml-filename] [-t
           t-flag- whatever] [-h hostname]

cibadmin (--delete|-D) -[Vrwlsmfbp]  [-i xml-object-id|
           -o xml-object-type] [-t t-flag-whatever] [-h hostname]

cibadmin (--delete_alt|-d) -[Vrwlsmfbp] -o
           xml-object-type [-X xml-string|-x xml-filename]
           [-t t-flag-whatever]  [-h hostname]

cibadmin --erase (-E)

cibadmin --bump (-B)

cibadmin --ismaster (-m)

cibadmin --master (-w)

cibadmin --slave (-r)

cibadmin --sync (-S)

cibadmin --help (-?)
```

描述

cibadmin 是用于操作 Heartbeat CIB 的低级管理命令。它可以用来转储、更新或修改所有或部分 CIB，删除整个 CIB 或执行各种 CIB 管理操作。

cibadmin 在 CIB 的 XML 树上运行，基本上不清楚所执行的更新或查询的含义。这意味着对于理解 XML 树中的元素含义的人来说很常见的快捷方式无法用于 cibadmin。它在输入和输出时要求消除不明确性，并且只能处理有效的 XML 子树（标记和元素）。

注意

使用 cibadmin 应始终优先于手动编辑 cib.xml 文件，特别是在群集活动的情况下。群集会竭尽全力检测和阻止此做法，从而确保您的数据不会丢失或损坏。

选项

--obj_type *对象类型*，-o *对象类型*
: 指定要操作的对象类型。有效值为 nodes、resources、constraints、crm_status 和 status。

--verbose，-V
: 打开调试模式。附加的 -V 选项可增加输出的详细程度。

--help，-?
: 从 cibadmin 获取安全消息。

--xpath *指定路径*，-A *指定路径*
: 提供用于替换 obj_type 的有效 XPath。

命令

--bump, -B
递增 CIB 中的 epoch 版本计数器。通常在选举新的领导节点时群集会自动增加该值。当您想废弃旧的配置（如不活动的群集节点上储存的配置）时，手动增加该值可能会有用。

--create, -C
从自变量的 XML 内容创建新的 CIB。

--delete, -D
删除与提供的准则匹配的第一个对象，例如，<op id="rsc1_op1" name="monitor"/>。标记名称和所有属性必须匹配才能删除元素。

--erase, -E
删除整个 CIB 的内容。

--ismaster, -m
打印指示 CIB 软件的本地实例是否为主实例的消息。如果它是主实例，将退出并返回代码 0；如果不是，则返回代码 35。

--modify, -M
在 CIB 的 XML 树中的某处找到对象并更新它。

--query, -Q
查询 CIB 的一部分。

--replace, -R
递归替换 CIB 中的 XML 对象。

--sync, -S
强制所有节点与指定主机上的 CIB（如果使用 -h）或 DC（如果未使用 -h 选项）重新同步。

XML 数据

`--xml-text` *字符串*，`-X` *字符串*
指定 crmadmin 应操作的 XML 标记或片段。它必须是完整的标记或 XML 片段。

`--xml-file` *文件名*，`-x` *文件名*
从文件指定 cibadmin 应操作的 XML。它必须是完整的标记或 XML 片段。

`--xml_pipe`，`-p`
指定 `cibadmin` 应操作的 XML 来自标准输入。它必须是完整的标记或 XML 片段。

高级选项

`--host` *主机名*，`-h` *主机名*
将命令发送至指定的主机。仅适用于 `query` 和 `sync` 命令。

`--local`，`-l`
使命令在本地生效（较少使用的高级选项）。

`--no-bcast`，`-b`
即使命令改变了 CIB，也不会对它进行广播。

重要

使用此选项时请小心，以避免导致群集分散。

`--sync-call`，`-s`
返回前等待调用完成。

示例

要获取递送到 stdout 的整个活动 CIB（包含状态部分等）的副本，请发出以下命令：

```
cibadmin -Q
```

要向资源部分添加 IPaddr2 资源，首先要使用以下内容创建一个文件 `foo`：

```
<primitive id="R_10.10.10.101" class="ocf" type="IPaddr2"
 provider="heartbeat">
 <instance_attributes id="RA_R_10.10.10.101">
  <attributes>
   <nvpair id="R_ip_P_ip" name="ip" value="10.10.10.101"/>
   <nvpair id="R_ip_P_nic" name="nic" value="eth0"/>
  </attributes>
 </instance_attributes>
</primitive>
```

然后发出以下命令：

```
cibadmin --obj_type resources -U -x foo
```

要更改先前添加的 IPaddr2 资源的 IP 地址，请发出以下命令：

```
cibadmin -M -X '<nvpair id="R_ip_P_ip" name="ip" value="10.10.10.102"/>'
```

注意

此命令不会更改资源名称以匹配新的 IP 地址。要执行该操作，请先删除资源再重新添加带有新 ID 标记的资源。

要停止（禁用）先前添加的 IP 地址资源而不将它删除，请创建包含以下内容的名为 `bar` 的文件：

```
<primitive id="R_10.10.10.101">
 <instance_attributes id="RA_R_10.10.10.101">
  <attributes>
   <nvpair id="stop_R_10.10.10.101" name="target-role" value="Stopped"/>
  </attributes>
 </instance_attributes>
</primitive>
```

然后发出以下命令：

```
cibadmin --obj_type resources -U -x bar
```

要重启动上一步中停止的 IP 地址资源，请发出以下命令：

```
cibadmin -D -X '<nvpair id="stop_R_10.10.10.101">'
```

要将此 IP 地址资源从 CIB 中彻底删除，请发出以下命令：

```
cibadmin -D -X '<primitive id="R_10.10.10.101"/>'
```

要将 CIB 替换为新的手动编辑版本，请使用以下命令：

```
cibadmin -R -x $HOME/cib.xml
```

文件数

`/var/lib/heartbeat/crm/cib.xml`— 磁盘上的 CIB（去除状态部分）。

另请参见

crm_resource(8) [157]，crmadmin(8) [139]，lrmadmin(8)，heartbeat(8)

作者

`cibadmin` 由 Andrew Beekhof 编写。

此手册页最初由 Alan Robertson 编写。

警告

避免在本地磁盘上的 CIB 的自动维护副本中工作。当群集中的任何内容发生更改时，会更新 CIB。因此使用 CIB 的过期备份副本来传播配置更改可能会导致群集不一致。

crmadmin (8)

crmadmin — 控制群集资源管理器

大纲

```
crmadmin [-V|-q] [-i|-d|-K|-S|-E] node
crmadmin [-V|-q] -N -B
crmadmin [-V|-q] -D
crmadmin -v
crmadmin -?
```

描述

crmadmin 最初用于控制 CRM 守护程序的大多数操作。但是，其他一些工具（如 crm_attribute 和 crm_resource）已使其大部分功能过时。其剩余的功能大多与测试和 crmd 进程状态相关。

警告

一些 crmadmin 选项调整到面向测试，如果使用不当，将导致故障。特别是不要使用 --kill 或 --election 选项，除非您完全了解正在执行的操作。

选项

--help, -?
打印帮助文本。

--version, -v
打印 HA、CRM 和 CIB 功能集的版本细节。

--verbose, -V
打开命令调试信息。

注意

可通过提供更多实例来增加详细级别。

`--quiet, -q`
不提供任何调试信息并将输出减少到最小化。

`--bash-export, -B`
创建形式为 `export uname=uuid` 的 Bash 导出项。此选项仅适用于 `crmadmin -N` *节点*命令。

注意

`-B` 功能很少有用，在以后的版本中删除。

命令

`--debug_inc` *节点*`, -i` *节点*
递增指定节点上的 CRM 守护程序的调试级别。这也可以通过向 crmd 进程发送 USR1 信号来实现。

`--debug_dec` *节点*`, -d` *节点*
递减指定节点上的 CRM 守护程序的调试级别。这也可以通过向 crmd 进程发送 USR2 信号来实现。

`--kill` *节点*`, -K` *节点*
关闭指定节点上的 CRM 守护程序。

警告

请谨慎使用此命令。此操作通常由 Heartbeat 来执行，并可能产生不希望的副作用。

`--status` *节点*`, -S` *节点*
查询指定节点上的 CRM 守护程序的状态。

输出包含一个常规运行状况指标和crmd进程的内部FSM状态。这对于确定群集正在执行的操作非常有帮助。

`--election` *节点*, `-E` *节点*
从特定的节点启动选举。

警告

请谨慎使用此命令。此操作通常由内部启动，并可能产生不希望的副作用。

`--dc_lookup`, `-D`
查询当前 DC 的 uname。

DC 的位置在内部仅对 crmd 有意义，而对管理员很少有用，除非是决定在哪个节点上检查日志时。

`--nodes`, `-N`
查询所有成员节点的 uname。此查询的结果可能包含脱机模式下的节点。

注意

`-i`、`-d`、`-K` 和 `-E` 选项很少使用，在以后的版本中可能删除。

另请参见

crm_attribute(8) [142]，crm_resource(8) [157]

作者

`crmadmin` 由 Andrew Beekhof 编写。

crm_attribute (8)

crm_attribute — 允许查询、修改和删除节点属性和群集选项。

大纲

```
crm_attribute [options]
```

描述

crm_attribute命令用于查询和操作CIB中使用的节点属性和群集配置选项。

选项

--help, -?
打印帮助消息。

--verbose, -V
打开调试信息。

注意

可通过提供更多实例来增加详细级别。

--quiet, -Q
当使用 -G 执行属性查询时，只是将值打印到 stdout。将此选项与 -G 一起使用。

--get-value, -G
检索而不是设置自选设置。

--delete-attr, -D
删除属性，而不是设置属性。

--attr-id *字符串*, -i *字符串*
仅适用于高级用户。用于标识 ID 属性。

`--attr-value` *字符串*，`-v` *字符串*
要设置的值。当与 `-G` 一起使用时，忽略此选项。

`--node` *节点名*，`-N` *节点名*
要更改的节点的 uname

`--set-name` *字符串*，`-s` *字符串*
指定要读取或写入属性的属性集。

`--attr-name` *字符串*，`-n` *字符串*
指定要设置或查询的属性。

`--type` *字符串*，`-t` *类型*
确定应设置属性的 CIB 部分或查询的属性所属的 CIB 部分。可能的值有 `nodes`、`status` 或 `crm_config`。

示例

在 CIB 的主机 *myhost* 的 `nodes` 部分中查询 `location` 属性的值：

```
crm_attribute -G -t nodes -U myhost -n location
```

在 CIB 的 `crm_config` 部分中查询 `cluster-delay` 属性的值：

```
crm_attribute -G -t crm_config -n cluster-delay
```

在 CIB 的 `crm_config` 部分中查询 `cluster-delay` 属性的值。只打印值：

```
crm_attribute -G -Q -t crm_config -n cluster-delay
```

从 CIB 的 `nodes` 部分删除主机 *myhost* 的 `location` 属性：

```
crm_attribute -D -t nodes -U myhost -n location
```

将值为 `office` 的名为 `location` 的新属性添加到 CIB 中 `nodes` 部分的 *set* 子部分（设置将应用到主机 *myhost*）：

```
crm_attribute -t nodes -U myhost -s set -n location -v office
```

更改 *myhost* 主机的 `nodes` 部分中的 `location` 属性：

```
crm_attribute -t nodes -U myhost -n location -v backoffice
```

文件数

/var/lib/heartbeat/crm/cib.xml—磁盘上的 CIB（去除状态部分）。强烈建议您不要直接编辑此文件。

另请参见

cibadmin(8) [133]

作者

crm_attribute 由 Andrew Beekhof 编写。

crm_diff (8)

crm_diff — 识别对群集配置所做的更改，并将增补程序应用到配置文件

大纲

```
crm_diff [-?|-V] [-o filename] [-O string] [-p filename] [-n filename] [-N
string]
```

描述

crm_diff 命令协助创建和应用 XML 增补程序。这对可视化两个版本的群集配置之间的更改或保存更改可能非常有用，以便以后使用 cibadmin 应用更改。

选项

--help，-?
打印帮助消息。

--original *文件名*，-o *文件名*
指定要进行区分或应用增补程序的原始文件。

--new *文件名*，-n *文件名*
指定新文件的名称。

--original-string *字符串*，-O *字符串*
指定进行区分或应用增补程序的原始字符串。

--new-string *字符串*，-N *字符串*
指定新字符串。

--patch *文件名*，-p *文件名*
将增补程序应用于原始 XML。总是与 -o 一起使用。

`--cib, -c`
比较或增补 CIB 输入。总是使用 `-o` 指定基础版本，使用 `-p` 或 `-n` 分别提供增补程序文件或第二个版本。

`--stdin, -s`
从 stdin 读取输入。

示例

使用 crm_diff 确定各种 CIB 配置文件的区别并创建增补程序。通过增补程序的方式，轻松重用各个配置部分，而不必对每个部分使用 cibadmin 命令。

1. 通过对要比较的两个群集设置运行 cibadmin 命令，获取两个不同的配置文件：

   ```
   cibadmin -Q > cib1.xml
   cibadmin -Q > cib2.xml
   ```

2. 确定是区分所有文件还是只比较配置子集。

3. 要将文件间的区别打印到 stdout，请使用以下命令：

   ```
   crm_diff -o cib1.xml -n cib2.xml
   ```

4. 要将文件间的区别打印到某个文件并创建增补程序，请使用以下命令：

   ```
   crm_diff -o cib1.xml -n cib2.xml > patch.xml
   ```

5. 将增补程序应用于原始文件：

   ```
   crm_diff -o cib1.xml -p patch.xml
   ```

文件数

`/var/lib/heartbeat/crm/cib.xml`— 磁盘上的 CIB（去除状态部分）。强烈建议您不要直接编辑此文件。

另请参见

cibadmin(8) [133]

作者

`crm_diff` 由 Andrew Beekhof 编写。

crm_failcount (8)

crm_failcount — 管理记录每个资源的故障计数的计数器。

大纲

```
crm_failcount [-?|-V] -D -u|-U node -r resource
crm_failcount [-?|-V] -G -u|-U node -r resource
crm_failcount [-?|-V] -v string -u|-U node -r resource
```

描述

Heartbeat 实施了一种精密的计算方法，当资源在当前节点上趋向失败时强制将该资源故障转移到其他节点。资源携带了一个 `resource-stickiness` 属性以确定它希望在某个节点上运行的自愿程度。它还具有 `migration-threshold` 属性，可用于确定资源应故障转移到其他节点的阈值。

可将 `failcount` 属性添加到资源，它的值将根据资源监视到的故障而递增。将 `failcount` 的值与 `migration-threshold` 的值相乘，可确定该资源的*故障转移分数*。如果此数字超过该资源的自选设置，则该资源将被移到其他节点并且不会在原始节点上再次运行，直到重设置故障计数。

`crm_failcount` 命令可查询指定节点上每个资源的故障计数。此工具还可用于重设置故障计数，并允许资源在它多次失败的节点上再次运行。

选项

`--help, -?`
打印帮助消息。

`--verbose, -V`
打开调试信息。

注意

可通过提供更多实例来增加详细级别。

`--quiet, -Q`
当使用 `-G` 执行属性查询时，只是将值打印到 stdout。将此选项与 `-G` 一起使用。

`--get-value, -G`
检索而不是设置自选设置。

`--delete-attr, -D`
指定要删除的属性。

`--attr-id` *字符串*`, -i` *字符串*
仅适用于高级用户。用于标识 ID 属性。

`--attr-value` *字符串*`, -v` *字符串*
指定要使用的值。当与 `-G` 一起使用时，将忽略此选项。

`--node` *节点 uname*`, -U` *节点 uname*
指定要更改的节点 uname。

`--resource-id` *资源名称*`, -r` *资源名称*
指定要运行的资源的名称。

示例

重设置节点 `node1` 上资源 `myrsc` 的故障计数：

```
crm_failcount -D -U node1 -r my_rsc
```

查询节点 `node1` 上资源 `myrsc` 的当前故障计数：

```
crm_failcount -G -U node1 -r my_rsc
```

文件数

/var/lib/heartbeat/crm/cib.xml— 磁盘上的 CIB（去除状态部分）。强烈建议您不要直接编辑此文件。

另请参见

crm_attribute(8) [142]、cibadmin(8) [133] 和 Linux High Availability FAQ Web 站点 [http://www.linux-ha.org/v2/faq/forced_failover]

作者

crm_failcount 由 Andrew Beekhof 编写。

crm_master (8)

crm_master — 管理主/从属资源的自选设置，以在给定节点上提升

大纲

```
crm_master [-V|-Q] -D [-l lifetime]
crm_master [-V|-Q] -G [-l lifetime]
crm_master [-V|-Q] -v string [-l string]
```

描述

从资源代理脚本内调用 crm_master 以确定应提升哪些资源实例为主模式。它不应从命令行中使用，并且只是资源代理的帮助实用程序。RA 使用 crm_master 将特定实例提升为主模式，或从此实例中删除自选设置。通过指派有效期，可确定此设置在重引导节点后仍然存在（将有效期设置为 forever）或不再存在（将有效期设置为 reboot）。

资源代理需要确定 crm_master 应操作的资源。必须在资源代理脚本内处理这些查询。对 crm_master 的实际调用遵循类似于 crm_attribute 命令的语法。

选项

--help, -?
打印帮助消息。

--verbose, -V
打开调试信息。

注意

通过提供更多实例增加详细程度。

--quiet, -Q
使用 -G 进行属性查询时，只将值打印到 stdout。将此选项与 -G 一起使用。

--get-value, -G
获取而不是设置要提升的自选设置。

--delete-attr, -D
删除而不是设置属性。

--attr-id *string*, -i *string*
仅适用于高级用户。标识 id 属性。

--attr-value *string*, -v *string*
设置的值。与 -G 一起使用时将忽略此项。

--lifetime *string*, -l *string*
指定自选设置持续的时间。可能的值有 reboot 或 forever。

环境变量

OCF_RESOURCE_INSTANCE— 资源实例的名称

文件数

/var/lib/heartbeat/crm/cib.xml— 磁盘上的 CIB（去除状态部分）。

另请参见

cibadmin(8) [133]，crm_attribute(8) [142]

作者

crm_master 由 Andrew Beekhof 编写。

crm_mon (8)

crm_mon — 监视群集状态

大纲

```
crm_mon [-V] -d -pfilename -h filename
crm_mon [-V] [-1|-n|-r] -h filename
crm_mon [-V] [-n|-r] -X filename
crm_mon [-V] [-n|-r] -c|-1
crm_mon [-V] -i interval
crm_mon -?
```

描述

crm_mon 命令允许您监视群集的状态和配置。其输出包括节点数、uname、uuid、状态、群集中配置的资源及其各自的当前状态。crm_mon 的输出可以显示在控制台上或打印到 HTML 文件。当具有不包含状态部分的群集配置文件时，crm_mon 会按文件中所指定的方式创建节点和资源概览。

选项

--help，-?
提供帮助。

--verbose，-V
增加调试输出。

--interval *秒*，-i *秒*
确定更新频率。如果 -i 未指定，则采用 15 秒的默认值。

--group-by-node，-n
按节点对资源分组。

`--inactive, -r`
显示不活动的资源。

`--simple-status, -s`
以一行输出显示一次群集状态（适合于 nagios）。

`--one-shot, -1`
在控制台上显示一次群集状态，然后退出（不使用 ncurses）。

`--as-html` *文件名*`, -h` *文件名*
将群集的状态写入指定的文件。

`--web-cgi, -w`
带有输出的 Web 模式，适合于 CGI。

`--daemonize, -d`
在后台作为守护程序运行。

`--pid-file` *文件名*`, -p` *文件名*
指定守护程序的 pid 文件。

示例

显示群集的状态并每隔 15 秒获取一次更新后的列表：

```
crm_mon
```

显示群集的状态并在 `-i` 中指定的间隔后获取一次更新后的列表。如果 `-i` 未指定，则采用 15 秒的默认刷新间隔：

```
crm_mon -i interval[s]
```

在控制台上显示群集状态：

```
crm_mon -c
```

在控制台上显示一次群集状态，然后退出：

```
crm_mon -1
```

显示群集的状态并按节点对资源分组：

```
crm_mon -n
```

显示群集的状态，按节点对资源分组，并在列表中包括不活动的资源：

```
crm_mon -n -r
```

将群集的状态写入 HTML 文件：

```
crm_mon -h filename
```

在后台作为守护程序运行 crm_mon，指定守护程序的 pid 文件以更轻松地控制守护程序进程，并创建 HTML 输出。此选项允许您持续创建 HTML 输出，该输出可以由其他监视应用程序轻松处理：

```
crm_mon -d -p filename -h filename
```

在现有群集配置文件（*文件名*）中显示群集配置，按节点对资源分组，并包括不活动的资源：此命令用于群集在线前的群集配置测试。

```
crm_mon -r -n -X filename
```

文件数

/var/lib/heartbeat/crm/cib.xml— 磁盘上的 CIB（去除状态部分）。强烈建议您不要直接编辑此文件。

作者

crm_mon 由 Andrew Beekhof 编写。

crm_node (8)

crm_node — 列出群集的成员

大纲

```
crm_node [-V] [-p|-e|-q]
```

描述

列出群集的成员。

选项

`-V`
详细

`--partition, -p`
打印此分区的成员

`--epoch, -e`
打印此节点加入分区的时间点。

`--quorum, -q`
如果分区包含仲裁，请打印 1

crm_resource (8)

crm_resource — 执行与群集资源相关的任务

大纲

```
crm_resource  [-?|-V|-S] -L|-Q|-W|-D|-C|-P|-p [options]
```

描述

crm_resource命令对资源执行各种资源相关的操作。它可以修改已配置资源的定义、启动和停止资源，以及在节点间删除和迁移资源。

--help, -?
打印帮助消息。

--verbose, -V
打开调试信息。

注意

通过提供更多的实例增加详细程度。

--quiet, -Q
仅在 stdout 上打印值（与 -W 结合使用）。

命令

--list, -L
列出所有资源。

--query-xml, -x
查询资源。

必需：-r

--locate, -W
定位资源。

必需：-r

--migrate, -M
从当前位置迁移资源。使用 -N 指定目标。

如果 -N 未指定，通过创建当前位置规则和分数 -INFINITY，强制移动资源。

注意

这可防止资源在此节点上运行，直到使用 -U 删除约束为止。

必需：-r，可选：-N，-f。

--un-migrate, -U
删除所有通过 -M 创建的约束。

必需：-r

--delete, -D
从 CIB 删除资源。

必需：-r，-t

--cleanup, -C
从 LRM 删除资源。

必需：-r。可选：-H

--reprobe, -P
重新检查在 CRM 之外启动的资源。

可选：-H

--refresh, -R
从 LRM 刷新 CIB。

可选：-H

--set-parameter *字符串*，-p *字符串*
为资源设置指定的参数。

必需：-r，-v。可选：-i、-s 和 --meta

--get-parameter *字符串*，-g *字符串*
为资源获取指定的参数。

必需：-r。可选：-i、-s 和 --meta

--delete-parameter *字符串*，-d *字符串*
为资源删除指定的参数。

必需：-r。可选：-i 和 --meta

--list-operations *字符串* ，-O *字符串*
列出活动资源操作。可选择按资源和/或节点过滤。可选：-N，-r

--list-all-operations *字符串* ，-o *字符串*
列出所有资源操作。可选择按资源和/或节点过滤。可选：-N，-r

选项

--resource *字符串*，-r *字符串*
指定资源 ID。

--resource-type *字符串*，-t *字符串*
指定资源类型（原始、克隆和组等）。

--property-value *字符串*，-v *字符串*
指定属性值。

--node *字符串*，-N *字符串*
指定主机名。

--meta
修改资源的配置选项，而不是修改传递给资源代理脚本的选项。与 -p、-g 和 -d 结合使用。

--lifetime *字符串*, -u *字符串*
迁移约束的有效期。

--force, -f
通过创建当前位置规则和分数 -INFINITY 强制移动资源。

如果资源的黏性和约束总分超出 INFINITY（目前是 100,000），则应使用此命令。

注意

这可防止资源在此节点上运行，直到使用 -U 删除约束为止。

-s *字符串*
（仅限于高级用法）指定要更改的 instance_attributes 对象的 ID。

-i *字符串*
（仅限于高级用法）指定要更改或删除的 nvpair 对象的 ID。

示例

列出所有资源：

```
crm_resource -L
```

检查正在运行资源的位置（以及是否在运行）：

```
crm_resource -W -r my_first_ip
```

如果 my_first_ip 资源正在运行，此命令的输出中会显示正在运行资源的节点。如果资源没有在运行，输出中会显示此情况。

启动或停止资源：

```
crm_resource -r my_first_ip -p target_role -v started
crm_resource -r my_first_ip -p target_role -v stopped
```

查询资源的定义：

```
crm_resource -Q -r my_first_ip
```

将资源迁离当前位置：

```
crm_resource -M -r my_first_ip
```

将资源迁移到指定的位置：

```
crm_resource -M -r my_first_ip -H c001n02
```

允许资源返回其常规位置：

```
crm_resource -U -r my_first_ip
```

注意

`resource_stickiness` 和 `default_resource_stickiness` 的值可能会意味着资源没有移回。在这种情况下，应先使用 `-M` 将资源移回，再运行此命令。

从 CRM 删除资源：

```
crm_resource -D -r my_first_ip -t primitive
```

从 CRM 删除资源组：

```
crm_resource -D -r my_first_group -t group
```

为 CRM 中的资源禁用资源管理：

```
crm_resource -p is-managed -r my_first_ip -t primitive -v off
```

为 CRM 中的资源启用资源管理：

```
crm_resource -p is-managed -r my_first_ip -t primitive -v on
```

在手动清理后，重设置有故障的资源：

```
crm_resource -C -H c001n02 -r my_first_ip
```

重新检查所有节点，以找出从 CRM 之外启动的资源：

```
crm_resource -P
```

重新检查一个节点，以找出从 CRM 之外启动的资源：

```
crm_resource -P -H c001n02
```

文件数

/var/lib/heartbeat/crm/cib.xml— 磁盘上的 CIB（去除状态部分）。强烈建议您不要直接编辑此文件。

另请参见

cibadmin(8) [133]，crmadmin(8) [139]，lrmadmin(8)，heartbeat(8)

作者

crm_resource 由 Andrew Beekhof 编写。

crm_shadow (8)

crm_shadow — 更新在线群集前在 Sandbox 中执行配置更改

大纲

```
crm_shadow [-V] [-p|-e|-q]
```

描述

设置一种配置工具（cibadmin 和 crm_resource 等）脱机工作而非针对在线群集工作的环境，以便预览和测试更改是否会产生不利影响。

选项

--verbose, -V
   打开调试信息。实例更多会增加详细程度

--which, -w
   指示活动的阴影副本

--display, -p
   显示阴影副本的内容

--diff, -d
   显示阴影副本的更改

--create-empty, -e*名称*
   使用空的群集配置创建命名的阴影副本

--create, -c*名称*
   用指定的名称创建活动群集配置的阴影副本

--reset, -r*名称*
   根据活动的群集配置创建命名的阴影副本

`--commit, -c`*名称*
　将指定的阴影副本内容上载到群集

`--delete, -d`*名称*
　删除指定的阴影副本的内容

`--edit, -e`*名称*
　用喜爱的编辑器编辑指定的阴影副本内容

`--batch, -b`
　不衍生新的壳层

`--force, -f`
　不衍生新的壳层

`--switch, -s`
　切换到指定的阴影副本

内部命令

要使用阴影配置，需要先创建一个：

```
crm_shadow --create-empty YOUR_NAME
```

它为您提供内部壳层，如 `crm` 工具提供的壳层。使用 `help` 获取所有内部命令的概览，或使用 `help` *子命令*获取特定命令。

表 16.1 *内部命令概览*

命令	语法/描述
别名	`alias [-p] [name[=value] ... ]` `alias` 不带任何自变量或带有 -p 选项，在标准输出中按别名格式 `NAME=VALUE` 打印别名列表。否则，为每个指定了 VALUE 的 NAME 定义别名。VALUE 中的尾部空格会导致在展开别名时检查下一个要进行别名替换的单词。除非指定了 NAME，否则别名返回 true，因为尚未定义别名。

命令	语法/描述
bg	`bg [JOB_SPEC ...]` 将每个 JOB_SPEC 放置在后台，好像已使用 & 启动。如果 JOB_SPEC 不出现，则使用壳层的当前作业概念。

bind	`bind [-lpvsPVS] [-m keymap] [-f filename]`
	`[-q name] [-u name] [-r keyseq]`
	`[-x keyseq:shell-command]`
	`[keyseq:readline-function or readline-command]`
	将键序列与 Readline 函数或宏绑定，或设置 Readline 变量。非选项自变量语法等同于对 `~/.inputrc` 使用的语法，但必须以单个自变量 bind 传递：`"\C-x\C-r": re-read-init-file`。

break	`break [N]` 从 `for`、`while` 或 `until` 循环中退出。如果指定了 N，中断 N 级。
builtin	`builtin [shell-builtin [arg ...]]` 运行壳层 builtin。如果希望将壳层 builtin 重命名为某个函数，但在该函数本身中需要 builtin 函数的功能，则此命令非常有用。
caller	`caller [EXPR]` 返回当前子例程调用环境。不带 *EXPR*，返回 `$line $filename`。带有 *EXPR*，返回 `$line $subroutine $filename`；此附加的信息可用于提供堆栈跟踪。
case	`case WORD in [PATTERN [| PATTERN] [COMMANDS;;] ... esac` 根据*单词*与*模式*的匹配程度选择性地执行*命令*。“|”用于分隔多个模式。
cd	`cd [-L|-P] [dir]` 将当前目录更改为 DIR。

命令	语法/描述

命令	`command [-pVv]`
	`command [arg ...]`
	
	运行带有 *ARGS* 的*命令*忽略壳层函数。如果有一个壳层函数为“ls”，要调用“ls”命令，可以输入“command ls”。如果指定了 -p 选项，则对 PATH 使用默认值，以保证找到所有标准的实用程序。如果指定了 -V 或 -v 选项，则打印描述“命令”的字符串。-V 选项给出更加详细的描述。

compgen	`compgen [-abcdefgjksuv] [-o option] [-A action]`
	`[-G globpat] [-W wordlist] [-P prefix]`
	`[-S suffix] [-X filterpat] [-F function]`
	`[-C command] [WORD]`
	
	根据选项显示可能的补全。专门用于壳层函数中，生成可能补全。如果在可选的*单词*自变量中提供了值，则根据*单词*生成匹配项。

complete	`complete [-abcdefgjksuv] [-pr] [-o option]`
	`[-A action] [-G globpat] [-W wordlist] [-P prefix]`
	`[-S suffix] [-X filterpat] [-F function] [-C command]`
	`[name ...]`
	
	为每个*名称*指定补全自变量的方式。如果提供了 -p 选项，或如果未提供任何选项，则以允许重新用作输入的方式打印现有补全规范。-r 选项删除每个*名称*的补全规范，如果未提供任何*名称*，删除所有补全规范。

continue	`continue [N]` 结束 FOR、WHILE 或 UNTIL 循环后继续下一个迭代。如果指定了 *N*，在第 *N* 次结束循环后继续。
declare	`declare [-afFirtx] [-p] [name[=value] ...]` 声明变量和/或为变量提供属性。如果未指定*名称*，则改为显示变量值。-p 选项显示每个*名称*的属性和值。
dirs	`dirs [-clpv] [+N] [-N]`

命令	语法/描述
	显示目前记住的目录的列表。使用 pushd 命令可使目录显示在列表上；使用 popd 命令可从列表中取回目录。
disown	disown [-h] [-ar] [JOBSPEC ...] 默认情况下，从活动的作业表中删除每个 *JOBSPEC* 自变量。如果指定了 -h 选项，则作业不会从表中删除，但会加以标记，以便壳层在收到 SIGHUP 时不发送给作业。-a 选项，在未指定 *JOBSPEC* 时，表示从作业表中删除全部作业；-r 选项表示仅删除正在运行的作业。
echo	echo [-neE] [arg ...] 输出 ARG。如果指定了 -n，则不换行。如果指定了 -e 选项，则打开以下反斜杠转义字符解释功能： \a（警告响铃符） \b（退格键） \c（不换行） \E（转义符） \f（换页符） \n（换行符） \r（回车键） \t（水平制表符） \v（垂直制表符） \\（反斜杠） \0nnn（ASCII 码为 NNN（八进制）的字符，NNN 可以是 0 到 3 的八进制数字） 使用 -E 选项可以明确关闭上述字符的解释功能。
enable	enable [-pnds] [-a] [-f filename] [name...] 启用和禁用 builtin 壳层命令。它允许您使用与壳层 builtin 名称相同的磁盘命令，而不必指定完整路径名。如果使用 -n，则禁用 *名称*；否则启用 *名称*。例如，要在 $PATH 中使用 test 代替

命令	语法/描述
	壳层 builtin 版本，请输入 `enable -n test`。在支持动态装载的系统上，`-f` 选项可用于从共享的对象*文件名*中装载新的 builtin。`-d` 选项删除以前使用 `-f` 装载的 builtin。如果未指定任何非选项名称，或指定了 `-p` 选项，则打印 builtin 列表。`-a` 选项表示打印每个 builtin，并指示是否启用。`-s` 选项将输出限制到 POSIX.2“特殊”builtin 中。`-n` 选项显示所有已禁用 builtin 的列表。
`eval`	`eval [ARG ...]` 将 *ARG* 读作壳层输入并执行所生成的命令。
`exec`	`exec [-cl] [-a name] file [redirection ...]` Exec *文件*，将此壳层替换为指定的程序。如果未指定*文件*，则会使重定向在此壳层中生效。如果第一个自变量为 `-l`，则在传递给*文件*的 zeroth 自变量中放置一个破折号，如登录时那样。如果指定了 `-c` 选项，以空环境执行“文件”。`-a` 选项表示将已执行过程的 argv[0] 设置为*名称*值。如果文件无法执行且壳层不是交互的，则在没有设置 `execfail` 壳层选项的情况下将退出壳层。
`exit`	`exit [N]` 状态为 *N* 时退出壳层。如果省略 *N*，退出状态为上次执行命令的状态。

`export`	`export [-nf] [NAME[=value] ...]`
	`export -p`
	*名称*标记为自动导出至后续执行的命令环境中。如果指定了 `-f` 选项，*名称*指的是函数。如果指定了*名称*，或如果指定了 `-p`，则打印所有在此壳层导出的名称列表。`-n` 自变量表示从后续*名称*中删除导出属性。`--` 自变量禁用进一步的选项处理。

`false`	`false` 返回不成功的结果。

命令	语法/描述

fc	`fc [-e ename] [-nlr] [FIRST] [LAST]`
	`fc -s [pat=rep] [cmd]`
	
	fc用于列出或编辑和重新执行历史记录列表中的命令。*第一个*和*最后一个*可以是指定范围的数字，*第一个*也可以是一个字符串，表示以该字符串开始的最新命令。

fg	`fg [JOB_SPEC]` 将 *JOB_SPEC* 放置在前景中，使其成为当前作业。如果 *JOB_SPEC* 不出现，则使用壳层的当前作业概念。
for	`for NAME [in WORDS ... ;] do COMMANDS; done` for 循环为每个列表项成员连续执行命令。如果 `in WORDS ...;` 不出现，则采用 `in "$@"`。对于*单词*中的每个元素，*名称*设置为该元素，并执行*命令*。

function	`function NAME { COMMANDS ; }`
	`function NAME () { COMMANDS ; }`
	
	创建由运行*命令*的*名称*调用的简单命令。将命令行上的自变量和*名称*传递给函数，形式为 `$0 .. $n`.

getopts	`getopts OPTSTRING NAME [arg]` Getopts 由壳层过程用于解析位置参数。
hash	`hash [-lr] [-p PATHNAME] [-dt] [NAME...]` 对于每个*名称*，确定并记住命令的完整路径名。如果指定了 `-p` 选项，*路径名*用作*名称*的完整路径名，且不执行路径搜索。`-r` 选项使壳层忘记所有记住的位置。`-d` 选项使壳层忘记每个记住的*名称*位置。如果指定了 `-t` 选项，则打印与每个*名称*对应的完整路径名。如果指定 `-t` 的同时指定了多个*名称*自变量，则对完整路径名执行哈希运算之前打印*名称*。`-l` 选项使输出以能够重新用作输入的格式显示。如果未指定自变量，则显示已记住命令的信息。

命令	语法/描述
history	history [-c] [-d OFFSET] [n] history -ps arg [arg...] history -awrm [filename] 显示带有行编号的历史记录列表。已修改带 * 列出的行。*N* 自变量表示仅列出最后 *N* 行。-c 选项表示通过删除所有条目来清除历史记录列表。-d选项删除偏移量为*偏移量*值的历史记录条目。-w选项将当前历史记录写入历史记录文件；-r 表示读取文件并追加内容到历史记录列表。-a 表示将此会话的历史记录行追加到历史记录文件。自变量 -n 表示读取所有尚未从历史记录文件中读取的历史记录行，并将它们追加到历史记录列表。
jobs	jobs [-lnprs] [JOBSPEC ...] job -x COMMAND [ARGS] 列出活动作业。-l 选项列出常规信息和进程 id；-p 选项仅列出进程 id。如果指定了 -n，则仅打印自上次通知以来更改了状态的进程。*JOBSPEC* 将输出限制为该作业。-r 和 -s 选项分别将输出限制为正在运行的作业和已停止的作业。如果不带选项，则打印所有活动作业的状态。如果指定了 -x，*命令*在 *ARGS* 中的所有作业规范均替换为该作业进程组领导的进程 ID 后运行。
kill	kill [-s sigspec | -n signum | -sigspec] pid | JOBSPEC ... kill -l [sigspec] 向由 PID（或 *JOBSPEC*）命名的进程发送 SIGSPEC 信号。如果 SIGSPEC 不出现，则采用 SIGTERM。-l 自变量列出信号名称，如果自变量前面带有 -l，则视为应列出其名称的信号数。Kill 是壳层 builtin 的原因有二：它允许使用作业 ID 代替进程 ID；如果您达到了可以创建的进程数限值，则不必在启动一个进程时终止另一个进程。
let	let ARG [ARG ...] 每个 *ARG* 都是一个需要求值的算术表达式。求值在固定宽度的整数中完成，不检查是否溢出，即便除数 0 被捕获并标记为错误也是如此。以下运算符列表分为若干个相同优先级运算符级别。级别以降序顺序列出。

命令	语法/描述
local	`local NAME[=VALUE] ...` 创建名为*名称*的本地变量，并为其指定*值*。local 可以用在函数中，使变量*名称*的可见范围限制为该函数及其子函数。
logout	`logout` 注销登录壳层。
popd	`popd [+N | -N] [-n]` 删除目录堆栈中的条目。如果不带自变量，则从堆栈中删除顶层目录并 cd 到新的顶层目录。
printf	`printf [-v var] format [ARGUMENTS]` printf 在*格式*控制下格式化并打印*自变量*。*格式*是包含三种对象的字符型字符串：普通字符，只需复制到标准输出；字符转义序列，转换并复制到标准输出；格式规范，每种规范均会导致打印下一个连续的自变量。除了标准 printf(1) 格式外，%b 表示在相应的自变量中展开反斜杠转义序列，%q 表示以可重新用作壳层输入的方式引用自变量。如果指定了 -v 选项，输出放置在壳层变量 VAR 的值中，而不是发送至标准输出。
pushd	`pushd [dir | +N | -N] [-n]` 将目录添加到目录堆栈的顶层，或旋转堆栈，使新的堆栈顶层成为当前工作目录。如果不带自变量，交换两个顶层目录。
pwd	`pwd [-LP]` 打印当前工作目录。如果带有 -P 选项，pwd 打印物理目录，不带任何符号链接；-L 选项使 pwd 前面带有符号链接。
read	`read [-ers] [-u fd] [-t timeout] [-p prompt] [-a array] [-n nchars] [-d delim] [NAME ...]` 指定的*名称*标记为只读，并且这些*名称*的值可能无法通过随后的指派进行更改。如果指定了 -f 选项，与“名称”对应的函数会如

命令	语法/描述
	上标记。如果未指定自变量，或如果指定了 -p，则打印所有只读名称列表。-a 选项表示将每个“名称”视为数组变量。-- 自变量禁用进一步的选项处理。

readonly	`readonly [-af] [NAME[=VALUE] ...]`
	`readonly -p`
	指定的*名称*标记为只读，并且这些*名称*的值可能无法通过随后的指派进行更改。如果指定了 -f 选项，与*名称*对应的函数会如上标记。如果未指定自变量，或如果指定了 -p，则打印所有只读名称列表。-a 选项表示将每个*名称*视为数组变量。-- 自变量禁用进一步的选项处理。

return	`return [N]` 使函数退出，返回值由*N*指定。如果省略了*N*，返回状态为上次执行命令的状态。
select	`select NAME [in WORDS ... ;] do COMMANDS; done` 展开*单词*，生成单词列表。将一组展开的单词打印在标准错误上，每个单词前面带有编号。如果 in WORDS 不出现，则采用 in "$@"。显示 PS3 提示，并从标准输入中读取一行。如果该行包括与某个显示的单词对应的编号，则将*名称*设置为该单词。如果该行为空，则再次显示*单词*和提示。如果读到 EOF，则命令完成。读取任何其他的值都会使*名称*设置为空。读取的行保存在变量*答复*中。*命令*在每次选择后执行，直到执行中断命令为止。
set	`set [--abefhkmnptuvxBCHP] [-o OPTION] [ARG...]` 设置内部壳层选项。
Shift	`shift [n]` $N+1 ...的位置参数重命名为 $1 ...如果未指定 *N*，则假设为 1。

命令	语法/描述
shopt	shopt [-pqsu] [-o long-option] OPTNAME [OPTNAME...] 切换变量值以控制选项行为。-s 标志表示启用（设置）每个 *OPTNAME*；-u 标志取消设置每个 *OPTNAME*。-q 标志禁止输出，退出状态指示是设置还是取消设置每个 *OPTNAME*。-o 选项将 *OPTNAME* 限制为那些已定义为与 set -o 结合使用的名称。如果不带选项，或如果带有 -p 选项，则显示所有可设置的选项列表，并指示是否设置每个选项。
source	source FILENAME [ARGS] 从 *文件名* 读取并执行命令，然后返回。$PATH 中的路径名用于查找包含 *文件名* 的目录。如果指定了任何 *ARGS*，在执行 *文件名* 时，它们变为位置参数。
suspend	suspend [-f] 暂挂此壳层的执行，直到收到 SIGCONT 信号为止。如果指定了 -f，则表示不介意这是一个登录壳层（如果是），只将其暂挂。
test	test [expr] 退出状态为 0 (true) 或为 1 (false)，这取决于 *EXPR* 的求值结果。表达式可以是一元或二元的。一元表达式通常用于检查文件的状态。包括字符串运算符和数值比较运算符。
time	time [-p] PIPELINE 执行 *流水线*，并在流水线终止时打印执行 *流水线* 所用的实时用户 CPU 时间和系统 CPU 时间的摘要。返回状态是“流水线”的返回状态。-p 选项打印计时摘要，打印的格式稍有不同。它使用“时间格式”变量的值作为输出格式。
times	times 打印从壳层运行进程的累计用户时间和系统时间。
trap	trap [-lp] [ARG SIGNAL_SPEC ...]

命令	语法/描述
	在壳层接收到信号 *SIGNAL_SPEC* 时，读取和执行命令 *ARG*。如果 *ARG* 缺少（且提供了单个的 *SIGNAL_SPEC*）或为 -，则将每个指定的信号重设置为其原始值。如果 *ARG* 为空字符串，则壳层及其调用的命令会忽略每个 *SIGNAL_SPEC*。如果 *SIGNAL_SPEC* 为 EXIT (0)，则在从壳层退出时执行 ARG 命令。如果 *SIGNAL_SPEC* 为 DEBUG，则在执行每个简单的命令后执行 *ARG* 命令。如果指定了 -p 选项，则显示与每个 *SIGNAL_SPEC* 关联的 trap 命令。如果未指定任何自变量或仅指定了 -p，则 trap 打印与每个信号关联的命令列表。每个 *SIGNAL_SPEC* 可以是 signal.h 中的信号名称，也可以是信号编号。信号名称不区分大小写且 SIG 前缀是可选的。trap -l 打印信号名称及其相应编号的列表。注意，信号可以通过 kill -signal $$ 发送给壳层。
true	true 返回成功结果。
type	type [-afptP] NAME [NAME ...] 过时，请参见 declare。
typeset	typeset [-afFirtx] [-p] name[=value] 过时，请参见 declare。
ulimit	ulimit [-SHacdfilmnpqstuvx] [limit Ulimit 用于控制壳层启动的进程可以在允许这样控制的系统上使用哪些资源。
umask	umask [-p] [-S] [MODE] 将用户文件创建掩码设置为 *模式*。如果省略了 *模式*，或如果指定了 -S，则打印掩码的当前值。-S 选项使输出变为符号；否则，输出八进制数值。如果指定了 -p，且省略了 *模式*，输出采用可

命令	语法/描述
	以用作输入的格式。如果*模式*以数字开始，则它会解析为八进制数值，否则为符号模式字符串，如 chmod(1) 接受的字符串。
unalias	`unalias [-a] NAME [NAME ...]` 从定义的别名列表中删除*名称*。如果指定了 `-a` 选项，删除所有别名定义。
unset	`unset [-f] [-v] [NAME ...]` 对于每个*名称*，删除相应的变量或函数。如果指定了 `-v`，则仅对变量取消设置。如果指定了 `-f` 标志，则仅对函数取消设置。如果两个标志都未指定，则 unset 首先试图取消设置变量，如果失败，则试图取消设置函数。某些变量无法取消设置；另请参见 readonly。
until	`until COMMANDS; do COMMANDS; done` 只要 `until` *命令*中的最终命令的退出状态不为零，就展开并执行*命令*。
wait	`wait [N]` 等待指定的进程并报告其终止状态。如果未指定 *N*，则等待所有当前活动的子进程，返回代码是零。*N* 可以是进程 ID，也可以是作业规范；如果指定了作业规范，则等待作业流水线中的所有进程。
while	`while COMMANDS; do COMMANDS; done` 只要 `while` *命令*中的最终命令的退出状态为零，就展开并执行*命令*。

crm_standby (8)

crm_standby — 操作节点的备用属性以确定资源是否可在此节点上运行

大纲

```
crm_standby [-?|-V] -D -u|-U node -r resource

crm_standby  [-?|-V] -G -u|-U node -r resource

crm_standby [-?|-V] -v string -u|-U node -r resource [-l string]
```

描述

crm_standby 命令可操作节点的备用属性。备用模式中的所有节点都不再具备主管资源的资格，并且必须移动那里的所有资源。备用模式对于执行维护任务（如内核更新）很有用。当备用属性应再次成为群集的完全活动的成员时，将它从节点中删除。

通过为 standby 属性指派有效期，可确定该备用设置应免于重引导节点（有效期设置为 forever），还是应借助重引导重设置（有效期设置为 reboot）。或者，也可以删除 standby 属性，并手动使节点离开备用模式。

选项

--help, -?
   打印帮助消息。

--verbose, -V
   打开调试信息。

注意

可通过提供更多实例来增加详细级别。

`--quiet, -Q`
当使用 `-G` 执行属性查询时，只是将值打印到 stdout。将此选项与 `-G` 一起使用。

`--get-value, -G`
检索而不是设置自选设置。

`--delete-attr, -D`
指定要删除的属性。

`--attr-value` *字符串*`, -v` *字符串*
指定要使用的值。当与 `-G` 一起使用时，将忽略此选项。

`--attr-id` *字符串*`, -i` *字符串*
仅适用于高级用户。标识 ID 属性。

`--node` *节点 uname*`, -u` *节点 uname*
指定要更改的节点 uname。

`--lifetime` *字符串*`, -l` *字符串*
确定此自选设置的持续时间。可能的值包括 `reboot` 或 `forever`。

注意

如果 `forever` 值存在，则 CRM 将始终使用该值，而不使用任何 `reboot` 值。

示例

将本地节点转到备用模式：

```
crm_standby -v true
```

将节点 (`node1`) 转到备用模式：

```
crm_standby -v true -U node1
```

查询某个节点的备用状态：

```
crm_standby -G -U node1
```

从节点删除备用属性：

```
crm_standby -D -U node1
```

将节点无限期转到备用模式：

```
crm_standby -v true -l forever -U node1
```

将节点转到备用模式，直到下次重引导此节点：

```
crm_standby -v true -l reboot -U node1
```

文件数

`/var/lib/heartbeat/crm/cib.xml`— 磁盘上的 CIB（去除状态部分）。强烈建议您不要直接编辑此文件。

另请参见

cibadmin(8) [133]，crm_attribute(8) [142]

作者

`crm_standby` 由 Andrew Beekhof 编写。

crm_verify (8)

crm_verify — 检查 CIB 的一致性

大纲

```
crm_verify [-V] -x file
crm_verify [-V] -X string
crm_verify [-V] -L|-p
crm_verify [-?]
```

描述

crm_verify 用于检查配置数据库 (CIB) 的一致性和其他问题。它可用来检查包含配置的文件，并可连接到正在运行的群集。它报告两类问题：错误和警告。必须修复错误，Heartbeat 才能正常工作。但是，应让管理员来决定警告是否也应修复。

crm_verify 可帮助创建新的或已修改的配置。您可以在运行的群集中获取 CIB 的本地副本，编辑它，使用 crm_verify 验证它，然后使用 cibadmin 使新配置生效。

选项

--help, -h
   打印帮助消息。

--verbose, -V
   打开调试信息。

注意

可通过提供更多实例来增加详细级别。

`--live-check, -L`
连接到正在运行的群集并检查 CIB。

`--crm_xml` *字符串*, `-X` *字符串*
检查提供的字符串中的配置。仅传递完整的 CIB。

`--xml-file` *文件*, `-x` *文件*
检查命名的文件中的配置。

`--xml-pipe, -p`
使用通过 stdin 传送的配置。仅传递完整的 CIB。

示例

检查运行的群集中配置的一致性，并生成详细输出：

```
crm_verify -VL
```

检查指定文件中的配置的一致性，并生成详细输出：

```
crm_verify -Vx file1
```

将配置传送到 `crm_verify`，并生成详细输出：

```
cat file1.xml | crm_verify -Vp
```

文件数

`/var/lib/heartbeat/crm/cib.xml`— 磁盘上的 CIB（去除状态部分）。强烈建议您不要直接编辑此文件。

另请参见

cibadmin(8) [133]

作者

`crm_verify` 由 Andrew Beekhof 编写。

17 群集资源

本章总结了与群集资源相关的最重要信息：包括 High Availability Extension 支持的资源代理类、OCF 资源代理的错误代码以及群集对错误代码如何反应、可用的资源选项、资源操作和实例属性。在配置资源（可使用 `crm` 命令行工具手动配置或使用 Linux HA Management Client）时可使用此概述作为参考。

17.1 支持的资源代理类

您需要为添加的每个群集资源定义资源代理应符合的标准。资源代理可抽象化所提供的服务，并为群集提供精确的状态，使群集与它管理的资源无关。当指定启动、停止或监视命令时，群集会依赖资源代理来执行正确的操作。

资源代理通常以壳层脚本的形式提供。High Availability Extension 支持以下各种资源代理：

旧版 Heartbeat 1 资源代理
: Heartbeat 版本 1 附带自己的资源代理样式。由于很多用户已根据约定编写了自己的代理，所以同样也支持这些资源代理。但是，建议如有可能请将配置迁移到 High Availability OCF RA。有关详细信息，参见 http://wiki.linux-ha.org/HeartbeatResourceAgent。

Linux Standards Base (LSB) 脚本
: LSB 资源代理一般由操作系统/分发包提供，并可在 `/etc/init.d` 中找到。要想用于群集，这些代理必须遵守 LSB 规范。例如，根据 http://www.linuxbase.org/spec/refspecs/LSB_3.0.0/LSB-Core-generic/LSB-Core-generic/iniscrptact.html 中所述，它们必须

已实现一些操作，至少要包括启动、停止、重新启动、重新装载、强制重新装载和状态。

Open Cluster Framework (OCF) 资源代理
: OCF RA 代理最适合用于 High Availability，特别是在您需要主资源或特殊监视功能时。这些代理通常位于 `/usr/lib/ocf/resource.d/heartbeat/`。其功能与 LSB 脚本的功能相似。但是，它们始终使用环境变量来执行配置，这使它们可以轻松地接受和处理参数。OCF 规范（由于它与资源代理相关）可在 http://www.opencf.org/cgi-bin/viewcvs.cgi/specs/ra/resource-agent-api.txt?rev=HEAD_4 中找到。OCF 规范包含以下严格定义：操作必须返回退出代码。群集会严格遵守这些规范。有关详细信息，请参见 http://wiki.linux-ha.org/OCFResourceAgent。有关所有可用的 OCF RA 的详细列表，请参见第 18 章 *HA OCF 代理* [189]。

STONITH 资源代理
: 此类仅用于与屏障相关的资源。有关详细信息，参见第 8 章 *屏障和 STONITH* [75]。

随 High Availability Extension 提供的代理已写入 OCF 规范。

17.2 OCF 返回代码

根据 OCF 规范，有一些关于操作必须返回的退出代码的严格定义。群集会始终检查返回代码与预期结果是否相符。如果结果与预期值不匹配，则将操作视为失败，并将启动恢复操作。有三种类型的故障恢复：

表 17.1 *故障恢复类型*

恢复类型	描述	群集执行的操作
软	发生临时错误。	重启动资源或将它移到新位置。
硬	发生非临时错误。该错误可能特定于当前节点。	将资源移到别处，避免在当前节点上重试该资源。

恢复类型	描述	群集执行的操作
致命	发生所有群集节点共有的非临时错误。这意味着指定了错误的配置。	停止资源，避免在任何群集节点上启动该资源。

假定将某个操作视为已失败，下表概括了不同的 OCF 返回代码以及收到相应的错误代码时群集将启动的恢复类型。

表 17.2 *OCF 返回代码*

OCF 返回代码	OCF 别名	描述	恢复类型
0	OCF_SUCCESS	成功。命令成功完成。这是所有启动、停止、升级和降级命令的所需结果。	软
1	OCF_ERR_GENERIC	通用“出现问题”错误代码。	软
2	OCF_ERR_ARGS	此计算机上的资源配置无效（例如，它引用了节点上找不到的某个位置/工具）。	硬
3	OCF_ERR_UNIMPLEMENTED	请求的操作未实现。	硬
4	OCF_ERR_PERM	资源代理不具备完成该任务的足够特权。	硬
5	OCF_ERR_INSTALLED	资源所需的工具未安装在此计算机上。	硬
6	OCF_ERR_CONFIGURED	资源的配置无效（例如，缺少必需的参数）。	致命
7	OCF_NOT_RUNNING	资源未运行。群集将不会尝试停止为任何操作返回此代码的资源。	不适用

OCF 返回代码	OCF 别名	描述	恢复类型
		此 OCF 返回代码可能需要或不需要资源恢复，这取决于所需的资源状态。如果出现意外，则执行软恢复。	
8	OCF_RUNNING_MASTER	资源正在主节点中运行。	软
9	OCF_FAILED_MASTER	资源在主节点中，但已失败。资源将再次被降级、停止再重启动（然后也可能升级）。	软
其他	不适用	自定义错误代码。	软

17.3 资源选项

您可以为添加的每个资源定义选项。群集使用这些选项来决定资源的行为方式，它们会告知 CRM 如何对待特定的资源。可使用 `crm_resource --meta` 命令或 GUI 来设置资源选项，请参见。

表 17.3 *原始资源选项*

选项	描述
`priority`	如果不允许所有的资源都处于活动状态，群集会停止优先级较低的资源以便保持较高优先级资源处于活动状态。
`target-role`	群集应试图将此资源保持在何种状态？允许的值：`Stopped` 和 `Started`。

选项	描述
is-managed	是否允许群集启动和停止资源？允许的值：true 和 false。
resource-stickiness	资源留在所处位置的自愿程度如何？默认为 default- resource-stickiness 的值。
migration-threshold	节点上的此资源应发生多少故障后才能确定该节点没有资格主管此资源？默认值：none。
multiple-active	如果发现资源在多个节点上活动，群集该如何操作？允许的值：block（将资源标记为未受管）、stop_only 和 stop_start。
failure-timeout	在恢复为如同未发生故障一样正常工作（并允许资源返回它发生故障的节点）之前，需要等待几秒钟？默认值：never。

17.4 资源操作

默认情况下，群集将不会确保您的资源一直正常。要指示群集如此操作，需要向资源的定义中添加一个监视操作。可为所有类或资源代理添加监视操作。

表 17.4 *资源操作*

操作	描述
ID	您的操作名称。必须是唯一的。
name	要执行的操作。常见值：monitor、start 和 stop。
interval	执行操作的频率。单位：秒。
timeout	需要等待多久才能声明操作失败。

操作	描述
requires	需要满足什么条件才能发生此操作。允许的值：nothing、quorum 和 fencing。默认值取决于是否启用屏障和资源的类是否为 stonith。对于 STONITH 资源，默认值为 nothing。
on-fail	此操作失败时执行的操作。允许的值： • ignore：假装资源没有失败。 • block：不对资源执行任何进一步操作。 • stop：停止资源并且不在其他位置启动该资源。 • restart：停止资源并（可能在不同的节点上）重启动。 • fence：关闭资源失败的节点 (STONITH)。 • standby：将*所有*资源从资源失败的节点上移走。
enabled	如果值为 false，将操作视为不存在。允许的值：true、false。

17.5 实例属性

可为所有资源类的脚本指定参数，这些参数可确定脚本的行为方式和所控制的服务实例。如果资源代理支持参数，可以使用 crm_resource 命令添加参数，请参见。在 crm 命令行实用程序中，实例属性称为 params。通过以 root 身份执行以下命令可以找到 OCF 脚本支持的实例属性列表：

```
crm ra meta resource_agent class
```

例如

```
crm ra meta Ipaddr ocf heartbeat
```

显示内容

18 HA OCF 代理

所有 OCF 代理在启动时都需要设置若干参数。以下概述将说明如何手动操作这些代理。本附录中可用的数据直接来自从各 RA 调用的元数据。在 `/usr/lib/ocf/resource.d/heartbeat/` 中可找到所有这些代理。

配置 RA 时，请省略参数名的 `OCF_RESKEY_` 前缀。方括号中的参数在配置中可以省略。

ocf:anything_ra (7)

ocf:anything_ra — 任意

大纲

OCF_RESKEY_binfile=string [OCF_RESKEY_cmdline_options=string] [OCF_RESKEY_pidfile=string] [OCF_RESKEY_logfile=string] [OCF_RESKEY_errlogfile=string] [OCF_RESKEY_user=string] [OCF_RESKEY_monitor_hook=string] anything_ra [start | stop | monitor | meta-data | validate-all]

描述

这是一个几乎可以管理任何事务的通用 OCF RA。

支持的参数

OCF_RESKEY_binfile=要执行的二进制文件的完整路径名
要执行的二进制文件的全名。预期它会以同一 pid 保持运行，而不只是执行某种操作后退出。

OCF_RESKEY_cmdline_options=命令行选项
传递到二进制文件的命令行选项。

OCF_RESKEY_pidfile=要写入 STDOUT 的文件
要读取 PID 或向其写入 PID 的文件。

OCF_RESKEY_logfile=要写入 STDOUT 的文件
要写入 STDOUT 的文件。

OCF_RESKEY_errlogfile=要写入 STDERR 的文件
要写入 STDERR 的文件。

OCF_RESKEY_user=用户运行命令时的身份
    用户运行命令时的身份。

OCF_RESKEY_monitor_hook=要在监视程序操作中运行的命令
    要在监视程序操作中运行的命令。

ocf:apache (7)

ocf:apache — Apache web 服务器

大纲

OCF_RESKEY_configfile=string [OCF_RESKEY_httpd=string] [OCF_RESKEY_port=integer] [OCF_RESKEY_statusurl=string] [OCF_RESKEY_testregex=string] [OCF_RESKEY_options=string] [OCF_RESKEY_envfiles=string] apache [start | stop | status | monitor | meta-data | validate-all]

描述

这是 Apache web 服务器的资源代理。该资源代理可操作 V1.x 和 V2.x Apache 服务器。启动操作以循环结束，在这个循环中反复调用监视程序，以确保服务器已启动并可用。因此，如果监视程序操作未在启动操作超时前成功，apache 资源将以错误状态结束。监视程序操作时，默认装载服务器状态页，后者取决于 mod_status 模块和对应的配置文件（通常是 /etc/apache2/mod_status.conf）。确保服务器状态页可用，并且*只*允许从本地主机（地址 127.0.0.1）访问。细节请查看 statusurl 和 testregex 属性。另请参见 http://httpd.apache.org/

支持的参数

OCF_RESKEY_configfile=配置文件路径
: Apache 配置文件的完整路径名。本文件经过分析，为各种其他资源代理参数提供默认值。

OCF_RESKEY_httpd=httpd 二进制文件路径
: httpd 二进制文件的完整路径名（可选）。

OCF_RESKEY_port=httpd 端口
: 可以用 statusurl 探测状态信息的端口号。默认是配置文件中的端口号，如果配置文件里没有就是 80。

`OCF_RESKEY_statusurl`=url 名称
要监视的 URL（默认的 apache 服务器状态页）。如果不指定，将从 apache 配置文件推断。如果已设置，要确保它*只*继承自本地主机 (127.0.0.1)。否则，可能会出现群集报告资源在多个节点上处于活动状态的情况。

`OCF_RESKEY_testregex`=监视正则表达式
要在 statusurl 的输出中匹配的正则表达式。它不区分大小写。

`OCF_RESKEY_options`=命令行选项
启动 apache 时要应用的额外选项。请参见 man httpd(8)。

`OCF_RESKEY_envfiles`=环境设置文件
包含额外环境变量（例如 /etc/apache2/envvars）的文件（一个或多个）。

ocf:AudibleAlarm (7)

ocf:AudibleAlarm — AudibleAlarm 资源代理

大纲

[OCF_RESKEY_nodelist=string] AudibleAlarm [start | stop | restart | status | monitor | meta-data | validate-all]

描述

AudibleAlarm 的资源脚本。它设置按所设间隔鸣响的音频响铃。

支持的参数

OCF_RESKEY_nodelist=节点列表
从不鸣响铃的节点列表。

ocf:ClusterMon (7)

ocf:ClusterMon — ClusterMon 资源代理

大纲

[OCF_RESKEY_user=string] [OCF_RESKEY_update=integer]
[OCF_RESKEY_extra_options=string] OCF_RESKEY_pidfile=string
OCF_RESKEY_htmlfile=string ClusterMon [start | stop | monitor | meta-data | validate-all]

描述

这是 ClusterMon 资源代理。它会将当前群集状态输出到 html。

支持的参数

OCF_RESKEY_user=用户运行 crm_mon 时的身份
用户运行 crm_mon 时的身份。

OCF_RESKEY_update=更新间隔
更新群集状态的频率。

OCF_RESKEY_extra_options=其他选项
要传递到 crm_mon 的其他选项。例如-n -r。

OCF_RESKEY_pidfile=PID 文件
PID 文件位置，以确保只有一个实例在运行。

OCF_RESKEY_htmlfile=HTML 输出
写入 HTML 输出的位置。

ocf:db2 (7)

ocf:db2 — db2 资源代理

大纲

[OCF_RESKEY_instance=string] [OCF_RESKEY_admin=string] db2 [start | stop | status | monitor | validate-all | meta-data | methods]

描述

db2 的资源脚本。它以 HA 资源的形式管理 DB2 Universal Database。

支持的参数

OCF_RESKEY_instance=实例
    数据库的实例。

OCF_RESKEY_admin=admin
    该实例的 admin 用户。

ocf:Delay (7)

ocf:Delay — 延迟资源代理

大纲

[OCF_RESKEY_startdelay=integer] [OCF_RESKEY_stopdelay=integer] [OCF_RESKEY_mondelay=integer] Delay [start | stop | status | monitor | meta-data | validate-all]

描述

该脚本是引入延迟的测试资源。

支持的参数

OCF_RESKEY_startdelay=启动延迟
　　启动操作中延迟时间的长短（单位：秒）。

OCF_RESKEY_stopdelay=停止延迟
　　停止操作中延迟的秒数。如果未指定，默认为“startdelay”。

OCF_RESKEY_mondelay=监视程序延迟
　　监视程序操作中延迟的秒数。如果未指定，默认为“startdelay”。

ocf:drbd (7)

ocf:drbd — 该资源代理将 Distributed Replicated Block Device (DRBD) 对象作为主/从属资源管理。DRBD 是一种复制储存的机制；有关设置细节，请参见文档。

大纲

OCF_RESKEY_drbd_resource=string [OCF_RESKEY_drbdconf=string]
[OCF_RESKEY_clone_overrides_hostname=boolean]
[OCF_RESKEY_clone_max=integer] [OCF_RESKEY_clone_node_max=integer]
[OCF_RESKEY_master_max=integer]
[OCF_RESKEY_master_node_max=integer] drbd [start | promote | demote | notify | stop | monitor | monitor | meta-data | validate-all]

描述

DRBD 主/从属 OCF 资源代理。

支持的参数

OCF_RESKEY_drbd_resource=drbd 资源名称
来自 drbd.conf 文件的 drbd 资源名称。

OCF_RESKEY_drbdconf=drbd.conf 的路径
drbd.conf 文件的完整路径。

OCF_RESKEY_clone_overrides_hostname=覆盖 drbd 主机名
是否用克隆成员覆盖主机名。它可用于创建浮动对等配置；将告诉 drbd 用 node_<cloneno> 而不是真实 uname 作为主机名，随即便可用于 drbd.conf 中。

OCF_RESKEY_clone_max=克隆数
该 drbd 资源的克隆数。请勿篡改默认值。

OCF_RESKEY_clone_node_max=节点数
    每节点克隆数。请勿篡改默认值。

OCF_RESKEY_master_max=主节点数
    活动主节点数。请勿篡改默认值。

OCF_RESKEY_master_node_max=每个节点的主节点数
    每个节点的主节点最大数。请勿篡改默认值。

ocf:Dummy (7)

ocf:Dummy — 虚设资源代理

大纲

OCF_RESKEY_state=string Dummy [start | stop | monitor | reload | migrate_to | migrate_from | meta-data | validate-all]

描述

这是虚设资源代理。它只跟踪是否在运行，不执行其他任何操作。它在运行中的作用是测试并用作 RA 编写程序的模板。

支持的参数

OCF_RESKEY_state=状态文件
    储存资源状态的位置。

ocf:eDir88 (7)

ocf:eDir88 — eDirectory 资源代理

大纲

OCF_RESKEY_eDir_config_file=string
[OCF_RESKEY_eDir_monitor_ldap=boolean]
[OCF_RESKEY_eDir_monitor_idm=boolean]
[OCF_RESKEY_eDir_jvm_initial_heap=integer]
[OCF_RESKEY_eDir_jvm_max_heap=integer]
[OCF_RESKEY_eDir_jvm_options=string] eDir88 [start | stop | monitor | meta-data | validate-all]

描述

管理 eDirectory 实例的资源脚本。将 eDirectory 的单独实例作为 HA 资源管理。在 v8.8 中添加了“多实例”功能或 eDirectory。该脚本不适用于早于 8.8 的任何 eDirectory 版本。该 RA 可用于在同一主机上装载多个 eDirectory 实例。强烈建议将 eDir 配置文件（按照 eDir_config_file 参数）放在每个节点的本地储存上。该 RA 必须能处理共享储存不可用的情形。如果 eDir 配置文件不可用，该 RA 将失效，检测信号将无法管理资源。不利影响包括 STONITH 操作、无法管理的资源等... 强烈建议设置一个较高的操作超时值。带有 IDM 的 eDir 可能需要 10 多分钟才能启动。如果检测信号在 eDir 有机会正常启动前已超时，可能会发生损害。LDAP 模块可能正是最后启动的模块之一。因此，如果启动了 IDM 和/或 LDAP 的监视，带有 IDM 和 LDAP 的安装上该脚本启动会较慢，因为启动命令要等待 IDM 和 LDAP 可用。

支持的参数

OCF_RESKEY_eDir_config_file=eDir 配置文件
eDirectory 实例配置文件的路径

OCF_RESKEY_eDir_monitor_ldap=eDir 监视 ldap
要监视 LDAP 是否在为 eDirectory 实例而运行吗？

OCF_RESKEY_eDir_monitor_idm=eDir 监视 IDM
要监视 IDM 是否在为 eDirectory 实例而运行吗？

OCF_RESKEY_eDir_jvm_initial_heap=DHOST_INITIAL_HEAP 值
DHOST_INITIAL_HEAP java 环境变量的值。如果未设置，将使用 java 默认值。

OCF_RESKEY_eDir_jvm_max_heap=DHOST_MAX_HEAP 值
DHOST_MAX_HEAP java 环境变量的值。如果未设置，将使用 java 默认值。

OCF_RESKEY_eDir_jvm_options=DHOST_OPTIONS 值
DHOST_OPTIONS java 环境变量的值。如果未设置，将使用原始值。

ocf:Filesystem (7)

ocf:Filesystem — 文件系统资源代理

大纲

[OCF_RESKEY_device=string] [OCF_RESKEY_directory=string]
[OCF_RESKEY_fstype=string] [OCF_RESKEY_options=string] Filesystem
[start | stop | notify | monitor | validate-all | meta-data]

描述

文件系统的资源脚本。它管理共享储存媒体上的文件系统。

支持的参数

OCF_RESKEY_device=块设备
    文件系统块设备的名称、装入的 -U、-L 选项或 NFS 装入规范。

OCF_RESKEY_directory=安装点
    文件系统的安装点。

OCF_RESKEY_fstype=文件系统类型
    要装入的文件系统的可选类型。

OCF_RESKEY_options=选项
    要以 -o选项给出的任何其他待装入选项。对于绑定装入，在这里添加“bind”，
    将 fstype 设置为“none”。能正确处理“bind,ro”之类的选项。

ocf:ICP (7)

ocf:ICP — ICP 资源代理

大纲

[OCF_RESKEY_driveid=string] [OCF_RESKEY_device=string] ICP [start | stop | status | monitor | validate-all | meta-data]

描述

ICP 的资源脚本。它将 ICP Vortex 群集主机驱动器作为 HA 资源管理。

支持的参数

OCF_RESKEY_driveid=ICP 群集驱动器 ID
    ICP 群集驱动器 ID。

OCF_RESKEY_device=设备
    设备名称。

ocf:ids (7)

ocf:ids — IBM 的数据库服务器的 OCF 资源代理称为 Informix Dynamic Server (IDS)

大纲

[OCF_RESKEY_informixdir=string] [OCF_RESKEY_informixserver=string] [OCF_RESKEY_onconfig=string] [OCF_RESKEY_dbname=string] [OCF_RESKEY_sqltestquery=string] ids [start | stop | status | monitor | validate-all | meta-data | methods | usage]

描述

OCF 资源代理用于将 IBM Informix Dynamic Server (IDS) 实例作为高可用性资源管理。

支持的参数

OCF_RESKEY_informixdir= INFORMIXDIR 环境变量
IDS 的典型安装完成后环境变量 INFORMIXDIR 所具有的值。换句话说，就是安装 IDS 的路径（结尾无 / ）。如果该参数未指定，脚本将尝试从壳层环境获取该值。

OCF_RESKEY_informixserver= INFORMIXSERVER 环境变量
IDS 的典型安装完成后环境变量 INFORMIXSERVER 所具有的值。换句话说，就是要管理的 IDS 服务器实例的名称。如果该参数未指定，脚本将尝试从壳层环境获取该值。

OCF_RESKEY_onconfig=ONCONFIG 环境变量
IDS 的典型安装完成后环境变量 ONCONFIG 所具有的值。换句话说，就是 INFORMIXSERVER 中为 IDS 实例指定的配置文件的名称。将在 /etc/ 搜索指定的配置文件。如果该参数未指定，脚本将尝试从壳层环境获取该值。

OCF_RESKEY_dbname=用于监视的数据库，默认为“sysmaster”

该参数定义用于监视 IDS 实例的数据库。如果该参数未指定，脚本将默认使用“sysmaster”数据库。

OCF_RESKEY_sqltestquery=用于监视的 SQL 测试查询，默认为“SELECT COUNT(*) FROM systables;”

在参数“dbname”指定的数据库上运行的 SQL 测试查询，用于监视 IDS 实例，并判断其功能是否正常。如果该参数未指定，脚本将默认使用“SELECT COUNT(*) FROM systables;”。

ocf:IPaddr2 (7)

ocf:IPaddr2 — 管理虚拟 IPv4 地址

大纲

OCF_RESKEY_ip=string [OCF_RESKEY_nic=string]
[OCF_RESKEY_cidr_netmask=string] [OCF_RESKEY_broadcast=string]
[OCF_RESKEY_iflabel=string] [OCF_RESKEY_lvs_support=boolean]
[OCF_RESKEY_mac=string] [OCF_RESKEY_clusterip_hash=string]
[OCF_RESKEY_unique_clone_address=boolean]
[OCF_RESKEY_arp_interval=integer] [OCF_RESKEY_arp_count=integer]
[OCF_RESKEY_arp_bg=string] [OCF_RESKEY_arp_mac=string] IPaddr2 [start
| stop | status | monitor | meta-data | validate-all]

描述

这一 Linux 专用资源管理 IP 别名 IP 地址。它可以添加和删除 IP 别名。此外，它作为克隆资源调用时，可实现群集别名 IP 功能。

支持的参数

OCF_RESKEY_ip=IPv4 地址
IPv4 地址以由“...”连接的四个小于 255 的数字表示，如“192.168.1.1”。

OCF_RESKEY_nic=网络接口
IP 地址用来联机的基本网络接口。如果保留空白，脚本将从路由选择表尝试和判断。请勿在这里以 eth0:1 形式指定别名接口；而是只指定基本接口。

OCF_RESKEY_cidr_netmask=CIDR 网络屏蔽
CIDR 格式的接口网络屏蔽（例如 24，而不是 255.255.255.0）。如果未指定，脚本也将从路由选择表尝试判断。

OCF_RESKEY_broadcast=广播地址
与 IP 关联的广播地址。如果保留空白，脚本将从网络屏蔽判断。

OCF_RESKEY_iflabel=接口标签
您可以在这里为 IP 地址指定其他标签。该标签将附加到您的接口名称之后。如果标签以网络接口卡名称指定，该参数无效。

OCF_RESKEY_lvs_support=启用对 LVS DR 的支持
启用对 LVS Direct Routing 配置的支持。停用 IP 地址时，只将其移到回写设备，让本地节点能继续响应请求，但不再把它发布到网络上。

OCF_RESKEY_mac=群集 IP MAC 地址
显式设置接口的 MAC 地址。目前只用于群集 IP 别名的情况。若保留空白，将自动选择。

OCF_RESKEY_clusterip_hash=群集 IP 哈希函数
指定用于群集 IP 功能的哈希算法。

OCF_RESKEY_unique_clone_address=为克隆实例创建唯一地址
如果为 true，将把克隆 ID 添加到 ip 的提供值，以创建用于管理的唯一地址。

OCF_RESKEY_arp_interval=ARP 包间隔（以毫秒为单位）
指定未请求 ARP 包之间的间隔（以毫秒为单位）。

OCF_RESKEY_arp_count=ARP 包计数
要发送的未请求 ARP 包数。

OCF_RESKEY_arp_bg=来自后台的 ARP
是否发送后台中的 arp 包。

OCF_RESKEY_arp_mac=ARP MAC
同时发送 ARP 包的 MAC 地址。不建议您使用。

ocf:IPaddr (7)

ocf:IPaddr — 管理虚拟 IPv4 地址

大纲

OCF_RESKEY_ip=string [OCF_RESKEY_nic=string]
[OCF_RESKEY_cidr_netmask=string] [OCF_RESKEY_broadcast=string]
[OCF_RESKEY_iflabel=string] [OCF_RESKEY_lvs_support=boolean]
[OCF_RESKEY_local_stop_script=string]
[OCF_RESKEY_local_start_script=string]
[OCF_RESKEY_ARP_INTERVAL_MS=integer]
[OCF_RESKEY_ARP_REPEAT=integer]
[OCF_RESKEY_ARP_BACKGROUND=boolean]
[OCF_RESKEY_ARP_NETMASK=string] IPaddr [start | stop | monitor | validate-all | meta-data]

描述

该脚本管理 IP 别名 IP 地址 它可以添加或删除 IP 别名。

支持的参数

OCF_RESKEY_ip=IPv4 地址
IPv4 地址以由“...”连接的四个小于 255 的数字表示，如 192.168.1.1。

OCF_RESKEY_nic=网络接口
IP 地址用来联机的基本网络接口。如果保留空白，脚本将从路由选择表尝试和判断。不要在这里以 eth0:1 形式指定别名接口；而是只指定基本接口。

OCF_RESKEY_cidr_netmask=网络屏蔽
接口网络屏蔽，CIDR 格式。（即 24），或以由“...”连接的四个小于 255 的数字表示：255.255.255.0）。如果未指定，脚本还将尝试从路由选择表判断。

OCF_RESKEY_broadcast=广播地址
与 IP 关联的广播地址。如果保留空白，脚本将从网络屏蔽判断。

OCF_RESKEY_iflabel=接口标签
您可以在这里为您的 IP 地址指定其他标签。

OCF_RESKEY_lvs_support=启用对 LVS DR 的支持
启用对 LVS Direct Routing 配置的支持。停用 IP 地址时，只将其移到回写设备，让本地节点能继续响应请求，但不再把它发布到网络上。

OCF_RESKEY_local_stop_script=释放该 IP 时调用的脚本
释放该 IP 时调用的脚本。

OCF_RESKEY_local_start_script=添加该 IP 时调用的脚本
添加该 IP 时调用的脚本。

OCF_RESKEY_ARP_INTERVAL_MS=无偿 ARP 之间的毫秒数
无偿 ARP 之间的毫秒数。

OCF_RESKEY_ARP_REPEAT=重复计数
打开新地址时，发送的无偿 ARP 数量。

OCF_RESKEY_ARP_BACKGROUND=在后台运行
在后台运行（这样做不再需要任何理由）。

OCF_RESKEY_ARP_NETMASK=ARP 的网络屏蔽
ARP 的网络屏蔽 - 以非标准十六进制格式表示。

ocf:IPsrcaddr (7)

ocf:IPsrcaddr — IPsrcaddr 资源代理

大纲

[OCF_RESKEY_ipaddress=string] IPsrcaddr [start | stop | stop | monitor | validate-all | meta-data]

描述

IPsrcaddr 的资源脚本。它管理首选源地址修改。

支持的参数

OCF_RESKEY_ipaddress=IP 地址
    IP 地址。

ocf:IPv6addr (7)

ocf:IPv6addr — 管理 IPv6 别名

大纲

[OCF_RESKEY_ipv6addr=string] IPv6addr [start | stop | status | monitor | validate-all | meta-data]

描述

该脚本管理 IPv6 别名 IPv6 地址，它可以添加或删除 IP6 别名。

支持的参数

OCF_RESKEY_ipv6addr=IPv6 地址
该 RA 将管理的 IPv6 地址。

ocf:iscsi (7)

ocf:iscsi — iscsi 资源代理

大纲

[OCF_RESKEY_portal=string] OCF_RESKEY_target=string
[OCF_RESKEY_discovery_type=string] [OCF_RESKEY_iscsiadm=string]
[OCF_RESKEY_udev=string] iscsi [start | stop | status | monitor | validate-all |
methods | meta-data]

描述

iSCSI 的 OCF 资源代理。 添加（启动）或删除（停止）iSCSI 目标。

支持的参数

OCF_RESKEY_portal=门户
iSCSI 门户地址的格式是：{ip 地址|主机名}[":"端口]

OCF_RESKEY_target=目标
iSCSI 目标。

OCF_RESKEY_discovery_type=发现类型
发现类型。目前，open-iscsi 只支持 sendtargets 类型。

OCF_RESKEY_iscsiadm=iscsiadm
iscsiadm 程序路径。

OCF_RESKEY_udev=udev
如果下个资源取决于创建设备的 udev，则等待它完成。在正常装载的主机上，该步骤要快速完成，否则可能会遇到问题。如果您不用 udev，将该选项设置为“no”，否则会无限循环直到超时。

ocf:Ldirectord (7)

ocf:Ldirectord — ldirectord 的封装程序 OCF 资源代理

大纲

OCF_RESKEY_configfile=string [OCF_RESKEY_ldirectord=string]
Ldirectord [start | stop | monitor | meta-data | validate-all]

描述

它是一个简单的 ldirectord 的 OCF RA 封装程序，用 ldirectord 接口创建符合 OCF 的接口。 您就可以监视 ldirectord 了。注意：查询 ldirectord 状态是个开销很大的操作。

支持的参数

OCF_RESKEY_configfile=配置文件路径
ldirectord 配置文件的完整路径名。

OCF_RESKEY_ldirectord=ldirectord 二进制文件路径
ldirectord 的完整路径名。

ocf:LinuxSCSI (7)

ocf:LinuxSCSI — LinuxSCSI 资源代理

大纲

[OCF_RESKEY_scsi=string] LinuxSCSI [start | stop | methods | status | monitor | meta-data | validate-all]

描述

这是 LinuxSCSI 的资源代理。它从 linux 内核的视角管理 SCSI 设备的可用性。
它使 Linux 认为该设备已卸载或再次装载。

支持的参数

OCF_RESKEY_scsi=SCSI 实例
    要管理的 SCSI 实例。

ocf:LVM (7)

ocf:LVM — LVM 资源代理

大纲

[OCF_RESKEY_volgrpname=string] LVM [start | stop | status | monitor | methods | meta-data | validate-all]

描述

LVM 的资源脚本。它将 HA 资源作为 Linux Volume Manager 卷 (LVM) 管理。

支持的参数

OCF_RESKEY_volgrpname=卷组名
卷组名。

ocf:MailTo (7)

ocf:MailTo — MailTo 代理

大纲

[OCF_RESKEY_email=string] [OCF_RESKEY_subject=string] MailTo [start | stop | status | monitor | meta-data | validate-all]

描述

这是 MailTo 的资源代理。它会在每次接管时向 sysadmin 发送电子邮件。

支持的参数

OCF_RESKEY_email=电子邮件地址
sysadmin 的电子邮件地址。

OCF_RESKEY_subject=主题
电子邮件的主题。

ocf:ManageRAID (7)

ocf:ManageRAID — 管理 RAID 设备

大纲

[OCF_RESKEY_raidname=string] ManageRAID [start | stop | status | monitor | validate-all | meta-data]

描述

管理在 /etc/conf.d/HB-ManageRAID 中预配置的 RAID 设备的启动、停止和监视。

支持的参数

OCF_RESKEY_raidname=RAID 名称
要管理的 RAID 的名称（区分大小写）。（在 /etc/conf.d/HB-ManageRAID 中预配置）。

ocf:ManageVE (7)

ocf:ManageVE — OpenVZ VE 资源代理

大纲

[OCF_RESKEY_veid=integer] ManageVE [start | stop | status | monitor | validate-all | meta-data]

描述

这一符合 OCF 的资源代理管理 OpenVZ VE，因此需要包含最新 vzctl util 的适当 OpenVZ 安装。

支持的参数

OCF_RESKEY_veid=VE 的 OpenVZ ID
虚拟环境的 OpenVZ ID（请参见所有指派 ID 的 vzlist -a 的输出）。

ocf:mysql (7)

ocf:mysql — MySQL 资源代理

大纲

[OCF_RESKEY_binary=string] [OCF_RESKEY_config=string]
[OCF_RESKEY_datadir=string] [OCF_RESKEY_user=string]
[OCF_RESKEY_group=string] [OCF_RESKEY_log=string]
[OCF_RESKEY_pid=string] [OCF_RESKEY_socket=string]
[OCF_RESKEY_test_table=string] [OCF_RESKEY_test_user=string]
[OCF_RESKEY_test_passwd=string]
[OCF_RESKEY_enable_creation=integer]
[OCF_RESKEY_additional_parameters=integer] mysql [start | stop | status | monitor | validate-all | meta-data]

描述

MySQL 的资源脚本。它以 HA 资源的形式管理 MySQL 数据库实例。

支持的参数

OCF_RESKEY_binary=MySQL 二进制文件
MySQL 二进制文件的位置。

OCF_RESKEY_config=MySQL 配置文件
配置文件。

OCF_RESKEY_datadir=MySQL 数据目录
包含数据库的目录。

OCF_RESKEY_user=MySQL 用户
运行 MySQL 守护程序的用户。

`OCF_RESKEY_group`=MySQL 组
运行 MySQL 守护程序的组（日志文件和目录访问权限）。

`OCF_RESKEY_log`=MySQL 日志文件
用于 myspld 的日志文件。

`OCF_RESKEY_pid`=MySQL pid 文件
用于 mysqld 的 pid 文件。

`OCF_RESKEY_socket`=MySQL 套接字
用于 mysqld 的套接字。

`OCF_RESKEY_test_table`=MySQL 测试表
监视语句测试的表（database.table 表示法）。

`OCF_RESKEY_test_user`=MySQL 测试用户
MySQL 测试用户。

`OCF_RESKEY_test_passwd`=MySQL 测试用户密码
MySQL 测试用户密码。

`OCF_RESKEY_enable_creation`=创建数据库（如果不存在）
如果 MySQL 数据库不存在，将创建它。

`OCF_RESKEY_additional_parameters`=要传递到 mysqld 的其他参数
启动时传递到 mysqld 的其他参数。（例如 --skip-external-locking 或 --skip-grant-tables）。

ocf:nfsserver (7)

ocf:nfsserver — nfsserver

大纲

[OCF_RESKEY_nfs_init_script=string]
[OCF_RESKEY_nfs_notify_cmd=string]
[OCF_RESKEY_nfs_shared_infodir=string] [OCF_RESKEY_nfs_ip=string]
nfsserver [start | stop | monitor | meta-data | validate-all]

描述

Nfsserver 有助于将 Linux nfs 服务器作为 Linux-HA 中可故障转移的资源进行管理。它取决于特定于 Linux 的 NFS 实现细节，因此尚不能移植到其他平台。

支持的参数

OCF_RESKEY_nfs_init_script= nfsserver 的 init 脚本
随 Linux distro 提供的默认 init 脚本。 nfsserver 资源代理将启动/停止/监视的工作转移到 init 脚本，因为启动/停止/监视 nfsserver 的步骤随 Linux distro 而异。

OCF_RESKEY_nfs_notify_cmd=用于发出通知的工具。
用于发送 NSM 重引导通知的工具。nfsserver 的故障转移可视为重引导到其他计算机。nfsserver 资源代理使用该命令通知所有客户端所发生的故障转移。

OCF_RESKEY_nfs_shared_infodir=用于储存 nfs 相关信息的目录。
nfsserver 资源代理将在这一特定目录中保存 nfs 的相关信息。 该目录必须能够在 nfsserver 本身故障转移前实现故障转移。

OCF_RESKEY_nfs_ip=IP 地址。
用于访问 nfs 服务的浮动 IP 地址。

ocf:oracle (7)

ocf:oracle — oracle 资源代理

大纲

OCF_RESKEY_sid=string [OCF_RESKEY_home=string]
[OCF_RESKEY_user=string] [OCF_RESKEY_ipcrm=string]
[OCF_RESKEY_clear_backupmode=boolean]
[OCF_RESKEY_shutdown_method=string] oracle [start | stop | status | monitor | validate-all | methods | meta-data]

描述

oracle 的资源脚本。将 Oracle 数据库实例作为 HA 资源管理。

支持的参数

OCF_RESKEY_sid=sid
Oracle SID (aka ORACLE_SID)。

OCF_RESKEY_home=用户主目录
Oracle 用户主目录 (aka ORACLE_HOME)。如果未指定，则 SID 及其用户主目录应列在 /etc/oratab 中 。

OCF_RESKEY_user=用户
Oracle 拥有者 (aka ORACLE_OWNER)。如果未指定，则将它设置为文件 $ORACLE_HOME/dbs/*${ORACLE_SID}.ora 的拥有者。如果这对您不适用，只需显式设置它。

OCF_RESKEY_ipcrm=ipcrm
有时属于 Oracle 实例的 IPC 对象（共享的内存段和信号）可能会留到后面，导致实例无法启动。很难判断哪些共享段属于哪个实例，特别是很多实例以同一用户身份运行时。在这里使用的是“oradebug”功能及其“ipc”跟踪实用程序。它并不很适合分析调试信息，但我找不到其他办法来找出 IPC 信息。跟

踪报告的格式或措辞更改时，分析可能会失败。但是，也有些预防措施可以避免误操作。还有一个 dumpinstipc 选项可用于打印属于该实例的 IPC 对象。用它查看对跟踪文件的分析是否正确。有三个设置可用：- none：不干预 IPC，希望最佳情况出现（注意：可能迟早会失败）- instance：试图找出属于该实例的 IPC 材料并删除（默认；应该是安全的）- orauser：删除运行该实例的用户拥有的所有 IPC（如果您以一个用户身份运行多个实例，或者有其他以该用户身份运行的应用程序使用 IPC，则不使用该选项）默认设置“instance”应该是安全的，但在那种情况下不能确保实例会启动。 若 IPC 对象已经被丢下（例如由于某人无情地终止了 Oracle 进程），则无法找出应删除的 IPC 对象。在那种情况下，人工干预是必需的，可能需要终止以同一用户运行的所有实例。第三个设置 orauser 确保 IPC 对象可删除，但它执行删除完全是基于 IPC 对象的所有权，因此只能在每个实例都以独立用户运行时使用。请报告出现的任何问题。欢迎提供建议/修复。

OCF_RESKEY_clear_backupmode=clear_backupmode
清除 ORACLE 的备份模式。

OCF_RESKEY_shutdown_method=关闭方法
如何停止 Oracle 似乎是个人习惯问题。默认方式（“checkpoint/abort”）是：更改系统检查点；关闭中止；这可能是关闭实例最安全的方法。如果您不喜欢“shutdown abort”可将该属性设置为“immediate”，这样就会立即关闭；如果您仍认为有更好的方法可以关闭 Oracle，欢迎提供。

ocf:oralsnr (7)

ocf:oralsnr — oralsnr 资源代理

大纲

OCF_RESKEY_sid=string [OCF_RESKEY_home=string]
[OCF_RESKEY_user=string] OCF_RESKEY_listener=string oralsnr [start | stop | status | monitor | validate-all | meta-data | methods]

描述

Oracle Listener 的资源脚本。它将 Oracle Listener 实例作为 HA 资源管理。

支持的参数

OCF_RESKEY_sid=sid
Oracle SID (aka ORACLE_SID)。对监视操作（即执行 tnsping SID）是必需的。

OCF_RESKEY_home=用户主目录
Oracle 用户主目录 (aka ORACLE_HOME)。如果未指定，则 SID 应列在 /etc/oratab 中。

OCF_RESKEY_user=用户
以该用户身份运行该监听程序。

OCF_RESKEY_listener=监听程序
要启动的监听程序实例（如同在 listener.ora 中定义的）。默认为 LISTENER。

ocf:pgsql (7)

ocf:pgsql — pgsql 资源代理

大纲

[OCF_RESKEY_pgctl=string] [OCF_RESKEY_start_opt=string]
[OCF_RESKEY_ctl_opt=string] [OCF_RESKEY_psql=string]
[OCF_RESKEY_pgdata=string] [OCF_RESKEY_pgdba=string]
[OCF_RESKEY_pghost=string] [OCF_RESKEY_pgport=string]
[OCF_RESKEY_pgdb=string] [OCF_RESKEY_logfile=string]
[OCF_RESKEY_stop_escalate=string] pgsql [start | stop | status | monitor | meta-data | validate-all | methods]

描述

PostgreSQL 的资源脚本。它将 PostgreSQL 作为 HA 资源管理。

支持的参数

OCF_RESKEY_pgctl=pgctl
pg_ctl 命令的路径。

OCF_RESKEY_start_opt=start_opt
启动选项（pgi_ctl 中的 -o start_opt）。例如 -i -p 5432。

OCF_RESKEY_ctl_opt=ctl_opt
其他 pg_ctl 选项（-w，-W 等）。默认值是 ""

OCF_RESKEY_psql=psql
psql 命令的路径。

OCF_RESKEY_pgdata=pgdata
PostgreSQL 数据目录路径。

`OCF_RESKEY_pgdba`=pgdba
拥有 PostgreSQL 的用户。

`OCF_RESKEY_pghost`=pghost
PosrgeSQL 监听的主机名/IP 地址。

`OCF_RESKEY_pgport`=pgport
PosrgeSQL 监听的端口。

`OCF_RESKEY_pgdb`=pgdb
将用于监视的数据库。

`OCF_RESKEY_logfile`=logfile
PostgreSQL 服务器日志输出文件的路径。

`OCF_RESKEY_stop_escalate`=停止增加
采取 -m immediate 前允许重试的次数（使用 -m fast）。

ocf:pingd (7)

ocf:pingd — pingd 资源代理

大纲

[OCF_RESKEY_pidfile=string] [OCF_RESKEY_user=string]
[OCF_RESKEY_dampen=integer] [OCF_RESKEY_set=integer]
[OCF_RESKEY_name=integer] [OCF_RESKEY_section=integer]
[OCF_RESKEY_multiplier=integer] [OCF_RESKEY_host_list=integer]
pingd [start | stop | monitor | meta-data | validate-all]

描述

这是 pingd 资源代理。它记录（在 CIB 中）某节点可连接到的 ping 节点的当前数。

支持的参数

OCF_RESKEY_pidfile=PID 文件
PID 文件

OCF_RESKEY_user=运行 pingd 的用户身份
运行 pingd 的用户身份。

OCF_RESKEY_dampen=衰减间隔
等待（衰减）后续更改的时间。

OCF_RESKEY_set=设置名称
设置用来放置值的实例属性的名称极少需要指定。

OCF_RESKEY_name=属性名称
要设置的属性的名称。这是用于约束的名称。

`OCF_RESKEY_section`=部分名称
放置值的部分。极少需要指定。

`OCF_RESKEY_multiplier`=值乘数
用来乘以连接的 ping 节点数的数字。

`OCF_RESKEY_host_list`=主机列表
要计数的 ping 节点列表。默认为所有已配置的 ping 节点。极少需要指定。

ocf:portblock (7)

ocf:portblock — portblock 资源代理

大纲

[OCF_RESKEY_protocol=string] [OCF_RESKEY_portno=integer] [OCF_RESKEY_action=string] portblock [start | stop | status | monitor | meta-data | validate-all]

描述

portblock 的资源脚本。它用于临时阻止使用 ip 表的端口。

支持的参数

OCF_RESKEY_protocol=协议
    用于已阻止/解除阻止的协议。

OCF_RESKEY_portno=端口号
    用于已阻止/解除阻止的端口号。

OCF_RESKEY_action=操作
    要在 protocol::portno 上完成的操作（阻止/解除阻止）。

ocf:Pure-FTPd (7)

ocf:Pure-FTPd — 适用于 OCF 资源代理的 FTP 脚本。

大纲

OCF_RESKEY_script=string OCF_RESKEY_conffile=string
OCF_RESKEY_daemon_type=string [OCF_RESKEY_pidfile=string]
Pure-FTPd [start | stop | monitor | validate-all | meta-data]

描述

该脚本管理 Active-Passive 设置中的 Pure-FTPd。

支持的参数

OCF_RESKEY_script=带完整路径的脚本名称
Pure-FTPd 启动脚本的完整路径。例如“/sbin/pure-config.pl”

OCF_RESKEY_conffile=带完整路径的配置文件名
带完整路径的 Pure-FTPd 配置文件名。例如“/etc/pure-ftpd/pure-ftpd.conf”

OCF_RESKEY_daemon_type=带完整路径的配置文件名
由 pure-ftpd 封装程序调用的 Pure-FTPd 守护程序。有效选项是：“”用于 pure-ftpd，“mysql”用于 pure-ftpd-mysql，“postgresql”用于 pure-ftpd-postgresql，“ldap”用于 pure-ftpd-ldap

OCF_RESKEY_pidfile=PID 文件
PID 文件

ocf:Raid1 (7)

ocf:Raid1 — RAID1 资源代理

大纲

[OCF_RESKEY_raidconf=string] [OCF_RESKEY_raiddev=string] [OCF_RESKEY_homehost=string] Raid1 [start | stop | status | monitor | validate-all | meta-data]

描述

RAID1 的资源脚本。它管理共享储存媒体上的软件 Raid1 设备。

支持的参数

OCF_RESKEY_raidconf=RAID 配置文件
RAID 配置文件。例如 /etc/raidtab 或 /etc/mdadm.conf。

OCF_RESKEY_raiddev=块设备
要使用的块设备。

OCF_RESKEY_homehost=用于 mdadm 的 homehost
homehost 指令的值；这是避免 RAID 意外激活的 mdadm 功能。建议创建由“homehost”设置为特殊值的群集管理的 RAID，这样它们不会被不应拥有它们的节点意外地自动组合。

ocf:Route (7)

ocf:Route — 管理网络路由

大纲

OCF_RESKEY_destination=string OCF_RESKEY_device=string OCF_RESKEY_gateway=string OCF_RESKEY_source=string Route [start | stop | monitor | reload | meta-data | validate-all]

描述

启用和禁用网络路由。支持主机和网络路由、通过网关地址的路由，以及使用特定源地址的路由。如果节点的路由表需要根据节点角色指派来操纵，则该资源代理很有用。请考虑以下示范用例：一个群集节点用作 IPsec 隧道端点。- 所有其他节点都使用IPsec隧道抵达特定远程网络中的主机。然后，以下是您如何利用路由资源代理实现该模式：配置 ipsec LSB 资源。- 配置克隆路由 OCF 资源。- 创建顺序约束，确保该 ipsec 在路由前启动。 - 创建 ipsec 和路由资源之间的排列约束，确保您的克隆路由资源实例不会在隧道端点本身启动之前启动。

支持的参数

OCF_RESKEY_destination=目标网络
　要为路由配置的目标网络（或主机）。以 CIDR 表示法指定网络屏蔽后缀（例如“/24”）。如果没有给出后缀，将创建主机路由。如果希望该资源设置系统默认路由，请指定“0.0.0.0/0”或“default”。

OCF_RESKEY_device=出去的网络设备
　用于该路由的出去的网络设备。

OCF_RESKEY_gateway=网关 IP 地址
　用于该路由的网关 IP 地址。

OCF_RESKEY_source=源 IP 地址
　为该路由配置的源 IP 地址。

ocf:rsyncd (7)

ocf:rsyncd — 符合 OCF 资源代理的 rsync 守护程序脚本。

大纲

[OCF_RESKEY_binpath=string] [OCF_RESKEY_conffile=string] [OCF_RESKEY_bwlimit=string] rsyncd [start | stop | monitor | validate-all | meta-data]

描述

该脚本管理 rsync 守护程序。

支持的参数

OCF_RESKEY_binpath=rsync 二进制文件的完整路径
rsync 二进制文件路径。例如“/usr/bin/rsync”

OCF_RESKEY_conffile=带完整路径的配置文件名
带完整路径的 rsync 守护程序配置文件名。例如“/etc/rsyncd.conf”

OCF_RESKEY_bwlimit=限制 I/O 带宽，单位每秒千字节
该选项可用于指定最大传输速率，单位每秒千字节。该选项在将 rsync 用于大文件（数兆字节以上）时最有效。由于 rsync 传输的性质，将发送数据块，然后若 rsync 判断传输过快，将在发送下个数据块之前等待。结果使得平均传输速率等于指定的上限。值为零表示没有限制。

ocf:SAPDatabase (7)

ocf:SAPDatabase — SAP 数据库资源代理

大纲

OCF_RESKEY_SID=string OCF_RESKEY_DIR_EXECUTABLE=string
OCF_RESKEY_DBTYPE=string OCF_RESKEY_NETSERVICENAME=string
OCF_RESKEY_DBJ2EE_ONLY=boolean OCF_RESKEY_JAVA_HOME=string
OCF_RESKEY_STRICT_MONITORING=boolean
OCF_RESKEY_AUTOMATIC_RECOVER=boolean
OCF_RESKEY_DIR_BOOTSTRAP=string OCF_RESKEY_DIR_SECSTORE=string
OCF_RESKEY_DB_JARS=string OCF_RESKEY_PRE_START_USEREXIT=string
OCF_RESKEY_POST_START_USEREXIT=string
OCF_RESKEY_PRE_STOP_USEREXIT=string
OCF_RESKEY_POST_STOP_USEREXIT=string SAPDatabase [start | stop | status | monitor | validate-all | meta-data | methods]

描述

SAP 数据库的资源脚本。它将任何类型的 SAP 数据库作为 HA 资源管理。

支持的参数

OCF_RESKEY_SID=SAP 系统 ID
唯一的 SAP 系统标识符。例如 P01。

OCF_RESKEY_DIR_EXECUTABLE=sapstartsrv 和 sapcontrol 的路径
sapstartsrv 和 sapcontrol 的完全限定路径。

OCF_RESKEY_DBTYPE=数据库供应商
您使用的数据库供应商名称。设置为：ORA、DB6 或 ADA。

OCF_RESKEY_NETSERVICENAME=监听程序名称
Oracle TNS 监听程序名称。

OCF_RESKEY_DBJ2EE_ONLY=只安装了 JAVA 堆栈
如果在 SAP 数据库中没有安装 ABAP 堆栈，将它设置为 TRUE。

OCF_RESKEY_JAVA_HOME=Java SDK 的路径
仅当 DBJ2EE_ONLY 参数设置为 true 时才需要它。输入 SAP WebAS Java 使用的 Java SDK 的路径。

OCF_RESKEY_STRICT_MONITORING=激活应用程序级监视
将控制资源代理监视数据库的方式。如果设置为 true，将使用 SAP 工具测试与数据库的连接。不要与 Oracle 一起使用，因为这会导致在存档程序卡住时发生不需要的故障转移。

OCF_RESKEY_AUTOMATIC_RECOVER=启用或禁用自动启动恢复
SAPDatabase 资源代理自动尝试恢复一次失败的启动。这可以通过运行 RDBMS 的强制中止和/或执行恢复命令来完成。

OCF_RESKEY_DIR_BOOTSTRAP=j2ee bootstrap 目录的路径
J2EE 实例 bootstrap 目录的完全限定路径。例如 /usr/sap/P01/J00/j2ee/cluster/bootstrap

OCF_RESKEY_DIR_SECSTORE=j2ee 安全储存目录的路径
J2EE 安全储存目录的完全限定路径。例如 /usr/sap/P01/SYS/global/security/lib/tools。

OCF_RESKEY_DB_JARS=jdbc 驱动程序的文件名
数据库连接测试的 jdbc 驱动程序的完全限定文件名。在 Java 引擎 6.40 和 7.00 中，它会自动从 bootstrap.properties 文件读取。对 Java 引擎 7.10，该参数是必需的。

OCF_RESKEY_PRE_START_USEREXIT=启动前脚本的路径
该资源启动前应执行的脚本或程序的完全限定路径。

OCF_RESKEY_POST_START_USEREXIT=启动后脚本的路径
该资源启动后应执行的脚本或程序的完全限定路径。

OCF_RESKEY_PRE_STOP_USEREXIT=启动前脚本的路径
该资源停止前应执行的脚本或程序的完全限定路径。

OCF_RESKEY_POST_STOP_USEREXIT=启动后脚本的路径
该资源停止后应执行的脚本或程序的完全限定路径。

ocf:SAPInstance (7)

ocf:SAPInstance — SAP 实例资源代理

大纲

OCF_RESKEY_InstanceName=string OCF_RESKEY_DIR_EXECUTABLE=string
OCF_RESKEY_DIR_PROFILE=string OCF_RESKEY_START_PROFILE=string
OCF_RESKEY_START_WAITTIME=string
OCF_RESKEY_AUTOMATIC_RECOVER=boolean
OCF_RESKEY_PRE_START_USEREXIT=string
OCF_RESKEY_POST_START_USEREXIT=string
OCF_RESKEY_PRE_STOP_USEREXIT=string
OCF_RESKEY_POST_STOP_USEREXIT=string SAPInstance [start | recover | stop | status | monitor | validate-all | meta-data | methods]

描述

SAP 的资源脚本。它将 SAP 实例作为 HA 资源管理。

支持的参数

OCF_RESKEY_InstanceName=实例名称：SID_INSTANCE_VIR-HOSTNAME
完全限定 SAP 实例名称。例如 P01_DVEBMGS00_sapp01ci。

OCF_RESKEY_DIR_EXECUTABLE=sapstartsrv 和 sapcontrol 的路径
sapstartsrv 和 sapcontrol 的完全限定路径。

OCF_RESKEY_DIR_PROFILE=启动配置文件的路径
SAP 启动配置文件的完全限定路径。

OCF_RESKEY_START_PROFILE=启动配置文件名称
SAP 启动配置文件的名称。

OCF_RESKEY_START_WAITTIME=检查该时间后的成功启动次数（不等待J2EE-Addin）

该秒数过后，资源代理将执行监视操作。如果监视程序返回 SUCCESS，启动将处理为SUCCESS。这对于解决时序问题很有用，例如在J2EE-Addin实例中。

OCF_RESKEY_AUTOMATIC_RECOVER=启用或禁用自动启动恢复

SAPInstance 资源代理自动尝试恢复一次失败的启动。这是通过中止运行中的实例进程并执行 cleanipc 实现的。

OCF_RESKEY_PRE_START_USEREXIT=启动前脚本的路径

该资源启动前应执行的脚本或程序的完全限定路径。

OCF_RESKEY_POST_START_USEREXIT=启动后脚本的路径

该资源启动后应执行的脚本或程序的完全限定路径。

OCF_RESKEY_PRE_STOP_USEREXIT=启动前脚本的路径

该资源停止前应执行的脚本或程序的完全限定路径。

OCF_RESKEY_POST_STOP_USEREXIT=启动后脚本的路径

该资源停止后应执行的脚本或程序的完全限定路径。

ocf:scsi2reserve (7)

ocf:scsi2reserve — scsi-2 保留

大纲

[OCF_RESKEY_scsi_reserve=string] [OCF_RESKEY_sharedisk=string] [OCF_RESKEY_start_loop=string] scsi2reserve [start | stop | monitor | meta-data | validate-all]

描述

scsi-2-reserve 资源代理是为 SCSI-2 保留的占位符。scsi-2-reserve 资源的正常实例，表示指定 SCSI 设备的所有权。该资源代理取决于来自 scsires 包的 scsi_reserve，后者是特定于 Linux 的。

支持的参数

OCF_RESKEY_scsi_reserve= scsi_reserve 命令

scsi_reserve 是来自 scsires 包的命令。它帮助实现 SCSI 设备上 SCSI-2 的保留。

OCF_RESKEY_sharedisk=共享磁盘。

可保留的共享磁盘。

OCF_RESKEY_start_loop=放弃前重试的次数。

放弃前将尝试若干次。Start_loop 表示将重试的次数。

ocf:SendArp (7)

ocf:SendArp — SendArp 资源代理

大纲

[OCF_RESKEY_ip=string] [OCF_RESKEY_nic=string] SendArp [start | stop | monitor | meta-data | validate-all]

描述

该脚本发出 IP 地址的无偿 Arp。

支持的参数

OCF_RESKEY_ip=IP 地址
发送 arp 包的 IP 地址。

OCF_RESKEY_nic=NIC
发送 arp 包的 nic。

ocf:ServeRAID (7)

ocf:ServeRAID — ServeRAID 资源代理

大纲

[OCF_RESKEY_serveraid=integer] [OCF_RESKEY_mergegroup=integer]
ServeRAID [start | stop | status | monitor | validate-all | meta-data | methods]

描述

ServeRAID 的资源脚本。它启用/禁用共享 ServeRAID 合并组。

支持的参数

OCF_RESKEY_serveraid=serveraid
ServeRAID 适配器的编号。

OCF_RESKEY_mergegroup=合并组
涉及的逻辑驱动器。

ocf:sfex (7)

ocf:sfex — SF-EX 资源代理

大纲

[OCF_RESKEY_device=string] [OCF_RESKEY_index=integer]
[OCF_RESKEY_collision_timeout=integer]
[OCF_RESKEY_monitor_interval=integer]
[OCF_RESKEY_lock_timeout=integer] sfex [start | stop | monitor | meta-data]

描述

SF-EX 的资源脚本。它以独占方式管理共享储存媒体。

支持的参数

OCF_RESKEY_device=块设备
储存独占控制数据的块设备路径。

OCF_RESKEY_index=索引
储存独占控制数据的块设备中的位置。指定 1 或更多。默认值是 1。

OCF_RESKEY_collision_timeout=锁定获取的等待时间
检测到锁定获取有冲突时的等待时间。默认为 1 秒。

OCF_RESKEY_monitor_interval=监视间隔
监视间隔（以秒为单位）。默认为 10 秒

OCF_RESKEY_lock_timeout=有效锁定期
有效锁定期（以秒为单位）。默认为 20 秒。

ocf:SphinxSearchDaemon (7)

ocf:SphinxSearchDaemon — searchd 资源代理

大纲

OCF_RESKEY_config=string [OCF_RESKEY_searchd=string]
[OCF_RESKEY_search=string] [OCF_RESKEY_testQuery=string]
SphinxSearchDaemon [start | stop | monitor | meta-data | validate-all]

描述

这是 searchd 资源代理。它管理 Sphinx 搜索守护程序。

支持的参数

OCF_RESKEY_config=配置文件
searchd 配置文件。

OCF_RESKEY_searchd=searchd 二进制文件
searchd 二进制文件。

OCF_RESKEY_search=搜索二进制文件
监视操作中搜索二进制文件以进行功能测试。

OCF_RESKEY_testQuery=测试查询
监视操作中查询测试以进行功能测试。该查询无需匹配索引中的任何文档。
目的只是测试搜索守护程序是否能查询索引并正确响应。

ocf:Squid (7)

ocf:Squid — Squid 的 RA

大纲

[OCF_RESKEY_squid_exe=string] OCF_RESKEY_squid_conf=string
OCF_RESKEY_squid_pidfile=string OCF_RESKEY_squid_port=integer
[OCF_RESKEY_squid_stop_timeout=integer]
[OCF_RESKEY_debug_mode=string] [OCF_RESKEY_debug_log=string] Squid
[start | stop | status | monitor | meta-data | validate-all]

描述

Squid 的资源代理。它将 Squid 实例作为 HA 资源管理。

支持的参数

OCF_RESKEY_squid_exe=可执行文件
这是必需的参数。该参数指定 squid 的可执行文件。

OCF_RESKEY_squid_conf=配置文件
这是必需的参数。该参数为由该 RA 管理的 squid 实例指定配置文件。

OCF_RESKEY_squid_pidfile=Pid 文件
这是必需的参数。该参数为由该 RA 管理的 squid 实例指定进程 id 文件。

OCF_RESKEY_squid_port=端口号
这是必需的参数。该参数为由该 RA 管理的 squid 实例指定端口号。如果使用多个端口，只能指定其中之一。

OCF_RESKEY_squid_stop_timeout=确认正常停止方法等待的秒数
此参数为可省略参数。在停止操作时，首先使用正常停止方法。随后等待该参数指定的秒数后确认其完成。默认值是 10。

OCF_RESKEY_debug_mode=调试模式

此参数为可选参数。该参数包含 x 或 v 时，该 RA 运行于调试模式。如果包含 x，STDOUT 和 STDERR 都重定向到 debug_log 指定的日志文件，然后打开 builtin 壳层选项 x。对 v 也是相似的。

OCF_RESKEY_debug_log=调试日志的目标

这是可选且可省略的参数。该参数指定调试日志的目标文件，仅适用于 RA 运行于调试模式时。关于调试模式，请参见 debug_mode。如果未给出值但值是必需的，将按以下规则确定：“/var/log/”为目录部分，配置文件的基本名称为“syslog_ng_conf”，“.log”作为后缀。

ocf:Stateful (7)

ocf:Stateful — 监控状态的资源代理示例

大纲

`OCF_RESKEY_state`=string `Stateful` [start | stop | monitor | meta-data | validate-all]

描述

这是实现两种状态的资源代理示例。

支持的参数

`OCF_RESKEY_state`=状态文件
储存资源状态的位置。

ocf:SysInfo (7)

ocf:SysInfo — SysInfo 资源代理

大纲

[OCF_RESKEY_pidfile=string] [OCF_RESKEY_delay=string] SysInfo [start | stop | monitor | meta-data | validate-all]

描述

这是 SysInfo 资源代理。它记录（以 CIB 方式）节点 Sample Linux 输出的各种属性：arch: i686 os: Linux-2.4.26-gentoo-r14 free_swap: 1999 cpu_info: Intel(R) Celeron(R) CPU 2.40GHz cpu_speed: 4771.02 cpu_cores: 1 cpu_load: 0.00 ram_total: 513 ram_free: 117 root_free: 2.4 Sample Darwin output: arch: i386 os: Darwin-8.6.2 cpu_info: Intel Core Duo cpu_speed: 2.16 cpu_cores: 2 cpu_load: 0.18 ram_total: 2016 ram_free: 787 root_free: 13 Units: free_swap: Mb ram_*: Mb root_free: Gb cpu_speed (Linux): bogomips cpu_speed (Darwin): Ghz

支持的参数

OCF_RESKEY_pidfile=PID 文件
    PID 文件。

OCF_RESKEY_delay=衰减延迟
    让值稳定下来的间隔。

ocf:tomcat (7)

ocf:tomcat — tomcat 资源代理

大纲

OCF_RESKEY_tomcat_name=string OCF_RESKEY_script_log=string
[OCF_RESKEY_tomcat_stop_timeout=integer]
[OCF_RESKEY_tomcat_suspend_trialcount=integer]
[OCF_RESKEY_tomcat_user=string] [OCF_RESKEY_statusurl=string]
[OCF_RESKEY_java_home=string] OCF_RESKEY_catalina_home=string
OCF_RESKEY_catalina_pid=string
[OCF_RESKEY_tomcat_start_opts=string]
[OCF_RESKEY_catalina_opts=string]
[OCF_RESKEY_catalina_rotate_log=string]
[OCF_RESKEY_catalina_rotatetime=integer] tomcat [start | stop | status | monitor | meta-data | validate-all]

描述

tomcat 的资源脚本。它将 Tomcat 实例作为 HA 资源管理。

支持的参数

OCF_RESKEY_tomcat_name=资源名称
资源的名称。

OCF_RESKEY_script_log=该脚本日志的目标
该脚本日志的目标。

OCF_RESKEY_tomcat_stop_timeout=停止时超时
停止时超时。

OCF_RESKEY_tomcat_suspend_trialcount=等待停止的重试次数
等待停止的重试次数。

OCF_RESKEY_tomcat_user=启动资源的用户名
启动资源的用户名。

OCF_RESKEY_statusurl=状态确认的 URL
状态确认的 URL。

OCF_RESKEY_java_home=Java 的用户主目录
Java 的用户主目录。

OCF_RESKEY_catalina_home=Tomcat 的用户主目录
Tomcat 的用户主目录。

OCF_RESKEY_catalina_pid=Tomcat 的 PID 文件名
Tomcat 的 PID 文件名。

OCF_RESKEY_tomcat_start_opts=Tomcat 启动选项
Tomcat 启动选项。

OCF_RESKEY_catalina_opts=Catalina 选项
Catalina 选项。

OCF_RESKEY_catalina_rotate_log=旋转 catalina.out 标志
旋转 catalina.out 标志。

OCF_RESKEY_catalina_rotatetime=旋转 catalina.out 的时间范围
旋转 catalina.out 的时间范围。

ocf:VIPArip (7)

ocf:VIPArip — RIP2 协议的虚拟 IP 地址

大纲

OCF_RESKEY_ip=string [OCF_RESKEY_nic=string] VIPArip [start | stop | monitor | validate-all | meta-data]

描述

RIP2 协议的虚拟 IP 地址。该脚本用 quagga/ripd 管理不同子网中的 IP 别名。它可以添加和删除 IP 别名。

支持的参数

OCF_RESKEY_ip=不同子网中的 IP 地址
不同子网中的 IPv4 地址，例如“192.168.1.1”。

OCF_RESKEY_nic=广播路由信息的 nic
广播路由信息的 nic。该 ripd 使用这一 nic 将路由信息广播至他处。

ocf:VirtualDomain (7)

ocf:VirtualDomain — 管理虚拟域

大纲

OCF_RESKEY_config=string [OCF_RESKEY_hypervisor=string]
[OCF_RESKEY_force_stop=boolean]
[OCF_RESKEY_migration_transport=string]
[OCF_RESKEY_monitor_scripts=string] VirtualDomain [start | stop | status | monitor | migrate_from | migrate_to | meta-data | validate-all]

描述

libvirtd 管理的虚拟域（a.k.a. domU、虚拟机、虚拟环境等，具体取决于环境）的资源代理。

支持的参数

OCF_RESKEY_config=虚拟域配置文件
该虚拟域中 libvirt 配置文件的绝对路径。

OCF_RESKEY_hypervisor=超级管理程序 URI
连接到的超级管理程序 URI。有关支持的 URI 格式的细节，请参见 libvirt 文档。该默认值取决于系统。

OCF_RESKEY_force_stop=停止时强制关闭
停止时强制关闭（“清空”）该域。仅当您的虚拟域（或您的虚拟化后端）不支持正常关闭时才启用它。

OCF_RESKEY_migration_transport=远程超级管理程序传输
迁移时用于连接到远程超级管理程序的传输。关于可用传输的细节，请参见 libvert 文档。如果省略该参数，该资源将使用 libvert 的默认传输连接到远程超级管理程序。

OCF_RESKEY_monitor_scripts=以空格分隔的监视程序脚本列表
若还要监视虚拟域内的服务，请同时添加该参数和要监视的脚本列表。注意：使用监视脚本时，启动和“迁移自”操作仅在所有监视脚本都顺利完成时才完成。请确保设置这些操作的超时，以适应该延迟。

ocf:WAS6 (7)

ocf:WAS6 — WAS6 资源代理

大纲

[OCF_RESKEY_profile=string] WAS6 [start | stop | status | monitor | validate-all | meta-data | methods]

描述

WAS6 的资源脚本。它将 Websphere Application Server (WAS6) 作为 HA 资源管理。

支持的参数

OCF_RESKEY_profile=配置文件名称
    WAS 配置文件名称。

ocf:WAS (7)

ocf:WAS — WAS 资源代理

大纲

[OCF_RESKEY_config=string] [OCF_RESKEY_port=integer] WAS [start | stop | status | monitor | validate-all | meta-data | methods]

描述

WAS 的资源脚本。它将 Websphere Application Server (WAS) 作为 HA 资源管理。

支持的参数

OCF_RESKEY_config=配置文件
    WAS 配置文件。

OCF_RESKEY_port=端口
    WAS (snoop) 端口号。

ocf:WinPopup (7)

ocf:WinPopup — WinPopup 资源代理

大纲

[OCF_RESKEY_hostfile=string] WinPopup [start | stop | status | monitor | validate-all | meta-data]

描述

WinPopup 的资源脚本。它会在每次发生接管时向 sysadmin 的工作站发送 WinPopup 消息。

支持的参数

OCF_RESKEY_hostfile=主机文件
包含要发送 WinPopup 消息的主机的文件。

ocf:Xen (7)

ocf:Xen — 管理 Xen DomU

大纲

[OCF_RESKEY_xmfile=string] [OCF_RESKEY_name=string]
[OCF_RESKEY_allow_migrate=boolean]
[OCF_RESKEY_shutdown_timeout=boolean]
[OCF_RESKEY_allow_mem_management=boolean]
[OCF_RESKEY_reserved_Dom0_memory=string]
[OCF_RESKEY_monitor_scripts=string] Xen [start | stop | migrate_from | migrate_to | monitor | meta-data | validate-all]

描述

Xen 超级管理程序的资源代理。通过将群集资源启动和停止分别映射到 Xen 创建和关闭，从而管理 Xen 虚拟机实例。名称的说明。我们将试图从配置文件提取名称（xmfile属性）。如果您使用简单的指派语句，则应该没有问题。否则，如果涉及到某些依赖其他变量的 python 特效（例如动态指派名称），并要尝试检测它，请设置名称属性。如果存在任何不正常状况（可能缺少配置文件，例如该文件位于共享储存），也应该这样做。如果这些都失败，就只能靠实例 ID 来保留向后兼容性。半虚拟 guest 也可以通过启用 meta_attribute allow_migrate 来迁移。

支持的参数

OCF_RESKEY_xmfile=Xen 控制文件
　该虚拟机中 Xen 控制文件的绝对路径。

OCF_RESKEY_name=Xen DomU 名称
　虚拟机的名称。

OCF_RESKEY_allow_migrate=使用在线迁移
　该 bool 参数允许对半虚拟机使用在线迁移。

OCF_RESKEY_shutdown_timeout=关闭升级超时
Xen 代理会先尝试用 xm 关闭有序关闭。如果在超时时间内未成功，该代理将进而采用 xm 清空，强制关闭节点。如果未设置，它默认为停止操作超时的三分之二。将该值设置为 0 将强制立即清空。

OCF_RESKEY_allow_mem_management=使用动态内存管理
该操作将对用于 Dom0 和 DomU 的启动和停止操作启用内存的动态调整。默认设置不允许动态调整内存。

OCF_RESKEY_reserved_Dom0_memory=最小 Dom0 内存
使用内存管理时，该参数将定义为 dom0 保留的最小内存。默认的最小内存是 512 MB。

OCF_RESKEY_monitor_scripts=空格分隔的监视程序脚本列表
若还要监视非特权域内的服务，请在同时添加该参数和要监视的脚本列表。注意：在这种情况下，务必要将监视操作的启动延迟设置为至少是 DomU 启动所有服务所需的时间。

ocf:Xinetd (7)

ocf:Xinetd — Xinetd 资源代理

大纲

[OCF_RESKEY_service=string] Xinetd [start | stop | restart | status | monitor | validate-all | meta-data]

描述

Xinetd 的资源脚本。它启动/停止 xinetd 管理的服务。请注意，xinetd 守护程序本身必须在运行：将不会启动或停止它。重要：如果群集管理的服务是唯一启用的服务，应为 xinetd 指定 -stayalive 选项，否则会在检测信号停止时跳出。或者，您也可以启用某些内部服务，例如 echo。

支持的参数

OCF_RESKEY_service=服务名
    xinetd 管理的服务名。

部分 V. 附录

GNU 许可证

A

此附件包含 GNU 通用公共许可证和 GNU 自由文档许可证。

GNU General Public License

Version 2, June 1991

Preamble

The licenses for most software are designed to take away your freedom to share and change it. By contrast, the GNU General Public License is intended to guarantee your freedom to share and change free software--to make sure the software is free for all its users. This General Public License applies to most of the Free Software Foundation's software and to any other program whose authors commit to using it. (Some other Free Software Foundation software is covered by the GNU Library General Public License instead.) You can apply it to your programs, too.

When we speak of free software, we are referring to freedom, not price. Our General Public Licenses are designed to make sure that you have the freedom to distribute copies of free software (and charge for this service if you wish), that you receive source code or can get it if you want it, that you can change the software or use pieces of it in new free programs; and that you know you can do these things.

To protect your rights, we need to make restrictions that forbid anyone to deny you these rights or to ask you to surrender the rights. These restrictions translate to certain responsibilities for you if you distribute copies of the software, or if you modify it.

For example, if you distribute copies of such a program, whether gratis or for a fee, you must give the recipients all the rights that you have. You must make sure that they, too, receive or can get the source code. And you must show them these terms so they know their rights.

We protect your rights with two steps: (1) copyright the software, and (2) offer you this license which gives you legal permission to copy, distribute and/or modify the software.

Also, for each author's protection and ours, we want to make certain that everyone understands that there is no warranty for this free software. If the software is modified by someone else and passed on, we want its recipients to know that what they have is not the original, so that any problems introduced by others will not reflect on the original authors' reputations.

Finally, any free program is threatened constantly by software patents. We wish to avoid the danger that redistributors of a free program will individually obtain patent licenses, in effect making the program proprietary. To prevent this, we have made it clear that any patent must be licensed for everyone's free use or not licensed at all.

The precise terms and conditions for copying, distribution and modification follow.

GNU GENERAL PUBLIC LICENSE TERMS AND CONDITIONS FOR COPYING, DISTRIBUTION AND MODIFICATION

0. This License applies to any program or other work which contains a notice placed by the copyright holder saying it may be distributed under the terms of this General Public License. The “Program”, below, refers to any such program or work, and a “work based on the Program” means either the Program or any derivative work under copyright law: that is to say, a work containing the Program or a portion of it, either verbatim or with modifications and/or translated into another language. (Hereinafter, translation is included without limitation in the term “modification”.) Each licensee is addressed as “you”.

Activities other than copying, distribution and modification are not covered by this License; they are outside its scope. The act of running the Program is not restricted, and the output from the Program is covered only if its contents constitute a work based on the Program (independent of having been made by running the Program). Whether that is true depends on what the Program does.

1. You may copy and distribute verbatim copies of the Program’s source code as you receive it, in any medium, provided that you conspicuously and appropriately publish on each copy an appropriate copyright notice and disclaimer of warranty; keep intact all the notices that refer to this License and to the absence of any warranty; and give any other recipients of the Program a copy of this License along with the Program.

You may charge a fee for the physical act of transferring a copy, and you may at your option offer warranty protection in exchange for a fee.

2. You may modify your copy or copies of the Program or any portion of it, thus forming a work based on the Program, and copy and distribute such modifications or work under the terms of Section 1 above, provided that you also meet all of these conditions:

a) You must cause the modified files to carry prominent notices stating that you changed the files and the date of any change.

b) You must cause any work that you distribute or publish, that in whole or in part contains or is derived from the Program or any part thereof, to be licensed as a whole at no charge to all third parties under the terms of this License.

c) If the modified program normally reads commands interactively when run, you must cause it, when started running for such interactive use in the most ordinary way, to print or display an announcement including an appropriate copyright notice and a notice that there is no warranty (or else, saying that you provide a warranty) and that users may redistribute the program under these conditions, and telling the user how to view a copy of this License. (Exception: if the Program itself is interactive but does not normally print such an announcement, your work based on the Program is not required to print an announcement.)

These requirements apply to the modified work as a whole. If identifiable sections of that work are not derived from the Program, and can be reasonably considered independent and separate works in themselves, then this License, and its terms, do not apply to those sections when you distribute them as separate works. But when you distribute the same sections as part of a whole which is a work based on the Program, the distribution of the whole must be on the terms of this License, whose permissions for other licensees extend to the entire whole, and thus to each and every part regardless of who wrote it.

Thus, it is not the intent of this section to claim rights or contest your rights to work written entirely by you; rather, the intent is to exercise the right to control the distribution of derivative or collective works based on the Program.

In addition, mere aggregation of another work not based on the Program with the Program (or with a work based on the Program) on a volume of a storage or distribution medium does not bring the other work under the scope of this License.

3. You may copy and distribute the Program (or a work based on it, under Section 2) in object code or executable form under the terms of Sections 1 and 2 above provided that you also do one of the following:

a) Accompany it with the complete corresponding machine-readable source code, which must be distributed under the terms of Sections 1 and 2 above on a medium customarily used for software interchange; or,

b) Accompany it with a written offer, valid for at least three years, to give any third party, for a charge no more than your cost of physically performing source distribution, a complete machine-readable copy of the corresponding source code, to be distributed under the terms of Sections 1 and 2 above on a medium customarily used for software interchange; or,

c) Accompany it with the information you received as to the offer to distribute corresponding source code. (This alternative is allowed only for noncommercial distribution and only if you received the program in object code or executable form with such an offer, in accord with Subsection b above.)

The source code for a work means the preferred form of the work for making modifications to it. For an executable work, complete source code means all the source code for all modules it contains, plus any associated interface definition files, plus the scripts used to control compilation and installation of the executable. However, as a special exception, the source code distributed need not include anything that is normally distributed (in either source or binary form) with the major components (compiler, kernel, and so on) of the operating system on which the executable runs, unless that component itself accompanies the executable.

If distribution of executable or object code is made by offering access to copy from a designated place, then offering equivalent access to copy the source code from the same place counts as distribution of the source code, even though third parties are not compelled to copy the source along with the object code.

4. You may not copy, modify, sublicense, or distribute the Program except as expressly provided under this License. Any attempt otherwise to copy, modify, sublicense or distribute the Program is void, and will automatically terminate your rights under this License. However, parties who have received copies, or rights, from you under this License will not have their licenses terminated so long as such parties remain in full compliance.

5. You are not required to accept this License, since you have not signed it. However, nothing else grants you permission to modify or distribute the Program or its derivative works. These actions are prohibited by law if you do not accept this License. Therefore, by modifying or distributing the Program (or any work based on the Program), you indicate your acceptance of this License to do so, and all its terms and conditions for copying, distributing or modifying the Program or works based on it.

6. Each time you redistribute the Program (or any work based on the Program), the recipient automatically receives a license from the original licensor to copy, distribute or modify the Program subject to these terms and conditions. You may not impose any further restrictions on the recipients' exercise of the rights granted herein. You are not responsible for enforcing compliance by third parties to this License.

7. If, as a consequence of a court judgment or allegation of patent infringement or for any other reason (not limited to patent issues), conditions are imposed on you (whether by court order, agreement or otherwise) that contradict the conditions of this License, they do not excuse you from the conditions of this License. If you cannot distribute so as to satisfy simultaneously your obligations under this License and any other pertinent obligations, then as a consequence you may not distribute the Program at all. For example, if a patent license would not permit royalty-free redistribution of the Program by all those who receive copies directly or indirectly through you, then the only way you could satisfy both it and this License would be to refrain entirely from distribution of the Program.

If any portion of this section is held invalid or unenforceable under any particular circumstance, the balance of the section is intended to apply and the section as a whole is intended to apply in other circumstances.

It is not the purpose of this section to induce you to infringe any patents or other property right claims or to contest validity of any such claims; this section has the sole purpose of protecting the integrity of the free software distribution system, which is implemented by public license practices. Many people have made generous contributions to the wide range of software distributed through that system in reliance on consistent application of that system; it is up to the author/donor to decide if he or she is willing to distribute software through any other system and a licensee cannot impose that choice.

This section is intended to make thoroughly clear what is believed to be a consequence of the rest of this License.

8. If the distribution and/or use of the Program is restricted in certain countries either by patents or by copyrighted interfaces, the original copyright holder who places the Program under this License may add an explicit geographical distribution limitation excluding those countries, so that distribution is permitted only in or among countries not thus excluded. In such case, this License incorporates the limitation as if written in the body of this License.

9. The Free Software Foundation may publish revised and/or new versions of the General Public License from time to time. Such new versions will be similar in spirit to the present version, but may differ in detail to address new problems or concerns.

Each version is given a distinguishing version number. If the Program specifies a version number of this License which applies to it and "any later version", you have the option of following the terms and conditions either of that version or of any later version published by the Free Software Foundation. If the Program does not specify a version number of this License, you may choose any version ever published by the Free Software Foundation.

10. If you wish to incorporate parts of the Program into other free programs whose distribution conditions are different, write to the author to ask for permission. For software which is copyrighted by the Free Software Foundation, write to the Free Software Foundation; we sometimes make exceptions for this. Our decision will be guided by the two goals of preserving the free status of all derivatives of our free software and of promoting the sharing and reuse of software generally.

NO WARRANTY

11. BECAUSE THE PROGRAM IS LICENSED FREE OF CHARGE, THERE IS NO WARRANTY FOR THE PROGRAM, TO THE EXTENT PERMITTED BY APPLICABLE LAW. EXCEPT WHEN OTHERWISE STATED IN WRITING THE COPYRIGHT HOLDERS AND/OR OTHER PARTIES PROVIDE THE PROGRAM "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESSED OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE. THE ENTIRE RISK AS TO THE QUALITY AND PERFORMANCE OF THE PROGRAM IS WITH YOU. SHOULD THE PROGRAM PROVE DEFECTIVE, YOU ASSUME THE COST OF ALL NECESSARY SERVICING, REPAIR OR CORRECTION.

12. IN NO EVENT UNLESS REQUIRED BY APPLICABLE LAW OR AGREED TO IN WRITING WILL ANY COPYRIGHT HOLDER, OR ANY OTHER PARTY WHO MAY MODIFY AND/OR REDISTRIBUTE THE PROGRAM AS PERMITTED ABOVE, BE LIABLE TO YOU FOR DAMAGES, INCLUDING ANY GENERAL, SPECIAL, INCIDENTAL OR CONSEQUENTIAL DAMAGES ARISING OUT OF THE USE OR INABILITY TO USE THE PROGRAM (INCLUDING BUT NOT LIMITED TO LOSS OF DATA OR DATA BEING RENDERED INACCURATE OR LOSSES SUSTAINED BY YOU OR THIRD PARTIES OR A FAILURE OF THE PROGRAM TO OPERATE WITH ANY OTHER PROGRAMS), EVEN IF SUCH HOLDER OR OTHER PARTY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.

END OF TERMS AND CONDITIONS

How to Apply These Terms to Your New Programs

If you develop a new program, and you want it to be of the greatest possible use to the public, the best way to achieve this is to make it free software which everyone can redistribute and change under these terms.

To do so, attach the following notices to the program. It is safest to attach them to the start of each source file to most effectively convey the exclusion of warranty; and each file should have at least the "copyright" line and a pointer to where the full notice is found.

one line to give the program's name and an idea of what it does. Copyright (C) yyyy name of author

```
This program is free software; you can redistribute it and/or
modify it under the terms of the GNU General Public License
as published by the Free Software Foundation; either version 2
of the License, or (at your option) any later version.


This program is distributed in the hope that it will be useful,
but WITHOUT ANY WARRANTY; without even the implied warranty of
MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE. See the
GNU General Public License for more details.


You should have received a copy of the GNU General Public License
along with this program; if not, write to the Free Software
Foundation, Inc., 59 Temple Place - Suite 330, Boston, MA 02111-1307, USA.
```

Also add information on how to contact you by electronic and paper mail.

If the program is interactive, make it output a short notice like this when it starts in an interactive mode:

```
Gnomovision version 69, Copyright (C) year name of author
Gnomovision comes with ABSOLUTELY NO WARRANTY; for details
type `show w'. This is free software, and you are welcome
to redistribute it under certain conditions; type `show c'
for details.
```

The hypothetical commands `show w' and `show c' should show the appropriate parts of the General Public License. Of course, the commands you use may be called something other than `show w' and `show c'; they could even be mouse-clicks or menu items--whatever suits your program.

You should also get your employer (if you work as a programmer) or your school, if any, to sign a "copyright disclaimer" for the program, if necessary. Here is a sample; alter the names:

```
Yoyodyne, Inc., hereby disclaims all copyright
interest in the program `Gnomovision'
(which makes passes at compilers) written
by James Hacker.


signature of Ty Coon, 1 April 1989
Ty Coon, President of Vice
```

This General Public License does not permit incorporating your program into proprietary programs. If your program is a subroutine library, you may consider it more useful to permit linking proprietary applications with the library. If this is what you want to do, use the GNU Lesser General Public License [http://www.fsf.org/licenses/lgpl.html] instead of this License.

GNU Free Documentation License

Version 1.2, November 2002

PREAMBLE

The purpose of this License is to make a manual, textbook, or other functional and useful document "free" in the sense of freedom: to assure everyone the effective freedom to copy and redistribute it, with or without modifying it, either commercially or noncommercially. Secondarily, this License preserves for the author and publisher a way to get credit for their work, while not being considered responsible for modifications made by others.

This License is a kind of "copyleft", which means that derivative works of the document must themselves be free in the same sense. It complements the GNU General Public License, which is a copyleft license designed for free software.

We have designed this License in order to use it for manuals for free software, because free software needs free documentation: a free program should come with manuals providing the same freedoms that the software does. But this License is not limited to software manuals; it can be used for any textual work, regardless of subject matter or whether it is published as a printed book. We recommend this License principally for works whose purpose is instruction or reference.

APPLICABILITY AND DEFINITIONS

This License applies to any manual or other work, in any medium, that contains a notice placed by the copyright holder saying it can be distributed under the terms of this License. Such a notice grants a world-wide, royalty-free license, unlimited in duration, to use that work under the conditions stated herein. The “Document”, below, refers to any such manual or work. Any member of the public is a licensee, and is addressed as “you”. You accept the license if you copy, modify or distribute the work in a way requiring permission under copyright law.

A “Modified Version” of the Document means any work containing the Document or a portion of it, either copied verbatim, or with modifications and/or translated into another language.

A “Secondary Section” is a named appendix or a front-matter section of the Document that deals exclusively with the relationship of the publishers or authors of the Document to the Document’s overall subject (or to related matters) and contains nothing that could fall directly within that overall subject. (Thus, if the Document is in part a textbook of mathematics, a Secondary Section may not explain any mathematics.) The relationship could be a matter of historical connection with the subject or with related matters, or of legal, commercial, philosophical, ethical or political position regarding them.

The “Invariant Sections” are certain Secondary Sections whose titles are designated, as being those of Invariant Sections, in the notice that says that the Document is released under this License. If a section does not fit the above definition of Secondary then it is not allowed to be designated as Invariant. The Document may contain zero Invariant Sections. If the Document does not identify any Invariant Sections then there are none.

The “Cover Texts” are certain short passages of text that are listed, as Front-Cover Texts or Back-Cover Texts, in the notice that says that the Document is released under this License. A Front-Cover Text may be at most 5 words, and a Back-Cover Text may be at most 25 words.

A “Transparent” copy of the Document means a machine-readable copy, represented in a format whose specification is available to the general public, that is suitable for revising the document straightforwardly with generic text editors or (for images composed of pixels) generic paint programs or (for drawings) some widely available drawing editor, and that is suitable for input to text formatters or for automatic translation to a variety of formats suitable for input to text formatters. A copy made in an otherwise Transparent file format whose markup, or absence of markup, has been arranged to thwart or discourage subsequent modification by readers is not Transparent. An image format is not Transparent if used for any substantial amount of text. A copy that is not “Transparent” is called “Opaque”.

Examples of suitable formats for Transparent copies include plain ASCII without markup, Texinfo input format, LaTeX input format, SGML or XML using a publicly available DTD, and standard-conforming simple HTML, PostScript or PDF designed for human modification. Examples of transparent image formats include PNG, XCF and JPG. Opaque formats include proprietary formats that can be read and edited only by proprietary word processors, SGML or XML for which the DTD and/or processing tools are not generally available, and the machine-generated HTML, PostScript or PDF produced by some word processors for output purposes only.

The “Title Page” means, for a printed book, the title page itself, plus such following pages as are needed to hold, legibly, the material this License requires to appear in the title page. For works in formats which do not have any title page as such, “Title Page” means the text near the most prominent appearance of the work’s title, preceding the beginning of the body of the text.

A section “Entitled XYZ” means a named subunit of the Document whose title either is precisely XYZ or contains XYZ in parentheses following text that translates XYZ in another language. (Here XYZ stands for a specific section name mentioned below, such as “Acknowledgements”, “Dedications”, “Endorsements”, or “History”.) To “Preserve the Title” of such a section when you modify the Document means that it remains a section “Entitled XYZ” according to this definition.

The Document may include Warranty Disclaimers next to the notice which states that this License applies to the Document. These Warranty Disclaimers are considered to be included by reference in this License, but only as regards disclaiming warranties: any other implication that these Warranty Disclaimers may have is void and has no effect on the meaning of this License.

VERBATIM COPYING

You may copy and distribute the Document in any medium, either commercially or noncommercially, provided that this License, the copyright notices, and the license notice saying this License applies to the Document are reproduced in all copies, and that you add no other conditions whatsoever to those of this License. You may not use technical measures to obstruct or control the reading or further copying of the copies you make or distribute. However, you may accept compensation in exchange for copies. If you distribute a large enough number of copies you must also follow the conditions in section 3.

You may also lend copies, under the same conditions stated above, and you may publicly display copies.

COPYING IN QUANTITY

If you publish printed copies (or copies in media that commonly have printed covers) of the Document, numbering more than 100, and the Document’s license notice requires Cover Texts, you must enclose the copies in covers that carry, clearly and legibly, all these Cover Texts: Front-Cover Texts on the front cover, and Back-Cover Texts on the back cover. Both covers must also clearly and legibly identify you as the publisher of these copies. The front cover must present the full title with all words of the title equally prominent and visible. You may add other material on the covers in addition. Copying with changes limited to the covers, as long as they preserve the title of the Document and satisfy these conditions, can be treated as verbatim copying in other respects.

If the required texts for either cover are too voluminous to fit legibly, you should put the first ones listed (as many as fit reasonably) on the actual cover, and continue the rest onto adjacent pages.

If you publish or distribute Opaque copies of the Document numbering more than 100, you must either include a machine-readable Transparent copy along with each Opaque copy, or state in or with each Opaque copy a computer-network location from which the general network-using public has access to download using public-standard network protocols a complete Transparent copy of the Document, free of added material. If you use the latter option, you must take reasonably prudent steps, when you begin distribution of Opaque copies in quantity, to ensure that this Transparent copy will remain thus accessible at the stated location until at least one year after the last time you distribute an Opaque copy (directly or through your agents or retailers) of that edition to the public.

It is requested, but not required, that you contact the authors of the Document well before redistributing any large number of copies, to give them a chance to provide you with an updated version of the Document.

MODIFICATIONS

You may copy and distribute a Modified Version of the Document under the conditions of sections 2 and 3 above, provided that you release the Modified Version under precisely this License, with the Modified Version filling the role of the Document, thus licensing distribution and modification of the Modified Version to whoever possesses a copy of it. In addition, you must do these things in the Modified Version:

A. Use in the Title Page (and on the covers, if any) a title distinct from that of the Document, and from those of previous versions (which should, if there were any, be listed in the History section of the Document). You may use the same title as a previous version if the original publisher of that version gives permission.

B. List on the Title Page, as authors, one or more persons or entities responsible for authorship of the modifications in the Modified Version, together with at least five of the principal authors of the Document (all of its principal authors, if it has fewer than five), unless they release you from this requirement.

C. State on the Title page the name of the publisher of the Modified Version, as the publisher.

D. Preserve all the copyright notices of the Document.

E. Add an appropriate copyright notice for your modifications adjacent to the other copyright notices.

F. Include, immediately after the copyright notices, a license notice giving the public permission to use the Modified Version under the terms of this License, in the form shown in the Addendum below.

G. Preserve in that license notice the full lists of Invariant Sections and required Cover Texts given in the Document's license notice.

H. Include an unaltered copy of this License.

I. Preserve the section Entitled "History", Preserve its Title, and add to it an item stating at least the title, year, new authors, and publisher of the Modified Version as given on the Title Page. If there is no section Entitled "History" in the Document, create one stating the title, year, authors, and publisher of the Document as given on its Title Page, then add an item describing the Modified Version as stated in the previous sentence.

J. Preserve the network location, if any, given in the Document for public access to a Transparent copy of the Document, and likewise the network locations given in the Document for previous versions it was based on. These may be placed in the "History" section. You may omit a network location for a work that was published at least four years before the Document itself, or if the original publisher of the version it refers to gives permission.

K. For any section Entitled "Acknowledgements" or "Dedications", Preserve the Title of the section, and preserve in the section all the substance and tone of each of the contributor acknowledgements and/or dedications given therein.

L. Preserve all the Invariant Sections of the Document, unaltered in their text and in their titles. Section numbers or the equivalent are not considered part of the section titles.

M. Delete any section Entitled "Endorsements". Such a section may not be included in the Modified Version.

N. Do not retitle any existing section to be Entitled "Endorsements" or to conflict in title with any Invariant Section.

O. Preserve any Warranty Disclaimers.

If the Modified Version includes new front-matter sections or appendices that qualify as Secondary Sections and contain no material copied from the Document, you may at your option designate some or all of these sections as invariant. To do this, add their titles to the list of Invariant Sections in the Modified Version's license notice. These titles must be distinct from any other section titles.

You may add a section Entitled "Endorsements", provided it contains nothing but endorsements of your Modified Version by various parties--for example, statements of peer review or that the text has been approved by an organization as the authoritative definition of a standard.

You may add a passage of up to five words as a Front-Cover Text, and a passage of up to 25 words as a Back-Cover Text, to the end of the list of Cover Texts in the Modified Version. Only one passage of Front-Cover Text and one of Back-Cover Text may be added by (or through arrangements made by) any one entity. If the Document already includes a cover text for the same cover, previously added by you or by arrangement made by the same entity you are acting on behalf of, you may not add another; but you may replace the old one, on explicit permission from the previous publisher that added the old one.

The author(s) and publisher(s) of the Document do not by this License give permission to use their names for publicity for or to assert or imply endorsement of any Modified Version.

COMBINING DOCUMENTS

You may combine the Document with other documents released under this License, under the terms defined in section 4 above for modified versions, provided that you include in the combination all of the Invariant Sections of all of the original documents, unmodified, and list them all as Invariant Sections of your combined work in its license notice, and that you preserve all their Warranty Disclaimers.

The combined work need only contain one copy of this License, and multiple identical Invariant Sections may be replaced with a single copy. If there are multiple Invariant Sections with the same name but different contents, make the title of each such section unique by adding at the end of it, in parentheses, the name of the original author or publisher of that section if known, or else a unique number. Make the same adjustment to the section titles in the list of Invariant Sections in the license notice of the combined work.

In the combination, you must combine any sections Entitled “History” in the various original documents, forming one section Entitled “History”; likewise combine any sections Entitled “Acknowledgements”, and any sections Entitled “Dedications”. You must delete all sections Entitled “Endorsements”.

COLLECTIONS OF DOCUMENTS

You may make a collection consisting of the Document and other documents released under this License, and replace the individual copies of this License in the various documents with a single copy that is included in the collection, provided that you follow the rules of this License for verbatim copying of each of the documents in all other respects.

You may extract a single document from such a collection, and distribute it individually under this License, provided you insert a copy of this License into the extracted document, and follow this License in all other respects regarding verbatim copying of that document.

AGGREGATION WITH INDEPENDENT WORKS

A compilation of the Document or its derivatives with other separate and independent documents or works, in or on a volume of a storage or distribution medium, is called an “aggregate” if the copyright resulting from the compilation is not used to limit the legal rights of the compilation’s users beyond what the individual works permit. When the Document is included in an aggregate, this License does not apply to the other works in the aggregate which are not themselves derivative works of the Document.

If the Cover Text requirement of section 3 is applicable to these copies of the Document, then if the Document is less than one half of the entire aggregate, the Document’s Cover Texts may be placed on covers that bracket the Document within the aggregate, or the electronic equivalent of covers if the Document is in electronic form. Otherwise they must appear on printed covers that bracket the whole aggregate.

TRANSLATION

Translation is considered a kind of modification, so you may distribute translations of the Document under the terms of section 4. Replacing Invariant Sections with translations requires special permission from their copyright holders, but you may include translations of some or all Invariant Sections in addition to the original versions of these Invariant Sections. You may include a translation of this License, and all the license notices in the Document, and any Warranty Disclaimers, provided that you also include the original English version of this License and the original versions of those notices and disclaimers. In case of a disagreement between the translation and the original version of this License or a notice or disclaimer, the original version will prevail.

If a section in the Document is Entitled “Acknowledgements”, “Dedications”, or “History”, the requirement (section 4) to Preserve its Title (section 1) will typically require changing the actual title.

TERMINATION

You may not copy, modify, sublicense, or distribute the Document except as expressly provided for under this License. Any other attempt to copy, modify, sublicense or distribute the Document is void, and will automatically terminate your rights under this License. However, parties who have received copies, or rights, from you under this License will not have their licenses terminated so long as such parties remain in full compliance.

FUTURE REVISIONS OF THIS LICENSE

The Free Software Foundation may publish new, revised versions of the GNU Free Documentation License from time to time. Such new versions will be similar in spirit to the present version, but may differ in detail to address new problems or concerns. See http://www.gnu.org/copyleft/.

Each version of the License is given a distinguishing version number. If the Document specifies that a particular numbered version of this License “or any later version” applies to it, you have the option of following the terms and conditions either of that specified version or of any later version that has been published (not as a draft) by the Free Software Foundation. If the Document does not specify a version number of this License, you may choose any version ever published (not as a draft) by the Free Software Foundation.

ADDENDUM: How to use this License for your documents

To use this License in a document you have written, include a copy of the License in the document and put the following copyright and license notices just after the title page:

```
Copyright (c) YEAR YOUR NAME.
```

```
Permission is granted to copy, distribute and/or modify this document
under the terms of the GNU Free Documentation License, Version 1.2
only as published by the Free Software Foundation;
with the Invariant Section being this copyright notice and license.
A copy of the license is included in the section entitled “GNU
Free Documentation License”.
```

If you have Invariant Sections, Front-Cover Texts and Back-Cover Texts, replace the “with...Texts.” line with this:

```
with the Invariant Sections being LIST THEIR TITLES, with the
Front-Cover Texts being LIST, and with the Back-Cover Texts being LIST.
```

If you have Invariant Sections without Cover Texts, or some other combination of the three, merge those two alternatives to suit the situation.

If your document contains nontrivial examples of program code, we recommend releasing these examples in parallel under your choice of free software license, such as the GNU General Public License, to permit their use in free software.

术语

主动/主动、主动/被动
: 一个有关服务在节点上如何运行的概念。主动-被动方案是指一个或多个服务在主动节点上运行，而被动节点等待主动节点出现故障。另一方面，主动-主动是指每个节点同时是主动的和被动的。

群集
: 一个高性能的群集是指一组共享应用程序负载以快速完成工作的计算机（真实或虚拟）。高可用性群集主要是为确保服务最高可用性而设计的。

群集分区
: 当一个或多个节点与群集的剩余节点之间的通讯失败时，即会发生群集分区。群集分区的节点仍是活动的且能够相互通讯，但它们无法感知不能与其通讯的节点。由于无法确认其他分区的丢失，所以开发了一种节点分裂方案（另请参见节点分裂 [271]）。

一致群集成员资格 (CCM)
: CCM 确定组成群集的节点并在群集中共享此信息。任何节点或仲裁人数的新增和丢失都由 CCM 提供。群集的每个节点上都运行 CCM 模块。

群集信息库 (CIB)
: 整个群集配置和状态（节点成员资格、资源、约束等等）的表示。它用XML编写，位于内存中。主 CIB 在 DC 上保存和维护并复制到其他节点。

群集资源管理器 (CRM)
: 负责协调所有非本地交互的主要管理实体。群集的每个节点都有自己的 CRM，但在 DC 上运行的那一个是选为将决策传播到其他非本地 CRM 并处理其输入的 CRM。CRM 会与许多组件交互：自己的节点和其他节点上的本地资源管理器、非本地 CRM、管理命令、屏障功能及成员资格层。

指定协调器 (DC)
: “主”节点。在此节点上保存着 CIB 的主副本。所有其他节点都从当前 DC 获取他们的配置和资源分配信息。DC 是在成员资格更改后从群集的所有节点中选出的。

分布式复制块设备 (drbd)
: DRBD 是为构建高可用性群集而设计的块设备。整个块设备通过专用网络镜像，且视作网络 RAID-1。

failover
: 指资源或节点在某台服务器上出现故障、受影响的资源在另一个节点上启动的情况。

屏障
: 描述了防止非群集成员访问共享资源的概念。通过终止（关闭）“有故障”的节点以防止其引起问题、使资源远离状态不确定的节点或多种其他方式均可以达到此目的。另外，节点屏障和资源屏障是有区别的。

Heartbeat 资源代理
: Heartbeat 第 1 版中广泛地使用了 Heartbeat 资源代理。第 2 版中已废弃对它们的使用，但仍然支持。Heartbeat 资源代理可以执行*启动*、*停止*和*状态*操作，它位于 `/etc/ha.d/resource.d` 或 `/etc/init.d` 下。有关 Heartbeat 资源代理的更多信息，请参见 http://www.linux-ha.org/HeartbeatResourceAgent。

本地资源管理器 (LRM)
: 本地资源管理器 (LRM) 负责对资源执行操作。它使用资源代理脚本执行工作。LRM 是“哑”的，它自己无法了解任何策略。它需要 DC 告诉它做什么。

LSB 资源代理
: LSB 资源代理是标准 LSB init 脚本。LSB init 脚本不仅用于高可用性环境中。任何兼容 LSB 的 Linux 系统使用 LSB init 脚本控制服务。任何 LSB 资源代理支持 `start`、`stop`、`restart`、`status` 和 `force-reload` 选项，并可能可选地提供 `try-restart` 和 `reload`。LSB 资源代理位于 `/etc/init.d`。在 http://www.linux-ha.org/LSBResourceAgent 和 http://www.linux-foundation.org/spec/refspecs/LSB_3.0.0/LSB-Core-generic/LSB-Core-generic/iniscrptact.html 可了解有关 LSB 资源代理和实际规范的更多信息。

节点
: 任何作为群集成员并对用户可见的计算机（真实或虚拟）。

pingd
: ping 守护程序。它使用 ICMP ping 持续联系一个或多个群集外的服务器。

策略引擎 (PE)
: 策略引擎计算要实现 CIB 中的策略更改而需要执行的操作。此信息随后传递到事务引擎，它在群集设置中依次实施策略更改。PE 始终在 DC 上运行。

OCF 资源代理
: OCF 资源代理类似于 LSB 资源代理（init 脚本）。任何 OCF 资源代理必须支持 `start`、`stop` 和 `status`（有时候称为 `monitor`）选项。另外，它支持以 XML 返回资源代理类型描述的*元数据*选项。它可能支持更多选项，但不是强制的。OCF 资源代理位于 `/usr/lib/ocf/resource.d/`*提供程序*。在 http://www.linux-ha.org/OCFResourceAgent 和 http://www.opencf.org/cgi-bin/viewcvs.cgi/specs/ra/resource-agent-api.txt?rev=HEAD 可了解有关 OCF 资源代理和规范草稿的更多信息。

仲裁人数
: 在群集中，如果群集分区具有多数节点（或投票），则它定义为具有仲裁人数（是“具有仲裁人数的”）。仲裁人数准确地区分了一个分区。它是算法的组成部分，用于防止多个断开的分区或节点继续运行而导致数据和服务损坏（节点分裂）。仲裁人数是屏障的先决条件，而屏障随后确保仲裁人数确实是唯一的。

资源
: Heartbeat 已知的任何类型的服务或应用程序。例如，IP 地址、文件系统或数据库。

资源代理 (RA)
: 资源代理 (RA) 是一种脚本，作为代理管理资源。有三种不同的资源代理：OCF（开放群集框架）资源代理、LSB 资源代理（标准 LSB init 脚本）和 Heartbeat 资源代理（Heartbeat v1 资源）。

单一故障点 (SPOF)
: 单一故障点 (SPOF) 是群集的任何如下的组件：如果它出现故障，则会触发整个群集的故障。

节点分裂
: 一种将群集节点分为两个或多个互不了解的组的方案（通过软件或硬件故障）。为防止节点分裂情况严重影响整个群集，必须靠 STONITH 来救援。也称为“分区的群集”方案。

STONITH
: “Shoot the other node in the head（关闭其他节点）”的首字母缩写，它关闭功能不正常的节点以防止其在群集中造成故障。

事务引擎 (TE)

事务引擎 (TE) 从 PE 取得策略指令并执行它们。TE 始终在 DC 上运行。它从那里指示其他节点上的本地资源管理器应执行的操作。