

IBM PowerHA SystemMirror for AIX

Standard Edition

V7.2

IBM

规划 PowerHA SystemMirror

IBM PowerHA SystemMirror for AIX

Standard Edition

V7.2



规划 PowerHA SystemMirror

注意:

在使用本资料及其支持的产品前，请阅读第 95 页的『声明』中的信息。

此版本适用于 IBM PowerHA SystemMirror 7.2 Standard Edition for AIX 以及所有后续发行版和修订版，直到在新版本中另有声明为止。

© Copyright IBM Corporation 2015.

目录

关于本文档	v
重要事项	v
AIX 区分大小写	v
ISO 9000	v
相关信息	v
规划 PowerHA SystemMirror	1
“规划 PowerHA SystemMirror”中的新增内容	1
规划过程概述	1
规划准则	1
消除单点故障: 配置 PowerHA SystemMirror 支持的冗余组件	2
规划过程概述	3
初始集群规划	5
规划集群节点	5
规划存储库磁盘和集群多点广播 IP 地址	6
规划磁盘电子篱笆	7
规划集群站点	8
规划集群安全性	9
应用程序规划	9
绘制集群图	13
主机名要求	14
规划集群网络连接	15
PowerHA SystemMirror 的一般网络注意事项	15
在 PowerHA SystemMirror 中监视	18
设计网络拓扑	19
规划通过 IP 别名的 IP 地址接管	22
规划其他网络条件	25
避免网络冲突	28
向集群图添加网络拓扑	29
规划共享磁盘和磁带设备	29
共享磁盘和磁带设备概述	29
选择共享磁盘技术	30
磁盘电源注意事项	30
规划非共享磁盘存储	30
规划共享磁盘安装	31
向集群图添加磁盘配置	32
将磁带机作为集群资源进行规划	32
规划共享的 LVM 组件	34
规划 LVM 组件	35
规划 LVM 镜像	36
规划磁盘访问	40
使用快速磁盘接管	41

使用配额和联机来增加数据可用性	42
将 NFS 与 PowerHA SystemMirror 结合使用	45
规划资源组	51
资源组概述	51
资源和资源组的常规规则	52
资源组的类型: 并发和非并发	52
资源组的启动、故障转移和回退策略	53
资源组属性	53
将资源组移动到其它节点	60
规划集群网络和资源组	61
规划资源组的并行或串行处理顺序	61
规划具有站点的集群中的资源组	62
规划复制资源	68
规划 Workload Manager	69
处理集群事件	70
规划站点和节点事件	71
规划 node_up 和 node_down 事件	72
网络事件	75
网络接口事件	76
集群范围状态事件	77
资源组事件处理和恢复	77
定制集群事件处理	79
定制事件的远程通知	83
定制出现警告前的事件持续时间	84
用户定义的事件	84
事件摘要和前同步信号	86
规划 PowerHA SystemMirror 客户机	87
运行 Clinfo 的客户机	87
不运行 Clinfo 的客户机	87
网络组件	87
应用程序和 PowerHA SystemMirror	88
应用程序和 PowerHA SystemMirror 概述	88
应用程序自动化: 最低程度手动干预	88
应用程序依赖性	90
应用程序干扰	91
应用程序的强大性	92
应用程序实施策略	92

声明	95
隐私策略注意事项	96
商标	97
索引	99

关于本文档

本文档介绍了 PowerHA[®] SystemMirror[®] for AIX[®] 软件。操作系统随附的文档 CD 上也提供了本信息。

重要事项

在本文档中使用了以下突出显示的约定：

粗体	标识命令、子例程、关键字、文件、结构、目录和系统预定义名称的其他项。也标识图形对象，例如用户选择的按钮、标签以及图标。
斜体字	标识由用户提供其实际名称或值的参数。
等宽字体	标识特定数据值示例、与您所看到的显示的文本相类似的文本示例、与您作为程序员所写的相类似的部分程序代码示例、来自系统的消息或您应实际输入的信息。

AIX 区分大小写

AIX 操作系统中的所有内容都区分大小写，即表示它区分大写和小写字母。例如，可以使用 **ls** 命令列出文件。如果您输入 **LS**，那么系统的响应会是找不到该命令。同样，**FILEA**、**FiLea** 和 **filea** 是三个不同的文件名，即使它们位于同一个目录中也是如此。为了避免引起执行不想要的操作，要始终确保使用正确的大小写字母。

ISO 9000

在本产品的开发和制造过程中，使用了 ISO 9000 注册质量体系。

相关信息

- PowerHA SystemMirror PDF 文档在以下主题提供：PowerHA SystemMirror 7.2 PDFs。
- PowerHA SystemMirror 发行说明在以下主题提供：PowerHA SystemMirror 7.2 release notes。

规划 PowerHA SystemMirror

配置和安装 PowerHA SystemMirror 之前，必须针对 AIX 操作系统规划其实现。

“规划 PowerHA SystemMirror”中的新增内容

阅读“规划 PowerHA SystemMirror”主题集合的新增信息或已进行显著更改的信息。

如何查看新增或更改的内容

在此 PDF 文件中，您可以会在左页边距中看到用来标识新增信息和更改信息的修订栏 (I)。

2015 年 12 月

以下信息是本主题集合的更新摘要：

- 已在第 6 页的『规划存储库磁盘和集群多点广播 IP 地址』主题中添加有关可确定的备份存储库磁盘数的信息。
- 已在第 10 页的『规划按需扩容』主题中添加有关资源优化高可用性 (ROHA) 功能的信息。
- 已在第 7 页的『规划磁盘电子篱笆』主题中添加有关磁盘电子篱笆功能的信息。

规划过程概述

可以使用 PowerHA SystemMirror Standard Edition for AIX 来规划数据中心中的高可用性，并且可以使用 PowerHA SystemMirror Enterprise Edition for AIX 来规划多站点高可用性和灾难恢复。

您在整个规划过程中的主要目标是消除单点故障。当某个重要集群功能由单个组件提供时，将存在单点故障。如果该组件发生故障，那么集群没有其他方法提供该功能，并且依赖该组件的应用程序或服务将变为不可用。

例如，如果某个关键应用程序的所有数据都位于单个磁盘上，且该磁盘发生故障，那么该磁盘将是整个集群的单点故障。客户机将无法访问此应用程序，直到该磁盘上的数据恢复为止。同样，如果动态应用程序数据存储在内部磁盘而非外部磁盘上，那么将不可能通过使其他集群节点接管这些磁盘来恢复应用程序。因此，要规划成功的集群，确定应用程序需要的必备逻辑组件（例如，文件系统和目录，它们可能包含应用程序数据和配置变量）是重要的前提条件。

应明确的一点是，虽然您的目标是消除所有单点故障，但是您可能必须作出一些让步。通常消除单点故障会关系到成本。例如，购买额外的硬件设备来作为主设备的备件会增加成本。可以将消除单点故障的成本与该组件发生故障时失去服务的成本相比较。再次重申，PowerHA SystemMirror 的目的是提供可满足未来处理需求的高性价比的高可用计算平台。

注：应尽快对集群组件故障采取补救措施，这一点很重要。根据您的配置，PowerHA SystemMirror 可能由于缺少资源而无法处理第二个故障。

规划准则

要设计可为贵组织提供最佳解决方案的集群需要进行仔细周密的规划。事实上，进行充分规划是构建成功 PowerHA SystemMirror 集群的关键。相比规划不足的集群，经过良好规划的集群更易于安装，可提供更高的应用程序可用性，并且需要更少的维护。

您可能需要规划环境中的其他过程。例如，如果您希望环境用于处理各种类型的故障，那么补丁管理和过程管理过程非常重要。

要使关键应用程序具有高可用性，所有相关资源都不应是单点故障。在您设计 PowerHA SystemMirror 集群时，您的目标是确定并解决所有潜在的单点故障。应提出的问题包括：

- 哪些应用程序服务需要具有高可用性？这些服务的优先级如何？
- 发生某个故障的代价与用于消除该故障可能性的必需硬件相比如何？
- PowerHA SystemMirror 可以支持的冗余硬件和软件组件的最大数量是多少？
- 这些服务的所需可用性如何？它们是否需要 24 小时/天 x 7 天/周始终保持可用，还是只要 8 小时/天 x 5 天/周保持可用就足矣？
- 如果中断这些服务的可用性会发生什么后果？
- 为替换发生故障的资源分配了多长时间？在发生故障后进行操作时，可以接受什么程度的性能降级？
- 会将哪些故障自动检测为集群事件？需要为哪些故障编写定制代码来检测故障并触发集群事件？
- 团队实施和维护集群的熟练程度如何？

要规划、实施和维护成功的 PowerHA SystemMirror 集群，需要在贵组织的很多团队中进行持续沟通。理想情况是，您应召集以下代表（适当时）来协助 PowerHA SystemMirror 规划会议：

- 网络管理员
- 系统管理员
- 数据库管理员
- 应用程序编程人员
- 支持人员
- 最终用户

PowerHA SystemMirror 支持多种配置，这为您提供了极大的灵活性。有关为您集群设计最高可用性级别的信息，请参阅 IBM 白皮书 *High Availability Cluster Multiprocessing Best Practices*。

相关参考：

『消除单点故障：配置 PowerHA SystemMirror 支持的冗余组件』
PowerHA SystemMirror 软件提供了很多可避免单点故障的选项。

相关信息：

 [High Availability Cluster Multiprocessing Best Practices](#)

消除单点故障：配置 PowerHA SystemMirror 支持的冗余组件

PowerHA SystemMirror 软件提供了很多可避免单点故障的选项。

下表汇总了潜在的单点故障，并描述了如何通过配置冗余硬件和软件集群组件来消除它们。

集群组件	消除成为单点故障	PowerHA SystemMirror 支持
节点	使用多个节点	最多 16 个。
电源	使用多个电路或不间断电源	需要的数量。
网络	使用多个网络连接节点	最多 48 个。
网络接口、设备和标签	使用冗余网络适配器	最多 256 个。
TCP/IP 子系统	使用网络连接相邻节点和客户机	需要的数量。
磁盘适配器	使用冗余磁盘适配器	需要的数量。
控制器	使用冗余磁盘控制器	需要的数量。
磁盘	使用冗余硬件和磁盘镜像和/或条带化	需要的数量。
应用程序	分配一个节点以进行应用程序接管，用于配置应用程序监视器或配置节点位于多个站点上的集群。	灵活的配置策略，以在站点内和站点之间获得高可用性。
站点	使用多个站点进行灾难恢复。	最多两个站点。
资源组	使用资源组指定一组实体的执行方式。	每个集群最多 64 个。
集群资源	使用多个集群资源。	对于 Clinfo 守护程序最多 128 个（集群中可以存在更多个）。
虚拟 I/O 服务器 (Virtual I/O Server, VIOS)	使用冗余 VIOS	需要的数量。
HMC	使用冗余 HMC	2
主管集群节点的受管系统	为每个集群节点使用单独的受管系统	16 个节点
集群存储库磁盘	使用 RAID 保护	对于能够在故障后更换磁盘的每个站点，支持一个活动存储库磁盘。您必须具有一个备用磁盘，可用于替换当前集群中发生故障的存储库磁盘。

相关参考:

第 1 页的『规划准则』

要设计可为贵组织提供最佳解决方案的集群需要进行仔细周密的规划。事实上，进行充分规划是构建成功 PowerHA SystemMirror 集群的关键。相比规划不足的集群，经过良好规划的集群更易于安装，可提供更高的应用程序可用性，并且需要更少的维护。

规划过程概述

本主题描述 PowerHA SystemMirror 集群的规划步骤。

步骤 1: 规划高可用应用程序

在此步骤中，您将规划集群的核心：将具有高可用性的应用程序、应用程序需要的资源的类型、节点数、共享 IP 地址，以及磁盘共享方式（非并发或并发访问）。您的目标是提出系统的高级视图，作为集群设计的起点。在作出这些最初的决策后，请开始绘制集群图。“初始集群规划”描述了此规划过程步骤。

步骤 2: 规划集群拓扑

在此步骤中，您将决定集群和节点的名称。您还可以选择确定站点的名称，以及确定哪些节点属于哪个站点。“初始集群规划”描述了此规划过程步骤。

步骤 3: 规划站点

在此步骤中，您将确定站点将使用延伸集群还是链接集群。延伸集群包含相同地理位置的站点中的节点。延伸集群必须共享一个存储库磁盘。链接集群包含位于不同地理位置的站点中的节点。链接集群使用单独的存储库磁盘。

步骤 4: 规划集群网络连接

在此步骤中，您将规划连接系统中节点的网络。您应先研究在 PowerHA SystemMirror 环境中与 TCP/IP 和点到点网络相关的问题。“规划集群网络连接”描述了此规划过程步骤。

步骤 5: 规划共享磁盘设备

在此步骤中，您将规划集群的共享磁盘设备。您应决定要在集群中使用的磁盘存储技术，并研究在 PowerHA SystemMirror 环境中与这些技术相关的问题。“规划共享磁盘和磁带设备”描述了此规划过程步骤。

步骤 6: 规划共享 LVM 组件

在此步骤中，您将规划集群的共享卷组。您将先检查在 PowerHA SystemMirror 环境中与 LVM 组件相关的问题。“规划共享 LVM 组件”描述了此规划过程步骤。

步骤 7: 规划资源组

规划资源组整合了您在先前步骤中生成的所有信息。此外，您需要决定要使用依赖资源组还是特定运行时策略来使特定相关资源组保留在同一节点或不同节点上。“规划资源组”描述了此规划过程步骤。

步骤 8: 规划集群事件处理

在此步骤中，您将规划集群的事件处理。“规划集群事件”描述了此规划过程步骤。

步骤 9: 规划 PowerHA SystemMirror 客户机

在此步骤中，您将研究与 PowerHA SystemMirror 客户机相关的问题。“规划 PowerHA SystemMirror 客户机”描述了此规划过程步骤。

相关参考:

第 5 页的『初始集群规划』

本部分描述为规划 PowerHA SystemMirror 集群以使应用程序具有高可用性所应采取的初始步骤。

第 15 页的『规划集群网络连接』

本节描述为 PowerHA SystemMirror 集群规划网络支持。

第 51 页的『规划资源组』

以下主题描述如何规划 PowerHA SystemMirror 集群内的资源组。

第 34 页的『规划共享的 LVM 组件』

本节描述为 PowerHA SystemMirror 集群规划共享卷组。

第 29 页的『规划共享磁盘和磁带设备』

本节讨论在 PowerHA SystemMirror 集群中配置共享外部磁盘之前要考虑的信息，并提供有关将磁带机作为集群资源进行规划和配置的信息。

第 70 页的『处理集群事件』

以下主题描述 PowerHA SystemMirror 集群事件。

第 87 页的『规划 PowerHA SystemMirror 客户机』

以下主题讨论 PowerHA SystemMirror 客户机的规划注意事项。本步骤是在继续安装 PowerHA SystemMirror 软件之前的最后一步。

第 8 页的『规划集群站点』

PowerHA SystemMirror 集群可以在单个站点或多个站点中用于灾难恢复。

相关信息:

集群事件期间的资源组行为

初始集群规划

本部分描述为规划 PowerHA SystemMirror 集群以使应用程序具有高可用性所应采取的初始步骤。

在开始 PowerHA SystemMirror 规划之前，请确保您了解与 PowerHA SystemMirror 相关的概念和术语。

PowerHA SystemMirror 集群为会话关键型应用程序提供高可用性环境。在很多组织中，这些应用程序必须始终保持可用。例如，PowerHA SystemMirror 集群可以运行为客户机应用程序提供服务的数据库服务器程序，使该程序对于向服务器程序发送查询的客户机具有高可用性。

相关信息：

PowerHA SystemMirror 概念

规划集群节点

对于每个重要的应用程序，请留意该应用程序所需的资源，包括应用程序的处理和数据存储要求。

例如，当规划集群的大小时，请包括足够的节点以处理当某个节点发生故障后应用程序的处理要求。

确定集群节点数量时，请记住以下注意事项：

- PowerHA SystemMirror 集群可以由 IBM® Power Systems™ 服务器的任意组合组成。请确保所有集群节点均不共享可能成为单点故障的组件（例如，电源）。同样，也不要将节点放在同一机架中。
- 创建由执行类似功能或共享资源的节点组成的小型集群。简单的小集群易于设计、实施和维护。
- 出于性能原因，可能需要使用多个节点来支持同一应用程序。为提供相互接管服务，在某种情况下必须将应用程序设计为允许在同一节点上运行应用程序的多个实例。

例如，如果应用程序需要动态数据位于称为 */data* 的目录中，那么应用程序可能无法在同一处理器上支持多个实例。对于此类应用程序（在非并发环境中运行），请尝试对数据分区，以便可以运行应用程序的多个实例，每个实例访问一个唯一的数据库。

并且，如果应用程序支持的配置文件允许管理员指定应用程序的 *instance1* 的动态数据位于 *data1* 目录中，*instance2* 动态数据位于 *data2* 目录中，以此类推，那么可能支持应用程序的多个实例。

- 某些配置（包括集群设计中的其他节点）会提高集群提供的可用性级别。特定配置还在规划节点故障转移和重新集成方面提供更大灵活性。

最可靠的集群节点配置至少具有一个备用节点。

- 选择具有足够 I/O 插槽的集群节点，用以支持冗余网络接口卡和磁盘适配器。

请记住，虽然由多个节点组成的集群仍然比单个节点组成的集群昂贵，但是如果不进行规划以支持冗余硬件（例如，用于网络和磁盘适配器的足够 I/O 插槽），那么集群将无法具有更好的可用性。

- 请使用具有相似处理速度的节点。
- 请使用具有足够 CPU 周期和 I/O 带宽的节点，以使生产应用程序在负载达到峰值时也能够运行。请记住，节点应具有使 PowerHA SystemMirror 能够运行的足够容量。

要为此进行规划，请为生产应用程序设定基准或建模，并列示最大期望负载的参数。然后选择当运行生产应用程序时繁忙率不超过 85% 的 PowerHA SystemMirror 集群的节点。

当创建集群时，请向其分配名称。PowerHA SystemMirror 会将该名称与 PowerHA SystemMirror 分配的集群标识相关联。

规划存储库磁盘和集群多点广播 IP 地址

PowerHA SystemMirror 集群可通过多点广播或单点广播联网执行。要选择集群的多点广播方式，可规划和提供多点广播 IP 地址以在集群中通信。缺省情况下，如果在部署集群时未提供多点广播 IP 地址，那么 PowerHA 会部署基于单点广播（正常 TCP/IP 套接字通信）的集群。

集群存储库磁盘

- 1 针对标准集群和延伸集群，每个集群必须具有一个活动存储库磁盘。针对标准集群和延伸集群，可为每个集群确定多达 6 个备份存储库磁盘。针对链接集群，每个站点必须具有一个活动存储库磁盘。针对链接集群，可为每个站点确定多达 6 个备份存储库磁盘。

PowerHA SystemMirror 使用共享磁盘来存储 Cluster Aware AIX (CAA) 集群配置信息。必须为集群存储库磁盘分配至少 512 MB 并且不超过 460 GB 的磁盘空间。此配置可自动在所提供的磁盘保持高可用性。此功能需要对于集群包含的所有节点都可用的专用共享磁盘。此磁盘不能用于应用程序存储或任何其他用途。

规划要用作存储库磁盘的磁盘时，必须规划备份或替换磁盘，当主要存储库磁盘发生故障时就可以使用这些磁盘。备份磁盘的大小和类型必须与主要磁盘相同，但它可以位于不同的物理存储磁盘中。使用备份磁盘信息来更新管理过程和文档。还可以将正在使用的存储库磁盘替换为一个新磁盘，以增加大小或更改为其他存储子系统。要替换存储库磁盘，可以使用 SMIT 界面。

注：如果用作存储库磁盘的共享磁盘为映射的虚拟 SCSI (vSCSI) 磁盘，那么必须将磁盘作为 vSCSI 磁盘映射到集群中的所有节点。在集群中所有节点之间，vSCSI 磁盘的映射必须相同。例如，无法使用 vSCSI 方法将存储库磁盘映射到集群中一个节点，也无法使用 N 端口标识虚拟化 (NPIV) 方法将相同磁盘映射到集群中其他节点。

集群多点广播 IP 地址

您可以使用多点广播 IP 地址，进行集群监视和通信。您可以在创建集群时指定此地址，也可以在同步初始集群配置时自动生成一个地址。

注：缺省机制使用单点广播通信，不需要进行额外配置。但是，如果需要使用多点广播通信，那么必须继续读取，并确保针对网络设备启用了多点广播通信。

如果决定使用多点广播，那么 PowerHA SystemMirror 在集群中的主机之间使用基于多点广播的通信。您环境的网络必须允许多点广播 IP 包在集群中的主机间流动。要验证环境中的节点是否支持基于多点广播的通信，请使用 **mping** 命令。在您开始在环境中使用 PowerHA SystemMirror 之前运行 **mping** 命令。

注：一些网络交换机允许多点广播信息包在停止之前先流动一小段时间。因此，至少执行 **mping** 测试 5 分钟并确保网络光纤允许多点广播信息包流动而不产生任何问题会很关键。此外，对交换机进行级联操作时，通常，交换机需要进行其他配置以路由多点广播信息包。要配置多点广播信息包流，请参阅交换机供应商提供的文档，以配置多点广播信息包流。

多点广播地址也称为 D 类地址。目标地址以 1110 开头的 IP 数据报都是 IP 多点广播数据报。其余 28 位标识了发送数据报所在的多点广播组。您必须将内核配置为接收发送至特定多点广播组的包，这会使主机在所指定的接口加入组。

请勿使用以下多点广播组：

224.0.0.1

这是全主机组。如果 ping 该组，那么网络上所有支持多点广播的主机都会回答，原因是每个支持多点广播的主机都必须在启动时在其支持多点广播的所有接口上加入该组。

224.0.0.2

这是全路由器组。所有多点广播路由器都必须在其所有支持多点广播的接口上加入该组。

224.0.0.4

这是全部 DVMRP 路由器。

224.0.0.5

这是全部 OSPF 路由器。

224.0.0.13

这是全部 PIM 路由器。

注：范围 224.0.0.0 - 224.0.0.255 留作本地用途（例如，管理和维护任务），并且多点广播路由器始终不会转发它们接收到的数据。与此类似，239.0.0.0 - 239.255.255.255 之间的范围预留作管理作用域限定。这些特殊的多点广播组定期在 Assigned Numbers RFC 中发布。

PowerHA SystemMirror 7.1.2 或更高版本支持 IP V6 (IPv6)，但是，您无法显式指定 IPv6 多点广播地址。CAA 使用源自 IP V4 (IPv4) 多点广播地址的 IPv6 多点广播地址。要确定 IPv6 多点广播地址，将使用逻辑 OR 运算符将标准前缀 `0xFF05` 与 IPv4 地址的十六进制表示形式结合。例如，IPv4 多点广播地址是 `228.8.16.129` 或 `0xE4081081`。逻辑 OR 操作与标准前缀完成的变换为 `0xFF05:: | 0xE4081081`。因此，生成的 IPv6 多点广播地址为 `0xFF05::E408:1081`。

相关信息：

存储库磁盘故障

使用 SMIT 替换存储库磁盘

对多点广播进行故障诊断

测试网络中的多点广播

规划磁盘电子篱笆

磁盘电子篱笆是 PowerHA SystemMirror 中提供的隔离策略的其中一个功能。

要使用磁盘电子篱笆选项，必须确定关键资源组，且存储子系统必须支持 SCSI-3 持久性预留和 PR_shared 的 ODM reserve_policy。此策略适用于属于卷组和资源组的所有磁盘。

预留类型

磁盘必须具有所需预留类型 Write Exclusive All Registrant (WEAR)。您可以运行以下 **clmgr** 命令以检查磁盘是否支持 SCSI-3 功能 (WEAR - 类型 7h) 以支持 PowerHA SystemMirror 磁盘电子篱笆：

```
clmgr scsivr_capability query physical_volume <disk>
clmgr scsivr_capability query volume_group <vg>
```

其中，*disk* 是磁盘的名称，*vg* 是卷组的名称。

reserve_policy

定义是否在磁盘上运行预留方法。PR_shared 的 reserve_policy 是为磁盘应用共享主机方法的所需策略。要查看磁盘的属性，请运行 **lsattr -RI <diskname> -a reserve_policy** 命令。

要使用磁盘电子篱笆功能，必须指定关键资源组。指定的关键资源组必须满足以下条件：

- 无法将关键资源组作为子代添加到任何 parent_child、start_after 或 stop_after 依赖性关系中。
- 关键资源组必须具有集群中作为参与节点的所有节点。
- 关键资源组无法使用仅在主节点上联机启动策略。关键资源组可使用任何其他启动策略。

针对具有正在运行的集群服务和启用了磁盘电子篱笆的卷组磁盘的 reserve_policy，必须使用以下设置：

表 1. 磁盘设置

选项	值
配置的预留策略	PR_shared
有效预留策略	PR_shared
预留状态	SCSI PR 预留 (Write Exclusive All Registrant)

SCSI-3 持久性预留在 HyperSwap® 磁盘上不受支持。如果正使用 HyperSwap 磁盘，那么验证检查会显示错误消息。

使用 **Majority** 选项的合并策略现在可能无法正确用于 SCSI-3 持久性预留，因为来自主要站点的节点可能已被隔离。在此情况下，使用**主要**选项的合并策略不受支持。

相关信息：

配置隔离策略

对磁盘电子篱笆进行故障诊断

规划集群站点

PowerHA SystemMirror 集群可以在单个站点或多个站点中用于灾难恢复。

如果您有多个站点，那么可以将以下 PowerHA SystemMirror Enterprise Edition 功能用于灾难恢复：

- PowerHA SystemMirror Enterprise Edition for AIX 包括用于支持各种存储子系统提供的复制技术的选项。
- 用于 GLVM 的 PowerHA SystemMirror Enterprise Edition for AIX 提供通过 TCP/IP 网络完成的基于主机的复制。

PowerHA SystemMirror 7.1.2 或更高版本支持站点的不同类型定义和特定于站点的高可用性和灾难恢复 (HADR) 策略。您可以在 PowerHA SystemMirror Standard Edition for AIX 和 PowerHA SystemMirror Enterprise Edition for AIX 中定义多个站点。

PowerHA SystemMirror 使用 Cluster Aware AIX (CAA) 来进行集群通信和集群运行状况管理。

可以使用 PowerHA SystemMirror 管理界面来创建以下多站点解决方案：

延伸集群

包含位于同一地理位置的站点中的节点。延伸集群必须在该站点的所有节点中共享存储库磁盘。延伸集群不支持通过存储复制管理实现的 HADR。要使用延伸集群，您的网络环境必须支持基于多点广播的通信。

链接集群

包含位于不同地理位置的站点中的节点。链接集群使用单独的存储库磁盘。链接集群支持跨站点 LVM 镜像和 HyperSwap。在链接集群中，CAA 使用单点广播信息包来对独立 CAA 集群之间的站点进行通信和管理。

规划资源和站点策略

PowerHA SystemMirror 尝试确保在一个站点上使资源组的主要实例保持联机状态，而在另一个站点上使次要实例保持联机状态。规划在哪个站点上配置哪些节点，以及希望在何处运行活动应用程序，以便可以对资源组策略进行相应规划。

注： 站点仅在 PowerHA SystemMirror Enterprise Edition 和 Standard Edition 7.1.2 或更高版本中受支持。复制管理仅在 PowerHA SystemMirror Enterprise Edition 中受支持。

对 PowerHA SystemMirror 定义的所有资源必须具有唯一的名称。服务 IP 标签、卷组和资源组名称也必须在集群内唯一，并且互不相同。资源的名称应该与其服务的应用程序相关，还应该与任何相应设备相关。例如，运行 WebSphere® 实例的资源组的服务地址可以命名为 `websphere_service_address`。

规划集群安全性

PowerHA SystemMirror 通过控制用户对 PowerHA SystemMirror 的访问以及为节点间通信提供安全性来提供集群安全性。

连接认证

PowerHA SystemMirror 提供连接认证来保护集群节点间的 PowerHA SystemMirror 通信。此连接认证也称为标准认证。标准认证包括按 IP 地址和主机名的已验证连接，并限制可使用 root 特权运行的命令。此方式为远程命令执行使用最小特权原则，这确保没有任意命令能够使用 root 特权在远程节点上运行。会将一组选定的 PowerHA SystemMirror 命令视为可信，并且允许以 root 用户身份运行这些命令。所有其他命令均以 nobody 用户身份运行。这就消除了节点间通信对于 `/rhosts` 的依赖性。

您还可以为节点间通信配置虚拟专用网 (VPN)。如果使用 VPN，请为 VPN 隧道使用持久 IP 标签。

安全性配置

PowerHA SystemMirror 使用 Cluster Aware AIX (CAA) 功能来为集群中节点间的脉动信号和同步创建安全通信路径。

可以使用以下 CAA 方法来为集群中的节点创建集群安全凭证。

自签名 PowerHA SystemMirror 生成安全凭证。

安全证书和专用密钥对

PowerHA SystemMirror 使用您提供的现有安全证书和专用密钥对。

Secure Shell (SSH)

PowerHA SystemMirror 使用已为环境中的 SSH 通信配置的密钥。

消息认证和加密

PowerHA SystemMirror 为在集群节点之间发送的 PowerHA SystemMirror 消息提供安全性，如下所述：

- 消息认证可确保消息的生成和完整性。
- 消息加密可在传输数据时更改数据的外观，并在认证消息的节点接收数据时将其恢复为原始格式。
- 将根据安全性级别“低”、“中”或“高”对消息进行加密或散列。“低”安全性级别将仅对几条消息进行散列，而“高”安全级别将对消息进行加密。

PowerHA SystemMirror 支持以下类型的用于消息认证和加密的加密密钥：

- 使用数据加密标准 (DES) 的消息摘要 5 (MD5)
- 使用三重 DES 的 MD5
- 使用高级加密标准 (AES) 的 MD5。

请选择与贵组织所用安全性方法兼容的加密算法。

应用程序规划

在开始规划应用程序之前，请务必了解您应用程序的数据资源以及这些资源在集群中的位置，才能提供在某个节点失败时可正确处理这些资源的解决方案。

要避免故障，您必须充分了解应用程序在单节点和多节点环境中的行为。请勿假设应用程序在不利条件下的性能。

请使用具有足够 CPU 周期和 I/O 带宽的节点，以使生产应用程序在负载达到峰值时也能够运行。请记住，节点应具有使 PowerHA SystemMirror 能够运行的足够容量。

要为此进行规划，请为生产应用程序设定基准或建模，并列示最大期望负载的参数。然后选择当运行生产应用程序时繁忙率不超过 85% 的 PowerHA SystemMirror 集群的节点。

您可以为应用程序配置多个应用程序监视器并指示 PowerHA SystemMirror 执行以下两项操作：

- 监视进程的终止或影响应用程序的较细微问题
- 自动尝试重新启动应用程序并在重新启动尝试失败时采取适当操作（通知或故障转移）。

本部分说明如何记录有关您应用程序的所有关键信息，并开始绘制集群图。

请记住以下准则，以确保在 PowerHA SystemMirror 集群环境中为您的应用程序正确提供服务：

- 对应用程序及其数据进行布局，以便仅数据位于共享外部磁盘上。此安排不仅可避免软件许可证违例，而且可简化故障恢复。
- 如果您打算在集群中的父子依赖资源组中包括多层应用程序，那么请参阅“多层应用程序规划注意事项”一节。如果您打算使用位置依赖性使某些应用程序保持在同一节点或在不同节点上，那么请参阅“资源组依赖性”一节。
- 编写用于在集群节点上启动和停止应用程序的强大脚本。特别是启动脚本必须能够将应用程序从异常结束（例如，电源故障）中恢复。在包括 PowerHA SystemMirror 软件之前，请确保该脚本在单节点环境中正常运行。
- 确认应用程序许可要求。某些供应商针对每个运行应用程序的处理器需要一个唯一许可证，这意味着您必须通过在安装应用程序时将特定于处理器的信息合并到应用程序来保护针对应用程序的许可证。这样，即使 PowerHA SystemMirror 软件正确处理了节点故障，它也可能由于集群内可用应用程序许可证数量的限制，而无法在故障转移节点上重新启动应用程序。要避免此问题，请确保您对于可能运行应用程序的集群内的每个系统单元均具有一个许可证。
- 确保应用程序在单节点环境中成功运行。在集群中调试应用程序比在单处理器上进行调试更困难。
- 如果您需要并发访问，请验证应用程序是否使用专用锁定机制。

相关参考：

第 13 页的『多层应用程序的规划注意事项』

使用多层应用程序的业务配置可以使用父子依赖性资源组。例如，数据库必须在应用程序控制器之前联机。在此情况下，如果数据库关闭并且移动到其他节点，那么必须先将包含应用程序控制器的资源组脱机，再将其在集群的任何节点上联机。

第 56 页的『资源组依赖性』

PowerHA SystemMirror 提供了多种配置，您可在这些配置中指定希望在启动、故障转移和回退时维护的资源组之间的关系。

1 规划按需扩容

1 按需扩容 (CoD) 是一种资源优化高可用性 (ROHA) 功能，PowerHA SystemMirror 使用此功能动态管理应用程序的硬件资源。通过 CoD，您可以激活预先安装的非活动处理器，且不必由于资源需求更改而支付费用。

1 CoD 资源由 On/Off CoD 资源和 Enterprise Pool CoD (EPCoD) 资源组成。这两种资源都可将补充资源动态传递至您的环境，这些补充资源通过普通 DLPAR 管理使用（将资源分配或发布给 LPAR）。

其他处理器和内存尽管实际存在，但是只有在 PowerHA SystemMirror 认为需要的附加容量抵得上成本时，才会使用。您可以使用 ROHA 功能，快速并轻松地获取额外资源，以满足环境中的峰值或非预期工作负载。

PowerHA SystemMirror 与 DLPAR、On/Off CoD 和 EPCoD 功能集成。这些功能的集合称为 ROHA。活动节点由框架上具有足够永久资源的 LPAR 托管。备用节点由框架上具有最少永久资源的 LPAR 托管，且依赖于 ROHA 动态添加额外的资源。

需要在备用节点上运行应用程序时，PowerHA SystemMirror 会使用 ROHA 功能。ROHA 功能验证节点是否具有成功运行应用程序所需的足够资源，并分配必需资源。可以从以下源中分配资源：

On/Off CoD 供应的资源

如果 CEC 空闲池具有的资源（这些资源可通过 DLPAR 分配给节点）不足，那么 On/Off CoD 功能会为 CEC 提供更多资源。这些附加资源将添加到 CEC 空闲池，并可通过 LPAR 操作使用。如果应用程序需要更多内存或处理器，那么 PowerHA SystemMirror 可通过 ROHA 功能自动将这些资源分配给备用节点。

EPCoD 供应的资源

如果 CEC 空闲池具有的资源（这些资源可通过 DLPAR 分配给节点）不足，那么 EPCoD 功能会为 CEC 提供更多资源。这些附加资源将添加到 CEC 空闲池，并可通过 DLPAR 操作使用。如果应用程序需要更多内存或处理器，那么 PowerHA SystemMirror 可通过 ROHA 功能自动将这些资源分配给备用节点。

CEC 空闲池

DLPAR 功能通过分配空闲池中可用的资源来为备用节点提供资源。这些资源从 On/Off CoD 池或 EPCoD 池提供。

将 PowerHA SystemMirror 配置为通过 ROHA 功能使用资源后，在应用程序需要资源之前，集群中的 LPAR 节点将不使用更多资源。

在空闲池耗尽之前，PowerHA SystemMirror 不会激活任何 CoD 资源。在空闲池耗尽之后，PowerHA SystemMirror 将激活更多硬件资源。CoD 硬件资源将激活，并且动态分配给 LPAR，直至满足应用程序的需求。应用程序不再需要分配的硬件资源时，会将其释放给空闲池，此时 PowerHA SystemMirror 会取消激活硬件资源。这些硬件资源会返回给它们源自的池（On/Off CoD 池或 EPCoD 池）。

相关参考：

第 18 页的『在 PowerHA SystemMirror 中监视』

PowerHA SystemMirror 的主要任务是识别故障并作出响应。PowerHA SystemMirror 使用 Cluster Aware AIX 基础结构监视其网络接口、设备和 IP 标签的活动。

相关信息：

管理 PowerHA SystemMirror

应用程序控制器

要让 PowerHA SystemMirror 控制应用程序，可创建一个应用程序控制器资源，该资源将用户定义的名称与专门编写的用于启动和停止应用程序的脚本的名称相关联。

通过定义应用程序控制器，PowerHA SystemMirror 可在发生故障转移时启动接管节点上的另一个应用程序实例。这可保护您的应用程序，使其不会变为单点故障。

定义应用程序控制器后，您可以将其添加到资源组。资源组是您定义的一组资源，这样 PowerHA SystemMirror 软件可将这些资源作为单个单元处理。

相关参考：

第 51 页的『规划资源组』

以下主题描述如何规划 PowerHA SystemMirror 集群内的资源组。

与 PowerHA SystemMirror 集成的应用程序

某些应用程序（包括 Workload Manager）无需应用程序控制器或其他脚本，即可直接配置为高可用资源。此外，PowerHA SystemMirror 集群验证可在某些方面确保 Workload Manager 配置的正确性和一致性。

PowerHA SystemMirror 提供了以下 PowerHA SystemMirror Smart Assist 应用程序，可帮助您将这些应用程序集成到 PowerHA SystemMirror 集群中：

Smart Assist for WebSphere

将现有 PowerHA SystemMirror 配置扩展为包括针对各种 WebSphere 组件的监视和恢复支持。

Smart Assist for DB2®

将现有 PowerHA SystemMirror 配置扩展为包括针对 DB2 Universal Database™ (UDB) 企业服务器版的监视和恢复支持。

Smart Assist for Oracle

为在 IBM AIX 操作系统上安装 Oracle Application Server 10g (9.0.4) (AS10g) 冷故障转移集群 (Cold Failover Cluster, CFC) 解决方案所涉及的那些项提供帮助。

Smart Assist for FileNet® P8

提供企业级可伸缩性和灵活性，可处理最富挑战性的苛刻内容、最复杂的业务流程，以及与所用环境中现有系统的集成。

Smart Assist for SAP MaxDB

设置 MaxDB 和 liveCache 数据库实例以实现高可用性。

Smart Assist for Lotus® Domino® Server

自动为已配置了 Lotus Domino 的环境配置 PowerHA SystemMirror。

Smart Assist for Tivoli® Storage Manager

利用 Tivoli Storage Manager 的三个不同部分：服务器、客户机和管理中心，为所用环境提供高可用解决方案。

Smart Assist for SAP

通过保护 SAP Netweaver 2004s 的单点故障，设置 SAP Netweaver 2004s 以实现高可用性。

Tivoli Directory Server 的迅捷帮助

在已安装了 Tivoli Directory Server 的情况下，自动配置 PowerHA SystemMirror。

SAP liveCache Hot Standby 的智能帮助

提供帮助您为环境中的工作负载堆栈配置 PowerHA SystemMirror 策略以及部署启动方法、停止方法和监视器方法的管理界面。

Websphere MQSeries® 的迅捷帮助

使程序能够与由不同组件（例如，处理器、子系统、操作系统和通信协议）组成的网络中的各项相互通信。

相关信息：

PowerHA SystemMirror 的 Smart Assist 应用程序

应用程序监视

PowerHA SystemMirror 可以监视向应用程序控制器定义的应用程序。

PowerHA SystemMirror 用以下两种方式之一监视应用程序：

- **进程监视**使用 Reliable Scalable Cluster Technology (RSCT) 和资源监视和控制 (Resource Monitoring and Control, RMC) 功能检测进程的终止。
- **定制监视**使用您定义的监视器方法监视应用程序的运行状况。

可配置多个应用程序监视器并将它们与一个或多个应用程序控制器相关联。您可以在 SMIT 中为每个监视器分配一个唯一名称。由于对每个应用程序支持多个监视器，PowerHA SystemMirror 可支持更复杂的配置。例如，可为使用中的每个 Oracle 并行服务器实例配置一个监视器。否则，您可以配置一个定制监视器以检查数据库的运行状况，以及一个进程停止监视器以即时检测到数据库进程的结束。

您可以使用应用程序可用性分析工具来度量任何由 PowerHA SystemMirror 定义的应用程序的确切可用时长。PowerHA SystemMirror 软件将收集和记录以下信息并对其添加时间戳记：

- 应用程序监视器已定义、更改或删除
- 应用程序启动、停止或发生故障
- 节点发生故障或关闭，或者启动
- 资源组脱机或移动
- 通过多个监视器的应用程序监视已暂挂或恢复。

相关信息：

配置 PowerHA SystemMirror 集群拓扑和资源（扩展）

监视 PowerHA SystemMirror 集群

多层应用程序的规划注意事项

使用多层应用程序的业务配置可以使用父子依赖性资源组。例如，数据库必须在应用程序控制器之前联机。在此情况下，如果数据库关闭并且移动到其他节点，那么必须先将包含应用程序控制器的资源组脱机，再将其在集群的任何节点上联机。

环境（例如，服务访问点 (SAP)）要求无论数据库何时发生故障，应用程序都能够循环（停止，然后再次启动）。很多应用程序服务由环境（例如，SAP）提供，且各个应用程序组件通常需要以特定的顺序进行控制。

当需要系统服务来支持应用程序环境时，在资源组之间建立相互依赖性也很有用。用于修剪日志文件或启动备份的**定时**作业之类的服务需要与应用程序一起从一个节点移到另一个节点，但通常在建立应用程序之前不启动。这些服务可以构建到应用程序控制器启动和停止脚本中，或者可通过前置事件和后置事件处理对其进行控制。但是，依赖资源组可简化将系统服务配置为依赖其所服务的应用程序的方法。

注：为最大程度减小在应用程序停止和重新启动进程期间丢失数据的可能性，请定制应用程序控制器脚本，以确保在应用程序停止进程期间所有未落实的数据都将临时存储到共享磁盘，并将在应用程序重新启动进程期间读回到应用程序。使用共享磁盘是很重要的，原因是应用程序重新启动时所在的节点与其停止时所在的节点可能不同。

您还可以配置具有位置依赖性的资源组，以便在启动、故障转移和回退时使特定资源组在同一节点或在不同节点上保持联机。

相关参考：

第 51 页的『规划资源组』

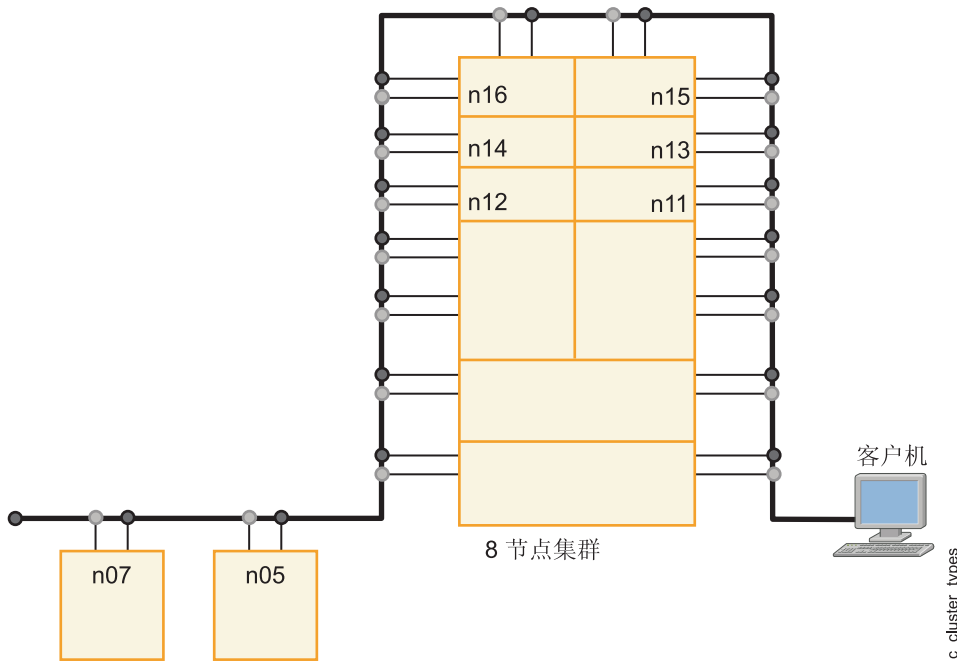
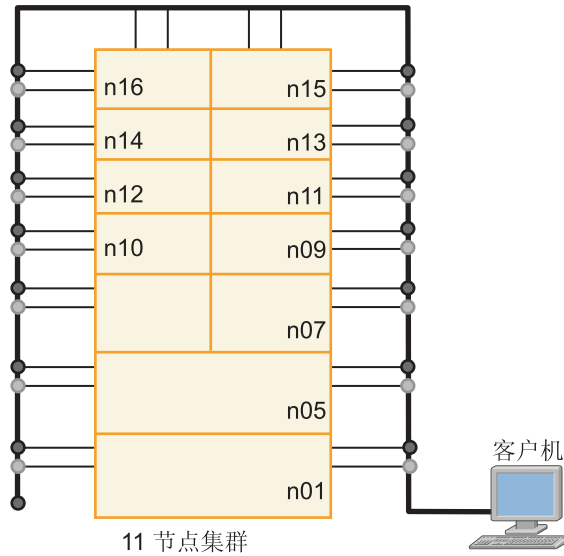
以下主题描述如何规划 PowerHA SystemMirror 集群内的资源组。

绘制集群图

集群图将每一规划步骤中的信息组合到显示集群功能和结构的一个图中。

下图显示了包括机架安装系统和独立系统的混合集群。该图使用长方形框代表节点支持的插槽。如果您的集群使用窄节点，那么请将节点的边框颜色加深并在抽屉中包括两个节点。对于宽节点，请使用整个抽屉。对于高节点，请使用与两个宽节点相当的量。请记住，每个窄节点均包含一个集成以太网连接。

绘制此图时，请从确定集群名称和将具有高可用性的应用程序开始。然后，对将组成集群的节点的边框颜色加深。包括每个节点的名称。



主机名要求

从 PowerHA SystemMirror 7.1.1 开始，由于存在新 Cluster Aware AIX (CAA) 层要求，针对哪个接口可为主机名具有特定要求。

请在选择主机名时记住以下要求：

- 在 `/etc/hosts` 文件中，主机名不能为别名。
- 主机名的名称解析必须通过两种方式工作。因此，仅可使用一组有限的字符。
- 在服务器上，属于主机名的 IP 地址必须可达，甚至在 PowerHA 处于关闭状态时也是如此。
- 主机名不能为服务地址。
- 主机名不能为网络上存在的地址，此地址在 PowerHA 中定义为专用。
- 主机名、CAA 节点名和 **COMMUNICATION_PATH**（即，节点的通信路径）必须相同。
- 缺省情况下，PowerHA 节点名、CAA 节点名和 **COMMUNICATION_PATH**（即，节点的通信路径）设置为相同。
- 节点名和 PowerHA 节点名可不同。
- 完成集群配置后，无法更改主机名。

注： 这些要求保持基本地址和持久性地址作为主机名的候选者。仅当在配置集群拓扑之前手动设置持久性别名时，可将持久性地址用作主机名。

规划集群网络连接

本节描述为 PowerHA SystemMirror 集群规划网络支持。

先决条件

在“初始集群规划”主题中，您已开始规划集群，并确定节点数量和要使其高可用的关键应用程序。您已开始绘制集群图。该图是您将在本节中进行规划的起点。

此外，此时您应已决定是否将使用 IP 地址接管 (IPAT) 来维护特定服务 IP 地址。

概述

您的主要目标是使用冗余来设计一种集群拓扑，消除具有作为单点故障潜在可能性的网络组件。

下表列示了这些网络组件和解决方案：

集群网络对象	消除成为单点故障的方式...
网络	使用多个网络来连接节点
网络接口卡 (NIC)	在每个网络上使用冗余 NIC

在本节中，您将完成以下规划任务：

- 设计集群网络拓扑，即连接集群节点的 IP 网络以及每个节点所具有的与每个网络的连接数。
- 向集群图添加网络。

相关参考：

第 5 页的『初始集群规划』

本部分描述为规划 PowerHA SystemMirror 集群以使应用程序具有高可用性所应采取的初始步骤。

第 18 页的『在 PowerHA SystemMirror 中监视』

PowerHA SystemMirror 的主要任务是识别故障并作出响应。PowerHA SystemMirror 使用 Cluster Aware AIX 基础结构监视其网络接口、设备和 IP 标签的活动。

PowerHA SystemMirror 的一般网络注意事项

PowerHA SystemMirror 允许与以太网网络进行节点间通信。

IP 别名

IP 别名是除网络接口控制器 (NIC) 中通常配置的 IP 标签或地址之外，配置到 NIC 上的 IP 标签或地址。使用 IP 别名是 PowerHA SystemMirror 支持的一种 AIX 功能。AIX 在一个 NIC 上支持多个 IP 别名。NIC 上的每个 IP 别名都可以位于单独的子网上。AIX 还允许为接口配置具有不同子网掩码的 IP 别名。PowerHA SystemMirror 尚未支持此功能。

IP 别名在 PowerHA SystemMirror 中用作 IP 地址接管的服务地址。

网络连接

PowerHA SystemMirror 要求集群中的每个节点与其他每个节点具有至少一个直接非路由网络连接。该软件使用这些网络连接在集群节点中传递脉动信号消息，从而确定所有集群节点、状态和网络接口的状态。

PowerHA SystemMirror 要求给定集群网络的所有通信接口在同一物理网络上定义，并且相互路由数据包。缺省情况下，PowerHA SystemMirror 针对脉动信号使用单点广播通信。如果选择改为使用多点广播脉动信号，那么需要确保网络支持多点广播。通信接口还必须能够接收来自对方的响应而不受任何网络设备的干扰。PowerHA SystemMirror 还要求站点中的所有节点必须至少与相同站点中每隔一个节点具有一个直接网络连接。

始终在链接的集群中站点之间使用单点广播通信。在站点中，可选择单点广播（缺省值）或多点广播通信。

在集群节点之间，请仅放置可以透明方式向所有集群节点传递多点广播数据包和其他数据包的智能型交换机、路由器或其他网络设备。此要求包括优化协议的设备。

如果将此类设备放置在集群节点和客户机之间的路径中，那么您可能需要在 **clinfo.rc** 文件中配置 ping 客户机列表，以帮助向客户机通知 IP 地址移动。特定网络拓扑可能需要其他解决方案，以确保客户机可在 IP 地址接管之后继续访问服务器。

对于不会修改数据包流的网桥、集线器和其他被动设备，可将其安全地放置在集群节点之间，以及节点和客户机之间。

相关信息：

为 Clinfo API 客户机应用程序编写程序

IP 标签

在不使用 PowerHA SystemMirror 的环境中，主机名通常标识一个系统，主机名还是系统中某个网络接口的 IP 标签。因此，可以通过使用系统的主机名作为用于连接的 IP 标签来访问该系统。

主机名解析

在 PowerHA SystemMirror 中，所有节点主机名都必须使用 **/etc/hosts** 文件在本地解析。当向集群定义节点时，您必须指定在本地解析为主机名的 IP 地址或标签，并且在您同步初始集群配置后，节点的主机名不能更改。

TCP/IP 网络中的 IP 标签

对于 TCP/IP 网络，IP 标签及其相关联的 IP 地址必须出现在 **/etc/hosts** 文件中。

服务 IP 标签或地址的名称在集群内必须唯一，并且不同于卷组和资源组名称。它应该与它服务的应用程序以及任何相应的设备（例如，**websphere_service_address**）相关。

当您向接口分配服务 IP 标签时，请使用有助于标识该接口在集群中的角色的命名约定。**/etc/hosts** 文件中的相关条目将与如下类似：


```
100.100.50.1 net1_en0
100.100.60.1 net2_en1
```

您可遵循相关 AIX 文档中的指示信息来配置网络接口控制器 (NIC)。当配置 NIC 时, AIX 会向其分配一个接口名称。接口名称指示 NIC 的类型, 由 2 或 3 个字符组成, 后跟 AIX 为特定类型的每个适配器依序分配的一个数字。例如, AIX 为它配置的第一个以太网 NIC 分配例如 **en0** 的一个接口名称, 为第二分配 **en1**, 以此类推。

相关信息:

配置集群事件

集群分区

分区也称为节点隔离, 发生在网络或网络接口控制器 (NIC) 故障将集群节点相互隔离的情况下。

当某个 PowerHA SystemMirror 节点停止接收来自另一节点的网络流量时, 将认为对方节点已发生故障。根据 PowerHA SystemMirror 配置, 该节点可能开始从发生故障的节点获取磁盘, 并使应用程序和 IP 标签可用。如果发生故障的节点实际仍处于启动状态, 那么从该节点中取下磁盘时, 可能发生数据损坏。如果网络再次变为可用, 那么 PowerHA SystemMirror 将停止其中一个节点, 以避免进一步磁盘争用和网络中出现重复 IP 地址。

PowerHA SystemMirror 脉动信号机制依赖于 IP 子系统和网络基础结构。因此, 如果网络阻塞或某个节点阻塞, 那么 IP 子系统可以静默方式放弃脉动信号。将尝试调整监视特性, 以考虑网络阻塞并防止集群分区。

相关参考:

第 25 页的『监视集群』

Cluster Aware AIX 基础结构监视所有可用和受支持的网络和存储器接口。集群节点上的集群管理器还通过这些接口之间的连接相互发送消息。

示例: 常规网络连接

一个正确的 PowerHA SystemMirror 网络连接示例, 包括两个单独的以太网网络, 其中每个网络在每个节点上具有两个网络接口。

由两个路由器将网络连接在一起, 它们在集群和客户机之间传送包, 但不在两个网络之间传送。集群中每个节点上均安装了 **clinfo.rc** 文件, 其中包含一些客户机系统的 IP 地址。

交换网络中的 PowerHA SystemMirror 配置

如果未正确定义或配置网络和交换机, 那么在使用交换网络的 PowerHA SystemMirror 配置中可能发生意外的网络接口故障事件。

当配置交换网络时, 请遵循以下准则:

- **虚拟局域网 (VLAN)**。如果使用 VLAN, 那么必须在同一 VLAN 上配置特定网络上的所有接口 (每个 VLAN 一个网络)。如果正将多点广播用于脉动信号, 那么所有接口和网络设备都必须能够在节点之间传送多点广播信息包。
- **自动协商设置**。某些以太网网络接口卡 (NIC) 能够自动协商它们的速度和其他特性, 例如, 半双工或全双工。请将 NIC 配置为不使用**自动协商**, 但以期望速度和双工值运行。请与 NIC 连接的交换机端口设置为相同的固定速度和双工值。

PowerHA SystemMirror 和虚拟以太网

PowerHA SystemMirror 支持由 Virtual I/O Server (VIOS) 或集成虚拟以太网 (IVE) 工具提供并且已安装适用 APAR 的虚拟以太网。PowerHA SystemMirror 支持对于 VIOS 和 IVE 是相同的。

PowerHA SystemMirror 中的 PCI 热插拔实用程序不适用于虚拟以太网的接口。此实用程序仅处理物理接口卡。由于虚拟以太网使用虚拟 I/O 适配器，因此无法使用该实用程序。

以下列表包含具有虚拟以太网的 PowerHA SystemMirror 的其他注意事项：

- 如果 VIOS 在同一网络中定义了多个物理接口，或者在同一框架中有两个或两个以上使用 VIOS 的 PowerHA SystemMirror 节点，那么出现单个物理接口故障时将不通知 PowerHA SystemMirror，因此它也不会对这些故障做出任何反应。这不会限制整个集群的可用性，因为 VIOS 会围绕着该故障安排流量。在这点上，VIOS 支持与以太网通道类似。使用不基于 VIOS 的方法提供单个物理接口故障的通知。
- 如果 VIOS 在网络中只有一个物理接口，那么 PowerHA SystemMirror 将检测到一个该物理接口的故障。但是，该故障会将节点与网络隔离。

注：在 VIOS 2.2.0.11 或更高版本中，通过借助每个 VIOS 客户机上的虚拟以太网适配器建立虚拟局域网，可以在逻辑分区之间使用存储区域网络 (SAN) 通信。可以通过 VIOS 为 NPIV 和 vSCSI 环境设置 SAN 通信。

- 在 VIOS 环境中，可能无法可靠地检测到虚拟化网络以外的物理网络适配器和网络组件的故障。要检测外部网络故障，必须使用虚拟化网络以外的一个或多个地址来配置 netmon.cf 文件。

对虚拟以太网连接进行故障诊断

要对 PowerHA SystemMirror 的已定义虚拟以太网接口进行故障诊断并检测接口故障，请将这些接口视为单个适配器网络上定义的接口。

注：针对以太网，PowerHA 支持对相同网络名称使用任何虚拟和物理适配器组合。

尤其是，在 `/usr/es/sbin/cluster/etc/clinfo.rc` 脚本的 `PING_CLIENT_LIST` 变量中或在 `etc/cluster/ping_client_list` 中列出属于 VLAN 的网络接口，然后运行 `clinfo` 程序。这样，只要发生集群事件，`clinfo` 程序就监控并检测到列出的网络接口的故障。由于 VLAN 的性质，用于检测网络接口故障的其他机制无效。

相关概念：

第 27 页的『PowerHA SystemMirror 中的虚拟网络』

在 PowerHA SystemMirror V7.1.0 或更高版本中，Cluster Aware AIX (CAA) 提供的适配器监视器无法始终确定虚拟适配器是否丢失其相应物理适配器。

在 PowerHA SystemMirror 中监视

PowerHA SystemMirror 的主要任务是识别故障并作出响应。PowerHA SystemMirror 使用 Cluster Aware AIX 基础结构监视其网络接口、设备和 IP 标签的活动。

必须存在监视连接，因为它们使 PowerHA SystemMirror 能够识别出网络故障与节点故障之间的区别。例如，如果与 PowerHA SystemMirror 网络（在资源组中使用此网络的 IP 标签）失去连接，并且您具有另一个基于 TCP/IP 的网络，那么 PowerHA SystemMirror 将识别出其集群网络发生故障，并且采取可避免集群分区的恢复操作。

要避免集群分区，您应在 PowerHA SystemMirror 集群中配置冗余网络。

PowerHA SystemMirror 会自动监视以下组件上的接口：

- TCP/IP 网络
- 存储区域网络
- 存储库磁盘

设计网络拓扑

将集群节点和客户机相链接的 IP 网络和非 IP（点到点）网络的组合称为集群网络拓扑。PowerHA SystemMirror 软件在每个节点上支持大量 IP 和点到点设备，从而为设计网络配置提供灵活性。

当设计网络拓扑时，请确保客户机对于其应用程序具有高可用网络访问权。这要求以下各网络接口均不是单点故障：

- IP 子系统
- 单个网络
- 单个 NIC

消除网络成为单点故障

在单网络设置中，集群中的每个节点都仅连接一个网络且对于客户机仅具有一个可用的服务接口。在此设置中，网络是整个集群的单点故障，并且每个服务接口都是单点故障。

下图显示了单网络配置。

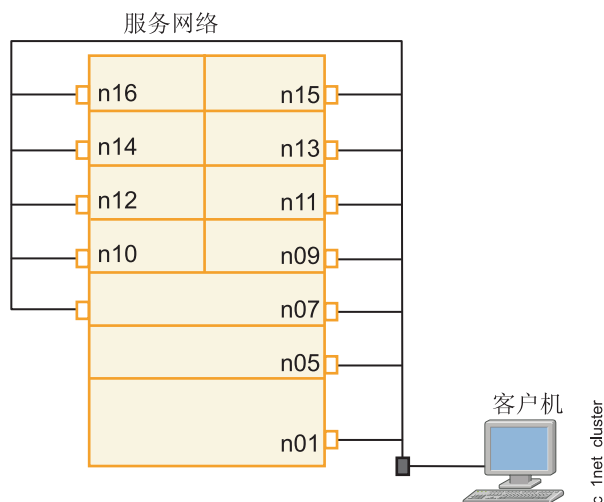


图 1. 单网络和单 NIC 设置

要消除网络成为单点故障，请配置多个网络，以便 PowerHA SystemMirror 在所有集群节点中具有多条路径。请记住，如果某个客户机仅连接一个网络，那么网络就是该客户机的单点故障。在多网络设置中，如果一个网络发生故障，那么其余网络仍然可以工作以便连接节点并为客户机提供访问权。

您可配置为用于在所有集群节点中传送脉动信号和其他信息的网络越多，系统的可用性就越高。

下图显示了双网络设置，其中每个集群节点有多条路径。

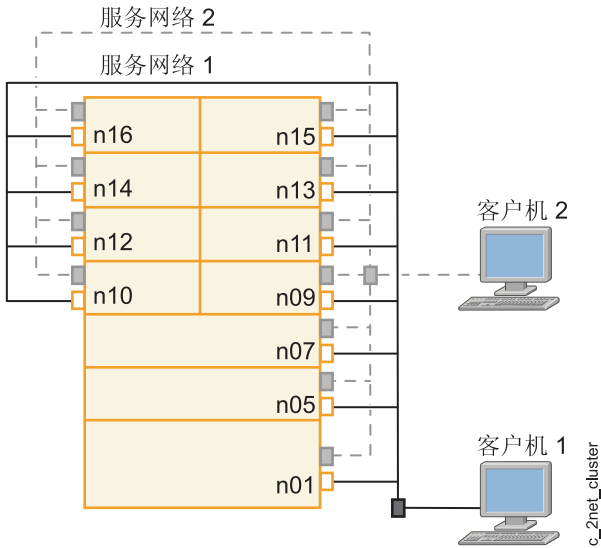


图 2. 双网络设置

注：对于用于为一个 PowerHA SystemMirror IP 网络配置两个接口的双端口以太网适配器，当前不支持对其进行热更换。

消除网络接口卡成为单点故障

网络接口卡 (NIC) 以物理方式将节点连接到网络。

当每个网络配置为带有单个 NIC 时，NIC 将变为潜在的单点故障。要为此问题采取补救措施，请将节点配置为与其连接的至少带有两个 NIC 的网络。在下图中，每个集群节点与每个网络具有两个连接。

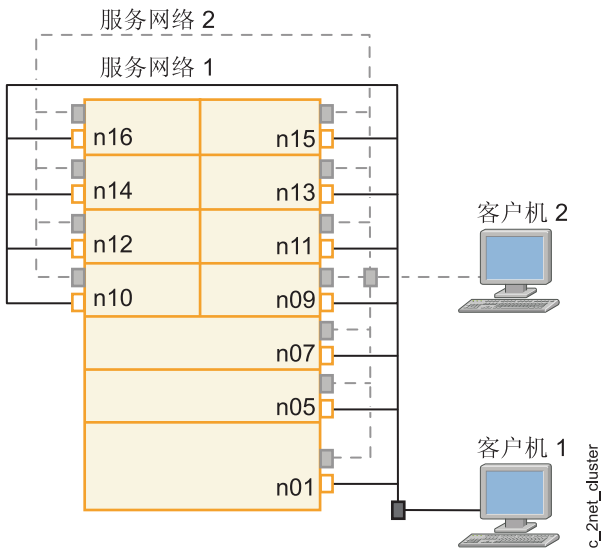


图 3. 双网络和双 NIC 配置

注：对于用于为一个 PowerHA SystemMirror IP 网络配置两个接口的双端口以太网适配器，当前不支持对其进行热更换。

相关信息：

集群事件期间的资源组行为

网络接口功能:

在将节点配置为具有与单个网络的多个连接后，网络接口将在 PowerHA SystemMirror 中具有不同功能。

服务接口

服务接口是使用 PowerHA SystemMirror 服务 IP 标签配置的网络接口。客户机使用服务 IP 标签来访问应用程序。仅当响应的资源组联机时，服务 IP 才可用。

持久节点 IP 标签

持久节点 IP 标签是可分配至集群网络的某个特定节点的 IP 别名。持久节点 IP 标签始终位于同一节点上（节点绑定），并且同时存在于始终定义了服务或引导 IP 标签的 NIC 上。持久节点 IP 标签不需要在该节点上安装附加的物理 NIC，并且不属于任何资源组。

通过分配持久节点 IP 标签，提供了您可用于管理目的的节点绑定地址，因为与持久节点 IP 标签的连接始终通往集群中的某个特定节点。每个节点每个网络可以具有一个持久节点 IP 标签。

PowerHA SystemMirror 的一种最佳作法是，您必须为每个集群节点配置一个持久 IP 标签。这在例如以下的情况中是很有用的：为运行报告或进行诊断的目的，您必须访问 PowerHA SystemMirror 集群中的某个特定节点。配置持久 IP 标签的优势是，即使单个 NIC 发生故障，PowerHA SystemMirror 仍然可以访问节点上的持久 IP 标签（假设网络上有多余 NIC）。

在指定的网络节点上配置了持久节点 IP 标签后，即使 PowerHA SystemMirror 在该节点上关闭，该标签仍会在引导时变为可用并保持已配置状态。

您可以在以太网上创建持久节点 IP 标签。

以下列表描述了在配置持久节点 IP 标签后，PowerHA SystemMirror 对于故障的响应：

- 如果配置了服务 IP 标签的 NIC 发生故障，并且在此 NIC 上还定义了一个持久标签，那么该持久标签将故障转移至服务 IP 标签会故障转移至的同一引导接口。
- 如果指定节点的集群网络上的所有 NIC 均发生故障，那么持久节点 IP 标签将变为不可用。持久节点 IP 标签始终保留在同一节点和同一网络上。它不会在集群的节点之间移动。

相关信息:

配置 PowerHA SystemMirror 集群拓扑和资源（扩展）

通过 IP 别名的 IP 地址接管:

PowerHA SystemMirror 使用通过 IP 别名的 IPAT 来使服务 IP 地址具有高可用性。

当 PowerHA SystemMirror 在节点上启动时，服务 IP 标签将作为别名添加到向 PowerHA SystemMirror 定义的其中一个引导接口上。如果该接口发生故障，并且同一网络中有某个可用接口，那么 IP 标签将作为别名添加到另一接口上。要使用通过 IP 别名的 IPAT，网络必须支持免费 ARP。

持久节点 IP 标签的子网注意事项

当您在集群网络上配置持久节点 IP 标签时，与持久 IP 标签相关联的 IP 地址必须位于不同于可能与其共享接口的任何服务地址的子网上。对于每个节点一个接口的网络不需要单独的子网。

在某些情况下，您可能需要将同一子网上的某个持久 IP 标签配置为服务 IP 标签。在此情况下，为避免从任一地址发送的网络包发生问题，请考虑为服务 IP 别名配置分配首选项。此首选项允许您配置适合 VPN 防火墙外部连接需求的分配首选项类型。

注：如果您计划配置网络文件系统 (NFS)，那么有不同的子网注意事项。

相关参考：

第 48 页的『NFS 交叉安装和 IP 标签』

为启用 NFS 交叉安装，每个集群节点都可以作为 NFS 客户机。其中每个节点都必须具有到 NFS 服务器节点的服务 IP 标签的有效路径。即，要启用 NFS 交叉安装，客户机节点上必须存在 IP 标签，并且必须在与 NFS 服务器节点的服务 IP 标签相同的子网上配置此 IP 标签。

第 24 页的『服务 IP 标签别名的分配类型』

您可以在 SMIT 中为服务 IP 标签别名安排指定不同的分配首选项。

『规划通过 IP 别名的 IP 地址接管』

通过向 NIC 分配 IP 别名，您可以在同一网络接口上创建多个 IP 标签。

规划通过 IP 别名的 IP 地址接管

通过向 NIC 分配 IP 别名，您可以在同一网络接口上创建多个 IP 标签。

在通过 IP 别名的 IP 地址接管期间，当某个 IP 标签从一个 NIC 移动到另一个 NIC 时，目标 NIC 将接收新 IP 标签作为 IP 别名，并保留原始 IP 标签和硬件地址。

为通过 IP 别名的 IPAT 配置网络可简化 PowerHA SystemMirror 中的网络配置。您可以为 NIC 配置服务地址和一个或多个引导地址。

为通过 IP 别名的 IPAT 分配 IP 标签

PowerHA SystemMirror 使用一项称为“通过 IP 别名的 IP 地址接管”技术，使 IP 地址保持高可用性。

当规划通过 IP 别名的 IP 地址接管时，请复审以下信息：

- 每个网络接口必须具有一个定义至 PowerHA SystemMirror 的引导 IP 标签。定义至 PowerHA SystemMirror 的接口将用于保持服务 IP 地址的高可用性。
- 对于通过 IP 别名判别使用 IP 地址接管的网络，无法为其配置硬件地址接管 (Hardware Address Takeover, HWAT)。
- 当节点上有多个接口连接到同一网络时，以下子网要求适用：
 - 必须在不同子网上定义所有引导地址。
 - 服务地址所在的子网必须不同于所有引导地址和持久地址所在的子网。

注：这些子网要求可避免 AIX 操作系统的 IP 路径条带分割功能，该功能允许多条路径连至同一子网，并且可能导致向发生故障的接口发送应用程序流量。这些要求不适用于使用以太网聚集或 EtherChannel 合并为单个逻辑接口的多个接口。

- 可以在所有非并发资源组中包括为“通过 IP 别名的 IP 地址接管”配置的服务地址标签。
- 多个服务标签可以作为给定接口上的别名共存。
- PowerHA SystemMirror 网络中所有 IP 标签的网络掩码必须相同。
- 如果有多个服务标签和持久标签，那么 PowerHA SystemMirror 会尝试将它们均匀分配在所有可用网络接口中。您可以指定一个位置首选项，以便持久别名和服务别名始终映射至同一接口。有关服务标签的更多信息，请参阅服务 IP 标签别名的分发首选项。

节点上的引导地址（在系统重新引导后且 PowerHA SystemMirror 软件启动前，由 AIX 操作系统指定的基地地址）将作为引导地址定义至 PowerHA SystemMirror。当您首次配置集群时，PowerHA SystemMirror 将自动发

现并配置引导地址。当 PowerHA SystemMirror 软件在节点上启动时，节点的服务 IP 标签将作为别名添加至定义为引导地址的某个 NIC 上。如果托管服务 IP 的 NIC 发生故障，那么 PowerHA SystemMirror 会尝试将其移动至同一节点上的其他活动引导 NIC。

在节点故障转移事件期间，已移动的服务 IP 标签将作为别名，与目标节点的 NIC 上可能已配置的任何其他服务标签一起放置在该 NIC 上。

例如，如果节点 A 发生故障，那么 PowerHA SystemMirror 会尝试将所有资源组移动至节点 B。如果资源组包含服务 IP 地址，那么 PowerHA SystemMirror 会将服务 IP 作为别名放置在节点 B 的适当 NIC 上，并且任何其他现有标签将原样保留在节点 B 的 NIC 上。这样，现在节点 B 上的一个 NIC 可以接收先前定向至节点 A 上的服务地址的客户机流量。稍后，在节点 A 重新启动后，它将在引导地址上启动，并且能够在节点 B 发生故障的情况下接管服务地址。当节点 B 释放所请求的服务 IP 标签时，将在节点 B 上删除服务 IP 标签的别名。节点 A 会再次将服务 IP 标签作为别名放置在适当网络的其某个引导地址接口上。

当使用“通过 IP 别名的 IPAT”时，将通过使用所有可用且适当的接口来获取服务 IP 标签。如果多个接口可用于主管服务 IP 标签，那么将根据配置的分配策略来选择接口。缺省情况下，将使用反并置策略，并将尝试在所有可用接口中均匀分配服务 IP 标签。

在 PowerHA SystemMirror 中，如果您使用 **ifconfig** 命令除去接口的引导时地址，那么不会影响该接口上作为别名的任何服务 IP 标签。但是，如果您使用 **chdev** 命令（或 **chinet** 快速路径）替换了接口上的引导地址，那么此命令将删除任何服务 IP 别名，并且在 PowerHA SystemMirror 中不会提示发生此删除操作。如果您需要更改或临时取消分配正在主管服务地址的接口的引导地址，那么最好先将服务地址移动到其他接口上。如果没有其他可用的接口，那么请使用 **ifconfig** 命令别名来替换接口上的引导地址，而无需中断应用程序访问。

相关参考：

第 47 页的『PowerHA SystemMirror 中的 NFS 交叉安装』

NFS 交叉安装是特定于 PowerHA SystemMirror 的 NFS 配置，其中集群中的每个节点都可以同时作为 NFS 服务器和 NFS 客户机。在从一个节点导出文件系统期间，文件系统将以 NFS 方式安装在资源组的所有节点上，其中包括导出文件系统的节点。还可以从其他节点导出其他文件系统，并且会将其以 NFS 方式安装在所有节点上。

规划服务 IP 标签别名安排

如果您使用“通过 IP 别名的 IPAT”，那么您可为 PowerHA SystemMirror 中配置的服务 IP 标签的安排配置分配首选项。

PowerHA SystemMirror 允许您为服务 IP 标签别名指定分配首选项。这些是属于 PowerHA SystemMirror 资源组且属于“通过 IP 别名的 IPAT”网络的服务 IP 标签。

服务 IP 标签别名的分配首选项是用于控制在集群节点的物理网络接口卡上安排服务 IP 标签别名的网络范围的属性。

请使用 IP 别名的分配首选项来解决以下集群要求：

- 防火墙注意事项
- 使用 VLAN 的集群配置（当期望应用程序接收来自特定网络接口的包时）
- 在集群中安排 IP 标签的特定要求

服务 IP 标签别名的分配首选项

为服务 IP 标签别名配置分配首选项能够：

- 允许您为集群中的服务 IP 标签定制负载均衡，同时将先前在节点上分配的持久 IP 标签考虑在内。

- 使 PowerHA SystemMirror 根据指定的首选项重新分配别名服务 IP 标签。
- 允许您配置适合 VPN 防火墙外部连接要求的分配首选项类型。

虽然服务 IP 标签可能会移动到另一网络接口，但是 PowerHA SystemMirror 可确保根据指定的分配首选项继续分配标签。也就是说，分配首选项是在启动和后续集群事件（例如故障转移、回退或相同节点上的接口更改）期间维护的。例如，如果将标签指定为映射至同一接口，那么即使最初配置的服务 IP 标签移动到另一个节点，标签仍会映射至同一接口上。

只要可接受的网络接口可用，就会在集群中应用分配首选项。PowerHA SystemMirror 总是保持服务 IP 标签处于活动状态，即使首选项不能得到满足。

服务 IP 标签别名的分配类型

您可以在 SMIT 中为服务 IP 标签别名安排指定不同的分配首选项。

分配首选项类型遵循。

分配首选项的类型	描述
Anti-collocation	这是缺省值。PowerHA SystemMirror 通过使用“最少装入”选择过程，在所有引导 IP 标签中分配所有服务 IP 标签别名。
Collocation	PowerHA SystemMirror 在同一网络接口卡 (NIC) 上分配所有服务 IP 标签别名。
Anti-collocation with persistent labels	PowerHA SystemMirror 在不托管持久 IP 标签的所有活动物理接口上分配所有服务 IP 标签别名。仅当没有任何其他网络接口可用时，PowerHA SystemMirror 才会在托管持久标签的接口上安排服务 IP 标签别名。 注：如果您未配置持久 IP 标签，那么 PowerHA SystemMirror 允许您选择“Anti-collocation with persistent”分配首选项，但是它会发出警告并且在缺省情况下使用常规“Anti-collocation”首选项。
Collocation with persistent labels	在托管持久 IP 标签的同一 NIC 上分配所有服务 IP 标签别名。此选项在以下 VPN 防火墙配置中可能很有用：仅向一个接口授予外部连接，并且所有 IP 标签（持久和服务）必须都在同一接口卡上分配。 注：如果您未配置持久 IP 标签，那么可使用 PowerHA SystemMirror 选择“Collocation with persistent”分配首选项，但是它会发出警告并且在缺省情况下使用常规“Collocation”首选项。
Anti-collocation with source	使用“Anti-Collocation”首选项来映射服务标签。如果没有足够适配器，那么可在一个适配器上放置一个以上的服务标签。如果选择此选项，那么其中一个标签将选择为用于传出通信的源地址。在 Source IP Label for outgoing packets 字段中选择的接口标签是源地址。
Collocation with source	使用“Collocation”首选项来映射服务标签。如果选择此选项，那么其中一个服务标签将选择为用于传出通信的源地址。在 Source IP Label for outgoing packets 字段中选择的接口标签是源地址。
Anti-Collocation with persistent label and source	使用“Anti-Collocation with persistent”首选项来映射服务标签。当服务地址数量超过引导适配器的数量时，可以选择一个服务地址作为源地址。

以下规则适用于分配首选项:

- 如果没有可满足首选项的足够接口，那么 PowerHA SystemMirror 会向现有活动网络接口卡分配服务 IP 标签别名和持久 IP 标签。
- 如果您未配置持久标签，那么 PowerHA SystemMirror 允许您选择“Collocation with persistent”和“Anti-Collocation with persistent”分配首选项，但是会发出警告并且在缺省情况下使用常规“Collocation”或“Anti-collocation”首选项。
- 您可以动态更改 IP 标签分配首选项：新选择将在后续集群事件期间激活。在首选项更改时，PowerHA SystemMirror 通过当前重新定位服务 IP 标签从而不中断处理。

如果某个服务 IP 标签故障并且另一个服务 IP 标签在同一节点上可用，那么 PowerHA SystemMirror 将恢复服务 IP 标签别名，方法是通过将其移动到同一节点上的另一个 NIC。在此事件期间，您指定的分配首选项依然有效。

相关信息:

管理 PowerHA SystemMirror

规划特定于站点的服务 IP 标签

您可以具有一个服务 IP 标签，该标签可在多个节点上配置，并且与可在节点或站点间移动的资源组关联。

注： 站点仅在 PowerHA SystemMirror Enterprise Edition 和 Standard Edition 7.1.2 或更高版本中受支持。复制管理仅在 PowerHA SystemMirror Enterprise Edition 中受支持。

在一个站点上有效的 IP 地址由于子网问题可能在另一个站点上无效，您可以将可在多个节点上配置的服务 IP 标签与特定站点关联。特定于站点的服务 IP 标签在 PowerHA SystemMirror 中配置，并且可以在采用或不采用 PowerHA SystemMirror Enterprise Edition for AIX 网络的情况下使用。此标签与资源组关联，并且仅当该资源组在关联站点上处于 ONLINE PRIMARY 状态时才活动。

规划其他网络条件

最常见的网络规划注意事项包括 PowerHA SystemMirror 环境中的名称供给，以及设置集群监视和故障检测。

将 PowerHA SystemMirror 与 NIS 和 DNS 结合使用

用于在发生故障后对网络和接口问题进行故障诊断的某些命令需要进行 IP 查找，以确定与指定的 IP 标签相关联的 IP 地址。

如果 Network Information Service (NIS) 或 Domain Name Server (DNS) 正在运行，那么 IP 查找将缺省为查找用于名称和地址解析的名称服务器系统。但是，如果通过发生故障的接口访问名称服务器，那么请求不会完成，并最终超时。此超时会严重减慢 PowerHA SystemMirror 事件处理速度。

为确保集群事件成功快速地完成，PowerHA SystemMirror 通过在服务 IP 标签交换期间设置以下 AIX 环境变量，禁用 NIS 或 DNS 主机名解析。

```
NSORDER = local
```

因此，每个集群节点的 `/etc/hosts` 文件都必须包含 PowerHA SystemMirror 为所有集群节点定义的所有 IP 标签。

由进程而非 PowerHA SystemMirror 发送的 DNS 请求

禁用 NIS 或 DNS 主机名解析是 PowerHA SystemMirror 事件脚本环境所特有的。当 PowerHA SystemMirror 连接服务 IP 标签或当它在接口上交换 IP 标签时，它会将 NSORDER 变量设置为 `local`。

其他进程会继续使用缺省系统名称解析设置（例如，在 PowerHA SystemMirror 外部需要 DNS IP 地址解析的应用程序）。如果这些进程请求进行 IP 查找，那么在 PowerHA SystemMirror 的网络接口重新配置事件期间，进程可能仍无法联系外部名称服务器。向 DNS 发送的请求将在 PowerHA SystemMirror 完成网络接口重新配置事件之后成功。

监视集群

Cluster Aware AIX 基础结构监视所有可用和受支持的网络和存储器接口。集群节点上的集群管理器还通过这些接口之间的连接相互发送消息。

集群中所有可用和受支持的网络和存储器接口都用于监视接口、确保对等集群的连接，并在连接发生故障时报告。

有关受支持适配器、磁盘和多路径驱动程序当前列表，请联系您的 IBM 代表。PowerHA SystemMirror 支持监视以下类型的接口并通过这些接口通信：

- 以太网
- 4 GB 和 8 GB Emulex 光纤通道适配器
- 存储库磁盘（支持 SAN 和 SAS 磁盘）
- FCoE（基于以太网的光纤通道）

相关信息：

 [PowerHA Hardware Support Matrix](#)

规划 PowerHA SystemMirror 中的 VPN 防火墙网络配置

PowerHA SystemMirror 允许您为服务 IP 标签别名指定分配首选项。这些是属于 PowerHA SystemMirror 资源组且属于“通过 IP 别名的 IPAT”的服务 IP 标签。

特定 VPN 防火墙配置允许一次从外部只连接到一个 NIC。如果防火墙以此方式配置，那么会在同一接口分配所有 PowerHA SystemMirror 服务和持久 IP 标签。

要使 PowerHA SystemMirror 管理 IP 标签以满足这种 VPN 防火墙的需要，请执行以下操作：

- 为集群中的每个节点指定持久 IP 标签。持久 IP 标签会映射到所选网络上的可用接口。
- 为包含服务 IP 标签的网络指定持久分配首选项的并置。这确保了所有服务 IP 标签别名分配在主管持久 IP 标签的同一物理接口上。

相关参考：

第 24 页的『服务 IP 标签别名的分配类型』

您可以在 SMIT 中为服务 IP 标签别名安排指定不同的分配首选项。

相关信息：

[管理 PowerHA SystemMirror](#)

使用 PowerHA SystemMirror 规划 IP V6 地址

PowerHA SystemMirror 7.1.2 或更高版本中支持因特网协议 V6 (IPv6)。

在环境中实现 IPv6 之前，必须考虑 PowerHA SystemMirror 的以下区域：

- Cluster Aware AIX (CAA) 将为节点中配置的任何 IP 地址自动使用脉动信号。要阻止 CAA 在特定接口上以脉动信号传递任何 IPv6 地址，必须将您使用 PowerHA SystemMirror 的网络标识为专用网络。
- IPv6 使用适配器地址和其他网络属性的动态配置。缺省情况下，在系统重新引导操作前后 IPv6 地址不会持续保留，但是，您可以配置系统以在启动期间运行 **autoconf6** 命令。您需要规划在您的系统环境中运行 **autoconf6** 命令的方式和位置。
- PowerHA SystemMirror 使用链接本地地址作为引导地址。您可以配置另一个别名地址以用作 PowerHA SystemMirror 引导地址。链接本地地址用作引导地址极为方便，因为它们始终针对使用 IPv6 的任何接口配置。您必须规划可以在环境中使用的 IPv6 引导地址。

相关信息：

[通过因特网协议网络传递脉动信号](#)

[在系统重新引导前后持续保留 IPv6 地址](#)

规划与 Oracle 进行节点间通信的网络

Oracle 使用 **private** 网络属性设置来选择用于 Oracle 节点间通信的网络。PowerHA SystemMirror 不使用此属性，并且此属性将不会对 PowerHA SystemMirror 产生任何影响。缺省属性是 **public**。

将网络属性更改为 **private** 将使网络与 Oracle 兼容，途径是将所有接口更改为服务。

在创建集群网络（通过手动方式或通过使用发现功能）后，您可以通过遵循以下 SMIT 路径来更改网络属性：

Cluster Nodes and Networks > Manage Networks and Network Interfaces > Networks > Change/Show a Network

选择要更改的网络，然后将 Network Attribute 设置更改为 **private**。在进行此更改后，请同步集群。

专用网络的配置规则

请遵循以下步骤来配置 Oracle 使用的专用网络：

1. 配置网络并添加所有接口。如果网络没有任何接口，那么您无法更改属性。
2. 将网络属性更改为 **private**。
3. 验证专用网络是否具有所有引导接口或所有服务接口。如果网络具有所有引导接口（使用发现时的缺省情况），那么 PowerHA SystemMirror 会将这些接口转换为服务（Oracle 仅查看服务接口）。
4. 在更改属性后同步集群。

注：将网络属性定义为 **private** 后，便无法再将其更改回 **public**。您必须删除网络，然后向 PowerHA SystemMirror 重新定义此网络（缺省为 **public**）。

相关信息：

验证和同步 PowerHA SystemMirror 集群

PowerHA SystemMirror 中的虚拟网络

在 PowerHA SystemMirror V7.1.0 或更高版本中，Cluster Aware AIX (CAA) 提供的适配器监视器无法始终确定虚拟适配器是否丢失其相应物理适配器。

例如，如果从 Virtual I/O Server (VIOS) 拔出网络电缆，那么它无法与外部网络通信。因此，当 VIOS 分区无法达到虚拟网络外的任何外部 LAN 时，可能会在其单个虚拟接口可用时进行报告。您可以使用 APAR IV14422 解决此问题。此问题与 APAR IZ01331（针对 HACMP 6.1）相似。如果从 HACMP 6.1 或先前版本迁移，且应用了 APAR IZ01331，那么不需要更改现有 netmon.cf 文件。但是，必须在迁移后应用 APAR IV14422。

如果这是您首次在 PowerHA SystemMirror 中设置虚拟网络或集成虚拟因特网 (IVE) 网络，那么必须在 /usr/es/sbin/cluster 目录中创建 netmon.cf 文件。netmon.cf 文件中的内容必须遵循 HACMP 6.1 中使用以及 APAR IZ01331 中描述的格式。在 netmon.cf 文件中，针对每个虚拟界面，必须至少具有使用以下格式的一行：

```
!REQD owner target
```

以下列表描述 netmon.cf 文件中使用的变量。

!REQD

显式字符串，它必须位于行的开始（无任何前导空格）。

owner 接口，它的联机或脱机状态由它是否可对任何指定目标执行 ping 操作确定。可以将 owner 指定为主机名、IP 地址或接口名称。如果使用主机名，那么它必须解析为 IP 地址，否则会忽略此行。您可以指定 !ALL 字符串以指示所有适配器使用指定目标。

target 您希望 owner 尝试进行 ping 操作的 IP 地址或主机名。要使用主机名，目标必须可解析为 IP 地址。

创建或更改 netmon.cf 文件时，请考虑以下信息：

- 必须选择位于虚拟网络环境以外的目标。
- 必须通过网络环境中的更改维护您可确定的目标。

- 针对每行，仅可提供一个目标，但在文件中可通过多达 32 个不同行指定相同所有者条目。针对在多行上列出的拥有适配器，如果适配器可对任何提供的目标执行 ping 操作，那么此适配器视为可用。
- 请勿使用相同物理系统上的全部目标。此外，请勿使所有目标变为来自相同 PowerHA SystemMirror 集群的适配器。否则，此集群中的任何节点为唯一联机节点时，无法保持其适配器可用。
- 每个虚拟适配器在 netmon.cf 文件中必须至少具有一行，指定可在此接口上从引导 IP 地址或配置的持久 IP 别名执行 ping 操作的目标。
- 可以对其执行 ping 操作的网络硬件（例如，网关和路由器）可用作目标地址，因为 PowerHA SystemMirror 节点已使用这些网络硬件。

如果相同网络上的相同适配器是虚拟适配器，其他不是虚拟适配器，那么最好使用 **!REQD** 格式。针对 netmon.cf 文件中的虚拟适配器 en0，使用 **!REQD** 格式。针对物理适配器 en1，可以选择在 netmon.cf 文件中包含 **!REQD**，还可以选择使用 **!REQD** 格式。

在 PowerHA 7.1 (CAA) 中，仅使用 **!REQD** 条目。如果在 netmon.cf 文件中具有任何其他条目，那么这些条目将不再有用，将其忽略。但是，使用 **!REQD** 值的方式仍相同。必须可以至少针对一个目标执行 ping 操作（如果针对相同适配器具有多个行）。

注：此格式还适用于 IVE 网络。但是，无法使用属于相同物理系统中 IVE 网络的目标。

示例

以下示例说明 netmon.cf 文件中的内容：

1. 在此示例中，拥有 host1.ibm 的适配器仅在可对 100.12.7.9 或 host4.ibm 可解析为的任何地址执行 ping 操作时才可用。拥有 100.12.7.20 的适配器仅在可对 100.12.7.10 或 host5.ibm 可解析为的任何地址执行 ping 操作时可用。如果 100.12.7.20 为 host1.ibm 解析为的 IP 地址，那么四个目标全部属于此相同适配器。

```
!REQD host1.ibm 100.12.7.9
!REQD host1.ibm host4.ibm
!REQD 100.12.7.20 100.12.7.10
!REQD 100.12.7.20 host5.ibm
```

2. 在此示例中，所有适配器仅当其可对 100.12.7.9、110.12.7.9 或 111.100.1.10 IP 地址执行 ping 操作时才可用。en1 所有者条目具有一个额外目标 9.12.11.10。


```
!REQD !ALL 100.12.7.9
!REQD !ALL 110.12.7.9
!REQD !ALL 111.100.1.10
!REQD en1 9.12.11.10
```


相关参考：

第 17 页的『PowerHA SystemMirror 和虚拟以太网』

PowerHA SystemMirror 支持由 Virtual I/O Server (VIOS) 或集成虚拟以太网 (IVE) 工具提供并且已安装适用 APAR 的虚拟以太网。PowerHA SystemMirror 支持对于 VIOS 和 IVE 是相同的。

相关信息：

 [APAR IV14422: VIOS cable pull does not lead to events run by PowerHA](#)

 [APAR IZ01331: New netmon functionality to support PowerHA SystemMirror for AIX on VIO](#)

避免网络冲突

您可以避免 IP 地址上的网络冲突。验证操作将通知您重复的 IP 地址。请更正重复的地址，并重新同步集群。

向集群图添加网络拓扑

将网络规划为包括所有 TCP/IP 网络。以名称和属性标识每个网络。在代表插槽的每个节点的框中，记录接口标签。

现在，您可以向在“规划过程概述”中开始绘制的样本集群图添加网络。

相关参考:

第 3 页的『规划过程概述』

本主题描述 PowerHA SystemMirror 集群的规划步骤。

规划共享磁盘和磁带设备

本节讨论在 PowerHA SystemMirror 集群中配置共享外部磁盘之前要考虑的信息，并提供有关将磁带机作为集群资源进行规划和配置的信息。

先决条件

您已完成了规划集群网络连接和规划应用程序和应用程序控制器各节中的规划步骤。

请参阅 AIX 文档，以了解适合您磁盘和磁带设备的一般硬件和软件设置。

共享磁盘和磁带设备概述

在 PowerHA SystemMirror 集群中，共享磁盘是与多个用于应用程序共享存储器的集群节点连接的外部磁盘。

在非并发配置中，每次只有一个节点拥有磁盘。如果所有者节点发生故障，那么资源组节点列表中具有次最高优先级的集群节点将获取共享磁盘的所有权，并重新启动应用程序以将关键服务恢复至客户机。这可确保客户机应用程序可访问存储在磁盘上的数据。

接管通常在 30 - 300 秒内发生。此范围取决于所使用的磁盘的数量和类型、卷组的数量、文件系统（共享还是网络文件系统 (NFS) 交叉安装）以及集群配置中的关键应用程序的数量。

为集群规划共享外部磁盘时，目标是消除磁盘存储子系统中的单点故障。下表列示了磁盘存储子系统组件，以及消除其成为单点故障的建议方式。

集群对象	消除成为单点故障的方式...
磁盘适配器	使用冗余磁盘适配器
控制器	使用冗余磁盘控制器
磁盘	使用冗余硬件和 LVM 磁盘镜像或 RAID 镜像

在本节中，您将执行以下规划任务:

- 选择一种共享磁盘技术。
- 规划共享磁盘存储器的安装。这包括:
 - 确定处理设计的存储器容量所需的磁盘数。您需要多个用于存放镜像逻辑卷的物理磁盘。将镜像逻辑卷放置在同一物理设备上，将无法实现创建副本。有关创建镜像逻辑卷的更多信息，请参阅“规划共享 LVM 组件”。
 - 确定每个节点将包含的用于连接磁盘或磁盘子系统的磁盘适配器数。

包含逻辑卷副本的物理磁盘应位于单独的适配器上。如果所有逻辑卷副本均连接到一个适配器，那么该适配器有可能成为单点故障。如果该唯一的适配器发生故障，那么 PowerHA SystemMirror 会将卷组移动到替代节点。单独的适配器将不需要进行此移动。

– 了解每种类型的磁盘技术的连线要求。

- 向集群图添加所选磁盘配置
- 为将直接光纤通道磁带单元附件配置为集群资源进行规划。

相关参考:

第 34 页的『规划共享的 LVM 组件』

本节描述为 PowerHA SystemMirror 集群规划共享卷组。

选择共享磁盘技术

PowerHA SystemMirror 软件支持在高可用性集群中作为应用程序共享外部磁盘的磁盘技术。

有关 PowerHA SystemMirror 特定版本和 AIX 操作系统所支持的磁盘技术的特定信息，请参阅 PowerHA 硬件支持矩阵。

相关信息:

OEM 磁盘、卷组和文件系统设备

磁盘电源注意事项

可靠的电源对于高可用性集群非常重要。集群中的每个镜像磁盘链都应具有单独的电源。在规划集群时，请确保任一电源来源（PDU、电源或建筑电路）发生故障都不会导致一个以上的节点或镜像链无法工作。

因为 IBM DS4000® 系列具有冗余电源，所以它们不易发生电源问题。

规划非共享磁盘存储

请记住有关非共享磁盘存储的一些注意事项。

这些注意事项包括:

- 内部磁盘。集群中每个节点上的内部磁盘必须为以下各项提供足够的空间：
 - AIX 软件（大约 500 MB）
 - PowerHA SystemMirror 软件（对于服务器节点大约为 50 MB）
 - 高可用应用程序的可执行文件模块
- 根卷组。每个节点的根卷组都不得位于标准 SCSI 总线上。
- AIX 错误通知设施。使用 AIX 错误通知设施可监视每个节点上的设施和适配器。您可以通过在 PowerHA SystemMirror 中启用自动错误通知来监视共享和非共享磁盘，方法是在命令行中输入 `smit sysmirror`，然后选择 **Problem Determination Tools > PowerHA SystemMirror Error Notification > Configure Automatic Error Notification > Add Error Notify Methods for Cluster Resources**。
- 磁盘适配器使用。因为共享磁盘需要它们自己的适配器，所以您无法为共享和非共享磁盘使用相同的适配器。除集群内任何其他适配器外，每个节点上的内部磁盘还需要一个 SCSI 适配器。
- 卷组使用。内部磁盘必须位于与外部共享磁盘不同的卷组中。

高可用应用程序的可执行文件模块应位于内部磁盘上，而不应位于共享外部磁盘上，原因如下:

- Licensing
- 应用程序启动

相关信息:

为 PowerHA SystemMirror 配置 AIX

Licensing

供应商可能要求您为每个应用程序购买针对可能运行应用程序的每个处理器或多处理器的单独副本，并且通过在安装该应用程序时将特定于处理器的信息集成到应用程序中来保护应用程序。

因此，如果您从共享磁盘运行应用程序可执行文件，那么在发生故障转移后，例如，由于新节点上的处理器标识与安装应用程序的节点的标识不匹配，PowerHA SystemMirror 将无法在其他节点上重新启动应用程序。

应用程序可能还要求您购买所谓的节点绑定许可证，即在每个节点上的许可证文件，其中包含特定于该节点的信息。

对于在该应用程序集群内的浮动许可证（对任何集群节点均可用），可能也有数量限制。要避免此问题，请确保对于集群中可能同时运行某个应用程序的所有处理器具有足够的许可证。

启动应用程序

应用程序可能包含您可在安装期间定制并与应用程序文件一起存储的配置文件。这些配置文件通常存储在应用程序启动时使用的信息（例如，路径名和日志文件）。

您可能需要定制配置文件，如果您的配置具有以下所有要求：

- 您计划将这些配置文件存储在共享文件系统上。
- 应用程序无法在每个故障转移节点上使用相同配置。

例如，在双节点互接管配置中，两个节点可能运行同一应用程序的不同实例，并且相互支持。每个节点都必须了解应用程序两个实例的配置文件位置，并且必须能够在故障转移后访问这些配置文件。否则，故障转移将失败，从而使关键应用程序对于客户机不可用。

要减少定制配置文件所需的工作量，请在每个节点的本地文件系统上为应用程序放置略有不同的启动文件。这使初始应用程序参数可保持静态。应用程序将不需要在其每次调用时重新计算参数。

规划共享磁盘安装

本节总结了设置 PowerHA SystemMirror 集群所需的基本硬件组件。

您的集群要求取决于您指定的配置。要确保您会考虑到所有所需组件，请绘制一份系统图。此外，请查阅硬件信息，以获取有关所配置的特定设备的详细布线和连接信息。

PowerHA SystemMirror 和虚拟 SCSI

PowerHA SystemMirror 支持安装有适当 APAR 的虚拟 SCSI (VSCSI)。

以下限制适用于在集群配置中使用 VSCSI:

- 如果文件系统在备用节点上使用，那么在故障转移前不会安装这些文件系统，这样就不会意外使用备用节点上的数据。
- 如果在增强并发方式下直接访问共享卷（而不通过文件系统），那么可以从多个节点访问这些卷，因此必须在更高层（如数据库）控制访问权。
- 从虚拟 I/O 服务器 (VIOS) 的角度出发，将共享物理磁盘，而不共享逻辑卷或卷组。
- 这些共享磁盘上所有卷组的构造和维护都通过 PowerHA SystemMirror 节点（而非 VIOS）完成。

磁盘适配器

卸下适配器上的任何 SAS 端接器。请在 PowerHA SystemMirror 集群中使用外部端接器。如果您终止适配器上的共享 SAS，那么当包含适配器的集群节点发生故障时，终止将丢失。

电缆

用于连接集群中节点的所需电缆取决于所配置的 SCSI 总线的类型。请选择与磁盘适配器和控制器兼容的电缆。有关所需 SCSI 电缆的类型和长度的信息，请参阅您要在 SCSI 总线中包括的每个设备随附的硬件文档。

示例: DS4000 Storage Server 配置

本示例显示了在 PowerHA SystemMirror 环境中使用 IBM DS4000 Storage Server 时的一种高可用性配置。

请参阅下图。

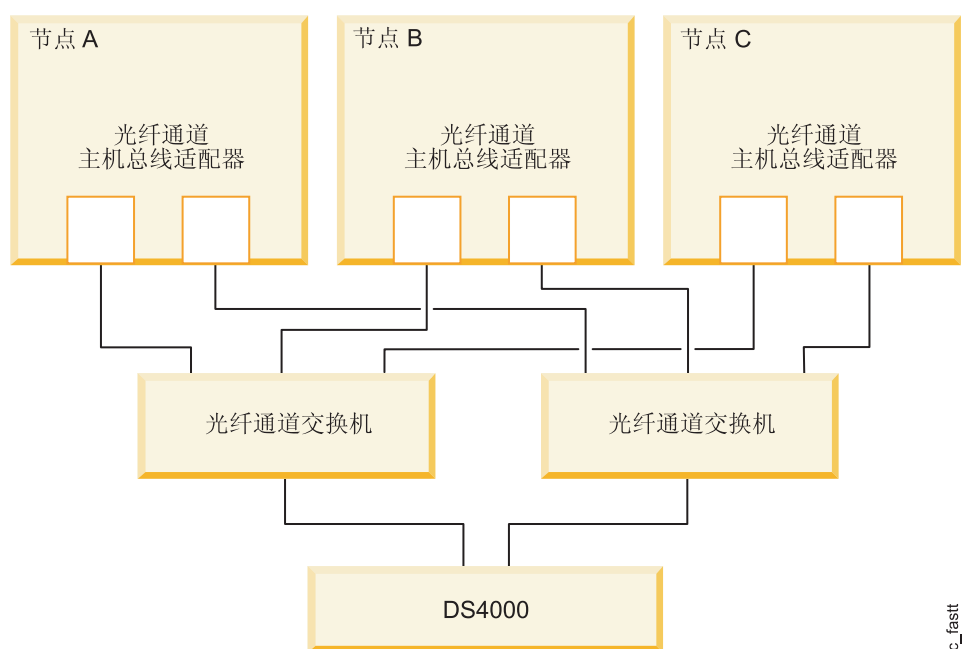


图 4. DS4000 Storage Server 环境

向集群图添加磁盘配置

在选择一项磁盘技术后，请将磁盘配置添加到您在初始集群规划中开始的集群图中。

对于集群图，请绘制代表每个共享磁盘的方框。然后，为每个方框标注共享磁盘名称。

相关参考:

第 5 页的『初始集群规划』

本部分描述为规划 PowerHA SystemMirror 集群以使应用程序具有高可用性所应采取的初始步骤。

将磁带机作为集群资源进行规划

您可以将磁带机配置为集群资源，从而使其对于集群中的多个节点高可用。

支持直接光纤通道磁带单元附件。以下 PowerHA SystemMirror 功能简化了共享磁带机的管理:

- 使用 SMIT 配置磁带机
- 磁带机正确配置的验证
- 资源组启动和停止操作期间磁带机的自动管理
- 节点故障和节点恢复时磁带机的重新分配
- 集群关闭时磁带机的受控重新分配
- 在动态重新配置期间控制磁带机的重新分配

局限性

当您计划将磁带机作为集群资源而包括时，请记住以下几点：

- PowerHA SystemMirror 将磁带装入器或堆叠器像简单磁带机一样处理。
- 可共享磁带资源的集群节点不能超过两个。
- 磁带资源不能属于并发资源组。
- 磁带机在共享磁带设备的两个节点上必须具有相同名称（例如，`/dev/rmt0`）。
- 当磁带特殊文件关闭时，缺省操作是释放磁带机。在应用程序打开磁带后，PowerHA SystemMirror 不负责磁带机的状态。
- 绝对不会同步磁带操作和应用程序控制器。如果您认为应以异步方式执行磁带预留和释放操作，那么请提供一种方法用于通知应用程序控制器等待，直到预留和释放操作完成。

保留和释放共享磁带机

当激活带有磁带资源的资源组时，将预留磁带机以允许其独占使用。

在应用程序释放此预留，或者节点从集群中除去之前，将一直维持此预留：

- 当磁带的特殊文件关闭时，缺省操作是释放磁带机。应用程序可以打开带有“关闭时不释放”标记的磁带机。PowerHA SystemMirror 将不负责在应用程序启动后保留预留。
- 当在某个节点上停止集群服务并将资源组脱机时，将释放磁带机，从而允许通过其他节点进行访问。
- 在发生意外节点故障时，在接管节点上将执行强制释放。然后，磁带机会在资源组激活期间被预留。

将磁带机设置为以同步或异步方式运行

如果在发起磁带预留或释放时某个磁带操作正在进行，那么在预留或释放操作完成之前会花费几分钟。PowerHA SystemMirror 允许同步或异步预留和释放操作。应分别为预留和释放指定同步和异步操作。

同步操作

对于同步操作（缺省值），PowerHA SystemMirror 会等待预留或释放操作（包括执行用户定义的恢复过程）完成，然后再继续。

异步操作

对于异步操作，PowerHA SystemMirror 会创建一个子进程来执行预留或释放操作（包括执行用户定义的恢复过程），并立即继续。

恢复过程

恢复过程高度依赖访问磁带机的应用程序。

对于以下操作，PowerHA SystemMirror 并非尝试预测可能的场景并提出恢复过程，而是执行用户定义的恢复脚本：

- 磁带启动
- 磁带停止

磁带启动脚本和停止脚本

在节点启动和停止、节点故障转移和重新集成，以及动态重新配置期间发生磁带启动和停止操作。当激活资源组（磁带启动）或取消激活资源组（磁带停止）时，将调用这些脚本。可在 `/usr/es/sbin/cluster/samples/tape` 目录中找到样本启动和停止脚本：

tape_resource_stop_example

- 在磁带启动期间，PowerHA SystemMirror 预留磁带机并在需要时强制释放，然后调用用户提供的磁带启动脚本。
- 在磁带停止期间，PowerHA SystemMirror 调用用户提供的磁带停止脚本，然后释放磁带机。

注：您负责正确放置磁带、终止进程或应用程序、向磁带机写入，以及在脚本内编写磁带结束标记。

其他特定于应用程序的过程应包括在启动服务器和停止服务器脚本中。

适配器故障转移和恢复

不支持具有多个 SCSI 接口的磁带机。因此，在节点在和磁带机之间仅存在一个连接。关于适配器故障转移的通常观念不适用。

节点故障转移和恢复

如果节点具有的磁带资源属于 PowerHA SystemMirror 资源组并且该节点发生故障，那么接管节点将预留磁带机并在需要时强制释放，然后调用用户提供的磁带启动脚本。

在节点重新集成时，接管节点将运行磁带停止脚本，然后释放磁带机。重新集成的节点将预留磁带机并调用用户提供的磁带启动脚本。

网络故障转移和恢复

PowerHA SystemMirror 不提供针对网络故障的磁带故障转移和恢复过程。

规划共享的 LVM 组件

本节描述为 PowerHA SystemMirror 集群规划共享卷组。

先决条件

您还应该熟悉如何使用逻辑卷管理器 (LVM)。

概述

应根据共享磁盘设备的类型和共享磁盘访问方法对 PowerHA SystemMirror 集群的共享逻辑卷管理器 (LVM) 组件进行规划。

为避免数据存储器的单点故障，请使用 LVM 或存储系统支持的数据冗余。

相关信息：

OEM 磁盘、卷组和文件系统设备
操作系统和设备管理

规划 LVM 组件

逻辑卷管理器 (LVM) 通过在物理存储器和逻辑存储器之间映射数据来控制磁盘资源。

物理存储器指数据在磁盘上的实际位置。逻辑存储器控制如何向用户提供数据。逻辑存储器可以不连续，可以进行扩展和复制，并且可以跨多个物理磁盘。这些功能提高了数据的可用性。

物理卷

物理卷是以存储器阵列方式呈现的单个物理磁盘或逻辑单元。

物理卷经过分区后向 AIX 操作系统提供映射到卷的数据映射管理方法。下图显示了在物理卷内使用物理分区的常规用法。

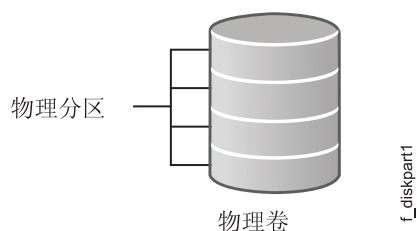


图 5. 物理卷上的物理分区:

当规划共享物理卷时，请确保：

- 在有权访问共享物理卷的所有集群节点上，该卷组的 PVID 列表相同
- 卷组的并发属性设置在所有相关集群节点中一致

卷组

卷组是一组物理卷，AIX 操作系统将其视为连续的、可寻址的磁盘区域。可以在同一卷组中放置多个物理卷。实际的数目取决于卷组的创建方法。

下图显示了一个包含三个物理卷的卷组：



图 6. 三个物理卷的卷组

在 PowerHA SystemMirror 环境中，共享卷组是完全位于由集群节点共享的外部磁盘上的卷组。一次只能使一个节点的非并发共享卷组联机。

使用共享卷组时：

- 不要在共享卷组中包含内部磁盘，因为其不能由其他节点访问。如果在共享卷组中包含了一个内部磁盘，**varyonvg** 命令会失败。

- 不要在系统引导时在 PowerHA SystemMirror 集群中手动激活共享卷组。使用集群事件脚本执行此操作。
- 确保 AIX ODM 中资源组内列出的共享卷组的自动联机属性设置为 **No**。PowerHA SystemMirror 集群验证实用程序将在集群资源验证时自动为您更正此属性，并将自动联机属性设置为 **No**。
- 如果将一个卷组定义到 PowerHA SystemMirror，那么 PowerHA SystemMirror 正在其他节点上运行时，不要在 PowerHA SystemMirror 之外的任何节点上手动对该卷组进行管理。否则将导致不可预测的结果。如果想要独立于 PowerHA SystemMirror 对卷组执行操作，请停止集群服务，执行手动卷组管理任务，保持卷组脱机，然后重新启动 PowerHA SystemMirror。要简化 PowerHA SystemMirror 使用物理卷的规划，验证实用程序将检查：
 - 卷组一致性
 - 磁盘可用性

相关信息：

mkvg 命令

varyon 命令

逻辑卷

逻辑卷是逻辑分区的集合，AIX 操作系统使其可用作单个存储单元，即磁盘的逻辑视图。

逻辑分区是物理分区的逻辑视图。逻辑分区可映射到一个、两个或三个物理分区以实现镜像。

在 PowerHA SystemMirror 环境中，逻辑卷可用于支持日志文件系统或原始设备。

文件系统

文件系统将写入到单个逻辑卷。

通常，可组织一个文件集合作为文件系统以方便快速地管理数据。

在 PowerHA SystemMirror 系统中，共享文件系统是完全位于共享逻辑卷中的日志文件系统。

您希望将共享文件系统规划为放置在集群节点共享的外部磁盘上。数据位于这些外部共享磁盘上的文件系统中，以使数据具有高可用性。

文件系统的安装顺序通常不重要。但是，如果该顺序对您的集群很重要，那么需要进行一些规划：

- 对于存在于单个资源组内的文件系统，当资源组联机时，将按字母数字顺序安装。当资源组脱机时，还将按字母数字顺序反向卸载这些文件系统。
- 如果您具有共享嵌套文件系统，需要额外小心。如果您具有在单个资源组内的共享嵌套文件系统，那么必须将该资源组的文件系统恢复方法设置为顺序，才能保证安装顺序正确。
- 如果您具有位于不同资源组的嵌套文件系统，那么必须额外地为这些资源组规划父子关系，才能保证安装顺序正确。

规划 LVM 镜像

逻辑卷管理器 (LVM) 镜像能够分配物理分区的多个副本，从而提高数据的可用性。当某个磁盘发生故障并且其物理分区变为不可用时，您仍然可以访问可用磁盘上的镜像数据。LVM 在逻辑卷内执行镜像。

在 PowerHA SystemMirror 集群内，您可以镜像以下各项：

- 共享卷组中的逻辑卷数据
- 具有文件系统的每个共享卷组的日志逻辑卷

镜像物理分区

要提高逻辑卷的可用性，您可为物理分区分配一个、两个或三个副本，镜像该分区中包含的数据。

如果某个副本由于错误而丢失，那么将访问其他未损坏的副本，并且 AIX 操作系统将使用某个正确的副本继续处理。当发生故障的物理分区的访问权恢复后，AIX 将使用一致的镜像副本的内容（数据）重新同步该物理分区的内容（数据）。

下图显示了由具有三个镜像副本的两个逻辑分区组成的逻辑卷。在该图中，每个逻辑分区均映射至三个物理分区。每个物理分区都应指定为位于单个卷组内单独的物理卷上。此配置向镜像副本提供了最大数量的备用路径，因此提供了最佳的可用性。

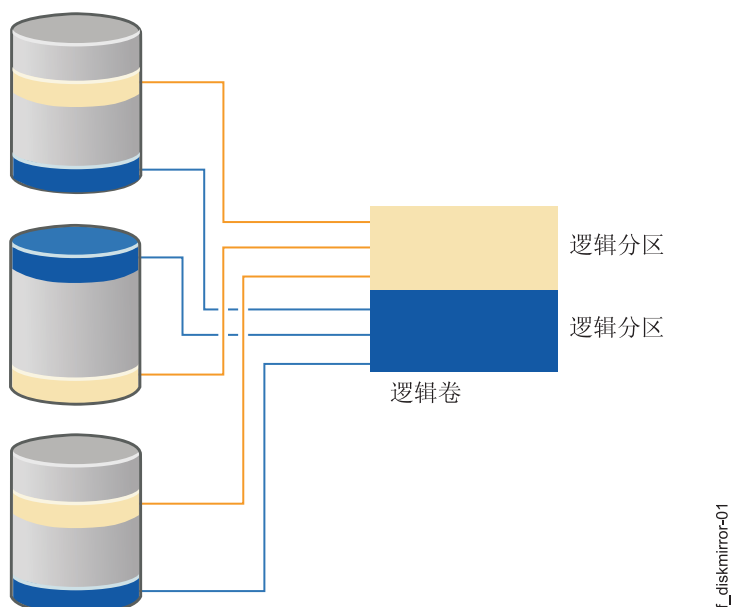


图 7. 由具有三个镜像副本的两个逻辑分区组成的逻辑卷

镜像副本是透明的，这表明您无法从这些副本中分离出其中一个。例如，如果您从具有多个副本的某个逻辑卷中删除了一个文件，那么将从该逻辑卷的所有副本中除去所删除的文件。

以下配置可增加数据可用性:

- 为逻辑分区分配三个副本，而不要分配一个或两个副本
- 将逻辑分区的副本分配在不同物理卷上，而不要将副本分配在同一物理卷上
- 在可能的情况下将逻辑分区的副本分配在不同逻辑磁盘仓中，不要分配在同一仓中
- 将逻辑分区的副本分配在不同磁盘适配器中。而不要使用单个磁盘适配器

虽然将跨多个磁盘（位于不同电源上）的镜像副本一起用于多个磁盘适配器可确保任何磁盘都不会成为集群的单点故障，但是这些配置可能增加写操作的时间。

为已为其指定强制联机的卷组中的逻辑卷指定 **superstrict** 磁盘分配策略。此配置:

- 保证逻辑卷的副本总是驻留在单独的磁盘上
- 增加在一个或多个磁盘发生故障后强制联机成功的机会。

如果计划为逻辑卷使用强制联机，那么请为集群中的磁盘仓应用 **superstrict** 磁盘分配策略。

有关强制联机的更多信息，请参阅“使用配额和联机以增加数据可用性”一节。

相关参考:

第 42 页的『使用配额和联机来增加数据可用性』

您如何为卷组配置配额和联机会增加镜像数据的可用性。

镜像日志记录

非并发访问配置支持日志文件系统和增强型日志文件系统。

AIX 操作系统为其文件系统使用日志记录功能。通常，这表明文件系统在启动时的内部状态（从块列表和空闲列表角度）与关闭时的状态相同。以实际角度而言，这表明当 AIX 启动时，任何文件的损坏程度都不能超过关闭时的损坏程度。

每个卷组包含一个 **jfslog** 或 **jfs2log** 日志，该日志本身是一个逻辑卷。此日志在卷组中通常位于与日志文件系统不同的物理磁盘上。但是，如果对该磁盘的访问权丢失，那么在这之后对文件系统进行更改是很危险的。

为避免该物理磁盘成为单点故障的可能性，您可以为每个 **jfslog** 或 **jfs2log** 日志指定镜像副本。请将这些副本放置在单独的物理卷上。

规划 LVM 分割站点镜像

您可以使用存储区域网络 (SAN) 来设置位于两个或三个不同位置的磁盘，以进行远程逻辑卷管理器 (LVM) 镜像。例如，分割站点镜像使用 LVM 来在位于每个不同位置的磁盘子系统之间复制数据以进行灾难恢复。

注: PowerHA SystemMirror 仅支持两个站点的配置。

SAN 是允许您的环境在存储设备与系统（节点）之间建立直接连接的高速网络。因此，位于不同位置的两个或多个系统可以使用 SAN 网络连接来访问相同物理磁盘。可以使用 LVM 将远程磁盘合并为一个卷组。可以将此卷组导入到位于不同位置的节点。

卷组中包含远程磁盘的逻辑卷可以具有最多三个远程镜像。您可以在每个位置至少设置一个远程镜像。逻辑卷中存储的数据具有高可用性。因此，如果环境中发生故障，例如某个位置的所有节点都不可用，那么另一个位置的远程镜像具有最新数据。

发生磁盘或节点故障后，PowerHA SystemMirror 将自动同步所有远程镜像，并且节点将重新恢复联机状态。即使某个磁盘处于 PVREMOVED 状态或 PVMISSING 状态，也会发生自动同步。自动同步并非对于 LVM 分割站点镜像的所有情况都可用。如果不可用，那么可以使用 C-SPOC 来同步数据。

规划 LVM 分割站点镜像配置时，还必须规划集群中使用的存储库磁盘。必须验证是否存在准备就绪的另一个磁盘，可在主磁盘发生故障时用作存储库磁盘。

注: Cluster Aware AIX 支持实时存储库替换，而不会影响重要的集群功能。

示例

下图是使用 SAN 的 LVM 分割站点镜像的配置示例。

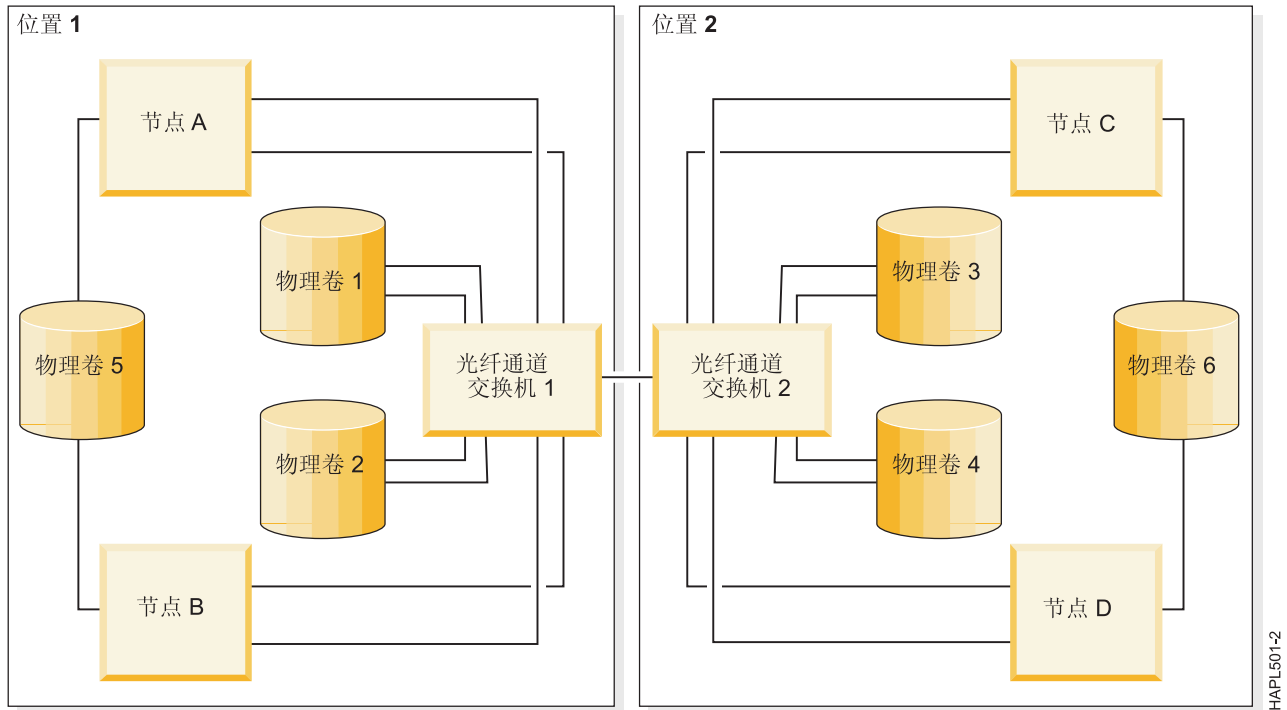


图 8. 使用 SAN 的 LVM 分割站点镜像配置

您可以对连接到两个位置中每个位置的至少一个节点的磁盘进行镜像。在此示例中，物理卷 4 通过光纤通道交换机 1 和光纤通道交换机 2 连接，对于位置 1 上的节点 A 和节点 B 以及位置 2 上的节点 C 可用。您可以在位置 1 上具有物理卷 4 的镜像。仅连接到一个位置上的节点的磁盘（物理卷 5 和物理卷 6）无法在位置间进行镜像。

可以使用 AIX LVM 镜像池功能来确保在两个位置之间正确并完整地数据镜像。如果物理卷 1 和物理卷 2 位于一个镜像池中，并且物理卷 3 和物理卷 4 位于另一个镜像池中，那么 LVM 允许在每个位置上存在一个完整数据副本。您必须使用 `superstrict` 镜像池来保证每个位置存在一个完整数据副本。

在此示例中，即使在添加、除去或替换卷组中的磁盘时，您也可以使用镜像池来帮助维护镜像。您还可以使用 C-SPOC 功能来定义镜像池，并将它们与每个位置上的磁盘关联。

相关信息:

管理 LVM 分割站点镜像

跨站点镜像

您可以使用存储区域网络 (SAN) 来设置位于两个不同站点的磁盘，以使用远程逻辑卷管理器 (LVM) 镜像。将在每个站点上的磁盘子系统之间复制数据以进行灾难恢复。

SAN 是允许在存储设备与处理器之间建立直接连接的高速网络。因此，位于不同站点的两个或多个节点可以通过公共 SAN 来访问由一定距离分隔开的相同物理磁盘。可以使用 LVM 将这些远程磁盘合并为一个卷组，并且可以将此卷组导入到位于不同站点上的节点。此卷组中的逻辑卷可以有最多三个镜像。因此，您可以在每个站点上至少设置一个镜像。此逻辑卷中存储的信息将保持高度可用，如果发生某些故障，另一个站点上的远程镜像仍将具有最新信息，这样就可以在另一个站点上继续操作。

发生磁盘或节点故障以及进行后续的重新集成后，PowerHA SystemMirror 将自动同步镜像。即使某个磁盘处于 PVREMOVED 或 PVMISSING 状态，PowerHA SystemMirror 也会处理镜像的自动同步。自动同步并非对于所有情况都可用，但是发生磁盘或节点故障以及进行后续的重新集成后，您可以使用 C-SPOC 来将数据从继续存在的镜像同步至过时镜像。

注：在 PowerHA SystemMirror Enterprise Edition 中，还可以通过使用地理逻辑卷管理器 (GLVM) 镜像功能，在跨两个站点的集群中使用镜像。

规划磁盘访问

您可以将磁盘配置为增强型并发访问还是非并发访问。

- **增强型并发访问。** 磁盘上的数据对于以并发方式连接的所有节点和对磁盘上的元数据具有访问权的所有节点均可用。此访问方式支持快速磁盘接管，原因是可以在读取元数据之前将卷组联机。

无论共享卷组是否以并发方式访问，必须将所有共享卷组配置为增强型并发模式卷组。您可以将现有卷组迁移为增强型并发模式。

如果您将相应的增强型卷组（在“资源组”策略中）配置为以并发方式通过两个或多个节点访问，那么不支持使用日志文件系统 (JFS) 和增强型日志文件系统 (JFS2)。

使用 IBM TotalStorage DS Series 或 IBM 2105 Enterprise Storage Servers 的并发访问配置不使用配置管理器 (LVM) 镜像。相反，这些系统可提供它们自己的数据冗余。

- **非并发访问。** 每次只有一个节点可以访问磁盘上的信息。

如果包含这些磁盘的资源组移动到其他节点，那么新节点可以访问这些磁盘、读取元数据（有关卷组和其他组件当前状态的信息）、在卷组上运行 **varyon** 命令，以及安装任何关联的文件系统。

非并发访问配置通常使用日志文件系统。在有些情况下，在非并发环境中运行的数据库应用程序可能绕开日志文件系统并直接访问原始逻辑卷。

相关参考：

『增强型并发访问』

对于连接多个节点的受 PowerHA SystemMirror 支持的任何磁盘都必须放置在增强型并发模式卷组中，并且可以在并发环境而非并发环境中使用（由资源组的类型指定）。

增强型并发访问

对于连接多个节点的受 PowerHA SystemMirror 支持的任何磁盘都必须放置在增强型并发模式卷组中，并且可以在并发环境而非并发环境中使用（由资源组的类型指定）。

- **并发。** 应用程序同时所有活动集群节点上运行。

为使此类应用程序能够访问其数据，并发卷组对于所有活动集群节点都处于联机状态。应用程序负责确保数据访问一致。

- **非并发。** 应用程序每次在一个节点上运行。

不以并发方式对卷组进行访问。在任何给定时间，它们仍然只由一个节点访问。

当您在集群中拥有卷组的所有节点上将卷组以增强型并发模式联机时，LVM 将允许访问所有节点上的卷组。但是，这将限制在所有节点上进行较高级别连接（例如，NFS 安装和 JFS 安装），而仅允许在当前拥有 PowerHA SystemMirror 中的卷组的节点上进行这些连接。

您可以使用 AIX MPIO 功能通过多条路径访问磁盘子系统。多条路径提供了比使用单条路径更大的吞吐量和更高的可用性。在使用多条路径的情况下，特别是当一条路径由于适配器、电缆或交换机故障而发生故障时，将不会导致应用程序丢失对数据的访问权。虽然 PowerHA SystemMirror 将尝试从完全丢失卷组访问权的状况中恢复，但是该丢失状况本身将造成临时破坏。AIX MPIO 功能可以避免单个组件故障造成应用程序停运。

在使用快速磁盘接管时，不使用磁盘预留功能。如果集群发生分区，那么每个分区中的节点会使卷组意外地以主动状态联机。卷组以主动状态联机，将允许安装文件系统和更改物理卷，所以此状况会导致同一卷组中出现不同副本。有关快速磁盘接管和使用多网络的更多信息，请参阅“使用快速磁盘接管”一节。

由 MPIO 访问的磁盘的并发访问要求

增强型并发模式是创建并发卷组的唯一选项。在 PowerHA SystemMirror 中，增强型并发模式卷组不使用磁盘预留量。在 PowerHA SystemMirror 中自动提供了由 MPIO 访问的磁盘所需的并发访问。

关于增强型并发模式

缺省情况下，所有并发卷组都创建为增强型并发模式卷组。对于增强型并发卷组，并发逻辑卷管理器 (CLVM) 通过 AIX 操作系统中 Reliable Scalable Cluster Technology (RSCT) 功能的“组服务”组件协调节点间的更改。“组服务”协议通过集群节点间的通信链路流动。

相关任务:

将卷组转换为增强式并行模式

相关参考:

『使用快速磁盘接管』

PowerHA SystemMirror 可自动检测发生故障的卷组并对作为非并发资源组中的资源而包括在内的增强型并发模式卷组启动快速磁盘接管。

使用快速磁盘接管

PowerHA SystemMirror 可自动检测发生故障的卷组并对作为非并发资源组中的资源而包括在内的增强型并发模式卷组启动快速磁盘接管。

快速磁盘接管对包含大量磁盘的增强型并发模式卷组的故障转移特别有用。这种磁盘接管机制比用于包括在非并发资源组中的标准资源组的磁盘接管机制更快速。在快速磁盘接管期间，PowerHA SystemMirror 将跳过打破磁盘预留量所需的额外处理，或通过运行懒惰更新来更新和同步逻辑卷管理器 (LVM) 信息所需的额外处理。

对于带有两个磁盘的卷组，已观察到快速磁盘接管可在 10 秒内完成。对于大量的磁盘和卷组，期望此时间缓慢增加。在任意配置中观察到的实际时间取决于 PowerHA SystemMirror 控制范围之外的因素，例如，节点的处理能力以及发生故障转移时无关活动的数量。完成故障转移处理的实际观测值取决于更多因素，例如，是否需要文件系统检查，以及重新启动应用程序所需的时间。

注：不以并发方式对增强型并发模式卷组进行访问。在任何给定时间，它们只由一个节点访问。快速磁盘接管机制在卷组级别工作，因此与使用的磁盘数无关。

快速磁盘接管以及主动和被动联机

可以使增强型并发卷组在节点上处于活动状态，或使其主动或被动联机。

为启用快速磁盘接管，PowerHA SystemMirror 可以主动和被动状态激活增强型并发卷组。

主动联机

主动联机的行为与普通联机相同，并且使逻辑卷可用。当增强型并发卷组在节点上以主动状态联机时：

- 对文件系统的操作，如文件系统安装
- 对应用程序的操作
- 对逻辑卷的操作，如创建逻辑卷
- 同步卷组

被动联机

当增强型并发卷组以被动状态联机时，LVM 将在 LVM 级别为卷组提供等效的磁盘电子篱笆。

被动状态联机仅允许对卷组执行有限数量的只读操作：

- LVM 对卷组的特殊文件进行只读访问
- LVM 对卷组拥有的所有逻辑卷的前 4 Kb 进行只读访问。

当卷组以被动状态联机时，不允许以下操作：

- 对文件系统的操作，如文件系统安装
- 逻辑卷上的任何操作，如打开逻辑卷
- 同步卷组

PowerHA SystemMirror 以及主动和被动联机

PowerHA SystemMirror 可正确地将卷组在拥有资源组的节点上以主动状态联机，并根据资源组状态和位置的更改，相应地更改主动和被动状态。

- 在集群启动时：
 - 在拥有资源组的节点上，PowerHA SystemMirror 以主动状态激活资源组。PowerHA SystemMirror 每次只在一个节点上以主动状态激活卷组。
 - PowerHA SystemMirror 在集群中的所有其他节点上以被动状态激活卷组。
- 故障转移时：
 - 如果节点释放资源组，或者如果资源组由于任何其它原因移动到另一个节点上，那么 PowerHA SystemMirror 会在释放资源组的节点上将卷组的联机状态从主动切换为被动。然后，PowerHA SystemMirror 在获取到资源组的节点上以主动状态激活卷组。
 - 卷组在集群中的所有其他节点上仍处于被动状态。
- 当发生节点重新集成时，PowerHA SystemMirror 将执行以下过程：
 - 在释放资源组的节点上将卷组的联机状态从主动更改为被动。
 - 在连接节点上以主动状态将卷组联机。
 - 在集群中的所有其他节点上以被动状态激活该卷组。

注：为防止同时在一个以上的节点上安装文件系统，主动和被动状态之间的切换很有必要。

使用配额和联机来增加数据可用性

您如何为卷组配置配额和联机会增加镜像数据的可用性。

使用配额

配额可确保在卷组中有一半以上的可用物理磁盘。

配额不跟踪逻辑卷镜像，因此不能用于确保数据可用性。当您仍然具有所有数据时，会丢失配额。反之，您会丧失对一些数据的访问权，但不丢失配额。

Quorum 对于 RAID 阵列（例如，ESS 和 IBM TotalStorage DS Series）上的卷组很有利。RAID 设备提供数据可用性并可从单个磁盘丢失中恢复。镜像通常不用于完全包含在单个 RAID 设备内的卷组。如果某个卷组在 RAID 设备之间镜像，那么即使其中一个 RAID 设备丢失，强制联机功能仍可以将卷组联机。

请为每个卷组决定要启用还是禁用配额。下表显示了当卷组联机和脱机时配额的影响方式：

	使卷组联机的条件	使卷组脱机的条件
启用配额	在卷组中有超过 50% 的磁盘可用	无法访问 50% 或更多磁盘
禁用配额	卷组中所有磁盘均可用	无法访问任何磁盘

缺省情况下，启用配额检查。您可以通过使用 `chvg -Qn vgrname` 命令或通过使用 `smit chvg` 快速路径来禁用配额。

相关信息：

chvg 命令

并发访问配置中的配额：

必须为 PowerHA SystemMirror 并发访问配置启用配额。禁用配额将导致数据损坏。对于多个故障会导致在集群节点之间没有通用共享磁盘的任何并发访问都有数据损坏或不一致的潜在可能。

下图显示为消除单点故障而配置了两组 IBM 磁盘子系统的集群。逻辑卷通过子系统镜像，并且每个磁盘子系统均连接到具有单独 NIC 的每个节点。

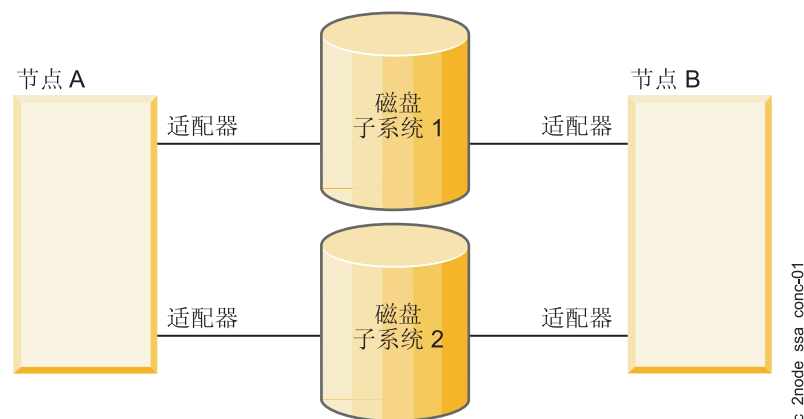


图 9. IBM 磁盘子系统并行访问配置

如果多个故障导致在每个节点和一组磁盘失去通信的情况如下：节点 A 可访问子系统 1 但不能访问子系统 2，而节点 B 可访问子系统 2 但不能访问子系统 1。两个节点都将通过它们可访问的镜像副本操作同一数据基线。但是，每个节点都不会看到另一节点对磁盘数据所作的修改。因此，在两个节点间数据变为不一致。

如果启用了配额保护，通信故障将导致其中一个或两个节点将卷组脱机。虽然应用程序无权访问已脱机的卷组上的数据，但是数据一致性得以保留。

由丢失配额触发的选择性故障转移:

PowerHA SystemMirror 可选择性恢复受特定资源故障影响的非并发资源组（启动策略不是 Online on All Available Nodes）。PowerHA SystemMirror 将自动回应与某个在集群节点上脱机的卷组相关联的 LVM_SA_QUORCLOSE 丢失配额错误。作为对此错误的响应，非并发资源组将在发生错误的节点上脱机。

如果 AIX Logical Volume Manager 由于资源组中的某个卷组丢失配额而将该卷组在节点上脱机，那么 PowerHA SystemMirror 会选择性地将资源组移动到另一个节点。您可以通过定制资源恢复来更改此缺省行为，以使用通知方法来替换故障转移。

作为对 LVM_SA_QUORCLOSE 错误的响应，PowerHA SystemMirror 将启动选择性故障转移并恢复受影响的资源组。此错误由 AIX LVM 针对特定错误状况生成，即使未将卷组定义为启用配额也会如此。AIX LVM 可能还会生成其他类型的错误通知；但是，在缺省情况下，PowerHA SystemMirror 不会对这些错误作出回应。在所有其他情况中，您仍然需要配置定制的错误通知方法，或使用 AIX 自动错误通知方法来对卷组故障作出回应。

您可以使用 rootvg 系统事件来监视 rootvg 访问权的丢失。如果系统丢失访问权，在缺省情况下，PowerHA SystemMirror 将在系统错误日志中记录一个事件并重新引导系统。您可以通过使用 SMIT 将此设置更改为记录一个事件但不重新引导系统。

仅可当 rootvg 磁盘正使用本机 AIX 多路径 I/O (MPIO) 驱动程序且 rootvg 磁盘不是内部并行 SCSI 磁盘时，可监视 rootvg 事件。要验证 rootvg 磁盘是否正使用 MPIO 驱动程序，请在命令行上，输入 `lspath -l hdiskname`，其中，*hdiskname* 是 rootvg 磁盘的名称。如果 rootvg 磁盘未在使用 MPIO 驱动程序，那么会显示以下错误消息：

```
lspath: 0514-538 Cannot perform the requested function because the
          specified device does not support multiple paths.
```

相关信息:

用于卷组损失的错误通知方法

PowerHA SystemMirror 监视系统事件

用于处理资源组的选择性故障转移

使用强制联机

PowerHA SystemMirror 提供了强制联机功能，以与 AIX 自动错误通知方法联合使用。强制联机功能使您具有可能的最高数据可用性。

通过强制卷组联机，只要有一个可用的有效数据副本，您就可以使卷组保持联机状态。请仅对具有镜像逻辑卷的卷组使用强制联机。

注：当使用此功能时，请小心使用以避免分区集群。

您可以在以下情况下使用 SMIT 在某个节点上强制卷组联机：由于缺少配额而导致该卷组上的正常 **varyon** 命令失败，但是具有一个可用的有效数据副本。如果在两个磁盘仓之间镜像数据，并且其中一个磁盘仓变为不可用，那么使用 SMIT 来强制联机对于本地灾难恢复是很有用的。

注：对于使用逻辑卷管理器 (LVM) 镜像的 SCSI 磁盘上的卷组，以及在不同 RAID 或 ESS 设备之间镜像的卷组，可以指定强制联机属性。

当磁盘不可用时，如果希望强制卷组联机，请使用 **varyonvg -f**，无论是否有您数据的副本，这都会强制卷组联机。您可以在 SMIT 中为资源组中的卷组指定强制联机。

强制联机 and 集群分区

如果您使用强制联机，那么在共享存储器的节点之间应存在多个网络连接，这一点很重要。多个网络连接有助于确保在某个网络发生故障的情况下，每个节点始终具有到其他节点的通信路径。具有多个网络连接可避免您的集群发生分区。否则，网络故障可能导致节点尝试接管在其他节点上仍处于活动状态的资源组。在此情况下，如果您设置了强制联机设置，那么可能发生数据丢失或离散。

将 NFS 与 PowerHA SystemMirror 结合使用

PowerHA SystemMirror 软件增强了网络文件系统 (NFS) 处理的可用性。

这些增强功能包括:

- 可靠 NFS 服务器功能，允许某个备份处理器在主 NFS 服务器故障的情况下恢复当前 NFS 活动，同时保留对 NFS 文件系统和重复请求高速缓存的锁定。如果资源组包含 NFS V2 或 V3 导出，那么此功能仅限于两个节点的资源组。仅包含 NFS V4 或更高版本的资源组最多可支持 16 个节点的配置。
- NFS 配置助手，用于简化设置和配置。
- 预配置的应用程序控制器和应用程序监视器 (clam_nfsv4)，用于监视 NFS V4 导出以及 NFS 守护程序。
- 能够为 NFS 安装指定一个网络。
- 能够在目录级别定义 NFS 导出和安装。
- 能够指定 NFS 导出的目录和文件系统的导出选项。

要使 NFS 可在 PowerHA SystemMirror 集群上按预期工作，有一些特定的配置要求。因此，您必须规划以下任务:

- 创建共享卷组
- 导出 NFS 文件系统
- NFS 安装和故障转移

PowerHA SystemMirror 脚本处理缺省 NFS 行为。您可能需要将脚本修改为处理您的特定配置。

您可以将所有资源组中的 NFS 配置为以非并发方式行动；即，它们没有“Online on All Available Nodes”启动策略。

在 PowerHA SystemMirror 集群中放弃对 NFS 文件系统的控制

在您配置包含 NFS 文件系统的资源组后，您就可将对 NFS 文件系统的控制权转交给 PowerHA SystemMirror。

NFS 文件系统加入属于活动 PowerHA SystemMirror 集群的资源组后，PowerHA SystemMirror 将在集群事件期间（如，包含文件系统的资源组故障转移到集群中的其他节点）负责交叉安装以及卸载文件系统。

如果由于某些原因，您停止了集群服务并且必须手动管理 NFS 文件系统，那么在您重新启动集群服务之前，必须先卸载文件系统。这样，在节点连接集群后，就可通过 PowerHA SystemMirror 来管理 NFS 文件系统。

可靠 NFS 服务器功能

PowerHA SystemMirror 集群可以利用标准 NFS 功能的 AIX 扩展功能，通过这些扩展功能，它可以在 NFS 服务器故障转移和重新集成期间正确处理重复请求并恢复锁状态。

当 NFS 客户机使用 NFS 锁定来仲裁对标准 NFS 文件系统的访问权时，每个资源组限制为两个节点。使用可靠 NFS 的每个资源组包含一对 PowerHA SystemMirror 节点。

集群中的独立节点可提供可靠 NFS 服务。例如，在四节点的集群中，您可以设置两个 NFS 客户机和服务器对（例如，节点 A 和节点 B 提供一组可靠 NFS 服务，节点 C 和节点 D 可提供另一组可靠 NFS 服务）。对 1 可为一组 NFS 文件系统提供可靠 NFS 服务，对 2 可向另一组 NFS 文件系统提供可靠 NFS 服务。无论是否配置了 NFS 交叉安装都是如此。只要参与资源组的节点遵循此示例概述的限制，PowerHA SystemMirror 就不限制资源组或 NFS 文件系统的数量。

为 NFS 指定 IP 地址

主机名必须能够解析为始终存在于节点上并且在接口上始终处于活动状态的 IP 地址。主机名不能是可以移动至其他节点的服务 IP 地址。

要确保 NFS 将使用的 IP 地址始终存在于节点上，您可以：

- 使用与持久标签 相关联的 IP 地址
- 对于通过别名的 IPAT 配置，请使用在引导时使用的 IP 地址
- 使用存在于不受 PowerHA SystemMirror 控制的接口上的 IP 地址

共享卷组

当创建共享卷组时，通常您可以将 **Major Number** 字段留空，并让系统提供缺省值。但是，NFS 使用卷组主数来帮助唯一标识导出的文件系统。因此，在包含以 NFS 方式导出的文件系统的资源组中，所有要包括的节点针对该文件系统所在的卷组都必须具有相同的主数。

在发生节点故障时，与 PowerHA SystemMirror 集群连接的 NFS 客户机将以标准 NFS 服务器发生故障并重新引导时相同的方式运行。对文件系统的访问将暂挂，然后在文件系统再次变为可用时恢复。但是，如果主数不同，当其他集群节点接管文件系统并重新导出文件系统时，客户机应用程序将不会恢复。因为节点导出的文件系统看起来将与发生故障的节点导出的文件系统不同，所以客户机应用程序将不会恢复。

导出 NFS 文件系统和目录

在 PowerHA SystemMirror 中导出 NFS 文件系统和目录的过程与在 AIX 操作系统中的过程不同。

当在 PowerHA SystemMirror 中规划导出 NFS 文件系统和目录时，请记住以下几点：

- 要导出的 NFS 文件系统和目录：

在 AIX 中，您通过使用 **smit mknfsexp** 命令（该命令将创建 **/etc/exports** 文件）指定要导出的 NFS 文件系统和目录。在 PowerHA SystemMirror 中，您通过将要导出的 NFS 文件系统和目录包括在一个 PowerHA SystemMirror 资源组中来指定这些文件系统和目录。

- 针对以 NFS 方式导出的文件系统和目录的导出选项：

如果要在 PowerHA SystemMirror 中指定用于导出 NFS 的特殊选项，那么可以创建 **/usr/es/sbin/cluster/etc/exports** 文件。此文件与常规 AIX **/etc/exports** 文件具有相同格式。

注：使用该备用导出文件是可选的。导出 NFS 文件系统或目录时，PowerHA SystemMirror 将检查 **/usr/es/sbin/cluster/etc/exports** 文件。如果在该文件中存在针对该文件系统或目录的条目时，PowerHA SystemMirror 将使用所列的选项。如果要导出的 NFS 文件系统或目录未在该文件中列出，或者如果不存在备用文件，那么将使用针对所有集群节点的缺省选项“根目录访问权”从 NFS 中导出文件系统或目录。

- 指定要导出的文件系统的资源组：

在 SMIT 中，请针对资源组将 **Filesystems Mounted before IP Configured** 字段设置为 **true**。通过将 **Filesystems Mounted before IP Configured** 字段设置为 **true**，可确保在导出文件系统后执行 IP 地址接管。如果 IP 地址先受管，那么 NFS 服务器将拒绝客户机请求，直到文件系统已导出为止。

- 用于 NFS V4 导出的稳定存储器：

请为 NFS V4 导出使用稳定存储器，并使其对于资源组的所有参与节点均可访问。NFS V4 使用此文件系统空间存储与 NFS 客户机事务相关的状态信息。稳定存储器中的状态信息对于平稳完成故障转移、回退以及将资源组的选项从一个节点移动到另一节点而言至关重要，同时还使 NFS V4 客户机的状态不受影响。

在资源组处于联机时，无法更改稳定存储器的位置。

相关参考:

『PowerHA SystemMirror 中的 NFS 交叉安装』

NFS 交叉安装是特定于 PowerHA SystemMirror 的 NFS 配置，其中集群中的每个节点都可以同时作为 NFS 服务器和 NFS 客户机。在从一个节点导出文件系统期间，文件系统将以 NFS 方式安装在资源组的所有节点上，其中包括导出文件系统的节点。还可以从其他节点导出其他文件系统，并且会将其以 NFS 方式安装在所有节点上。

NFS 和故障转移

要使 PowerHA SystemMirror 和 NFS 可一起正常工作，必须将 NFS 服务器的 IP 地址配置为实现高可用性的资源组中的资源。

为确保最佳的 NFS 性能，PowerHA SystemMirror 使用的 NFS 文件系统应在 `/etc/filesystems` 文件的 `options` 字段中包括条目 `vers = <version number>`。

相关参考:

第 48 页的『NFS 交叉安装和 IP 标签』

为启用 NFS 交叉安装，每个集群节点都可以作为 NFS 客户机。其中每个节点都必须具有到 NFS 服务器节点的服务 IP 标签的有效路径。即，要启用 NFS 交叉安装，客户机节点上必须存在 IP 标签，并且必须在与 NFS 服务器节点的服务 IP 标签相同的子网上配置此 IP 标签。

PowerHA SystemMirror 中的 NFS 交叉安装

NFS 交叉安装是特定于 PowerHA SystemMirror 的 NFS 配置，其中集群中的每个节点都可以同时作为 NFS 服务器和 NFS 客户机。在从一个节点导出文件系统期间，文件系统将以 NFS 方式安装在资源组的所有节点上，其中包括导出文件系统的节点。还可以从其他节点导出其他文件系统，并且会将其以 NFS 方式安装在所有节点上。

当环境使用 NFS V2 和 V3 在资源组中导出时，交叉安装功能将仅限于两个节点的资源组。如果资源组仅包含 NFS V4（或更高版本）导出，那么交叉安装功能扩展至支持资源组的任意数量的节点。

资源组中的每个节点都属于互接管（或主动-主动）集群配置的一部分，可提供和安装 NFS 文件系统。

缺省情况下，包含已导出的 NFS 文件系统的资源组将自动交叉安装这些文件系统（如果导出和导入均已配置），如下所述：

- 在当前托管资源组的节点上，组中的所有 NFS 文件系统都将以 NFS 方式导出。
- 可能托管此资源组的每个节点都将以 NFS 方式安装资源组中的所有 NFS 文件系统。

应用程序访问属于该资源组的任何节点上的 NFS 文件系统。

使用接管 IP 地址配置的资源组发生故障转移时，接管节点会在本地安装 NFS 文件系统，然后重新导出此文件系统。资源组中的所有其他节点将保持它们的 NFS 文件系统安装。

在 NFS 交叉安装配置中，在资源组中定义的任何 NFS 安装在资源组中必须具有相应 NFS 导出。如果您的 NFS 交叉安装不遵循此配置，那么会显示以下消息：

```
claddress: WARNING: NFS mounts were specified for the resource group
'<RG name>';however no NFS exports have been specified.
```

如果 NFS 交叉安装字段包含值，那么相应 NFS 导出文件系统也必须包含值。

NFS 交叉安装和 IP 标签:

为启用 NFS 交叉安装，每个集群节点都可以作为 NFS 客户机。其中每个节点都必须具有到 NFS 服务器节点的服务 IP 标签的有效路径。即，要启用 NFS 交叉安装，客户机节点上必须存在 IP 标签，并且必须在与 NFS 服务器节点的服务 IP 标签相同的子网上配置此 IP 标签。

如果 NFS 客户机节点具有位于同一网络上的服务 IP 标签，那么这不算问题。但是，在某些集群配置中，您需要创建有效路径。

NFS 服务器路径的创建方法:

确保可访问 NFS 服务器的最简单方法是使客户机节点上的 IP 标签与 NFS 服务器节点上的服务 IP 标签位于同一子网上。

要在 NFS 客户机节点和导出文件系统的节点之间创建有效路径，您可以用以下其中一种方法配置环境:

- 在服务 IP 网络和子网上配置单独一个具有 IP 标签的 NIC
- 在服务 IP 网络和子网上具有一个持久节点 IP 标签

请注意，由于在 PowerHA SystemMirror 中为 NFS 文件系统设置的缺省导出选项，这些解决方法不会向文件系统提供自动根许可权。

要允许对客户机节点上以 NFS 方式安装的文件系统进行根级别访问，请将节点的所有 IP 标签或地址都添加到集群导出文件 `/usr/es/sbin/cluster/etc/exports` 的 `root =` 选项中。您可以在第一个节点上执行此操作，原因是通过同步集群资源可将此信息传播到其他集群节点。

相关参考:

第 46 页的『导出 NFS 文件系统和目录』

在 PowerHA SystemMirror 中导出 NFS 文件系统和目录的过程与在 AIX 操作系统中的过程不同。

为 NFS V4 导出创建并配置稳定存储器:

稳定存储器是 NFS V4 服务器用于保存状态信息的文件系统空间。这对于保存 NFS V4 客户机的状态信息，以促进资源组平稳且透明地从一个节点故障转移/回退/移动到另一个节点而言是极其至关重要的。

稳定存储器的要求:

- 建议大小为 512 MB。
- 建议您具有用于稳定存储器的专用文件系统，但是可以接受现有文件系统的子目录。
- 稳定存储器所在的文件系统和卷组应属于资源组。
- PowerHA SystemMirror 会尽量尝试验证为稳定存储器指定的路径是否属于资源组中的文件系统；但是，这些检查无法将符号链接考虑在内，原因是发生验证时文件系统可能未在本地安装。请避免使用会干扰 PowerHA SystemMirror 验证的符号链接。
- 在将稳定存储器目录添加到资源组之前，确保稳定存储器目录为空（没有任何预先数据）。
- 虽然必须将稳定存储器存储在资源组管理的文件系统中，但是不应将其存储在资源组导出的 NSF 目录中。
- 配置助手提供了 AUTO_SELECT 选项。如果您选择此选项，那么 PowerHA SystemMirror 将使用属于给定资源组的卷组列表中的卷组。然后，PowerHA SystemMirror 会创建一个逻辑卷和一个文件系统，以用作稳定存储器位置。

示例: 两个节点的 NFS 交叉安装配置:

在此示例中, 节点 A 当前托管一个非并发资源组 RG1, 该资源组包括 /fs1 (作为导出的 NFS 文件系统) 和 service1 (作为服务 IP 标签)。

在此示例中, 节点 B 当前托管一个非并发资源组 RG2, 该资源组包括 /fs2 (作为导出的 NFS 文件系统) 和 service2 (作为服务 IP 标签)。在重新集成时, 会将 /fs1 传回至节点 A, 在本地进行安装, 然后导出。节点 B 将通过 NFS 再次安装 /fs1。

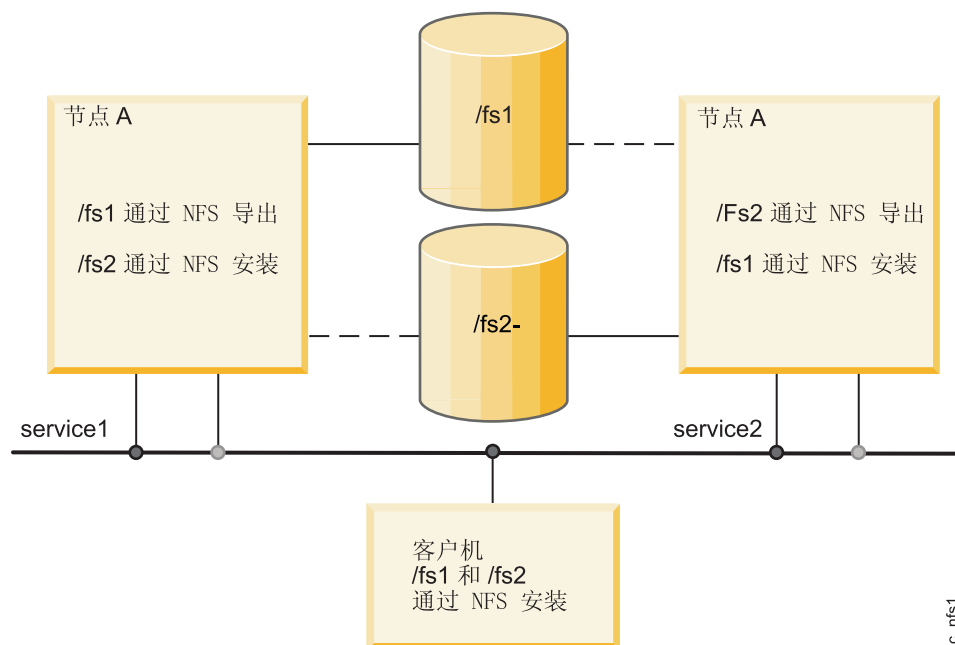


图 10. 两个节点的 NFS 交叉安装

应在 SMIT 中按如下所述定义两个资源组。

Resource group	RG1	RG2
Participating node names	Node A Node B	Node B Node A
File systems 当前拥有资源组的节点将在本地安装的文件系统。	/fs1	/fs2
File systems to export 当前拥有资源组的节点将以 NFS 方式导出的文件系统。该文件系统是上面列示的文件系统的子集。	/fs1	/fs2
File systems to NFS mount 资源组中的所有节点都将以 NFS 方式安装的文件系统和目录。第一个值是 NFS 安装点。第二个值是本地安装点。	/mnt1;/fs1	/mnt2;/fs2
File systems mounted before IP configured	true	true

在此场景中:

- 节点 A 在本地安装并导出 /fs1, 然后在 /mnt1 上交叉安装。
- 节点 B 通过 NFS 从节点 A 在 /mnt1 上安装 /fs1。

设置与此相似的资源组可确保期望的缺省节点到节点 NFS 行为。

当节点 A 发生故障时，节点 B 将关闭节点 A /fs1 中的所有打开文件，将其卸载，在本地进行安装，然后将其重新导出至等待客户机。

在接管后，节点 B:

- 已在本地安装 /fs2
- 已以 NFS 方式导出 /fs2
- 已在本地安装 /fs1
- 已以 NFS 方式导出 /fs1
- 已通过 /mnt1 以 NFS 方式安装 service1:/fs1
- 已通过 /mnt2b 以 NFS 方式安装 service2:/fs2

两个资源组均将这两个节点包含为可能的资源组所有者。

对于交叉安装的 NFS 文件系统的资源组接管

本节描述如何设置具有交叉安装的 NFS 文件系统的非并发资源组，以便在接管和重新集成期间正确处理 NFS 文件系统。此外，非并发资源组在故障转移期间支持跨服务器的自动 NFS 安装。

设置与本地安装点不同的 NFS 安装点:

PowerHA SystemMirror 以两种方式处理非并发资源组中的 NFS 安装

这些限制包括:

- 当前拥有资源组的节点将通过文件系统的本地安装点安装文件系统，并且此节点将导出 NFS 文件系统。
- 资源组中的所有节点（包括该组的当前所有者）均通过不同的安装点安装 NFS 文件系统。

因此，该组的所有者会两次安装文件系统。一个文件系统安装作为本地安装进行安装，另一个为 NFS 安装。

注: NFS 安装点必须在本地安装点的目录树之外。

因为 IPAT 在具有以 NFS 方式安装的文件系统的资源组中使用，所以在故障转移期间，节点不会卸载并重新安装文件系统。当资源组故障转移至新节点后，获取到的节点将在本地安装文件系统并以 NFS 方式将其导出。（在故障转移期间，以 NFS 方式安装的文件系统将临时对于集群节点不可用。）只要新节点获取到 IPAT 标签，NFS 文件系统的访问权就会恢复。

所有应用程序都必须通过以 NFS 方式安装的文件系统来引用文件系统。如果所使用的应用程序必须始终以同一安装点名称来引用文件系统，那么您可以更改本地文件系统安装的安装点（例如，将其更改为本地安装点，并使用前一本地安装点作为新的 NFS 安装点）。

PowerHA SystemMirror 的缺省 NFS 装配选项:

当执行 NFS 安装时，PowerHA SystemMirror 使用的缺省选项是 hard 和 intr。

要针对 NFS 安装设置软安装或任何其他选项:

1. 输入 `smit mknfsmnt`。
2. 在 **MOUNT now, add entry to /etc/filesystems or both?** 字段中，选择 **file systems** 选项。
3. 在 **/etc/filesystems entry will mount the directory on system RESTART** 字段中，接受缺省值 **no**。

该过程将向已创建的 **/etc/filesystems** 条目添加选择的选项。然后，PowerHA SystemMirror 脚本将读取此条目以使用您可能已选择的任何选项。

在客户机上创建和配置 NFS 安装点:

要使用 NFS 安装文件系统, 需要 NFS 安装点。在非并发资源组中, 该资源组中的所有节点都安装 NFS 文件系统。您应在该资源组的每个节点上都创建一个 NFS 安装点。NFS 安装点必须在本地安装点的目录树之外。

在资源组的所有节点上都创建 NFS 安装点后, 请为该资源组配置 **NFS Filesystem to NFS Mount** 属性。

要创建 NFS 安装点并为 NFS 装配配置资源组:

1. 在资源组的每个节点上, 通过执行以下命令创建 NFS 安装点:

```
mkdir /mount point
```

其中, *mount point* 是本地 NFS 安装点的名称, 通过该安装点装配远程文件系统。

2. 在 **Change/Show Resources and Attributes for a Resource Group** SMIT 面板中, **Filesystem to NFS Mount** 字段必须指定两个安装点。

请指定 NFS 安装点和本地安装点, 并用分号分隔。例如:

```
/nfspoint;/localpoint
```

如果有多个条目, 请以空格分隔:

```
/nfspoint1;/local1 /nfspoint2;/local2
```

3. 可选: 如果有嵌套的安装点, 请以与本地安装点相同的方式嵌套 NFS 安装点, 以使它们能够正确匹配。
4. 可选: 当交叉装配 NFS 文件系统时, 请为资源组将 SMIT 中的 **Filesystems Mounted before IP Configured** 字段设置为 **true**。

规划资源组

以下主题描述如何规划 PowerHA SystemMirror 集群内的资源组。

资源组概述

PowerHA SystemMirror 将资源组织为资源组。每个资源组均作为包括共享资源 (例如, IP 标签、应用程序、文件系统和卷组) 的单个单元处理。您可为每个资源组定义策略, 这些策略定义获取或释放该资源组的时间和方式。

在“初始集群规划”中, 您初步为资源组节点列表中的每个节点选择了资源组策略和接管优先级。在本节中, 您将执行以下各项:

- 确定组成每个资源组的各个资源。
- 对于每个资源组, 确定其组类型: 并发还是非并发。
- 为资源组定义参与节点列表。节点列表由分配为参与接管给定资源组的节点组成。
- 确定资源组启动、故障转移和回退策略。
- 确定您要为其设置位置依赖性和/或父子依赖性的应用程序及其资源组。
- 确定资源组的站点间管理策略。是否有需要考虑的复制资源?
- 确定用于优化资源组行为的其他属性和运行时策略。

本节中使用了以下术语:

- **参与节点列表**。可以托管特定资源组的节点的列表, 如在 SMIT 中针对资源组的参与节点名的定义。

请注意, 不同资源组策略和当前集群条件的组合也会影响资源组在集群中节点上的位置。

- 主节点（或此资源组的最高优先级节点）。在任何非并发资源组的参与节点列表中列示的第一个节点。

PowerHA SystemMirror 资源组支持 NFS 文件系统。

相关参考:

第 5 页的『初始集群规划』

本部分描述为规划 PowerHA SystemMirror 集群以使应用程序具有高可用性所应采取的初始步骤。

第 48 页的『NFS 交叉安装和 IP 标签』

为启用 NFS 交叉安装，每个集群节点都可以作为 NFS 客户机。其中每个节点都必须具有到 NFS 服务器节点的服务 IP 标签的有效路径。即，要启用 NFS 交叉安装，客户机节点上必须存在 IP 标签，并且必须在与 NFS 服务器节点的服务 IP 标签相同的子网上配置此 IP 标签。

资源和资源组的常规规则

有一些针对资源和资源组的常规规则和限制。

以下规则和限制适用于资源和资源组:

- 要使 PowerHA SystemMirror 可保持集群资源的高可用性，它必须属于某个资源组。如果您希望将某个资源保持分开，请单独为该资源定义一个组。可以为资源组定义一个或多个资源。
- 不能将某个资源包括在多个资源组中。
- 资源组的组成项必须是唯一的。请将应用程序及其需要的资源放置在同一资源组中。
- 服务 IP 标签、卷组和资源组名称也必须在集群内唯一，并且互不相同。资源的名称应该与其服务的应用程序有关，还应该与任何相应设备有关，例如，websphere_service_address。
- 如果您将同一节点包括在多个资源组的参与节点列表中，请确保该节点具有同时管理所有资源组所必需的内存和网络接口。

资源组的类型：并发和非并发

要对资源组行为分类并进行描述，您必须先将资源组分为两种类型：并发和非并发。

并发资源组

并发资源组可以在多个节点上联机。资源组节点列表中的所有节点都可以在它们加入集群时获取该资源组。在节点中没有优先级。可以将并发资源组配置为在集群中的所有节点上运行。

并发资源组中包括的资源仅限于具有原始逻辑卷的卷组、原始磁盘和使用磁盘的应用程序控制器。定义了这些逻辑存储器实体的设备必须支持并发访问。

并发资源组具有“Online on All Available Nodes”启动策略，并且不会从一个节点故障转移或回退到另一节点。

非并发资源组

非并发资源组无法在多个节点上联机。您可以为这些资源组定义多种启动、故障转移和回退策略。

您可以对节点首选项中非并发资源组在以下情况下的行为进行微调：节点启动、资源组在发生节点故障时故障转移到另一节点，或者资源组回退到重新集成节点。

相关参考:

第 53 页的『资源组的启动、故障转移和回退属性』

每个属性都会影响资源组的启动、资源组在发生节点故障时故障转移到另一节点，或者资源组回退到重新集成节点。

资源组的启动、故障转移和回退策略

资源组行为分为三种类型的节点策略。

这些策略是:

- **启动策略**定义当节点加入集群并且资源组在任何节点上均不处于活动状态时，资源组将在哪个节点上激活。
- **故障转移策略**定义当资源组由于故障状况（或者，如果您使用故障转移选项停止节点上的集群服务）必须离开其当前联机的节点时，资源组将故障转移到哪个节点上。
- **回退策略**定义当节点加入并且资源组已经在另一节点上处于活动状态时，资源组将回退到哪个节点上。

PowerHA SystemMirror 仅允许您配置资源组启动、故障转移和回退行为的有效组合。下表汇总了您可在 PowerHA SystemMirror 中为资源组配置的基本启动、故障转移和回退行为。

启动行为	故障转移行为	回退行为
仅在主节点（节点列表中的第一个节点）上联机	任何以下各项: <ul style="list-style-type: none">• 故障转移到列表中的下一个优先级节点• 使用动态节点优先级的故障转移	任何以下各项: <ul style="list-style-type: none">• 从不回退• 回退到列表中的较高优先级节点
使用节点分配策略联机	任何以下各项: <ul style="list-style-type: none">• 故障转移到列表中的下一个优先级节点• 使用动态节点优先级的故障转移	从不回退
在第一个可用节点上联机	任何以下各项: <ul style="list-style-type: none">• 故障转移到列表中的下一个优先级节点• 使用动态节点优先级的故障转移	任何以下各项: <ul style="list-style-type: none">• 从不回退• 回退到列表中的较高优先级节点
在所有可用节点上联机	脱机（仅在错误节点上）	从不回退

除上表中描述的节点策略外，其他问题也将决定节点获取的资源组。

相关参考:

第 70 页的『处理集群事件』

以下主题描述 PowerHA SystemMirror 集群事件。

资源组属性

本节概述了您可用于对资源组的启动、故障转移和回退策略进行微调的资源组属性。

资源组的启动、故障转移和回退属性

每个属性都会影响资源组的启动、资源组在发生节点故障时故障转移到另一节点，或者资源组回退到重新集成节点。

下表汇总了受给定属性或运行时策略影响的资源组启动、故障转移或回退策略。并非所有可用资源组都在下表中列出。

属性	启动策略	故障转移策略	回退策略
稳定时间	X		
节点分配策略	X		
动态节点优先级		X	
延迟回退计时器			X
资源组父子依赖性	X	X	X
资源组位置依赖性	X	X	X

相关参考:

第 57 页的『父子依赖资源组』

不同资源组中的相关应用程序将配置为以逻辑顺序进行处理。

第 57 页的『资源组位置依赖性』

不同资源组中的某些应用程序在一个节点上一起保持联机状态，或在不同节点上保持联机状态。

相关信息:

PowerHA SystemMirror 资源和资源组

启动的稳定时间

通过为当前脱机且具有“Online on First Available Node”启动策略的资源组指定稳定时间，您可以修改资源组的启动行为。

指定稳定时间后，您可以避免当多个节点同时启动集群服务时资源组在第一个可用节点上激活。资源组的较高优先级节点也可能在此时间段内加入集群。

稳定时间使集群管理器在激活资源组之前等待指定的时间量。使用此属性可确保在资源组优先级不断增加的节点陆续联机期间，资源组不会在节点中反弹。

如果启动的节点是此资源组节点列表中的第一个节点，那么将跳过稳定时间段并且 PowerHA SystemMirror 会立即尝试在此节点上获取资源组。

稳定时间具有以下特性:

- 仅影响当前脱机且您已将其启动策略指定为“Online on First Available Node”的那些资源组。您为所有此类资源组配置一个稳定时间。
- 当可获取资源组的第一个节点加入集群时激活，除非该节点是节点列表中的第一个节点（那么将忽略稳定时间并获取资源组）。

如果加入集群并且可能能够获取集群的第一个节点发生故障，那么此状况会取消稳定时间或重置稳定时间。

- 在 **node_up** 事件中延迟激活组，以便较高优先级的节点加入集群。

注: 如果为资源组指定了稳定时间段并且资源组当前处于 **ERROR** 状态，那么集群管理器在 **node_up** 事件期间尝试将该资源组联机之前，会等待该稳定时间段。

配置了稳定时间的节点重新集成

通常，当节点加入集群时，它可以获取资源组。以下列表描述了稳定时间在此过程中的作用:

- 如果节点是特定资源组的最高优先级节点，那么节点将立即获取该资源组且将忽略稳定时间。只有在此情况下，PowerHA SystemMirror 才会忽略该设置。
- 如果节点能够获取一些资源组，但不是这些组的最高优先级节点，那么不会在该节点上获取资源组。相反，它们会在稳定时间间隔期间等待，看是否有较高优先级节点加入集群。

当稳定时间间隔结束时，PowerHA SystemMirror 会将资源组移动到当前可用且可以取得资源组的最高优先级节点。如果 PowerHA SystemMirror 未找到适当的节点，那么资源组将保持脱机。

节点分配策略

您可以将资源组的启动行为配置为在启动期间使用节点分配策略。此策略可确保在启动期间一个节点上仅获取到一个启用了此策略的资源组。

您可以将节点分配策略用于集群启动，以确保 PowerHA SystemMirror 在每个节点上仅激活一个启用了此策略的资源组。此策略可帮助您将 CPU 密集型应用程序分配在不同节点上。

有关节点分配策略的实际情况如下：

- 如果您计划使用将配置“通过替换的 IPAT”的单适配器网络，那么资源组的启动策略应设置为“Online using Distribution Policy”。
- 如果在特定节点连接时，有两个启用了此策略的资源组脱机，那么在节点上将只获取这两个资源组的其中一个。PowerHA SystemMirror 将首选在节点列表中具有较少节点的资源组，然后按字母顺序对资源组列表排序。
- 如果其中一个资源组是父资源组（具有子资源组），那么 PowerHA SystemMirror 将首选父资源组，并将其在节点上激活。
- 要确保资源组不仅在启动时分配，还会针对恢复事件（故障转移和回退）分配，请使用位置依赖性。

相关参考：

第 56 页的『资源组依赖性』

PowerHA SystemMirror 提供了多种配置，您可在这些配置中指定希望在启动、故障转移和回退时维护的资源组之间的关系。

动态节点优先级策略

您可以将资源组的故障转移行为配置为使用动态节点优先级。这样，您可以使用预定义的 Reliable Scalable Cluster Technology 资源变量（例如 **lowest CPU load**）来选择接管节点，或者使用用户定义的动态节点优先级变量（例如 **cl_highest_udscript_rc**）来选择接管节点。

通过设置动态节点优先级策略，您可使用预定义的 Resource Monitoring and Control (RMC) 资源变量（例如 **lowest CPU load**）来选择接管节点。通过启用动态优先级策略，接管节点列表的顺序将由发生事件时集群的状态确定（由所选 RMC 资源变量度量）。可为不同组设置不同策略，或者为多个组设置相同策略。

另一选项是使用用户定义的动态节点优先级变量，例如 **cl_highest_udscript_rc**。如果您使用此选项，那么必须提供一个脚本和执行超时值，当发生事件时将在所有候选故障转移节点上调用此脚本。然后从每个候选节点中收集返回值，并根据脚本的返回值和所选动态节点优先级变量选择接管节点。

如果您决定使用 RMC 资源变量来定义动态节点优先级策略以确定资源组的故障转移节点，那么请考虑以下几点：

- 动态节点优先级策略在所有节点均具有相同处理能力和内存的集群中最有用。
- 动态节点优先级策略不适合少于三个节点的集群。
- 动态节点优先级策略不适合并行资源组。

- 配置站点时，动态节点优先级策略不受支持。

请记住，选择接管节点还取决于诸如该节点上网络接口的可用性等条件。

延迟回退计时器

通过指定和分配延迟回退计时器，您可以配置资源组的回退行为，以使其按一种预定义的重现时间（每天、每周、每月和每年，或者在特定日期和时间）发生。

您可以使用延迟回退计时器来设置资源组回退至较高优先级节点的时间。您可以配置资源组的回退行为，以使其在预定义的重现时间（每天、每周、每月或特定日期）发生。

延迟回退计时器具有以下特性：

- 指定位于非主节点或低优先级节点的联机资源组回退到其主节点或较高优先级节点的时间。
- 影响资源组移动到另一个节点。例如，如果您通过使用资源组管理 (RGM) 实用程序 (**cIRGmove**) 将非并发资源组（具有回退计时器属性）移动到目标节点，那么该组将位于目标节点上（除非您重新引导集群，但很少会执行此操作）。如果目标节点停机并在稍后重新集成，那么资源组也会在指定时间回退到此节点。

设置了延迟回退计时器的节点重新集成

在以下情况下，资源组不会立即回退到其较高优先级节点。

- 您已为资源组配置了延迟回退计时器。
- 较高优先级节点加入集群。

在“延迟回退计时器”属性指定的时间，将发生以下两种情况之一：

- **找到较高优先级节点。** 如果较高优先级节点对于资源组可用，那么在回退计时器到期时，PowerHA SystemMirror 会尝试将资源组移动到该节点。如果获取操作成功，那么资源组将获取到该节点上。

但是，如果将资源组获取到该节点的操作失败，那么 PowerHA SystemMirror 会尝试将资源组移动到组节点列表中的次较高优先级节点，以此类推。如果将资源组获取到最后一个可用节点的操作失败，那么该资源组将转入 **ERROR** 状态。您必须采取操作来修复错误，并将此类资源组恢复联机。

- **未找到较高优先级节点。** 如果对于资源组没有较高优先级的可用节点，那么该资源组将在同一节点上保持联机，直到回退计时器再次到期为止。例如，如果每日回退计时器在晚上 11:00 到期，且没有更高优先级节点可用于要回退的资源组，那么回退计时器会在第二天晚上 11:00 再次到期。

设置为特定日期的回退计时器不会再次到期。

资源组依赖性

PowerHA SystemMirror 提供了多种配置，您可在这些配置中指定希望在启动、故障转移和回退时维护的资源组之间的关系。

您可以配置以下资源组依赖性：

- **父子依赖资源组。** 不同资源组中的相关应用程序和其他资源将配置为以正确的顺序进行处理。
- **“后启动”依赖性。** 指定仅当一个或多个其他资源组在集群中处于活动状态后，某个资源组才能启动。
- **“后停止”依赖性。** 指定仅当将一个或多个其他资源组脱机后，某个资源组才能停止。
- **资源组位置依赖性。** 不同资源组中的某些应用程序在一个节点上一起保持联机状态，或在不同节点上保持联机状态。

当规划如何配置这些依赖性时，请记住以下几点：

- 虽然在缺省情况下所有资源组都以并行方式处理，但是 PowerHA SystemMirror 会根据依赖性指示的顺序处理依赖资源组，而不一定以并行方式处理。资源组依赖性在集群范围内受认可，并会覆盖为依赖性中包括的任何资源组定制的串行处理顺序。
- 资源组之间的依赖性提供了一种可预测且可靠的方法来构建具有多层应用的集群。

以下限制适用于合并了依赖性的配置。如果您具有不止一个同时属于“Online on Same Node”依赖性集合和“Online On Different Nodes”依赖性集合的资源组，那么验证将失败。

父子依赖资源组:

不同资源组中的相关应用程序将配置为以逻辑顺序进行处理。

通过配置资源组依赖性，可以更好地控制具有多层应用程序的集群，在此类集群中一个应用程序依赖于另一个应用程序的成功启动，并且两个应用程序都需要使用 PowerHA SystemMirror 来保持高可用性。

以下示例说明了父子依赖行为:

- 如果资源组 A 依赖资源组 B，那么在集群的任何节点获取到资源组 A 之前，资源组 B 必须处于联机状态。请注意，资源组 A 定义为子资源组，资源组 B 定义为父资源组。
- 如果子资源组 A 依赖父资源组 B，那么在节点启动或节点重新集成期间，在父资源组 B 联机之前，子资源组 A 无法联机。如果将父资源组 B 脱机，那么子资源 A 将先脱机，因为它依赖父资源组 B。

使用多层应用程序的业务配置可以利用父子依赖资源组。例如，数据库必须在应用程序控制器之前联机。在此情况下，如果数据库移到其他节点，那么必须先将包含应用程序控制器的资源组脱机，然后将其在集群的任一节点上恢复联机。

如果子资源组包含的某个应用程序依赖父资源组中的资源，并且父资源组故障转移至其他节点，那么子资源组将临时停止并自动重新启动。与此类似，如果子资源组处于并发状态，那么 PowerHA SystemMirror 会将其在所有节点上临时脱机，然后将其在所有可用节点上恢复联机。如果父资源组的故障转移未成功，那么父资源组和子资源组都将转入 ERROR 状态。

在规划父子依赖资源组时，请考虑以下事项:

- 规划您需要保持高可用性的应用程序，请考虑在某个应用程序可以启动之前，您的业务环境是否要求另一个应用程序先运行。
- 确保对顺序有要求的这些应用程序包括在不同资源组中。这样，您可以在这些资源组之间建立依赖性。
- 请为您计划要包括在子资源组或父资源组中的每个应用程序规划应用程序监视器。对于父资源组中的应用程序，请以**监视启动**方式配置监视器。

为最大程度减小在应用程序停止和重新启动进程期间丢失数据的可能性，请定制应用程序控制器脚本，以确保在应用程序停止进程期间所有未落实的数据都将临时存储到共享磁盘，并将在应用程序重新启动进程期间读回到应用程序。使用共享磁盘是很重要的，原因是应用程序启动时所在的节点与其停止时所在的节点可能不同。

相关参考:

第 13 页的『多层应用程序的规划注意事项』

使用多层应用程序的业务配置可以使用父子依赖性资源组。例如，数据库必须在应用程序控制器之前联机。在此情况下，如果数据库关闭并且移动到其他节点，那么必须先将包含应用程序控制器的资源组脱机，再将其在集群的任何节点上联机。

资源组位置依赖性:

不同资源组中的某些应用程序在一个节点上一起保持联机状态，或在不同节点上保持联机状态。

如果在一定时间确实发生了故障，那么 PowerHA SystemMirror 会分配资源组，以便它们保持可用，但不一定位于您最初指定的节点上，除非它们具有同一主节点以及相同的故障转移和回退策略。

资源组位置依赖性为您提供了一种直观的方法，用于指定某些资源组必须始终在同一节点上联机，或者某些资源组必须始终在不同节点上联机。您可以将这两个位置策略与父子依赖性、“开始后”依赖性和“后停止”依赖性相组合，从而使所有子资源组或源资源组在同一节点上联机，而父资源组或目标资源组在另一节点上联机。您还可以使所有子资源组或源资源组在不同节点上联机，以取得更好的性能。

注：站点仅在 PowerHA SystemMirror Enterprise Edition 和 Standard Edition 7.1.2 或更高版本中受支持。复制管理仅在 PowerHA SystemMirror Enterprise Edition 中受支持。

如果您具有复制资源，那么可以将资源组合并为一个站点依赖性来使这些组在同一站点上保持联机状态。

PowerHA SystemMirror 在资源组之间支持以下类型的资源组位置依赖性：

- **在同一节点上联机**

以下规则和限制适用于资源组的“在同一节点上联机”依赖性集合。如果您不遵循以下准则，验证将失败：

- 配置为包括在给定“同一节点”依赖性集合中的所有资源组都必须具有相同的节点列表（相同顺序的相同节点）。
- “同一节点”依赖性集合中的所有非并发资源组都必须具有相同的启动、故障转移和回退策略。
- 不允许将“Online Using Node Distribution Policy”用于启动。
- 如果将“Dynamic Node Priority Policy”配置为故障转移策略，那么集合中的所有资源组都必须具有相同策略。
- 如果为一个资源组配置了回退计时器，那么该计时器适用于资源组集。集合中的所有资源组都必须具有相同的回退时间设置。
- 可以包括并发和非并发资源组。
- 您可以在集群中具有一个以上的“同一节点”依赖性集合。
- PowerHA SystemMirror 施行以下条件：在“同一节点”依赖性集合中处于活动状态 (ONLINE) 的所有资源组需要位于同一节点上。该集合中的一些资源组可以脱机或处于 ERROR 状态。
- 如果“同一节点”依赖性集合中的一个或多个资源组故障，那么 PowerHA SystemMirror 会尝试将集合中的所有资源组放在可以托管所有资源组（当前处于 ONLINE 状态，即仍处于活动状态的资源组，加上一个或多个发生故障的资源组）的节点上。

- **在同一站点上联机**

以下规则和限制适用于资源组的“在同一站点上联机”依赖性集合。如果您不遵循以下准则，验证将失败：

- “同一站点”依赖性集合中的所有资源组都必须具有相同的站点间管理策略，但可以具有不同的启动策略、故障转移策略和回退策略。如果使用回退计时器，那么对于该集合中的所有资源组，这些计时器必须完全相同。
- 回退计时器不适用于跨站点边界移动资源组。
- 必须对“同一站点”依赖性集合中的所有资源组进行配置，以便将能够拥有资源组的节点分配至相同的主要站点和辅助站点。
- 可以包括并发和非并发资源组。
- 可以在集群中具有多个“同一站点”依赖性集合。
- “同一站点”依赖性集合中所有处于活动状态 (ONLINE) 的资源组都必须在同一站点上处于 ONLINE 状态，即使该集合中的某些资源组可能处于 OFFLINE 或 ERROR 状态也是如此。

- 如果将包含在“同一节点”依赖性集中的资源组添加到“同一站点”依赖性集合中，那么必须将该“同一节点”依赖性集中的所有其他资源组都添加到该“同一站点”依赖性集合中。

- **在不同节点上联机**

以下规则和限制适用于资源组的“在不同节点上联机”依赖性集合。如果您不遵循以下准则，验证将失败：

- 每个集群仅允许一个“在不同节点上联机”依赖性集合。
- 规划启动策略，以便集合中的每个资源组都将在不同节点上启动。
- 如果指定了父依赖性或子依赖性，那么子资源组不能具有高于其父资源组的优先级。

当配置有这些组的集群运行后，请注意：

- 如果高优先级资源组在节点上联机，那么“不同节点”依赖性集合中的任何其他较低优先级资源组都无法在该节点上联机。
- 具有更高优先级的资源组故障转移或回退到给定节点，那么具有较高优先级的资源组将联机，并且集群管理器会将较低优先级资源组脱机并将其移动到其它节点（如果可行）。
- 具有相同优先级的资源组无法在同一节点上联机（启动）。同一优先级级别内的节点的资源组优先级由组的字母顺序确定。
- 具有相同优先级的资源组不会导致对方在故障转移或回退之后从节点中移出。

相关参考：

第 64 页的『使用站点的特殊注意事项』

以下信息与当资源组已指定依赖性时的特殊注意事项相关。

“后启动”依赖性：

在“后启动”依赖性中，目标资源组必须在集群中的任一节点上联机，然后源（依赖项）资源组才能在节点上激活。当释放资源组并且以并行方式释放组时，没有任何依赖性。

以下是“后启动”依赖性的准则和限制。

- 资源组既可以充当目标资源组也可以充当源资源组，具体取决于资源组处于给定依赖性链接的哪一端。
- 可为资源组指定三级依赖性。
- 不能在资源组间指定循环依赖性。
- 此依赖性仅在资源组获取时适用。在资源组释放期间这些资源组之间没有任何依赖性。
- 在源资源组的目标资源组完全运行之前，无法在节点上获取源资源组。如果目标资源组未完全运行，那么源资源组将由于目标的脱机状态而变为脱机。如果您发现某个资源组处于此状态，那么您可能需要诊断哪些资源可能需要手动联机以解决资源组依赖性。
- 当充当目标角色的资源组从一个节点故障转移到另一个节点时，对依赖于该资源组的其他资源组没有任何效果。
- 源资源组联机后，对目标资源组的任何操作（将资源组脱机，或移动资源组）都不会影响源资源组。
- 如果目标资源组脱机，那么无法在源资源组上手动移动资源组或将资源组联机。

注：您应配置多个应用程序监视器，特别应为目标资源组中包括的应用程序配置一个检查应用程序启动的监视器。此过程将验证目标资源组中的应用程序是否成功启动。

“后停止”依赖性：

在“后停止”依赖性中，目标资源组必须在集群中的任一节点上脱机，然后源（依赖项）资源组才能在节点上脱机。当获取资源组并且以并行方式获取组时，没有任何依赖性。

以下是“后停止”依赖性的限制和准则。

- 资源组既可以充当目标资源组也可以充当源资源组，具体取决于资源组处于给定依赖性链接的哪一端。
- 可为资源组指定三级依赖性。
- 不能在资源组间指定循环依赖性。
- 此依赖性仅在资源组释放时适用。在资源组获取期间这些资源组之间没有任何依赖性。
- 在源资源组的目标资源组脱机之前，无法在节点上释放源资源组。
- 充当源角色的资源组从一个节点故障转移到另一个节点时，将先释放目标资源组，然后释放源资源组。此后，将以并行方式获取这两个资源组（假定这些资源组之间没有“后启动”依赖性或父子依赖性）。
- 如果目标资源组脱机，那么无法在源资源组上手动移动资源组或将资源组脱机。

将资源组移动到其他节点

当 PowerHA SystemMirror 将资源组移动到其他节点时，您有一些选项。

这些选项包括：

- 资源组保留在它们所移动到的节点上。

您可以将具有 Never Fallback 策略的资源组移动到其他节点。当执行此操作时，您可以告知 PowerHA SystemMirror 将资源组留在目标节点上，直到您决定再次移动该组为止。

- 当您使用 **RG_move** 命令移动资源组时，资源组将无限期地保留在它所移动到的节点上（直到您告知 PowerHA SystemMirror 将其移动到其他节点）或一直保留到您重新引导集群为止。

如果您不停止集群服务（很少必须这样做），并且您希望永久性更改资源组的节点列表和最高优先级节点，请更改资源组的属性并重新启动集群。

- 如果您在任何节点上将资源组联机或脱机，那么在该集群重新引导之前或您手动将该组在集群的其他位置上联机之前，该组将一直保持联机或脱机。
- 如果资源组具有回退至最高优先级节点的策略，那么在移动该资源组后，该组会回退至其目标节点。

例如，如果组配置了节点 A 作为其最高优先级节点，并且您将该组移动到节点 B，那么此组将保留在节点 B 上，并在此时将此节点视为其最高优先级节点。您可随时选择将该组再次移动到节点 A 上。当您使用 SMIT 执行此操作时，PowerHA SystemMirror 将通知您原始最高优先级节点（节点 A）现在是否可托管该组。

您可通过使用 **clRGinfo -p** 命令跟踪手动移动的所有资源组。

使用 **clRGmove** 命令移动资源组

您可以使用 **clRGmove** 命令将资源组移动到其他节点，或者将资源组联机或脱机。您可以通过使用 SMIT 或从命令行运行 **clRGmove** 命令。

如果您对具有“永不回退”回退策略的资源组使用 **clRGmove**，那么资源组将保留在该节点上，直到您将其移动到别处。

以下段落描述使用 **clRGmove** 来管理具有不同策略的资源组时所应用的规则。

移动父子依赖资源组

- 如果父资源组由于您通过 **clRGmove** 命令发出的请求而脱机，那么 PowerHA SystemMirror 将拒绝手动尝试将依赖于这些资源组的子资源组联机。错误消息列出了您必须首先将其联机的父资源组。
- 如果您具有联机的父资源组和子资源组，并且希望将父资源组移动到另一个节点或将其脱机，那么将子资源组脱机之前，PowerHA SystemMirror 将阻止您执行此操作。

移动“后启动”依赖资源组

在此类型的依赖性中，目标资源组必须在集群中的任何节点上联机，然后源（依赖项）资源组才能在节点上激活。以下规则适用于具有“后启动”依赖性的资源组：

- 如果目标资源组由于您通过 **cIRGmove** 命令发出的请求而脱机，那么 PowerHA SystemMirror 将拒绝手动尝试将依赖于这些资源组的源资源组联机。错误消息列出了您必须首先将其联机的目标资源组。

移动“后停止”依赖资源组

在此类型的依赖性中，目标资源组必须在集群中的任何节点上脱机，然后源（依赖项）资源组才能在节点上脱机。以下规则适用于具有“后停止”依赖性的资源组：

- 如果您具有联机的目标资源组和源资源组，并且希望将源资源组移动到另一个节点或将其脱机，那么将目标资源组脱机之前，PowerHA SystemMirror 将阻止您执行此操作。

移动位置依赖资源组

- 如果您将某个同站点依赖资源组移动到另一个站点，那么该依赖关系中的整个资源组集都将移动到该另一个站点。
- 如果您将某个同节点依赖资源组移动到其他节点，那么该依赖关系中的整个资源组集都将移动。
- 您无法将资源组移动到主管属于不同节点依赖性的联机资源组的任何节点。您必须先在所选节点上将不同节点依赖性中包括的资源组脱机。

规划集群网络和资源组

大部分非并发资源组使用服务 IP 标签通过网络向客户机提供对应用程序的访问权。PowerHA SystemMirror 使用“通过 IP 别名的 IP 地址接管 (IPAT)”来保持这些服务地址在集群中可用。

IPAT 不适用于并发资源组或者“在所有可用节点上联机”的资源组。

请考虑当环境中配置了防火墙或 VPN 时，客户机如何通过集群网络到达应用程序地址，这一点很重要。

相关参考：

第 22 页的『规划通过 IP 别名的 IP 地址接管』

通过向 NIC 分配 IP 别名，您可以在同一网络接口上创建多个 IP 标签。

规划资源组的并行或串行处理顺序

缺省情况下，PowerHA SystemMirror 将以并行方式来获取和释放集群中配置的所有个体资源。但是，您可以指定特定串行顺序，将根据该顺序获取或释放某些或全部个体资源组。

在获取期间完成以下过程：

1. PowerHA SystemMirror 按您在列表中指定顺序，以串行方式获取资源组。
2. PowerHA SystemMirror 以并行方式获取其余资源组。

在释放资源组期间，过程相反：

1. PowerHA SystemMirror 以并行方式释放您未定义特定串行顺序的资源组。
2. 集群中的其余资源组按您在列表中为这些资源组指定的顺序进行处理。
3. 如果您将集群从 PowerHA SystemMirror 先前版本中升级，请参阅“升级 PowerHA SystemMirror 集群”以获取有关在此情况下使用的处理顺序的更多信息。

注：即使您针对单个节点指定了资源组处理顺序，实际资源组故障转移也可能由不同策略触发。因此，不能保证资源组在集群范围内按指定顺序处理，原因是资源组的串行定制处理顺序仅适用于它们在特定节点上的处理。

4. 以并行方式处理资源组时，集群中发生的集群事件较少。如果资源组以并行方式处理，尤其不会发生例如 **node_up_local** 或 **get_disk_vg_fs** 等事件。
5. 因此，使用并行处理可减少特定集群事件（您可为这类事件创建定制的前置事件或后置事件脚本）的数量。如果您开始对配置中的某些资源组使用并行处理，请注意，您的现有前置事件或后置事件脚本可能对这些资源组不起作用。
6. 在 **hacmp.out** 文件的事件摘要中反映了资源组的并行和串行处理。

有关如何配置资源组的定制顺序获取和释放顺序，请参阅“为资源组配置处理顺序”。

依赖资源组以及并行或串行顺序

虽然，PowerHA SystemMirror 在缺省情况下以并行方式处理资源组，但是如果您在集群的某些资源组之间建立依赖性，那么花费的处理时间可能比处理没有依赖资源组的集群更长，原因是需要执行更多处理来处理一个或多个 **rg_move** 事件

在获取时，将先获取父资源组或优先级较高的资源组，然后获取子资源组。在释放时，顺序相反。集群中的其余资源组（本身没有依赖性的资源组）将以并行方式处理。

此外，如果您指定了串行处理顺序并且配置了依赖资源组，请确保串行顺序不会与指定的依赖性抵触。资源组依赖性将覆盖集群中的任何串行顺序。

相关参考：

第 70 页的『处理集群事件』

以下主题描述 PowerHA SystemMirror 集群事件。

相关信息：

配置资源组的处理顺序

升级 PowerHA SystemMirror 集群

规划具有站点的集群中的资源组

站点间管理策略与您选择的节点启动策略、故障转移策略和回退策略的组合确定了资源组启动、故障转移和回退行为。

注：站点仅在 PowerHA SystemMirror Enterprise Edition 和 Standard Edition 7.1.2 或更高版本中受支持。复制管理仅在 PowerHA SystemMirror Enterprise Edition 中受支持。

PowerHA SystemMirror 中的站点支持允许多种资源组配置。

并发资源组和站点

可以为并发资源组使用以下策略：

策略类型	描述
站点间管理策略	在两个站点上都联机 在任何一個站点上联机 首选主要站点 忽略
启动策略	在所有可用节点上联机
故障转移策略	脱机（仅在错误节点上）
回退策略	永不回退

非并发资源组和站点

对于非并发资源组，可以使用以下策略：

策略类型	描述
站点间管理策略	在任何一個站点上联机 首选主要站点 忽略
启动策略	在主节点上联机 在第一个可用节点上联机 使用节点分配策略联机
故障转移策略	故障转移至节点列表中的下一个优先级节点
回退策略	回退至节点列表中的较高优先级节点 永不回退

包含站点的集群中的常规资源组行为

当集群在运行时，使用“首选主要站点”或“在任何一個站点上联机”的站点间管理策略定义的非并发资源组具有两个实例。

注： 站点仅在 PowerHA SystemMirror Enterprise Edition 和 Standard Edition 7.1.2 或更高版本中受支持。复制管理仅在 PowerHA SystemMirror Enterprise Edition 中受支持。

以下实例在非并发资源组中运行：

- 位于主要站点的节点上的主要实例
- 位于辅助站点的节点上的次要实例

cIRGinfo 命令将这些实例显示为：

- ONLINE
- ONLINE SECONDARY

当集群在两个位置上运行时，站点间管理策略为“在两个站点上都联机”的并发资源组（在所有节点上联机）具有多个 ONLINE 实例，并且没有 ONLINE SECONDARY 实例。

站点间管理策略为“首选主要站点”或“在任何一個站点上联机”的并发资源组在主要站点的每个节点上具有主要实例，并在辅助站点的节点上具有次要实例。

发生 **node_up** 事件时，将根据相应站点管理策略和节点启动策略对具有复制资源的资源组进行并行处理，并且会将任何依赖性考虑在内。当资源组在站点间发生故障转移时，将在新备份站点上可用的最高优先级节点上获取次要实例，并将这些实例的状态变为 **ONLINE SECONDARY**。同一节点上可以有多个次要实例。将在新活动站点上可用于托管资源组的最高优先级节点上获取该资源组的主要实例，并将实例的状态变为 **ONLINE**。此处理顺序将确保当主要实例启动时，备份站点已准备好接收备份数据。

如果无法将次要实例变为 **ONLINE SECONDARY** 状态，那么仍将在可能的情况下将主要实例变为 **ONLINE**。

使用站点的特殊注意事项

以下信息与当资源组已指定依赖性时的特殊注意事项相关。

依赖资源组和站点

注： 站点仅在 PowerHA SystemMirror Enterprise Edition 和 Standard Edition 7.1.2 或更高版本中受支持。复制管理仅在 PowerHA SystemMirror Enterprise Edition 中受支持。

对于位于不同站点上的节点，您可以指定这些节点上存在的两个或多个资源组之间的依赖性。在这种情况下，如果父组或子组移至另一个站点，那么依赖组也将移动。如果无法在故障转移站点上激活父组，那么子资源组也将保持不活动状态。

依赖性仅适用于资源组的主要实例的状态。如果父组的主要实例为 **OFFLINE**，而次要实例在某个节点上为 **ONLINE SECONDARY**，那么子组的主要实例将为 **OFFLINE**。

在资源组恢复期间，资源组可以故障转移至任何一个站点上的节点。获取依赖资源组的顺序与没有站点的集群的获取顺序相同，即首先获取父资源组，然后获取子资源组。释放逻辑相反，即先释放子资源组，然后释放父资源组。

如果您具有在没有站点的集群中定义的站点，那么必须为具有依赖性的资源组中所包括的应用程序配置应用程序监视。

相关参考：

第 57 页的『资源组位置依赖性』

不同资源组中的某些应用程序在一个节点上一起保持联机状态，或在不同节点上保持联机状态。

示例：具有站点的集群中的资源组行为

回退策略适用于资源组的 **ONLINE** 和 **ONLINE SECONDARY** 实例。

注： 站点仅在 PowerHA SystemMirror Enterprise Edition 和 Standard Edition 7.1.2 或更高版本中受支持。复制管理仅在 PowerHA SystemMirror Enterprise Edition 中受支持。

资源组的站点间管理策略确定了站点间资源组的 **ONLINE** 实例的回退行为，该行为控制次要实例的位置。

ONLINE SECONDARY 实例位于没有 **ONLINE** 实例的站点上。下表根据启动和站点间管理策略显示站点事件期间资源组的预期行为。

注： 下表中的大多数限制仅在每个节点的多个接口都位于同一网络上时适用。当每个节点只有一个接口位于网络上时（使用光纤通道或虚拟以太网时这是常见情况），大多数这些限制都不适用。

节点启动策略（在站点中适用）	站点间管理策略	启动、故障转移或回退行为
<p>仅在主节点上联机</p>	<p>首选主要站点</p>	<p>集群启动</p> <p>主要站点 主节点获取处于 ONLINE 状态的资源组。非主节点保留该资源组。</p> <p>辅助站点 连接此站点的第一个节点获取处于 ONLINE SECONDARY 状态的资源组。</p> <p>站点间故障转移 当本地站点上的任何节点都无法获取资源组时，ONLINE 实例将在站点间转移。次要实例将移至另一个站点，并将在可用的最高优先级节点上变为 ONLINE SECONDARY 状态（如有可能）。</p> <p>站点间回退 当来自主要站点的节点连接时，ONLINE 实例回退至该主要站点。次要实例将移至另一个站点，并将在可用的最高优先级节点上变为 ONLINE SECONDARY 状态（如有可能）。</p>
<p>在第一个可用节点上联机</p> <p>或者</p> <p>使用节点分发策略联机</p>	<p>首选主要站点</p>	<p>集群启动</p> <p>主要站点 首先从主要站点连接并满足条件的节点获取处于 ONLINE 状态的资源组。该资源组在主要站点上的所有其他节点上为 OFFLINE。节点分发策略仅适用于资源组的主要实例。</p> <p>辅助站点 加入此站点中集群的第一个节点获取处于 ONLINE_SECONDARY 状态并具有此启动策略的资源组的所有次要实例（无分发）。</p> <p>站点间故障转移 当本地站点上的任何节点都无法获取资源组时，ONLINE 实例将在站点间转移。次要实例将移至另一个站点，并将在可用的最高优先级节点上变为 ONLINE SECONDARY 状态（如有可能）。</p> <p>站点间回退 当来自主要站点的节点连接时，ONLINE 实例回退至该主要站点。次要实例将移至另一个站点，并将在可用的最高优先级节点上变为 ONLINE SECONDARY 状态（如有可能）。</p>
<p>在所有可用节点上联机</p>	<p>首选主要站点</p>	<p>集群启动</p> <p>主要站点 所有节点获取处于 ONLINE 状态的资源组。</p> <p>辅助站点 所有节点获取处于 ONLINE_SECONDARY 状态的资源组。</p> <p>站点间故障转移 当本地站点上的所有节点都变为 OFFLINE 或无法启动资源组时，ONLINE 实例将在站点间转移。次要实例将移至另一个站点，并将在可能的位置变为 ONLINE SECONDARY 状态。</p> <p>站点间回退 当主要站点上的节点重新连接时，ONLINE 实例回退至该主要站点。辅助站点上的节点获取处于 ONLINE_SECONDARY 状态的资源组。</p>

节点启动策略（在站点中适用）	站点间管理策略	启动、故障转移或回退行为
<p>仅在主节点上联机</p>	<p>在任何一个站点上联机</p>	<p>集群启动</p> <p>主要站点 （从任何一个站点）加入集群的主节点获取处于 ONLINE 状态的资源组。非主节点将该资源组保留为 OFFLINE。</p> <p>辅助站点 从另一个站点连接的第一个节点获取处于 ONLINE_SECONDARY 状态的资源组。</p> <p>站点间故障转移 当本地站点上的任何节点都无法获取资源组时，ONLINE 实例将在站点间转移。次要实例将移至另一个站点，并将在可用的最高优先级节点上变为 ONLINE_SECONDARY 状态（如有可能）。</p> <p>站点间回退 当主要站点上的节点重新连接时，ONLINE 实例不回退至该主要站点。最高优先级的重新连接的节点获取处于 ONLINE_SECONDARY 状态的资源组。</p>
<p>在第一个可用节点上联机</p> <p>或者</p> <p>使用节点分发策略联机</p>	<p>在任何一个站点上联机</p>	<p>集群启动</p> <p>主要站点 首先从任何一个站点连接并满足分发条件的节点获取处于 ONLINE 状态的资源组。</p> <p>辅助站点 资源组处于 ONLINE 后，从另一个站点连接的第一个节点获取处于 ONLINE_SECONDARY 状态的资源组。</p> <p>站点间故障转移 当本地站点上的任何节点都无法获取资源组时，ONLINE 实例将在站点间转移。</p> <p>站点间回退 当主要站点连接时，ONLINE 实例不回退至该主要站点。重新连接的节点获取处于 ONLINE_SECONDARY 状态的资源组。</p>
<p>在所有可用节点上联机</p>	<p>在任何一个站点上联机</p>	<p>集群启动</p> <p>主要站点 从任何一个站点连接的第一个节点获取处于 ONLINE 状态的资源组。当该组的实例处于活动状态后，同一站点中的其余节点也将激活处于 ONLINE 状态的组。</p> <p>辅助站点 所有节点获取处于 ONLINE_SECONDARY 状态的资源组。</p> <p>站点间故障转移 当本地站点上的所有节点都变为 OFFLINE 或无法启动资源组时，ONLINE 实例将在站点间转移。</p> <p>站点间回退 当主要站点连接时，ONLINE 实例不回退至该主要站点。重新连接的节点获取处于 ONLINE_SECONDARY 状态的资源组。</p>

节点启动策略（在站点中适用）	站点间管理策略	启动、故障转移或回退行为
在所有可用节点上联机	在两个站点上都联机	<p>集群启动 两个站点上的所有节点激活处于 ONLINE 状态的资源组。</p> <p>站点间故障转移 不发生故障转移。资源组处于 OFFLINE 状态或 ERROR 状态。</p> <p>站点间回退 不发生回退。</p>

定制站点间资源组恢复

可以将站点间资源组恢复配置为进行自动处理或仅进行故障通知。

资源组的特定实例可以在一个站点内进行故障转移，但它无法在站点间移动。如果受影响实例所在的站点上没有可用的节点，那么该实例将进入 ERROR 或 ERROR_SECONDARY 状态。它不会停留在发生故障的节点上。此行为适用于主要实例和次要实例。

如果发生 `node_down` 或 `node_up` 事件，那么即使已禁用站点间的故障转移，集群管理器也会移动资源组。您还可以在站点间手动移动资源组。

启用或禁用站点间的故障转移

如果是从 PowerHA SystemMirror 的前发行版迁移而来，那么可以更改资源组恢复策略，以允许集群管理器将资源组移动到其他站点，从而避免使该资源组进入 ERROR 状态。

在站点间恢复复制资源组的主要实例

启用站点间的故障转移后，在连接到站点间网络的接口发生故障或变为不可用的情况下，PowerHA SystemMirror 将尝试恢复资源组的主要实例。

在站点间恢复复制资源组的次要实例

启用站点间的故障转移后，PowerHA SystemMirror 将在以下情况下尝试恢复资源组的次要实例和主要实例：

- 如果在获取资源组的次要实例时发生获取失败，那么集群管理器将尝试恢复该资源组的次要实例，如其为主要实例所执行操作一样。如果没有节点可用于获取，那么该资源组的次要实例将进入全局 ERROR_SECONDARY 状态。
- 如果触发了配额丢失，并且资源组的次要实例在受影响节点上处于 ONLINE 状态，那么 PowerHA SystemMirror 将尝试在另一个可用的节点上恢复该次要实例。
- 如果所有 XD_data 网络都发生故障，那么 PowerHA SystemMirror 会将具有 GLVM 资源的所有 ONLINE 资源组移至该站点上的另一个可用节点。主要实例的此功能已镜像至次要实例，因此可以通过选择性故障转移来恢复次要实例。

使用 SMIT 来启用或禁用站点间资源组恢复

要启用或禁用站点间资源组恢复，请完成以下步骤：

1. 从命令行输入 `smit sysmirror`。
2. 在 SMIT 界面中，选择 **Custom Cluster Configurations > Resources > PowerHA SystemMirror Extended Resources Configuration > Customize Inter-site Resource Group Recovery**，然后按 Enter 键。

规划复制资源

PowerHA SystemMirror 提供复制资源的集成支持。

通过 PowerHA SystemMirror 复制资源，可以使用以下功能：

- 使您可以使用 PowerHA SystemMirror Enterprise Edition for AIX 和 PowerHA SystemMirror 站点配置来动态重新配置包含复制资源的资源组。
- 通过自动检测和调用已安装的 PowerHA SystemMirror Enterprise Edition for AIX 产品的验证实用程序，将 PowerHA SystemMirror Enterprise Edition for AIX 验证合并到标准集群验证中。

注： 站点仅在 PowerHA SystemMirror Enterprise Edition 和 Standard Edition 7.1.2 或更高版本中受支持。复制管理仅在 PowerHA SystemMirror Enterprise Edition 中受支持。

配置复制资源

如果您已安装 PowerHA SystemMirror Enterprise Edition for AIX 产品，那么可以使用集成支持来配置不同的复制技术。

以下复制资源是受支持的 PowerHA SystemMirror Enterprise Edition for AIX 配置：

- 使用并发节点策略的资源组可以具有非并发站点管理策略。
- 缺省情况下，PowerHA SystemMirror 的新安装允许对包含 PowerHA SystemMirror Enterprise Edition for AIX 复制资源的资源组进行站点间恢复。从前发行版中更新和迁移的配置将保留先前存在的行为。您可以针对集群启动的资源组移动将此行为配置为**故障转移**或**通知**选项。如果选择**通知**选项，那么需要配置前置事件或后置事件脚本，或远程通知方法。
- 复制资源组的父、子和位置依赖性配置。
- 对于 PowerHA SystemMirror 站点的资源组，基于节点的资源组分发启动策略。

注： 站点仅在 PowerHA SystemMirror Enterprise Edition 和 Standard Edition 7.1.2 或更高版本中受支持。复制管理仅在 PowerHA SystemMirror Enterprise Edition 中受支持。

无法将资源组配置为使用非并发节点策略和并发站点间管理策略。

处理复制资源

PowerHA SystemMirror 事件处理自动支持已配置的复制资源。

复制资源包括以下功能：

- 只要可能，PowerHA SystemMirror 就以并行方式处理事件。将对事件以动态方式划分阶段，这样 PowerHA SystemMirror 就可以用正确的顺序（`release_primary`、`release_secondary`、`acquire_secondary`、`acquire_primary`）来处理资源组的主要实例和次要实例。
- 如果其他节点或网络可用，那么在卷组丢失、获取故障和 **local_network_down** 事件期间，PowerHA SystemMirror 将在恢复复制资源组的主要实例的同时也恢复次要实例。
- 发生站点故障转移时，比起任何其他实例，PowerHA SystemMirror 更可能获取资源组的次要实例。PowerHA SystemMirror 可以将辅助站点上的所有节点视为目标，而不仅仅是先前托管主要实例的节点（这是 PowerHA SystemMirror 的先前版本中的行为）。

注： 站点仅在 PowerHA SystemMirror Enterprise Edition 和 Standard Edition 7.1.2 或更高版本中受支持。复制管理仅在 PowerHA SystemMirror Enterprise Edition 中受支持。

移动具有复制资源的资源组

可以将具有复制资源的资源组的主要实例移至其他站点，并且 PowerHA SystemMirror 将在同一操作中自动处理该组的次要实例。

注： 站点仅在 PowerHA SystemMirror Enterprise Edition 和 Standard Edition 7.1.2 或更高版本中受支持。复制管理仅在 PowerHA SystemMirror Enterprise Edition 中受支持。

集群管理器使用动态事件阶段划分，因为如果有节点可以托管次要实例，那么它会将次要实例从该站点移至另一个站点。将尽一切努力使次要实例保持 SECONDARY_ONLINE 状态。即使已对给定站点上的节点进行配置，使其无法托管多个主要实例，它也可以托管多个次要实例以使这些实例全部保持 SECONDARY_ONLINE 状态。

规划 Workload Manager

IBM offers AIX Workload Manager (WLM) 作为系统管理资源随附在 AIX 操作系统中。

通过 WLM，您可以针对不同进程和应用程序而为 CPU、物理内存使用及磁盘 I/O 带宽设置目标和限制。这可以更好地控制在峰值负载时关键系统资源的使用情况。PowerHA SystemMirror 允许您将 WLM 类配置到 PowerHA SystemMirror 资源组中，以便 WLM 的启动和停止以及活动的 WLM 配置可以由集群进行控制。

PowerHA SystemMirror 不会验证 WLM 配置的每个方面，因此，您有责任确保 WLM 配置文件的完整性。在将 WLM 类添加到 PowerHA SystemMirror 资源组后，验证实用程序将仅检查所需的 WLM 类是否存在。因此，必须仔细了解 WLM 工作原理及配置方法。如果配置参数不正确但是可接受，会降低系统的生产力和可用性。

有关如何设置和使用 Workload Manager 的完整信息，请参阅 *IBM AIX Workload Manager (WLM) Redbooks*® 出版物。

Workload Manager 根据所属的类在请求系统资源的各个进程之间分配系统资源。根据类分配规则将进程分配给特定的类。规划 WLM 与 PowerHA SystemMirror 的集成包括两个基本步骤：

1. 使用 AIX SMIT 面板定义与高度可用的应用程序相关的 WLM 类和类分配规则。
2. 使用 PowerHA SystemMirror SMIT 面板在 WLM 配置和 PowerHA SystemMirror 资源组之间建立关联。

相关信息：

 [AIX Workload Manager \(WLM\) Redbooks](#)

Workload Manager 类

Workload Manager 根据向其分配了进程的类，在请求系统资源的进程中分配系统资源。

类的属性包括：

- 类名称。长度不超过 16 个字符的唯一字母数字字符串。
- 类级别。数字 0 到 9。此数字决定了某个类在最重要（级别 0）到最不重要（级别 9）之间的相对重要性。
- CPU 和物理内存份额的数目。为每个类分配的实际资源数量取决于所有类中的总份额数（因此，如果系统上定义了两个类，一个类占目标 CPU 使用的两份，另一个类占三份，那么第一个类将接收 2/5 的 CPU 时间并且第二个类将接收 3/5 的 CPU 时间）。
- 配置限制。可供进程访问的 CPU 时间、物理内存和磁盘 I/O 带宽的最小和最大百分比限制。

您应设置类分配规则，告知 WLM 如何根据进程的组标识 (GID)、用户标识 (UID) 和完整路径名来为所有新进程（以及在 WLM 启动时已经运行的进程）分类。

Workload Manager 重新配置、启动和关闭

本节描述了使 WLM 受控于 PowerHA SystemMirror 之后重新配置、启动或停止 WLM 的方式。

Workload Manager 重新配置

WLM 类添加到 PowerHA SystemMirror 资源组后，如果在该节点上同步集群，PowerHA SystemMirror 将重新配置 WLM 以使用与该节点关联的类所需的规则。如果节点上发生了动态资源重新配置，那么将根据对 WLM 类（与某个资源组关联）进行的任何更改来重新配置 WLM。

Workload Manager 启动

当节点连接集群或者当进行 WLM 配置的动态重新配置时，WLM 便会启动。

配置特定于节点，并取决于节点所参与到的资源组。如果节点无法获取与 WLM 类相关联的任何资源组，那么 WLM 将不会启动。

对于启动策略不是“Online Using Node Distribution”策略的非并发资源组，启动脚本将确定资源组是在主节点还是辅助节点上运行，并将对应的 WLM 类分配规则添加到 WLM 配置。对于所有其他非并发资源组和节点可以获取的并发访问资源组，会将与每个资源组关联的主 WLM 类置于 WLM 配置中。对应的规则会添加到规则表。

最后，如果 WLM 当前正在运行，且没有由 PowerHA SystemMirror 启动，那么启动脚本会通过用户指定的配置来重新启动 WLM，同时保存之前配置。当 PowerHA SystemMirror 停止时，会将 WLM 恢复为之前的配置。

启动 WLM 失败会生成一条记录在 `hacmp.out` 日志文件中的错误消息，但是节点启动和资源重新配置会继续进行。

Workload Manager 关闭

当节点离开集群或者进行动态集群重新配置时，WLM 将关闭。如果 WLM 当前正在运行，那么关闭脚本将确定在由 PowerHA SystemMirror 启动 WLM 之前，WLM 是否正在运行，并确定 WLM 使用的配置。然后，不执行任何操作（如果 WLM 当前未在运行）、停止 WLM（如果 PowerHA SystemMirror 启动之前其未在运行）或停止 WLM 并按照之前配置重新启动 WLM（如果 WLM 之前正在运行）。

限制和注意事项

规划 Workload Manager 配置时记住一些限制和注意事项

这些限制和注意事项包括：

- 一些 WLM 配置可能会妨碍 PowerHA SystemMirror 性能。设计类和规则时小心，并警惕这些类和规则对 PowerHA SystemMirror 可能的影响。
- 您在整个集群中可以有不超 27 个非缺省 WLM 类，因为跨集群节点共享一个配置。
- PowerHA SystemMirror Workload Manager 配置不支持子类，即使 WLM 允许也是如此。如果为放置在资源组中的任何 WLM 类配置了子类，将在集群验证时发出警告，且同步期间不会将子类传播到其他节点。
- 在任何指定节点上，只有与某个节点可以获取的资源组关联的类的规则在该节点上才是活动的。

处理集群事件

以下主题描述 PowerHA SystemMirror 集群事件。

概述

PowerHA SystemMirror 提供了以下事件处理管理方式:

- 定制预定义事件
- 定义新事件

在 PowerHA SystemMirror 中, 除非您为集群中的所有或某些资源组指定了定制串行处理顺序, 否则在可能情况下, 缺省为以并行方式处理资源组。

示例所描述的事件的逻辑和顺序可能未列示所有事件。

有关以下内容的信息, 请参阅“启动和停止集群服务”:

- 启动和停止集群服务的所执行步骤
- 与 AIX **shutdown** 命令的交互, 以及 PowerHA SystemMirror 集群服务与 Reliable Scalable Cluster Technology (RSCT) 的交互

相关参考:

第 61 页的『规划资源组的并行或串行处理顺序』

缺省情况下, PowerHA SystemMirror 将以并行方式来获取和释放集群中配置的所有个体资源。但是, 您可以指定特定串行顺序, 将根据该顺序获取或释放某些或全部个体资源组。

规划站点和节点事件

您可以在 PowerHA SystemMirror Standard Edition for AIX 或 PowerHA SystemMirror Enterprise Edition 中定义站点。必须定义站点以启用 PowerHA SystemMirror Enterprise Edition 存储器复制支持, 包括 Geographic Logical Volume Manager (GLVM) 和 Metro Mirror。将节点和存储设备与站点关联时, 可使用 PowerHA SystemMirror 实现分割站点逻辑卷管理器 (LVM) 镜像配置。PowerHA SystemMirror 确定基于站点信息的相应选择, 并验证站级别的镜像配置的一致性。

注: 站点仅在 PowerHA SystemMirror Enterprise Edition 和 Standard Edition 7.1.2 或更高版本中受支持。复制管理仅在 PowerHA SystemMirror Enterprise Edition 中受支持。

站点事件脚本包含在 PowerHA SystemMirror 软件中。如果未定义站点, 那么将不会生成任何站点事件。如果已定义站点, 那么 PowerHA SystemMirror **site_event** 脚本将按如下所示运行:

- 站点中的第一个节点在完成 **node_up** 事件处理之前运行 **site_up** 事件。**site_up_complete** 事件在 **node_up_complete** 事件之后运行。
- 当站点的最后一个节点停机时, **site_down** 事件将在 **node_down** 事件之前运行, 并且 **site_down_complete** 事件将在 **node_down_complete** 事件之后运行。

无需安装 PowerHA SystemMirror Enterprise Edition, 就可以定义当站点更改状态时要运行的前置事件和后置事件。在这种情况下, 您可以定义所有与站点相关的过程。

站点事件 (包括 **check_for_site_up** 事件和 **check_for_site_down** 事件) 记录在 **hacmp.out** 日志文件中。

如果已定义站点, 那么当站点中的第一个节点开始正常运行时, **site_up** 事件将运行; 当站点中的最后一个节点停机时, **site_down** 事件将运行。用于处理资源组的事件脚本顺序通常为:

site_up

site_up_remote

node_up

rg_move 事件用于处理资源组操作

node_up_complete

site_up_complete

site_up_remote_complete

site_down

site_down_remote

node_down

rg_move 事件用于处理资源组操作

node_down_complete

site_down

site_down_remote_complete

规划 node_up 和 node_down 事件

node_up 事件由集群启动时即加入集群的节点启动，或由在稍后时间重新加入集群的节点启动。

建立初始集群成员资格

本主题描述在集群启动以及建立集群的尝试成员资格时，集群管理器在每个节点上执行的步骤。本主题显示集群管理器如何在成员节点中建立通信，以及随着集群成员信息发展集群资源如何分布。

第一个节点加入集群

1. PowerHA SystemMirror 集群服务在节点 A 上启动。Reliable Scalable Cluster Technology (RSCT) 子系统检查网络接口的状态，并开始与其他集群节点上的 RSCT 子系统通信。节点 A 上的集群管理器会累积初始状态信息，然后将表示其准备加入集群的信息广播至与其连接的所有已配置网络 (**node_up**)。
2. 节点 A 将没有响应解释为它是该集群中的第一个节点。
3. 节点 A 启动 **process_resources** 脚本，该脚本处理该节点的资源配置信息。
4. 当事件处理完成时，节点 A 变为集群的成员。PowerHA SystemMirror 运行 **node_up_complete**。

此时，为节点 A 定义的所有资源组对于客户机均可用。

如果将 **Online on First Available Node** 启动策略指定为资源组的启动行为，那么节点 A 将取得所有这些资源组的控制权。

如果节点 A 定义为具有 **Online Using Node Distribution** 启动策略的非并发资源组的一部分，那么此节点将取得节点环境中列示的第一个资源组的控制权。

如果节点 A 定义为并发访问资源配置的一部分，那么它将使这些并发资源可用。

对于具有 **Online on First Available Node** 启动策略并且配置了稳定时间的资源组，节点 A 在获取此类资源组之前将等待该稳定时间间隔。稳定时间有利于等待优先级较高的节点以加入集群。

第二个节点加入集群

5. PowerHA SystemMirror 集群服务在节点 B 上启动。节点 B 将表示其准备加入集群的消息广播至与其连接的所有已配置网络 (**node_up**)。
6. 节点 A 接收消息并发送应答。
7. 节点 A 将节点 B 添加至活动节点列表，并启动与节点 B 的保持活动通信。
8. 节点 B 从节点 A 接收应答。该消息包括将节点 A 标识为集群的唯一其他成员的信息。（如果有其他成员，那么节点 B 将接收到成员列表。）

9. 节点 B 处理 **process_resources** 脚本并发送消息以使其他节点了解它已经结束的时间。

对 **process_resources** 脚本的处理可能包括节点 A 释放其当前持有的资源。如果节点 A 和节点 B 均处于一个或多个资源组的资源组节点列表中，并且节点 B 针对一个或多个这些资源具有更高优先级。这仅对于支持回退的资源组才成立。

请注意，如果配置了延迟回退计时器，那么对于在节点 A 上处于联机状态的任何资源组，如果节点 B 是较高优先级的节点，该资源组将在延迟回退计时器指定的时间回退至节点 B。

10. 同时，将一直监视节点 B 并向其发送保持活动，并且节点 B 将等待接收有关集群成员信息更改的消息。当节点 B 完成其自己的 **process_resources** 脚本后，它会通知节点 A。

在其 **node_up** 处理期间，节点 B 将声明为其配置的所有资源组（请见步骤 3）。请注意，如果配置了延迟回退计时器，那么资源组将在该计时器指定的时间回退至较高优先级的节点。

11. 两个节点同时拥有 **node_up_complete** 事件。

此时，节点 B 将节点 A 包括在其成员节点列表和其保持活动列表中。

12. 节点 B 向集群中所有可能的节点发送“新成员”消息。
13. 当节点 A 获取到消息时，它会将节点 B 从其活动节点列表移到其成员节点列表。

此时，为节点 A 和节点 B 配置的所有资源组对于集群客户机均可用。

其余节点加入集群

14. 随着 PowerHA SystemMirror 集群服务在每个其余集群节点上启动，将以概述的顺序对发送和接收控制消息以及处理事件的每个成员节点重复步骤 4 到步骤 9。请特别注意，在完成事件处理并将新节点移动到集群成员列表之前，所有节点都必须确认 **node_up_complete** 事件。

随着新节点加入，每个节点上的 RSCT 子系统都将建立通信并开始发送脉动信号。节点和适配器将加入根据 PowerHA SystemMirror 配置中的定义形成的 RSCT 脉动信号环。当网络接口卡 (NIC) 或节点的状态发生更改时，集群管理器将接收到状态更改并生成适当的事件。

重新加入集群

当节点重新加入集群时，在现有节点上运行的集群管理器将发起 **node_up** 事件作为返回节点已启动的应答。当这些节点完成处理其 **process_resources** 脚本后，新节点将处理 **node_up** 事件，这样它可以继续提供集群服务。

此过程对于确保正确均衡集群资源是必要的。只要现有集群管理器先应答了重新加入集群的节点，它们就可以在需要时释放属于该节点的任何资源组。在此情况下是否将实际释放资源组取决于资源组的接管配置方式（或依赖性）。然后新节点可以启动其操作。

node_up 事件的顺序

以下列表描述了 **node_down** 事件的顺序：

node_up

当节点加入或重新加入集群时，将发生此事件。

process_resources

此脚本调用节点获取服务地址（或共享地址）所需的子事件，获取其拥有（或共享）的所有资源，并取得这些资源。这包括使磁盘可用、在卷组上联机、安装文件系统、导出文件系统、安装 NFS 文件系统，以及在并发访问卷组上联机。

process_resources_complete

当资源已处理后，每个节点都会运行此脚本。

node_up_complete

在处理资源后以及在成功完成 **node_up** 事件后，将发生此事件。根据节点是本地节点还是远程节点，此事件会调用 **start_server** 脚本以在本地节点上启动应用程序控制器，或者仅当远程节点已完全启动后才允许本地节点安装 NFS 文件系统。

针对依赖资源组的 node_up 事件

如果在集群中任何资源组之间配置了依赖性，那么 PowerHA SystemMirror 将使用 **rg_move** 事件（发生 **node_up** 事件时，针对所有资源组而启动）来处理与集群中资源组相关的所有事件。

然后，集群管理器将考虑所有节点策略，特别是资源组的依赖性配置以及资源组在所有节点上的当前分布和状态，以正确处理对资源组的任何获取、释放、联机或脱机操作，然后 **node_up_complete** 事件才能运行。

资源组之间的父子依赖性和位置依赖性为构建具有多层应用程序的集群提供了一种可预测的可靠方法。但是，在具有依赖性的集群中进行 **node_up** 处理会比在没有资源组依赖性的集群进行并行处理花费更长时间。您可能需要调整 **node_up** 事件的 **config_too_long** 警告计时器。

node_down 事件

当所有网络接口均关闭，或者某个节点不响应脉动信号时，集群管理器将运行 **node_down** 事件。根据集群配置，对等节点将采取必要操作，以使关键应用程序启动并运行，并确保数据保持可用。

node_down 事件可由节点启动：

- 停止集群服务并将资源组脱机
- 停止集群服务并将资源组移动到其他节点
- 停止集群服务并将资源组置于 **unmanaged** 状态。
- 发生故障。

停止集群服务并将资源组脱机

在您停止集群并将资源组脱机后，PowerHA SystemMirror 将在 **node_down_complete** 事件释放已停止节点的资源之后在本地节点上停止。其他节点将运行 **node_down_complete** 事件并且不接管已停止节点的资源。

停止集群服务并移动资源组

在您停止集群服务并将资源组移动到其他节点后，PowerHA SystemMirror 将在 **node_down_complete** 事件在本地节点上释放其资源组之后停止。资源组节点列表中的其余节点将接管这些资源组。

停止集群服务并将资源组置于 unmanaged 状态

在您停止集群服务并将资源组置于 **unmanaged** 资源组后，PowerHA SystemMirror 软件将在本地节点上立即停止。**node_down_complete** 事件将在已停止节点上运行。远程节点上的集群管理器将处理 **node_down** 节点，但不会接管任何资源组。已停止的节点不会释放它的资源组。

节点发生故障

在节点发生故障后，节点上的集群管理器没有时间生成 **node_down** 事件。在此情况下，其余节点上的集群管理器可识别出已发生了 **node_down** 事件（当它们意识到发生故障的节点不再进行通信时），它们将触发 **node_down** 事件。

这将启动一系列重新配置集群的子事件以处理发生故障的节点。根据集群配置，资源组节点列表中的其余节点将接管资源组。

node_down 事件的顺序

以下列表描述了 **node_down** 事件的缺省并行顺序：

1. **node_down**
2. 当节点有目的地离开集群或发生故障时，将发生此事件。
3. 在有些情况下，**node_down** 事件可接收 **forced** 参数。
4. 所有节点都运行 **node_down** 事件。
5. 所有节点都运行 **node_down** 事件。
6. 所有节点都运行 **process_resources** 脚本。在集群管理器对受影响资源组的状态和配置进行评估后，它将启动一系列子事件，以根据针对故障转移或回退的配置来重新分配资源。
7. 所有节点都运行 **process_resources_complete** 脚本。
8. **node_down_complete**

网络事件

PowerHA SystemMirror 可区分两种类型的网络故障：本地和全局，并且可为每种类型的故障使用不同的网络故障事件。网络故障事件脚本通常定制为发送邮件。

网络事件的顺序

下表显示了网络事件。

表 2. 网络事件的顺序

事件名称	描述
network_down (本地)	<p>仅当特定节点与网络失去联系时，才发生此事件。该事件具有以下格式：</p> <pre>network_down node_name network_name</pre> <p>如果在资源组中配置了服务 IP，那么集群管理器将选择性地采取恢复操作，以将受影响的资源组移动到其他节点。恢复操作的结果记录在 hacmp.out 中。</p>
network_down (全局)	<p>当与网络连接的所有节点均与网络失去联系时，将发生此事件。在此情况下，假设发生了与网络相关的故障，而不是发生与节点相关的故障。此事件具有以下格式：</p> <pre>network_down -1 network_name</pre> <p>注： -1 参数是负一。此参数指示 network_down 事件是全局事件。</p> <p>全局网络故障事件会向系统管理员发送通知邮件，但不采取进一步操作，原因是适当的操作取决于本地网络配置。</p>
network_down_complete (本地)	<p>此事件在本地网络故障事件完成后发生。它具有以下格式：</p> <pre>network_down_complete node_name network_name</pre> <p>当发生本地网络故障事件时，集群管理器将针对包含连接到该网络的服务网络接口卡 (NIC) 的资源组采取选择性的恢复操作。</p>
network_down_complete (全局)	<p>此事件在全局网络故障事件完成后发生。它具有以下格式：</p> <pre>network_down_complete -1 network_name</pre> <p>此事件的缺省处理不采取任何操作，原因是适当的操作取决于网络配置。</p>
network_up	<p>当集群管理器确定网络已变为可供使用时，将发生此事件。只要网络再次变为可用，PowerHA SystemMirror 就会尝试将包含服务 IP 标签的资源组在该网络上恢复联机。</p>

表 2. 网络事件的顺序 (续)

事件名称	描述
network_up_complete	仅当 network_up 事件已成功完成时，才会发生此事件。此事件通常定制为通知系统管理员某个事件需要人工关注。只要网络再次变为可用，PowerHA SystemMirror 就会尝试将包含服务 IP 标签的资源组在该网络上恢复联机。

网络接口事件

集群管理器通过启动一个事件来对故障、无效性或网络接口的连接作出回应。

下表显示了网络接口事件。

表 3. 网络接口事件

网络接口事件	事件描述
swap_adapter	当节点上托管服务 IP 标签的接口发生故障时，将发生此事件。 swap_adapter 事件会将服务 IP 标签移动到同一 PowerHA SystemMirror 网络上的引导接口，然后重新构造路由表。如果服务 IP 标签是 IP 别名，那么它将成为附加 IP 标签放置在引导接口上。否则，将从该接口中除去引导 IP 标签并将其放置在发生故障的接口上。如果当前存放服务 IP 标签的接口在稍后发生故障，那么 swap_adapter 可以切换至另一引导接口（如果存在）。如果向发生故障的接口分配了持久节点 IP 标签，那么会将该标签移动到引导接口。 注： PowerHA SystemMirror 将在关闭时从接口中除去 IP 别名。当网络变为可运行时，它会再次创建别名。 hacmp.out 文件会记录这些更改。
swap_adapter_complete	仅当 swap_adapter 事件已成功完成时，才会发生此事件。 swap_adapter_complete 事件确保已通过删除条目和 ping 集群 IP 地址更新了本地地址解析协议 (ARP) 高速缓存。
fail_standby	如果某个引导接口因为 IP 地址接管而发生故障或变为不可用，将发生此事件。 fail_standby 事件会显示一条控制台消息，该消息指示某个引导接口已发生故障或不再可用。
join_standby	如果某个引导接口变为可用，将发生此事件。 join_standby 事件会显示一条控制台消息，该消息指示某个引导接口已变为可用。在 PowerHA SystemMirror 中，只要某个网络接口变为可用，PowerHA SystemMirror 就会尝试将资源组恢复联机。
fail_interface	如果某个接口发生故障并且没有可用于恢复服务地址的引导接口，将发生此事件。会监视接管服务地址。可能有接口故障且没有用于恢复的可用接口，并且在同一网络上另一接口处于启动状态。此事件适用于所有网络，包括那些使用 IP 别名进行恢复的网络。当配置为“通过 IP 别名的 IPAT”的网络上的引导 NIC 发生故障时， fail_interface 事件将运行。如果发生故障的接口是服务标签，那么将触发 rg_move 事件。
join_interface	如果某个引导接口变为可用或恢复，将发生此事件。此事件适用于所有网络，包括那些使用“通过 IP 别名的 IPAT”进行恢复的网络。根据定义使用 IP 别名的网络未定义引导接口，所以在此情况下运行的 join_interface 事件将仅指示某个引导接口加入集群。

单个网络接口的故障不会生成事件

如果在网络上只有一个网络接口处于活动状态，那么集群管理器无法针对该网络接口生成故障事件，原因是它没有可与之通信从而确定该接口运行状况的对等项。具有单个网络接口的情况如下：

- 单节点集群
- 仅具有一个活动节点的多节点集群
- 具有虚拟以太网接口的多节点集群
- 全部发生故障，但每次每个网络上仅一个接口发生故障

例如，启动所有服务或引导接口均断开连接的集群将生成以下结果：

- 第一个节点活动：不生成故障事件。
- 第二个节点活动：生成一个故障事件。
- 第三个节点活动：生成一个故障事件。

集群范围状态事件

缺省情况下，集群管理器可识别用于重新配置集群和处理拓扑更改的时间限制。如果达到该时间限制，那么集群管理器将发起 **config_too_long** 事件。

所有集群状态事件包括：

表 4. 集群范围状态事件

集群范围状态事件名称	事件描述
config_too_long	每次集群事件的完成时间超过指定的超时周期时，就会显示此系统警告。此消息会记录在 hacmp.out 文件中。缺省情况下，所有事件的超时周期均设置为 360 秒。您可以使用 SMIT 界面来定制在 PowerHA SystemMirror 为集群事件发出 config_too_long 警告之前，为等待该集群事件完成留出的时间段。
reconfig_topology_start	此事件标记集群拓扑动态重新配置的开始。
reconfig_topology_complete	此事件指示集群拓扑动态重新配置已完成。
reconfig_resource_acquire	此事件指示受动态重新配置影响的集群资源正在由相应节点获取。
reconfig_resource_release	此事件指示受动态重新配置影响的集群资源正在被相应节点释放。
reconfig_resource_complete	此事件指示集群资源动态重新配置已成功完成。
cluster_notify	当自动集群配置监视在集群配置中检测到错误时，验证操作会触发此事件。该事件的输出将记录在整个集群中正运行集群服务的每个节点上的 hacmp.out 日志文件中。
event_error	如果任何节点具有致命错误，那么所有集群节点都将运行 event_error 事件。所有节点都会在 hacmp.out 日志文件中记录错误，并且注明发生故障的节点名。

资源组事件处理和恢复

集群管理器会跟踪资源组节点优先级策略、所有已配置的依赖性，以及必要的拓扑信息和资源组状态，这样它可以采取大量的恢复操作，通常不需要用户干预。事件记录包括每个高级别事件的详细摘要，从而帮助您确切了解在故障处理期间对每个资源组采取的操作。

有关 PowerHA SystemMirror 如何处理资源组的更多信息，请参阅“集群事件期间的资源组行为”。此主题包含有关以下 PowerHA SystemMirror 功能的信息：

- 用于处理资源组的选择性故障转移
- 处理资源组获取失败
- 处理使用服务 IP 资源配置的资源组
- 处理 PowerHA SystemMirror Enterprise Edition 资源组

相关信息：

集群事件期间的资源组行为

资源组事件

集群管理器可能会由于在处理例如节点关闭等事件期间采取的恢复操作而移动资源组。

注：

- 如果已指定资源组或站点之间的依赖性，那么 PowerHA SystemMirror 会以与通常不同的顺序处理事件。
- 下表中的列表未包括所有可能的资源组状态。资源组实例还可以在获取或释放进程中。此处未列示相应的资源组状态，但是提供了说明所采取的操作的描述性名称。

表 5. 资源组事件

资源组事件名称	事件描述
rg_move	此事件将指定资源组从某个节点移动到另一个节点。
rg_move_complete	此操作指示 rg_move 事件已成功完成。
resource_state_change	如果集群中配置了资源组依赖性，那么此触发器事件用于资源组恢复。此操作指示集群管理器需要更改一个或多个资源组的状态，或者集群管理器所管理的资源的状态发生更改。如果发生以下其中一种情况，那么此事件将在所有节点上运行： <ul style="list-style-type: none"> • 应用程序监视器故障 • 针对卷组丢失的选择性故障转移 • 本地网络关闭 • WAN 故障 • 资源组获取失败 • 基于 IP 接口可用性的资源组恢复 • 资源组的稳定计时器到期 • 资源组的回退计时器到期。
resource_state_change_complete	当 resource_state_change 事件成功完成时，将运行此事件。您可以根据需要，在此处添加前置事件和后置事件。例如，您可能希望收到有关资源状态更改的通知。
external_resource_state_change	当您移动资源组，并且因为在集群中配置了资源组依赖性，所以 PowerHA SystemMirror 使用动态处理路径处理请求时，将运行此事件。
external_resource_state_change_complete	当 external_resource_state_change 事件成功完成时，将运行此事件。

资源组子事件

在处理事件期间对个体资源的处理可能包括以下操作。例如，当文件系统处于卸载和安装的过程中，会将其脱机，然后由一个节点释放。然后，另一个节点获取该文件系统并将其联机。

下表包括了资源组的某些（并非全部）可能状态：

表 6. 资源组子事件

资源组子事件	事件描述
releasing	此操作指示正释放的资源组将脱机或在其他节点上获取。
acquiring	当资源组在节点上获取时，将使用此操作。
rg_up	此操作指示资源组联机。
rg_down	此操作指示资源组脱机。
rg_error	此操作指示资源组处于 ERROR 状态。
rg_acquiring_secondary	此操作指示资源组在目标站点上联机（仅复制资源联机）。
rg_up_secondary	此操作指示资源组在目标站点上的辅助角色中联机（仅复制资源联机）。
rg_error_secondary	此操作指示资源组在接收镜像数据的站点上处于 ERROR 状态。
rg_temp_error_state	此操作指示资源组处于临时 ERROR 状态。例如，由于本地网络或应用程序故障而发生了此问题。此状态通知集群管理器对此资源组启动 rg_move 事件。当集群稳定时，资源组不应处于此状态。

在某个事件完成后，集群管理器将具有该事件中涉及的资源和资源组的状态。然后集群管理器将分析它在内部维护的资源组信息，并确定是否需要为任何资源组将恢复事件放入队列。集群管理器还使用资源组中个体资源的状态将全面的事件摘要打印输出到 **hacmp.out** 日志文件。

对于每个资源组，集群管理器将跟踪资源组已尝试在其中联机但失败的节点。当处理恢复事件时，将更新此信息。一旦资源组变为联机或错误状态，集群管理器便会立即重置资源组的节点列表。

在 PowerHA SystemMirror 中，所显示的资源组 **ERROR** 状态带有详细信息：

表 7. 处于 *ERROR* 状态的资源组

资源组 <i>ERROR</i> 状态的原因	PowerHA SystemMirror 显示此信息
父组为 NOT ONLINE; 因此, 子资源组不可用	OFFLINE due to parent offline
优先级较高的不同节点依赖性组为 ONLINE	OFFLINE due to lack of available node
已获取到其他分配组	OFFLINE
组正在故障转移并且临时处于 OFFLINE 状态	OFFLINE

仅当资源组在事件处理结束后仍保持 *ERROR* 状态时, 才需要手动干预。

当资源组处于移动过程中, 应用程序监视将相应地暂挂和恢复。在处理事件期间, 应用程序监视器可识别出应用程序处于恢复状态。

表 8. 应用程序监视事件

应用程序监视事件	事件描述
resume_appmon	此操作由应用程序监视器用于恢复对应用程序的监视。
suspend_appmon	此操作由应用程序监视器用于暂挂对应用程序的监视。

相关信息:

集群事件期间的资源组行为

定制集群事件处理

集群管理器能够识别特定的一系列事件和子事件, 从而能够实现灵活的定制方案。使用 PowerHA SystemMirror 事件定制工具, 您可以为您的站点定制集群事件处理。使用定制事件处理, 您能够在发生故障时为最重要的资源提供高效的路径。但是, 此效率取决于您的配置。

在规划过程期间, 您需要确定是否定制事件处理。如果缺省脚本采取的操作足以满足您的需求, 那么在配置过程中不需要执行任何其他操作来配置事件。

如果您确实决定要针对所用环境来定制事件处理, 那么请使用本节中描述的 PowerHA SystemMirror 事件定制设施。如果您定制事件处理, 那么请在配置过程期间向 PowerHA SystemMirror 注册这些用户定义脚本。

事件定制设施包括以下功能:

- 事件通知
- 前置事件和后置事件处理
- 事件恢复和重试。

完整的事件定制包括 (在事件处理之前和之后) 发送至系统管理员的通知, 以及在事件处理之前和之后运行的用户定义命令或脚本, 如下例所示:

```
通知系统管理员将对事件进行处理
事件前脚本或命令 PowerHA SystemMirror 事件脚本
事件后脚本或命令 通知系统管理员事件处理已完成
```

事件通知

您可以指定一个通知命令, 用于发送邮件以指示某一事件即将发生 (或刚发生) 以及有关事件脚本的成功或失败。

您可在 SMIT 的 **Custom Cluster Configuration > Events > Cluster Events > Change/Show Pre-Defined Events** 菜单下为集群事件配置通知方法。例如，集群可能希望使用网络故障通知事件来通知系统管理员可能必须重新路由流量。然后，您可以使用 **network_up** 通知事件告知系统管理员可以再次通过已恢复的网络来处理流量。

PowerHA SystemMirror 集群中的事件通知也可以通过使用前置事件和后置事件脚本来实现。

您还可以配置定制远程通知来响应事件。

相关参考:

第 83 页的『定制事件的远程通知』

您可以通过 SMIT 界面定义通知方法，以发出定制页面来响应集群事件。您可以向包括手机在内的联系号码发送文本消息传递通知，也可以向电子邮件地址发送通知。

前置事件和后置事件脚本

您可以指定在集群管理器调用事件脚本之前或之后执行的命令或多个用户定义脚本。

例如，您可以指定在处理 **node_down** 事件脚本之前运行的一个或多个前置事件脚本。当集群管理器识别出某个远程节点关闭时，它会先处理这些用户定义的脚本。其中一个脚本可以指定向所有用户发送消息以指示性能可能受到影响（在适配器交换时以及在应用程序控制器停止并重新启动时）。在 **node_down** 事件脚本后面，可以包括一个用于 **network_up** 通知的后置处理事件脚本，以向所有用户广播某个特定系统目前在另一网络地址上可用的消息。

以下场景是前置事件和后置事件处理有用的其他示例:

- 如果发生 **node_down** 事件，那么此脚本可以通知将接管已关闭的应用程序控制器的服务器上的用户性能可能变化，或者他们应为特定应用程序寻求备用系统。
- 由于网络关闭，定制安装也许能够通过创建新 IP 路径重新路由通过其他机器的流量。**network_up** 和 **network_up_complete** 事件脚本可将过程反向，从而确保在所有网络都正常工作后存在正确的路径。
- 如果网络在本地节点上发生故障（但网络应正常工作），那么您可以后置事件脚本的形式停止集群服务，并将资源组移动到其他节点。

请注意，在编写 PowerHA SystemMirror 前置事件或后置事件脚本时，在 **/etc/environment** 中定义的所有 shell 环境变量都不可用于您的程序。如果您需要使用其中任何变量，请通过在脚本中添加以下行来显式获取这些变量:

```
". /etc/environment"
```

如果您计划为集群创建前置事件或后置事件脚本，请注意，向您脚本传递的参数将与您指定的 PowerHA SystemMirror 事件脚本使用的参数相同。对于前置事件和后置事件脚本，向事件命令传递的参数为事件名称、事件退出状态和向事件命令传递的尾随参数。

所有 PowerHA SystemMirror 事件脚本都保留在 **/usr/es/sbin/cluster/events** 目录中。向脚本传递的参数将列示在事件脚本头中。

注意:

请小心不要终止您脚本中的任何 **PowerHA SystemMirror** 进程。如果您使用 **ps** 命令的输出并使用 **grep** 搜索特定模式，请确保该模式与任何 **PowerHA SystemMirror Cluster Aware AIX (CAA)** 或 **Reliable Scalable Cluster Technology (RSCT)** 进程均不匹配。

可能不必要的前置事件和后置事件脚本

如果您从 PowerHA SystemMirror 的先前版本迁移，那么可能不再需要某些现有前置事件和后置事件脚本。PowerHA SystemMirror 本身可处理更多情况。

使用强制联机属性来代替前置事件或后置事件脚本

如果为卷组指定了强制联机属性，那么将不再需要强制执行联机操作的特殊脚本。

event_error 事件指示在远程节点上发生故障

从历史角度，不可恢复的事件脚本故障会导致 **event_error** 事件在发生故障的集群节点上运行。其余集群节点不会指示该故障。使用 PowerHA SystemMirror，如果任何节点具有不可恢复错误，那么所有集群节点都将运行 **event_error** 事件。所有节点都会在 **hacmp.out** 日志文件中记录错误，并且记录发生故障的节点。

如果您为 **event_error** 事件添加了前置或后置事件，请注意，将在所有节点上调用这些事件方法，而不仅仅是在发生故障的节点上。

Korn shell 程序环境变量 **EVENT_FAILED_NODE** 指示事件脚本失败的节点，此环境变量设置事件失败的节点的名称。请在前置事件或后置事件脚本中使用此变量来确定是否发生了故障。

变量 **LOCALNODENAME** 标识本地节点。如果 **LOCALNODENAME** 与 **EVENT_FAILED_NODE** 不同，那么在远程节点上发生了故障。

以并行方式处理并使用前置事件和后置事件脚本的资源组

在 PowerHA SystemMirror 中，除非您为集群中的所有或某些资源组指定了定制串行处理顺序，否则缺省为以并行方式处理资源组。

以并行方式处理资源组时，集群中发生的集群事件较少并且出现在事件摘要中。

使用并行处理可减少特定集群事件（您可为这类事件创建定制的前置事件或后置事件脚本）的数量。如果您开始对配置中的一系列资源组使用并行处理，请注意，您的现有前置事件或后置事件脚本可能对这些资源组不起作用。

特别是在资源组并行处理期间，仅会发生以下事件：

- acquire_svc_addr**
- acquire_takeover_addr**
- node_down**
- node_up**
- release_svc_addr**
- release_takeover_addr**
- start_server**
- stop_server**

注：在并行处理中，这些事件适用于以并行方式处理的资源组的整个列表，而不仅限于如串行处理中的单个资源组。如果您为这些事件配置了前置事件和后置事件脚本，那么在迁移后，这些事件脚本将不是针对单个资源组启动，而是针对一系列资源组启动，并且可能不按预期工作。

在资源组的并行处理中不会发生以下事件：

- get_disk_vg_fs**

node_down_local
node_down_remote
node_down_local_complete
node_down_remote_complete
node_up_local
node_up_remote
node_up_local_complete
node_up_remote_complete
release_vg_fs

如果您具有前置事件和后置事件脚本并且计划升级到当前版本，那么请考虑在并行处理中不会发生的那些事件。

如果您要继续使用前置事件和后置事件脚本，那么可能遇到以下其中一种情况。

方案	您要执行的操作
您想要对新添加的资源组使用前置事件和后置事件脚本。	<p>所有新添加的资源组将以并行方式处理，这将导致较少的集群事件。因此，您可为其创建前置事件和后置事件脚本的可供选择的事件有限。</p> <p>在此情况下，如果资源组中具有需要为特定集群事件而编写的前置事件和后置事件脚本处理的资源，请在 SMIT 中将资源组包括在串行处理列表中，以确保特定前置事件和后置事件脚本可用于这些资源。</p> <p>有关指定资源组串行或并行处理的信息，请参阅“为资源组配置处理顺序”一节。</p>
您可升级到 PowerHA SystemMirror V4.5 或更高版本，并为配置中预先存在的资源组选择并行处理。	<p>如果在迁移之前，您已在集群中配置了定制前置事件或后置事件脚本，那么此时这些资源组在迁移后将以并行方式处理，一些事件的事件脚本将无法用于这些资源组，原因是这些事件在并行处理中不会发生。</p> <p>如果您希望现有事件脚本继续处理资源组，请在 SMIT 中将资源组包括在串行顺序列表中，以确保前置事件和后置事件脚本可用于这些资源。</p> <p>有关指定资源组串行或并行处理的信息，请参阅“为资源组配置处理顺序”。</p>

相关参考:

第 44 页的『使用强制联机』

PowerHA SystemMirror 提供了强制联机功能，可与 **AIX** 自动错误通知方法联合使用。强制联机功能使您具有可能的最高数据可用性。

依赖资源组以及前置事件和后置事件脚本

过去，要完成资源组和应用程序排序，系统管理员必须在其前置事件和后置事件处理脚本中构建应用程序恢复逻辑。应为每个集群配置一个针对所有集群事件的前置事件脚本和一个针对所有集群事件的后置事件脚本。**PowerHA SystemMirror** 的最新发行版包括更多配置选项，它们使客户可以使用内置策略以实现相同排序和安排的资源组。如果您当前正在使用前置事件和后置事件来实现资源组排序，那么您可能希望查看实现以及向内置机制的转换。

这些脚本可变为涵盖各种情况的语句。例如，如果您要对特定节点上的特定事件采取操作，那么需要编辑该单个情况，为前置事件脚本和后置事件脚本添加所需代码，还应确保这些脚本在所有节点中相同。

总之，虽然此类脚本的逻辑可捕获到集群的期望行为，但是这些脚本难以定制并且以后在集群配置发生更改时更难以维护。

如果您在使用前置事件和后置事件脚本或其他方法（例如，资源组处理排序）来建立集群所支持的应用程序之间的依赖性，那么这些方法可能不再需要，或者可以极大简化。您可以改为指定集群中资源组之间的依赖性。有关规划依赖资源组的更多信息，请参阅“资源组依赖性”。

如果您将应用程序包括在依赖资源组中，并且除依赖性外仍然打算使用前置事件和后置事件脚本，那么可能需要对前置事件和后置事件脚本进行额外定制。为最大程度减小在应用程序停止和重新启动进程期间丢失数据的可能性，请定制应用程序控制器脚本，以确保在应用程序停止进程期间所有未落实的数据都将临时存储到共享磁盘，并将在应用程序重新启动进程期间读回到应用程序。使用共享磁盘是很重要的，原因是应用程序重新启动时所在的节点与其停止时所在的节点可能不同。

相关参考:

第 56 页的『资源组依赖性』

PowerHA SystemMirror 提供了多种配置，您可在这些配置中指定希望在启动、故障转移和回退时维护的资源组之间的关系。

事件恢复和重试

您可以指定尝试从事件脚本故障中恢复的命令。如果恢复命令成功并且事件脚本的重试计数大于零，那么事件脚本将重新运行。还可以指定尝试执行恢复命令的次数。

例如，恢复命令可能包括在用户注销后重试卸载某个文件系统，并确保当前没有任何人访问该文件系统。

如果确定了影响处理集群的给定事件的某种情况（例如，计时问题），那么您可以插入具有足够大重试计数的恢复命令以确保可从该问题中恢复。

定制事件的远程通知

您可以通过 SMIT 界面定义通知方法，以发出定制页面来响应集群事件。您可以向包括手机在内的联系号码发送文本消息传递通知，也可以向电子邮件地址发送通知。

可以使用验证自动监视 **cluster_notify** 事件来配置 PowerHA SystemMirror 远程通知方法，用于在集群配置中检测到错误时向外发送消息。该事件的输出将记录在整个集群中运行集群服务的每个节点上的 **hacmp.out** 文件中。

可以针对不同事件来配置任意数量的通知方法，并且这些通知方法可以具有不同的文本或数字消息和要拨打的电话号码。相同的通知方法可用于多个不同事件，只要关联的文本消息能够传递足够的信息来响应可能触发该通知的所有事件。

在配置通知方法后，您可以发送一条测试消息以确保所有设置均配置正确，并且将为给定事件发送预期消息。

规划定制远程通知

远程通知需要以下条件:

- 任何指定的端口都必须针对 AIX 操作系统进行定义，并且必须可用。
- 可以发送寻呼或文本消息的每个节点必须均安装并启用了适当的调制解调器。

注: 当通知方法已配置并且在发出寻呼之前，PowerHA SystemMirror 将检查端口的可用性。不会检查调制解调器状态。

- 可以从 SMIT 面板通过使用 AIX 邮件发送电子邮件消息的每个节点，都必须具有与因特网的 TCP/IP 连接。
- 对于可以向手机发送文本消息的每个节点，都必须安装并启用了与 Hayes 兼容的适当拨号器调制解调器。

定制出现警告前的事件持续时间

根据集群配置、集群节点的速度以及集群事件期间要移动的资源数量和类型，完成某些特定事件所需的时间可能不同。对于此类事件，您可能希望定制发出 **config_too_long** 警告消息之前，PowerHA SystemMirror 等待事件完成的时间段。

包括获取和释放资源组的集群事件需要花费较长时间完成。以下集群事件被视为慢速事件：

- **node_up**
- **node_down**
- **reconfig_resource**
- **rg_move**

通过为慢速集群事件定制事件持续时间，可避免在正常的集群操作期间收到不必要的系统警告。

所有其他集群事件都被视为快速事件。这些事件通常花费较短时间完成，并且不涉及获取或释放资源。快速事件的示例包括：

- **swap_adapter**
- 不处理资源组的事件

您可以为快速事件定制在接收到警告之前的事件持续时间，以更快地采取纠正操作。

如果在慢速集群事件中，PowerHA SystemMirror 过于频繁地发出警告消息，那么请考虑定制**出现警告前的事件持续时间**。在快速事件中，您将希望加快可能问题事件的检测速度。

注：资源组间的依赖性为构建具有多层应用程序的集群提供了一种可预测且可靠的方法。但是，对于处理具有依赖性的集群中的某些集群事件（例如，**node_up**），可能比以并行方式处理所有资源组的情况下处理这些事件花费更多时间。只要资源组依赖性允许，PowerHA SystemMirror 就会以并行方式处理多个非并发资源组，并且一次性处理所有节点上的多个并发资源组。但是，对于依赖于其他资源组的资源组，只有在其他资源组先启动之后，该资源组才能启动。应将 **node_up** 事件的 **config_too_long** 警告计时器设置为足够大，以允许此操作。

用户定义的事件

您可以定义自己的事件，PowerHA SystemMirror 可以针对这些事件运行您指定的恢复程序。此过程会向预定义的 PowerHA SystemMirror 前置事件和后置事件脚本定制设施添加新维度。

您可以通过 SMIT 界面在所定义的事件和定义事件恢复操作的程序之间指定映射。进行此映射后，您将控制每个恢复操作的作用域和在所有节点中同步的事件步骤数。

RMC 资源指向系统中其他一些组件提供服务的物理或逻辑实体的实例。术语“资源”广泛用于指代软件和硬件实体。例如，资源可以是特定的文件系统或特定的主机。资源类指相同类型的所有资源，例如，处理器或主机。

资源管理器（守护程序）将实际实体映射为 RMC 的抽象。每个资源管理器都代表特定一组管理任务或系统功能。资源管理器识别与该管理任务或系统功能组相关的关键物理或逻辑实体类型，并定义资源类以代表这些实体类型。

例如，主机资源管理器包含一组资源类，用于代表单个主机的各个方面。它定义了代表以下各项的资源类：

- 各个机器 (IBM.Host)
- 页面调度设备 (IBM.PagingDevice)
- 物理卷 (IBM.PhysicalVolume)

- 处理器 (IBM.Processor)
- 主机的标识标记 (IBM.HostPublic)
- 在主机上运行的程序 (IBM.Program)
- 主机支持的每种以太网适配器类型

AIX 资源监视器将针对与操作系统相关的资源状况（例如，空闲的 CPU 百分比 (IBM.Host.PctTotalTimeIdle, 或者已用磁盘空间百分比 (IBM.PhysicalVolume.PctBusy)) 生成事件。程序资源监视器针对所发生的与进程相关的情况（例如，进程意外结束）生成事件。程序资源监视器使用资源属性 IBM.Program.ProgramName。

写恢复程序

恢复程序具有一个恢复命令规范序列，该序列中可能还分布着 **barrier** 命令。

这些规范的格式遵循：

```
:node_set recovery_command expected_status NULL
```

其中：

- *node_set* 是要运行恢复程序的一组节点
- *recovery_command* 是一个用于指定可执行程序完整路径的引号分隔字符串。命令不能包括任何参数。需要参数的任何可执行程序都必须是一个单独的脚本。恢复程序在集群的所有节点上都必须使用此路径。程序必须指定退出状态。
- *expected_status* 是恢复命令成功完成时要返回的整数状态。集群管理器将返回的实际状态与期望的状态比较。不匹配表示恢复不成功。如果在预期状态字段中指定字符 X，那么集群管理器将忽略比较。
- *NULL* 表示当前未使用。

您按照动态关系指定节点集。PowerHA SystemMirror 支持以下动态关系：

All 恢复命令在当前包括的所有节点上运行。

事件 发生事件的节点。

其他 除发生事件的节点之外的所有节点。

指定动态关系将生成与原始恢复命令集相同的恢复命令集，不同之处在于节点标识替换了每个命令集中的 *node_set*。

用户定义事件命令的命令字符串必须以斜杠 (/) 开头。 **clcallev** 命令运行不以斜杠开头的命令。

RMC 信息的有用命令和参考

要列出 IBM.Host RMC 资源的所有持久属性定义 (*selection string* 字段)，请执行以下命令：

```
lsrsrcdef -e -A p IBM.Host
```

要列出 IBM.Host RMC 资源的所有动态属性定义 (*Expression* 字段)，请执行以下命令：

```
lsrsrcdef -e -A d IBM.Host
```

示例：恢复程序

样本程序将发送消息到 **/tmp/r1.out**，其调页空间在发生事件的节点上较低。对于恢复程序 **r1.rp**，SMIT 字段将按如下填充。

表 9. 示例: 恢复程序字段

字段	值
事件名称	E_page_space (用户定义名称)
恢复程序路径	/r1.rp
资源名称	IBM.Host (集群节点)
选择字符串	Name = ?" (节点名称)
表达式	TotalPgSpFree < 256000 (VMM 在 200 MB 的调页空间警告级别以内)。
重整表达式	资源属性加上要标记的条件。 TotalPgSpFree >256000 资源属性加上调整条件。

其中，恢复程序 **r1.rp** 如下：

```
#format:
#relationship >command to run >expected status NULL
#
event "/tmp/checkpagingspace" 0 NULL
```

恢复程序本身不执行带有参数的命令。而是指向 shell 脚本 **/tmp/checkpagingspace**，该脚本包含：

```
#!/bin/ksh
/usr/bin/echo "Paging Space LOW!" > /tmp/r1.out
exit 0
```

node_up 事件的恢复程序示例

下例是 **node_up** 事件的恢复程序：

```
#format:
#relationshipcommand to run expected status NULL
#
other "node_up" 0 NULL
#
barrier
#
event "node_up" 0 NULL
#
barrier
#
all "node_up_complete" X NULL
```

Barrier 命令

可以在恢复程序中放置任意数量的 **barrier** 命令。**barrier** 命令之前的所有恢复命令都并行启动。某个节点遇到 **barrier** 命令后，所有节点都必须到达该命令，恢复程序才能继续执行。

barrier 命令的语法是 **barrier**。

事件上卷

如果多个事件同时突显，只能看到优先级最高的事件。节点事件的优先级比网络事件的优先级更高。但是用户定义的事件优先级最低，完全不会上卷，因此您会看到全部。

事件摘要和前同步信号

当事件记录到节点的 **hacmp.out** 日志文件后，详细输出将包含很多行的事件详细信息，后跟简明事件摘要。通过事件摘要，可方便地浏览重要集群事件的日志。

通过使用 **Problem Determination Tools** SMIT 面板中的 **View Event Summaries** 选项，您可以查看过去七天 **hacmp.out** 日志文件中仅事件摘要部分的汇总。即使您将 **hacmp.out** 文件重定向到某个非缺省位置，也可以汇总事件摘要。**Display Event Summaries** 报告还包括 **clRGinfo** 命令生成的资源组信息。您还可以将事件摘要保存至指定的文件，而不是通过 SMIT 进行查看。

当事件处理具有依赖性的资源组时，会将一个前同步信号写入到 **hacmp.out** 日志文件，此日志文件列出了用于处理资源组的子事件的规划。

规划 PowerHA SystemMirror 客户机

以下主题讨论 PowerHA SystemMirror 客户机的规划注意事项。本步骤是在继续安装 PowerHA SystemMirror 软件之前的最后一步。

PowerHA SystemMirror 客户机是可访问 PowerHA SystemMirror 集群中节点的最终用户设备。为进行规划，从客户机角度评估集群很重要。

运行 Clinfo 的客户机

只要发生网络或节点事件，Clinfo 程序都将调用 **/usr/es/sbin/cluster/etc/clinfo.rc** 脚本。缺省情况下，此操作将更新系统的地址解析协议 (ARP) 高速缓存，以反映对网络地址的更改。如果需要更多操作，那么您可以定制此脚本。

重新连接到集群

运行 Clinfo 守护程序的客户机可以在发生集群事件后快速重新连接到集群。如果在集群和客户机之间存在除 IBM System p[®] 之外的硬件，请确保您可以在发生集群事件后更新这些网络组件的 ARP 高速缓存。

如果您将集群配置为交换硬件地址和 IP 地址，那么不必考虑更新 ARP 高速缓存。但是，请注意此选项会导致较长的延迟。

如果您正在使用“通过 IP 别名的 IPAT”，请确保所有客户机均支持 TCP/IP 免费 ARP。

定制 clinfo.rc 脚本

对于运行 Clinfo 守护程序的客户机，请确定是否要将 **/usr/es/sbin/cluster/etc/clinfo.rc** 脚本定制为在发生集群事件时执行除更新 ARP 高速缓存之外的更多操作。

不运行 Clinfo 的客户机

在不运行 Clinfo 守护程序的客户机上，您也许必须通过从集群节点 ping 客户机来间接更新本地地址解析协议 (ARP) 高速缓存。

在集群节点上，将您要通知的客户机主机的名称或地址添加到 **clinfo.rc** 脚本的 **PING_CLIENT_LIST** 变量中。当发生集群事件时，**clinfo.rc** 脚本将对 **PING_CLIENT_LIST** 变量中指定的每个主机运行以下命令：

```
ping -c1 $host
```

这会假定客户机已直接连接到某个集群网络。

网络组件

如果您将网络配置为网络在一侧与客户机连接并在另一侧与路由器、网桥或网关连接，而不是与集群的本地网络连接，请确保您可以在发生集群事件后更新这些组件的 ARP 高速缓存。

应用程序和 PowerHA SystemMirror

本主题提出在 PowerHA SystemMirror 下使您的应用程序具有高可用性时应考虑的某些关键问题。

借助 PowerHA SystemMirror，您可以通过在包含不同应用程序的资源组间建立依赖性来使用多层应用程序配置集群。以下主题描述资源组依赖性以及它们可以如何帮助保持依赖应用程序的高可用性。

相关参考:

第 5 页的『初始集群规划』

本部分描述为规划 PowerHA SystemMirror 集群以使应用程序具有高可用性所应采取的初始步骤。

应用程序和 PowerHA SystemMirror 概述

除了解使集群具有高可用性所需的硬件和软件外，您还必须清楚规划 PowerHA SystemMirror 环境时的应用程序可用性注意事项。集群的目标是即使在任意单点发生故障时仍使重要应用程序保持可用状态。要实现此目标，请考虑使应用程序可在 PowerHA SystemMirror 下恢复的各方面。

应用程序必须满足一些要求，才能在 PowerHA SystemMirror 下恢复良好。本文也将介绍一些必需的特性和一些建议。根据所有 PowerHA SystemMirror 环境均适用的关键点对这些内容进行分组。本主题涉及以下应用程序注意事项:

- **自动化。**请确保您应用程序的启动和停止无需用户干预
- **依赖性。**了解会影响应用程序的 PowerHA SystemMirror 外部因素
- **干扰。**了解应用程序本身可能干扰 PowerHA SystemMirror 运行
- **强大性。**选择强大且稳定的应用程序
- **实施。**使用适当的脚本、文件位置和 cron 调度。

您应添加应用程序监视器来检测应用程序启动的问题。在启动监视方式下，应用程序监视器检查应用程序控制器是否在指定的稳定时间间隔内成功启动并在稳定周期结束后退出。

您可以在不停止应用程序的情况下在节点上启动 PowerHA SystemMirror 集群服务，方法是从 SMIT 面板中选择选项 **PowerHA SystemMirror Services > Start Cluster Services**。在启动时，PowerHA SystemMirror 依赖于应用程序启动脚本和已配置的应用程序监视器来确保 PowerHA SystemMirror 可知晓正在运行的应用程序且不会启动应用程序的第二个实例。

您可以类似方式停止 PowerHA SystemMirror 集群服务并使应用程序在节点上保持运行状态。当已停止并已置于非受管状态的节点重新连接集群时，除非用户发起了 PowerHA SystemMirror 资源组命令将资源组置于其他状态（例如，在活动节点上联机），否则将认为资源处于相同状态。

应用程序自动化：最低程度手动干预

应用程序在 PowerHA SystemMirror 下可成功运行的一个关键要求是应用程序能够无需任何手动干预地启动和停止。

应用程序启动脚本

创建启动应用程序的启动脚本。启动脚本应执行为确保应用程序正常启动以及正确管理需要启动的应用程序实例数目所必需的任何清除或准备操作。将应用程序控制器添加到资源组后，PowerHA SystemMirror 将在处理资源组期间调用此脚本以使应用程序处于联机状态。因为集群守护程序会调用启动脚本，所以没有交互选项。此外，在 PowerHA SystemMirror 故障转移时，恢复进程将调用此脚本以使应用程序在备用节点上处于联机状态。这实现了完全自动恢复，也是此脚本应包括任何必要清除或准备操作的原因。

PowerHA SystemMirror 以 root 用户身份调用启动脚本。可能需要更改为其他用户，才能启动应用程序。**su** 命令可以完成此操作。对于在后台启动并且可能在 shell 退出时结束的命令，可能还需要运行 **nohup** 命令。

例如，某个 PowerHA SystemMirror 集群节点可能是网络信息服务（Network Information Service, NIS）环境中的客户机。如果出现此情况，并且您需要使用 **su** 命令更改用户标识，那么必须始终存在至 NIS 服务器的路径。如果不存在该路径并且尝试了 **su** 命令，那么应用程序脚本将暂挂。您可以通过允许 PowerHA SystemMirror 集群节点作为 NIS 客户机来避免此情况。这样，集群节点能够访问它自己的 NIS 映射文件来验证用户标识。

启动脚本还应检查是否存在所需资源或进程。这将确保应用程序可以成功启动。如果所需资源不可用，那么会向管理团队发送一条消息，要求更正此情况并重新启动应用程序。

应编写启动脚本，以便它们可以确定某个应用程序实例是否已正在运行，并且除非需要多个实例，否则不再启动另一实例。请记住，启动脚本可能在主节点失败后运行。可能需要对备份节点执行恢复操作，以便重新启动应用程序。此情况在数据库应用程序中常见。恢复操作必须也能够无需任何管理员交互即可运行。

应用程序停止脚本

应用程序停止脚本最重要的一点是它可完全停止应用程序。如果此操作失败，可能使 PowerHA SystemMirror 无法成功完成使备份节点接管资源。在停止期间，脚本可能需要解决一些与启动脚本解决的相同的问题，例如 NIS 和 **su** 命令。

应用程序停止脚本应使用分阶段方法。第一阶段应尝试停止集群服务并将资源组脱机。如果进程拒绝结束，那么应使用第二个阶段来强制确保所有处理均已停止。最后，第三个阶段可使用循环来重复任何所需步骤以确保应用程序已完全结束。

请确保当应用程序已成功停止时，应用程序停止脚本以值 0 退出。特别是，如果在应用程序已停止的情况下运行停止脚本，请检查发生的结果。在此情况下，脚本也必须以 0 退出。如果停止脚本以其他值退出，这表明“PowerHA SystemMirror 应用程序虽然可能处于某种被损害状态，但仍在运行”。**event_error** 事件将运行，并且集群将进入 ERROR 状态。此检查将警告管理员集群未正常运行。

请记住，在缺省情况下，PowerHA SystemMirror 将留出 360 秒以使事件完成处理。在集群完成重新配置并返回至稳定状态后，将显示一条消息，指示集群处于重新配置状态已太久。此警告可能指示脚本已暂挂并且需要手动干预。如果有可能，您可能希望考虑在停止 PowerHA SystemMirror 前手动停止应用程序。

您可以在调用 **config_too_long** 事件之前更改该时间段。

应用程序启动和停止脚本及依赖资源组

在 PowerHA SystemMirror 中，对依赖资源组的支持允许您配置以下选项：

- 资源组之间的三个依赖性级别，例如以下一种配置：节点 A 依赖节点 B，节点 B 依赖节点 C。PowerHA SystemMirror 禁止配置循环依赖性。
- 这是一种依赖性类型，其中，父资源必须在集群中的任何节点上联机，然后子（依赖项）资源组才能在节点上激活。

如果两个应用程序必须在同一节点上运行，那么这两个应用程序必须均位于同一资源组中。

如果子资源组包含的某个应用程序依赖父资源组中的资源，那么在故障转移条件下，如果父资源组故障转移至其他节点，子资源组将临时停止并自动重新启动。与此类似，如果子资源组处于并发状态，那么 PowerHA SystemMirror 会将其在所有节点上临时脱机，然后将其在所有可用节点上恢复联机。如果父资源组的故障转移未成功，那么父资源组和子资源组都将转入 ERROR 状态。

请注意，当子资源组临时停止并重新启动时，属于子资源组的应用程序也将停止并重新启动。因此，为最大程度减小在应用程序停止和重新启动进程期间丢失数据的可能性，请定制应用程序控制器脚本，以确保在应用程序停止进程期间所有未落实的数据都将临时存储到共享磁盘，并将在应用程序重新启动进程期间读回到应用程序。使用共享磁盘是很重要的，原因是应用程序重新启动时所在的节点与其停止时所在的节点可能不同。

应用程序层问题

应用程序通常具有多层体系结构（例如，数据库层、应用程序层和客户机层）。如果通过使用 PowerHA SystemMirror 使其中一层或多层具有高可用性，那么请考虑体系结构的所有层。

例如，如果使数据库具有高可用性，并且发生了故障转移，那么请考虑是否应对较高的层采取操作，以便使应用程序自动恢复服务。如果是，那么可能需要停止和重新启动应用程序或客户机层。可采用以下两种方式中的一种来完成此操作。一种方式是在层上运行 `cli_on_node` 命令，另一种方式是使用远程执行命令，例如 `rsh`、`rexec` 或 `ssh`。

注：某些方法（例如，使用 `~/rhosts` 文件）可能造成安全性风险。

使用依赖资源组

要配置具有多层应用程序的复合集群，您可以使用父子依赖资源组。您可能还希望考虑使用位置依赖性。

使用 Clinfo API

Clinfo API 是集群信息守护程序。您可以使用 Clinfo API 编写可在任何层上运行的程序，该程序将在故障转移成功完成后停止并重新启动应用程序。在此意义上，层（或应用程序）将变为对集群敏感，并将对集群中发生的事件作出响应。

使用前置事件脚本和后置事件脚本

解决多层体系结构问题的另一种方式是在某个集群事件的前后使用前置事件脚本和后置事件脚本。这两个脚本将调用远程执行命令，例如，`rsh`、`rexec` 或 `ssh` 来停止和重新启动应用程序。

相关概念：

第 88 页的『应用程序和 PowerHA SystemMirror』

本主题提出在 PowerHA SystemMirror 下使您的应用程序具有高可用性时应考虑的某些关键问题。

相关参考：

第 92 页的『编写有效脚本』

编写巧妙的应用程序启动脚本还有助于在将应用程序联机时减少出现问题的可能性。

第 51 页的『规划资源组』

以下主题描述如何规划 PowerHA SystemMirror 集群内的资源组。

『应用程序依赖性』

过去，要完成资源组和应用程序排序，系统管理员必须在其前置事件和后置事件处理脚本中构建应用程序恢复逻辑。应为每个集群配置一个针对所有集群事件的前置事件脚本和一个针对所有集群事件的后置事件脚本。

应用程序依赖性

过去，要完成资源组和应用程序排序，系统管理员必须在其前置事件和后置事件处理脚本中构建应用程序恢复逻辑。应为每个集群配置一个针对所有集群事件的前置事件脚本和一个针对所有集群事件的后置事件脚本。

这些脚本可变为涵盖各种情况的语句。例如，如果您要对特定节点上的特定事件采取操作，那么需要编辑该个别情况，为前置事件脚本和后置事件脚本添加所需代码，还应确保这些脚本在所有节点中相同。

总之，虽然这些脚本的逻辑可捕获到集群的期望行为，但是这些脚本难以定制并且以后在集群配置发生更改时更难以维护。

如果您在使用前置事件脚本和后置事件脚本或其他方法（例如，资源组处理排序）来建立集群所支持的应用程序之间的依赖性，那么这些方法可能不再需要，或者可以极大简化。您可以改为指定集群中资源组之间的依赖性。

注：在很多情况下，应用程序所依赖的远不止数据和 IP 地址。要使 PowerHA SystemMirror 下的任何应用程序都能够成功，应了解应用程序为正常运行而不应该依赖的事项，这一点很重要。本主题概述了很多主要的依赖性问题。请记住，这些依赖性可能来自 PowerHA SystemMirror 和应用程序环境外部。它们可能是不兼容的产品或外部资源冲突。请跳出应用程序本身，在企业内查找潜在的问题。

本地连网设备

本地连网设备可能造成明显的依赖性问题。若出现失败转移，如果这些设备未连接至备用节点并且不能对其进行访问，那么应用程序可能无法正常运行。这些设备可能包括 CD-ROM 设备、磁带设备或光盘库。请考虑您的应用程序是否依赖于任何这些设备，并且这些设备是否可在集群节点间共享。

硬编码

将应用程序硬编码到特定位置的特定设备中会导致潜在的依赖性问题。例如，控制台通常指定为 `/dev/tty0`。虽然通常情况下会分配此名称，但并不保证一定如此。如果您的应用程序采用名称 `/dev/tty0`，那么请确保所有可能的备用节点均具有相同配置。

主机名依赖性

有些应用程序编写为依赖于 AIX 主机名。它们将发出一个命令，以便验证许可证或名称文件系统。主机名不是 IP 地址标签。主机名特定于节点，并且 PowerHA SystemMirror 不对主机名进行故障转移。主机名无法更改。PowerHA SystemMirror 不支持任何更改主机名的应用程序。

发放软件许可证

另一个可能的问题是软件许可。软件的许可证可能针对特定 CPU 标识。如果您的应用程序是这种情况，那么软件的故障转移将不会成功重新启动。您也许能够通过所有集群节点上创建软件副本来避免此问题。请了解您的应用程序是否使用许可证是针对特定 CPU 标识的软件。

相关参考：

第 13 页的『多层应用程序的规划注意事项』

使用多层应用程序的业务配置可以使用父子依赖性资源组。例如，数据库必须在应用程序控制器之前联机。在此情况下，如果数据库关闭并且移动到其他节点，那么必须先将包含应用程序控制器的资源组脱机，再将其在集群的任何节点上联机。

应用程序干扰

有时，应用程序或应用程序环境可能干扰 PowerHA SystemMirror 的正常运行。某个应用程序可能在主节点和备用节点上均正常运行。但是，当 PowerHA SystemMirror 启动时，可能与应用程序或环境发生冲突，该冲突会使 PowerHA SystemMirror 无法成功运行。

处理网络路径的产品

此外，处理网络路径的产品可能使 PowerHA SystemMirror 无法按其设计运行。这些产品可能找到经过已发生原始故障的网络的第二条路径。此路由可能使 PowerHA SystemMirror 无法正确诊断故障以及采取适当恢复操作。

相关参考:

第 5 页的『初始集群规划』

本部分描述为规划 PowerHA SystemMirror 集群以使应用程序具有高可用性所应采取的初始步骤。

应用程序的强大性

任何应用程序能够成功的首要一点是应用程序的良好运行状况，或称强大性。如果应用程序不稳定或间歇性崩溃，那么请在将应用程序放入高可用性环境之前解决这些问题。

除基本稳定性外，PowerHA SystemMirror 下的应用程序还应满足其他强大性特性。

在硬件发生故障后成功启动

PowerHA SystemMirror 的良好候选应用程序应能够在硬件故障后成功重新启动。在使用 PowerHA SystemMirror 管理应用程序之前，请对其运行测试。在高负载和使节点发生故障的条件下运行应用程序。在节点恢复联机后，它采取了哪些恢复操作？此恢复操作是否能完全自动执行？如果不能，那么该应用程序可能不是实现高可用性的良好候选。

实内存丢失的存活

应用程序应定期将重新启动所必需的任何信息保存到磁盘。如果发生故障，那么应用程序可从它之前所在位置开始，而不是完全重新开始。

应用程序实施策略

当您规划在 PowerHA SystemMirror 下实施某个应用程序时，应考虑该应用程序的一些方面。

请考虑诸如启动时间、发生故障后重新启动的时间，以及停止时间等特性。您在一些方面的决定（例如，脚本编写、文件存储、`/etc/inittab` 文件和 cron 调度问题）会提高应用程序成功实施的可能性。

编写有效脚本

编写巧妙的应用程序启动脚本还有助于在将应用程序联机时减少出现问题的可能性。

启动脚本的一种好的作法是在启动应用程序之前检查先决条件状况。这些先决条件可能包括对文件系统的访问权、足够的调页空间和可用文件系统空间。如果不满足需要，启动脚本应退出并运行一个命令来通知系统管理员。

当启动数据库时，应考虑在同一集群中是否有多个实例，这一点很重要。如果是这样，那么应仅启动适合每个节点的实例。某些数据库启动命令会读取配置文件，并且同时启动所有已知数据库。此行为可能不是适合所有环境的期望配置。

请小心不要终止您脚本中的任何 PowerHA SystemMirror 进程。如果您使用 `ps` 命令的输出并使用 `grep` 搜索特定模式，请确保该模式与任何 PowerHA SystemMirror 或 Reliable Scalable Cluster Technology (RSCT) 进程均不匹配。

考虑文件存储位置

请考虑配置文件所在的位置。配置文件可以在一个共享磁盘上，这样才有可能由具有联机卷组的任何节点访问，或者在每个节点的内部磁盘上。对于应用程序的各方面而言均应如此。某些文件必须在共享驱动器上。这些文件包括数据、日志和可通过应用程序执行而更新的任何内容。诸如配置文件或应用程序二进制文件等文件可位于其中任意一种位置。

将可选文件存储在其中任意一种位置有利有弊。如果将文件存储在每个节点的内部磁盘，那么意味着您具有应用程序的多个副本，并可能具有多个许可证。保持这些文件同步需要更多成本和维护。但是，如果应用程序需要升级，那么不需要将整个集群停止运行。可以升级一个节点，而使其他节点保持运行。最适合特定环境的作法才是最佳作法。

考虑 `/etc/inittab` 和 `cron` 表问题

还应考虑通过 `/etc/inittab` 文件或 `cron` 表启动的应用程序或应用程序所需的资源。

当您启动系统时，`inittab` 文件会启动应用程序。如果应用程序运行需要集群资源，那么在 PowerHA SystemMirror 启动之前，这些集群资源将不会变为可用。最好使用 PowerHA SystemMirror 应用程序控制器设施，该设施可使应用程序成为在所有依赖资源联机后才启动的资源。

注：在 `/etc/inittab` 文件中，以下设置正确很重要：

```
hacmp:2:once:/usr/es/sbin/cluster/etc/rc.init
```

- `clinit` 和 `pst_clinit` 条目必须是最后两个运行级别为“2”的条目。
- `clinit` 条目后面必须是 `pst_clinit` 条目。

如果其中任何条目错误，将使 PowerHA SystemMirror 无法启动。

在 `cron` 表中，作业是根据该表中设置的调度和节点上的日期设置启动的。此信息在内部磁盘上维护，因此无法与备用节点共享。请同步这些 `cron` 表，以便备用节点可以在适当时间执行必需的操作。此外，请确保在主节点和其任何备用节点上设置的日期相同。

示例：Oracle 数据库和 SAP R/3

以下两个实例说明了为使应用程序 Oracle 数据库和 SAP R/3 在 PowerHA SystemMirror 下工作正常应考虑的问题。

示例 1：Oracle 数据库

Oracle 数据库和很多数据库一样，可在 PowerHA SystemMirror 下正常工作。这个强大的应用程序也能很好地处理故障。它可以在故障转移后回滚未落实的事务并及时返回到服务状态。但是，在 PowerHA SystemMirror 下使用 Oracle 数据库时，有几点应当记住。

启动 Oracle

必须以 Oracle 用户标识启动 Oracle。这样，启动脚本应包含以下代码：`su - oracleuser`。短横线 (-) 很重要，因为 `su` 命令需要具备 Oracle 用户的所有特性并位于 Oracle 用户主目录中。该命令将类似以下内容：

```
su - oracleuser -c /apps/oracle/startup/dbstart
```

`dbstart` 命令和 `dbshut` 命令读取 `/etc/oratabs` 文件中有关哪些数据库实例已知且应启动的指示信息。在某些情况下，启动所有实例是不正确的，因为它们可能由其他节点拥有。两个 Oracle 实例相互接管就属于这种情况。`oratabs` 文件通常位于内部磁盘上，因此无法共享。在适当的情况下，请考虑其他启动不同 Oracle 实例的方法。

停止 Oracle

应特别关心 Oracle 停止过程。有多种不同方法可确保 Oracle 已完全停止。建议顺序如下：首先，执行正常关闭；其次，立即调用 `shutdown`，此方法具有一定的强制性；最后，创建一个循环来检查进程表，以确保所有 Oracle 进程均已退出。

Oracle 文件存储

Oracle 产品数据库包含一些文件和数据。需要将数据和重做日志存储在共享磁盘上，以便两个节点都可以访问信息。但是，Oracle 二进制文件和配置文件可以位于内部或共享磁盘上。请考虑最适合您环境的方法。

示例 2: SAP R/3, 多层应用程序

SAP R/3 是三层应用程序的一个示例。它具有数据库层、应用程序层和客户机层。最常见的情况是使数据库层具有高可用性。在此情况下，当发生故障转移并且数据库重新启动时，需要停止并重新启动 SAP 应用程序层。您可用以下两种方式之一来完成此操作：

- 使用远程执行命令，例如 **rsh**、**rexec** 或 **ssh**

注：某些方法（例如，使用 `~/rhosts` 文件）可能造成安全性风险。

- 使应用程序层节点可感知集群。

使用远程执行命令

停止和启动 SAP 应用程序层的第一种方式是创建可在应用程序节点上执行远程命令执行的脚本。SAP 的应用程序层将停止，然后会重新启动。将对应用程序层中的每个节点执行此过程。使用远程执行命令需要允许数据库节点可访问应用程序节点的一种方法。

注：某些方法（例如，使用 `~/rhosts` 文件）可能造成安全性风险。

使应用程序层节点可感知集群

停止和启动应用程序层的第二种方法是使应用程序层节点可感知集群。这意味着应用程序层节点可感知到集群化数据库并了解发生故障转移的时间。您可以通过使应用程序层节点成为 PowerHA SystemMirror 服务器或客户机来实现此目的。如果应用程序节点是服务器，那么它将运行与数据库节点相同的集群事件来指示发生故障。然后，可以编写前置事件和后置事件脚本来停止和重新启动 SAP 应用程序层。如果应用程序节点是 PowerHA SystemMirror 客户机，那么将使用 SNMP 通过集群信息守护程序 (Clinfo) 向其发送数据库故障转移通知。可以使用 Clinfo API 编写一个程序来停止和重新启动 SAP 应用程序层。

相关信息：

为 Clinfo API 客户机应用程序编写程序

声明

本信息是为在美国提供的产品和服务编写的。

IBM 可能在其他国家或地区不提供本文中讨论的产品、服务或功能特性。有关您所在区域当前可获得的产品和服务的信息，请向您当地的 IBM 代表咨询。任何对 IBM 产品、程序或服务的引用并非意在明示或暗示只能使用 IBM 的产品、程序或服务。只要不侵犯 IBM 的知识产权，任何同等功能的产品、程序或服务，都可以代替 IBM 产品、程序或服务。但是，评估和验证任何非 IBM 产品、程序或服务的操作，由用户自行负责。

IBM 可能已拥有或正在申请与本文档内容有关的各项专利。提供本文档并不意味着授予用户使用这些专利的任何许可。您可以用书面形式将许可查询寄往：

*IBM Director of Licensing
IBM Corporation
North Castle Drive, MD-NC119
Armonk, NY 10504-1785
United States of America*

有关双字节字符集（DBCS）信息的许可查询，请与您所在国家或地区的 IBM 知识产权部门联系，或用书面方式将查询寄往：

*Intellectual Property Licensing
Legal and Intellectual Property Law
IBM Japan Ltd.
19-21, Nihonbashi-Hakozakicho, Chuo-ku
Tokyo 103-8510, Japan*

以下段落对于英国和与当地法律有不同规定的其他国家或地区均不适用：INTERNATIONAL BUSINESS MACHINES CORPORATION“按现状”提供本出版物，不附有任何种类的（无论是明示的还是暗含的）保证，包括但不限于暗含的有关非侵权、适销和适用于某特定用途的保证。某些国家或地区在某些交易中不允许免除明示或暗含的保证。因此本条款可能不适用于您。

本信息可能包含技术方面不够准确的地方或印刷错误。本信息将定期更改；这些更改将编入本信息的新版本中。IBM 可以随时对本出版物中描述的产品和/或程序进行改进和/或更改，而不另行通知。

本资料中对非 IBM Web 站点的任何引用都只是为了方便起见才提供的，不以任何方式充当对那些 Web 站点的认可。那些 Web 站点中的资料不是 IBM 产品资料的一部分，使用那些 web 站点带来的风险将由您自行承担。

IBM 可以按它认为适当的任何方式使用或分发您所提供的任何信息而无须对您承担任何责任。

本程序的被许可方如果要了解有关程序的信息以达到如下目的：(i) 使其能够在独立创建的程序和其它程序（包括本程序）之间进行信息交换，以及 (ii) 使其能够对已经交换的信息进行相互使用，请与下列地址联系：

*IBM Corporation
Dept. LRAS/Bldg. 903*

11501 Burnet Road
Austin, TX 78758-3400
USA

只要遵守适当的条款和条件，包括某些情形下的一定数量的付费，都可获得这方面的信息。

本文中描述的许可程序及其所有可用的许可资料均由 IBM 依据 IBM 客户协议、IBM 国际程序许可协议或任何同等协议中的条款提供。

此处包含的任何性能数据都是在受控环境中测得的。因此，在其他操作环境中获得的结果可能会有明显的不同。有些测量可能是在开发级的系统上进行的，因此不保证与一般可用系统上进行的测量结果相同。此外，有些测量是通过推算而估计的，实际结果可能会有差异。本文档的用户应当验证其特定环境的适用数据。

涉及非 IBM 产品的信息可从这些产品的供应商、其出版说明或其他可公开获得的资料中获取。IBM 没有对这些产品进行测试，也无法确认其性能的精确性、兼容性或任何其他关于非 IBM 产品的声明。有关非 IBM 产品性能的问题应当向这些产品的供应商提出。

所有关于 IBM 未来方向或意向的声明都可随时更改或收回，而不另行通知，它们仅仅表示了目标和意愿而已。

显示的所有 IBM 的价格均是 IBM 当前的建议零售价，可随时更改而不另行通知。经销商价格可能会有所不同。

本信息仅用于规划的目的。在所描述的产品上市之前，此处的信息会有更改。

本信息包含在日常业务操作中使用的数据和报告的示例。要尽可能对它们进行完整地举例说明，示例应包含个人、公司、商标和产品的名称。所有这些名字都是虚构的，若现实生活中实际业务企业使用的名字和地址与此相似，纯属巧合。

版权许可证:

本信息包含源语言形式的样本应用程序，用以阐明在不同操作平台上的编程技术。如果是为按照在编写样本程序的操作平台上的应用程序编程接口 (API) 进行应用程序的开发、使用、经销或分发为目的，您可以任何形式对这些样本程序进行复制、修改、分发，而无须向 IBM 付费。这些示例尚未在所有条件下经过全面测试。因此，IBM 不能担保或暗示这些程序的可靠性、可维护性或功能。这些实例程序“按现状”提供，不附有任何种类的保证。对于因使用样本程序所引起的任何损害，IBM 概不负责。

凡这些实例程序的每份拷贝或其任何部分或任何衍生产品，都必须包括如下版权声明:

此部分代码是根据 IBM Corp. 公司的样本程序衍生出来的。

© Copyright IBM Corp. (输入年份)。All rights reserved.

隐私策略注意事项

IBM 软件产品 (“软件产品”，其中包括作为服务解决方案的软件) 可能使用 cookie 或其他技术来收集产品使用信息，以帮助改进最终用户体验、定制与最终用户的交互或实现其他目的。在许多情况下，软件产品不会收集任何个人可标识信息。我们的某些软件产品可以帮助您收集个人可标识信息。如果此软件产品使用 cookie 来收集个人可标识信息，那么会在下面列出有关此产品使用 cookie 的特定信息

此软件产品不会使用 cookie 或其他技术来收集个人可标识信息。

如果为此软件产品部署的配置使您能够作为客户通过 cookie 和其他技术从最终用户收集个人可标识信息，那么您应该向您自己的法律顾问咨询有关适用于这种数据收集（其中包括对于通知和同意的任何需求）的任何法律。

有关为这些目的使用各种技术（其中包括 cookie）的更多信息，请参阅“IBM 隐私策略”（网址为 <http://www.ibm.com/privacy>）和“IBM 在线隐私声明”（网址为 <http://www.ibm.com/privacy/details>）中标题为“cookie、Web 信标和其他技术”和“软件产品和 Software-as-a 服务”（网址为 <http://www.ibm.com/software/info/product-privacy>）的部分。

商标

IBM、IBM 徽标和 [ibm.com](http://www.ibm.com) 是 International Business Machines Corp. 在全世界许多管辖区域注册的商标或注册商标。其他产品和服务名称可能是 IBM 或其他公司的商标。当前最新的 IBM 商标列表在以下 Web 站点提供版权和商标信息 (www.ibm.com/legal/copytrade.shtml)。

索引

[A]

安全 9
按需容量升级
 请参阅 CoD

[B]

并发
 资源组 52

[C]

初始集群规划 5
处理顺序
 资源组 61
磁带机 32
磁带设备 29
磁盘
 电缆 32
 电源注意事项 30
 非共享磁盘存储 30
 概述 29
 共享磁盘安装 31
 共享磁盘技术 30
 适配器 32
 虚拟 SCSI 31
 IBM DS4000 Storage Server
 sample 32
磁盘访问 40
 增强型并发 40
磁盘配置
 添加 32

[F]

非并发
 资源组 52
分割站点镜像
 规划 38
复制资源 68

[G]

概述
 磁盘 29
 规划过程 3
 集群事件 71
 应用程序 88
 资源组 51

概述 (续)
 AIX Workload Manager 69
共享磁盘 29
共享 LVM 组件 34
共享 SCSI 磁盘
 安装 31
规划
 IPv6 26
 LVM 分割站点镜像 38
规划过程
 概述 3

[J]

集群
 分区 17
 图 14
集群事件
 请参阅 事件
监视
 集群 25
节点 5
 事件 71
 node_down 事件 72
 node_up 事件 72
节点隔离 17
卷组 35

[K]

快速磁盘接管 41

[L]

联机 42
 强制 44
逻辑卷 36

[M]

脉动信号传递 18

[R]

日志记录
 制作镜像 38

[S]

事件

- 概述 71
- 集群范围状态 77
- 节点 71
- 前置事件和后置事件脚本 80
- 通知 80
- 网络接口 76
- 用户定义的 84
- 摘要 87
- 站点 71
- 资源组 77
- network 75

示例

- 网络连接 17

[T]

添加

- 磁盘配置 32
- 网络拓扑 29

拓扑结构

- network 19

[W]

网络接口

- 事件 76

文件系统 36

物理分区

- 制作镜像 37

物理卷 35

[X]

虚拟适配器 27

虚拟网络 27

虚拟以太网 18

虚拟 SCSI 31

[Y]

移动

- 资源组 60

应用程序 10, 88

- 编写脚本 92

- 多层 13

- 概述 88

- 干扰 91

- 相关性 90

应用程序监视 12

应用程序控制器 11

[Z]

站点 8

- 事件 71

- 制作镜像 39

- 资源组 62

制作镜像

- 日志记录 38

- 物理分区 37

- 站点 39

准则 2

资源组 51

- 处理顺序 61

- 复制资源 68

- 概述 51

- 类型 52

- 使用 clRGmove 移动 60

- 事件 77

- 属性 53

- 网络 61

- 移动 60

- 站点 62

- policies 53

A

AIX Workload Manager

- 概述 69

C

client 87

- 不运行 clinfo 87

- 网络组件 88

- clinfo 87

clinfo 87

- 不运行 87

clRGmove

- 资源组 60

CoD 10

D

DNS 25

H

hacmp.out 87

I

IBM DS4000 Storage Server

- sample 32

IP 标签 16

IP 别名 16

IP 地址接管
 IP 别名 22
IPv6
 规划 26

L

LVM 分割站点镜像
 规划 38
LVM 镜像 36
LVM 组件 35

N

netmon.cf 27
network
 避免冲突 29
 集群分区 17
 监视集群 25
 交换网络 17
 连接 16
 连接性 (connectivity) 15
 脉动信号传递 18
 事件 75
 示例 17
 添加拓扑 29
 通过 IP 别名的 IP 地址接管 22
 拓扑结构 19
 虚拟以太网 18
 资源组 61
 client 88
 DNS 25
 IP 标签 16
 IP 别名 16
 NIS 25
 Oracle 27
 VPN 防火墙 26
NFS 45
NIS 25

O

Oracle
 规划网络 27

Q

quorum 42, 43

S

sample
 IBM DS4000 Storage Server 32
 Oracle 数据库和 SAP R/3 93

script
 前置事件和后置事件 80
 写 92
status
 事件 77

V

VPN 防火墙 26

W

WLM
 请参阅 AIX Workload Manager



Printed in China