

Veritas Storage Foundation™ and High Availability Solutions 故障排除指南

AIX

6.0.1

Veritas Storage Foundation™ and High Availability Solutions 故障排除指南

本手册所述软件是根据许可协议而提供，仅可按该协议的条款使用。

产品版本：6.0.1

文档版本：6.0.1 Rev 0

法律声明

Copyright © 2012 Symantec Corporation. © 2012 Symantec Corporation 版权所有。All rights reserved. 保留所有权利。

Symantec、Symantec 徽标、Veritas、Veritas Storage Foundation、CommandCentral、NetBackup、Enterprise Vault 和 LiveUpdate 是 Symantec Corporation 或其附属公司在美国和其他国家/地区的商标或注册商标。“Symantec”和“赛门铁克”是 Symantec Corporation 在中国的注册商标。其他名称可能为其各自所有者的商标，特此声明。

本档中介绍的产品根据限制其使用、复制、分发和反编译/逆向工程的授权许可协议分发。未经 Symantec Corporation 及其特许人（如果存在）事先书面授权，不得以任何方式任何形式复制本文档的任何部分。

本档按“现状”提供，对于所有明示或暗示的条款、陈述和保证，包括任何适销性、针对特定用途的适用性或无侵害知识产权的暗示保证，均不提供任何担保，除非此类免责声明的范围在法律上视为无效。Symantec Corporation 不对任何与提供、执行或使用本文档相关的伴随或后果性损害负责。本文档所含信息如有更改，恕不另行通知。

根据 FAR 12.212 中的定义，授权许可的软件和文档被视为“商业计算机软件”，受 FAR 第 52.227-19 节“Commercial Computer Software - Restricted Rights”（商业计算机软件受限权利）和 DFARS 第 227.7202 节“Rights in Commercial Computer Software or Commercial Computer Software Documentation”（商业计算机软件或商业计算机软件文档权利）中的适用规定，以及所有后续法规中规定的权利的制约。美国政府仅可根据本协议的条款对授权许可的软件和文档进行使用、修改、发布复制、执行、显示或披露。

Symantec Corporation
350 Ellis Street
Mountain View, CA 94043
<http://www.symantec.com>

技术支持

Symantec 技术支持具有全球性支持中心。技术支持的主要任务是响应有关产品特性和功能的特定查询。技术支持小组还负责编写我们的联机知识库文章。技术支持小组与 Symantec 内的其他职能部门相互协作，及时解答您的问题。例如，技术支持小组与产品工程和 Symantec 安全响应中心协作，提供警报服务和病毒定义更新服务。

Symantec 提供的维护服务包括：

- 一系列支持服务，使您能为任何规模的单位选择适用的支持服务
- 通过电话和 Web 支持快速响应并提供最新信息
- 升级保证可保证软件顺利升级
- 全天候提供全球支持
- 高级功能，包括“客户管理服务”

有关 Symantec 维护计划的更多信息，请访问我们的网站：

<http://www.symantec.com/zh/cn/support/index.jsp>

与技术支持联系

具有有效维护协议的客户可以通过以下网址访问技术支持信息：

<http://www.symantec.com/zh/cn/support/index.jsp>

在联系技术支持之前，请确保您的计算机符合产品文档中所列的系统要求。而且您应当坐在发生问题的计算机旁边，以便需要时重现问题。

联系技术支持时，请准备好以下信息：

- 产品版本信息
- 硬件信息
- 可用内存、磁盘空间和 NIC 网卡信息
- 操作系统
- 版本和补丁程序级别
- 网络结构
- 路由器、网关和 IP 地址信息
- 问题说明：
 - 错误消息和日志文件
 - 联系 Symantec 之前执行过的故障排除操作

- 最近所做的软件配置更改和网络更改

授权许可与产品注册

如果您的 Symantec 产品需要注册或许可证密钥，请访问我们的技术支持网页：

<https://licensing.symantec.com/>

客户服务

可从以下网站获得客户服务信息：

<http://www.symantec.com/zh/cn/support/index.jsp>

客户服务可帮助您解决一些非技术性问题，例如以下几类问题：

- 有关产品许可或序列号的问题
- 产品注册更新（例如，更改地址或名称）
- 一般产品信息（功能、可用的语言、当地经销商）
- 有关产品更新和升级的最新信息
- 有关升级保障和维护合同的信息
- Symantec 采购计划的相关信息
- 有关 Symantec 技术支持选项的建议
- 非技术性的售前问题
- 与光盘或手册相关的问题

维护协议资源

如果想就现有维护协议事宜联络 Symantec，请通过以下方式联络您所在地区的维护协议管理部门：

国家/地区	销售热线	电子邮件
中国大陆	800 810 8826	China-Sales@symantec.com
中国台湾	0080 1611 391	Taiwan-Sales@symantec.com
中国香港特别行政区	800 963 421	HongKong-Sales@symantec.com

文档

介质中提供了 PDF 格式的产品指南。请确保您使用的是文档的最新版本。每个指南的第 2 页上提供了文档版本信息。从 Symantec 网站可以获取最新的产品文档。

<https://sort.symantec.com/documents>

您对产品文档的反馈对我们很重要。请发送改进建议和有关错误或疏漏的报告。请在您的报告中包括所报告的文本内容的文档标题和文档版本（位于第二页上）以及章节标题。请将反馈发送到：

doc_feedback@symantec.com

如需最新 HOWTO 文章、文档更新的信息，或询问有关产品文档的问题，请访问 Symantec Connect 上的 Storage and Clustering Documentation（存储和集群文档）论坛。

<https://www-secure.symantec.com/connect/storage-management/forums/storage-and-clustering-documentation>

关于 Symantec Connect

Symantec Connect 是为 Symantec 企业客户提供的点对点技术社区网站。参与者可以与其他产品用户联络并共享信息，包括发布论坛帖子、文章、视频、下载、博客和提出建议，并可与 Symantec 产品团队和技术支持进行交流。内容会由社区进行评分，成员可凭其贡献获得奖励积分。

<http://www.symantec.com/connect/storage-management>

目录

技术支持	4
第 1 章	简介 13
	关于 Veritas Storage Foundation and High Availability Solutions 产品故障排除 13
	关于 Symantec Operations Readiness Tools 13
	用于收集 VxExplorer 故障排除存档的增强功能 14
	关于 SORT 数据收集器 15
	关于唯一消息标识符 15
	使用 Symantec Operations Readiness Tools 查找唯一消息标识符说明和解决方案 15
部分 1	Veritas File System 故障排除 17
第 2 章	诊断消息 19
	文件系统对问题的响应 19
	恢复禁用的文件系统 20
	关于内核消息 20
部分 2	Veritas Volume Manager 故障排除 21
第 3 章	硬件故障后恢复 23
	关于硬件故障后恢复 23
	列出不可启动的卷 24
	显示卷和 Plex 状态 24
	Plex 状态循环 25
	恢复不可启动的镜像卷 28
	恢复不可启动的卷（其中禁用的 Plex 为 RECOVER 状态） 28
	强制重新启动禁用的卷 29
	清除磁盘上的 failing 标志 30
	重新挂接出错的磁盘 31
	在 Plex 挂接或同步操作失败后恢复 31
	RAID-5 卷上的故障 32

	系统故障	32
	磁盘故障	33
	RAID-5 的默认启动恢复过程	34
	恢复 RAID-5 卷	35
	在移动 RAID-5 子磁盘后恢复	37
	不可启动的 RAID-5 卷	38
	在不完整的磁盘组移动后恢复	40
	在 DCO 卷出现故障后恢复	40
	恢复 0 版 DCO 卷	42
	恢复即时快照 DCO 卷（版本 20 或更高版本）	45
第 4 章	在即时快照故障后恢复	47
	在 vxsnap prepare 失败后恢复	47
	在对完整空间即时快照执行 vxsnap make 失败后恢复	48
	在对断开即时快照执行 vxsnap make 失败后恢复	48
	在对优化空间即时快照执行 vxsnap make 失败后恢复	49
	在 vxsnap restore 失败后恢复	49
	在 vxsnap refresh 失败后恢复	49
	在“写入时复制”失败后恢复	50
	发生重新同步 I/O 错误后恢复	50
	DCO 卷发生 I/O 故障后恢复	51
	在对即时快照数据更改对象 (DCO) 执行 vxsnap upgrade 时出现的故障 中恢复	51
第 5 章	在引导磁盘故障后恢复	53
	VxVM 和引导磁盘故障	53
	通过重新安装来恢复	53
	一般重新安装信息	53
	重新安装系统并恢复 VxVM	54
第 6 章	管理命令和事务	63
	命令日志	63
	事务日志	65
	命令日志和事务日志的关联	66
	将从从属节点发出的 CVM 命令关联到主节点	67
	未启用命令完成	69
第 7 章	备份和还原磁盘组配置	71
	关于磁盘组配置备份	71
	备份磁盘组配置	72
	还原磁盘组配置	73

	解决磁盘组的冲突备份	75
第 8 章	对磁盘组导入问题进行故障排除	77
	清除非克隆磁盘的 <code>udid_mismatch</code> 标志	77
第 9 章	从 CDS 错误恢复	79
	CDS 错误代码和恢复操作	79
第 10 章	错误消息	83
	关于错误消息	83
	如何记录错误消息	83
	在启动脚本中配置日志记录	84
	消息类型	85
	消息	86
第 11 章	Veritas Volume Replicator 故障排除	87
	RLINK 连接出现问题后恢复	87
	发生配置错误后恢复	90
	RLINK 挂接期间的错误	90
	修改 RVG 期间出现错误	93
	主节点或辅助节点上的恢复	97
	关于从主节点主机崩溃中恢复	97
	主节点数据卷发生错误后恢复	97
	主节点 SRL 卷错误清除和重新启动	100
	重新启动时主节点 SRL 卷出错	101
	主节点 SRL 卷溢出恢复	101
	主节点 SRL 头错误清理和恢复	101
	辅助节点数据卷错误清除和恢复	102
	辅助节点 SRL 卷错误清除和恢复	104
	辅助节点 SRL 头错误清除和恢复	104
	重新启动时发生辅助节点 SRL 头错误	106
部分 3	Dynamic Multi-Pathing 故障排除	109
第 12 章	Dynamic Multi-Pathing 故障排除	111
	升级到 DMP 6.0 后显示扩展属性	111
	降级阵列支持	112

部分 4	Veritas Storage Foundation Cluster File System High Availability 故障排除	113
第 13 章	Veritas Storage Foundation Cluster File System High Availability 故障排除	115
	关于 Veritas Storage Foundation Cluster File System High Availability 故障排除	115
	CFS 故障排除	115
	root 用户的 <library> 路径中的顺序不正确	116
	非 root 用户运行 CFS 命令时 CFS 命令可能会挂起 (2403263)	116
	防护配置故障排除	117
	已有网络分裂 (裂脑) 的示例	117
	从已有网络分裂 (裂脑) 中恢复	117
	对 SFHA Solutions 集群中的 Cluster Volume Manager 进行故障排除	119
	将节点添加到 SFHA Solutions 集群后, CVM 组仍处于脱机状态	119
	在 SFHA Solutions 集群中无法导入共享磁盘组	120
	在 SFHA Solutions 集群中导入共享磁盘组出错	121
	无法在 SFHA Solutions 集群中启动 CVM	121
	即使 CVMCluster 在 SFHA Solutions 集群中处于联机状态, CVMVolDg 也不联机	121
部分 5	Veritas Cluster Server 故障排除	123
第 14 章	VCS 的故障排除与恢复	125
	VCS 消息日志记录	125
	GAB 消息日志记录	127
	为代理启用调试日志	127
	启用 IMF 的调试日志	128
	对 VCS 引擎启用调试日志	129
	关于调试日志标记用法	129
	收集 VCS 信息以进行支持分析	130
	收集 LLT 和 GAB 信息以进行支持分析	131
	收集 IMF 信息以进行支持分析	132
	消息目录	132
	VCS 引擎故障排除	133
	HAD 诊断	133
	HAD 连续重新启动	134
	DNS 配置问题导致 GAB 终止 HAD	134

种子设定和 I/O 防护	134
联机前 IP 检查	135
Low Latency Transport (LLT) 故障排除	135
LLT 启动脚本显示错误	135
LLT 检测到使用了交叉链路	136
LLT 链路状态消息	136
Group Membership Services/Atomic Broadcast (GAB) 故障排除	138
GAB 计时器问题	138
端口在重新打开时出现延迟	139
由于客户端进程故障，导致节点混乱	139
VCS 启动故障排除	140
VCS:10622 local configuration missing	140
VCS:10623 local configuration invalid	140
VCS:11032 registration failed.Exiting	140
Waiting for cluster membership.	141
智能监视框架 (IMF) 故障排除	141
疑难解答服务组	142
VCS 不会自动启动服务组	142
系统未处于 RUNNING 状态	143
服务组未配置为在该系统上运行	143
服务组未配置为自动启动	143
服务组已冻结	143
故障转移服务组在另一个系统上处于联机状态	143
某个关键资源出现故障	143
已自动禁用服务组	143
服务组正在等待资源联机/脱机	144
服务组正在等待某个依赖关系得到满足	144
未完全探查服务组	144
疑难解答资源	145
由于故障转移而使服务组联机	145
正在等待服务组状态	145
正在等待子资源	145
正在等待父资源	145
正在等待资源响应	145
代理未运行	146
磁盘组代理的 Monitor 入口点返回 ONLINE，即使磁盘组已被禁用也是如此	146
I/O 防护故障排除	147
正在逐出某个节点时，另一个节点不能加入集群	147
当 SCSI TEST UNIT READY 命令失败时，vxfsntsthdw 实用程序将失败	147
手动删除 SCSI-3 磁盘中的现有键	148
为防止潜在数据损坏，系统发生混乱	149

协调器磁盘的 I/O 防护键上的集群 ID 与本地集群的 ID 不匹 配	153
防护启动报告已有的裂脑	154
注册键会在协调器磁盘中丢失	156
集群脱机时替换损坏的磁盘	156
如果 rcp 命令或 scp 命令不能正常运行, vxfenswap 实用程序会 退出	158
CP 服务器故障排除	159
SFHA Solutions 集群节点上基于服务器的防护的故障排除	160
协调点联机迁移期间发生的问题	161
疑难解答通知	161
已配置通知程序, 但在 SNMP 控制台上看不到陷阱	161
全局集群的故障排除与恢复	162
灾难声明	162
丢失的心跳和查询机制	162
VCS 警报	163
对 steward 进程进行故障排除	164
授权许可故障排除	165
验证许可证密钥	165
授权错误消息	166
 索引	 169

简介

本章节包括下列主题：

- [关于 Veritas Storage Foundation and High Availability Solutions 产品故障排除](#)
- [关于 Symantec Operations Readiness Tools](#)
- [关于唯一消息标识符](#)

关于 Veritas Storage Foundation and High Availability Solutions 产品故障排除

本文档介绍在使用 Veritas Storage Foundation and High Availability Solutions 时可能会出现常见问题并提供针对这些问题的可能解决方案。除了本文档中的故障排除信息之外，还可以参见相应的 Symantec 产品“版本说明”文档以了解已知问题和软件限制。

关于 Symantec Operations Readiness Tools

[Symantec Operations Readiness Tools \(SORT\)](#) 是一个网站，可自动处理和简化某些最耗时的管理任务。SORT 有助于您更高效地管理数据中心，并充分利用 Symantec 产品。

SORT 可以帮助您执行以下操作：

- | | |
|--------------|---|
| 为下一次安装或升级做准备 | <ul style="list-style-type: none">■ 列出产品安装和升级要求，包括操作系统版本、内存、磁盘空间和体系结构。■ 分析系统以确定是否已做好安装或升级 Symantec 产品的准备。■ 从中央储存库下载最新的修补程序、文档和高可用性代理。■ 访问硬件、软件、数据库和操作系统的最新兼容性列表。 |
| 管理风险 | <ul style="list-style-type: none">■ 从中央储存库获取有关对修补程序、阵列特定模块 (ASL/APM/DDI/DDI) 和高可用性代理所做更改的自动电子邮件通知。■ 确定并降低系统和环境风险。■ 显示数百个 Symantec 错误代码的说明和解决方案。 |
| 提高效率 | <ul style="list-style-type: none">■ 根据产品版本和平台查找并下载修补程序。■ 列出已安装的 Symantec 产品和许可证密钥。■ 调整并优化您的环境。 |

注意： SORT 的某些功能并非对所有产品都可用。访问 SORT 不需要额外费用。

要访问 SORT，请转到：

<https://sort.symantec.com>

用于收集 VxExplorer 故障排除存档的增强功能

SORT 数据收集器包含收集和提交 VxExplorer 存档的功能。您可以将此存档发送给 Symantec 技术支持以便诊断并排除故障。VxExplorer 不收集客户数据。

旧 VxExplorer 脚本已废弃，因此未在 Veritas Storage Foundation and High Availability Solutions 6.0.1 版本中提供。

您可以使用 SORT 数据收集器选择性地执行下列操作：

- 使用基于操作系统的远程 Shell (RSH) 或安全 Shell (SSH) 远程功能从多台服务器收集 VxExplorer 数据。
- 将数据自动上传到 <ftp.veritas.com> 的 `/incoming` 目录。
- 确定指定服务器是否为集群的一部分，并将其他集群节点添加到数据收集/VxExplorer 进程（部分集群检查）。
- 自动升级数据收集器脚本。如果服务器没有 Internet 连接，则会自动禁用此功能。您还可以通过更新 `sortdc.conf` 配置文件来手动禁用自动升级。

关于 SORT 数据收集器

SORT 数据收集器包含在 VRTSpt 软件包中，并在安装 SFHA Solutions 时安装在系统上。

标准数据收集器功能包括下列功能：

- 采用可读的标准 Perl 脚本编写。
- 使用同一工具生成 SORT 自定义报告并收集 VxExplorer 存档。
- 数据收集器日志和 SFHA Solutions 安装日志具有相同格式。数据收集器日志存储在 ~/sort/log/ 目录中。

要了解有关使用数据收集器收集 VxExplorer 存档的详细信息，请参见：

<http://www.symantec.com/docs/HOWTO32575>

关于唯一消息标识符

Symantec 企业产品会显示唯一消息标识符 (UMI) 消息代码。UMI 包括错误、警告和信息性消息。如果您收到 UMI，可以使用 Symantec Operations Readiness Tools (SORT) 查找消息说明和解决方案。

请参见第 13 页的“关于 Symantec Operations Readiness Tools”。

请参见第 15 页的“使用 Symantec Operations Readiness Tools 查找唯一消息标识符说明和解决方案”。

使用 Symantec Operations Readiness Tools 查找唯一消息标识符说明和解决方案

您可以使用 Symantec Operations Readiness Tools (SORT) 查找唯一消息标识符 (UMI) 说明和解决方案。

查找唯一消息标识符说明和解决方案

- 1 将 Web 浏览器指向以下 URL：
<http://sort.symantec.com>
- 2 在任意 SORT 页的“search(搜索)”字段中，输入 UMI 代码，然后单击“search(搜索)”图标。

- 3 在“**Search Result(搜索结果)**”页的“**Error codes(错误代码)**”窗格中，单击指向您的消息代码的链接。如果有大量搜索结果，请使用页面顶部的复选框只显示可轻松查找您的代码的错误代码。

此时将显示 UMI 代码的“**Error Code details(错误代码详细信息)**”页，其中提供了说明以及任何可能的解决方案。

- 4 如果该页上的信息未提供适合您问题的解决方案，您可以单击页面上的任一链接，执行下列操作之一：
 - 对 UMI 或其解决方案进行评论。
 - 请求解决方案。
 - 添加您自己的解决方案。

部分

1

Veritas File System 故障排除

- 2. 诊断消息

诊断消息

本章节包括下列主题：

- [文件系统对问题的响应](#)
- [关于内核消息](#)

文件系统对问题的响应

当文件系统遇问题时，它将按下列方式之一进行响应：

标记错误的 <code>inode</code>	如果 <code>inode</code> 更新或目录块更新失败，则将 <code>Inode</code> 可以标记为出错。在这些类型的失败中，文件系统不知道磁盘上的具体信息，但会考虑它查到所有无效信息。在将 <code>inode</code> 标记为出错后，内核仍然允许对文件名的访问，但是任何访问文件中的数据或更改 <code>inode</code> 的尝试均会失败。
禁用事务处理	如果文件系统在写入意向日志时检测到错误，则它将禁用事务处理。禁用事务处理之后，仍可以读取或写入文件系统文件，但不允许任何块或 <code>inode</code> 释放或分配、结构更改、目录条目更改或其他对元数据的更改。
禁用文件系统	如果发生了错误进而损害了文件系统的完整性，则 <code>VxFS</code> 将禁用本身。如果意向日志失败或 <code>inode</code> 列表发生错误，超级块通常会更新（设置 <code>VX_FULLFSCK</code> 标志），以便下一 <code>fsck</code> 执行整个结构检查。如果此超级块更新失败，任何对文件系统的更改都可能造成不一致，而意向日志重放无法检查到这些不一致。要避免此情形，文件系统将禁用本身。

恢复禁用的文件系统

禁用文件系统后，不能向磁盘写入任何数据。尽管一些次要的文件系统仍可工作，但是大多数文件系统只返回EIO。禁用文件系统之后，唯一可做的事是执行 `umount` 命令和完全运行 `fsck`。

虽然日志重放可以生成一个干净的文件系统，但是执行整个结构检查会更安全。

文件系统通常会由于磁盘错误而遭到禁用。磁盘故障将会导致禁用文件系统，因此，应尽快将其修复。

请参见 `fsck_vxfs(1M)` 手册页。

执行整个结构检查

- ◆ 使用 `fsck` 命令执行整个结构检查：

```
# fsck -V vxfs -o full -y /dev/vx/rdisk/diskgroup/volume
```

警告：运行此命令时，应小心。通过指定 `-y` 选项，所有 `fsck` 用户提示均以 `yes` 回答，如果它执行了整个文件系统检查，则会造成不可撤消的更改。

关于内核消息

内核消息是 Veritas File System (VxFS) 内核生成的诊断性或错误消息。每个消息均有说明和如何处理或改正基本错误的建议。

当 VxFS 内核消息显示在系统控制台上时，它位于 `msgcnt` 字段显示的数字 ID 之后。此 ID 编号随着每个消息实例而增大，以保证在分析文件系统问题时事件的顺序是已知的。

Veritas Volume Manager 故障排除

- 3. 硬件故障后恢复
- 4. 在即时快照故障后恢复
- 5. 在引导磁盘故障后恢复
- 6. 管理命令和事务
- 7. 备份和还原磁盘组配置
- 8. 对磁盘组导入问题进行故障排除
- 9. 从 CDS 错误恢复
- 10. 错误消息
- 11. Veritas Volume Replicator 故障排除

硬件故障后恢复

本章节包括下列主题：

- 关于硬件故障后恢复
- 列出不可启动的卷
- 显示卷和 Plex 状态
- Plex 状态循环
- 恢复不可启动的镜像卷
- 恢复不可启动的卷（其中禁用的 Plex 为 RECOVER 状态）
- 强制重新启动禁用的卷
- 清除磁盘上的 failing 标志
- 重新挂接出错的磁盘
- 在 Plex 挂接或同步操作失败后恢复
- RAID-5 卷上的故障
- 在不完整的磁盘组移动后恢复
- 在 DCO 卷出现故障后恢复

关于硬件故障后恢复

Symantec 的 Veritas Volume Manager (VxVM) 可以保护系统免受磁盘和其他硬件故障的影响，并帮助您从这类事件中恢复。恢复过程可帮助您防止因磁盘和其他硬件故障导致数据丢失或无法访问系统。

如果一个卷发生磁盘 I/O 故障（例如，由于磁盘有无法纠正的错误），VxVM 可分离此故障所涉及的 Plex。I/O 在该 Plex 上停止，但在卷的其他 Plex 上继续。

如果磁盘彻底损坏，VxVM 可将该磁盘从所在的磁盘组中分离出来。该磁盘上的所有 Plex 都被禁用。如果分离时磁盘上有任何未镜像的卷，这些卷也被禁用。

注意：明显的磁盘故障可能不是由物理磁盘介质或磁盘控制器中的故障引起的，而可能是由中间组件或辅助组件（如电缆、主机总线适配器或电源）的故障引起的。

VxVM 中的热重定位功能自动检测磁盘故障，并用电子邮件将故障通知给系统管理员和其他指定用户。热重定位还尝试使用备用磁盘和空闲磁盘空间来还原冗余，并保持对镜像卷和 RAID-5 卷的访问能力。

有关管理热重定位的详细信息，请参见《Veritas Storage Foundation 管理指南》。

有关在 LVM 引导 (rootvg) 磁盘故障后恢复的信息，请参考 LVM 文档。

请参见第 53 页的“VxVM 和引导磁盘故障”。

列出不可启动的卷

不可启动的卷可能未正确配置或存在其他妨碍其启动的错误或情况。要显示不可启动的卷，请使用 `vxinfo` 命令。此命令用于显示卷的可访问性和可用性信息。

列出不可启动的卷

- ◆ 键入以下命令：

```
# vxinfo [-g diskgroup] [volume ...]
```

以下示例输出表明卷 `mkting` 不可启动：

```
home          fsgen         Started
mkting        fsgen         Unstartable
src           fsgen         Started
rootvol       root          Started
swapvol       swap          Started
```

显示卷和 Plex 状态

要显示有关卷配置（包括卷的状态和卷的 Plex 的状态）的详细信息，请使用 `vxprint` 命令。

显示卷和 Plex 状态

- ◆ 键入以下命令：

```
# vxprint [-g diskgroup] -hvt [volume ...]
```

以下示例显示了一个禁用的卷 `vol`，它有两个干净的 Plex（`vol-01` 和 `vol-02`），每个 Plex 都有一个子磁盘：

```
# vxprint -g mydg -hvt vol
```

```
Disk group: mydg
```

V	NAME	RVG/VSET/CO	KSTATE	STATE	LENGTH	READPOL	PREFPLEX	UTYPE
PL	NAME	VOLUME	KSTATE	STATE	LENGTH	LAYOUT	NCOL/WID	MODE
SD	NAME	PLEX	DISK	DISKOFFS	LENGTH	[COL/]OFF	DEVICE	MODE
SV	NAME	PLEX	VOLNAME	NVOLLAYR	LENGTH	[COL/]OFF	AM/NM	MODE
SC	NAME	PLEX	CACHE	DISKOFFS	LENGTH	[COL/]OFF	DEVICE	MODE
DC	NAME	PARENTVOL	LOGVOL					
SP	NAME	SNAPVOL	DCO					
v	vol	-	DISABLED	ACTIVE	212880	SELECT	-	fsgen
pl	vol-01	vol	DISABLED	CLEAN	212880	CONCAT	-	RW
sd	mydg11-01	vol-01	mydg11	0	212880	0	hdisk11	ENA
pl	vol-02	vol	DISABLED	CLEAN	212880	CONCAT	-	RW
sd	mydg12-01	vol-02	mydg12	0	212880	0	hdisk12	ENA

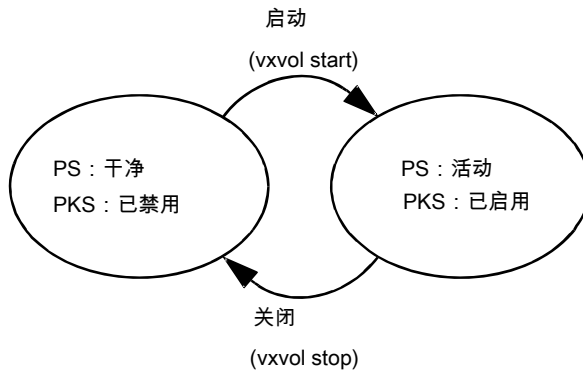
有关 Plex 和卷可能状态的说明，请参见《Veritas Storage Foundation 管理指南》。

Plex 状态循环

更改 Plex 状态是正常操作的一部分，并不一定表示出现了必须纠正的异常情况。清楚地了解各种 Plex 状态及其相互之间的关系对于正确执行任何恢复步骤是十分必要的。

图 3-1 显示了发生在 VxVM 中的 Plex 状态之间的主要转换。

图 3-1 主要的 Plex 状态循环



PS = plex state (Plex 状态)

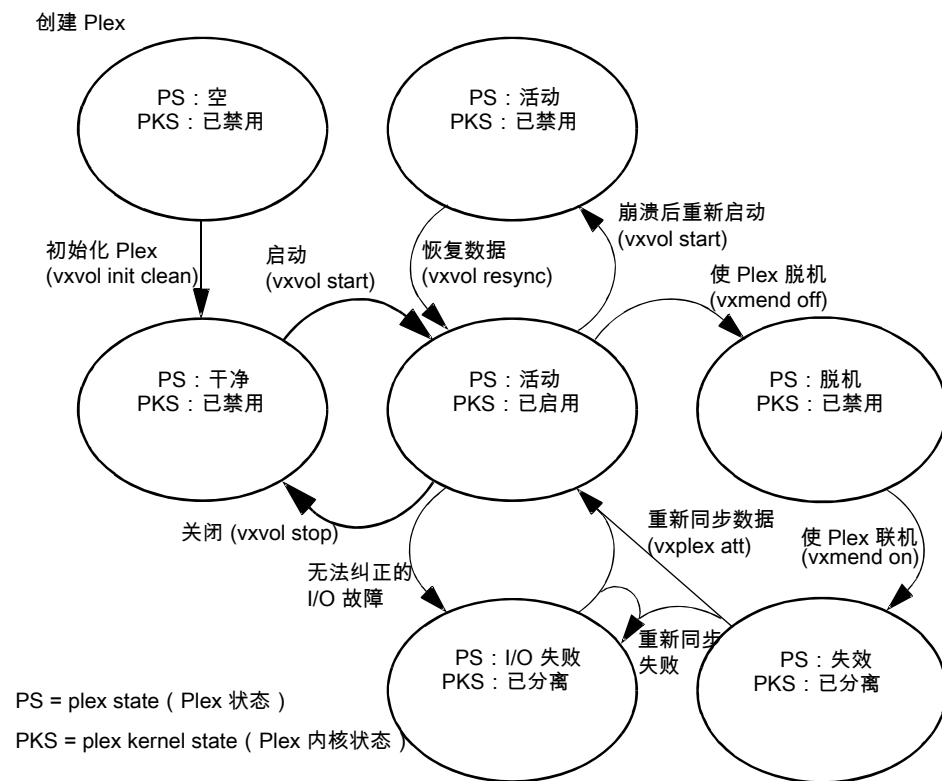
PKS = plex kernel state (Plex 内核状态)

有关 Plex 状态的详细信息，请参见《Veritas Storage Foundation 管理指南》。

系统启动时，卷自动启动并且 vxvol start 任务使所有 CLEAN Plex 变为 ACTIVE 状态。关机时，vxvol stop 任务将所有 ACTIVE Plex 标记为 CLEAN。如果启动时所有 Plex 为 CLEAN 状态，则表明曾发生了受控关机，系统将优化启动卷所需的时间。

图 3-2 显示了因硬件问题、非正常关机和系统管理员干预而可能导致的 Plex 状态之间的附加转换。

图 3-2 附加的 Plex 状态转换



Plex 刚创建时的状态为 EMPTY，且一直保持到挂接该 Plex 的卷被初始化为止。然后其状态被设置为 CLEAN。其 Plex 内核状态保持为 DISABLED，直到卷启动后才设置为 ENABLED。

在系统崩溃并重新启动后，卷的所有 Plex 均为 ACTIVE 状态，但 Plex 内核状态仍为 DISABLED，并且一直保持到 Plex 数据被 vxvol resync 任务恢复为止。

可以使用 vxmend off 命令使 Plex 脱机，使用 vxmend on 命令使其重新联机，并在使用 vxplex att 重新挂接 Plex 时使其数据与其他 Plex 重新同步。失败的重新同步或无法纠正的 I/O 故障会将 Plex 置为 IOFAIL 状态。

在系统崩溃或 I/O 错误导致镜像卷的所有 Plex 均退出 CLEAN 和 ACTIVE 状态时可以采取多种措施。

请参见第 28 页的“[恢复不可启动的镜像卷](#)”。

请参见第 32 页的“[RAID-5 卷上的故障](#)”。

恢复不可启动的镜像卷

系统崩溃或 I/O 错误会损坏镜像卷的一个或多个 Plex，并使所有 Plex 退出 CLEAN 或 ACTIVE 状态。可以将其中一个 Plex 标记为 CLEAN，并指示系统使用该 Plex 作为修复其他 Plex 的来源。

恢复不可启动的镜像卷

- 1 使用以下命令将所需的 Plex 置为 CLEAN 状态：

```
# vxmend [-g diskgroup] fix clean plex
```

例如，将 Plex vol101-02 置为 CLEAN 状态：

```
# vxmend -g mydg fix clean vol01-02
```

- 2 要从 CLEAN Plex 恢复卷中的其他 Plex，必须禁用该卷，并且其他 Plex 必须处于 STALE 状态。可根据需要对所有其他 CLEAN 或 ACTIVE Plex 依次运行以下命令，使其状态变为 STALE：

```
# vxmend [-g diskgroup] fix stale plex
```

承载卷的所有镜像 Plex 的几个磁盘或其他相关子系统发生严重的硬件故障后，用户可能无法使用 vxmend 恢复卷。在这种情况下，可以先删除该卷，在正常工作的硬件上重新创建它，然后从备份或快照映像中还原该卷的内容。

请参见 vxmend(1M) 手册页。

- 3 要启用 CLEAN Plex 并从中恢复 STALE Plex，请使用以下命令：

```
# vxvol [-g diskgroup] start volume
```

例如，要恢复卷 vol101，请使用以下命令：

```
# vxvol -g mydg start vol01
```

请参见 vxvol(1M) 手册页。

恢复不可启动的卷（其中禁用的 Plex 为 RECOVER 状态）

如果某个 Plex 的内容相对于卷来说已过时，则该 Plex 会显示为 RECOVER 状态。当包含 Plex 的一个或多个子磁盘的磁盘被替换或重新挂接时，会发生这种情况。如果某个 Plex 显示为这种状态，可通过 vxmend 和 vxvol 命令恢复：

恢复不可启动的卷（其中禁用的 Plex 为 RECOVER 状态）

- 1 使用以下命令将 Plex 强制设为 OFFLINE 状态：

```
# vxmend [-g diskgroup] -o force off plex
```

- 2 使用以下命令将 Plex 置于 STALE 状态：

```
# vxmend [-g diskgroup] on plex
```

- 3 如果卷中有其他 ACTIVE 或 CLEAN 状态的 Plex，请使用以下命令将 Plex 重新挂接到卷中：

```
# vxplex [-g diskgroup] att volume plex
```

如果卷已经启用，Plex 的重新同步将立即开始。

如果卷中没有其他 CLEAN 状态的 Plex，请使用以下命令将 Plex 设为 DISABLED 和 CLEAN 状态：

```
# vxmend [-g diskgroup] fix clean plex
```

- 4 如果尚未启用卷，请使用以下命令启动卷，并在后台执行所有的 Plex 重新同步：

```
# vxvol [-g diskgroup] -o bg start volume
```

当 Plex 中的数据已损坏，并且卷中没有可以从中重新同步卷内容的 ACTIVE 或 CLEAN 冗余 Plex 时，必须从备份或快照映像中还原卷。

强制重新启动禁用的卷

如果由于磁盘故障导致某个卷被禁用，并且该卷中未包含任何有效的冗余 Plex，则必须在更换出错的磁盘后从备份还原该卷。对于所有被列为 Unstartable 的卷，必须先使用 vxvol 命令重新启动这些卷，然后才能从备份中还原其内容。

强制重新启动禁用的卷

- ◆ 键入以下命令：

```
# vxvol [-g diskgroup] -o bg -f start volume
```

-f 选项将强制重新启动卷，而 -o bg 选项则作为后台任务重新同步其 Plex。例如，要重新启动卷 myvol，以便能够从备份中还原它，可使用以下命令：

```
# vxvol -g mydg -o bg -f start myvol
```

清除磁盘上的 failing 标志

如果 I/O 错误是间歇性的而非持久性的，则 Veritas Volume Manager 会在磁盘上设置 failing 标志，而不是分离磁盘。临时拔掉电缆、控制器出错、磁盘阵列中有部分出错的 LUN 或者磁盘中有少量损坏的扇区或磁道，都会引发此类错误。

如果硬件错误不是由磁盘本身造成的（例如，是由控制器出问题或磁盘电缆出问题造成的），则可以在从根本上纠正 I/O 错误之后，使用 vxedit 命令清除 failing 标志。

警告： 如果不知道引发 I/O 错误的原因，请不要清除 failing 标志。如果磁盘硬件确实出现了问题，而您又清除了 failing 标志，则会有丢失数据的危险。

清除磁盘上的 failing 标志

- 1 使用 vxdisk list 命令找出出错的磁盘：

```
# vxdisk list
```

```
DEVICE      TYPE          DISK      GROUP    STATUS
  hdisk10    auto:simple   mydg01    mydg     online
  hdisk11    auto:simple   mydg02    mydg     online failing
  hdisk12    auto:simple   mydg03    mydg     online
  . . .
```

- 2 使用 vxedit set 命令为每个被标记为 failing 的磁盘（在本例中为 mydg02）清除标志：

```
# vxedit -g mydg set failing=off mydg02
```

- 3 使用 vxdisk list 命令确认已清除 failing 标志：

```
# vxdisk list
```

```
DEVICE      TYPE          DISK      GROUP    STATUS
  hdisk10    auto:simple   mydg01    mydg     online
  hdisk11    auto:simple   mydg02    mydg     online
  hdisk12    auto:simple   mydg03    mydg     online
  . . .
```

重新挂接出错的磁盘

如果系统在启动时找不到某个磁盘，或者在 VxVM 启动时某些磁盘驱动程序被卸载或无法装载（从而导致磁盘进入 `failed` 状态），则可以执行重新挂接操作。解决根本问题（如电缆或控制器错误）之后，使用 `vxreattach` 命令重新挂接磁盘，重新挂接之后 Plex 不会被标记为 `STALE`。但是，重新挂接操作必须在启动磁盘上的卷之前进行。

作为磁盘恢复过程的一部分，可以在引导期间从 `vxdiskadm` 菜单中调用 `vxreattach` 命令。如果可能，`vxreattach` 将出错的磁盘介质记录重新挂接到具有同一设备名称的磁盘上。重新挂接操作会将磁盘放在它以前所属的同一磁盘组中，并保留其原有磁盘介质名。

重新挂接出错的磁盘

- 1 如下例所示，使用 `vxdisk list` 命令查看出错的磁盘：

```
# vxdisk list
DEVICE      TYPE          DISK      GROUP    STATUS
hdisk10    auto:simple  mydg01   mydg     online
hdisk11    auto:simple  mydg02   mydg     online
-          -            mydg03   mydg     failed was: hdisk12
-          -            mydg04   mydg     failed was: hdisk13
```

- 2 一旦纠正错误，便可使用如下命令重新扫描设备列表，进而重新挂接磁盘：

```
# /usr/sbin/vxdctl enable
```

- 3 使用不带任何选项的 `vxreattach` 命令重新挂接磁盘：

```
# /etc/vx/bin/vxreattach
```

执行重新挂接操作之后，除非磁盘出错且必须替换，否则不必进行恢复。如果最初的磁盘故障起因（或其他原因）仍然存在，则重新挂接操作可能失败。

可以不执行重新挂接操作，而是先使用 `vxreattach -c` 命令来检查该操作是否可行。该命令显示可以重新挂接磁盘的磁盘组和磁盘介质名称。

请参见 `vxreattach(1M)` 手册页。

在 Plex 挂接或同步操作失败后恢复

Plex 挂接操作要求 Plex 与卷中的现有 Plex 保持同步。其他操作（例如为卷创建镜像）也要求 Plex 同步。Plex 同步可能是长时间运行的操作，具体取决于卷的大小以及需要同步的数据量。

此版本中的恢复行为已在以前版本的 VxVM 基础上进行了更改。过去，如果 Plex 同步中断，则在卷恢复期间 VxVM 会从卷中分离 Plex。这时需要手动重新开始 Plex 同步。如果磁盘组版本低于 170，则会保留此行为。

如果磁盘组版本为 170 或 180 并由 `vxplex att`、`vxassist mirror`、`vxsnap addmir` 或 `vxsnap reattach` 命令触发 Plex 重新同步，则在卷恢复期间，Plex 会保持与卷关联。VxVM 会检测到同步已中断，然后恢复同步。如果卷具有关联的 DCO（20 版本或更高版本），则 VxVM 会跟踪 Plex 同步期间所发生的更改；当同步因系统崩溃或 `vxconfigd` 故障而失败时，VxVM 会从失败所在点恢复同步。同步在后台执行，因此使用卷时没有延迟。如果卷没有关联的 DCO，但有 170 或更高版本的磁盘组，则同步会从头重新开始。

在创建卷并通过一个操作 (`vxassist make nmirror=2`) 添加镜像时，不会为自动恢复跟踪该镜像的同步。要确保同步从它失败所在点恢复，请首先创建卷，然后使用 `vxassist mirror` 命令添加镜像。

在某些情况下，VxVM 可自动恢复卷。如果您需要手动恢复卷，则 `vxrecover` 命令会为同步过程失败的所有 Plex 触发同步。这些 Plex 的状态为 TEMP、TEMPRM 或 TEMPMSD。

在 CVM 环境中，如果在执行 Plex 同步期间主节点崩溃，则在恢复主节点后新的主节点会从主节点失败所在点重新开始同步。磁盘组版本必须为 170 版或更高版本。如果磁盘组为 170 版本，但未挂接 DCO，则同步会从头重新开始。

可以使用 `Ctrl-C` 或 `vxtask abort` 命令中止 Plex 挂接操作或同步操作。在这种情况下，VxVM 会将 Plex 从卷中分离。

RAID-5 卷上的故障

故障有两种：系统故障与磁盘故障。系统故障是指因操作系统混乱或电源故障而使系统突然停止运行。磁盘故障是指因系统故障（如磁头划盘、磁盘上的电子元件故障或磁盘控制器故障）而使一些磁盘上的数据不可用。

系统故障

RAID-5 卷的设计使其能够在磁盘发生故障时以最低的磁盘空间消耗来保持可用性。但是，在发生系统故障后，许多 RAID-5 形式的卷都会有数据丢失。导致数据丢失的原因是系统故障致使 RAID-5 卷中的数据和奇偶校验不再同步。而失去同步则是因故障发生时无法确定未完成的写操作的状态所致。

如果在访问 RAID-5 卷时失去同步，则将该卷描述为具有失效的奇偶校验。此时必须通过以下方法重新构建奇偶校验：读取每个条带中的所有非奇偶列，重新计算奇偶校验，然后写出条带中的奇偶校验条带单元。由于必须对卷中的每个条带执行此操作，因此这需要很长时间才能完成。

警告：当对没有日志 Plex 的 RAID-5 卷进行重新同步时，卷中磁盘的任何故障都会导致其数据丢失。

重同步过程除易出故障之外，还会占用系统资源并使系统运行速度变慢。

RAID-5 日志保留故障发生时所写入数据的副本，因而可减少因系统故障导致的损失。重同步过程包括从日志中读取数据及奇偶校验，然后将它们写入 RAID-5 卷的相应区域。这可以大大缩短数据和奇偶校验重同步所需的时间。它还意味着卷永远不会真正失效。任何时候卷中所有条带的数据和奇偶校验均为已知，因此，单个磁盘的故障不会导致卷中的数据丢失。

磁盘故障

当磁盘故障、电缆连接或其他问题导致磁盘上的数据不可用时，将出现无法纠正的 I/O 错误。对于 RAID-5 卷，则意味着有一个子磁盘不可用。该子磁盘不能用来存储数据，并且被视为失效且已分离。即使基础磁盘变为可用或被更换，该子磁盘仍被视为失效并且不被使用。

当系统尝试读取失效子磁盘上的数据时，它将通过条带中所有其他条带单元上的数据重新构建这些数据。此操作称为重构读取。与单纯读取数据相比，这是一个开销很大的操作，会导致读取性能下降。当 RAID-5 卷包含失效的子磁盘时，它被视为处于退化模式。

处于退化模式的 RAID-5 卷可从 `vxprint -ht` 命令的输出中识别出来，如下所示：

```

V  NAME          RVG/VSET/COKSTATE STATE    LENGTH  READPOL  PREFPLEX  UTYPE
PL NAME          VOLUME    KSTATE  STATE    LENGTH  LAYOUT   NCOL/WID  MODE
SD NAME          PLEX      DISK    DISKOFFS LENGTH  [COL/]OFF DEVICE    MODE
SV NAME          PLEX      VOLNAME NVOLLAYR LENGTH  [COL/]OFF AM/NM     MODE
...
v   r5vol        -          ENABLED  DEGRADED 204800  RAID     -          raid5
pl  r5vol-01     r5vol     ENABLED  ACTIVE   204800  RAID     3/16      RW
sd  disk01-01   r5vol-01disk01 0        102400  0/0     hdisk3    ENA
sd  disk02-01   r5vol-01disk02 0        102400  1/0     hdisk4    dS
sd  disk03-01   r5vol-01disk03 0        102400  2/0     hdisk5    ENA
pl  r5vol-02     r5vol     ENABLED  LOG      1440    CONCAT   -          RW
sd  disk04-01   r5vol-02disk04 0        1440    0        hdisk6    ENA
pl  r5vol-03     r5vol     ENABLED  LOG      1440    CONCAT   -          RW
sd  disk05-01   r5vol-03disk05 0        1440    0        hdisk7    ENA

```

如卷状态所示，卷 `r5vol` 处于退化模式，它被标为 `DEGRADED`。出错的子磁盘是 `disk02-01`，如 `MODE` 标志所示；`d` 表示子磁盘已分离，`s` 表示子磁盘的内容失效。

警告：不要对处于退化模式的 RAID-5 卷运行 `vxr5check` 命令。

包含 RAID-5 日志 Plex 的磁盘也可能会发生故障。如果 RAID-5 日志已镜像，则单个 RAID-5 日志 Plex 出故障对卷的操作没有直接影响。但是，卷中所有 RAID-5 日志 Plex 丢失会使卷容易彻底失败。在 `vxprint -ht` 命令的输出中，有故障的 RAID-5 日志 Plex 是通过显示为 BADLOG（而不是 LOG）的 Plex 状态来指示。

在以下示例中，RAID-5 日志 Plex `r5vol-02` 失败：

```
V  NAME          RVG/VSET/COKSTATE  STATE  LENGTH  READPOL  PREFPLEX  UTYPE
PL  NAME          VOLUME    KSTATE  STATE  LENGTH  LAYOUT    NCOL/WID  MODE
SD  NAME          PLEX      DISK    DISKOFFS  LENGTH  [COL/]OFF  DEVICE    MODE
SV  NAME          PLEX      VOLNAME  NVOLLAYR  LENGTH  [COL/]OFF  AM/NM     MODE
...
v   r5vol         -          ENABLED  ACTIVE  204800  RAID       -          raid5
pl  r5vol-01      r5vol     ENABLED  ACTIVE  204800  RAID       3/16      RW
sd  disk01-01    r5vol-01disk01  0       102400  0/0       hdisk3     ENA
sd  disk02-01    r5vol-01disk02  0       102400  1/0       hdisk4     ENA
sd  disk03-01    r5vol-01disk03  0       102400  2/0       hdisk5     ENA
pl  r5vol-02      r5vol     DISABLED  BADLOG  1440     CONCAT     -          RW
sd  disk04-01    r5vol-02disk04  0       1440    0         hdisk6     ENA
pl  r5vol-03      r5vol     ENABLED  LOG     1440     CONCAT     -          RW
sd  disk05-01    r5vol-12disk05  0       1440    0         hdisk7     ENA
```

RAID-5 的默认启动恢复过程

VxVM 可能需要执行若干操作才能完全还原 RAID-5 卷的内容并使它可用。无论何时启动卷，所有 RAID-5 日志 Plex 均在卷启动之前被设置为零。这样可以防止随机数据被解释为日志项并破坏卷的内容。此外，还可能需恢复某些子磁盘，或可能需要重同步奇偶校验（如果 RAID-5 日志出错）。

启动 RAID-5 卷时，VxVM 执行下列步骤：

- 如果 RAID-5 卷未正常关闭，则检查它是否存在有效的 RAID-5 日志 Plex。
- 如果存在有效的日志 Plex，则重放它们。重放过程是通过将卷内核状态置为 DETACHED、将卷状态设置为 REPLAY 并启用 RAID-5 日志 Plex 来完成的。
- 如果没有有效的日志，则必须重同步奇偶校验。重新同步是通过将卷内核状态置为 DETACHED 并将卷状态设置为 SYNC 来完成的。所有的日志 Plex 都保持 DISABLED Plex 内核状态。

由于重同步奇偶校验过程中任何子磁盘失败都会导致卷无法使用，因此重同步期间卷被设为不可用。通过将 `-o unsafe start` 选项和 `vxvol` 命令一起使用可改写此设置。如果存在任何失效的子磁盘，则 RAID-5 卷无法使用。

警告： `-o unsafe start` 选项被视为是有害选项，因为它会使卷的内容不可用。因此建议不要使用该选项。

- 将现有的所有日志 Plex 置零并启用这些日志。如果此过程中所有日志都失败，则启动进程中止。
- 如果不存在失效子磁盘或现有的失效子磁盘可以恢复，则卷被置为 ENABLED 卷内核状态，卷状态被设置为 ACTIVE。至此，卷已启动。

恢复 RAID-5 卷

对于 RAID-5 卷，可能需要执行以下类型的恢复：

- 重新同步奇偶校验
- 重新挂接出错的 RAID-5 日志 Plex
- 恢复失效子磁盘

通常在 RAID-5 卷启动时或系统引导之后不久执行奇偶校验重同步和失效子磁盘恢复。也可通过运行 `vxrecover` 命令来执行它们。

请参见第 38 页的“不可启动的 RAID-5 卷”。

如果磁盘发生故障时热重定位功能是激活的，除非没有合适的磁盘空间用于重定位，否则不需要系统管理员干预。热重定位功能通过故障触发，并且系统管理员会收到有关故障的电子邮件通知。

热重定位操作自动尝试重定位出错的 RAID-5 Plex 的子磁盘。发生重定位后，热重定位后台驻留程序 (`vxrelocd`) 还会启动奇偶校验重同步。

如果 RAID-5 日志 Plex 失败，则仅当该日志 Plex 已被镜像时才发生重定位；`vxrelocd` 后台驻留程序此时启动镜像重同步来重新创建 RAID-5 日志 Plex。如果在故障发生时热重定位功能被禁用，则需要系统管理员来启动重同步或恢复操作。

注意： 承载 RAID-5 Plex 的几个磁盘或其他相关子系统发生严重的硬件故障后，可能只能通过删除该卷、在正常工作的硬件上重新创建，然后从备份中还原该卷的内容才能还原该卷。

重新同步 RAID-5 卷中的奇偶校验

大多数情况下，RAID-5 卷没有失效的奇偶校验。仅当 RAID-5 卷的所有 RAID-5 日志 Plex 都失败并且存在系统故障时，才出现失效的奇偶校验。即使 RAID-5 卷有失效的奇偶校验，通常也会在卷启动过程中修复它。

如果启动了没有有效的 RAID-5 日志的卷并且该进程在卷重新同步之前停止，则会产生带失效的奇偶校验的活动卷。

以下是 `vxprint -ht` 命令对于失效的 RAID-5 卷的示例输出：

V	NAME	RVG/VSET/COKSTATE	STATE	LENGTH	READPOL	PREFPLEX	UTYPE	
PL	NAME	VOLUME	KSTATE	STATE	LENGTH	LAYOUT	NCOL/WID	MODE

```

SD  NAME          PLEX          DISK          DISKOFFS     LENGTH      [COL/]OFF    DEVICE        MODE
SV  NAME          PLEX          VOLNAME       NVOLLAJR     LENGTH      [COL/]OFF    AM/NM         MODE
...
v   r5vol         -             ENABLED       NEEDSYNCS    204800      RAID         -             raid5
pl  r5vol-01      r5vol        ENABLED       ACTIVE        204800      RAID         3/16         RW
sd  disk01-01    r5vol-01     disk01        0             102400      0/0          hdisk3       ENA
sd  disk02-01    r5vol-01     disk02        0             102400      1/0          hdisk4       dS
sd  disk03-01    r5vol-01     disk03        0             102400      2/0          hdisk5       ENA
...

```

此输出将卷状态显示为 `NEEDSYNCS`，表明需要重新同步奇偶校验。卷状态也可能是 `SYNCS`，表明启动时尝试进行同步，并且同步进程应正在执行同步。如果不存在这样的进程，或者卷处于 `NEEDSYNCS` 状态，则可以通过对 `vxvol` 命令使用 `resync` 关键字来手动启动同步操作。

通过向 RAID-5 卷发出 `VOL_R5_RESYNCS ioctl`s 可重新生成奇偶校验。重新同步过程从 RAID-5 卷的起始处开始，对大小等于 `-o iosize` 选项指定的扇区数的区域进行重新同步。如果未指定 `-o iosize` 选项，则使用默认的最大 I/O 大小。然后，`resync` 操作移到下一个区域进行，直到整个 RAID-5 卷长都被重新同步为止。

对于较大的卷，奇偶校验重建需要很长时间，在操作完成前有可能发生系统关机或崩溃。如果系统关机，奇偶校验重建过程必须在系统重新启动后继续进行。否则，该过程必须从头开始。

为了避免从头再来，奇偶校验重建中将被设置检查点。这意味着奇偶校验重建过程所到达的偏移位置将被存储在配置数据库中。`-o checkpoint=size` 选项控制保存检查点的频率。如果未指定该选项，则使用默认的检查点大小。

由于保存检查点偏移要求进行事务处理，因此检查点太小会延长奇偶校验重建所需的时间。在重新启动系统后，检查点偏移小于卷长的 RAID-5 卷会在检查点偏移处启动奇偶校验重新同步。

重新同步 RAID-5 卷中的奇偶校验

◆ 键入以下命令：

```
# vxvol -g diskgroup resync r5vol
```

重新挂接出错的 RAID-5 日志 Plex

磁盘故障可能会使 RAID-5 日志 Plex 分离。在 `vxplex` 命令中使用 `att` 关键字可以重新挂接这些 RAID-5 日志。

重新挂接出错的 RAID-5 日志 Plex

- ◆ 键入以下命令：

```
# vxplex -g diskgroup att r5vol r5vol-plex
```

恢复 RAID-5 卷中的失效子磁盘

失效子磁盘的恢复通常在卷启动时完成。但是，执行恢复的进程可能会崩溃，或者卷可能是通过 `-o delayrecover` 这样的妨碍子磁盘恢复的选项启动的。此外，还有可能未执行恢复操作就替换了子磁盘所在的磁盘。在这些情况下，可以使用 `vxvol recover` 命令执行子磁盘恢复。

恢复 RAID-5 卷中的失效子磁盘

- ◆ 键入以下命令：

```
# vxvol -g diskgroup recover r5vol subdisk
```

具有多个失效子磁盘的 RAID-5 卷只需一个操作即可恢复。要恢复多个失效子磁盘，请对卷使用 `vxvol recover` 命令：

```
# vxvol -g diskgroup recover r5vol
```

在移动 RAID-5 子磁盘后恢复

当移动或更换了 RAID-5 子磁盘时，新的子磁盘被标记为 `STALE` 以待恢复。如果卷是活动的，则可以使用 `vxsd` 命令对卷进行恢复。如果卷处于非活动状态，则当它下次启动时会进行恢复。恢复操作期间的 RAID-5 卷是退化的。

移动操作所涉及的条带中的任何故障都会使卷不可用。如果奇偶校验失效，则 RAID-5 卷也会变为无效。

为避免卷不可用，在下列情况下，`vxsd` 命令不允许移动子磁盘：

- 失效子磁盘占用了正在移动的子磁盘所用的任何条带。
- RAID-5 卷停止但未正常关闭；即，奇偶校验被视为失效。
- RAID-5 卷活动并且没有有效的日志区域。

只有第三种情况可通过使用 `-o force` 选项来改写。

RAID-5 卷的子磁盘还可以使用 `vxsd split` 命令和 `vxsd join` 命令拆分或结合。这些操作的工作方式与在镜像卷上相同。

RAID-5 子磁盘移动和其他卷类型的子磁盘移动的执行方式相同，但没有冗余退化的缺点。

不可启动的 RAID-5 卷

在 RAID-5 卷启动时，它的状态可以是许多状态中的一种。在系统正常关闭后，卷应是干净的并且不需要恢复。但是，如果卷没有关闭，或者在崩溃之前没有卸载，则在启动 RAID-5 卷并使它可用之前会要求进行恢复。

正常条件下，在重新启动后和自动或通过 `vxrecover` 命令手动执行任何恢复后，卷都自动启动。

如果 RAID-5 Plex 的某些部分未映射 RAID-5 卷长度，则该卷在下列情况下不可用：

- 与 RAID-5 卷长度相比，RAID-5 Plex 是稀疏 Plex。
- RAID-5 Plex 未映射一个条带中有两个子磁盘出现故障的区域，故障原因是子磁盘失效，或者是子磁盘建立在出现故障的磁盘上。

发生这种情况时，`vxvol start` 命令返回以下错误消息：

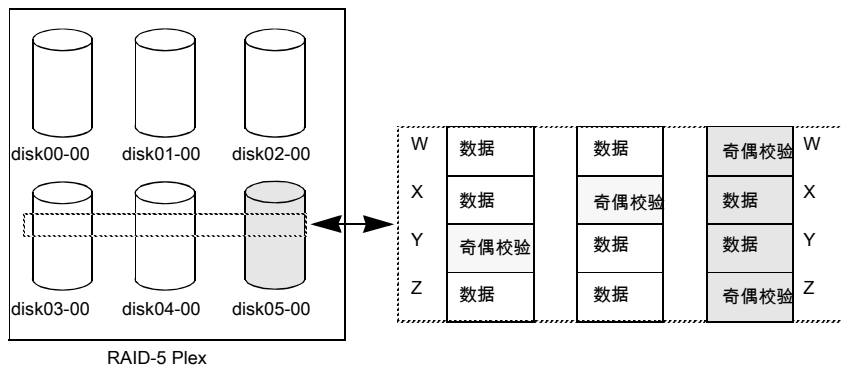
```
VxVM vxvol ERROR V-5-1-1236 Volume r5vol is not startable; RAID-5 Plex does not map entire volume length.
```

此时，RAID-5 卷的内容不可用。

RAID-5 卷无法启动的另一种可能的原因是，奇偶校验失效且子磁盘分离或失效。出现这种情况的原因在于包含故障子磁盘的条带内奇偶校验条带单元无效（因为奇偶校验失效）以及故障子磁盘上的条带单元也无效。

图 3-3 显示由于失效的奇偶校验以及故障子磁盘而导致无效的 RAID-5 卷。

图 3-3 无效的 RAID-5 卷



RAID-5 阵列中有四个条带。所有奇偶校验均失效且子磁盘 `disk05-00` 发生故障。这使条带 X 和 Y 因条带中发生两种故障而不可用。

这是一个条带内发生两种故障从而导致卷无法使用的情况。在这种情况下，`vxvol start` 命令的输出为：

```
VxVM vxvol ERROR V-5-1-1237 Volume r5vol is not startable;  
some subdisks are unusable and the parity is stale.
```

如果始终在 RAID-5 卷中使用两个或更多的 RAID-5 日志 Plex，则可避免这种情况。RAID-5 日志 Plex 可防止卷内的奇偶校验失效，因而避免了这种情况的发生。

请参见第 32 页的“系统故障”。

强制启动具有失效子磁盘的 RAID-5 卷

即使子磁盘被标记为失效也可以启动卷：例如，停止的卷有失效的奇偶校验，没有 RAID-5 日志，并且某个磁盘分离后重新挂接，就属于这种情况。

即使数据没有过期也认为子磁盘失效（原因是在子磁盘不可用时使用了卷），并认为 RAID-5 卷无效。为防止这种情况的发生，应尽可能地始终将多个有效的 RAID-5 日志与卷相关联。

强制启动具有失效子磁盘的 RAID-5 卷

- ◆ 指定 `vxvol start` 命令的 `-f` 选项。

```
# vxvol [-g diskgroup] -f start r5vol
```

这使所有失效子磁盘都标记为未失效。标记操作发生在 `start` 操作评估 RAID-5 卷的有效性并确定启动所需的条件之前。可通过使用以下命令将单个子磁盘标记为未失效：

```
# vxmend [-g diskgroup] fix unstale subdisk
```

如果某些子磁盘因失效而需要恢复，并且存在有效的日志，则通过将卷置为 ENABLED 内核状态可启用该卷，并且在子磁盘修复期间该卷可用。否则，卷的内核状态将设置为 DETACHED，并且该卷在子磁盘恢复期间不可用。这样做是因为，如果卷活动时系统崩溃或卷异常停止，则奇偶校验会变为失效，从而使卷不可用。如果不希望这种做，则可使用 `-o unsafe start` 选项来启动卷。

警告： `-o unsafe start` 选项被视为是有害选项，因为它会使卷的内容不可用。因此建议不要使用此选项。

卷状态设置为 RECOVER 并还原失效子磁盘。由于每个子磁盘上的数据变为有效，因此子磁盘不再标记为失效。如果有任何子磁盘的恢复失败且没有有效的日志，则因为子磁盘仍处于失效状态，卷的启动将中止，并且系统崩溃会使 RAID-5 卷不可用。这也可以通过使用 `-o unsafe start` 选项来覆盖。

如果卷有有效的日志，则子磁盘恢复故障会被记录下来，但不会停止启动过程。

当所有子磁盘均恢复后，卷变为 ENABLED 内核状态且标记为 ACTIVE。

在不完整的磁盘组移动后恢复

如果在磁盘组移动、拆分或结合操作正在进行时系统崩溃或子系统失败，则在系统重新启动或子系统修复时，VxVM 会尝试反转或完成此操作。具体是反转还是完成此操作取决于操作的进度。

能否自动恢复取决于是否能导入源磁盘组和目标磁盘组。但是，有时可能无法自动恢复，例如，如果已在另一主机上导入其中一个磁盘组，即是如此。

在不完整的磁盘组移动后恢复

- 1 使用 `vxprint` 命令检查这两个磁盘组的配置。未完成移动的磁盘组中的对象的 `TUTILO` 字段设置为 `MOVE`。
- 2 输入以下命令尝试完成移动：

```
# vxdg recover sourcedg
```

如果因为其中一个磁盘组已导入到另一主机或因为它不存在而不能导入，则此操作失败：

```
VxVM vxdg ERROR V-5-1-2907 diskgroup: Disk group does not exist
```

如果恢复失败，请根据需要执行下列步骤。

- 3 如果磁盘组已导入到另一主机，将它从那个主机上逐出，然后导入当前主机。如果在源磁盘组或目标磁盘组中，所有需要的对象均已存在，则使用以下命令在该磁盘组中重新设置 `MOVE` 标志：

```
# vxdg -o clean recover diskgroup1
```

在另一磁盘组上使用以下命令，将那些 `TUTILO` 字段标记为 `MOVE` 的对象删除：

```
# vxdg -o remove recover diskgroup2
```

- 4 如果只能导入一个磁盘组，则使用以下命令在此磁盘组上重新设置 `MOVE` 标志：

```
# vxdg -o clean recover diskgroup
```

在 DCO 卷出现故障后恢复

从数据更改对象 (DCO) 卷故障中恢复的步骤取决于 DCO 版本号。

有关 DCO 版本的信息，请参见《Veritas Storage Foundation 管理指南》。

持久性快速重新同步使用 DCO 卷来跟踪卷中的已更改区域。如果读取或写入 DCO 日志卷时发生错误，则分离该卷并在 DCO 上设置 badlog 标志。以后写入卷的所有内容将不被 DCO 跟踪。

以下 vxprint 命令的输出示例显示了一个具有分离 DCO 卷的完整卷（为清楚起见，省略了 TUTILO 和 PUTILO 字段）：

TY	NAME	ASSOC	KSTATE	LENGTH	PLOFFS	STATE ...
dg	mydg	mydg	-	-	-	-
dm	mydg01	hdisk10	-	35521408	-	-
dm	mydg02	hdisk11	-	35521408	-	-
dm	mydg03	hdisk12	-	35521408	-	FAILING
dm	mydg04	hdisk13	-	35521408	-	FAILING
dm	mydg05	hdisk14	-	35521408	-	-
v	SNAP-vol1	fsgen	ENABLED	204800	-	ACTIVE
pl	vol1-03	SNAP-vol1	ENABLED	204800	-	ACTIVE
sd	mydg05-01	vol1-03	ENABLED	204800	0	-
dc	SNAP-vol1_dco	SNAP-vol1	-	-	-	-
v	SNAP-vol1_dcl	gen	ENABLED	144	-	ACTIVE
pl	vol1_dcl-03	SNAP-vol1_dcl	ENABLED	144	-	ACTIVE
sd	mydg05-02	vol1_dcl-03	ENABLED	144	0	-
sp	vol1_snp	SNAP-vol1	-	-	-	-
v	vol1	fsgen	ENABLED	204800	-	ACTIVE
pl	vol1-01	vol1	ENABLED	204800	-	ACTIVE
sd	mydg01-01	vol1-01	ENABLED	204800	0	-
pl	vol1-02	vol1	ENABLED	204800	-	ACTIVE
sd	mydg02-01	vol1-01	ENABLED	204800	0	-
dc	vol1_dco	vol1	-	-	-	BADLOG
v	vol1_dcl	gen	DETACHED	144	-	DETACH
pl	vol1_dcl-01	vol1_dcl	ENABLED	144	-	ACTIVE
sd	mydg03-01	vol1_dcl-01	ENABLED	144	0	-
pl	vol1_dcl-02	vol1_dcl	DETACHED	144	-	IOFAIL
sd	mydg04-01	vol1_dcl-02	ENABLED	144	0	RELOCATE
sp	SNAP-vol1_snp	vol1	-	-	-	-

此输出显示了镜像卷 vol1、它的快照卷 SNAP-vol1 以及它们相应的 DCO vol1_dco 和 SNAP-vol1_dco。保存 vol1 的 DCO 卷 vol1_dcl 的 DCO Plex 的两个磁盘 mydg03 和 mydg04 发生故障。因此，分离了卷 vol1 的 DCO 卷 vol1_dcl，并将 vol1_dco 的状态设置为 BADLOG。为了将来参考，请记住分别指向 vol1 和 SNAP-vol1 的 snap 对象 vol1_snp 和 SNAP-vol1_snp 的项。

可以使用该输出推导出卷的 DCO 名称（在此例中为 `vol1_dco`），或者使用以下 `vxprint` 命令显示卷的 DCO 名称：

```
# vxprint [-g diskgroup] -F%dco_name volume
```

可以使用 `vxprint` 命令检查是否为卷的 DCO 设置了 `badlog` 标志，如下所示：

```
# vxprint [-g diskgroup] -F%badlog dco_name
```

如果设置了 `badlog` 标志，则此命令返回值 `on`。对于上面的输出示例，此命令应采用以下形式：

```
# vxprint -g mydg -F%badlog vol1_dco
```

```
on
```

可以使用以下命令验证 DCO 的版本号：

```
# vxprint [-g diskgroup] -F%version dco_name
```

返回值为 0 或 20。对于上面的输出示例，此命令应采用以下形式：

```
# vxprint -g mydg -F%version vol1_dco
```

DCO 版本号决定应该使用的恢复过程。

请参见第 42 页的“[恢复 0 版 DCO 卷](#)”。

请参见第 45 页的“[恢复即时快照 DCO 卷（版本 20 或更高版本）](#)”。

恢复 0 版 DCO 卷

恢复 0 版 DCO 卷

- 1 解决导致 I/O 故障的问题。
- 2 使用以下命令从 DCO 中删除 `badlog` 标志：

```
# vxdco [-g diskgroup] -o force enable dco_name
```

对于上面的输出示例，此命令应采用以下形式：

```
# vxdco -g mydg -o force enable vol1_dco
```

`vxprint` 输出中用于 `vol1_dco` 的项现在类似于：

```
dc vol1_dco          vol1          -          -          -          -
```

3 使用以下命令重新启动 DCO 卷：

```
# vxvol [-g diskgroup] start dco_log_vol
```

对于上面的输出示例，此命令应采用以下形式：

```
# vxvol -g mydg start vol1_dcl
```

- 4 使用 `vxassist snapclear` 命令清除原始卷及其所有快照的快速重新同步映像。这样可以确保在对快照进行快照合并（执行完全重新同步）时不会使用有可能失效的快速重新同步映像。任何后续的卷快照操作都重新启用快速重新同步跟踪。

警告：在将 `badlog` 标志从 DCO 中删除后，必须对该卷的所有快照使用 `vxassist snapclear` 命令。否则，在对快照进行快照合并时有可能丢失或损坏数据。

如果卷及其快照卷在同一磁盘组中，则以下命令将同时清除这两个卷的快速重新同步映像：

```
# vxassist [-g diskgroup] snapclear volume \  
    snap_obj_to_snapshot
```

其中，`snap_obj_to_snapshot` 是与指向快照卷的 `volume` 关联的 `snap` 对象的名称。

对于上面的输出示例，此命令应采用以下形式：

```
# vxassist -g mydg snapclear vol1 SNAP-vol1_snp
```

如果快照卷和原始卷在不同的磁盘组中，则必须分别对每个卷执行 `snapclear` 操作：

```
# vxassist -g diskgroup1 snapclear volume snap_obj_to_snapshot  
# vxassist -g diskgroup2 snapclear snapvol snap_obj_to_volume
```

其中，`snap_obj_to_volume` 是与指向原始卷的快照卷 `snapvol` 关联的 `snap` 对象的名称。

对于上面的输出示例，如果已将 `SNAP-vol1` 移到磁盘组 `snapg` 中，此命令应采用以下形式：

```
# vxassist -g mydg snapclear vol1 SNAP-vol1_snp  
# vxassist -g snapdg snapclear SNAP-vol1 vol1_snp
```

- 5 要对执行了 `snapclear` 操作的快照卷进行快照合并，请使用以下命令（如有必要，可在使用 `vx dg move` 命令将快照 Plex 移回原始磁盘组之后执行）：

```
# vxplex -f [-g diskgroup] snapback volume snapvol_plex
```

对于上面的输出示例，此命令应采用以下形式：

```
# vxplex -f -g mydg snapback vol1 vol1-03
```

不能使用 `vxassist snapback` 命令，因为 `snapclear` 操作会删除快照关联信息。

恢复即时快照 DCO 卷（版本 20 或更高版本）

恢复即时快照 DCO 卷

- 1 解决导致 I/O 故障的问题。
- 2 使用 `vxsnap` 命令分离每个与卷关联的完整空间即时快照卷：

```
# vxsnap [-g diskgroup] dis snapvol
```

对于上面的输出示例，此命令应采用以下形式：

```
# vxsnap -g mydg dis SNAP-vol1
```

- 3 使用以下命令取消卷的准备：

```
# vxsnap [-g diskgroup] unprepare volume
```

对于上面的输出示例，此命令应采用以下形式：

```
# vxsnap -g mydg unprepare vol1
```

- 4 使用 `vxvol` 命令启动卷：

```
# vxvol [-g diskgroup] start volume
```

对于上面的输出示例，此命令应采用以下形式：

```
# vxvol -g mydg start vol1
```

- 5 使用以下命令再次准备卷：

```
# vxsnap [-g diskgroup] prepare volume [ndcomirs=number]\  
[regionsize=size] [drl=yes|no|sequential]\  
[storage_attribute ...]
```

对于上面的输出示例，此命令可采用以下形式：

```
# vxsnap -g mydg prepare vol1 ndcomirs=2 drl=yes
```

这将为 DCO 卷添加 2 个 Plex，还将启用 DRL 和快速重新同步（如果得到授权）。

有关如何使用 `vxsnap prepare` 命令的全部详细信息，请参见《Veritas Storage Foundation 管理指南》和 `vxsnap(1M)` 手册页。

在即时快照故障后恢复

本章节包括下列主题：

- 在 `vxsnap prepare` 失败后恢复
- 在对完整空间即时快照执行 `vxsnap make` 失败后恢复
- 在对断开即时快照执行 `vxsnap make` 失败后恢复
- 在对优化空间即时快照执行 `vxsnap make` 失败后恢复
- 在 `vxsnap restore` 失败后恢复
- 在 `vxsnap refresh` 失败后恢复
- 在“写入时复制”失败后恢复
- 发生重新同步 I/O 错误后恢复
- DCO 卷发生 I/O 故障后恢复
- 从对即时快照数据更改对象 (DCO) 执行 `vxsnap upgrade` 时出现的故障中恢复

在 `vxsnap prepare` 失败后恢复

如果 `vxsnap prepare` 操作过早失败，`vxprint` 命令可能以 `INSTSNAPTMP` 状态显示新的 DCO 卷。VxVM 通常可以在没有干预的情况下恢复 DCO 卷。但在某些情况下，此恢复可能不会成功。发生这种情况后，必须删除 DCO 卷。

在 `vxsnap prepare` 命令失败后恢复

- ◆ 键入以下命令：

```
# vxedit [-g diskgroup] rm DCO_volume
```

或者选择系统下次重新启动时自动删除 DCO 卷。将 DCO 卷删除后，请再次运行 `vxsnap prepare` 命令。

在对完整空间即时快照执行 vxsnap make 失败后恢复

在创建完整空间即时快照的过程中，如果 vxsnap make 操作失败，快照卷可能会变为 DISABLED 状态，被标记为无效，且无法启动。可使用以下命令验证 inst_invalid 标志是否设置为 on：

```
# vxprint [-g diskgroup] -F%inst_invalid snapshot_volume
```

VxVM 通常在没有干预的情况下恢复快照卷。但在某些情况下，此恢复可能不会成功。发生这种情况后，必须删除 DCO 卷。

在对完整空间即时快照执行 vxsnap make 命令失败后恢复

- 1 清除快照卷的 tutil0 字段。输入以下命令：

```
# vxmend [-g diskgroup] clear tutil0 snapshot_volume
```

- 2 取消准备快照卷。输入以下命令：

```
# vxsnap [-g diskgroup] unprepare snapshot_volume
```

- 3 如果快照卷处于 DISABLED 状态，请启动快照卷。输入以下命令：

```
# vxvol [-g diskgroup] start snapshot_volume
```

- 4 再次为快照操作准备快照卷。输入以下命令：

```
# vxsnap [-g diskgroup] prepare snapshot_volume
```

- 5 清除卷的 tutil0 字段（如果已设置）。输入以下命令。

```
# vxmend [-g diskgroup] clear tutil0 original_volume
```

在对断开即时快照执行 vxsnap make 失败后恢复

在创建第三镜像断开即时快照的过程中，如果 vxsnap make 操作失败，则快照卷可能变为 INSTSNAPTMP 状态。VxVM 通常可以在没有干预的情况下恢复快照卷。但在某些情况下，此恢复可能不会成功。发生这种情况后，必须删除快照卷。

在对断开即时快照执行 vxsnap make 命令失败后恢复

- ◆ 键入以下命令：

```
# vxedit [-g diskgroup] rm snapshot_volume
```

或者选择系统下次重新启动时自动删除快照卷。

在对优化空间即时快照执行 vxsnap make 失败后恢复

在创建优化空间即时快照的过程中，如果 vxsnap make 操作失败，则快照卷可能变为 INSTSNAPTMP 状态。VxVM 通常可以在没有干预的情况下恢复快照卷。但在某些情况下，此恢复可能不会成功。发生这种情况后，必须删除快照卷。

在对优化空间即时快照执行 vxsnap make 命令失败后恢复

◆ 键入以下命令：

```
# vxedit [-g diskgroup] rm snapshot_volume
```

或者选择系统下次重新启动时自动删除快照卷。

如果通过指定 cache 属性，对准备好的缓存对象执行 vxsnap make 操作，则删除快照后缓存对象保持不变。如果使用 cachesize 属性指定了新的缓存对象，删除快照后该缓存对象便不继续存在。

在 vxsnap restore 失败后恢复

如果 vxsnap restore 操作失败，所还原的卷可能变为 DISABLED 状态。

在 vxsnap restore 命令失败后恢复

◆ 键入以下命令：

```
# vxvol [-g diskgroup] start volume
```

在 vxsnap refresh 失败后恢复

如果 vxsnap refresh 操作失败，则在恢复卷后 VxVM 会尝试重新开始同步。卷将启动，同步在后台进行。

此过程要求磁盘组版本为 170 或更高版本。

所刷新的卷可能变为 DISABLED 状态、被标记为无效且无法启动。

在 vxsnap refresh 命令失败后恢复

1 使用以下命令检查是否将 inst_invalid 标志设置为 on：

```
# vxprint [-g diskgroup] -F%inst_invalid volume
```

2 使用 vxmend 命令清除卷的 tutil0 字段：

```
# vxmend [-g diskgroup] clear tutil0 volume
```

- 3 使用以下命令启动卷：

```
# vxvol [-g diskgroup] start volume
```

- 4 再次运行失败的 `refresh` 命令。

这将完全重新同步卷。或者选择在需要时删除快照卷并重新创建它。

在“写入时复制”失败后恢复

如果执行内部的卷快照重新同步时遇到错误，则快照卷会变为 `INVALID` 状态，并且无法用于 I/O 存取和即时快照操作。

在“写入时复制”失败后恢复

- 1 使用 `vxsnap` 命令将卷从快照层次结构中分离出来：

```
# vxsnap [-g diskgroup] dis snapshot_volume
```

- 2 使用以下命令取消卷的准备：

```
# vxsnap [-g diskgroup] unprepare snapshot_volume
```

- 3 使用以下命令准备卷：

```
# vxsnap [-g diskgroup] prepare volume [ndcomirs=number] \  
[regionsize=size] [dr1=yes|no|sequential] \  
[storage_attribute ...]
```

该卷现在可以再次用于快照操作。

或者选择在需要时删除快照卷并重新创建它。

发生重新同步 I/O 错误后恢复

如果发生 I/O 错误，就会停止快照重新同步（通过 `vxsnap syncstart` 或为 `vxsnap` 指定 `sync=on` 启动），并且在系统控制台上显示以下消息：

```
VxVM vxsnap ERROR V-5-1-6840 Synchronization of the volume  
volume stopped due to I/O error
```

更正错误源后，重新启动重新同步操作。

发生重新同步 I/O 错误后恢复

- ◆ 键入以下命令：

```
# vxsnap [-b] [-g diskgroup] syncstart volume
```

DCO 卷发生 I/O 故障后恢复

如果 DCO 卷上发生 I/O 故障，则无法访问它的快速重同步映射和 DRL 日志，并且该 DCO 卷被用 BADLOG 标志做了标记。除非恢复卷的 DCO 卷，否则无法对该卷执行 DRL 日志记录与恢复以及即时快照操作。

如果 I/O 故障还影响数据卷，则必须先恢复数据卷，然后才能恢复其 DCO 卷。

请参见第 45 页的“恢复即时快照 DCO 卷（版本 20 或更高版本）”。

从对即时快照数据更改对象 (DCO) 执行 vxsnap upgrade 时出现的故障中恢复

当您升级即时快照数据更改对象 (DCO) 时，DCO 及 DCO 卷将就地升级。如果 DCO 中没有足够的空间来容纳升级后的新 DCO，此操作将失败。

如果此操作因缺乏空间而失败，请采用以下方法之一：

使用 `vxsnap upgrade` 命令的 `-f` 选项。指定了 `-f` 时，VxVM 将使用磁盘组中的任意可用空间来容纳升级后的新 DCO。因此，如果磁盘组中有足够的可用空间，此升级操作便可能会成功。

```
# vxsnap [-g diskgroup] -f upgrade  
[volume1|volset1][volume2|volset2...]
```

使用 `alloc` 属性指定要用来保存新 DCO 的存储。VxVM 将在指定的存储上创建新的 DCO。有关存储属性的信息，请参见 `vxsnap(m1)` 手册页。

```
# vxsnap -g diskgroup upgrade  
[volume1|volset1][volume2|volset2...]  
[alloc=storage_attributes]
```

如果既指定了 `alloc` 属性又指定了 `-f` 选项，VxVM 会对 `alloc` 属性使用指定的存储。

52 | 在即时快照故障后恢复
| 从对即时快照数据更改对象 (DCO) 执行 vxsnap upgrade 时出现的故障中恢复

在引导磁盘故障后恢复

本章节包括下列主题：

- [VxVM 和引导磁盘故障](#)
- [通过重新安装来恢复](#)

VxVM 和引导磁盘故障

本版 VxVM 不支持将引导磁盘置于 Veritas Volume Manager 控制之下的根目录可置性。但是，如果发生引导（根）磁盘发生故障，有一个过程可用于恢复 VxVM 配置。

请参见第 53 页的[“通过重新安装来恢复”](#)。

通过重新安装来恢复

如果引导 (root) 磁盘的所有副本均损坏，或某些关键文件因文件系统损坏而丢失，则必须重新安装。

如果发生这种故障，请尝试尽可能多地保留原始 VxVM 配置。故障中未直接涉及的卷不必重新配置。保留的卷也无须重新配置。

一般重新安装信息

在发生引导（根）磁盘故障后重新安装 VxVM 时，应该尽可能多地保留原始配置。

警告： 系统重新安装将破坏用于重新安装的所有磁盘上的内容。

重新安装期间将删除所有与 VxVM 相关的信息。删除的数据包括被删除磁盘上专用区域（含有磁盘标识符和 VxVM 配置的副本）中的数据。该信息的删除使磁盘无法用作 VM 磁盘。

在重新安装的过程中总会涉及系统的 root 磁盘。重新安装所涉及的任何其他磁盘，或者那些被删除和替换的磁盘，都会丢失 VxVM 配置数据（包括卷和镜像）。

如果磁盘在发生故障前不受 VxVM 控制，则在重新安装时不会丢失 VxVM 配置数据。

重新安装时，将只保存故障和重新安装未直接涉及的那些磁盘上的卷，或是在这些磁盘上有副本的卷。根磁盘及故障或重新安装所涉及的其他磁盘上的任何卷都将在重新安装中丢失。如果这些卷的备份副本可用，则可以在重新安装后还原卷。在某些系统中，root、stand 和 usr 文件系统属于例外情况，无法从备份还原。

重新安装系统并恢复 VxVM

要重新安装系统并恢复 Veritas Volume Manager 配置，需要完成以下步骤：

- 更换出错的磁盘或其他硬件，并分离重新安装中所不涉及的任何磁盘。
请参见第 54 页的“[为重新安装系统做准备](#)”。
- 重新安装基础系统和任何其他不相关的 Volume Manager 文件集。
请参见第 54 页的“[重新安装操作系统](#)”。
- 添加 Volume Manager 文件集，但不执行 vxinstall 命令。
请参见第 55 页的“[重新安装 Veritas 软件](#)”。
- 恢复 Veritas Volume Manager 配置。
请参见第 55 页的“[恢复 Veritas Volume Manager 配置](#)”。
- 还原受故障或重新安装影响的卷中的任何信息，并重新创建系统卷（rootvol、swapvol、usr 以及其他系统卷）。
请参见第 56 页的“[清理系统配置](#)”。

为重新安装系统做准备

为防止重新安装过程中未涉及的磁盘丢失数据，在重新安装过程中应只对根磁盘操作。

用于安装的几个自动选项无须得到管理员确认即可访问除根磁盘以外的磁盘。在重新安装操作系统之前，断开包含卷的所有其他磁盘与系统的连接。

断开其他磁盘，以确保其不受重新安装的影响。例如，如果最初安装操作系统时 home 文件系统位于第二个磁盘上，仍可以将其恢复。删除第二个磁盘可以确保 home 文件系统保持原样。

重新安装操作系统

更换任何出错的磁盘或正在发生故障的磁盘并分离重新安装中所不涉及的磁盘后，请按操作系统文档的说明重新安装操作系统。先安装操作系统，然后再安装 VxVM。

在安装操作系统时，必须确保不以任何方式访问除根磁盘外的任何其他磁盘。无论将什么内容写入除根磁盘以外的其他磁盘，都可能破坏该磁盘上的 Veritas Volume Manager 配置。

注意：在重新安装期间可以更改系统的主机名称（或主机ID）。以下几节中的处理过程假定您未更改主机名称，因此建议保留现有的主机名称。

重新安装 Veritas 软件

重新安装 Veritas Volume Manager

- 1 从安装光盘重新安装 Veritas 软件。

请参见您的 Storage Foundation 产品的“安装指南”。

警告：不要使用 `vxinstall` 来初始化 VxVM。

- 2 如果需要，使用 `vxlicinst` 命令安装 Veritas Volume Manager 许可证密钥。

请参见 `vxlicinst(1)` 手册页。

恢复 Veritas Volume Manager 配置

加载 Veritas Volume Manager 文件集并安装 VxVM 的许可证之后，恢复 Veritas Volume Manager 配置。

恢复 Veritas Volume Manager 配置

- 1 创建 (`touch`) `/etc/vx/reconfig.d/state.d/install-db` 文件。
- 2 关闭该系统。
- 3 重新挂接那些从系统中删除的磁盘。
- 4 重新启动系统。
- 5 系统启动时，使用下面的命令将系统引入单用户模式：

```
# exec init S
```
- 6 出现提示时，输入密码并按回车键继续。
- 7 使用以下命令删除安装过程中涉及的一些文件（这些文件是在加载 VxVM 时创建的，但已不再需要）：

```
# rm -rf /etc/vx/reconfig.d/state.d/install-db
```

- 8 使用以下命令启动一些 Veritas Volume Manager I/O 后台驻留程序：

```
# vxiod set 10
```

- 9 使用以下命令以禁用模式启动 Veritas Volume Manager 配置后台驻留程序

```
vxconfigd:
```

```
# vxconfigd -m disable
```

- 10 使用以下命令初始化 vxconfigd 后台驻留程序：

```
# vxdctl init
```

- 11 使用以下命令初始化 DMP 子系统：

```
# vxdctl initdmp
```

- 12 使用以下命令启用 vxconfigd：

```
# vxdctl enable
```

至此，保留在重新安装未涉及的磁盘上的配置已经恢复。但是，由于重新安装了根磁盘，因此，对 VxVM 来说，根磁盘不再是 VM 磁盘。保留磁盘的配置不把根磁盘包括在 VxVM 配置中。

如果系统根磁盘以及重新安装过程所涉及的其他任何磁盘在发生故障和重新安装时不受 VxVM 控制，那么此时即完成了重新配置。

如果根磁盘（或另一个磁盘）与重新安装有关，则该磁盘（或其他不再挂接到系统的磁盘）上的任何卷或镜像现在都无法访问。如果卷只有一个位于已重新安装、删除或更换的磁盘上的 Plex，那么该卷中的数据将丢失，用户必须从备份中还原它。

清理系统配置

重新安装 VxVM 之后，必须清理系统配置。

清理系统配置

- 1 恢复 VxVM 配置之后，必须确定需要从备份中还原的卷。要还原的卷包括那些所有镜像（卷的所有副本）都位于被重新安装或删除的磁盘上的卷。这些卷是无效的，必须将它们删除，然后重新创建它们并从备份中还原。如果在重新初始化或删除的磁盘上只存在卷的部分镜像，则必须删除这些镜像。可在以后重新添加这些镜像。

使用以下命令确定删除或重新安装了哪些 VM 磁盘：

```
# vxdisk list
```

此命令会显示系统磁盘设备及其状态的列表。例如，对于一个重新安装的系统（它有三个磁盘和一个重新安装的根磁盘），`vxdisk list` 命令的输出类似于以下内容：

DEVICE	TYPE	DISK	GROUP	STATUS
hdisk1	simple	-	-	online invalid
hdisk2	simple	mydg01	mydg	online
hdisk3	simple	mydg02	mydg	online
hdisk4	simple	mydg03	mydg	online

输出表明，重新安装的根设备 `hdisk1` 与 VM 磁盘不关联，并带有 `online invalid` 状态标记。重新安装中不涉及磁盘 `mydg01`、`mydg02` 和 `mydg03`，它们被 VxVM 识别并与其设备（`hdisk2`、`hdisk3` 和 `hdisk4`）关联。

如果重新安装过程中删除或更换了其他磁盘（其上有卷或镜像），这些磁盘也会有一个以 `online invalid` 状态列出的磁盘设备和一个未与设备关联的 VM 磁盘。

- 2 确定了删除或更换的磁盘后，使用以下命令定位出错的磁盘上的所有镜像：

```
# vxprint -sF "%vname" -e'sd_disk = "disk"
```

其中，`disk` 是具有 `failed` 状态的磁盘的访问名称。确保在此命令中将磁盘名称括在引号里面。否则，此命令将返回错误消息。`vxprint` 命令返回在出错的磁盘上有镜像的卷的列表。对每个具有 `failed` 状态的磁盘重复使用此命令。

3 检查每个卷的状态并使用以下命令输出卷信息：

```
# vxprint -th volume
```

其中，**volume** 是要检查的卷的名称。vxprint 命令显示该卷的状态、Plex 和构成那些 Plex 的磁盘部分。

例如，名为 v01 的卷只有一个 Plex 驻留在名为 mydg01 的重新安装的磁盘上。

vxprint -th v01 命令产生以下输出：

V	NAME	USETYPE	KSTATE	STATE	LENGTH	READPOL	PREFPLEX		
PL	NAME	VOLUME	KSTATE	STATE	LENGTH	LAYOUT	NCOL/WID	MODE	
SD	NAME	PLEX	DISK	DISKOFFS	LENGTH	[COL/]OFF	DEVICE	MODE	
v	v01	fsgen	DISABLED	ACTIVE	24000	SELECT	-		
pl	v01-01	v01	DISABLED	NODEVICE	24000	CONCAT	-	RW	
sd	mydg01-06	v0101	mydg01	245759	24000	0	hdisk2	ENA	

卷唯一的 Plex 显示在以 pl 开头的行中。名为 v01-01 的 Plex 的 STATE 字段为 NODEVICE。该 Plex 在被替换、删除或重新安装的磁盘上占用了空间。它不再有效，必须将其删除。

4 由于 v01-01 是卷的唯一 Plex，因此，除非从备份中还原，否则该卷的内容无法还原。该卷还必须被删除。如果卷有一个备份副本，则可以以后还原。保留该卷的名称及其长度的记录，因为在备份过程中会需要它。

使用以下命令删除不可恢复的卷（如 v01）：

```
# vxedit -r rm v01
```

- 5 有可能 Plex 只有一部分位于出错的磁盘上。如果卷有关联的条带 Plex，该卷就被划分到若干个磁盘上。例如，名为 v02 的卷有一个条带 Plex，该 Plex 条带分布在三个磁盘上，其中一个磁盘是重新安装的磁盘 mydg01。vxprint -th v02 命令产生以下输出：

V	NAME	USETYPE	KSTATE	STATE	LENGTH	READPOL	PREFPLEX		
PL	NAME	VOLUME	KSTATE	STATE	LENGTH	LAYOUT	NCOL/WID	MODE	
SD	NAME	PLEX	DISK	DISKOFFS	LENGTH	[COL/]OFF	DEVICE	MODE	
v	v02	fsgen	DISABLED	ACTIVE	30720	SELECT	v02-01		
pl	v02-01	v02	DISABLED	NODEVICE	30720	STRIPE	3/128	RW	
sd	mydg02-02v02-01		mydg01	424144	10240	0/0	hdisk3	ENA	
sd	mydg01-05v02-01		mydg01	620544	10240	1/0	hdisk2	DIS	
sd	mydg03-01v02-01		mydg03	620544	10240	2/0	hdisk4	ENA	

输出显示了 3 个磁盘，Plex v02-01 条带分布在这 3 个磁盘上（以 sd 开始的行代表这些条带）。其中一个条带区域位于出错的磁盘上。此磁盘不再有效，因此，名为 v02-01 的 Plex 处于 NODEVICE 状态。由于这是卷的唯一 Plex，因此该卷无效，必须将其删除。如果备份介质上有 v02 的一个副本，则可以以后还原它。若打算将卷从备份中还原，则要保留该卷名称和长度的记录。

使用以下命令删除无效卷（如 v02）：

```
# vxedit -r rm v02
```

- 6 在出错的磁盘上有一个镜像的卷可能在仍然有效的磁盘上还有其他镜像。在这种情况下，不需要从备份中还原该卷，因为有效磁盘上的数据仍然有效。

对于在出错的磁盘(mydg01)上有一个 Plex 并在有效磁盘(mydg02)上有另一个 Plex 的卷，`vxprint -th` 命令的输出类似于以下内容：

V	NAME	USETYPE	KSTATE	STATE	LENGTH	READPOL	PREFFLEX	
PL	NAME	VOLUME	KSTATE	STATE	LENGTH	LAYOUT	NCOL/WID	MODE
SD	NAME	PLEX	DISK	DISKOFFS	LENGTH	[COL/]OFF	DEVICE	MODE
v	v03	fsgen	DISABLED	ACTIVE	0720	SELECT	-	
pl	v03-01	v03	DISABLED	ACTIVE	30720	CONCAT	-	RW
sd	mydg02-01	v03-01	mydg01	620544	30720	0	hdisk3	ENA
pl	v03-02	v03	DISABLED	NODEVICE	30720	CONCAT	-	RW
sd	mydg01-04	v03-02	mydg03	262144	30720	0	hdisk2	DIS

此卷有两个 Plex v03-01 和 v03-02。第一个 Plex (v03-01) 不使用无效磁盘上的空间，所以它仍然可用。第二个 Plex (v03-02) 使用无效磁盘 mydg01 上的空间，其状态为 NODEVICE。必须将 Plex v03-02 删除。但是，该卷仍然有一个包含有效数据的有效 Plex。如果需要镜像该卷，可以以后添加另一个 Plex。记下该卷的名称，以便将来创建另一个 Plex。

要删除无效的 Plex，请使用 `vxplex` 命令将此 Plex 与卷分离，然后将其删除。例如，要分离并删除 Plex v03-02，可使用以下命令：

```
# vxplex -o rm dis v03-02
```

- 7 删除了所有无效卷和无效 Plex 后，就可以清理磁盘配置了。必须将每个被删除、重新安装或更换的磁盘（根据 `vxdisk list` 命令的输出来确定）从配置中删除。

要删除磁盘，请使用 `vxdg` 命令。要删除出错的磁盘 mydg01，请使用以下命令：

```
# vxdg rmdisk mydg01
```

如果 `vxdg` 命令返回错误消息，说明存在无效的镜像。

重复步骤 1 到步骤 6，直到删除完所有无效卷和无效镜像。

- 8 现在，应使用 `vxdiskadm` 命令添加所有其他已替换磁盘。

- 9 将所有磁盘添加到系统后，可以重新创建在配置清理过程中完全删除的任何卷，并可以从备份中还原其内容。可以使用 `vxassist` 命令或图形用户界面重新创建卷。

例如，要重新创建卷 `v01` 和 `v02`，请使用以下命令：

```
# vxassist make v01 24000
# vxassist make v02 30720 layout=stripe nstripe=3
```

创建卷之后，就可以使用正常的备份/还原过程从备份中还原它们。

- 10 为 Plex 被删除（在卷清理过程中）的那些卷重新创建所有 Plex。要恢复卷 `v03` 中被删除的 Plex，请使用以下命令：

```
# vxassist mirror v03
```

还原了卷和在重新安装过程中丢失的 Plex 后，就完成了恢复，系统也会配置为故障发生之前的状态。

- 11 如有必要，启动热重定位，可通过重新启动系统或手动启动重定位监视后台驻留程序 `vxrelocd`（这还会启动 `vxnotify` 进程）完成。

警告： 仅当确信热重定位不会干扰其他重配置过程时，才可以启动它。

要确定热重定位是否已启动，请使用以下命令查找其在进程表中的项：

```
# ps -ef | grep vxrelocd
```

请参见《Veritas Storage Foundation 管理指南》。

请参见 `vxrelocd(1M)` 手册页。

管理命令和事务

本章节包括下列主题：

- [命令日志](#)
- [事务日志](#)
- [命令日志和事务日志的关联](#)
- [将从从属节点发出的 CVM 命令关联到主节点](#)
- [未启用命令完成](#)

命令日志

使用 `vxcmdlog` 命令可以将其他 Veritas Volume Manager (VxVM) 命令的调用操作记录到文件中。

下面举例说明 `vxcmdlog` 的用法：

<code>vxcmdlog -l</code>	列出命令日志记录的当前设置。
<code>vxcmdlog -m on</code>	打开命令日志记录功能。
<code>vxcmdlog -s 512k</code>	将最大命令日志文件大小设置为 512K。
<code>vxcmdlog -n 10</code>	将最大历史命令日志文件数设置为 10。
<code>vxcmdlog -n no_limit</code>	取消对历史命令日志文件数的限制。
<code>vxcmdlog -m off</code>	关闭命令日志记录功能。

默认情况下，命令日志记录处于打开状态。命令行将记录到 `/etc/vx/log` 目录下的 `cmdlog` 文件中。此路径名是一个指向目录的符号链接，目录的位置与操作系统有关。必要时可以重新定义所链接的目录。

如果要保留 `vxcmdlog` 实用程序的设置，还必须将设置文件 `.cmdlog` 复制到新的目录。

警告： `.cmdlog` 文件是一个二进制文件，不应对其进行编辑。

在写入一个条目（因此导致实际大小可能略大于指定的大小）之后，系统将检查命令日志的大小。当日志达到最大空间时，当前命令日志文件 `cmdlog` 被重命名为下一个可用的历史日志文件 `cmdlog.number`（其中 `number` 是一个整数，其范围为从 1 到当前定义的最大历史日志文件数），并且新建一个当前日志文件。

为了避免写满文件系统，系统仅保留有限数量的历史日志文件。如果达到最大历史日志文件数，则删除最早的历史日志文件，并将当前的日志文件重命名为该文件。

每个日志文件都包含一个标头，其中记录了主机名、主机 ID 以及日志的创建日期和时间。

下面是摘自一个命令日志文件的示例条目：

```
# 0, 2329, Wed Feb 12 21:19:31 2003
    /usr/sbin/vxdctl mode
# 17051, 2635, Wed Feb 12 21:19:33 2003
    /usr/sbin/vxdisk -q -o alldgs list
# 0, 2722, Wed Feb 12 21:19:34 2003
    /etc/vx/diag.d/vxprivutil dumpconfig /dev/vx/rdmp/Disk_4
# 26924, 3001, Thu Feb 13 19:30:57 2003
    /usr/sbin/vxdisk list Disk_1
```

每个条目通常包含客户端 ID（标识 `vxconfigd` 后台驻留程序的命令连接）、当前运行的命令的进程 ID、时间戳以及包含任何参数的命令行。

如果客户端 ID 是 0（如上例中的第三条），则说明该命令未打开与 `vxconfigd` 的连接。

该客户端 ID 与在事务日志中为相应事务记录的客户端 ID 相同。

请参见第 65 页的“事务日志”。

请参见第 66 页的“命令日志和事务日志的关联”。

大多数命令脚本不会被记录，但它们调用的命令二进制文件会被记录。

`vxdisksetup`、`vxinstall` 和 `vxdiskunsetup` 脚本例外，系统会记录这些脚本。

如果读取设置文件时发生错误，命令日志记录将切换到内置的默认设置。这可能意味着某些情况，例如在使用 `vxcmdlog -m off` 命令禁用日志记录后，日志记录仍保持启用状态。如果发生这种情况，请使用 `vxcmdlog` 实用程序重新创建设置文件，或使用备份还原此文件。

请参见 `vxcmdlog(1M)` 手册页。

事务日志

可以使用 `vxtranslog` 命令将 VxVM 事务记录到文件中。

下面举例说明 `vxtranslog` 的用法：

<code>vxtranslog -l</code>	列出事务日志记录的当前设置。
<code>vxtranslog -m on</code>	打开事务日志记录功能。
<code>vxtranslog -s 512k</code>	将最大事务日志文件大小设置为 512K。
<code>vxtranslog -n 10</code>	将最大历史事务日志文件数设置为 10。
<code>vxtranslog -n no_limit</code>	取消对历史事务日志文件数的限制。
<code>vxtranslog -q on</code>	打开查询日志记录功能。
<code>vxtranslog -q off</code>	关闭查询日志记录功能。
<code>vxtranslog -m off</code>	关闭事务日志记录功能。

默认情况下，事务日志记录处于打开状态。事务将记录到 `/etc/vx/log` 目录下的 `translog` 文件中。此路径名是一个指向目录的符号链接，目录的位置与操作系统有关。必要时可以重新定义所链接的目录。如果要保留 `vxtranslog` 实用程序的设置，还必须将设置文件 `.translog` 复制到新的目录。

警告： `.translog` 文件是一个二进制文件，不应对其进行编辑。

在写入一个条目（因此导致实际大小可能略大于指定的大小）之后，系统将检查事务日志的大小。当日志达到最大空间时，当前事务日志文件 `translog` 被重命名为下一个可用的历史日志文件 `translog.number`（其中 `number` 是一个整数，其范围为从 1 到当前定义的最大历史日志文件数），并且新创建一个当前日志文件。

为了避免写满文件系统，系统仅保留有限数量的历史日志文件。如果达到最大历史日志文件数，则删除最早的历史日志文件，并将当前的日志文件重命名为该文件。

每个日志文件都包含一个标头，其中记录了主机名、主机 ID 以及日志的创建日期和时间。

下面是摘自一个事务日志文件的示例条目：

```
Fri Oct 17 13:23:30 2003
Clid = 23460, PID = 21240, Part = 0, Status = 0, Abort Reason = 0
    DA_GET    Disk_0
    DISK_GET_ATTRS  Disk_0
    DISK_DISK_OP  Disk_0 8
```

```
DEVNO_GET    Disk_0
DANAME_GET    0x160045 0x160072
GET_ARRAYNAME Disk DISKS
CTLR_PTOLNAME 11-08-01
GET_ARRAYNAME Disk DISKS
CTLR_PTOLNAME 21-08-01
DROPPED      <no request data>
```

每个日志条目的第一行是事务的时间戳。Clid 字段与该命令打开的 vxconfigd 连接的客户端 ID 相对应。PID 字段显示正在请求操作的实用程序的进程 ID。如果事务未正常完成，Status 和 Abort Reason 字段将包含错误代码。记录的其余部分将显示处理事务时使用的数据。

该客户端 ID 与在命令日志中为对应命令行所记录的客户端 ID 相同。

请参见第 63 页的“命令日志”。

请参见第 66 页的“命令日志和事务日志的关联”。

如果读取设置文件时发生错误，事务日志记录将切换到内置的默认设置。这可能意味着某些情况，例如在使用 vxtranslog -m off 命令禁用日志记录后，日志记录仍保持启用状态。如果发生这种情况，请使用 vxtranslog 实用程序重新创建设置文件，或使用备份还原此文件。

命令日志和事务日志的关联

为每个请求和命令记录的客户端 ID 和进程 ID 可帮助您将命令日志和事务日志中的条目关联起来。要确定事务日志中的某个请求是哪个命令发出的，可使用如下命令在命令日志中搜索进程 ID 和客户端 ID：

```
# egrep -n PID cmdlog | egrep Clid
```

在本例中，事务日志内记录了以下请求：

```
Wed Feb 12 21:19:36 2003
Clid = 8309, PID = 2778, Part = 0, Status = 0, Abort Reason = 0
  DG_IMPORT foodg
  DG_IMPORT foodg
  DISCONNECT <no request data>
```

要找到发出此请求的实用程序，可使用如下命令：

```
# egrep -n 2778 cmdlog | egrep 8309
7310:# 8309, 2778, Wed Feb 12 21:19:36 2003
```

该示例输出表明命令日志第 7310 行上存在匹配项。通过检查命令日志中的第 7310 行和 7311 行发现，系统对 foodg 磁盘组运行过 vxldg import 命令：

```
# sed -e '7310,7311!d' cmdlog
# 8309, 2778, Wed Feb 12 21:19:36 2003 7311
/usr/sbin/vxldg -m import foodg
```

如果客户端 ID 和进程 ID 的组合有多个匹配项，可以通过检查时间戳来确定正确的匹配项。

当实用程序打开与 vxconfigd 的条件连接时，其客户端 ID 在命令日志中显示为零，而在事务日志中显示为非零值。这种情况下，可以使用进程 ID 和时间戳关联日志条目。

将从从属节点发出的 CVM 命令关联到主节点

在 CVM 从属节点上运行用于更改共享磁盘组配置的命令时，CVM 会将这些命令传送到 CVM 主节点进行执行。

例如，在从属节点上运行以下命令，以在共享磁盘组中创建一个卷。CVM 会将该命令传送到主节点，且 CVM 会在主节点上执行该命令。

```
# vxassist -g shareddg make shared-vol1 200M
```

在 CVM 从属节点上，输入以下命令可从事务日志 (translog) 中识别传送的命令：

```
# egrep CMDSHIP_REQUEST translog
```

在本示例中，以下条目录录在从属节点上的事务日志中：

```
Thu Jul 15 06:30:16 2010
Clid = 5302, PID = 589906, Part = 0, Status = 0, Abort Reason = 0
    DG_SET_CURRENT_ID shareddg
    DG_SET_CURRENT shareddg
    DG_GETCFG_ID 0xdde49f shareddg
    DG_GETCFG_NAME 0xdde49f shareddg
    DG_SET_CURRENT_ID shareddg
    DG_SET_CURRENT shareddg
    DG_SET_CURRENT_ID shareddg
    DG_SET_CURRENT shareddg
    DG_GETCFG_ALL 0x420
    DG_GETCFG_ALL 0x420
    VOL_TRANS ds4700-0_7 ds4700-0_3 ds4700-0_
    DG_GET_DEFAULT <no request data>
    CMDSHIP_REQUEST Command Shipped = /usr/sbin/vxassist -g
```

```
shareddg make shared-vol1 200M
Default dg = nodg
DROPPED <no request data>
```

要查找在从属节点上发出此请求的实用程序，请输入以下命令：

```
# egrep -n PID cmdlog | egrep Clid
```

在本示例中，输入以下命令：

```
# egrep -n 589906 cmdlog | egrep 5302
```

```
7310#: 5302, 589906, Thu Jul 15 06:30:14 2010 /usr/sbin/vxassist -g
```

该示例输出表明命令日志第 7310 行上存在匹配项。通过检查命令日志中的第 7310 行和 7311 行发现，系统对 shareddg 磁盘组运行过 vxassist make 命令：

```
# sed -e '7310,7311!d' cmdlog
```

```
# 5302, 589906, Thu Jul 15 06:30:14 2010
/usr/sbin/vxassist -g shareddg make shared-vol1 200M
```

要确定从属节点的集群监视器 nodeid (CM nid)，请输入以下命令：

```
# /etc/vx/bin/vxclustadm nidmap
```

如果该命令使用磁盘访问 (DA) 名称，传送的命令会将 DA 名称转换为唯一磁盘 ID (UDID) 或磁盘介质 (DM) 名称。在 CVM 主节点上，vxconfigd 日志将显示接收到的命令的条目。要在主节点上确定自从属节点接收到的命令，请输入以下命令：

```
# egrep CMDSHIP_REQUEST /var/adm/messages
```

注意： vxconfigd 消息记录到的文件可能随重定向消息的位置而变化。

在本示例中，以下接收到的命令将记录在主节点上的 vxconfigd 日志中：

```
07/15 06:29:02: V-5-1-0 receive_cmdship_message:
CMDSHIP_REQUEST: Received command:
Text - /usr/sbin/vxassist -g shareddg make shared-vol1 200M len = 53
CLID = 5302 SlaveID = 0 Defaultdg = nodg
```

CVM 主节点将执行接收到的命令并将响应发送到从属节点。

要查找主节点发送到从属节点的响应，请在主节点上输入以下命令：

```
# egrep CMDSHIP_RESPONSE translog | egrep Slave-Clid
```

在本示例中，输入以下命令查找主节点发送的响应：

```
# egrep CMDSHIP_RESPONSE translog | egrep 5302
Thu Jul 15 06:29:03 2010
Clid = 27741, PID = 475212, Part = 0, Status = 0, Abort Reason = 0
  CMDSHIP_RESPONSE  SlaveCLID = 5302 SlaveCMID = 0
    ExitCode = 12 Flags = 1 stdoutlen = 0 stderrlen = 98  Response =

VxVM vxassist ERROR V-5-1-10127 creating volume shared-voll:

Record already exists in disk group
DROPPED <no request data>
```

未启用命令完成

如果 **Tab** 键未自动完成命令，请检查以下几点：

- 该命令必须位于支持的命令列表中。
有关支持的命令列表，请参见产品版本说明。
- 该 shell 必须为 **bash 2.4** 版或更高版本。

如果满足上述要求，则可能是 **bash** 完成条目已从 **bashrc** 或 **bash_profile** 中删除。

启用命令完成

- ◆ 运行以下命令：

```
# ./etc/bash_completion.d/vx_bash
```


备份和还原磁盘组配置

本章节包括下列主题：

- [关于磁盘组配置备份](#)
- [备份磁盘组配置](#)
- [还原磁盘组配置](#)

关于磁盘组配置备份

备份和还原磁盘组配置的功能用于备份和还原 Veritas Volume Manager (VxVM) 磁盘组以及在磁盘组内配置的 VxVM 对象（例如卷）的所有配置数据。利用此功能，您可以恢复损坏的、以元数据形式存储在 VM 磁盘专用区域中的磁盘组配置。当还原磁盘组配置并启用卷后，您不必从备份介质中还原，公共区域中的用户数据就可以使用了。

警告： 备份和还原实用程序仅对 VxVM 配置数据起作用。它们不备份或还原卷或其他 VxVM 对象中包含的任何用户数据或应用程序数据。如果对磁盘使用 `vxdiskunsetup` 和 `vxdisksetup`，并指定与配置备份中不同的属性，可能会损坏公共区域及其包含的所有数据。

`vxconfigbackupd` 后台驻留程序监视对 VxVM 配置所做的更改，并在大约一小时后自动记录发生的任何配置更改。`vxconfigbackup` 和 `vxconfigrestore` 这两个实用程序可以用来备份和还原磁盘组的 VxVM 配置。

导入磁盘组时，如果 `vxconfigd` 日志中存在以下任何错误，则说明磁盘组配置和/或磁盘专用区域头文件已损坏：

```
VxVM vxconfigd ERROR V-5-1-569 Disk group group,Disk disk:Cannot auto-import group: reason
```

该错误的原因通常是以下项目之一：

```
Configuration records are inconsistent
Disk group has no valid configuration copies
Duplicate record in configuration
Errors in some configuration copies
Format error in configuration copy
Invalid block number
Invalid magic number
```

如果 VxVM 因磁盘出错而无法更新磁盘组配置，它将禁用该磁盘组并显示以下错误：

```
VxVM vxconfigd ERROR V-5-1-123 Disk group group: Disabled by errors
```

如果出现这样的错误，可以在解决根本问题（如硬件故障或连接断开）后从备份还原磁盘组配置。

使用备份中的配置数据，您可以重新安装磁盘组中头文件损坏的 VxVM 磁盘的专用区域头文件，重新创建损坏的磁盘组配置，或者重新创建磁盘组和磁盘组中的 VxVM 对象。如果原来的系统不可用，您还可以使用配置数据在其他系统上重新创建磁盘组。

注意：要还原磁盘组配置，所使用的物理磁盘必须与在执行备份时在磁盘组中配置的物理磁盘相同。

请参见第 72 页的“[备份磁盘组配置](#)”。

请参见第 73 页的“[还原磁盘组配置](#)”。

备份磁盘组配置

VxVM 使用磁盘组配置后台驻留程序来监视磁盘组的配置，并在配置发生变化后一小时进行备份。执行此操作的目的是为了避免在较短的时间间隔内触发多次备份处理。默认情况下，系统将保留最近的五个备份。备份放置在以下位置：`/etc/vx/cbr/bk/diskgroup.dgid/bkp_YYYYMMDD_HHMMSS/`。此格式可帮助您根据备份时间从特定备份还原磁盘组。必要时您还可以通过运行 `vxconfigbackup` 命令来备份磁盘组配置。

以下文件记录磁盘组配置信息：

`/etc/vx/cbr/bk/diskgroup.dgid/dgid.dginfo` 磁盘组信息。

`/etc/vx/cbr/bk/diskgroup.dgid/dgid` 磁盘属性。
`.diskinfo`

```
/etc/vx/cbr/bk/diskgroup.dgid/dgid      二进制配置副本。  
.binconfig
```

```
/etc/vx/cbr/bk/diskgroup.dgid/dgid.cfgrec vxprint -m 格式的配置记录。
```

此处的 *diskgroup* 表示磁盘组名称，*dgid* 表示磁盘组 ID。如果要在其他系统上重新创建磁盘组，请将这些文件复制到该系统。

警告： 请注意不要重写目标系统上由该系统中的磁盘组使用的任何文件。

备份磁盘组配置

◆ 键入以下命令：

```
# /etc/vx/bin/vxconfigbackup [-f] [-l directory] [[diskgroup ...] |  
[dgid ...]]
```

diskgroup(s) 可以按名称或 ID 指定磁盘组

-f -f 选项用于强制完全备份

-l -l 选项用于为备份配置文件位置（而非默认位置
/etc/vx/cbr/bk）指定一个目录。

要备份所有磁盘组，请使用此版本的命令：

```
# /etc/vx/bin/vxconfigbackup [-f] [-l directory]
```

请参见 `vxconfigbackup(1M)` 手册页。

还原磁盘组配置

可以使用 `vxconfigrestore` 实用程序从磁盘组的配置备份中还原或重新创建磁盘组。还原过程包括预提交操作，后跟提交操作。在预提交阶段，您可以检查将从备份中还原的磁盘组配置。在您选择提交更改后，实际的磁盘组配置才永久还原。

警告： 在执行还原时，磁盘组中的任何磁盘或 VxVM 对象不能被任何应用程序打开或使用。

在预提交阶段，您可以选择是否重新安装损坏的磁盘头文件。如果有任何磁盘的专用区域头文件无效，就必须重新安装受影响的磁盘的头文件，否则还原可能无法进行。

请参见 `vxconfigrestore(1M)` 手册页。

执行预提交操作

- ◆ 请使用以下命令执行磁盘组配置状态的预提交分析，并在头文件已损坏的磁盘上重新安装磁盘头文件：

```
# /etc/vx/bin/vxconfigrestore -p [-l directory] \  
{diskgroup | dgid}
```

可以通过名称或 ID 来指定磁盘组。

使用 `-l` 选项可以为备份配置文件指定一个默认位置 `/etc/vx/cbr/bk` 以外的目录。

请参见第 72 页的“备份磁盘组配置”。

指定不重新安装磁盘头文件

- ◆ 键入以下命令：

```
# /etc/vx/bin/vxconfigrestore -n [-l directory] \  
{diskgroup | dgid}
```

在预提交阶段，可以使用 `vxprint` 命令检查所还原的磁盘组将具有的配置。可以选择继续提交更改并还原磁盘组配置，也可以在做出任何永久更改之前取消还原。

在预提交阶段放弃还原

- ◆ 键入以下命令：

```
# /etc/vx/bin/vxconfigrestore -d [-l directory] \  
{diskgroup | dgid}
```

执行提交操作

- ◆ 要提交还原磁盘组配置所需的更改，请使用以下命令：

```
# /etc/vx/bin/vxconfigrestore -c [-l directory] \  
  {diskgroup | dgid}
```

注意：在预提交和提交状态之间，不应尝试任何导致磁盘组配置发生更改的操作。这可能会导致意外行为。用户应放弃还原，或者提交此操作。

如果没有重新安装任何磁盘头文件，则磁盘专用区域中的配置副本会从为磁盘组保存的最新二进制配置副本进行更新。

如果重新安装了磁盘头文件，则会使用磁盘属性的已保存副本重新创建其专用区域和公共区域。系统还会为这些磁盘分配新的磁盘ID。然后，使用磁盘组的备份配置记录重新创建磁盘组中的VxVM对象。此过程还将在磁盘组中创建新配置副本。

卷在后台同步。对于大卷的配置，执行同步可能需要一些时间。可以使用`vxtask -l list`命令监视此操作的进度。

正在使用或布局已更改的磁盘不会包括在还原过程中。

如果备份是从共享磁盘组创建的，`vxconfigrestore`命令会将其还原为专用磁盘组。在还原磁盘组后，请运行以下命令以便共享磁盘组。

共享磁盘组

- 1 逐出磁盘组：

```
# vxdg deport dg_name
```

- 2 将磁盘组作为共享磁盘组导入：

```
# vxdg -s import dg_name
```

解决磁盘组的冲突备份

有时，系统上的磁盘已更换，磁盘组可能有几个冲突备份。在这种情况下，`vxconfigrestore`命令会显示如下信息：

```
VxVM vxconfigrestore ERROR V-5-1-6012 There are two backups that  
have the same diskgroup name with different diskgroup id :  
1047336696.19.xxx.veritas.com  
  
1049135264.31.xxx.veritas.com
```

解决办法是通过 ID（而不是名称）来指定要执行还原的磁盘组。备份文件 `/etc/vx/cbr/bk/diskgroup.dgid/ dgid.dginfo` 包含记录备份执行时间的时间戳。

下面是来自一个备份文件的示例，其中显示了时间戳和磁盘组 ID 信息：

```
TIMESTAMP
Tue Apr 15 23:27:01 PDT 2003
.
.
.
DISK_GROUP_CONFIGURATION
Group:      mydg
dgid: 1047336696.19.xxx.veritas.com
.
.
.
```

借助时间戳信息确定哪个备份包含相关信息，然后使用 `vxconfigrestore` 命令指定磁盘组 ID（而不是磁盘组名称）来还原配置。

对磁盘组导入问题进行故障排除

本章节包括下列主题：

- 清除非克隆磁盘的 `udid_mismatch` 标志

清除非克隆磁盘的 `udid_mismatch` 标志

当安装或升级新的 Veritas Volume Manager (VxVM) 软件包或阵列支持库 (ASL) 软件包后，对 ASL 进行更改可能会导致在非克隆磁盘上设置 `udid_mismatch` 标志或 `clone_disk` 标志。VxVM 使用这些标志指示 LUN 的硬件快照或副本。

如果磁盘是非克隆磁盘，则此行为可能会导致磁盘组导入期间出错，具体情况取决于导入标志。在导入磁盘组时，VxVM 可能会跳过克隆磁盘。此外，导入可能会因磁盘组同时包含克隆和非克隆磁盘而失败。

如果磁盘是非克隆磁盘，您必须手动清除该磁盘上的 `udid_mismatch` 标志，然后才能继续导入磁盘组，

注意：在集群中，对同一节点执行下列所有步骤。

从磁盘清除 `udid_mismatch` 标志

- 1 检索显示 `udid_mismatch` 标志或 `clone_disk` 标志的已刷新磁盘列表。使用下列方法之一：

- 运行以下命令：

```
# vxdisk scandisks
```

```
# vxdisk list | egrep "udid_mismatch|clone_disk"
```

■ 或者，运行下面一条命令：

```
# vxdisk -o alldgs list | egrep "udid_mismatch|clone_disk"
```

2 如果磁盘属于已导入磁盘组，则逐出该磁盘组。

```
# vxdg deport dgname
```

3 清除步骤 1 中标识的所有非克隆磁盘上的 `udid_mismatch` 标志。

```
# vxdisk updateudid diskname
```

4 清除非克隆磁盘上的 `clone` 标志。

```
# vxdisk set diskname clone=off
```

5 导入磁盘组（如果在步骤 2 中将其逐出）

```
# vxdg [-s] import dgname
```

从 CDS 错误恢复

本章节包括下列主题：

- [CDS 错误代码和恢复操作](#)

CDS 错误代码和恢复操作

表 9-1 列出了 CDS 错误代码和所需的操作。

表 9-1 错误代码和所需操作

错误编号	消息	操作
329	不能将非 CDS 磁盘组与 CDS 磁盘组结合	将非 CDS 磁盘组更改为 CDS 磁盘组（或以相反顺序），然后重试结合操作。
330	磁盘组用于不同的平台	将磁盘组导入到正确平台上。它无法导入到此平台上。
331	卷具有与 CDS 不兼容的日志	要获得兼容 CDS 的日志，如果卷当前处于活动状态，您需要停止卷，然后启动卷。卷成功启动之后，重试设置磁盘组的 CDS 属性。
332	用于 CDS 的许可证过期或不可用	从 Symantec 获取许可证，以开始使用 CDS 磁盘组。

错误编号	消息	操作
333	不能将非 CDS 磁盘放入 CDS 磁盘组	<p>执行以下操作之一：</p> <ul style="list-style-type: none"> ■ 向其他非 CDS 磁盘组的另一磁盘组添加磁盘。 ■ 将磁盘重新初始化为 CDS 磁盘，以便它可以被添加到 CDS 磁盘组。 ■ 将 CDS 磁盘组更改到非 CDS 磁盘组，然后添加磁盘。
334	磁盘组对齐值与 CDS 不兼容	将磁盘组的对齐值更改为 8k，然后重试设置磁盘组的 CDS 属性。
335	子磁盘长度违反磁盘组对齐值	确保子磁盘长度值是 8K 的倍数。
336	子磁盘偏移违反磁盘组对齐值	确保子磁盘偏移值是 8K 的倍数。
337	子磁盘 plex 偏移违反磁盘组对齐值	确保子磁盘 Plex 偏移值是 8K 的倍数。
338	Plex 条带宽度与磁盘组对齐值冲突	确保 Plex 条带宽度值是 8K 的倍数。
339	卷或日志长度与磁盘组对齐值冲突	<p>确保卷的长度值是 8K 的倍数。</p> <p>对于日志，将 <code>dgalignment_checking</code> 属性值设置为 <code>round</code>。这可以确保日志长度静默地舍入至有效值。</p>
340	最后一个磁盘介质偏移值与磁盘组对齐值冲突	在升级之前重新结合 DM 记录。
341	磁盘组中的设备节点太多	如果未达到最大数量，增加磁盘组中允许的设备节点数。否则，您可能需要通过拆分磁盘组从磁盘组中删除卷。
342	映射长度对于当前日志长度太大	为 DRL/DCM 日志使用更小的映射长度，或者增加日志长度并重试。
343	卷日志映射对齐值与磁盘组对齐值冲突	删除 DRL/DCM 日志，然后在更改磁盘组对齐值后重新添加它。
344	磁盘设备无法用作 CDS 磁盘	

错误编号	消息	操作
345	磁盘组包含无法在此平台上导入的旧式 RVG	在创建 RVG 的平台上导入磁盘组。要在此平台上导入磁盘组，请首先在创建平台上删除 RVG。
346	max_autogrow 执行的缓存对象的自动放大量与磁盘组对齐值冲突。	确保缓存属性值是 8K 的倍数。
347	该磁盘组的用户事务被禁用	重试命令，因为同时它由执行的 vxcdsconvert 命令临时禁止。
348	磁盘正在使用中	与技术支持联系。

错误消息

本章节包括下列主题：

- [关于错误消息](#)
- [如何记录错误消息](#)
- [消息类型](#)

关于错误消息

Veritas Volume Manager (VxVM) 配置后台驻留程序 (vxconfigd)、VxVM 内核驱动程序 vxio 和各种 VxVM 命令可能会在控制台上显示信息性消息、故障消息和其他错误消息。这些消息可能涉及一些不常见的和难以解决的错误。

注意： 这里描述的某些错误消息可能不适用于您的系统。

您可能会发现参考 VxVM 命令和事务日志对了解错误发生的环境很有用。

请参见第 63 页的“[命令日志](#)”。

如何记录错误消息

VxVM 提供了将调试消息记录到文件的选项。这种记录有用的原因在于系统崩溃之前的任何消息输出都可以从日志文件找到（假定崩溃并未导致文件系统损坏）。

如果启用，则默认调试日志文件为 `/var/adm/ras/vxconfigd.log`。

若要准许将调试输出内容记录到默认的调试日志文件，请编辑 vxconfigd 的启动脚本。

请参见第 84 页的“[在启动脚本中配置日志记录](#)”。

或者，可以使用以下命令更改调试级别：

```
# vxdctl debug level [pathname]
```

有 10 种可能的调试日志记录级别：级别值从 0 到 9。1 级提供的详细信息最少，9 级提供的详细信息最多。0 级关闭日志。如果指定了路径名称，就使用此文件而不是默认的调试日志文件来记录调试输出。如果使用 `vxdctl debug` 命令，则新的调试日志记录级别与调试日志文件在下次重新启动 VxVM 配置后台驻留程序 `vxconfigd` 之前一直有效。

请参见 `vxdctl(1M)` 手册页。

请参见 `vxconfigd(1M)` 手册页。

在启动脚本中配置日志记录

要永久性启用日志文件或 `syslog` 日志记录，可以编辑用来启动 VxVM 配置后台驻留程序 `vxconfigd` 的 `/etc/init.d/vxvm-startup2` 脚本。

在启动脚本中配置日志记录

- ◆ 在 `vxconfigd` 中注释掉或取消注释以下行，以启用或禁用相应功能：

```
#opts="$opts -x log"
# messages to vxconfigd.log
#opts="$opts -x logfile=/foo/bar" # specify an alternate log file
#opts="$opts -x timestamp"
# timestamp console messages

# To turn on debugging console output, uncomment the following line.
# The debug level can be set higher for more output.  The highest
# debug level is 9.

#debug=1
# enable debugging console output
```

通常，取消对 `opts="$opts -x syslog"` 字符串的注释，这样 `vxconfigd` 在默认情况下使用 `syslog` 日志记录。在行首插入 `#` 字符将禁用 `vxconfigd` 的 `syslog` 记录。

默认情况下，引导时使用 `-x syslog` 选项启动 `vxconfigd`。这会将 `vxconfigd` 控制台信息重定向到 `syslog`。如果希望在从命令行重新启动 `vxconfigd` 时保留此行为，请包括 `-x syslog` 参数，因为重新启动 `vxconfigd` 时不保留它以前运行时所使用的选项设置。同样，任何要求重新启动 `vxconfigd` 的 Veritas Volume Manager 操作都可能不保留以前由选项设置指定的行为。

消息类型

VxVM 具有容错能力，不需要系统管理员干预就可以解决大多数问题。如果配置后台驻留程序 `vxconfigd` 识别出必须执行的操作，就会将所需的事务进行排队。VxVM 提供系统配置的原子更改：即要么事务处理全部完成，要么系统保留在似乎从未尝试过处理事务的状态。如果 `vxconfigd` 无法识别和解决系统问题，系统管理员就需要处理一些任务，即使用软件返回的诊断信息来解决问题。以下各节介绍了可能出现的错误消息编号和错误消息类型，并提供了常见错误列表、问题的可能起因的详细说明以及可采取的措施的建议。

错误消息一般具有以下格式：

```
product component severity message_number message_text
```

对于 Veritas Volume Manager，`product` 设置为 `VxVM`。`component` 可以是内核模块或驱动程序的名称（如 `vxvmp`）、配置后台驻留程序（如 `vxconfigd`）或命令（如 `vxassist`）。

按照对系统影响程度由重到轻的顺序，将错误消息划分为以下几种严重程度类型：

- | | |
|------|---|
| 告急 | 告急是一种严重事件，因为此类事件会在系统正常操作时停止系统。来自内核模块或设备驱动程序的告急信息表明，硬件问题或软件不一致问题非常严重，以致系统无法继续运行。操作系统可能还提供了 CPU 寄存器内容转储及堆栈跟踪，以帮助确定告急的原因。下面是此类信息的示例：

<code>VxVM vxio PANIC V-5-0-239 Object association depth overflow</code> |
| 致命错误 | 来自配置后台驻留程序（如 <code>vxconfigd</code> ）的致命错误消息表明 VxVM 操作中出现妨碍其运行的严重问题。下面是此类信息的示例：

<code>VxVM vxconfigd FATAL ERROR V-5-0-591 Disk group bootdg: Inconsistency -- Not loaded into kernel</code> |
| 错误 | 来自命令的错误消息表明无法正确执行请求的操作。下面是此类信息的示例：

<code>VxVM vxassist ERROR V-5-1-5150 Insufficient number of active snapshot mirrors in snapshot_volume.</code> |
| 警告 | 来自内核的警告消息表明非关键操作失败，原因可能是某些资源不可用或者无法执行该操作。下面是此类信息的示例：

<code>VxVM vxio WARNING V-5-0-55 Cannot find device number for boot_path</code> |

通知	通知消息表明发生了应该监视的错误。尽管可能需要以后采取措施纠正错误，但不必马上关闭系统。下面是此类消息的示例： <i>VxVM vxio NOTICE V-5-0-252 read error on object subdisk of mirror plex in volume volume (start offset, length length) corrected.</i>
信息性	信息性消息不表示错误，因此不需要执行任何操作。

唯一消息编号由以字母“V”开头的字母数字字符串组成。例如，在消息编号 v-5-1-3141 中，“V”表示这是 Veritas 产品错误消息，第一个数字字段 (5) 表示产品代码（这里是 VxVM），第二个字段 (1) 表示有关产品组件的信息，第三个字段 (3141) 是消息索引。错误消息的正文在唯一消息编号后列出。

消息

本节提供在 Veritas Volume Manager 操作过程中可能遇到的消息的列表。但是，该列表的内容不够详尽，并且第二个字段中可能包含不同于这里显示的命令、驱动程序或模块的名称。

本章对引发具体消息的情况或问题进行了详细说明，并尽可能提供了有助于找出并解决问题的恢复过程。

如果遇到产品错误消息，请记录消息文本前的唯一消息编号。从以下 URL 中搜索消息编号以查找有关该消息的信息：

<http://sort.symantec.com/>

通过电话或访问 Veritas 技术支持网站与 VERITAS 技术支持部门联系时，请务必提供相关的消息编号。Veritas 技术支持部门将使用此消息编号迅速确定有无适合您的 TechNote 或其他信息。

Veritas Volume Replicator 故障排除

本章节包括下列主题：

- [RLINK 连接出现问题后恢复](#)
- [发生配置错误后恢复](#)
- [主节点或辅助节点上的恢复](#)

RLINK 连接出现问题后恢复

本节介绍连接 RLINK 时可能遇到的错误。要解决 RLINK 连接问题，必须了解 RLINK 连接过程。

连接主节点 RLINK 和辅助节点 RLINK 的操作分两个步骤。第一步是挂接 RLINK，通过执行 `vradmin startrep` 命令来完成。第二步是连接 RLINK，通过主节点主机和辅助节点主机上的内核来完成。

执行 `vradmin startrep` 命令时，VVR 将进行多项检查以确保操作成功，如果操作成功，该命令会将 RLINK 的状态从 DETACHED/STALE 更改为 ENABLED/ACTIVE。然后，该命令会返回成功信息。

如果该命令成功，主节点上的内核将收到通知，指出 RLINK 已启用并开始向请求进行连接的辅助节点发送消息。在正常情况下，辅助节点将收到此消息并进行连接。然后，RLINK 的状态会从 ENABLED/ACTIVE 更改为 CONNECT/ACTIVE。

如果 RLINK 的状态在短时间内没有更改为 CONNECT/ACTIVE，则可能发生了阻止连接的问题。本节将介绍可能存在的多种原因。在控制台上，可能会显示指出问题的错误消息。

- 如果控制台上显示以下错误：

```
VxVM VVR vxrlink INFO V-5-1-5298 Unable to establish connection  
with remote host <remote_host>, retrying
```

确保 `vradmind` 后台驻留程序正在主节点主机和辅助节点主机上运行；否则，通过执行以下命令启动 `vradmind` 后台驻留程序：

```
# /etc/init.d/vras-vradmind.sh start
```

对于共享磁盘组中的 **RLINK**，确保对日志所有者启用了 **RLINK** 的虚拟 IP 地址。

- 如果没有很明显的错误消息说明，请在主节点主机和辅助节点主机上执行以下命令：

```
# vxprint -g diskgroup -l rlink_name
```

在输出信息中检查以下项：

每台主机的 `remote_host` 与另一台主机的 `local_host` 是否相同。

每台主机的 `remote_dg` 与另一台主机上的 **RVG** 的磁盘组是否相同。

每台主机的 `remote_dg_dgid` 与另一台主机上的 **RVG** 的 `dgid`（磁盘组 ID）是否相同，如 `vxprint -l diskgroup` 命令的输出所示。

每台主机的 `remote_rlink` 与另一台主机上对应的 **RLINK** 的名称是否相同。

每台主机的 `remote_rlink_rid` 与另一台主机上对应的 **RLINK** 的 `rid` 是否相同。

确保网络按照预期运行。网络问题可能会影响 **VVR**，例如阻止 **RLINK** 进行连接或降低性能。问题可能包括高延迟、低带宽、高冲突计数以及丢失过多的数据包。

- 对于专用磁盘组中的 **RLINK**，在每个主机上执行以下命令。
对于共享磁盘组中的 **RLINK**，在主节点和辅助节点上对日志所有者执行以下命令：

```
# ping remote_host
```

应该不会丢失数据包或丢失的数据包很少。为确保网络可以传输大型数据包，请在每个主机上对专用磁盘组中的 **RLINK** 执行以下命令。

对于共享磁盘组中的 **RLINK**，在主节点和辅助节点上对日志所有者执行以下命令：

```
# ping -s 8192 remote_host
```

数据包丢失情况应该与早期的 `ping` 命令大致相同。

- 在每个主机上执行 `vxiod` 命令，以确保存在活动的 I/O 后台驻留程序。如果输出是 `0 volume I/O daemons running`，请执行以下命令激活 I/O 后台驻留程序：

```
# vxiod set 10
```

- VVR 使用公知端口与其他主机建立通信。

执行以下命令显示端口号：

```
# vxprint -g diskgroup -l rlink_name
```

执行以下命令，确保输出中的心跳端口号与 vxprint 命令显示的端口匹配：

```
# vrport
```

发出以下命令以确认心跳端口是否已开放：

```
# netstat -an | grep port-number
```

其中，port-number 是心跳服务器正在使用的端口号，与 vrport 命令显示的端口号相同。

输出与以下内容类似：

```
udp4      0      0  *.port-number *.*
```

- 检查主站点和辅助站点上的 VVR 端口。

运行 vrport 实用程序，验证两端的端口是否相同。

检查所需的 VVR 端口是否已打开。检查 UDP 4145、TCP 4145、UDP 8199、TCP 8199 和匿名端口。输入以下命令：

```
# netstat -an | grep tcp | grep 4145
```

```
tcp      0      0  *.*.4145          *.*          LISTEN
tcp4     0      0  10.209.117.183.32790  10.209.117.184.4145  ESTABLISHED
```

```
# netstat -an | grep udp | grep 4145
```

```
udp      0      0  *.*.4145          *.*          *
```

```
# netstat -an | grep tcp | grep 8199
```

```
# netstat -an | grep tcp | grep 8199
```

```
tcp      0      0  *.8199            *.*          LISTEN
tcp4     0      0  10.209.117.183.32787  10.209.117.184.8199  ESTABLISHED
```

执行 telnet 测试以检查打开的端口。例如，要确定端口 4145 是否已打开，请输入以下内容：

```
# telnet <remote> 4145
```

- 使用 netstat 命令可检查 vradmind 后台驻留程序是否能够在主站点与辅助站点之间建立连接。

```
# netstat -an | grep 8199 | grep ESTABLISHED
```

```
tcp4     0      0  10.209.117.183.32787  10.209.117.184.8199  ESTABLISHED
```

如果不存在建立的连接，请检查 `/etc/hosts` 文件中是否包含主站点与辅助站点条目。将所有参与的系统名称和 IP 地址添加到每个系统中的 `/etc/hosts` 文件，或将该信息添加到名称服务的名称服务器数据库。

发生配置错误后恢复

当主节点 RVG 和辅助节点 RVG 的配置不同时，就会发生配置错误。主节点 RVG 中的每个数据卷必须在辅助节点 RVG 中有对应的数据卷，且大小必须完全相同；否则，复制将不会继续。如果卷集与 RDS 相关联，则主节点和辅助节点上的卷集配置也必须匹配。

可通过两种方式检测配置错误：

- 首次挂接 RLINK 时，将会检查辅助节点是否存在配置错误。如果发现任何错误，`attach` 命令就会失败并输出错误消息指出问题。通过更正配置错误并重试挂接即可修复该问题。
- 影响主节点或辅助节点上配置的更改可能导致辅助节点进入已设置 `secondary_config_err` 标志的 PAUSE 状态。通过更正配置错误并恢复 RLINK 即可修复该问题。

RLINK 挂接期间的错误

在 RLINK 挂接期间，VVR 会检查数据卷的配置中是否存在错误。如果 RDS 具有与 RVG 关联的卷集，VVR 还会检查卷集的配置中是否存在错误。

RLINK 挂接期间的数据卷错误

挂接 RLINK 时，VVR 会检查对于与主节点 RVG 相关联的每个数据卷，辅助节点 RVG 是否有一个大小相同、映射到主节点上其对应项的相关联数据卷。以下示例说明了可能会发生各种问题的挂接尝试，以及如何修复这些问题。在执行 `attach` 操作之前，主节点具有以下配置：

TY	Name	Assoc	KSTATE	LENGTH	STATE
rv	hr_rvg	-	DISABLED	-	EMPTY
rl	rlk_london_hr_rvg	hr_rvg	DETACHED	-	STALE
v	hr_dv01	hr_rvg	ENABLED	12800	ACTIVE
pl	hr_dv01-01	hr_dv01	ENABLED	12800	ACTIVE
sd	disk01-05	hr_dv01-01	ENABLED	12800	-

v	hr_dv02	hr_rvg	ENABLED	12800	ACTIVE
pl	hr_dv02-01	hr_dv02	ENABLED	12880	ACTIVE
sd	disk01-06	hr_dv02-01	ENABLED	12880	
v	hr_dv03	hr_rvg	ENABLED	12880	ACTIVE
pl	hr_dv03-01	hr_dv03	ENABLED	12880	ACTIVE
sd	disk01-07	hr_dv03-01	ENABLED	12880	-
v	hr_srl	hr_rvg	ENABLED	12880	ACTIVE
pl	hr_srl-01	hr_srl	ENABLED	12880	ACTIVE
sd	disk01-08	hr_srl-01	ENABLED	12880 0	-

辅助节点具有以下配置：

TY	Name	Assoc	KSTATE	LENGTH	STATE
rv	hr_rvg	-	ENABLED	-	- ACTIVE
rl	rlk_seattle_hr_rvg	hr_rvg	ENABLED	-	- ACTIVE
v	hr_dv01	hr_rvg	ENABLED	12700	- ACTIVE
pl	hr_dv01-01	hr_dv01	ENABLED	13005	- ACTIVE
sd	disk01-17	hr_dv01-01	ENABLED	13005	0 -
v	hr_dv2	hr_rvg	ENABLED	12880	- ACTIVE
pl	hr_dv02-01	vol2	ENABLED	13005	- ACTIVE
sd	disk01-18	hr_dv02-01	ENABLED	13005	0 -
v	hr_srl	hr_rvg	ENABLED	12880	- ACTIVE
pl	hr_srl-01	hr_srl	ENABLED	13005	- ACTIVE
sd	disk01-19	hr_srl-01	ENABLED	13005	0 -

请注意，辅助节点上卷 `hr_dv01` 的大小比较小，`hr_dv2` 的名称有误（必须是 `hr_dv02`）并且缺少 `hr_dv03`。尝试使用 `attach` 命令将主节点 `RLINK` 挂接到此辅助节点失败。

```
# vxrlink -g hrdg -f att rlk_london_hr_rvg
```

将显示以下消息：

```
VxVM VVR vxrlink INFO V-5-1-3614 Secondary data volumes detected
with rvg hr_rvg as parent:
VxVM VVR vxrlink ERROR V-5-1-0 Size of secondary datavol hr_dv01
(len=12700) does not match size of primary (len=12800)
VxVM VVR vxrlink ERROR V-5-1-3504 primary datavol hr_dv02 is not
mapped on secondary, yet
VxVM VVR vxrlink ERROR V-5-1-3504 primary datavol hr_dv03 is not
mapped on secondary, yet
```

要修复该问题，请在辅助节点上发出以下命令：

- 1 调整数据卷 `hr_dv01` 的大小：

```
# vradmin -g hrdg resizevol hr_rvg hr_dv01 12800
```

- 2 将数据卷 `hr_dv2` 重命名为 `hr_dv02`：

```
# vxedit -g hrdg rename hr_dv2 hr_dv02
```

- 3 关联与主节点数据卷 `hr_dv03` 大小相同的新卷 `hr_dv03`。

```
# vxassist -g hrdg make hr_dv03 12800
# vxvol -g hrdg assoc hr_rvg hr_dv03
```

或者，通过更改主节点以匹配辅助节点或结合上述两种方法都可以修复该问题。当主节点和辅助节点匹配时，重试挂接。

在主节点上：

```
# vxrlink -g hrdg -f att rlk_london_hr_rvg
```

```
VxVM VVR vxrlink INFO V-5-1-3614 Secondary data volumes detected
with rvg hr_rvg as parent:
VxVM VVR vxrlink INFO V-5-1-0 voll: len=12800 primary_datavol=hr_dv01
VxVM VVR vxrlink INFO V-5-1-0 voll: len=12800 primary_datavol=hr_dv02
VxVM VVR vxrlink INFO V-5-1-0 voll: len=12800 primary_datavol=hr_dv03
```

RLINK 挂接期间的卷集错误

如果卷集与 RDS 相关联，则主节点上的卷集名称必须与辅助节点上的卷集名称相同，且卷集的组件卷必须具有相同的配置。

挂接 RLINK 时，VVR 会检查对于与主节点 RVG 相关联的每个卷集，辅助节点 RVG 是否有一个名称相同的相关联卷集。此外，VVR 还会检查主节点上的卷集是否与辅助节点上的卷集具有相同的组件卷（名称、长度和索引相同）。（如果主节点上的组件卷和辅助节点上的组件卷已如独立卷映射，则其卷名可以不同。）如果辅助节点上不存在任何组件卷或组件卷名称、长度或索引不匹配，则 RLINK 挂接命令会失败并返回相应的错误消息。

请参见第 95 页的“修改 RVG 期间出现的卷集配置错误”。

如果辅助节点上不存在卷集，但辅助节点上存在所有组件卷，并且都有正确的名称和长度，则 VVR 会在辅助节点上创建卷集并将其关联到 RDS。这不会导致配置错误。

修改 RVG 期间出现错误

在初次设置和挂接辅助节点 RLINK 之后，如果不正确的修改（如添加、重定大小和重命名卷）导致主节点上和辅助节点上的卷不匹配，则会产生配置错误。如果 RVG 有相关联的卷集，则对卷集的修改也会导致配置错误。这些修改包括不正确地添加、删除或重命名相关联卷集的组件卷；在主节点上和辅助节点上添加具有不同索引的组件卷；或重命名相关联的卷集。

当修改 RVG 导致配置错误时，受影响的 RLINK 会进入已设置 `secondary_config_err` 标志的 PAUSED 状态。这会禁止复制到辅助节点，直到问题更正为止。

请在任一节点上运行 `vxrlink verify rlink` 命令，检查是否发生此错误。配置错误得到更正时，可以恢复受影响的 RLINK。

在修改 RVG 期间出现缺少数据卷错误

如果向主节点 RVG 添加数据卷，但辅助节点没有对应的数据卷，则 RLINK 状态将更改为 PAUSED 并且设置 `secondary_config_err` 标志。执行 `vxrlink verify` 命令将生成以下内容：

在主节点上：

```
# vxrlink -g hrdg verify rlk_london_hr_rvg
RLINK          REMOTE_HOST    LOCAL_HOST    STATUS    STATE
rlk_london_hr_rvg  london        seattle      ERROR    PAUSE
ERROR: hr_dv04 does not exist on secondary (london)
```

在辅助节点上：

```
# vxlink -g hrdg verify rlk_seattle_hr_rvg
RLINK          REMOTE_HOST    LOCAL_HOST    STATUS
STATE
rlk_seattle_hr_rvg  seattle        london        ERROR
PAUSE
ERROR: hr_dv04 does not exist on secondary (local host)
```

要更正此问题，请在辅助节点上创建并关联 `hr_dv04`，或者使 `vol04` 与主节点分离，然后恢复辅助节点 `RLINK`。要恢复辅助节点 `RLINK`，请使用 `vradmin resumerep rvg_name` 命令。

如果主节点上的 `hr_dv04` 包含有效数据，请将其内容复制到辅助节点上的 `hr_dv04` 中，然后再将该卷与辅助节点 `RVG` 关联。

在修改 RVG 期间出现数据卷不匹配错误

如果增加了主节点数据卷大小，但未增加辅助节点数据卷大小，则会发生配置错误。

在主节点上：

```
# vxassist growby hr_dv04 100
# vxlink -g hrdg verify rlk_london_hr_rvg
RLINK          REMOTE_HOST    LOCAL_HOST    STATUS    STATE
rlk_london_hr_rvg  london         seattle       ERROR     PAUSE
ERROR: hr_dv04 too small (12800). Primary is 12900
```

在辅助节点上：

```
# vxlink -g hrdg verify rlk_seattle_hr_rvg
RLINK          REMOTE_HOST    LOCAL_HOST    STATUS    STATE
rlk_seattle_hr_rvg  seattle        london        ERROR     PAUSE
ERROR: hr_dv04 too small (12800). Primary is 12900
```

要更正此问题，请增加辅助节点数据卷的大小，或者减小主节点数据卷的大小：

```
# vradmin -g hrdg resizevol hr_rvg hr_dv04 12900
```

重定数据卷大小后，通过在 RDS 中的任意主机上执行以下命令来恢复辅助节点 `RLINK`：

```
# vradmin -g hrdg resumerep hr_rvg
```

在修改 RVG 期间出现数据卷名称不匹配错误

如果在主节点上重命名了某数据卷，但未在辅助节点上执行该操作，则会发生配置错误且 RLINK 将断开连接。请使用 `vxprint -lP` 命令查看 RLINK 标志。如果设置了 `secondary_config_err` 标志，请使用以下命令之一来确定是否存在数据卷名称不匹配错误。

在主节点上：

```
# vxrlink -g hrdg verify rlk_london_hr_rvg
RLINK          REMOTE_HOST LOCAL_HOST      STATUS      STATE
rlk_london_hr_rvg london      seattle          ERROR       PAUSE
ERROR: hr_dv04 on secondary has wrong primary_datavol name (hr_dv04,
should be hr_dv05)
```

在辅助节点上：

```
# vxrlink -g hrdg verify rlk_seattle_hr_rvg
RLINK          REMOTE_HOST LOCAL_HOST      STATUS      STATE
rlk_seattle_hr_rvg seattle      london          ERROR       PAUSE
ERROR: hr_dv04 on secondary has wrong primary_datavol name (hr_dv04,
should be hr_dv05)
```

要修复此错误，请执行以下操作之一：

- 重命名主节点数据卷或辅助节点数据卷，然后使用 `vradmin resumerep rvg_name` 命令恢复 RLINK。
或
- 将辅助节点数据卷上的 `primary_datavol` 字段设置为引用主节点数据卷的新名称（如下所示），然后使用 `vradmin resumerep rvg_name` 命令恢复 RLINK。
在辅助节点上：

```
# vxedit -g hrdg set primary_datavol=hr_dv05 hr_dv04
```

其中 `hr_dv05` 是主节点上的新名称

修改 RVG 期间出现的卷集配置错误

如果卷集与 RDS 相关联，则辅助节点上卷集的名称只有与主节点上卷集的名称相同，才能进行复制。此外，辅助节点上的卷集必须与主节点上的卷集具有相同的组件卷（名称、长度和索引相同）。

如果在主节点上调整了组件卷的大小，但没有在辅助节点上进行调整，则会导致数据卷不匹配的错误。请调整该卷的大小并恢复制。

请参见第 94 页的“在修改 RVG 期间出现数据卷不匹配错误”。

首次挂接 RLINK 时，将会检查辅助节点是否存在配置错误。如果发现任何错误，`vradmin startrep` 命令就会失败并输出错误消息指出问题。请更正配置错误，然后重试该命令。

修改卷集或其组件卷时，也可能发生配置错误。请在任一节点上运行 `vxrlink verify rlink` 命令，检查是否发生此错误。更正配置错误，然后恢复 RLINK。

卷集名称不匹配错误

如果主节点和辅助节点上的卷集名称不同，将显示以下错误：

```
VSet name vset_name of secondary datavol vol_name does not match  
VSet name vset_name of primary datavol vol_name
```

要更正此问题，请使用以下命令在主节点或者辅助节点上重命名卷集：

```
# vxedit -g diskgroup rename vset_name new_vset_name
```

卷索引不匹配错误

如果主节点卷集和辅助节点卷集上的组件卷的索引不同，将显示以下错误：

```
VSet index (index_name) of secondary datavol vol_name does not  
match VSet index (index_name) of primary datavol vol_name
```

要更正此问题，请在辅助节点上执行以下步骤：

- 1 使用以下命令从卷集中分离每个卷：

```
# vxvset -g diskgroup rmvol vset_name compvol_name
```

删除最后一个卷时，卷集也会被删除。

- 2 使用以下命令创建卷集：

```
# vxvset -g diskgroup -o index make vset_name \  
compvol_name index
```

- 3 将其余每个卷关联到卷集，并使用以下命令在主节点上指定相应卷的索引：

```
# vxvset -g diskgroup -o index addvol vset_name \  
compvol_name index
```

组件卷不匹配错误

如果仅将数据卷从主节点 RVG 上的卷集中删除，或者仅将数据卷添加到辅助节点 RVG 上的卷集中，将显示以下错误：

```
Secondary datavol vol_name is associated to VSet vol_name  
whereas primary datavol is not associated to any Vset
```

类似地，如果仅将数据卷从辅助节点 RVG 上的卷集中删除，或者仅将数据卷添加到主节点 RVG 上的卷集中，将显示以下错误：

```
Primary datavol vol_name is associated to VSet whereas secondary  
datavol vol_name is not associated to any Vset
```

要更正此问题，请在辅助节点卷集或主节点卷集中添加或删除数据卷。主节点和辅助节点上的卷集应该有相同的组件卷。

要向卷集中添加数据卷，请执行以下操作之一：

- 向 RVG 中的卷集添加数据卷：

```
# vradmin -tovset vset_name addvol rvg_name vol_name
```

- 从 RVG 中的卷集删除数据卷：

```
# vradmin -fromvset vset_name delvol rvg_name vol_name
```

主节点或辅助节点上的恢复

本节介绍如何在发生各类灾难（例如主节点主机崩溃或者主节点或辅助节点数据卷中发生错误）后恢复。

关于从主节点主机崩溃中恢复

当主节点主机在发生故障后恢复时，VVR 会自动恢复 RVG 配置。当主节点恢复时，VVR 会恢复 RVG 中的主节点 SRL 和所有卷。SRL 和数据卷最近活动的相关信息会在 SRL 头中进行维护。VVR 使用该信息可加快重新启动时自动执行的恢复的速度。

主节点数据卷发生错误后恢复

如果对主节点数据卷的写入失败，将会分离该数据卷。RVG 继续像之前一样运行，提供对 RVG 中其他卷的访问。向故障卷写入会返回错误，但不会记录在 SRL 中。

RLINK 不受数据卷故障影响。如果 SRL 在卷发生错误时空置，这些更新会继续从 SRL 流向辅助节点 RLINK。已由应用程序完成但未写入卷的所有向故障卷的写入都保存在 SRL 中。这些写入在 SRL 中标记为挂起，当该卷恢复后将重放到该卷中。如果从备份恢复该卷并重新启动，这些写入将被丢弃。

如果数据卷的故障是永久性的，例如硬件损坏，则必须从备份恢复。发生此故障后恢复包括两部分操作：

- 从备份还原主节点数据卷
- 重新同步所有辅助节点 RLINK

如果 RVG 包含数据库，则出现故障的数据卷的恢复必须与数据库的恢复要求一致。数据库恢复顺序的详细信息确定同步辅助节点 RLINK 必须执行的操作。

下面给出了恢复过程的详细示例：

- 请参见第 98 页的“[示例 - 使用分离的 RLINK 进行恢复](#)”。
- 请参见第 99 页的“[示例 - 以最小修复工作量进行恢复](#)”。
- 请参见第 99 页的“[示例 - 通过迁移主节点进行恢复](#)”。

如果数据卷故障是由于临时中断（例如电缆断开连接）所致，并且您确定不存在永久性的硬件损坏，则可以在不将数据卷与 RVG 分离的情况下启动数据卷。SRL 中暂停的写入将重放到该数据卷。

请参见第 99 页的“[示例 - 发生临时 I/O 错误后恢复](#)”。

示例 - 使用分离的 RLINK 进行恢复

在此示例中，在主节点上开始恢复故障之前，将分离所有 RLINK。完成故障恢复（包括所有数据库恢复过程）后，必须使用主节点存储检查点同步所有 RLINK。

请在主节点上执行这些步骤。在此示例中，主节点主机是 `seattle`。

故障后恢复

- 1 分离所有 RLINK

```
# vxrlink -g hrdg det rlink_london_hr_rvg
```

- 2 修复数据卷。

如果可以通过修复数据卷的基础子磁盘来修复数据卷，则不需要从 RVG 中分离数据卷。如果通过分离发生故障的卷并在其位置关联一个新卷来解决问题，则必须在停止 RVG 时进行分离和关联。

- 3 确保在重新启动 RVG 前启动数据卷。

```
# vxvol -g hrdg start hr_dv01  
# vxrvg -g hrdg start hr_rvg
```

- 4 还原数据库。
- 5 使用块级别备份和检查点同步所有 RLINK。

示例 - 以最小修复工作量进行恢复

本示例采取最少的操作修复数据卷错误，同时保持所有 **RLINK** 处于挂接状态。在本示例中，从备份还原出现故障的卷数据以及数据库恢复是通过活动 **RLINK** 实现的。由于会复制主节点中的全部更改，因此复制完更改后，所有辅助节点都必须与主节点保持一致。此方法可能不总是实用，因为可能需要复制大量的数据。此外，还必须在要支持的每个目标数据库中对修复的数据卷进行仔细测试。

请在主节点上执行这些步骤。在此示例中，主节点主机是 `seattle`。

故障后恢复

- 1 停止 RVG。

```
# vxrvg -g hrdg stop hr_rvg
```

- 2 将出现故障的数据卷与 RVG 分离。

- 3 修复该数据卷或使用新卷。

如果可以通过修复数据卷的基础子磁盘来修复数据卷，则不需要从 RVG 中分离数据卷。如果通过分离发生故障的卷并在其位置关联一个新卷来解决问题，则必须在停止 RVG 时进行分离和关联。

- 4 将卷与 RVG 关联。

- 5 确保在重新启动 RVG 前启动数据卷。如果该数据卷未启动，则启动数据卷：

```
# vxvol -g hrdg start hr_dv01
```

- 6 启动 RVG：

```
# vxrvg -g hrdg start hr_rvg
```

- 7 还原数据库。

示例 - 通过迁移主节点进行恢复

一种替代性的恢复方法是：将主节点角色转移到辅助节点主机。

接管后，具有故障数据卷的原始主节点将不会变为 `acting_secondary`，直到恢复或分离故障数据卷为止。

示例 - 发生临时 I/O 错误后恢复

如果数据卷上的 I/O 错误是临时错误，并且您确定所有现有数据都保持不变，则可以启动数据卷，而不必将其与 RVG 分离。例如，如果 SCSI 电缆断开连接或存储器断电，在这种情况下，请执行下列步骤。

发生临时 I/O 错误后恢复

- 1 修复临时故障。
- 2 启动数据卷：

```
# vxvol -g hrdg start hr_dv01
```

SRL 中的所有未完成写入均被写入到数据卷。

主节点 SRL 卷错误清除和重新启动

如果访问主节点 SRL 时出错，将分离该 SRL 以及 RLINK。主节点和辅助节点 RLINK 的状态会更改为 STALE。RVG 状态不会更改，但是 RVG 会被置于 PASSTHRU 模式，这使得在修复该错误之前可以继续更新主节点卷。

请参见第 100 页的“关于 RVG PASSTHRU 模式”。

必须手动修复 SRL，然后将其与 RVG 关联。在修复 SRL 时，不能尝试向 RLINK 发送数据。在替换 SRL 后，必须完全同步所有 RLINK。挂接 RLINK 并对辅助节点执行完全同步。

在主节点 (seattle) 上：

发生主节点 SRL 错误后清除

- 1 从 RVG 中分离 SRL。

```
# vxvol -g hrdg dis hr_srl
```

- 2 修复或替换 SRL 卷。
- 3 确保先启动已修复的 SRL，然后再将其与 RVG 关联。如果修复后的 SRL 未启动，请启动它：

```
# vxvol -g hrdg start hr_srl
```

- 4 将新的 SRL 与 RVG 关联。关联新 SRL 后，RVG PASSTHRU 模式将不再显示在 `vxprint -lv` 命令的输出中。

```
# vxvol -g hrdg aslog hr_rvg hr_srl
```

- 5 完全同步辅助节点。

关于 RVG PASSTHRU 模式

通常，写入关联 RVG 的数据卷中的内容先转至 RVG 的 SRL 上，然后再转至 RLINK 和数据卷。如果由于访问错误导致主节点 SRL 曾经被分离，则主节点 RVG 将被置

于 PASSTHRU 模式。在 PASSTHRU 模式下，写入数据卷的内容将绕过 SRL，直接传递到底层数据卷。RLINK 不接收写入内容。在 RVG 上使用 `vxprint -l` 可查看是否设置了 `passthru` 标志。关联新的 SRL 将清除 PASSTHRU 模式，且辅助节点 RVGs 必须保持同步。

重新启动时主节点 SRL 卷出错

如果在重新启动期间主节点 SRL 出错，那么可能是包含 SRL 的磁盘或阵列仍未联机所致。因此，VVR 不会恢复 RVG，而是将其置于 PASSTHRU 模式。当 SRL 变得可用时，可执行下列命令来恢复 RVG 和 RLINK：

```
# vxrvg -g diskgroup recover rvg_name  
# vxrlink -g diskgroup recover rlink_name
```

在发生错误并成功恢复 RVG 之后，如果从 RVG 中分离卷，则可能会看到以下消息：

```
Because there could be outstanding writes in the SRL, the data volume  
being dissociated should be considered out-of-date and inconsistent
```

可以忽略此消息。

如果 SRL 永久丢失，请创建一个新的 SRL。

请参见第 102 页的“在 SRL 头出错后恢复”。

在这种情况下，有可能在旧 SRL 上已成功完成且为应用程序所确认的写入仍未刷新到数据卷，并且现已丢失。因此，在继续操作之前必须从备份中还原数据卷。因为这会导致完全重写数据卷，建议在还原操作完成之后分离 RLINK 并使其同步。

主节点 SRL 卷溢出恢复

由于主节点 SRL 的大小是有限的，对任何 RLINK 的更新活动的延迟停止可能会超出日志维护所有必需的更新历史记录以使 RLINK 最新的能力。发生这种情况时，有问题的 RLINK 会被标记为 STALE 并且需要手动恢复，然后复制才能继续。只能使用自动同步或块级别备份以及存储检查点使 STALE RLINK 保持最新。其他 RLINK、RVG 和 SRL 卷都会依然处于运行中。

可以设置 SRL 溢出保护以防止 SRL 溢出，这是默认设置。不会允许 RLINK 成为 STALE 状态，将启动 `dcm` 日志记录。当稍后通信链路未过载时，可以使用 `vradmin resync rvg` 命令以递增方式重新同步 RLINK。

主节点 SRL 头错误清理和恢复

主节点上发生的 SRL 头故障是一个严重错误。所有 RLINK 都会丢失，必须使用主节点存储检查点进行恢复。由于数据卷错误的相关信息保存在 SRL 头中，因此在发

生此错误的任何情况下，都无法保证数据卷的状态是正确的。为此，建议您为 SRL 设置镜像。

如果 SRL 头错误在正常操作期间发生，而且您在重新启动之前注意到该错误，则可以确定同样（同步）出现故障的所有数据卷的状态都将是 DETACHED。如果在 vxprint 命令显示将处于 DETACHED 状态的卷之前重新启动系统，可能会丢失所有出现故障的数据卷的状态。这两种情况都涉及多个错误，因此不太可能发生，但了解使用此类错误可以质疑主节点数据卷的状态是很重要的。

发生主节点 SRL 头错误时，将会继续执行到 RVG 的写入；但是，所有 RLINK 都将置于 STALE 状态。RVG 在 PASSTHRU 模式下运行。

在 SRL 头出错后恢复

如果要在 SRL 头出错后恢复，则需要从 RVG 中分离 SRL、修复 SRL 以及完全同步所有 RLINK。

在 SRL 头出错后恢复

- 1 停止 RVG。

```
# vxrvg -g hrdg stop hr_rvg
```

- 2 从 RVG 中分离 SRL。

```
# vxvol -g hrdg dis hr_srl
```

- 3 修复或还原 SRL。即使可通过修复基础子磁盘来解决该问题，也必须对 SRL 进行分离和重新关联以便初始化 SRL 头。

- 4 确保 SRL 已启动，然后重新关联 SRL：

```
# vxvol -g hrdg start hr_srl  
# vxvol -g hrdg aslog hr_rvg hr_srl
```

- 5 启动 RVG：

```
# vxrvg -g hrdg start hr_rvg
```

- 6 如果需要，请从备份中还原数据卷。同步所有 RLINK。

辅助节点数据卷错误清除和恢复

如果在访问辅助节点数据卷期间发生 I/O 错误，则该数据卷将自动从 RVG 分离且 RLINK 断开连接。主节点到辅助节点的后续连接尝试将失败，并显示一条消息，说

明辅助节点数据卷已停止。主节点不受影响，写入会继续记录到 SRL 中。当辅助节点数据卷的错误得以修复且该数据卷启动后，RLINK 将自动重新连接。

如果没有合适的主节点或辅助节点存储检查点，将同时在主节点和辅助节点上分离 RLINK，然后再同步 RLINK。

使用辅助节点存储检查点恢复

本节介绍如何使用辅助节点存储检查点实现辅助节点数据卷错误恢复。

在辅助节点 (london) 上：

- 1 修复出现故障的数据卷。如果可以通过修复基础子磁盘修复该问题，则不必分离数据卷。
- 2 确保启动该数据卷：

```
# vxvol -g hrdg start hr_dv01
```

- 3 将辅助节点存储检查点备份中的数据还原到所有卷中。如果所有卷都已从备份还原，则辅助节点将在同步过程中保持一致。通过执行以下命令还原 RLINK：

```
# vxrlink -g hrdg -c sec_chkpt restore rlk_seattle_hr_rvg
```

使用主节点存储检查点清除

在辅助节点 (london) 上：

- 1 如上所示，修复出现故障的数据卷。确保先启动数据卷，再继续：

```
# vxvol -g hrdg start hr_dv01
```

- 2 分离 RLINK 以启用对辅助节点数据卷的写入：

```
# vxrlink -g hrdg det rlk_seattle_hr_rvg
```

- 3 将主节点存储检查点备份中的数据还原到所有数据卷中。与从辅助节点存储检查点还原不同，主节点存储检查点数据必须加载到所有辅助节点数据卷，而不仅仅是出现故障的卷。如果可用主节点存储检查点尚不存在，请创建一个新的存储检查点。
- 4 重新挂接 RLINK。

```
# vxrlink -g hrdg att rlk_seattle_hr_rvg
```

在主节点 (seattle) 上：

先分离 RLINK，然后再使用以下命令从主节点存储检查点重新挂接：

```
# vxrlink -g hrdg det rlk_london_hr_rvg  
# vxrlink -g hrdg -c primary_checkpoint att rlk_london_hr_rvg
```

辅助节点 SRL 卷错误清除和恢复

只有在对 RLINK 进行原子恢复的过程中且 IBC 处于活动状态时，才能使用辅助节点 SRL。如果在辅助节点 SRL 的恢复过程中发生 I/O 错误，则恢复将失败，SRL 卷会被自动分离，且 RLINK 被强制设置为暂停状态。需要手动干预来修复此物理问题、重新挂接 SRL，然后恢复 RLINK。恢复完成时，将会自动重试对 RVG 的恢复，如果成功，则更新活动可继续。唯一的问题是如果在修复完成之前主节点 SRL 溢出，则需要完全同步。

如果在 SRL 的数据部分发生错误，则 RLINK 会被强制设置为 PAUSE 状态，同时设置 `secondary_paused` 标志。不会分离 SRL。

如果在 SRL 头中发生错误，则辅助节点 RVG 会被强制设置为 FAIL 状态，并将分离 SRL。

在辅助节点 (london) 上：

- 1 分离 SRL，对其进行修复，然后重新关联。即使可以通过修复基础子磁盘来修复此问题，分离和重新关联依然是必需的，原因是此序列操作将初始化 SRL 头。

```
# vxvol -g hrdg dis hr_srl
```

修复或替换 SRL。确保先启动 SRL，然后再关联：

```
# vxvol -g hrdg start hr_srl  
# vxvol -g hrdg aslog hr_rvg hr_srl
```

- 2 运行 RLINK 恢复操作以清除 `secondary_log_err` 标志。

```
# vxrlink -g hrdg resume rlk_seattle_hr_rvg
```

辅助节点 SRL 头错误清除和恢复

辅助节点上的 SRL 头故障会将辅助节点 RVG 置于失败状态，且将主节点和辅助节点上的 RLINK 状态设置为 PAUSE 状态。由于数据卷错误的相关信息保存在 SRL 头中，因此不能保证所有情况下数据卷状态都是正确的。如果在正常操作期间发生辅助节点 SRL 头故障，且您在重新启动之前注意到该故障，则同时发生故障的所有数据卷的状态都将为 DETACHED。如果在 `vxprint` 命令显示将处于 DETACHED 状态的卷之前重新启动系统，可能会丢失所有出现故障的数据卷的状态。这两种情况

都涉及多个错误，因此不太可能发生，但了解使用此类错误可以质疑辅助节点数据卷的状态是很重要的。

清除和恢复 SRL 头故障

- 1 分离 SRL 卷。

```
# vxvol -g hrdg dis hr_srl
```

- 2 修复 SRL 卷。即使可通过修复基础子磁盘来解决该问题，也必须对 SRL 卷进行分离和重新关联以便初始化 SRL 头。

- 3 启动 SRL 卷。然后，重新关联该卷。

```
# vxvol -g hrdg start hr_srl
```

```
# vxvol -g hrdg aslog hr_rvg hr_srl
```

4 启动 RVG。

```
# vxrvrg -g hrdg start hr_rvg
```

5 如果对数据卷的完整性没有疑问，则可恢复 RLINK。

```
# vxrlink -g hrdg resume rlk_seattle_hr_rvg
```

或

如果怀疑数据卷的完整性，且辅助节点存储检查点备份可用，请从辅助节点存储检查点还原。

```
# vxrlink -g hrdg det rlk_seattle_hr_rvg
# vxrlink -g hrdg -f att rlk_seattle_hr_rvg
# vxrlink -g hrdg -w pause rlk_seattle_hr_rvg
```

将辅助节点存储检查点备份数据还原到数据卷。

```
# vxrlink -g hrdg -c secondary_checkpoint restore \
rlk_seattle_hr_rvg
```

或

如果怀疑数据卷的完整性且没有辅助节点存储检查点可用，请使用块级备份和主节点存储检查点同步辅助节点。

此外，还可以使用自动同步。

```
# vxrlink -g hrdg det rlk_seattle_hr_rvg
```

在辅助节点上，将主节点存储检查点备份数据还原到数据卷。

```
# vxrlink -g hrdg -f att rlk_seattle_hr_rvg
```

在主节点 (seattle) 上：

```
# vxrlink -g hrdg -c primary_checkpoint att \
rlk_london_hr_rvg
```

重新启动时发生辅助节点 SRL 头错误

如果在重新启动后辅助节点 SRL 发生错误，将无法修复，即使 SRL 随后可用也一样。忽略以下消息：

```
VxVM VVR vxrvg ERROR V-5-1-0 RVG rvg_name cannot be recovered  
because SRL is not accessible. Try recovering the RVG after the  
SRL becomes available using vxrecover -s command
```

重置 SRL 卷

1 分离 SRL:

```
# vxvol -g hrdg -f dis srl
```

忽略以下消息:

```
VxVM vxvol WARNING V-5-1-0 WARNING: Rvg rvgname has not been  
recovered because the SRL is not available. The data volumes may  
be out-of-date and inconsistent  
VxVM vxvol WARNING V-5-1-0 The data volumes in the rvg rvgname  
cannot be recovered because the SRL is being dissociated.  
Restore the data volumes from backup before starting the applications
```

2 创建一个新 SRL 卷 *new_srl* 并继续, 如下所示:

```
# vxvol -g hrdg aslog rvg_name new_srl  
# vxrlink -g hrdg recover rlink_name  
# vxrlink -g hrdg -f att rlink_name  
# vxrvg -g hrdg start rvg_name
```

如果复制由于收到 IBC 而被冻结, 则 SRL 中的数据将丢失, 但不会指出发生此问题。要查看是否存在该问题, 请检查 `/var/adm/ras/conslog` 文件中的消息, 例如:

```
WARNING: VxVM VVR vxio V-5-0-259 Replication frozen for rlink  
<rlink>
```

如果这是 RLINK 的最后一条消息, 即如果没有后续消息说明复制已取消冻结, 则必须完全重新同步主节点 RLINK。

Dynamic Multi-Pathing 故障排除

- [12. Dynamic Multi-Pathing 故障排除](#)

Dynamic Multi-Pathing 故障排除

本章节包括下列主题：

- [升级到 DMP 6.0 后显示扩展属性](#)
- [降级阵列支持](#)

升级到 DMP 6.0 后显示扩展属性

当您从 Storage Foundation 5.1 版升级到 DMP 6.0 时，会发现功能有下列变化：

- 在 `vxdisk list` 输出中列出的设备名称不显示阵列卷 ID (AVID)。
- `vxdisk -e list` 输出不显示扩展属性。
- 主动/被动 (A/P) 或 ALUA 阵列被声明为主动/主动 (A/A)。

此行为可能是因为 LUN 由本机多径处理驱动程序 MPIO 控制。

检查 LUN 是否由本机多径处理驱动程序控制

- ◆ 检查以下命令的输出以确定 LUN 是否是 MPIO 设备：

```
# lsdev -Cc disk
```

可以将 LUN 从由本机多径处理驱动程序控制迁移到由 DMP 控制。

- 要使用 Veritas Volume Manager 迁移到 DMP，请参考《Veritas Storage Foundation 管理指南》中有关禁用 MPIO 的章节。
- 要使用 OS 本机卷支持迁移到 DMP，请参考《Veritas Dynamic Multi-Pathing 管理指南》中有关从 MPIO 迁移到 DMP 的章节。

降级阵列支持

单个文件集 `VRTSaslapm` 提供阵列支持，包括阵列支持库 (ASL) 和阵列策略模块 (APM)。Veritas Storage Foundation and High Availability Solutions 的每个主要版本均包括支持的 `VRTSaslapm` 文件集，它是在产品安装过程中安装的。在主要版本之间，Symantec 可能会通过 `VRTSaslapm` 文件集的更新来提供额外的阵列支持。

如果更新的 `VRTSaslapm` 文件集有问题，Symantec 会建议您降级到 ASL/APM 文件集的早期版本。您只能还原到已安装的 Veritas Storage Foundation and High Availability Solutions 版本所支持的文件集。为了在系统联机时执行降级，请不要删除安装的文件集，而是在新文件集的基础上安装该文件集的早期版本。此方法可防止安装 `VRTSaslapm` 文件集的多个实例。

使用以下方法可以使 `VRTSaslapm` 文件集降级。

在联机时使 ASL/APM 文件集降级

- ◆ 在以下命令中指定 `VRTSaslapm` 文件集的早期版本：

```
# installp -F -ad ./VRTSaslapm.bff VRTSaslapm
```

Veritas Storage Foundation Cluster File System High Availability 故障排除

- [13. Veritas Storage Foundation Cluster File System High Availability 故障排除](#)

Veritas Storage Foundation Cluster File System High Availability 故障排除

本章节包括下列主题：

- [关于 Veritas Storage Foundation Cluster File System High Availability 故障排除](#)
- [CFS 故障排除](#)
- [防护配置故障排除](#)
- [对 SFHA Solutions 集群中的 Cluster Volume Manager 进行故障排除](#)

关于 Veritas Storage Foundation Cluster File System High Availability 故障排除

本章中的信息可帮助您诊断可能出现的设置或配置问题。对于组件产品中出现的問題，可能需要参考相应的技术文档予以解决。

I/O 防护的故障排除信息也适用于 Veritas Storage Foundation Cluster File System High Availability (SFCFSHA) 的故障排除。

请参见第 147 页的[“I/O 防护故障排除”](#)。

CFS 故障排除

本节讨论 CFS 问题故障排除。

root 用户的 <library> 路径中的顺序不正确

如果 root 用户的 <library> 路径中的顺序不正确，则在更改集群文件系统或 RAC 集群中的主节点时，系统可能会挂起。

如果 root 用户的 <library> 路径在 /usr/lib 条目之前包含一个指向集群文件系统 (CFS) 文件系统的条目，则在尝试执行以下任务之一时系统可能会挂起：

- 更改 CFS 文件系统的主节点
- 在主节点上卸载 CFS 文件系统
- 在主节点上停止集群或服务组

此配置问题主要出现在在共享 CFS 文件系统上安装了 Oracle 二进制文件的 RAC 环境中。

下面是一个可能导致系统挂起的 <library path> 示例：

```
LIBPATH=/app/oracle/orahome/lib:/usr/lib:/usr/ccs/lib
```

在以上示例中， /app/oracle 是一个 CFS 文件系统，如果用户尝试更改该文件系统的主节点，系统将挂起。用户仍然能够对系统执行 ping 和 telnet 命令，但类似 ls 这样的简单命令不会响应。更改主节点过程中需要执行的初始步骤之一是在集群范围内冻结文件系统，随后立即执行 fsck 命令以重放意向日志。

由于 <library> 路径中的初始条目指向冻结的文件系统本身，因此 fsck 命令将进入死锁状态。实际上，依赖 <library> 路径的所有命令（包括 ls）从现在开始都会挂起。

建议用来纠正此问题的过程如下所示：将任何用户（尤其是 root 用户）的 <library> 路径中指向 CFS 文件系统的所有条目移动到列表末尾，位于 /usr/lib 条目之后。

这样，上面的 <library path> 示例将更改为以下形式：

```
LIBPATH=/usr/lib:/usr/ccs/lib:/app/oracle/orahome/lib
```

非 root 用户运行 CFS 命令时 CFS 命令可能会挂起 (2403263)

非 root 用户运行 CFS 命令时 CFS 命令可能会挂起。

Workaround

解决此问题

- ◆ 在非 root 用户会话中运行任何 CFS 命令之前，使用 halogin 命令保存身份验证信息。

当您运行 halogin 命令时，VCS 会将已加密的身份验证信息存储在用户主目录中。

防护配置故障排除

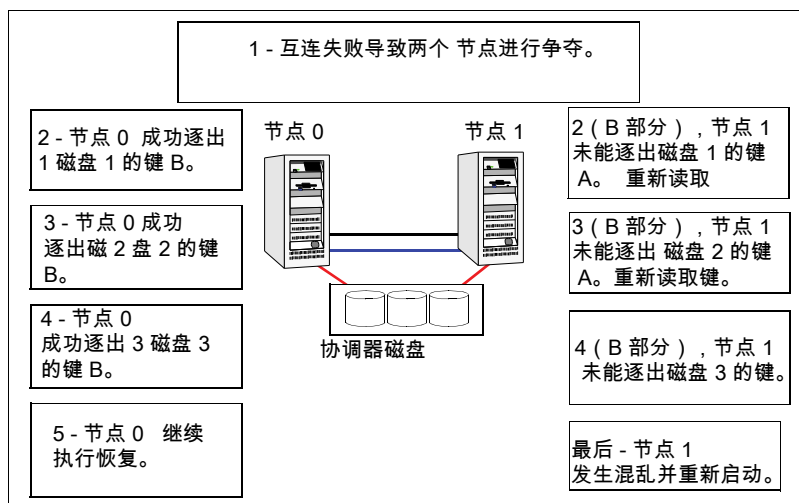
以下信息介绍了防护环境下的网络分裂。

请参见《Veritas Cluster Server 管理指南》。

已有网络分裂（裂脑）的示例

图 13-1 显示了一个双节点集群，在该集群中被切断的集群互联形成了潜在的裂脑情况。

图 13-1 已有的网络分裂（裂脑）



由于防护模块在各系统上的运行方式相同，因此两个节点假设其他节点已发生故障并执行防护操作，以确保逐出其他节点。各节点上的 VCS GAB 模块可确定因心跳丢失发生的对等节点故障，并将成员集更改传递给防护模块。

每一端都会争夺协调器磁盘的控制权。由于只有已注册节点能逐出另一节点的注册，因此只有一端能够成功地在各磁盘上完成命令。

从协调器磁盘主节点成功逐出对等节点的一端获胜。然后，获胜端的防护模块将成员集更改传递给 VCS 及其他已注册防护模块的高级别软件包，并允许 VCS 调用恢复操作。失败端则强制调用内核混乱并重新引导。

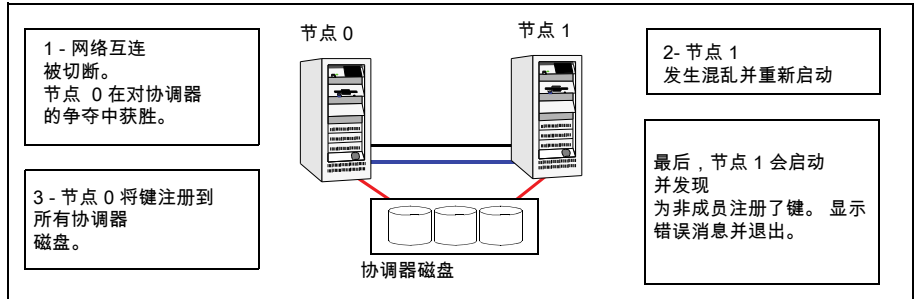
从已有网络分裂（裂脑）中恢复

防护模块 `vxfsen` 防止在发生网络分裂和接下来的混乱后启动节点，然后重新引导节点。

示例场景 I

图 13-2 的场景可能会产生与其中一个节点被关闭用于维护的双节点集群相似的症状。在中断期间，专用互连电缆是断开的。

图 13-2 示例场景 I



在示例场景 I 中，会出现下列情况：

- 节点 0 在网络故障后对协调器的争夺中获胜。
- 节点 1 发生混乱并重新启动。
- 节点 0 将键注册到协调器磁盘。节点 1 启动时，识别节点 0 键，但无法识别当前 GAB 成员集中的节点 0。它可以判断可能预先存在的裂脑，随后使 `vxfs` 模块在控制台上打印出一条错误消息。`vxfs` 模块将阻止防护启动，随后阻止 VCS 联机。

建议的解决方案：关闭节点 1，重新连接电缆，并重新启动节点 1。

示例场景 II

与示例场景 I 类似，如果在双节点集群中断开专用互连电缆，则节点 1 被防护在集群外，该节点将发生混乱并重新启动。如果在专用互连电缆修复之前，节点 1 重新加入集群，则节点 0 重新启动并远程登录（或只是重新启动）。在专用网络修复前，无法将节点写入数据磁盘。原因是无法形成 GAB 成员集，因此也无法形成集群。

与示例场景 I 类似，如果在双节点集群中断开专用互连电缆，则节点 1 被防护在集群外，该节点将发生混乱并重新启动。在专用网络修复前，无法将节点写入数据磁盘。原因是无法形成 GAB 成员集，因此也无法形成集群。

建议的解决方案：关闭两个节点，重新连接电缆，并重新启动节点。

示例场景 III

与示例场景 II 类似，如果在双节点集群中断开专用互连电缆，则节点 1 被防护在集群外，该节点将发生混乱并重新启动。如果在专用互联电缆修复前，节点 1 重新加入集群，但节点 0 因硬件故障出现混乱且无法恢复，则节点 1 无法重新加入。

建议的解决方案：关闭节点 1，重新连接电缆，并重新启动节点。然后必须从协调器磁盘中清除节点 0 的注册。

修复场景 III

- 1 在节点 1 上，键入以下命令：

```
# /opt/VRTSvcs/vxfen/bin/vxfenclearpre
```

- 2 重新启动节点。

- 3 运行：

```
# gabconfig -cx
```

对 SFHA Solutions 集群中的 Cluster Volume Manager 进行故障排除

本节讨论就 CVM 问题进行的故障排除。

将节点添加到 SFHA Solutions 集群后，CVM 组仍处于脱机状态

向集群中添加节点后，CVM 组处于脱机状态的可能原因如下：

- cssd 资源在 cvm 组中配置为关键资源。
- cvm 组中配置为关键资源的其他资源处于脱机状态。

在将 cssd 配置为关键资源时解决问题

- 1 以 root 用户身份登录现有集群中的其中一个节点。
- 2 在 cvm 组中将 cssd 资源配置为非关键资源：

```
# haconf -makerw  
# hares -modify cssd Critical 0  
# haconf -dump -makero
```

在组中的其他资源脱机时解决问题

- 1 以 root 用户身份登录现有集群中的其中一个节点。
- 2 使资源联机：

```
# hares -online resource_name -sys system_name
```

- 3 验证资源的状态：

```
# hastatus -resource resource_name
```

- 4 如果资源脱机，请将其配置为非关键资源：

```
# haconf -makerw  
# hares -modify resource_name Critical 0  
# haconf -dump -makero
```

在 SFHA Solutions 集群中无法导入共享磁盘组

比如显示类似以下内容的消息时：

```
vxvm:vxconfigd:ERROR:vold_pgr_register(/dev/vx/rdmp/disk_name):  
local_node_id<0  
Please make sure that CVM and vxfen are configured  
and operating correctly
```

首先，确保 CVM 正在运行。您可以通过运行 `vxclustadm nidmap` 命令查看集群中的 CVM 节点。

```
# vxclustadm nidmap  
Name          CVM Nid    CM Nid     State  
system01      1          0          Joined: Master  
system02      0          1          Joined: Slave
```

上述输出表明 CVM 运行正常，并且系统 `system01` 被用作 CVM 主节点。如果 CVM 运行正常，则当 CVM 无法从 `vxfen` 驱动程序检索本地系统的节点 ID 时，会显示上述输出。未配置端口 `b` 时，通常发生这种情况。

验证是否配置了 vxfen 驱动程序

- ◆ 使用以下命令检查 GAB 端口：

```
# gabconfig -a
```

本地系统上必须存在端口 `b`。

在 SFHA Solutions 集群中导入共享磁盘组出错

导入共享磁盘组时，可能出现以下消息：

```
VxVM vxdg ERROR V-5-1-587 Disk group disk group name: import
failed: No valid disk found containing disk group
```

可能需要删除已写入到磁盘的密钥。

有关删除写入到磁盘的密钥的信息：

无法在 SFHA Solutions 集群中启动 CVM

如果您无法启动 CVM，请检查 `/etc/llthosts` 和 `main.cf` 文件中节点 ID 的一致性。

可能需要删除已写入到磁盘的密钥。

有关删除写入到磁盘的密钥的信息：

即使 CVMCluster 在 SFHA Solutions 集群中处于联机状态，CVMVolDg 也不联机

当 CVMCluster 资源处于联机状态时，会自动导入已设置 `auto-import` 标志的所有共享磁盘组。如果由于某种原因磁盘组导入失败，则 CVMVolDg 资源会出现故障。清理 CVMVolDg 类型的资源并使之脱机不能解决此问题。

解决资源问题

- 1 解决导致共享磁盘组导入失败的问题。
- 2 使包含 CVMVolDg 资源类型的 `cvm` 组和包含 CVMCluster 资源类型的服务组脱机。
- 3 使包含 CVMCluster 资源的 `cvm` 组联机。
- 4 使包含 CVMVolDg 资源的 `cvm` 组联机。

Veritas Cluster Server 故障排除

- 14. VCS 的故障排除与恢复

VCS 的故障排除与恢复

本章节包括下列主题：

- VCS 消息日志记录
- VCS 引擎故障排除
- Low Latency Transport (LLT) 故障排除
- Group Membership Services/Atomic Broadcast (GAB) 故障排除
- VCS 启动故障排除
- 智能监视框架 (IMF) 故障排除
- 疑难解答服务组
- 疑难解答资源
- I/O 防护故障排除
- 疑难解答通知
- 全局集群的故障排除与恢复
- 对 `steward` 进程进行故障排除
- 授权许可故障排除

VCS 消息日志记录

VCS 生成两种类型的日志：引擎日志和代理日志。日志文件名中会追加字母后缀。字母 A 指示第一个日志文件，B 指示第二个日志文件，C 指示第三个日志文件，依此类推。

引擎日志位于 `/var/VRTSvcs/log/engine_A.log`。引擎日志消息的格式为：

Timestamp (Year/MM/DD) | Mnemonic | Severity | UMI | Message Text

- **Timestamp**: 消息生成的日期和时间。
- **Mnemonic**: 代表产品的字符串 ID (例如 VCS)。
- **Severity**: 一些级别, 包括 CRITICAL、ERROR、WARNING、NOTICE 和 INFO (分别表示从最严重到最不严重)。
- **UMI**: 唯一的消息 ID。
- **Message Text**: 由 VCS 生成的实际消息。

典型的引擎日志如下所示:

```
2011/07/10 16:08:09 VCS INFO V-16-1-10077 Received new
cluster membership
```

代理日志位于 `/var/VRTSvcs/log/<agent>.log`。代理日志消息的格式如下所示:

Timestamp (Year/MM/DD) | Mnemonic | Severity | UMI | Agent Type | Resource Name | Entry Point | Message Text

典型的代理日志如下所示:

```
2011/07/10 10:38:23 VCS WARNING V-16-2-23331
Oracle:VRT:monitor:Open for ora_lgwr failed, setting
cookie to null.
```

请注意, 并非所有节点上的日志都相同, 原因如下:

- VCS 记录本地节点上的本地事件。
- 当某个事件发生时, 并非所有节点都在运行。

VCS 将警告和错误消息输出至 `STDERR`。

如果 VCS 引擎、Command Server 或任何 VCS 代理遇到问题, 则会生成首次故障数据捕获 (FFDC) 日志, 并随其他核心转储和堆栈跟踪一起转储到以下位置:

- 对于 VCS 引擎: `$VCS_DIAG/diag/had`
- 对于 Command Server: `$VCS_DIAG/diag/CmdServer`
- 对于 VCS 代理: `$VCS_DIAG/diag/agents/type`, 其中 *type* 表示具体的代理类型。

变量 `$VCS_DIAG` 的默认值为 `/var/VRTSvcs/`。

如果没有打开调试日志记录, 则这些 FFDC 日志对于分析需要专业支持的问题很有用。

GAB 消息日志记录

如果 GAB 遇到某个问题，则还会生成并转储首次故障数据捕获 (FFDC) 日志。

配置 GAB 后，GAB 也会启动 GAB 日志记录后台驻留程序 (/opt/VRTSgab/gablogd)。默认情况下启用 GAB 日志记录后台驻留程序。可更改 GAB 可调参数 `gab_ibuf_count` 的值以禁用 GAB 日志记录后台驻留程序。

此 GAB 日志记录后台驻留程序在发生严重事件（例如 `iofence`），或在发生任何 GAB 端口的主节点故障时收集 GAB 相关的日志，并以压缩二进制格式存储数据。可使用以下 `gabread_ffdc` 实用程序读取 GAB 二进制日志文件：

```
/opt/VRTSgab/gabread_ffdc binary_logs_files_location
```

可更改下列控制 GAB 二进制日志文件的环境变量的值：

■ **GAB_FFDC_MAX_INDX**：定义最大 GAB 二进制日志文件数

GAB 日志记录后台驻留程序收集定义数量的日志文件，每个文件大小为 8 MB。默认值为 20，文件以从 `gablog.1` 到 `gablog.20` 的方式命名。无论在任何时候，最新的文件都是 `gablog.1` 文件。

■ **GAB_FFDC_LOGDIR**：为 GAB 二进制日志文件定义日志目录位置

默认位置为：

```
/var/adm/gab_ffdc
```

请注意，`gablog` 后台驻留程序将其日志写入位于相同目录中的 `glogd_A.log` 和 `glogd_B.log` 文件。

可以在以下 GAB 启动文件中定义这些变量，也可以使用 `export` 命令。要使更改生效，必须重新启动 GAB。

```
/etc/default/gab
```

为代理启用调试日志

本节介绍如何为 VCS 代理启用调试日志。

为代理启用调试日志

- 1 将配置设置为读写：

```
# haconf -makerw
```

- 2 启用日志记录并设置所需的日志级别。以下示例说明了用于 IPMultiNIC 资源类型的命令。

```
# hatype -modify IPMultiNIC LogDbg DBG_1 DBG_2 DBG_4 DBG_21
```

有关更多信息，请参见 `LogDbg` 属性的说明。

- 3 对于基于脚本的代理，运行 `halog` 命令将消息添加至引擎日志：

```
# halog -addtags DBG_1 DBG_2 DBG_4 DBG_21
```

- 4 保存该配置。

```
# haconf -dump -makero
```

如果设置了 `DBG_AGDEBUG`，代理实例的代理框架日志出现在运行代理的节点上的代理日志中。

启用 IMF 的调试日志

运行以下命令，启用智能监视框架 (IMF) 的其他调试日志。将在特定于代理的 `/var/VRTSvcs/log/agentname_A.log` 日志文件中记录这些消息。

请参见第 141 页的“[智能监视框架 \(IMF\) 故障排除](#)”。

启用其他调试日志

- 1 对于 `Process`、`Mount` 和应用程序代理：

```
# hatype -modify agentname LogDbg  
DBG_AGDEBUG DBG_AGTRACE DBG_AGINFO DBG_1 DBG_2  
DBG_3 DBG_4 DBG_5 DBG_6 DBG_7
```

- 2 对于 `Oracle` 和 `Netlsnr` 代理：

```
# hatype -modify agentname LogDbg  
DBG_AGDEBUG DBG_AGTRACE DBG_AGINFO DBG_1 DBG_2  
DBG_3 DBG_4 DBG_5 DBG_6 DBG_7  
DBG_8 DBG_9 DBG_10
```

- 3 对于 `CFSMount` 代理：

```
# hatype -modify agentname LogDbg  
DBG_AGDEBUG DBG_AGTRACE DBG_AGINFO DBG_1 DBG_2  
DBG_3 DBG_4 DBG_5 DBG_6 DBG_7  
DBG_8 DBG_9 DBG_10 DBG_11 DBG_12  
DBG_13 DBG_14 DBG_15 DBG_16  
DBG_17 DBG_18 DBG_19 DBG_20 DBG_21
```

- 4 对于 CVMvxconfigd 代理，不必启用其他任何调试日志。
- 5 对于 AMF 驱动程序内存跟踪缓冲区：

```
# amfconfig -S errlevel all all
```

如果已启用 AMF 驱动程序内存跟踪缓冲区，可使用 `amfconfig -p dbglog` 命令查看其他日志。

对 VCS 引擎启用调试日志

您可以通过以下两种方法对 VCS 引擎、VCS 代理和 HA 命令启用调试日志：

- 要在运行时启用调试日志，请使用 `halog -addtags` 命令。
- 要在启动时启用调试日志，请使用 `VCS_DEBUG_LOG_TAGS` 环境变量。必须先设置 `VCS_DEBUG_LOG_TAGS`，然后再启动 HAD 或运行 HA 命令。

例如：

```
# export VCS_DEBUG_LOG_TAGS="DBG_TRACE DBG_POLICY"  
# hstart
```

```
# export VCS_DEBUG_LOG_TAGS="DBG_AGINFO DBG_AGDEBUG DBG_AGTRACE"  
# hstart
```

```
# export VCS_DEBUG_LOG_TAGS="DBG_IPM"  
# hagrpl -list
```

注意：调试日志消息是十分详细的。如果启用调试日志，则日志文件可能会很快填满。

关于调试日志标记用法

下表说明了调试标记的用法：

实体	所使用的调试日志
代理函数	DBG_1 到 DBG_21
代理框架	DBG_AGTRACE DBG_AGDEBUG DBG_AGINFO

实体	所使用的调试日志
Icmp 代理	DBG_HBFW_TRACE DBG_HBFW_DEBUG DBG_HBFW_INFO
HAD	DBG_AGENT (用于与代理相关的调试日志) DBG_ALERTS (用于警报调试日志) DBG_CTEAM (用于 GCO 调试日志) DBG_GAB、DBG_GABIO (用于 GAB 调试消息) DBG_GC (用于对每个日志消息显示全局计数器) DBG_INTERNAL (用于内部消息) DBG_IPM (用于进程间消息传送) DBG_JOIN (用于结合逻辑) DBG_LIC (用于与授权相关的消息) DBG_NTEVENT (用于 NT 事件日志) DBG_POLICY (用于引擎策略) DBG_RSM (用于 RSM 调试消息) DBG_TRACE (用于跟踪消息) DBG_SECURITY (用于与安全相关的消息) DBG_LOCK (用于调试锁基元) DBG_THREAD (用于调试线程基元) DBG_HOSTMON (用于 HostMonitor 调试日志)

收集 VCS 信息以进行支持分析

遇到 VCS 问题时，必须运行 `hagetcf` 命令收集信息。Symantec 技术支持使用这些脚本的输出帮助分析和解决所有 VCS 问题。`hagetcf` 命令收集关于安装的软件、集群配置、系统、日志的信息和相关信息并创建 `gzip` 文件。

有关更多信息，请参见 `hagetcf(1M)` 手册页。

收集 VCS 信息以进行支持分析

- ◆ 在每个节点上运行以下命令：

```
# /opt/VRTSvcs/bin/hagetcf
```

该命令会提示您指定 `gzip` 文件的输出目录。可以将 `gzip` 文件保存到默认的 `/tmp` 目录或其他目录。

进行故障排除并解决问题。

请参见第 133 页的“VCS 引擎故障排除”。

请参见第 140 页的“VCS 启动故障排除”。

请参见第 142 页的“疑难解答服务组”。

请参见第 145 页的“疑难解答资源”。

请参见第 161 页的“疑难解答通知”。

请参见第 162 页的“全局集群的故障排除与恢复”。

请参见第 164 页的“对 `steward` 进程进行故障排除”。

如果问题无法解决，请联系 Symantec 技术支持部门，并提供 `hagetcf` 命令所生成的文件。

收集 LLT 和 GAB 信息以进行支持分析

当 LLT 和 GAB 出现问题时，必须运行 `getcomms` 脚本来收集 LLT 和 GAB 信息。`getcomms` 脚本收集 LLT 和 GAB 信息的同时，还收集核心转储和堆栈跟踪。

收集 LLT 和 GAB 信息以进行支持分析

- 1 如果已更改 GAB_FFDC_LOGDIR 参数的默认值，则必须重新导出同一变量才能运行 getcomms 脚本。

请参见第 127 页的“GAB 消息日志记录”。

- 2 运行以下命令以收集信息：

```
# /opt/VRTSgab/getcomms
```

默认情况下，该脚本使用 rsh。确保您配置了无密码的 rsh。如果在集群节点之间有无密码的 ssh，则可以使用 -ssh 选项。要在运行该命令的节点上收集信息，请使用 -local 选项。

进行故障排除并解决问题。

请参见第 135 页的“Low Latency Transport (LLT) 故障排除”。

请参见第 138 页的“Group Membership Services/Atomic Broadcast (GAB) 故障排除”。

如果问题无法解决，请联系 Symantec 技术支持部门，并提供 getcomms 脚本所生成的 /tmp/commslog.time_stamp.tar 文件。

收集 IMF 信息以进行支持分析

遇到 IMF（智能监视框架）问题时，必须运行 getimf 脚本以收集信息。

收集 IMF 信息以进行支持分析

- ◆ 在每个节点上运行以下命令：

```
# /opt/VRTSamf/bin/getimf
```

进行故障排除并解决问题。

请参见第 141 页的“智能监视框架 (IMF) 故障排除”。

如果无法解决问题，请与 Symantec 技术支持联系，并提供 getimf 脚本生成的文件。

消息目录

VCS 对消息目录提供多语言支持。这些二进制消息目录 (BMC) 存储在以下默认位置中。变量 *language* 代表一个两字母的语言名称缩写。

```
/opt/VRTS/messages/language/module_name
```

VCS 命令行界面以 VCS 支持的语言显示错误消息和成功消息。hamsg 命令以 VCS 支持的语言显示 VCS 引擎日志。

BMC 包括:

<code>gcoconfig.bmc</code>	gcoconfig 消息
<code>VRTSvcsHbfw.bmc</code>	心跳框架消息
<code>VRTSvcsTriggers.bmc</code>	VCS 触发器消息
<code>VRTSvcsWac.bmc</code>	广域连接器进程消息
<code>vxfen*.bmc</code>	防护消息
<code>gab.bmc</code>	GAB 命令行界面消息
<code>hagetcf.bmc</code>	hagetcf 消息
<code>llt.bmc</code>	LLT 命令行界面消息
<code>VRTSvcsAgfw.bmc</code>	代理框架消息
<code>VRTSvcsAlerts.bmc</code>	VCS 警报消息
<code>VRTSvcsApi.bmc</code>	VCS API 消息
<code>VRTSvcsCommon.bmc</code>	通用模块消息
<code>VRTSvcsHad.bmc</code>	VCS 引擎 (HAD) 消息
<code>VRTSvcsplatformAgent.bmc</code>	VCS 捆绑代理消息
<code>VRTSvcsplatformagent_name.bmc</code>	VCS Enterprise Agent 消息

VCS 引擎故障排除

本主题包含有关 VCS 引擎故障排除的信息。

请参见第 135 页的[“联机前 IP 检查”](#)。

HAD 诊断

当 VCS 引擎 HAD 转储核心时，会将核心写入到目录 `$VCS_DIAG/diag/had` 中。变量 `$VCS_DIAG` 的默认值是 `/var/VRTSvcs/`。

当 HAD 进行核心转储时，请查看 `$VCS_DIAG/diag/had` 目录的内容。请参见以下日志以获取更多信息：

- 操作系统控制台日志
- 引擎日志
- hashadow 日志

当 HAD 和 GAB 遇到心跳问题时，VCS 运行脚本 `/opt/VRTSvcs/bin/vcs_diag` 来收集诊断信息。诊断信息存储在 `$VCS_DIAG/diag/had` 目录中。

当 HAD 启动时，它会将此目录重命名为 `had.timestamp`，其中 *timestamp* 表示此目录的重命名时间。

HAD 连续重新启动

当您使用 `hastart` 命令启动 HAD 时，HAD 可能会连续重新启动。系统无法进入 **RUNNING** 状态并丢失其端口 **h** 成员集。

建议的操作：检查 `engine_A.log` 文件中是否存在带有标识符 **V-16-1-10125** 的消息。下面的消息是一个示例：

```
VCS INFO V-16-1-10125 GAB timeout set to 30000 ms (VCS INFO V-16-1-10125  
GAB 超时设置为 30000 毫秒)
```

该值指示为了使 HAD 向 GAB 注册心跳而设置的超时。如果系统负载较重，则 30 秒超时可能不足，无法使 HAD 向 GAB 注册心跳。请根据需要将该超时设置为适当的较大值。

DNS 配置问题导致 GAB 终止 HAD

如果 HAD 定期被 GAB 因不明原因而终止，请查看 HAD 核心文件以确定堆栈跟踪中是否出现了 DNS 解析函数（`res_send()`、`res_query()`、`res_search()` 等等）。出现 DNS 解析函数可能表示存在 DNS 配置问题。

VCS 高可用性后台驻留程序 (HAD) 使用 `gethostbyname()` 函数。在 UNIX 平台上，如果文件 `/etc/nsswitch.conf` 的主机条目中有 DNS，则调用 `gethostbyname()` 函数可导致调用 DNS 解析方法。

如果无法访问 `/etc/resolve.conf` 中指定的名称服务器或存在 DNS 配置问题，则调用的 DNS 解析方法可能阻止 HAD，从而导致 HAD 不能及时将心跳发送到 GAB。

种子设定和 I/O 防护

当 I/O 防护启动时，会执行一项检查，以确保在协调点上具有键的系统也位于 GAB 成员集中。如果 `/etc/gabtab` 中的 `gabconfig` 命令允许集群在少于集群系统总数的系统上设定种子，或者使用 `gabconfig -x` 命令强制集群设定种子，则很可能此项检查将不匹配。在这种情况下，防护模块将检测可能的裂脑情况，显示相应的错误，而且 HAD 不会启动。

当集群的所有成员之间可以互换心跳信号时，建议使集群自动设定种子。在这种情况下，所有系统在加入 GAB 成员集后都会执行 I/O 防护键放置。

联机前 IP 检查

您可以启用故障转移 IP 地址的联机前检查，以防止发生网络分裂。此项检查会对服务组的已配置 IP 地址运行 ping 命令，以验证它未在使用中。如果它已在使用中，则服务组不会进入联机状态。

另一项检查验证系统是否连接到其公共网络和专用网络。如果系统在对公共网络进行广播 ping 时和对专用网络进行检查时未收到响应，则会确定系统是独立的并且不会使服务组联机。

要启用联机前 IP 检查，请执行下列操作之一：

- 如果 preonline 触发器脚本不存在，请将 sample_triggers 目录中的 preonline 触发器脚本复制到 triggers 目录：

```
# cp/opt/VRTSvcs/bin/sample_triggers/VRTSvcs/preonline_ipc  
/opt/VRTSvcs/bin/triggers/preonline
```

更改文件权限使它成为可执行文件。

- 如果 preonline 触发器脚本已存在，请创建一个诸如 /preonline 的目录，并将现有 preonline 触发器以 T0preonline 的形式移到该目录。将 preonline_ipc 触发器以 T1preonline 的形式复制到同一目录。
- 如果您使用了多个触发器，请以 TNpreonline 的形式复制 preonline_ipc 触发器，其中 TN 是下一个更大的 TNumber。

Low Latency Transport (LLT) 故障排除

本节包含与 Low Latency Transport (LLT) 关联的错误消息，并提供了说明和建议的操作。

LLT 启动脚本显示错误

如果网络中有多个系统具有相同的 clusterid-nodeid 对和相同的以太网 SAP/UDP 端口，则 LLT 启动脚本会显示如下所示的错误消息：

```
LLT lltconfig ERROR V-14-2-15238 node 1 already exists  
in cluster 8383 and has the address - 00:18:8B:E4:DE:27  
LLT lltconfig ERROR V-14-2-15241 LLT not configured,  
use -o to override this warning  
LLT lltconfig ERROR V-14-2-15664 LLT could not  
configure any link  
LLT lltconfig ERROR V-14-2-15245 cluster id 1 is  
already being used by nid 0 and has the  
address - 00:04:23:AC:24:2D
```

```
LLT lltconfig ERROR V-14-2-15664 LLT could not
configure any link
```

建议的操作：请确保网络上的所有系统都具有唯一的 **clusterid-nodeid** 对。您可以使用 `lltdump -f device -D` 命令获取连接到网络的唯一 **clusterid-nodeid** 对的列表。该实用程序仅适用于 LLT over Ethernet。

LLT 检测到使用了交叉链路

如果 LLT 检测到系统的多个链路连接到了同一个网络，则 LLT 会在 **syslog** 中记录类似如下内容的警告消息：

```
LLT WARNING V-14-1-10498 recvarpack cross links? links 0 and 2 saw
the same peer link number 1 for node 1
```

建议的操作：这是一条信息性消息。LLT 支持交叉链路。但是，如果该交叉链路不是有意网络设置，请确保没有两个链路从同一系统连接至同一网络。也就是说，不同的 LLT 链路需要位于不同的网络。

LLT 链路状态消息

表 14-1 描述了 LLT 日志消息，例如 **syslog** 中链路的故障、活动、非活动或过期状态。

表 14-1 LLT 链路状态消息

消息	说明和建议的操作
<pre>LLT INFO V-14-1-10205 link 1 (link_name) node 1 in trouble</pre>	<p>此消息意味着 LLT 没有接收到来自 LLT peertrouble 时间的指示对等节点中指示链路上的任何心跳。对于高优先级链路，默认的 LLT peertrouble 时间是 2 秒，对于低优先级链路是 4 秒。</p> <p>建议的操作：如果这些消息只是偶尔出现在 syslog 中，可以将其忽略。如果这些消息大量充斥 syslog，请执行下列操作之一：</p> <ul style="list-style-type: none"> ■ 将 peertrouble 时间增加为更高的值（但是要远远低于 peerinact 值）。运行以下命令： <pre>lltconfig -T peertrouble:<value> for hipri link lltconfig -T peertroublelo:<value> for lopri links.</pre> <p>有关详细信息，请参见 <code>lltconfig(1m)</code> 手册页。</p> <ul style="list-style-type: none"> ■ 替换 LLT 链路。

消息	说明和建议的操作
<pre>LLT INFO V-14-1-10024 link 0 (link_name) node 1 active</pre>	<p>此消息意味着 LLT 启动后将在此链路上看到心跳。 建议的操作：不需要执行任何操作。该消息是信息性消息。</p>
<pre>LLT INFO V-14-1-10032 link 1 (link_name) node 1 inactive 5 sec (510)</pre>	<p>此消息意味着 LLT 未接收到来自指示时间段的指示对等节点中指示链路上的任何心跳。 如果对等节点实际上没有断开，请检查以下情况：</p> <ul style="list-style-type: none"> ■ 检查链路是否以物理方式断开了与系统或交换机的连接。 ■ 检查链路的运行状况，并在必要时替换链路。
<pre>LLT INFO V-14-1-10510 sent hbreq (NULL) on link 1 (link_name) node 1. 4 more to go. LLT INFO V-14-1-10510 sent hbreq (NULL) on link 1 (link_name) node 1. 3 more to go. LLT INFO V-14-1-10510 sent hbreq (NULL) on link 1 (link_name) node 1. 2 more to go. LLT INFO V-14-1-10032 link 1 (link_name) node 1 inactive 6 sec (510) LLT INFO V-14-1-10510 sent hbreq (NULL) on link 1 (link_name) node 1. 1 more to go. LLT INFO V-14-1-10510 sent hbreq (NULL) on link 1 (link_name) node 1. 0 more to go. LLT INFO V-14-1-10032 link 1 (link_name) node 1 inactive 7 sec (510) LLT INFO V-14-1-10509 link 1 (link_name) node 1 expired</pre>	<p>此消息意味着 LLT 在未接收到 LLT peerinact 时间以上的指示对等节点中指示链路上的任何心跳。LLT 会尝试请求心跳机制（向对等节点发送 5 次 hbreq），如果对等节点没有响应，LLT 将针对该对等节点将此链路标记为“过期”。 建议的操作：如果对等节点实际上没有断开，请检查以下情况：</p> <ul style="list-style-type: none"> ■ 检查链路是否以物理方式断开了与系统或交换机的连接。 ■ 检查链路的运行状况，并在必要时替换链路。
<pre>LLT INFO V-14-1-10499 recvarpreq link 0 for node 1 addr change from 00:00:00:00:00:00 to 00:18:8B:E4:DE:27</pre>	<p>当 LLT 获知对等节点的地址时，会记录此消息。 建议的操作：不需要执行任何操作。该消息是信息性消息。</p>

消息	说明和建议的操作
<p>在检测到链路故障的本地节点上:</p> <pre>LLT INFO V-14-1-10519 link 0 down LLT INFO V-14-1-10585 local link 0 down for 1 sec LLT INFO V-14-1-10586 send linkdown_ntf on link 1 for local link 0 LLT INFO V-14-1-10590 recv linkdown_ack from node 1 on link 1 for local link 0 LLT INFO V-14-1-10592 received ack from all the connected nodes</pre> <p>在对等节点上:</p> <pre>LLT INFO V-14-1-10589 recv linkdown_ntf from node 0 on link 1 for peer link 0 LLT INFO V-14-1-10587 send linkdown_ack to node 0 on link 1 for peer link 0</pre>	<p>如果您启用了 LLT 以更快地检测链路故障，则会显示这些消息。当链路出现故障或与节点断开连接时（电缆拔出、交换机故障等），本地节点上的 LLT 会检测到此情况并通过 LLT 隐藏链路将相关信息传播到所有对等节点。在本地节点上的 LLT 收到所有节点的确认之后，该 LLT 会将此链路标记为断开连接。</p>

Group Membership Services/Atomic Broadcast (GAB) 故障排除

本节包括与 Group Membership Services/Atomic Broadcast (GAB) 关联的错误消息，并提供了说明和建议的操作。

GAB 计时器问题

GAB 计时器仅唤醒执行真实工作的线程。如果操作系统没有机会安排线程，或者操作系统没有启用计时器，则 GAB 可能会记录类似如下内容的消息：

```
GAB INFO V-15-1-20124 timer not called for 11 seconds
```

建议的操作：如果此问题很少在高负载时发生，并且不会造成任何不利影响，则此问题是良性的。如果该问题重复发生，或者导致 VCS 后台驻留程序反复重新启动或导致节点混乱，则必须提高线程的优先级（使用 `gab_timer_pri` 可调参数）。

如果问题仍然存在，请收集以下信息并与 Symantec 支持联系：

- `sar` 或 `perfmgr` 数据
- 任何 VCS 后台驻留程序诊断

- 任何混乱的转储
- 所有节点上的 `commslog`

端口在重新打开时出现延迟

如果在 LLT 状态清除操作完成之前将 GAB 端口关闭并重新打开，则 GAB 会记录类似如下内容的消息：

```
GAB INFO V-15-1-20102 Port v: delayed reopen
```

建议的操作：如果在 GAB 重新配置期间发生此问题，并且没有重复发生，则此问题是良性的。如果问题仍然存在，请从每个节点中收集 `commslog`，并与 Symantec 支持联系。

由于客户端进程故障，导致节点混乱

如果 VCS 后台驻留程序在 `VCS_GAB_TIMEOUT`（默认值为 30 秒）环境变量指定的配置超时时间内没有向 GAB 发送心跳，则节点会发生混乱，并显示类似如下内容的消息：

```
GAB Port h halting node due to client process failure at 3:109
```

如果 VCS 停滞在处于不可中断状态的内核中，或者系统负载较重，使得 VCS 后台驻留程序无法使用 `SIGKILL` 停止，则 GAB 终止 VCS 后台驻留程序的尝试（重试五次）会失败。

建议的操作：

- 如果是性能问题，请增加 `VCS_GAB_TIMEOUT` 环境变量的值，以允许 VCS 有更长的时间发送心跳。
- 如果是内核问题，请将 GAB 配置为不发生混乱但继续尝试停止 VCS 后台驻留程序。

执行以下任务：

- 在每个节点上运行以下命令：

```
gabconfig -k
```

- 在 `/etc/gabtab` 文件中将 `-k` 选项添加到 `gabconfig` 命令：

```
gabconfig -c -k -n 6
```

- 如果问题仍然存在，请收集 `sar` 或类似输出以及崩溃转储，并与 Symantec 支持联系。

VCS 启动故障排除

本主题包括与启动 VCS 相关的错误消息（以粗体文本显示），并对每个错误及建议的操作进行了说明。

VCS:10622 local configuration missing

本地配置缺失。

建议的操作：在另一个具有有效配置文件的系统上启动 VCS 引擎 HAD。出现配置错误的系统从其他系统中提取有效配置。

另一种方法就是在本地系统上安装配置文件，并强制 VCS 重新读取该配置文件。如果该文件看起来有效，则验证它不是早期版本。

键入以下命令验证配置：

```
# cd /etc/VRTSvcs/conf/config  
# hacf-verify
```

VCS:10623 local configuration invalid

本地配置无效。

建议的操作：在另一个具有有效配置文件的系统上启动 VCS 引擎 HAD。出现配置错误的系统从其他系统中提取有效配置。

另一种方法就是更正本地系统上的配置文件，并强制 VCS 重新读取该配置文件。如果该文件看起来有效，则验证它不是早期版本。

键入以下命令验证配置：

```
# cd /etc/VRTSvcs/conf/config  
# hacf-verify
```

VCS:11032 registration failed.Exiting

GAB 没有注册过或者已经取消注册。

建议的操作：使用 /etc/gabtab 文件中的 `gabconfig` 命令注册 GAB。验证该文件是否存在以及它是否包含命令 `gabconfig -c`。

如果 LLT 设置不正确，则 GAB 可能会取消注册。验证 /etc/llttab 中的配置是否正确。如果 LLT 配置不正确，则进行适当的更改并重新启动。

Waiting for cluster membership.

此消息表明 GAB 无法设定种子。如果是这种情况，则 `gabconfig -a` 命令不显示任何成员，并可能在控制台上或事件日志中显示以下消息。

```
GAB: Port a registration waiting for seed port membership
GAB: Port h registration waiting for seed port membership
```

智能监视框架 (IMF) 故障排除

查看下列日志以对智能监视框架 (IMF) 的相关问题进行隔离和故障排除：

- 给定操作系统的系统控制台日志
- VCS 引擎日志： `/var/VRTSvcs/log/engine_A.log`
- 特定于代理的日志： `/var/VRTSvcs/log/agentname_A.log`
- AMF 内存中跟踪缓冲区：使用 `amfconfig -p dbglog` 命令查看内容

请参见第 128 页的“启用 IMF 的调试日志”。

请参见第 132 页的“收集 IMF 信息以进行支持分析”。

表 14-2 列出了智能资源监视的最常见问题，并提供了故障排除和解决问题的说明。

表 14-2 与 IMF 相关的问题和建议的操作

问题	说明和建议的操作
智能资源监视并未降低系统利用率	<p>如果在启用了智能资源监视后系统依然繁忙，请按如下所示进行故障排除：</p> <ul style="list-style-type: none"> ■ 检查代理日志文件以查看 <code>imf_init</code> 代理函数是否失败。 <p>如果 <code>imf_init</code> 代理函数失败，则执行以下操作：</p> <ul style="list-style-type: none"> ■ 确保将 <code>AMF_START</code> 环境变量值设置为 1。 ■ 确保已加载 AMF 模块。 <ul style="list-style-type: none"> ■ 确保为以下属性键设置了正确的 IMF 属性值： <ul style="list-style-type: none"> ■ 必须将 IMF 属性的 <code>Mode</code> 键的值设置为 1、2 或 3。 ■ 必须将 IMF 属性的 <code>MonitorFreq</code> 键的值设置为 0 或大于 0 的值。 <p>例如，可以将 <code>Process</code> 代理的 <code>MonitorFreq</code> 键的值设置为 0。有关与可识别 IMF 的代理对应的配置建议，请参考相应的代理文档。</p> <p>请注意，可以覆盖 IMF 属性。因此，如果针对单个资源设置属性，请检查单个资源的值。</p> ■ 验证是否已在 AMF 驱动程序中注册这些资源。检查 <code>amfstat</code> 命令输出。 ■ 检查代理的 <code>LevelTwoMonitorFreq</code> 属性设置。例如，<code>Process</code> 代理必须将此属性值设置为 0。 <p>有关与可识别 IMF 的代理对应的配置建议，请参考相应的代理文档。</p>

问题	说明和建议的操作
启用代理的智能监视不立即提供性能结果	<p>只有在达到稳定状态后才会启动资源的实际智能监视。因此，启用 IMF 后需要一段时间才能看到对性能的正向影响。这是预期的行为。</p> <p>有关何时达到稳定状态的更多信息，请参见以下主题：</p>
尽管将 IMF 模式设置为 3，但代理仍不执行智能监视	<p>对于使用 AMF 驱动程序进行 IMF 通知的代理，如果智能资源监视尚未生效，请执行以下操作：</p> <ul style="list-style-type: none"> ■ 确保将 IMF 属性的 Mode 键值设置为三 (3)。 ■ 查看代理日志以确认在 AMF 中注册 imf_init() 代理已成功。必须在启动代理之前加载 AMF 驱动程序，因为代理在启动时在 AMF 中注册。如果并非如此，则先启动 AMF 模块再重新启动代理。
尽管将 IMF 模式更改为 0，但 AMF 模块仍无法卸载	<p>即使将 Mode 键的值更改为 0 后，代理仍继续控制 AMF 驱动程序，直到终止代理为止。要卸载 AMF 模块，必须释放对它的所有控制。</p> <p>如果将 IMF 模式值更改为 0 后 AMF 模块仍无法卸载，请执行以下操作：</p> <ul style="list-style-type: none"> ■ 运行 <code>amfconfig -Uof</code> 命令。此命令强制删除对模块的所有控制并取消配置该模块。 ■ 然后，卸载 AMF。
尝试为代理启用 IMF 时， <code>haimfconfig -enable -agent <agent_name></code> 命令将返回一条消息，指出已为该代理启用 IMF。但是，当 VCS 和相关代理运行时， <code>haimfconfig -display</code> 命令会将 <code>agent_name</code> 的状态显示为 DISABLED。	<p>出现此行为的部分可能原因如下所示：</p> <ul style="list-style-type: none"> ■ 此代理可能需要一些手动步骤才能识别 IMF。有关这些手动步骤，请参考代理文档。 ■ 此代理为自定义代理，且无法识别 IMF。有关如何使自定义代理识别 IMF 的信息，请参见《Veritas Cluster Server Agent 开发指南》。 ■ 如果之前的步骤无法解决此问题，请联系 Symantec 技术支持部门。

疑难解答服务组

本主题介绍与使服务组联机和脱机相关的最常见问题。粗体文本提供问题的说明。还包括建议的操作（如果适用）。

VCS 不会自动启动服务组

如果集群中的 VCS 引擎 (HAD) 是由 `hashadow` 进程重新启动的，则 VCS 不会自动启动故障转移服务组。

此行为可防止服务组由于发生某些事件（如 GAB 因高负载而终止 HAD，或 HAD 自行终止以纠正意外的错误情况）而自动联机。

系统未处于 RUNNING 状态

建议的操作：键入 `hasys -display system` 以检查系统状态。当系统未处于运行状态时，如果未发现问题，就可以启动 VCS。

服务组未配置为在该系统上运行

服务组的 `SystemList` 属性可能不包含该系统的名称。

建议的操作：使用 `hagrp -display service_group` 命令的输出验证系统名称。

服务组未配置为自动启动

如果服务组在系统上不自动启动，则该组可能未配置为自动启动，或者可能未配置为在特定系统上自动启动。

建议的操作：使用 `hagrp -display service_group AutoStartList node_list` 命令的输出验证 `AutoStart` 和 `AutoStartList` 属性的值。

服务组已冻结

建议的操作：使用 `hagrp -display service_group` 命令的输出验证 `Frozen` 和 `TFrozen` 属性的值。使用 `hagrp -unfreeze` 命令对该组取消冻结。请注意，VCS 不会使冻结的服务组脱机。

故障转移服务组在另一个系统上处于联机状态

该组是故障转移组，并且在另一个系统上处于联机状态或部分联机状态。

建议的操作：使用 `hagrp -display service_group` 命令的输出验证 `State` 属性的值。使用 `hagrp -offline` 命令使组在另一个系统上脱机。

某个关键资源出现故障

`hagrp -display service_group` 命令的输出表明服务组出现故障。

建议的操作：使用 `hares -clear` 命令清除该故障。

已自动禁用服务组

当 VCS 不了解特定系统上服务组的状态时，它会自动禁用该系统上的服务组。在以下情况下会发生自动禁用：

- 当 VCS 引擎 HAD 未在系统上运行时。
在这些情况下，在其 `SystemList` 属性中包括该系统的所有服务组都会被自动禁用。这不适用于断电的系统。

- 当在系统上未探查服务组内的所有资源时。

建议的操作：使用 `hagrp -display service_group` 命令的输出验证 `AutoDisabled` 属性的值。

警告：要在 VCS 自动禁用某个组后手动使该组联机，请确保该组在 `AutoDisabled` 属性已由 VCS 设置为 1 的任何系统上未处于完全或部分活动状态。具体地说，验证可能会因在多个系统上处于活动状态而受损的所有资源在指定的系统上处于关闭状态。然后，为每个系统清除 `AutoDisabled` 属性：`# hagrp -autoenable service_group -sys system`

服务组正在等待资源联机/脱机

建议的操作：查看服务组中所有资源的 `IState` 属性，确定哪些资源正在等待进入联机状态（或哪些资源正在等待进入脱机状态）。使用 `hastatus` 命令可帮助识别这些资源。请参见 `/var/VRTSvcs/log` 中的引擎日志和代理日志，以获取有关资源无法联机或脱机的原因的信息。

要清除此状态，请确保正在等待进入联机/脱机状态的所有资源不会自行联机/脱机。使用 `hagrp -flush` 命令或 `hagrp -flush -force` 命令清除 VCS 的内部状态。然后，可以使服务组在另一个系统上联机或脱机。

有关 `hagrp -flush` 和 `hagrp -flush -force` 命令的更多信息：

警告：请谨慎使用 `-force` 选项。此选项可能导致无意中将资源状态返回为 `FAULTED`。在资源从 `WAITING TO GO ONLINE` 转换为 `NOT WAITING` 的时间间隔中，如果代理未完成 `offline` 代理函数，则该代理可能会将资源的状态返回为 `OFFLINE`。VCS 会将资源的意外脱机视为意外脱机视为 `FAULT`，并启动不必要的恢复操作。

服务组正在等待某个依赖关系得到满足

建议的操作：要查看未满足哪些依赖关系，请键入 `hagrp -dep service_group` 查看服务组依赖关系，或者键入 `hares -dep resource` 查看资源依赖关系。

未完全探查服务组

如果代理进程没有对服务组中的每一个资源都加以监视，则会出现这种情况。VCS 引擎 HAD 启动时，会立即“探查”以查找所有资源的初始状态。（如果代理没有返回值，则它不能进行探查。）在 VCS 尝试使服务组作为 `AutoStart` 的一部分进行联机之前，必须在 `SystemList` 属性中包含的所有系统上对该组进行探查。这可

确保即使服务组在 VCS 启动之前就处于联机状态，VCS 也不会无意间使该服务组在另一个系统上联机。

建议的操作：使用 `hagrp -display service_group` 的输出查看系统服务组的 **ProbesPending** 属性值。（该值应为 0。）要确定未探查到哪些资源，请验证指定系统上的每个资源的本地 **Probed** 属性。0 表示正在等待探查结果，1 表示已探查，2 表示未引导 VCS。请参见引擎日志和代理日志以获取有关信息。

疑难解答资源

本主题介绍与使资源联机和脱机相关的最常见问题。粗体文本提供问题的说明。还包括建议的操作（如果适用）。

由于故障转移而使服务组联机

VCS 尝试使已在出故障的系统上联机或正在联机的资源联机。每个父资源必须等待其子资源联机后才能启动。

建议的操作：验证子资源是否处于联机状态。

正在等待服务组状态

服务组的状态可防止 VCS 使资源联机。

建议的操作：查看服务组的状态。

正在等待子资源

父资源的一个或多个子资源处于脱机状态。

建议的操作：先使子资源联机。

正在等待父资源

一个或多个父资源处于联机状态。

建议的操作：先使父资源脱机。

请参见第 145 页的“[疑难解答资源](#)”。

正在等待资源响应

资源正根据指示等待进入联机或脱机状态。VCS 指示代理运行资源的 **online** 入口点。

建议的操作：验证资源的 IState 属性。请参见 `/var/VRTSvcs/engine_A.log` and `/var/VRTSvcs/agent_A.log` 中的引擎和代理日志，以获取有关资源无法联机的原因的信息。

代理未运行

资源的代理进程未运行。

建议的操作：使用 `hastatus -summary` 查看该代理是否被列为 `faulted`。重新启动该代理：

```
# haagent -start resource_type -sys system
```

无效的代理参数列表

脚本正在接收不正确的参数。

建议的操作：验证脚本的参数是否正确。使用 `hares -display resource` 的输出查看 `ArgListValues` 属性的值。如果动态地更改了 `ArgList` 属性，则停止代理并重新启动它。

停止代理：

```
◆ # haagent -stop resource_type -sys system
```

重新启动代理

```
◆ # haagent -start resource_type -sys system
```

磁盘组代理的 Monitor 入口点返回 ONLINE，即使磁盘组已被禁用也是如此

这是预期的代理行为。VCS 假定正在从卷读取数据或将数据写入卷中，并且不将资源声明为脱机。这可防止可能由在两个主机上导入的磁盘组所导致的潜在数据损坏。

当完成所有 I/O 操作或关闭所有卷后，可以逐出禁用的磁盘组。然后，可以将磁盘组重新导入到同一系统中。重新导入禁用的磁盘组可能需要重新启动系统。

注意：如果大量磁盘的专用区域中的数据（包括内核日志、配置副本或头文件）无效或不可访问，则禁用磁盘组。如果不需要对磁盘专用区域进行任何更改，则卷可以执行读写操作。

I/O 防护故障排除

以下各节介绍了 I/O 防护问题的故障排除。查看症状和建议的解决方案。

正在逐出某个节点时，另一个节点不能加入集群

如果某个集群当前正在隔离（逐出）某个节点，在防护操作完成之前，该集群将阻止新节点加入其中。下面是针对新节点在控制台上显示的消息示例：

```
...VxFEN ERROR V-11-1-25 ... Unable to join running cluster
since cluster is currently fencing
a node out of the cluster.
```

如果在引导新节点时出现了这些消息，该节点上的 `vxfen` 启动脚本便会做出五次加入集群的尝试。

当 I/O 防护尝试失败时手动将节点加入到此集群中

- ◆ 如果 `vxfen` 脚本尝试允许该节点加入此集群失败，请使用以下命令重新启动 `vxfen` 驱动程序：

```
# /etc/init.d/vxfen.rc stop
# /etc/init.d/vxfen.rc start
```

如果该命令失败，请重新启动新节点。

当 SCSI TEST UNIT READY 命令失败时，`vxfcntlsthdw` 实用程序将失败

运行 `vxfcntlsthdw` 实用程序时，可能会看到一条如下所示的消息：

```
Issuing SCSI TEST UNIT READY to disk reserved by other node
FAILED.
Contact the storage provider to have the hardware configuration
fixed.
```

当有其他主机使用 SCSI-3 持久性保留保留磁盘时，磁盘阵列不支持 `SCSI TEST UNIT READY` 命令返回成功。如果系统模式选项“186 位”未启用，Hitachi Data Systems 99XX 阵列将出现此情况。

手动删除 SCSI-3 磁盘中的现有键

查看从磁盘删除由其他节点创建的特定注册键和保留键的以下过程。

注意：如果要清除所有现有键，请使用 `vxfcntlclearpre` 实用程序。

从磁盘删除注册键和保留键

- 1 创建包含磁盘访问名的文件：

```
# vi /tmp/disklist
```

例如：

```
/dev/rhdisk74
```

- 2 读取现有键：

```
# vxfenadm -s all -f /tmp/disklist
```

此命令的输出显示该键：

```
Device Name: /dev/rhdisk74

Total Number Of Keys: 1
key[0]:
[Numeric Format]: 86,70,66,69,65,68,48,50
[Character Format]: VFBEAD02
[Node Format]: Cluster ID: 48813 Node ID: 2
Node Name: unknown
```

- 3 如果您知道创建键（假定为 A1）的节点，请登录到该节点并输入以下命令：

```
# vxfenadm -x -kA1 -f /tmp/disklist
```

这样就会删除 A1 键。

- 4 如果您不知道该键是在哪个节点上创建的，则执行 5 到 7 来删除该键。
- 5 临时向磁盘注册另一个键 A2：

```
# vxfenadm -m -k A2 -f /tmp/disklist
```

```
Registration completed for disk path /dev/rhdisk74
```

6 通过用第二个键先占磁盘，将第一个键从磁盘删除：

```
# vxfenadm -p -kA2 -f /tmp/disklist -vA1

key: A2----- preempted the key: A1----- on disk

/dev/rhdisk74
```

7 删除在 5 中注册的临时键。

```
# vxfenadm -x -kA2 -f /tmp/disklist

Deleted the key : [A2-----] from device /dev/rhdisk74

磁盘上将不再存在注册键。
```

为防止潜在数据损坏，系统发生混乱

节点遇到裂脑情况并从集群逐出时，该节点将发生混乱并显示以下控制台消息：

```
VXFEN:vxfen_plat_panic: Local cluster node ejected from cluster to
prevent potential data corruption.
```

当节点由于所有专用互连失败或节点挂起而无法与对等节点保持心跳时，将遇到裂脑情况。请检查不同情况下的 I/O 防护行为以及要采取的纠正措施。

请参见第 149 页的“[I/O 防护在不同事件情形下的工作方式](#)”。

I/O 防护在不同事件情形下的工作方式

表 14-3 介绍了在不同的故障事件情形下，I/O 防护如何防止数据损坏。针对各种情况，查看操作员应采取的纠正操作。

表 14-3 I/O 防护情形

事件	节点 A：出现的情况	节点 B：出现的情况	操作员操作
两个专用网络均出现故障。	节点 A 会争夺大部分协调点。 如果节点 A 在对协调点的争夺中获胜，则节点 A 会将节点 B 逐出共享磁盘，然后继续运行。	节点 B 会争夺大部分协调点。 如果节点 B 在对协调点的争夺中失败，节点 B 会发生混乱并自行从集群中删除。	节点 B 被逐出集群之后，请先修复专用网络，然后再尝试使节点 B 重新返回到集群中。

事件	节点 A：出现的情况	节点 B：出现的情况	操作员操作
发生上述事件之后，两个专用网络重新恢复运行。	节点 A 继续运行。	节点 B 已经崩溃，它不能启动数据库，因为无法向数据磁盘写入数据。	还原专用网络之后，重新启动节点 B。
一个专用网络发生故障。	节点 A 会在控制台上输出有关 IOFENCE 的消息，并继续运行。	节点 B 会在控制台上输出有关 IOFENCE 的消息，并继续运行。	修复专用网络。修复网络之后，两个节点就可自动使用网络了。
节点 A 挂起。	<p>节点 A 由于某种原因非常忙或节点 A 处于内核调试程序中。</p> <p>节点 A 不再挂起或不再处于内核调试程序中时，所有已排队的针对数据磁盘的写入操作均会失败，因为节点 A 已被逐出。当节点 A 从 GAB 收到有关被逐出的消息之后，会发生混乱并自行从集群中删除。</p>	<p>节点 B 失去与节点 A 之间的心跳链接，并争夺大部分协调点。</p> <p>节点 B 在对协调点的争夺中获胜，并将节点 A 逐出共享数据磁盘。</p>	修复或调试挂起的节点，并重新启动该节点以便重新加入集群。

事件	节点 A：出现的情况	节点 B：出现的情况	操作员操作
<p>节点 A 和节点 B 以及专用网络断电。而协调点和数据磁盘的电源正常供电。</p> <p>节点的电源恢复供电，节点重新启动，但专用网络仍然无法正常运行。</p>	<p>节点 A 重新启动，I/O 防护驱动程序 (vxfen) 检测到节点 B 向协调点进行了注册。驱动程序没有在集群成员表中看到节点 B，因为专用网络已断开。这会导致 I/O 防护设备驱动程序阻止节点 A 加入集群。节点 A 控制台会显示以下消息：</p> <p>Potentially a preexisting split brain. Dropping out of the cluster. Refer to the user documentation for steps required to clear preexisting split brain.</p>	<p>节点 B 重新启动，I/O 防护驱动程序 (vxfen) 检测到节点 A 向协调点进行了注册。驱动程序没有在集群成员表中看到节点 A，因为专用网络已断开。这会导致 I/O 防护设备驱动程序阻止节点 B 加入集群。节点 B 控制台会显示以下消息：</p> <p>Potentially a preexisting split brain. Dropping out of the cluster. Refer to the user documentation for steps required to clear preexisting split brain.</p>	<p>解决已有的裂脑情况。</p> <p>请参见第 154 页的“防护启动报告已有的裂脑”。</p>

事件	节点 A：出现的情况	节点 B：出现的情况	操作员操作
<p>节点 A 崩溃而节点 B 停机。节点 B 恢复正常运行而节点 A 仍然停机。</p>	<p>节点 A 崩溃。</p>	<p>节点 B 重新启动并且检测到节点 A 向协调点进行了注册。驱动程序没有在集群成员表中看到节点 A。I/O 防护设备驱动程序会在控制台上输出以下消息：</p> <pre>Potentially a preexisting split brain. Dropping out of the cluster. Refer to the user documentation for steps required to clear preexisting split brain.</pre>	<p>解决已有的裂脑情况。</p> <p>请参见第 154 页的“防护启动报告已有的裂脑”。</p>
<p>包含三个协调点中的两个协调点的磁盘阵列断电。</p> <p>没有任何节点脱离集群成员集</p>	<p>只要任何节点都没有脱离集群，节点 A 就会继续运行。</p>	<p>只要任何节点都没有脱离集群，节点 B 就会继续运行。</p>	<p>打开发生故障的磁盘阵列的电源，以便后续网络分裂不会导致集群关闭；或者更换协调点。</p>
<p>包含三个协调点中的两个协调点的磁盘阵列断电。</p> <p>节点 B 正常脱离集群，磁盘阵列仍处于断电状态。正常脱离意味着干净的关闭，从而可以正确取消配置 vxfs。</p>	<p>节点 A 在集群中继续运行。</p>	<p>节点 B 已脱离集群。</p>	<p>打开发生故障的磁盘阵列的电源，以便后续网络分裂不会导致集群关闭；或者更换协调点。</p>

事件	节点 A：出现的情况	节点 B：出现的情况	操作员操作
<p>包含三个协调点中的两个协调点的磁盘阵列断电。</p> <p>节点 B 突然崩溃，或者在节点 A 和节点 B 之间出现网络分裂，且磁盘阵列仍处于断电状态。</p>	<p>节点 A 会争夺大部分协调点。节点 A 发生故障，因为三个协调点中只有一个是可用的。节点 A 会发生混乱并自行从集群中删除。</p>	<p>节点 B 由于崩溃或网络分裂已脱离集群。</p>	<p>打开发生故障的磁盘阵列的电源，并重新启动 I/O 防护驱动程序，使节点 A 能够向所有协调点进行注册；或者更换协调点。</p> <p>请参见第 156 页的“集群脱机时替换损坏的磁盘”。</p>

协调器磁盘的 I/O 防护键上的集群 ID 与本地集群的 ID 不匹配

如果您意外地将一个集群的协调器磁盘分配给另一个集群，则在启动 I/O 防护时，防护驱动程序会显示类似如下内容的错误消息：

```
000068 06:37:33 2bdd5845 0 ... 3066 0 VXFEN WARNING V-11-1-56
Coordinator disk has key with cluster id 48813
which does not match local cluster id 57069
```

该警告意味着集群 ID 为 57069 的本地集群拥有键。但是，ID 为 48813 的集群也在磁盘上拥有键，这表明集群 ID 为 48813 的集群中的节点可能会使用同一个协调器磁盘。

您可以运行下列命令来验证这些磁盘是否由另一个集群使用。在本地集群的其中一个节点上运行以下命令。例如，在 system01 上：

```
system01> # lltstat -C
57069

system01> # cat /etc/vxfentab
/dev/vx/rdmp/disk_7
/dev/vx/rdmp/disk_8
/dev/vx/rdmp/disk_9

system01> # vxfenadm -s /dev/vx/rdmp/disk_7
Reading SCSI Registration Keys...
Device Name: /dev/vx/rdmp/disk_7
Total Number Of Keys: 1
key[0]:
[Numeric Format]: 86,70,48,49,52,66,48,48
```

```
[Character Format]: VFBEAD00  
[Node Format]: Cluster ID: 48813 Node ID: 0 Node Name: unknown
```

其中 *disk_7*、*disk_8* 和 *disk_9* 表示设置中的磁盘名称。

建议的操作：必须为每个集群使用唯一的协调器磁盘组。如果其他集群没有使用这些协调器磁盘，则在本地集群中将其用作协调器磁盘之前，请使用 `vxfcntlclearpre` 命令清除键。

防护启动报告已有的裂脑

运行 `vxfen` 驱动程序，可防止已逐出的节点在从专用网络链接断开后到这些链路修复之前的这段时间内重新加入该集群。

例如，假设专用网络链接断开后，系统 1 和系统 2 的集群仍可正常运行。同时假设系统 1 为已逐出的系统。当系统 1 在专用网络链接还原之前重新启动时，其成员集配置不会显示系统 2；但是，当它尝试向协调器磁盘进行注册时，会发现系统 2 已向协调器磁盘注册。假设出现了有关系统 2 的此冲突信息，系统 1 将不加入此集群并通过 `vxfenconfig` 返回类似以下内容的错误消息：

```
vxfenconfig: ERROR: There exists the potential for a preexisting  
split-brain. The coordinator disks list no nodes which are in  
the current membership. However, they also list nodes which are  
not in the current membership.
```

```
I/O Fencing Disabled!
```

此外，在控制台上还会显示以下信息：

```
<date> <system name> vxfen: WARNING: Potentially a preexisting  
<date> <system name> split-brain.  
<date> <system name> Dropping out of cluster.  
<date> <system name> Refer to user documentation for steps  
<date> <system name> required to clear preexisting split-brain.  
<date> <system name>  
<date> <system name> I/O Fencing DISABLED!  
<date> <system name>  
<date> <system name> gab: GAB:20032: Port b closed
```

但是，当专用网络链接正常运行，而这两个系统都停机时，也会显示同样的错误消息，这时系统 1 重新启动，而系统 2 则无法恢复。从系统 1 中的集群来看，系统 2 可能仍注册在协调点上。

假定下列情形以了解在基于服务器的防护中的已有裂脑：

- 有三个 CP 服务器充当协调点。然后，这三个 CP 服务器中的其中一个变得无法访问。在此状态下，一个客户端节点离开集群，其注册无法从不可访问的 CP 服

务器中删除。当不可访问的 CP 服务器重新启动时，它有一个来自脱离了 VCS 的节点的失效注册。在这种情况下，任何新节点都不能加入集群。每个尝试加入集群的节点会从 CP 服务器中获得一个注册列表。一个 CP 服务器包括一个（先前离开的节点）的额外注册。这使得加入者节点推断，在它与失效注册所表示的节点之间存在已有的裂脑。

- 所有客户端节点已同时崩溃，因而未从 CP 服务器清除防护键。因此，节点重新启动时，`vxfen` 配置无法报告已有的裂脑。

这些情形类似于协调器磁盘的已有裂脑情况，在那些情况下，您可以通过运行 `vxfenclearpre` 命令来解决问题。在基于服务器的防护中需要使用 `cpsadm` 命令解决类似的问题。

请参见第 155 页的“清除已有的裂脑情况”。

清除已有的裂脑情况

查看有关 VxFEN 驱动程序如何检查已有裂脑情况的信息。

请参见第 154 页的“防护启动报告已有的裂脑”。

表 14-4 描述了如何根据您遇到的情形解决已有的裂脑情况：

表 14-4 清除已有裂脑情况的建议解决方案

情况	解决方案
实际潜在裂脑情况-系统 2 正在运行，系统 1 被逐出	<ol style="list-style-type: none"> 1 确定系统 1 是否正在运行。 2 如果系统 1 在正常运行，请将其关闭并修复专用网络链接以消除裂脑情况。 3 重新启动系统 1。
表面潜在裂脑情况-系统 2 宕机，系统 1 被逐出（已配置基于磁盘的防护）	<ol style="list-style-type: none"> 1 物理验证系统 2 是否停机。 验证系统当前是否已注册到协调点。 对协调器磁盘使用以下命令： # <code>vxfenadm -s all -f /etc/vxfentab</code> 该命令的输出会标识注册到协调器磁盘的键。 2 使用 <code>vxfenclearpre</code> 命令清除协调器磁盘上的键和所有共享磁盘组中的数据磁盘。 该命令删除 SCSI-3 注册和保留。 3 对系统 2 进行所有必要的修复。 4 重新启动系统 2。

情况	解决方案
表面潜在裂脑情况-系统 2 宕机，系统 1 被逐出 (已配置基于服务器的防护)	<p>1 物理验证系统 2 是否停机。 验证系统当前是否已注册到协调点。 对 CP 服务器使用以下命令：</p> <pre># cpsadm -s <i>cp_server</i> -a list_membership -c <i>cluster_name</i></pre> <p>其中，<i>cp_server</i> 是配置 CP 服务器所在的虚拟 IP 地址或虚拟主机名，<i>cluster_name</i> 是 SFHA Solutions 集群（应用集群）的 VCS 名称。 该命令列出向 CP 服务器注册的系统。</p> <p>2 使用 cpsadm 命令清除 CP 服务器上的键。cpsadm 命令可清除 CP 服务器上的注册：</p> <pre># cpsadm -s <i>cp_server</i> -a unreg_node -c <i>cluster_name</i> -n <i>nodeid</i></pre> <p>其中，<i>cp_server</i> 是 CP 服务器正在侦听的虚拟 IP 地址或虚拟主机名，<i>cluster_name</i> 是 VCS 的 VCS 名称，而 <i>nodeid</i> 指定 VCS 节点的节点 ID。确保在 CP 服务器上清除节点注册之前未在节点上运行防护。 删除所有失效注册后，要加入的节点将可以加入集群。</p> <p>3 对系统 2 进行所有必要的修复。</p> <p>4 重新启动系统 2。</p>

注册键会在协调器磁盘中丢失

如果协调器磁盘丢失已注册的键，则发生集群重新配置时集群可能出现混乱。

刷新丢失的键

- ◆ 使用 vxfenswap 实用程序用相同的磁盘替换协调器磁盘。vxfenswap 实用程序会注册在磁盘替换过程中丢失的键。

集群脱机时替换损坏的磁盘

如果磁盘已损坏或无法运行，且您希望切换到脱机集群中的某个新磁盘组，请执行以下过程。

在联机集群中，可以使用 vxfenswap 实用程序替换这些磁盘。

查看以下信息以替换协调器磁盘组中的协调器磁盘，或者销毁协调器磁盘组。

执行以下过程时，请注意下列事项：

- 添加磁盘时，请将磁盘添加到磁盘组 `vxfcntl`，然后重新测试该组是否支持 SCSI-3 持久性保留。
- 可以销毁协调器磁盘组，以便磁盘中不再保留任何注册键。然后，该磁盘可另作他用。

集群脱机时替换协调器磁盘组中的磁盘

- 1 以超级用户身份登录到其中一个集群节点上。
- 2 如果 VCS 正在运行，则将其关闭：

```
# hstop -all
```

确保所有节点已关闭端口 `h`。运行下列命令以验证端口 `h` 是否已关闭：

```
# gabconfig -a
```

- 3 在每个节点上停止 VCSMM 驱动程序：

```
# /etc/init.d/vcsmm.rc stop
```

- 4 停止每个节点上的 I/O 防护：

```
# /etc/init.d/vxfen.rc stop
```

这会删除磁盘上的所有注册键。

- 5 导入协调器磁盘组。`/etc/vxfendg` 文件中包括磁盘组的名称（通常为 `vxfcntl`），该磁盘组包含协调器磁盘，因此请使用以下命令：

```
# vxdg -tfC import 'cat /etc/vxfendg'
```

其中：

-t 指定只有在重新启动节点之后才可以导入磁盘组。

-f 指定强制执行导入，适用于一个或多个磁盘不可访问的情况。

-C 指定删除所有导入锁。

- 6 要从磁盘组删除磁盘，请使用 VxVM 磁盘管理员实用程序 `vxdiskadm`。

您还可以销毁现有协调器磁盘组。例如：

- 验证协调器属性是否设置为启用状态。

```
# vxdg list vxfcntl | grep flags: | grep coordinator
```

- 销毁协调器磁盘组。

```
# vxdg -o coordinator destroy vxfcntl
```

- 7 向节点添加新磁盘并将其初始化为 VxVM 磁盘。

然后，将新磁盘添加到 `vxfcntl` 磁盘组：

- 如果在步骤 6 中销毁了该磁盘组，请再次创建该磁盘组，然后将新磁盘添加到其中。
- 如果该磁盘组已经存在，请将新磁盘添加到其中。

```
# vxdg -g vxfcntl -o coordinator adddisk disk_name
```

- 8 测试重新创建的磁盘组是否符合 SCSI-3 持久性保留标准。

- 9 替换了协调器磁盘组中的磁盘之后，逐出该磁盘组：

```
# vxdg deport 'cat /etc/vxfendg'
```

- 10 在每个节点上，启动 I/O 防护驱动程序：

```
# /etc/init.d/vxfen.rc start
```

- 11 在每个节点上，启动 VCSMM 驱动程序：

```
# /etc/init.d/vcsmm.rc start
```

- 12 验证 I/O 防护模块是否已启动并且已启用。

```
# gabconfig -a
```

请确保端口 `b` 成员集存在于集群中所有节点的输出中。

请确保端口 `b` 和端口 `o` 成员集存在于集群中所有节点的输出中。

```
# vxfenadm -d
```

请确保在输出中未禁用 I/O 防护模式。

- 13 如有必要，在每个节点上重新启动 VCS：

```
# hstart
```

如果 `rcp` 命令或 `scp` 命令不能正常运行，`vxfsnap` 实用程序会退出

如果 `rcp` 命令或 `scp` 命令不能正常运行，则 `vxfsnap` 实用程序将显示错误消息。

恢复 vxfenswap 实用程序故障

- ◆ 验证 `rcp` 或 `scp` 是否正常运行。
确保在节点的 `.bashrc` 文件中没有使用 `echo` 或 `cat` 输出消息。
如果 `vxfenswap` 操作失败，且需要回滚该实用程序所做的任何更改，请使用 `vxfenswap -a cancel` 命令。

CP 服务器故障排除

所有 CP 服务器操作和消息均以易读的格式详细记录在 `/var/VRTScps/log` 目录下的日志中。这些项按日期和时间排序。这些日志可用于故障排除，或者用于查看承载 CP 服务器的系统上任何可能的安全问题。

下列文件包含可能有助于了解 CP 服务器和 CP 服务器排除故障的日志和文本文件：

- `/var/VRTScps/log/cpsrvr_[ABC].log`
- `/var/VRTSvcs/log/vcsauthserver.log`（与安全相关）
- 如果 CP 服务器上的 `vxcperv` 进程失败，则请查看下列诊断文件：
 - `/var/VRTScps/diag/FFDC_CPS_pid_vxcperv.log`
 - `/var/VRTScps/diag/stack_pid_vxcperv.txt`

注意：如果 CP 服务器上的 `vxcperv` 进程失败，这些文件将成为核心文件的补充。在这种情况下，VCS 会自动重新启动 `vxcperv` 进程。

`/var/VRTSvcs/log/vxfen/vxfend_[ABC].log` 文件包含一些日志，可为您了解和排查 VCS（客户端集群）节点上与防护相关的问题提供帮助。

请参见第 159 页的“[CP 服务器服务组相关问题故障排除](#)”。

请参见第 160 页的“[检查 CP 服务器的连接](#)”。

请参见第 160 页的“[在已设置基于服务器的防护的 VCS 节点上启动防护期间发生的问题](#)”。

请参见第 161 页的“[协调点联机迁移期间发生的问题](#)”。

CP 服务器服务组相关问题故障排除

如果在 CP 服务器配置后无法启动 CPSSG 服务组，请执行以下步骤：

- 验证 CPSSG 服务组及其资源是否有效以及是否在 VCS 配置中正确配置。
- 检查 VCS 引擎日志 (`/var/VRTSvcs/log/engine_[ABC].log`) 以查看是否有 CPSSG 服务组资源为 `FAULTED` 状态。

- 查看依赖关系图示例以确保正确配置所需的资源。

检查 CP 服务器的连接

可以使用 `cpsadm` 命令测试 CP 服务器的连接性。

只有设置了环境变量 `CPS_USERNAME` 和 `CPS_DOMAINTYPE`，才能在 VCS（客户端集群）节点上运行 `cpsadm` 命令。

检查 CP 服务器的连接

- ◆ 运行以下命令以检查 CP 服务器是否已启动并在进程级别上运行：

```
# cpsadm -s cp_server -a ping_cps
```

其中，`cp_server` 是 CP 服务器正在监听的虚拟 IP 地址或虚拟主机名。

SFHA Solutions 集群节点上基于服务器的防护的故障排除

`/var/VRTSvcs/log/vxfen/vxfend_[ABC].log` 文件包含一些日志，可为您了解和排查 SFHA Solutions 集群（应用集群）节点上与防护相关的问题提供帮助。

在已设置基于服务器的防护的 VCS 节点上启动防护期间发生的问题

表 14-5 VCS（客户端集群）节点上的防护启动问题

问题	说明和解决方案
在 VCS 上使用 <code>cpsadm</code> 命令会导致连接错误	如果在 VCS 上发出 <code>cpsadm</code> 命令后收到连接错误消息，请执行以下操作： <ul style="list-style-type: none"> ■ 确保 CP 服务器可从所有的 VCS 节点进行访问。 ■ 检查 VCS 节点使用的 CP 服务器虚拟 IP 或虚拟主机名以及端口号是否正确。检查 <code>/etc/vxfenmode</code> 文件。 ■ 确保正在运行的 CP 服务器使用的是相同的虚拟 IP/虚拟主机名和端口号。
授权故障	如果未在 CP 服务器配置中添加 CP 服务器的节点或用户，会出现授权故障，从而不允许 VCS（客户端集群）节点上的防护访问 CP 服务器以及在 CP 服务器上注册自身。如果向大多数协调点注册失败，将无法启动防护。 要解决此问题，请在 CP 服务器配置中添加 CP 服务器节点和用户，然后重新启动防护。
身份验证故障	如果已在 CP 服务器和 VCS（客户端集群）节点之间配置了安全通信，以下原因可能会导致身份验证错误： <ul style="list-style-type: none"> ■ 未在 CP 服务器和/或 VCS 上正确配置 Symantec Product Authentication Services (AT)。 ■ CP 服务器和 VCS 节点使用不同的根代理，并且未在身份验证代理之间建立信任：

协调点联机迁移期间发生的问题

在使用 `vxfsnwap` 实用程序进行协调点联机迁移期间，如果在验证任意集群节点中的协调点期间遇到故障，则操作会自动回滚。

在以下情况中可能会发生新协调点组验证故障：

- 未在所有 VCS 节点上更新 `/etc/vxfenmode.test` 文件，因为节点上的新协调点是从旧的 `/etc/vxfenmode.test` 文件中取得的。必须使用当前详细信息更新 `/etc/vxfenmode.test` 文件。如果 `/etc/vxfenmode.test` 文件不存在，`vxfsnwap` 将从 `/etc/vxfenmode` 文件中复制新协调点的配置。
- 不同 VCS 节点上的 `/etc/vxfenmode` 文件中列出的协调点不相同。如果在集群节点上的 `/etc/vxfenmode` 文件中列出了不同的协调点，则会由于在协调点快照检查期间发生故障而导致操作失败。
- 没有从一个或多个 VCS 节点到 CP 服务器的网络连接。
- 尚未在新 CP 服务器上添加集群、节点或 VCS 节点的用户，从而导致授权故障。

在发出 `vxfsnwap` 命令后进行 Vxfen 服务组活动

协调点代理从 `vxfsnconfig -l` 输出中读取协调点详细信息，并开始监视这些协调点上的注册内容。

因此，在执行 `vxfsnwap` 命令期间，如果用户正在更改 `vxfenmode` 文件，则协调点代理不会变为 `FAULTED` 状态，而是继续监视旧的协调点组。

只要未提交对 `vxfenmode` 文件的更改，或者未在 `vxfsnconfig -l` 输出中反映出新协调点集，协调点代理就会继续监视在每个监视周期从 `vxfsnconfig -l` 输出中读取的旧协调点集。

协调点代理的状态（`ONLINE` 或者 `FAULTED`）根据协调点的可访问性、这些协调点上的注册以及容错值而定。

在提交对 `vxfenmode` 文件的更改并反映在 `vxfsnconfig -l` 输出中后，协调点代理会读取新协调点集，并继续在新监视周期内对其进行监视。

疑难解答通知

有时，您可能在使用 VCS 通知时遇到问题。本节介绍最常见的问题和建议的操作。粗体文本提供问题的说明。

已配置通知程序，但在 SNMP 控制台上看不到陷阱

建议的操作：验证控制台所支持的 SNMP 陷阱的版本；VCS 通知程序发送 SNMP v2.0 陷阱。如果您将 HP OpenView Network Node Manager 用作 SNMP，请验证已使用 `xnmevents` 配置了 VCS 事件。合并 `vcs_trapd` 中的 VCS 事件后，如果这些

事件未在 OpenView Network Node Manager Event 配置中列出，您还可以尝试重新启动 OpenView 后台驻留程序 (ovw)。

默认情况下，通知程序假定团体字符串是公共的。如果您的 SNMP 控制台是使用其他团体配置的，则根据通知程序配置重新配置它。有关 NotifierMngr 的更多信息，请参见《Veritas Cluster Server Bundled Agents 参考指南》。

全局集群的故障排除与恢复

本主题介绍灾难声明的概念，并提供对使用全局集群的配置的故障排除提示。

灾难声明

全局集群中的某个集群转换为 FAULTED 状态时，由于无法再访问该集群，所以将根据导致故障的原因是远程集群出现裂脑、临时中断还是永久灾难来执行故障转移。

如果您选择对全局集群中的集群故障采取措施，则 VCS 会提示您声明故障的类型。

- *Disaster* (灾难)，表示主数据中心永久丢失
- *Outage* (中断)，表示主数据中心可能在某些时间会返回到其当前形态
- *Disconnect* (断开)，表示裂脑情况；两个集群已启动，但它们之间的链接已断开
- *Replica* (副本)，表示接管目标上的数据已与备份源一致，当服务组联机后，RVGPrimary 可以启动接管。此选项只适用于 VVR 环境。

可以选择要故障转移到本地集群的组，在这种情况下，VCS 将根据所选组的 FailOverPolicy 属性使组在节点上联机。它还在其他集群中将这组标记为 OFFLINE。如果您未选择任何要进行故障转移的服务组，则 VCS 除在关闭的集群上隐式地将服务组标记为脱机外，不会采取任何措施。

丢失的心跳和查询机制

任意两个集群之间的内部心跳和所有外部心跳丢失表明远程集群出现故障，或者两个集群之间的所有通信链路已断开（广域裂脑）。

VCS 查询集群以便确认已丢失其心跳的远程集群是否已真正关闭。这种机制被称为查询。在双集群配置中，如果一个连接器丢失了与另一个连接器的所有心跳，则肯定是远程集群出现了故障。如果有两个以上的集群，并且一个连接器丢失了与另一个集群所有心跳，则它会在声明该集群出现故障之前查询剩余连接器。如果其他连接器将该集群视为处于运行状态，则进行查询的连接器会将该集群转换为 UNKNOWN 状态，此过程会最大程度地减少假集群故障。如果所有连接器均报告该集群出现故障，则进行查询的连接器也将该集群视为出现故障，并将远程集群状态转换为 FAULTED。

VCS 警报

VCS 警报由警报 ID 标识，警报 ID 由以下元素组成：

- `alert_type` - 警报类型
- `cluster` - 在其上生成警报的集群
- `system` - 在其上生成警报的系统
- `object` - 为其生成警报的 VCS 对象的名称。此对象可以是集群或服务组。

按以下格式生成警报：

```
alert_type-cluster-system-object
```

例如：

```
GNOFAILA-Cluster1-oracle_grp
```

这是在集群 `Cluster1` 上为服务组 `oracle_grp` 生成的 `GNOFAILA` 类型的警报。

警报类型

VCS 可生成下列类型的警报。

- `CFAULT` - 表明集群已出现故障
- `GNOFAILA` - 表明全局组无法在其处于联机状态的集群内进行故障转移。如果将 `ClusterFailOverPolicy` 属性设置为“**Manual (手动)**”，并且广域连接器 (`wac`) 已进行正确配置且在出现故障时正在运行，则会显示此警报。
- `GNOFAIL` - 表明全局组无法在集群或远程集群内故障转移到任何系统。全局组可能无法故障转移到远程集群的一些原因如下：
 - `ClusterFailOverPolicy` 设置为“**Auto (自动)**”或“**Connected (连接)**”，且 VCS 无法确定将该组自动故障转移到的有效远程集群。
 - `ClusterFailOverPolicy` 属性设置为“**Connected (连接)**”，并且组出现故障的集群无法与该组 `ClusterList` 中列出的一个或多个远程集群通信。
 - 广域连接器 (`wac`) 未处于联机状态，或在组出现故障的集群中未进行正确配置。

管理警报

警报需要用户干预。您可以按下列方式响应警报：

- 如果可以忽略警报的原因，则使用 `Java` 控制台中的“**Alerts (警报)**”对话框或 `haalert` 命令删除警报。您必须提供关于删除警报原由的注释；VCS 会将此注释记录到引擎日志中。

- 对具有相关操作的管理警报采取相应的操作。
- 当发生警报的取消事件时，VCS 将删除或取消某些警报。

如果未执行上述任何操作，并且 VCS 引擎 (HAD) 正在集群中的至少一个节点上运行，则管理警报将继续存在。如果 HAD 未在集群中的任何节点上运行，则管理警报将丢失。

与警报相关的操作

本节介绍可以对下列类型的警报执行的操作：

- CFAULT - 出现此警报时，单击“**Take Action (采取操作)**”可指导您完成对全局组（全局组在集群出现故障前在集群中处于联机状态）进行故障转移的过程。
- GNOFAILA - 出现此警报时，单击“**Take Action (采取操作)**”可指导您完成将全局组故障转移到远程集群（全局组在该远程集群中配置为运行）的过程。
- GNOFAIL - 控制台没有为此警报提供相关操作

取消事件

当出现故障的集群返回到运行状态时，VCS 将删除 CFAULT 警报。

VCS 将删除 GNOFAILA 和 GNOFAIL 警报以响应下列事件：

- 出现故障的组的状态由 FAULTED 更改为 ONLINE。
- 组的故障已清除。
- 已从生成警报的集群中删除组。

启动时发生并发冲突

当您添加集群到服务组的 ClusterList 中时，VCS 可能报告并发冲突。并发冲突意味着服务组同时在两个节点上联机。

建议的操作：使服务组成为全局服务组之前，验证该服务组在每个集群中的状态。

对 steward 进程进行故障排除

启动 steward 时，它会阻止命令提示符并将消息输出到标准输出。要停止 steward，请从同一系统中的不同命令提示符运行以下命令：

如果 steward 以安全模式运行：`steward -stop -secure`

如果 steward 不是以安全模式运行：`steward -stop`

除标准输出之外，steward 可以将日志记录到它自己的日志文件：

- steward_A.log

- 3 在命令输出中查找以下内容：
 - 确保 **Product Name** 中列出您购买的组件的名称，例如 **Veritas Cluster Server**。如果命令输出中未返回产品名称，则说明您没有安装 VCS 密钥。
 - 如果输出中将 VCS 密钥的许可证类型显示为 **DEMO**，请确保 **Demo End Date** 显示的不是过去的日期。
 - 确保 **Mode** 属性显示正确的值。
 - 如果您已购买 **Global Cluster Option** 的许可证密钥，请确保其状态为 **Enabled**。
- 4 启动 VCS。如果 HAD 拒绝许可证密钥，请参见 **engine_A** 日志文件末尾的授权错误消息。

授权错误消息

本节列出了与授权相关的错误消息。这些消息记录在 `/var/VRTSvcs/log/engine_A.log` 文件中。

[Licensing] Insufficient memory to perform operation ([授权] 内存不足，无法执行操作)

系统没有足够的资源来执行授权操作。

[Licensing] No valid VCS license keys were found ([授权] 找不到有效的 VCS 许可证密钥)

在系统上找不到有效的 VCS 密钥。

[Licensing] Unable to find a valid base VCS license key ([授权] 找不到有效的基本 VCS 许可证密钥)

在系统上找不到有效的基本 VCS 密钥。

[Licensing] License key cannot be used on this OS platform ([授权] 无法在此操作系统平台上使用该许可证密钥)

此消息表明该许可证密钥适用于其他平台。例如，适用于 Windows 的许可证密钥用在了 AIX 平台上。

[Licensing] VCS evaluation period has expired ([授权] VCS 评估期已过)

VCS 评估版已过期。

[Licensing] License key can not be used on this system ([授权] 无法在此系统上使用该许可证密钥)

表明您安装了一个适用于其他系统的密钥（即节点锁定的密钥）。

[Licensing] Unable to initialize the licensing framework ([授权] 无法初始化该授权框架)

这是一个 VCS 内部消息。请致电 Veritas 技术支持。

[Licensing] QuickStart is not supported in this release ([授权] 此版本不支持 QuickStart)

此版本的 VCS 不支持 VCS QuickStart。

[Licensing] Your evaluation period for the feature has expired. This feature will not be enabled the next time VCS starts ([授权] 该功能的评估期已过。VCS 下次启动后该功能将处于禁用状态)

指定的 VCS 功能的评估版已过期。

索引

符号

- .cmdlog 文件 63
- .translog 文件 65
- /etc/init.d/vxvm-sysboot 文件 84
- /etc/vx/cbr/bk/diskgroup.dgid
 - dgid .binconfig 文件 73
 - dgid .cfgrec 文件 73
 - dgid .diskinfo 文件 72
 - dgid.dginfo 文件 72
- /etc/vx/log 日志记录目录 63, 65
- /var/adm/ras/vxconfigd.log 文件 83

A

- ACTIVE Plex 状态 25
- ACTIVE 卷状态 34

B

- BADLOG Plex 状态 34
- badlog 标志
 - DCO 清除 42
- 备份
 - 主节点存储检查点 103

C

- 重新启动禁用的卷 29, 53
 - 失效子磁盘 33
- 重新挂接磁盘 31
 - RAID-5 奇偶校验 35
- CLEAN Plex 状态 25
- cmdlog 文件 63
- 磁盘

- failng 标志 30
- 故障 33
- 故障原因 23
- 重新挂接 31
- 重新挂接出错的 31
- 磁盘组
 - 备份配置 71-72
 - 还原配置 71, 73
 - 解决备份冲突 75

- 配置备份文件 72
- 移动失败后恢复
 - 拆分或结合 40

存储检查点

- 数据卷错误
 - 恢复使用 103
 - 清除使用 103

错误

- 初次 RLINK 挂接期间 90
- 名称不匹配
 - 数据卷 95-96

配置 90

- 缺少数据卷 93
- 数据卷大小不匹配 94-95
- 修改 RVG 期间 93

错误消息 85

- RAID-5 Plex 未映射整个卷长 38
- 磁盘组不存在 40
- 代理日志 125
- 记录 83
- 卷不可启动 38
- 卷的同步由于 I/O 错误而停止 50
- 默认日志文件 83
- 某些子磁盘不可用且奇偶校验失效 38
- 启动脚本 84
- 启动时 140
- 无法自动导入组 71
- 消息目录 132
- 因发生错误而禁用 71
- 引擎日志 125
- 有两个磁盘组名称相同而磁盘组ID不同的备份 75

D

DCO

- 恢复卷 40
- 删除 badlog 标志 42

DCO 卷

- I/O 故障后恢复 51

- DEGRADED 卷状态 33

- DETACHED 卷内核状态 34

- DISABLED Plex 内核状态 25, 34

代理日志

- 格式 125
- 位置 125

调试消息

- 记录 83

E

- EMPTY Plex 状态 25
- ENABLED Plex 内核状态 25
- ENABLED 卷内核状态 34
- 二进制消息目录
 - 关于 132
 - 位置 132

F

- failing 标志
 - 清除 30
- 防护配置
 - 故障排除 117
- 辅助节点 SRL
 - 卷错误 104
 - 头错误 104
- 辅助节点存储检查点
 - 恢复使用 103

G

- 告急消息 85
- 根磁盘
 - 恢复 53
- 故障
 - 磁盘 33
 - 系统 32
- 故障排除
 - CVMVolDg 121
 - VCS 启动 140
 - 防护配置 117
 - 日志记录 125
 - 无法导入共享磁盘组 120
 - 许可证密钥 165
 - 主题概述 115, 119

H

- HAD 诊断 133
- 恢复
 - 磁盘 31

I

- I/O 防护
 - 测试和情形 149
- inode 列表错误 19
- install-db 文件 55
- IOFAIL Plex 状态 25

J

- 记录调试错误消息 83
- 进程 ID
 - 在命令日志记录文件中 63
 - 在事务日志记录文件中 65
- 禁用的文件系统
 - 事务处理 19
- 警告消息 85
- 镜像卷
 - 恢复 28
- 卷
 - DCO 恢复 40
 - RAID-5 数据丢失 32
 - 恢复 RAID-5 35
 - 恢复镜像 28
 - 列出不可启动的 24
 - 失效的子磁盘
 - 启动 39
 - 显示状态 24
 - 重新启动禁用的 29
 - 卷内核状态
 - DETACHED 34
 - 启用 34
 - 卷状态
 - DEGRADED 33
 - NEEDSYNCH 35
 - REPLAY 34
 - SYNC 34-35
 - 主动 34

K

- 客户端 ID
 - 在命令日志记录文件中 63
 - 在事务日志记录文件中 65
- 快照重新同步
 - 错误后恢复 50

L

- LOG Plex 状态 34
- 列表
 - 不可启动的卷 24

M

- MOVE 标志
 - 设置于 TUTILO 字段中 40
- msgcnt 字段 20
- 命令
 - 日志记录 63
 - 与事务关联 66

N

- NEEDSYNC 卷状态 35
- notification
 - troubleshooting 161

P

- parity (奇偶校验)
 - 失效 32
- plex
 - 恢复镜像卷 28
- Plex
 - 为 RECOVER 状态 28
 - 显示状态 24
 - 已定义 25
 - 映射问题 38
- Plex 内核状态
 - DISABLED 25
 - ENABLED 25
 - 禁用 34
- plex 状态
 - STALE 28
- Plex 状态
 - ACTIVE 25
 - BADLOG 34
 - CLEAN 25
 - EMPTY 25
 - IOFAIL 25
 - LOG 34
- 配置
 - 备份文件 72
 - 解决备份冲突 75
 - 为磁盘组备份 71
 - 为磁盘组还原 71
 - 针对磁盘组进行备份 72
 - 针对磁盘组进行还原 73
- 配置错误
 - 恢复 90

Q

- 奇偶校验
 - 对 RAID-5 的重新同步 35
 - 重建检查点设置 35
- 全局消息 ID 20
- 缺少数据卷错误 93

R

- RAID-5
 - 不可启动的卷 38
 - 分离子磁盘 33
 - 故障 32
 - 恢复过程 34
 - 恢复卷 35
 - 恢复日志 Plex 36
 - 奇偶校验重新同步 35
 - 启动恢复过程 34
 - 启动卷 38
 - 强制启动 39
 - 热重定位 35
 - 日志 Plex 的重要性 32
 - 失效的奇偶校验 32
 - 子磁盘移动恢复 37
- RECOVER 状态 28
- REPLAY 卷状态 34
- resources
 - troubleshooting 145
- RLINK
 - STALE 101
- root file system (根文件系统)
 - 损坏的 53
- RVG
 - 关联以清除 PASSTHRU 100
- 热重定位
 - RAID-5 35
 - 已定义 23
- 日志 Plex
 - 对于 RAID-5 的重要性 32
 - 恢复 RAID-5 36
- 日志故障 19
- 日志记录
 - 代理日志 125
 - 关联命令和事务 66
 - 目录 63, 65
 - 消息标记 125
 - 引擎日志 125
- 日志文件 159
 - vxconfigd 83
 - 默认 83

S

secondary_log_err 标志 104
 secondary_paused 标志 104
 service groups
 troubleshooting 142
 SRL 头错误 102
 SRL 溢出保护
 防止 101
 SYNC 卷状态 34-35
 失效的奇偶校验 32
 事务
 日志记录 65
 与命令关联 66
 事务处理禁用 19
 数据丢失
 RAID-5 32
 数据卷错误
 调整大小 94
 名称不匹配 95
 缺少 93
 数据卷大小
 调整大小 95

T

translog 文件 65
 troubleshooting
 notification 161
 resources 145
 service groups 142
 TUTILO 字段
 清除 MOVE 标志 40
 通知消息 86
 退化模式
 RAID-5 33

U

unmount 20

V

V-5-1-123 71
 V-5-1-1236 38
 V-5-1-1237 38
 V-5-1-2907 40
 V-5-1-569 71
 V-5-1-6012 75
 V-5-1-6840 50
 VCS
 troubleshooting resources 145

 troubleshooting service groups 142
 日志记录 125
 VX_FULLFSCK 19
 vxcmdlog
 控制命令日志记录 63
 vxconfigbackup
 备份磁盘组配置 72
 vxconfigd
 日志文件 83
 vxconfigd.log 文件 83
 vxconfigrestore
 还原磁盘组配置 73
 vxdco
 从 DCO 中删除 badlog 标志 42
 vxdctl
 更改调试日志记录的级别 83
 vxdg
 磁盘组移动失败后恢复
 拆分或结合 40
 vxedit
 清除磁盘 failing 标志 30
 vxedit 命令
 rename 关键字
 在 RLINK 恢复期间对卷进行重命名 92
 vxinfo 命令 24
 vxmend 命令 28
 vxplex 命令 36
 vxprint
 显示卷和 Plex 状态 24
 vxreattach
 重新挂接出错的磁盘 31
 vxsnap make
 故障后恢复 48
 vxsnap prepare
 故障后恢复 47
 vxsnap refresh
 失败后恢复 49
 vxsnap restore
 失败后恢复 49
 vxtranslog
 控制事务日志记录 65
 VxVM
 RAID-5 恢复过程 34
 恢复配置 55
 重新安装 55
 vxvol recover 命令 37
 vxvol resync 命令 35
 vxvol start 命令 28

vxvol 命令
 aslog 关键字
 关联 SRL 100
 assoc 关键字
 关联新卷 92
 dis 关键字
 分离 SRL 100
 start 关键字
 启动 SRL 100
 启动数据卷 98

W

文件
 磁盘组配置备份 72

X

系统
 重新安装 53
 系统故障 32
 消息标记
 关于 125
 写入时复制
 故障后恢复 50
 信息性消息 86
 许可证密钥
 故障排除 165

Y

已分离的 RAID-5 日志 Plex 36
 引导磁盘
 恢复 53
 引擎日志
 格式 125
 位置 125
 硬件故障
 恢复 23

Z

致命错误消息 85
 主节点 SRL
 错误
 清除 100
 头错误 101
 溢出恢复 101
 主节点存储检查点
 清除使用 103
 主节点主机
 故障恢复 97

状态
 显示卷和 Plex 24
 子磁盘
 RAID-5 移动后恢复 37
 标记为未失效 39
 失效
 启动卷 39