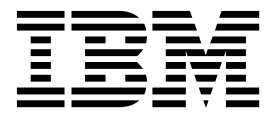


**IBM PowerHA SystemMirror for AIX
Standard Edition**

V7.2.1

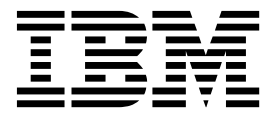
PowerHA SystemMirror 概念



**IBM PowerHA SystemMirror for AIX
Standard Edition**

V7.2.1

PowerHA SystemMirror 概念



注意：

在使用本资料及其支持的产品前，请阅读第 97 页的『声明』中的信息。

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
PowerHA SystemMirror 概念	1
PowerHA SystemMirror 概念中的新增内容	1
新功能部件和不再使用的功能部件	1
PowerHA SystemMirror for AIX	1
AIX 的高可用性集群	1
PowerHA SystemMirror 使用 Cluster Aware AIX	3
可用性成本和收益连续体	7
用 AIX 软件增强可用性	7
PowerHA SystemMirror 集群的物理组件	8
PowerHA SystemMirror 的目标：消除计划内停机时间	14
PowerHA SystemMirror 集群节点、网络和脉动信号传递概念	14
节点	14
站点	14
拆分策略	15
合并策略	16
拆分策略和合并策略的仲裁选项	18
隔离策略	18
PowerHA SystemMirror 多站点解决方案	21
集群网络	24
IP 地址接管	27
通过 TCP/IP 和存储区域网络传递脉动信号	28
PowerHA SystemMirror 多点广播	30
PowerHA SystemMirror 资源和资源组	32
识别和保持集群资源可用	32
集群资源类型	33
集群资源组	35
资源组策略和属性	37
资源组依赖关系	39
站点和资源组	42

支持 PowerHA SystemMirror 的硬件	42
PowerHA SystemMirror 集群软件	43
PowerHA SystemMirror 节点的软件组件	43
集群管理器	44
PowerHA SystemMirror 软件组件	44
补充集群软件	50
确保应用程序可用性	50
概述：应用程序可用性	50
消除 PowerHA SystemMirror 集群中的单点故障	51
用 PowerHA SystemMirror 使计划内停机时间最短	60
使计划外停机时间最短	65
使接管时间最短	66
最大限度地进行灾难恢复	66
集群事件	67
PowerHA SystemMirror 集群配置	69
备用配置	70
接管配置	72
具有多层应用程序的集群配置	76
具有资源组位置依赖关系的集群配置	78
用于灾难恢复的跨站点 LVM 镜像配置	79
具有动态 LPAR 的集群配置	80
PowerHA SystemMirror 配置过程和设施	83
您提供给 PowerHA SystemMirror 的信息	83
由 PowerHA SystemMirror 发现的信息	84
集群配置选项：标准和扩展	84
集群安全性	85
安装、配置和管理工具	85
监视工具	89
故障诊断工具	94
集群测试工具	96

声明	97
隐私策略注意事项	98
商标	99

索引	101
---------------------	------------

关于本文档

此文档介绍了 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

本文档中使用了下列重要约定：

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 V7.2.1 PDF 文档在以下主题提供：PowerHA SystemMirror 7.2.1 PDFs。
- PowerHA SystemMirror V7.2.1 发行说明在以下主题提供：PowerHA SystemMirror 7.2.1 release notes。

PowerHA SystemMirror 概念

以下信息介绍了必须在使用针对 AIX 操作系统的 PowerHA SystemMirror 软件前理解的重要概念。

PowerHA SystemMirror 概念中的新增内容

阅读"PowerHA SystemMirror 概念"主题集合的新增信息或显著更改的信息。

如何查看新增或更改的内容

在此 PDF 文件中，您可能会在左页边距处看到用于标识新增信息和已更改信息的修订线 (|)。

2016 年 12 月

以下信息是对此主题集所作更新的摘要：

- 添加了下列主题：
 - 第 18 页的『隔离策略』
 - 第 19 页的『活动节点暂停策略』
 - 第 20 页的『磁盘电子篱笆策略』
- 更新了下列主题中的策略信息：
 - 第 15 页的『拆分策略』
 - 第 16 页的『合并策略』
 - 第 18 页的『拆分策略和合并策略的仲裁选项』
- 更新了第 91 页的『集群状态实用程序 (clstat)』主题中有关使用 IPv6 回送地址时可能出现的问题的信息。

新功能部件和不再使用的功能部件

有关此发行版中的新功能部件和不再使用的功能部件的列表，请参阅 IBM® 产品声明材料。还可以在与产品软件一起安装的 `/usr/es/sbin/cluster/WhatsNewInThisRelease` 文件中找到附加信息。

PowerHA SystemMirror for AIX

以下信息将讨论高可用性和集群的概念，提供 PowerHA SystemMirror 集群图，并从功能的角度描述 PowerHA SystemMirror 集群。

AIX 的高可用性集群

IBM PowerHA SystemMirror 软件提供了低成本的商业计算环境，可确保任务关键型应用程序从硬件和软件故障中快速恢复。

利用 PowerHA SystemMirror 软件，关键资源将保持可用。例如，PowerHA SystemMirror 集群可以运行服务客户机应用程序的数据库服务器程序。客户机通过访问存储在共享外部磁盘上的数据库，来向响应客户机请求的服务器程序发送查询请求。

当系统、组件或应用程序发生故障时，该高可用性系统将定制软件与业界标准的硬件结合，通过快速恢复服务使停机时间最短。尽管不是瞬时的，服务仍然可以快速恢复，通常在 30 到 300 秒内。

在 PowerHA SystemMirror 集群中，要确保这些应用程序的可用性，可使 PowerHA SystemMirror 控制应用程序。PowerHA SystemMirror 将采取措施来确保即使集群中的组件发生故障，应用程序仍然可对客户机进程可用。要确保可用性，在组件发生故障的情况下，PowerHA SystemMirror 会将应用程序（与确保可对应用程序进行访问的资源一起）移到集群的其他节点上。

高可用性和硬件可用性

高可用性有时与简单的硬件可用性混淆。故障容错系统、冗余系统（例如 RAID）和动态切换技术（例如 DLPAR）提供某些硬件故障的恢复，但是不提供保持复杂应用程序高度可用所需的完整范围的错误检测和恢复。

当今的复杂应用程序需要对所有这些组件的访问权：

- 节点（CPU、内存）
- 网络接口（包括网络拓扑中的外部设备）
- 磁盘或存储设备。

最近的停机时间原因调查显示实际的硬件故障只在意外中断中占很小比例。导致意外中断的其他因素包括：

- 操作员错误
- 环境问题
- 应用程序和操作系统错误。

可靠和可恢复的硬件根本无法防止所有这些不同方面的配置故障。使这些各种各样的元素（从而使应用程序）保持高度可用需要执行下列操作：

- 对应用程序依赖的资源进行访问和操作的物理和逻辑过程进行彻底和全面的规划。这些过程有助于从一开始就避免故障。
- 实现自动检测和从错误中自动恢复的监视和恢复软件包。
- 用于在保持应用程序可用的同时，对集群配置进行硬件和软件方面维护的经过严格控制的流程。

高可用性与容错

容错与高可用性的区别是：容错环境不会造成服务中断，但成本明显较高，而高度可用的环境会有最短的服务中断。

容错依赖专业硬件来检测硬件故障并立即切换到冗余硬件组件 - 无论发生故障的组件是处理器、内存板、电源、I/O 子系统还是存储子系统。尽管这种移交显然是无缝的且提供不停止的服务，但在硬件成本和性能上都要支付高额的费用，因为冗余组件不进行处理。更重要的是，容错模型不会解决软件故障，这是迄今为止造成停机的最常见原因。

高可用性不将可用性视为一系列重复的物理组件，而是视为系统范围内的一组互相协作以保证提供关键服务的共享资源。当系统、组件或应用程序发生故障时，高可用性将软件与业界标准的硬件结合，通过快速恢复关键服务使停机时间最短。尽管不是即刻恢复服务，但是可以快速恢复，通常不超过一分钟。

很多站点都愿意用高可用性承担一点停机时间，而不是支付提供容错所需的高得多的成本。另外，在大多数高度可用的配置中，正常操作期间可以使用备份处理器。

高可用性系统对于必须快速恢复且在发生故障时可以承受短暂中断的应用程序来说是极好的解决方案。一些行业所拥有的应用程序时间至关重要，所以他们无法承受甚至几秒钟的停机时间。但是，很多其他行业却可以承受短暂的数据库不可用的时间段。对于那些行业，PowerHA SystemMirror 可以在无需全冗余的情况下提供必要的服务连续性。

PowerHA SystemMirror 的角色

PowerHA SystemMirror 有很多优点。

PowerHA SystemMirror 为您提供以下帮助：

- PowerHA SystemMirror 规划过程和文档包含了一些有关安装和维护高度可用的 PowerHA SystemMirror 集群的最佳实践的提示和建议。
- 如果集群在操作中，那么 PowerHA SystemMirror 将为应用程序所依赖的所有资源提供自动监视和恢复。
- PowerHA SystemMirror 提供了一个完整的工具集，用于在保持应用程序对客户机可用的同时维护集群。

PowerHA SystemMirror 允许您：

- 通过使用典型的初始集群配置 SMIT 路径或者应用程序配置助手 (Smart Assist) 来轻松快捷地设置基本的双节点集群。
- 通过使用集群测试工具来测试 PowerHA SystemMirror 配置。您可以评估在一组指定情况下（例如，当节点不可访问、网络不可访问等情况时）集群如何运行。
- 通过消除 PowerHA SystemMirror 环境中的单点故障来确保应用程序的高可用性。
- 利用 AIX 中提供的高可用性功能。
- 管理集群如何处理组件故障。
- 确保集群通信安全。
- 监视 PowerHA SystemMirror 组件并诊断可能发生的问题。

相关信息：



PowerHA SystemMirror for AIX

应用程序集群

应用程序集群是一组联网在一起的松散耦合的机器，它们共享磁盘资源。

在集群中，多个服务器机器相互协作以向客户机提供一组服务或资源。

使两个或更多服务器集群来备份关键应用程序是个非常划算的高可用性选择。您可在确保关键应用程序在经历由硬件或软件故障造成的短暂中断后恢复操作的同时，使用站点中的更多计算能力。

应用程序集群还可提供渐进的可扩展的增长路径。将处理器添加到集群以共享不断增长的工作负载非常容易。还可将集群中的一个或多个处理器升级为更强大的模型。如果您使用的是容错策略，那么必须添加两个处理器，一个作为在正常操作期间不执行处理的冗余备份。

PowerHA SystemMirror 使用 Cluster Aware AIX

PowerHA SystemMirror 基于在 AIX 操作系统中受支持的核心集群功能。所有支持 Cluster Aware AIX (CAA) 功能的 AIX 版本都支持 PowerHA SystemMirror。

CAA 和 PowerHA SystemMirror 使用通用标识 (UID 和 UUID) 来跟踪磁盘和节点。不支持动态更改 UID 和 UUID。在大多数情况下，UID 和 UUID 通常是不变量。但是，存在 UID 和 UUID 可以更改的已知情况，例如，重新安装操作系统。如果对 UID 和 UUID 进行更改，那么必须移除并重新创建 CAA 集群，以确保更新所有 UID 和 UUID。

在 AIX V7.2 或更高版本或者带有技术级别 4 的 IBM AIX 7 或更高版本中，CAA 将在 20 秒（缺省值）之后检测到网络故障并处理。要将缺省值从 20 秒进行更改，请运行 **clmgr modify cluster NETWORK_FAILURE_DETECTION_TIME=<xxx>** 命令，其中 xxx 是秒数，在 5 - 590 范围内。

以下信息是有关 Cluster Aware AIX 的关键组件的信息，这些组件可用作构建 PowerHA SystemMirror 解决方案堆栈的基础：

脉动信号管理

缺省情况下，PowerHA SystemMirror 使用单点广播通信来传递脉动信号。另一种方法是可以配置多点广播通信，而不配置单点广播。对于多点广播，您可以选择多点广播地址，或者让 Cluster Aware AIX (CAA) 自动指定多点广播地址。可在配置集群时指定多点广播地址，或者在基于网络环境进行配置期间由 Cluster Aware AIX (CAA) 设置多点广播。集群通信可通过在多个冗余的通信路径中通信来实现。以下冗余的通信路径提供了不容易集群分区的强健的集群基础：

TCP/IP 网络

PowerHA SystemMirror 和 Cluster Aware AIX 使用对集群通信可用的所有网络接口。缺省情况下所有这些接口都可被发现，且都可用于运行状况管理和其他集群通信。可使用 PowerHA SystemMirror 管理接口除去您不希望用于应用程序可用性的任何接口。您还可以使用 PowerHA SystemMirror 定义不想用作私有接口的接口。

基于 SAN 的通信

对于有限数量的适配器，CAA 支持基于存储区域网络 (SAN) 光纤网的集群通信，其中包括脉动信号传递。此类型的脉动信号传递是可选的，可能并不适用于大多数环境，这是因为需要进行网络分区，以允许使用小型计算机系统接口 (SCSI) 协议将包从一个客户机移动到另一个客户机。

基于中央集群存储库的通信

集群运行状况和其他集群通信通过中央存储库磁盘来实现。PowerHA SystemMirror 7.2.0 或更高版本提供了“自动更换存储库磁盘”(ARR) 功能，会将失败的存储库磁盘自动更换为备份存储库磁盘。仅当您使用 PowerHA SystemMirror 来配置和标识备份存储库磁盘时，ARR 功能才可用。

网络接口故障检测时间

PowerHA SystemMirror 依靠 CAA 来监视和检测网络接口故障以及节点故障。在带有技术级别 4 的 IBM AIX 7 或更低版本中，CAA 在固定的时间量（5 秒）内检测网络故障。如果在这些版本的 AIX 操作系统中发生硬件故障，那么将立即报告这些故障。此类型的报告称为快速故障检测过程。AIX 操作系统中的此检测和报告过程与 PowerHA SystemMirror V6.1 报告和检测故障的方式不同。在 PowerHA SystemMirror 6.1 中，在全部的网络故障检测时间过去之后，才会声明故障。此过程称为基于放宽故障检测的完全等待时间。

在 AIX V7.2 或更高版本或者带有技术级别 4 的 IBM AIX 7 或更高版本中，可以将 NETWORK_FAILURE_DETECTION_TIME 选项与 **clmgr** 命令配合使用，以设置网络接口的故障检测时间。NETWORK_FAILURE_DETECTION_TIME 选项的缺省值为 20 秒。在 AIX V7.2 或更高版本或者带有技术级别 4 的 IBM AIX 7 或更高版本中，在故障检测时间的完全等待时间段之后，将执行故障检测过程。这些版本的 AIX 操作系统不会使用快速故障检测过程。

要将 NETWORK_FAILURE_DETECTION_TIME 选项的缺省值从 20 秒进行更改，请运行 **clmgr modify cluster NETWORK_FAILURE_DETECTION_TIME=<xxx>** 命令，其中 xxx 是下列其中一个值：

0 如果指定此值，并且集群已同步，那么在 5 秒之后进行网络故障检测，并且使用快速故障检测过程。在带有技术级别 4 的 IBM AIX 7 或更低版本中使用此选项。

5 - 590 秒

如果指定此范围内的值，并且集群已同步，那么将在指定的值之后进行网络故障检测，并且使用完全等待时间过程。

节点故障检测时间

当网络通信和磁盘通信中缺少脉动信号时，PowerHA SystemMirror 和 CAA 可以检测集群中合作伙伴节点的故障。当失去这些通信信道时，将在设置的时间段内启用监视。此监视称为节点故障检测时间。

要配置节点故障检测时间，可以使用下列其中一个选项：

SMIT 要配置节点故障检测时间，请完成下列步骤：

1. 从命令行中，输入 **smit sysmirror**。
2. 在 SMIT 界面中，选择 **Custom Cluster Configurations > CCluster Nodes and Networks > Manage the Cluster > Cluster heartbeat settings**，然后按 Enter 键。
3. 填写所有必填字段，然后按 Enter 键。

命令行

从命令行运行 **clmgr modify cluster HEARTBEAT_FREQUENCY=<v1> GRACE_PERIOD=<v2>** 命令，其中 v1 和 v2 是以秒计的值。

HEARTBEAT_FREQUENCY 选项是节点通信超时值。此值是 CAA 接收来自合作伙伴节点的包将等待的秒数，然后完成该过程中的下一个步骤以确定合作伙伴节点是否发生了故障。HEARTBEAT_FREQUENCY 选项的有效值为 20 - 600 秒。缺省值为 30 秒。HEARTBEAT_FREQUENCY 选项的值必须比用于 NETWORK_FAILURE_DETECTION_TIME 选项的值多 10 秒。

GRACE_PERIOD 选项是 CAA 在对 HEARTBEAT_FREQUENCY 选项指定的时间过去之后将额外等待的时间。GRACE_PERIOD 选项的缺省值为 10 秒。

增强的事件管理

CAA 将生成精细的存储和网络事件，PowerHA SystemMirror 使用这些事件为高可用性管理提供更好的决策功能。

在节点间管理存储

PowerHA SystemMirror 使用 AIX 的存储保护功能，以便在集群中的节点之间更好地进行存储管理。只有配置了本机 AIX 多路径 I/O (MPIO) 的磁盘才支持保护功能。PowerHA SystemMirror 通过增强的并发卷管理方法来管理共享磁盘。

注：对于不能访问拥有的共享卷组的节点，PowerHA SystemMirror 尝试使用 CAA 存储框架保护功能来阻止这些节点访问共享磁盘。此保护功能可防止由于从多个节点无意中访问共享磁盘而造成数据破坏。但是，只有本机 AIX MPIO 才支持 CAA 存储框架保护功能。

相关概念:

第 21 页的『PowerHA SystemMirror 多站点解决方案』

PowerHA SystemMirror 支持站点和特定于站点的策略的不同类型的定义以实现高可用性灾难恢复 (HADR)。您可以在 PowerHA SystemMirror Standard Edition for AIX 和 PowerHA SystemMirror Enterprise Edition for AIX 中定义多个站点。

第 21 页的『PowerHA SystemMirror 链接集群』

对于每个站点位于不同地理位置的情况，最好是使用链接集群。通常，站点之间相隔足够远，以至于它们无法方便地共享公共存储区网络 (SAN)。每个站点都必须具有它自己的活动存储库磁盘以及任何备份存储库。链接集群始终使用单点广播在站点之间进行通信。链接集群是高可用性灾难恢复 (HADR) 的有用部分。

第 23 页的『PowerHA SystemMirror 延伸集群』

对于每个站点都位于同一地理位置的情况，最好是使用延伸集群。通常，所有节点都在公共存储区网络 (SAN) 中。至少，站点之间将共享活动存储库和任何备份存储库，就像资源组中可以在任一站点运行的任何磁盘一样。延伸集群可以支持 LVM 跨站点镜像和 HyperSwap®。

第 11 页的『集群存储库磁盘』

PowerHA SystemMirror 使用共享磁盘作为中央存储库，用来管理集群的配置。集群中的所有节点必须可以使用标准集群部署方法或延伸集群部署方法来访问此磁盘，或者站点中的所有节点必须可以使用链接集群部署方法来访问此磁盘。

『中央配置管理』

集群配置的一个重要元素是集群存储库磁盘。集群存储库磁盘用作集群配置数据的中央存储库。

相关信息:

设置集群存储器通信

中央配置管理

集群配置的一个重要元素是集群存储库磁盘。集群存储库磁盘用作集群配置数据的中央存储库。

必须可从集群中的所有节点访问集群存储库磁盘，对于 PowerHA SystemMirror 集群配置，集群存储库磁盘必须至少为 512 MB，但是不超过 460 GB。它必须由冗余且高度可用的 SAN 配置提供支持。必须为独立磁盘冗余阵列 (RAID) 适当地配置中央存储库磁盘以避免具有单个故障点。

集群存储库磁盘是集群的特殊设备。它具有特殊的特征，并且由 Cluster Aware AIX (CAA) 和 PowerHA SystemMirror 唯一管理。请勿直接使用 LVM 或文件系统命令和接口，因为它们会导致集群存储库磁盘损坏。集群存储库磁盘由 CAA 专门管理。不得在 CAA 环境或 PowerHA SystemMirror 环境外部管理集群存储库磁盘。

集群存储库用于以下目的的集群管理：

- 在存储库磁盘上集中维护和管理 CAA 集群的配置。
- 对于基于集群的通信，保留一部分存储库磁盘以进行节点到节点的脉动信号传递和消息通信。

注：当所有其他形式的通信都失败时，可使用这种形式的通信。

相关概念:

第 3 页的『PowerHA SystemMirror 使用 Cluster Aware AIX』

PowerHA SystemMirror 基于在 AIX 操作系统中受支持的核心集群功能。所有支持 Cluster Aware AIX (CAA) 功能的 AIX 版本都支持 PowerHA SystemMirror。

设置集群 SAN 通信

PowerHA SystemMirror 和 Cluster Aware AIX 支持主机之间使用其他通信方法，此方法称为存储区域网络 (SAN) 通信。

SAN 通信利用存在于典型数据中心的 SAN 物理链路来提供集群节点间的高速通信。如果在集群节点间的网络通信路径中发生故障，那么该通信机制将提供回退功能。

SAN 通信可用于一组磁盘适配器。您必须使用 SAN 分区设置，以使小型计算机系统接口 (SCSI) 数据包通过每个主机间的多个交换机从主机流向主机。

设置 SAN 通信涉及到在受支持磁盘适配器上调用一组命令以使其可以进行 SAN 通信。

在 VIOS 2.2.0.11 或更高版本中，您可以通过每个 VIOS 客户机上的虚拟以太网适配器来建立虚拟局域网，从而在逻辑分区之间使用 SAN 通信。您可以通过 VIOS 对 NPIV 和 vSCSI 环境建立 SAN 通信。

相关信息:

设置集群存储器通信

定义虚拟以太网适配器

可用性成本和收益连续体

PowerHA SystemMirror 可用性有很多成本和收益。

下图显示了可用性技术的成本和收益。

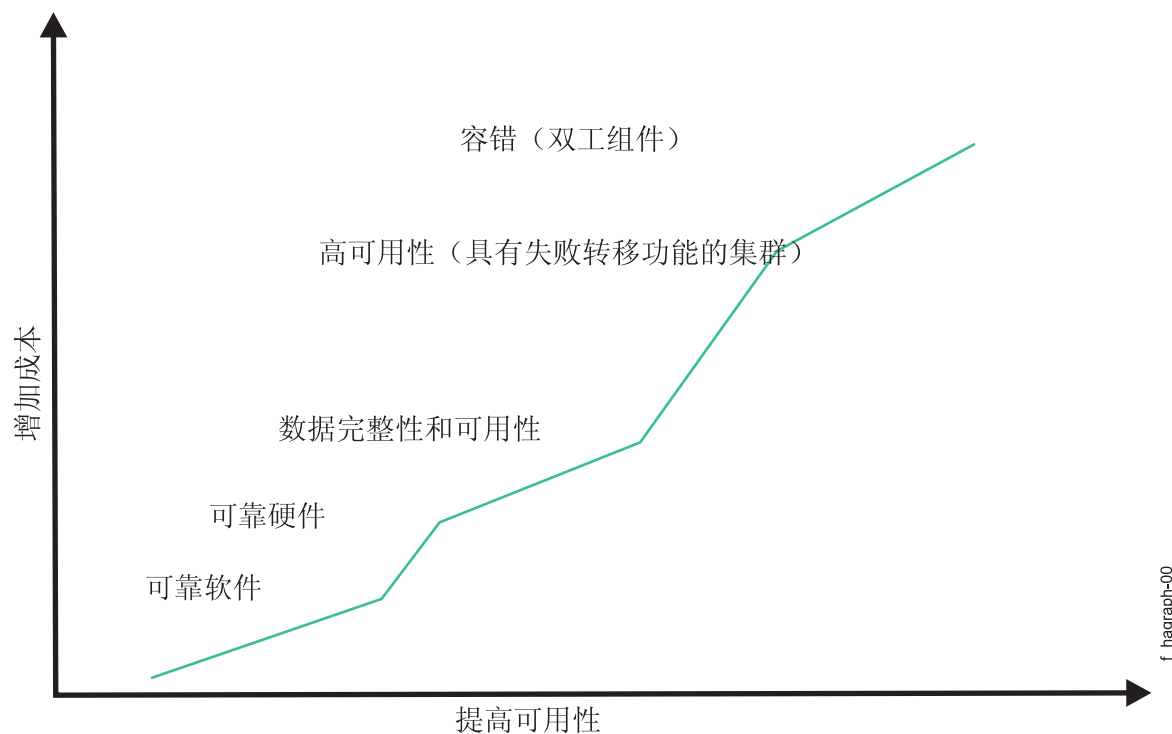


图 1. 可用技术的成本和收益

正如您所看到的，可用性不是一个非此即彼的命题。将可用性看作一个连续体。可靠的硬件和软件提供基本级别的可用性。高级功能部件（例如，RAID 设备）提供增强级别的可用性。高可用性软件提供对数据 and 应用程序几乎连续的访问。容错系统确保整个系统稳定的可用性，但成本却较高。

用 AIX 软件增强可用性

PowerHA SystemMirror 利用了 AIX 中的功能部件 - 高性能的 UNIX 操作系统。AIX 添加了新的功能以进一步提高安全性和系统可用性。

这包括提高的镜像数据可用性以及对 Workload Manager 的增强，其中 Workload Manager 可以通过动态地提供对关键应用程序的资源可用性来帮助解决混合负载的问题。PowerHA SystemMirror 可提供水平可伸缩性和垂直可伸缩性而无需关闭系统。AIX 操作系统提供了大量功能部件，旨在通过减小计划内（数据备份、系统管理）和计划外（硬件或软件故障）停机时间的影响来增加系统可用性并通过使用 Capacity on Demand (CoD) 功能来增加硬件资源管理的灵活性。

AIX 操作系统提供以下功能：

- 日志文件系统和增强型日志文件系统
- 磁盘镜像
- 进程控制
- DLPAR 和 CoD
- 工作负载分区

日志文件系统和增强型日志文件系统

AIX 本机文件系统，日志文件系统 (JFS)，使用数据库日志记录技术来维护其结构完整性。系统或软件故障不会使文件系统无法管理。在重大故障后重新构建文件系统时，AIX 将使用 JFS 日志将文件系统恢复到它最后的一致状态。因此，与标准 UNIX 文件系统一致性检查 (fsck) 实用程序相比，日志记录可提供更快速的恢复。另外，AIX 提供了增强型日志文件系统 (JFS2)。

磁盘镜像

磁盘镜像软件可提供数据完整性和联机备份能力。它可预防磁盘故障导致的数据丢失，其方法是通过分别在不同的磁盘上维护最多三份数据副本，从而使数据在任何单个磁盘发生故障后仍然可用。磁盘镜像对应用程序是透明的。因为镜像和常规磁盘对应用程序来说是相同的，所以无需进行应用程序修改。

进程控制

AIX System Resource Controller (SRC) 监视并控制关键进程。SRC 可检测进程何时异常终止，记录终止，向通知程序传递消息并在备份处理器上重新启动该进程。

动态 LPAR 管理

PowerHA SystemMirror 可以在 LPAR 之间移动应用程序资源，可以通过资源优化高可用性 (ROHA) 功能来执行必需的动态资源调整。ROHA 使用 IBM Power Systems™ 服务器提供的功能来动态管理以下类型的硬件资源：

- Capacity on Demand (CoD) 功能（包括 On/Off CoD 和 Enterprise Pool CoD）在框架 (CEC) 级别管理内存和 CPU 资源。
- DLPAR 功能在逻辑分区级别管理内存和 CPU 资源。

通过 RPHA 功能，可以使用 PowerHA SystemMirror 来优化每个应用程序的大量资源。例如，在接管期间，会从活动节点动态释放硬件资源（CPU 和内存），并由备用节点动态获取和分配。

工作负载分区

工作负载分区允许多个应用程序实例在相同的操作系统环境内运行，同时提供这些环境间的隔离，从而提供实例间的保护和隔离。

Cluster Aware AIX (CAA)

PowerHA SystemMirror 可使用 AIX 集群能力来提供高级的高可用性解决方案。CAA 为 AIX 内核提供了脉动信号传递和运行状况管理。PowerHA SystemMirror 监视精细的存储和网络事件，并处理 AIX 操作系统中的重要情况。PowerHA SystemMirror 可发现硬件组件，因此使管理和部署集群更加容易。

PowerHA SystemMirror 集群的物理组件

PowerHA SystemMirror 通过识别一组对不中断处理至关重要的资源来提供高度可用的环境。它还定义了一个协议，节点使用该协议相互协作以确保这些资源可用。

PowerHA SystemMirror 通过在相互协作的处理器间定义关系来扩展集群模型，如果其中一个处理器无法提供服务，那么对等处理器将提供该处理器所提供的服务。如下图所示，PowerHA SystemMirror 集群由以下物理组件组成：

- 节点
- 共享外部磁盘设备

- 网络
- 网络接口
- 集群存储库磁盘
- 客户机。

PowerHA SystemMirror 软件允许您将物理组件组合到范围广泛的集群配置中，在构建满足处理需求的集群方面为您提供灵活性。该图显示了 PowerHA SystemMirror 集群的一个示例。其他 PowerHA SystemMirror 集群看起来可能大不一样 - 取决于处理器数目、网络的选择和磁盘技术等。

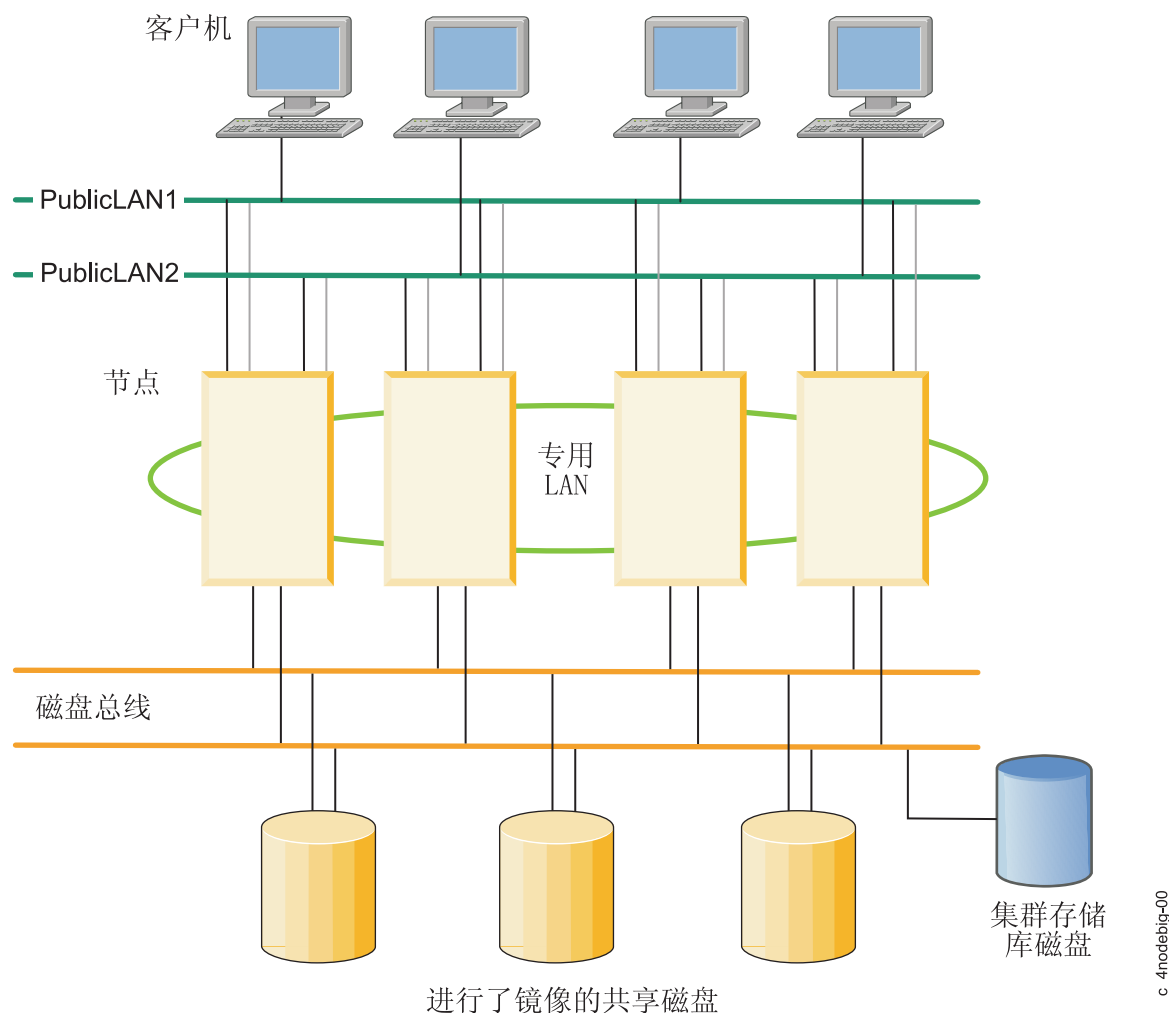


图 2. PowerHA SystemMirror 集群的示例

PowerHA SystemMirror 集群配置提供了 PowerHA SystemMirror 软件支持的集群配置的示例。

相关概念:

第 69 页的『PowerHA SystemMirror 集群配置』

本章提供了由 PowerHA SystemMirror 软件支持的多种集群配置类型的示例。

PowerHA SystemMirror 节点

节点构成 PowerHA SystemMirror 集群的核心。节点是运行 AIX、PowerHA SystemMirror 软件和应用软件的处理器。

使这些服务器集群来备份关键应用程序是个非常划算的高可用性选择。业务可在确保其关键应用程序在经历由硬件或软件故障造成的短暂中断后恢复运行的同时，使用更多计算能力。

在 PowerHA SystemMirror 集群中，每个节点都由一个唯一的名称识别。节点可拥有一组资源：磁盘、卷组、文件系统、网络、网络地址和应用程序。通常，节点运行服务器或可访问共享外部磁盘上数据的后端应用程序。

相关概念:

第 50 页的『确保应用程序可用性』

本节描述了 PowerHA SystemMirror 软件如何通过确保集群组件的可用性来确保应用程序可用性。

PowerHA SystemMirror 可消除所有关键系统组件的单点故障，并消除针对大部分例程集群维护任务的计划内停机时间需求。

第 32 页的『PowerHA SystemMirror 资源和资源组』

请阅读本节以了解整个文档以及 PowerHA SystemMirror 用户界面中使用的与资源相关的概念和定义。

共享外部磁盘设备

每个节点都可访问一个或多个共享外部磁盘设备。共享外部磁盘设备是物理连接到多个节点的磁盘。

该共享磁盘存储任务关键型数据，通常对该数据进行镜像或 RAID 配置以使数据冗余。PowerHA SystemMirror 集群中的节点还必须具有存储操作系统和应用程序二进制的内部磁盘，但是这些磁盘不共享。根据使用的磁盘类型，PowerHA SystemMirror 软件支持对共享外部磁盘设备的以下类型的访问 - 非并发访问和并发访问。

- 在非并发访问环境中，在任何给定的时间只有一个连接是活动的，且具有该活动连接的节点拥有磁盘。当节点发生故障时，当前拥有磁盘的节点将离开集群，发生磁盘接管且剩余节点承担共享磁盘的所有权。这种典型的集群配置用于大部分应用程序。
- 在并发访问环境中，共享磁盘主动地同时连接到多个节点。因此，当节点发生故障时不需要磁盘接管。这种访问类型仅由可管理和协调从多个节点对共享数据进行同时访问的应用程序所使用。

注：在这样的环境中，要么集群中定义的所有节点都可以是并发访问的一部分，要么仅部分集群节点可以访问共享磁盘。在第二种情况下，仅可将资源组配置在那些具有共享磁盘访问权的节点上。这些方法间的区别在"共享外部磁盘访问"这一节中有更全面的解释。

相关概念:

第 48 页的『共享外部磁盘访问』

PowerHA SystemMirror 软件支持两种共享外部磁盘访问的方法：非并发访问和并发访问。

网络

作为 AIX 操作系统的独立分层组件，PowerHA SystemMirror 软件设计为与任何基于 TCP/IP 的网络配合使用。

PowerHA SystemMirror 集群中的节点使用网络以：

- 允许客户机访问集群节点
- 使集群节点能够交换脉动信号消息
- 使对（并发访问环境中的）数据的访问序列化

已使用以太网和存储区域网络对 PowerHA SystemMirror 软件进行测试。

网络类型

PowerHA SystemMirror 软件定义两种类型的通信网络，特点是这些网络使用基于 TCP/IP 子系统（基于 TCP/IP）网络的通信接口，还是使用基于存储区域网络的通信接口。

基于 TCP/IP 的网络

连接两个或多个服务器节点，并选择性地允许客户机使用 TCP/IP 协议访问这些集群节点。缺省情况下，PowerHA SystemMirror 使用单点广播通信来传递脉动信号。（可选）如果网络配置为支持多点广播，那么可以选择使用多点广播通信。

存储区域网络 (SAN)

在 PowerHA SystemMirror 集群节点间提供通信，以控制消息和脉动信号流量。这是 TCP/IP 网络的备用通信路径。

集群存储库磁盘

PowerHA SystemMirror 使用共享磁盘作为中央存储库，用来管理集群的配置。集群中的所有节点必须可以使用标准集群部署方法或延伸集群部署方法来访问此磁盘，或者站点中的所有节点必须可以使用链接集群部署方法来访问此磁盘。

使用系统管理界面工具 (SMIT) 界面来管理和配置集群存储库磁盘和备份存储库磁盘。

对于标准集群和延伸集群，每个集群必须至少具备一个活动的存储库磁盘。对于标准集群和延伸集群，可以对每个集群标识多达六个备份存储库磁盘。对于链接集群，每个站点必须有一个活动的存储库磁盘。对于链接集群，可以对每个站点标识多达六个备份存储库磁盘。

借助 Cluster Aware AIX，集群存储库磁盘用于下列用途：

- 集群范围的配置管理
- 集群消息传递和脉动信号传递。存储库磁盘用作节点之间进行通信的另一种冗余方法。

您必须为集群存储库磁盘至少分配 512 MB，但是不超过 460 GB 的磁盘空间。

请验证您选择作为存储库磁盘的磁盘不具备预留策略。要检查磁盘预留策略，请运行 **lsattr -El hdisk - a reserve_policy** 命令。要将磁盘预留策略更改为 no_reserve，请运行 **chdev -a reserve_policy=no_reserve -l hdisk** 命令。

将磁盘指定为存储库磁盘之后，无法将该磁盘用于任何其他用途。请验证您用作存储库磁盘的磁盘不包含任何用户数据。当 PowerHA SystemMirror 将磁盘标识为存储库磁盘时，会擦除该磁盘上的所有信息。

当您计划将磁盘用作存储库磁盘时，必须计划备份磁盘或更换磁盘，可供主存储库磁盘发生故障时使用。备份磁盘必须与主磁盘具有相同大小和类型，但是可以位于不同的物理存储器磁盘中。使用备份磁盘信息来更新管理过程和文档。集群可以容许缺少存储库磁盘或者存储库磁盘故障。没有存储库磁盘，关键操作仍会继续运行。但是，您必须快速解决存储库磁盘存在的问题。如果您正在使用 PowerHA SystemMirror 7.2.0 或更低版本，那么您必须将失败的存储库磁盘手动更换为备份列表中的存储库磁盘。您还可以将正在工作的存储库磁盘更换为新磁盘，以增大大小或者更改为另一存储子系统。

可以配置 PowerHA SystemMirror 7.2.0 或更高版本，以在活动存储库磁盘发生故障期间使用 Cluster Aware AIX (CAA) 中的“自动更换存储库磁盘”(ARR) 功能。ARR 功能会将失败的存储库磁盘自动更换为备份存储库磁盘中的磁盘。该列表中的第一个备份存储库磁盘将更换失败的存储库磁盘。

您应该对所有存储库磁盘配置 AIX 本机多路径 I/O (MPIO)。AIX MPIO 使 PowerHA SystemMirror 能够监视和更好地响应使用存储库磁盘时产生的磁盘故障。

镜像池和存储库磁盘

如果您使用 PowerHA SystemMirror 和逻辑卷管理器 (LVM) 来管理位于不同位置的镜像池和存储库磁盘故障，那么新的存储库磁盘不得来自于发生故障的位置。

相关概念:

第 21 页的『PowerHA SystemMirror 链接集群』

对于每个站点位于不同地理位置的情况，最好是使用链接集群。通常，站点之间相隔足够远，以至于它们无法方便地共享公共存储区网络 (SAN)。每个站点都必须具有它自己的活动存储库磁盘以及任何备份存储库。链接集群始终使用单点广播在站点之间进行通信。链接集群是高可用性灾难恢复 (HADR) 的有用部分。

第 23 页的『PowerHA SystemMirror 延伸集群』

对于每个站点都位于同一地理位置的情况，最好是使用延伸集群。通常，所有节点都在公共存储区网络 (SAN) 中。至少，站点之间将共享活动存储库和任何备份存储库，就像资源组中可以在任一站点运行的任何磁盘一样。延伸集群可以支持 LVM 跨站点镜像和 HyperSwap。

第 3 页的『PowerHA SystemMirror 使用 Cluster Aware AIX』

PowerHA SystemMirror 基于在 AIX 操作系统中受支持的核心集群功能。所有支持 Cluster Aware AIX (CAA) 功能的 AIX 版本都支持 PowerHA SystemMirror。

相关信息:

规划存储库磁盘

将存储库磁盘替换为 SMIT

存储库磁盘和多路径:

在配置集群期间，AIX 操作系统将以一种特殊方式来管理存储库磁盘。

在 AIX 设备配置过程中，AIX 集群服务识别存储库磁盘，并且可以对该磁盘执行特殊操作。例如，集群服务使用存储库磁盘来启用在主机之间进行基于磁盘的通信。磁盘通信涉及到一台写入磁盘的主机以及通过轮询机制来检索消息的其他主机。

集群服务还会生成与存储库磁盘的运行状况有关的详细事件。如果存储库磁盘失败，那么 CAA 会使用“自动更换存储库”(ARR) 功能将失败的存储库磁盘自动更换为备份存储库磁盘。仅当您使用 PowerHA SystemMirror 来标识备份存储库磁盘时，ARR 功能才可用。

您可以将下列类型的磁盘配置为存储库磁盘：

AIX 多路径磁盘 (AIX 本地 MPIO)

可以自动创建这些磁盘并将其用作存储库磁盘。这是建议用作存储库磁盘的磁盘类型。

虚拟 SCSI (vSCSI) 磁盘

可以通过 Virtual I/O Server (VIOS) 将这些磁盘作为 vSCSI 磁盘映射至客户机逻辑分区 (LPAR)。

将 vSCSI 磁盘用作存储库磁盘时，必须使用相同的 vSCSI 方法将它映射至不同的 PowerHA SystemMirror 节点或者不同的 Cluster Aware AIX (CAA) 节点。

第三方多路径磁盘

这些磁盘遵循 AIX 多路径概念的准则，但是提供它们自己的多路径设备驱动程序和软件。当相关的对象数据管理器 (ODM) 信息可用时，可以将这些磁盘配置为存储库磁盘。使用此方法可以将 EMC PowerPath 和 Hitachi Dynamic Link Manager (HDLM) 管理的磁盘配置为存储库磁盘。

要升级 MPIO 软件或者第三方多路径软件，您必须通过输入 **clmgr offline cluster STOP_CAA=yes** 命令来停止 CAA 集群服务。

存储库磁盘故障:

您必须正确计划存储库磁盘故障以及更正与存储库磁盘故障相关的问题所需要执行的操作。

PowerHA SystemMirror 集群容许存储库磁盘故障。如果集群中的任何节点遇到与存储库磁盘有关或者与访问存储库磁盘有关的错误，那么该集群进入受限制方式。在此操作方式下，您无法使用大多数与拓扑相关的操作。例如，无法添加节点，或者节点无法连接集群。但是，可以执行关键的集群功能。例如，可以将资源组从活动节点移到备用节点。

当存储库磁盘失败时，会将磁盘故障通知管理员。PowerHA SystemMirror 会继续将存储库磁盘故障通知管理员，直到解决此故障为止。要在存储库磁盘发生故障时获取通知或者定制事件处理，请参阅配置事件前处理和事件后处理主题。

PowerHA SystemMirror 和 Cluster Aware AIX (CAA) 支持动态更换存储库磁盘，可以使用此功能来更换已失败或者正在运行的存储库磁盘。CAA 会使用集群信息来重新填充该新磁盘，并且开始将该磁盘用作存储库磁盘。

PowerHA SystemMirror 7.2.0 或更高版本支持"自动更换存储库磁盘"(ARR) 功能。ARR 使用 CAA 并将失败的存储库磁盘自动更换为备份存储库磁盘。仅当您使用 PowerHA SystemMirror 来配置备份存储库磁盘时，ARR 功能才可用。

要使用 ARR 功能，您所在环境必须满足下列需求：

- 集群或站点标识了备份存储库磁盘。
- 已安装 PowerHA SystemMirror V7.2.0 或更高版本。
- 已安装下列其中一个版本的 AIX 操作系统：
 - AIX V7.1.4 或更高版本
 - AIX V7.2.0 或更高版本

CAA 通过检查 I/O 错误以及验证磁盘是否处于活动状态来监视存储库磁盘故障。会定期执行这些验证检查，而不是每次读取或写入存储库磁盘时都进行验证检查。即使是进行测试，也不要直接写入存储库磁盘。直接异步写入存储库磁盘可能会导致操作系统和 CAA 操作突然中断，从而导致不可预测的结果。

相关信息：

dd 命令

规划存储库磁盘

将存储库磁盘替换为 SMIT

客户机

客户机是可通过局域网访问集群中节点的处理器。

每个客户机都运行一个"前端"或运行用于查询运行在集群节点上的服务器应用程序的客户机应用程序。PowerHA SystemMirror 软件为集群节点上的关键数据和应用程序提供高度可用的环境。PowerHA SystemMirror 软件不会使客户机本身高度可用。AIX 客户机可使用集群信息 (Clinfo) 服务来接收集群事件的通知。Clinfo 提供显示集群状态信息的 API。PowerHA SystemMirror 提供集群状态实用程序：**/usr/es/sbin/cluster/clstat**。它基于 Clinfo 并报告集群关键组件的状态，这些组件包括：集群本身、集群中的节点、已连接到节点的网络接口和每个节点上的资源组。**clstat** 实用程序的集群状态接口包括基于 Web、基于 Motif 和基于 ASCII 的版本。

相关概念：

第 47 页的『集群信息程序』

集群信息程序 (Clinfo)，**clinfo** 守护程序，是基于 SNMP 的监视器。

PowerHA SystemMirror 的目标：消除计划内停机时间

高可用性集群软件的主要目标是最小化维护和重新配置活动期间使资源脱离服务的需求，或者理想情况下是消除该需求。

PowerHA SystemMirror 软件通过允许动态重新配置正在运行的集群来优化可用性。大多数例程集群维护任务（例如，添加或删除节点或更改加入资源组的节点的优先级）都可在无需停止并重新启动集群服务的情况下应用到活动的集群。此外，在更改配置时，可以使用“集群单点控制 (C-SPOC)”功能使 PowerHA SystemMirror 集群保持联机。C-SPOC 使得更容易管理集群，因为您可以从单个节点对整个集群中的共享卷组、用户和组进行更改。这些更改可透明地传播到其他集群节点。

PowerHA SystemMirror 集群节点、网络 and 脉动信号传递概念

本节介绍了与集群拓扑相关的主要概念，以及在整个文档和 PowerHA SystemMirror 用户界面中使用的定义。

节点

节点是同时运行 AIX 和 PowerHA SystemMirror 软件的处理器。

节点可以共享一组资源，例如，磁盘、卷组、文件系统、网络、网络 IP 地址和应用程序。PowerHA SystemMirror 软件在一个集群中最多支持 16 个节点。在 PowerHA SystemMirror 集群中，每个节点都由一个唯一的名称识别。在 PowerHA SystemMirror 中，节点名和主机名通常是一样的。节点作为 PowerHA SystemMirror 集群的核心物理组件。有关节点和硬件的更多信息，请参阅“节点”这一节。定义了两种类型的节点：

- 服务器节点构成 PowerHA SystemMirror 集群的核心。服务器节点运行服务器或可访问共享外部磁盘上数据的后端应用程序。
- 客户机节点运行从服务器节点所提供的服务中检索数据的前端应用程序。客户机节点可运行 PowerHA SystemMirror 软件以监视节点的运行状况，并对故障作出反应。

服务器节点

集群服务器节点通常运行可访问共享外部磁盘上数据的应用程序。服务器节点运行 PowerHA SystemMirror 守护程序并保持资源高度可用。通常，应用程序运行，存储器在这些节点间共享，并且客户机通过服务 IP 地址连接到服务器节点。

客户机节点

完整的高可用性解决方案通常包括使用服务器所提供的服务的客户机机器。客户机节点可分为两类：简单型和智能型。

- 简单型客户机将集群看作单个实体。如果服务器发生故障，那么必须重新启动客户机，或者至少必须重新连接至服务器。
- 智能型客户机可以集群感知。集群感知的客户机可对服务器故障作出适当的反应，连接到备用服务器，也许还对用户掩饰故障。这样的智能型客户机必须了解集群状态。

PowerHA SystemMirror 通过提供动态集群配置报告和集群状态更改的通知（例如，子系统更改或节点故障），以向客户机扩展集群模型。

站点

您可以将由一个或多个服务器节点组成的组定义为属于某个站点。

站点成为 PowerHA SystemMirror 软件知道的一个组件（例如，节点或网络）。PowerHA SystemMirror 支持划分为两个站点的集群。

可以通过使特定 LVM 镜像池与物理站点相关联来配置拆分站点逻辑卷管理器 (LVM) 镜像。如果您对 PowerHA SystemMirror 指定哪些物理卷位于每个站点中，那么从在站点中选择用于 LVM 镜像池的卷时，C-SPOC 将显示站点信息。进行集群验证期间，PowerHA SystemMirror 将执行其他检查以确保镜像定义与站点定义一致。

此外，PowerHA SystemMirror Enterprise Edition 的“延伸距离”功能为灾难恢复提供了两种截然不同的软件解决方案。这些解决方案使 PowerHA SystemMirror 集群能够在两个相隔更远距离的站点上运行。

PowerHA SystemMirror Enterprise Edition for Metro Mirror 提高了使用 Peer-to-Peer Remote Copy (PPRC) 将数据复制到远程站点以实现灾难恢复的 DS8000® 卷、DS6000™ 卷和 IBM TotalStorage Enterprise Storage Server® (ESS) 卷的数据可用性。PowerHA SystemMirror Enterprise Edition for Metro Mirror 利用 PPRC 失败转移和回退功能以及 PowerHA SystemMirror 集群管理来缩短进行灾难恢复期间的停机时间和恢复时间。

将 PPRC 用于站点之间的数据镜像时，站点之间的物理距离受到 ESS 硬件功能的限制。

PowerHA SystemMirror Enterprise Edition for Geographic Logical Volume Manager (GLVM) 提高了使用 GLVM 将数据复制到远程站点以实现灾难恢复的 IBM 卷的数据可用性。PowerHA SystemMirror Enterprise Edition for GLVM 利用下列组件来缩短进行灾难恢复期间的停机时间和恢复时间：

AIX and PowerHA SystemMirror Enterprise Edition for GLVM 数据镜像和同步。可以使用 GLVM 实用程序从地理位置上镜像标准卷组和增强的并发卷组。

基于 TCP/IP 的距离不受限制的网络支持最多可以配置四个 XD_data 数据镜像网络。

如果某个组件失败，那么 PowerHA SystemMirror 将确保在本地站点或远程站点都可以访问数据的镜像副本。可以在使用 GLVM 的 PowerHA SystemMirror 集群中配置并发资源组和非并发资源组。但是，无法同时使用站点间策略。

相关信息：

规划指南

高速镜像：规划和管理

地理 LVM 规划和管理

拆分策略

当一组节点无法与集群中的其余节点进行通信时，站点之间可能会发生集群拆分事件。例如，在链接集群中，如果两个站点之间的所有通信链路都失败，那么将进行拆分。集群拆分事件会将集群拆分为两个或两个以上的分区。

可以使用 PowerHA SystemMirror 来配置用于指定对集群拆分事件的响应的拆分策略。

下列选项可用于配置拆分策略：

无 此选项指示在检测到集群拆分事件时不执行操作。由集群拆分事件创建的每个分区都将成为独立集群。每个分区可以独立于其他分区启动工作负载。如果正在使用共享卷组，那么可能会导致数据被破坏。此选项是缺省设置，因为需要进行手动配置来建立备用策略。如果您所在环境配置为使用 HyperSwap for PowerHA SystemMirror，请勿使用此选项。

I 仲裁 可以使用此选项来指定磁盘或 NFS 文件。

如果您指定磁盘用于仲裁，那么每个分区都会尝试通过锁定仲裁磁盘来获取仲裁磁盘。如果您指定磁盘用于仲裁，那么将使用集群中所有节点都可访问的 SCSI 磁盘。按照操作计划中所指定那样，会重新引导无法锁定磁盘的分区。

如果您指定了 NFS 文件用于仲裁，那么集群中的每个节点上都必须已从所选 NFS 服务器安装了 NFS。第一个保留 NFS 文件的分区将继续起作用。按照操作计划中所指定那样，会重新引导无法锁定 NFS 文件的分区。

手动 此选项指示发生集群拆分事件时，要手动解决问题。

分区中的每个节点提供一条消息，以选择继续运行集群服务或者恢复集群服务（将重新启动该节点）。使用此选项，可以指定需要您的输入的尝试次数和尝试频率。还可以指定在达到需要您的输入的尝试次数、但是您尚未提供任何输入之后要执行的缺省操作。

对于指定了手动选项的链接集群，发生集群拆分事件时，将显示以下消息：

```
04:09:48 来自 root@e08m138.ausprv.stglabs.ibm.com (tty) 的广播报文...
检测到集群拆分事件。
您必须决定分区集群的此端是否要继续。
要让此端继续，请输入
    /usr/es/sbin/cluster/utilities/cl_sm_continue
要在此分区的所有节点上执行"重新引导"恢复操作，请输入
    /usr/es/sbin/cluster/utilities/cl_sm_recover
LOCAL_PARTITION 1 e08m138 OTHER_PARTITION 2 e08m140
```

在此示例中，可以使用手动选项来检查拆分或合并是否正在等待从 **Problem Determination Tools > Manual Response to Split or Merge > Display any needed Manual Response** SMIT 菜单进行手动响应。

如果要将手动选项用于延伸集群和标准集群，那么环境必须正在运行下列版本的软件：

- 带有技术级别 1 的 IBM AIX 7.2 或更高
- PowerHA SystemMirror V7.2.1 或更高

注：对于任何使用手动选项的集群类型，在达到指定的尝试次数之后，如果您尚未提供任何输入，那么将选择节点标识最小的分区作为获胜分区。

相关概念：

『合并策略』

根据集群拆分策略，集群可能具有两个互相独立运行的分区。可以使用 PowerHA SystemMirror 来配置合并策略，合并策略允许分区在恢复分区之间的通信之后再次一起运行。

第 18 页的『拆分策略和合并策略的仲裁选项』

可以使用仲裁选项来指定拆分策略和合并策略所使用的 SCSI 磁盘或网络文件系统 (NFS) 文件。

第 21 页的『PowerHA SystemMirror 链接集群』

对于每个站点位于不同地理位置的情况，最好是使用链接集群。通常，站点之间相隔足够远，以至于它们无法方便地共享公共存储区网络 (SAN)。每个站点都必须具有它自己的活动存储库磁盘以及任何备份存储库。链接集群始终使用单点广播在站点之间进行通信。链接集群是高可用性灾难恢复 (HADR) 的有用部分。

相关信息：

配置拆分策略和合并策略

合并策略

根据集群拆分策略，集群可能具有两个互相独立运行的分区。可以使用 PowerHA SystemMirror 来配置合并策略，合并策略允许分区在恢复分区之间的通信之后再次一起运行。

下列选项可用于配置合并策略：

大多数

具有最大节点数的分区将保持联机。如果每个分区的节点数相同，那么将选择节点标识最小的分区。按照所选择的操作计划指定那样，会重新引导不保持联机的分区。要使延伸集群能够使用"大多数"选项，环境必须正在运行下列其中一个版本的 AIX 操作系统：

- 带有技术级别 4 的 IBM AIX 7 或更高
- AIX V7.2 或更高

仲裁

可以将磁盘或 NFS 文件用于仲裁。如果您将磁盘用于仲裁，那么每个分区将尝试通过对仲裁磁盘施加锁定来获取仲裁。仲裁是集群中的所有节点都可以访问的 SCSI 磁盘。按照所选择的操作计划指定那样，会重新引导无法锁定磁盘的分区，或者重新启动集群服务。

如果您将 NFS 文件用于仲裁，那么在所选择的 NFS 服务器中，集群中的每个节点上都必须进行 NFS 安装。第一个保留 NFS 文件的分区将继续起作用。按照所选择的操作计划指定那样，会重新引导无法锁定 NFS 文件的分区，或者重新启动集群服务。如果您使用此选项，那么拆分策略配置也必须使用该仲裁选项。

手动

此策略选项要求您在发生合并事件期间选择获胜站点。

分区中的每个节点提供一条消息，以选择继续运行集群服务或者恢复集群服务（将重新启动该节点）。使用此选项，可以指定需要您的输入的尝试次数和尝试频率。还可以指定在达到需要您的输入的尝试次数、但是您尚未提供任何输入之后要执行的缺省操作。

对于指定了手动选项的链接集群，发生集群拆分事件时，将显示以下消息：

```
04:09:48 来自 root@e08m138.ausprv.stglabs.ibm.com (tty) 的广播报文...
```

```
检测到集群拆分事件。
```

```
您必须决定分区集群的此端是否要继续。
```

```
要让此端继续，请输入
```

```
    /usr/es/sbin/cluster/utilities/cl_sm_continue
```

```
要在此分区的所有节点上执行"重新引导"恢复操作，请输入
```

```
    /usr/es/sbin/cluster/utilities/cl_sm_recover
```

```
LOCAL_PARTITION 1 e08m138 OTHER_PARTITION 2 e08m140
```

在此示例中，可以使用手动选项来检查拆分或合并是否正在等待从 **Problem Determination Tools > Manual Response to Split or Merge > Display any needed Manual Response** SMIT 菜单进行手动响应。

如果要将手动选项用于延伸集群和标准集群，那么环境必须正在运行下列版本的软件：

- 带有技术级别 1 的 IBM AIX 7.2 或更高
- PowerHA SystemMirror V7.2.1 或更高

注：对于任何使用手动选项的集群类型，在达到指定的尝试次数之后，如果您尚未提供任何输入，那么将选择节点标识最小的分区作为获胜分区。

| 无 此选项指示在发生集群合并事件时不执行操作。要避免在发生合并事件之后毁坏任何数据，必须手动重新引导失败的分区节点。只能在 **clmgr** 命令中使用此选项。如果对合并策略指定"无"，那么必须对拆分策略选择"无"。

| 如果要将"无"选项用于延伸集群和标准集群，那么环境必须正在运行下列版本的软件：

- |
- 带有技术级别 1 的 IBM AIX 7.2 或更高
 - PowerHA SystemMirror V7.2.1 或更高
- |

相关概念：

第 15 页的『拆分策略』

当一组节点无法与集群中的其余节点进行通信时，站点之间可能会发生集群拆分事件。例如，在链接集群中，如果两个站点之间的所有通信链路都失败，那么将进行拆分。集群拆分事件会将集群拆分为两个或两个以上的

分区。

『拆分策略和合并策略的仲裁选项』

可以使用仲裁选项来指定拆分策略和合并策略所使用的 SCSI 磁盘或网络文件系统 (NFS) 文件。

第 21 页的『PowerHA SystemMirror 链接集群』

对于每个站点位于不同地理位置的情况，最好是使用链接集群。通常，站点之间相隔足够远，以至于它们无法方便地共享公共存储区网络 (SAN)。每个站点都必须具有它自己的活动存储库磁盘以及任何备份存储库。链接集群始终使用单点广播在站点之间进行通信。链接集群是高可用性灾难恢复 (HADR) 的有用部分。

相关信息:

配置拆分策略和合并策略

拆分策略和合并策略的仲裁选项

可以使用仲裁选项来指定拆分策略和合并策略所使用的 SCSI 磁盘或网络文件系统 (NFS) 文件。

当集群中的站点不再能够互相通信时，将使用仲裁磁盘或 NFS 文件。此通信故障会导致集群将站点拆分为两个独立分区。如果由于集群通信链路未响应而发生故障，那么两个分区都将尝试锁定仲裁磁盘或 NFS 文件。获得仲裁磁盘的分区将继续起作用，而另一个分区将重新引导或者重新启动集群服务，视所选择的操作计划而定。

确定为仲裁的磁盘或者通过 NFS 安装的文件必须可供集群中的所有节点访问。

当集群中包含的分区在发生通信故障之后重新联机时，这些分区必须能够与拥有仲裁磁盘或者 NFS 文件的分区进行通信。如果重新联机的分区无法与仲裁磁盘或 NFS 文件进行通信，那么它不会加入该集群。当配置中的所有节点重新加入集群时，会释放仲裁磁盘或 NFS 文件。

配置仲裁磁盘以进行拆分和合并恢复处理时，该磁盘还必须受 **devsrc** 命令支持，此命令是 AIX 操作系统的一部分。用于选择仲裁磁盘的 SMIT 界面会滤出所有不满足此要求的磁盘。EMC 磁盘不满足此要求。

不支持将 EMC PowerPath 磁盘用作仲裁磁盘。

针对集群中具有两个以上的节点的仲裁选项的局限性

- | 在具有两个以上的节点的集群中，进行集群拆分之后，将由获胜分区保留仲裁磁盘或 NFS 文件。在获胜分区中，单个节点将保留仲裁磁盘或 NFS 文件。如果此节点发生故障，但是未释放对仲裁磁盘或 NFS 文件的保留，那么其余节点无法获得对仲裁磁盘或 NFS 文件的保留，将失败。因此，会重新引导所有节点。

进行集群拆分之后，您必须快速解决失败节点存在的问题，以便释放对仲裁磁盘的保留。

相关概念:

第 15 页的『拆分策略』

当一组节点无法与集群中的其余节点进行通信时，站点之间可能会发生集群拆分事件。例如，在链接集群中，如果两个站点之间的所有通信链路都失败，那么将进行拆分。集群拆分事件会将集群拆分为两个或两个以上的分区。

第 16 页的『合并策略』

根据集群拆分策略，集群可能具有两个互相独立运行的分区。可以使用 PowerHA SystemMirror 来配置合并策略，合并策略允许分区在恢复分区之间的通信之后再次一起运行。

| 隔离策略

- | 隔离策略将在发生集群拆分事件或节点故障之后，隔离先前处于活动状态并在托管关键资源组的节点。隔离策略将确保应用程序数据不会损坏或丢失。

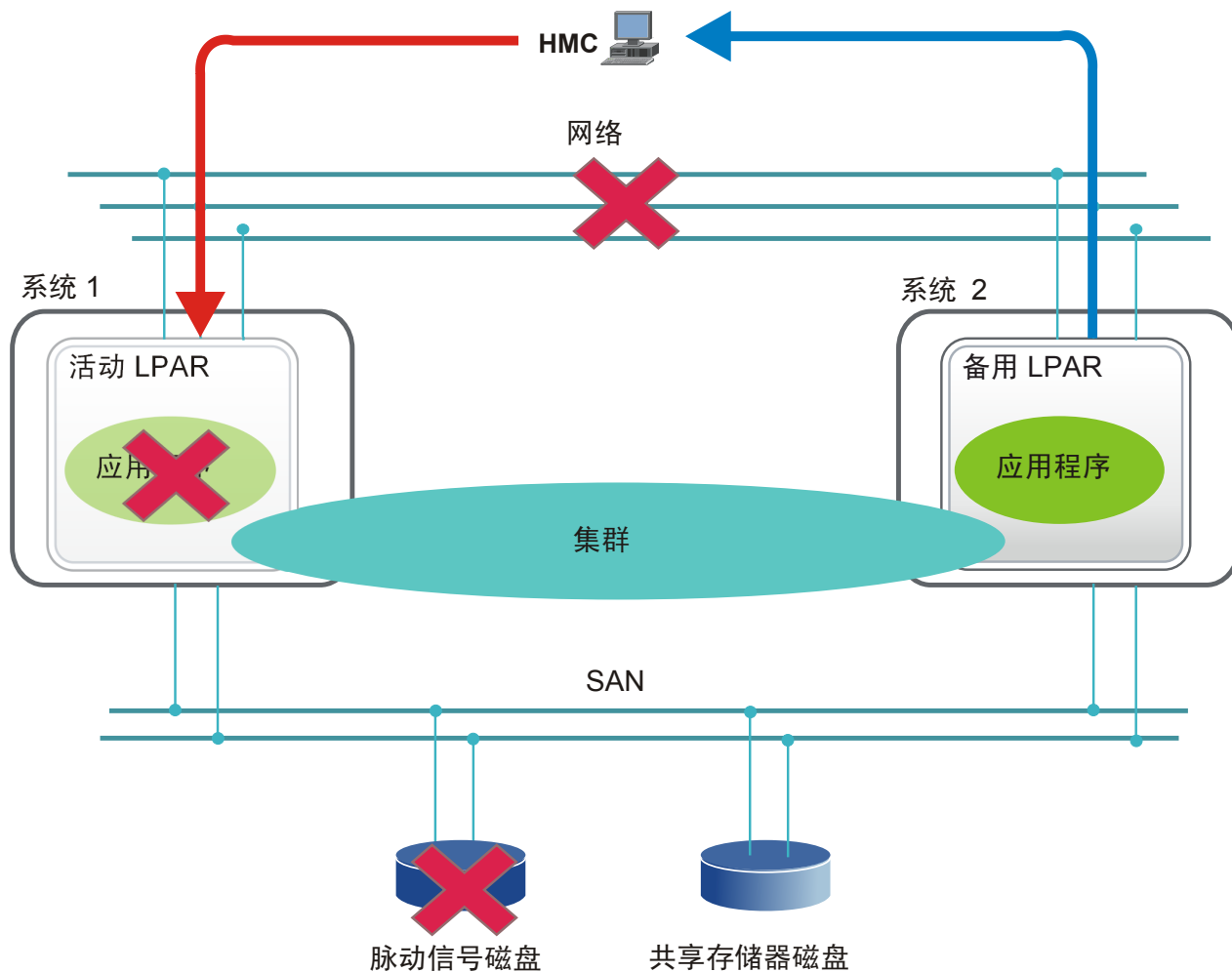
| 隔离策略包含两个选项，即节点暂停策略和磁盘电子篱笆。

| 活动节点暂停策略

| 如果发生集群拆分事件，那么活动节点暂停策略将停止先前的活动 LPAR，以便在备用 LPAR 使应用程序重新联机之前，托管应用程序的 LPAR 已完全停止。

| 此过程确保应用程序每次只在集群中的一个节点上运行。PowerHA SystemMirror 通过从硬件管理控制台 (HMC) 运行命令使活动 LPAR 脱机。如果您配置了 LPAR 策略以在 AIX 操作系统中重新引导，那么必须配置要手动重新启动设置的 PowerHA SystemMirror 资源组引导时策略。

| 下图显示了基于 HMC 的活动节点暂停策略实现。在下图中，备用 LPAR 在 HMC 上运行命令以停止活动 LPAR。当该命令成功使活动 LPAR 脱机时，备用 LPAR 启动备用节点上的资源组。



| 图 3. 活动节点暂停策略

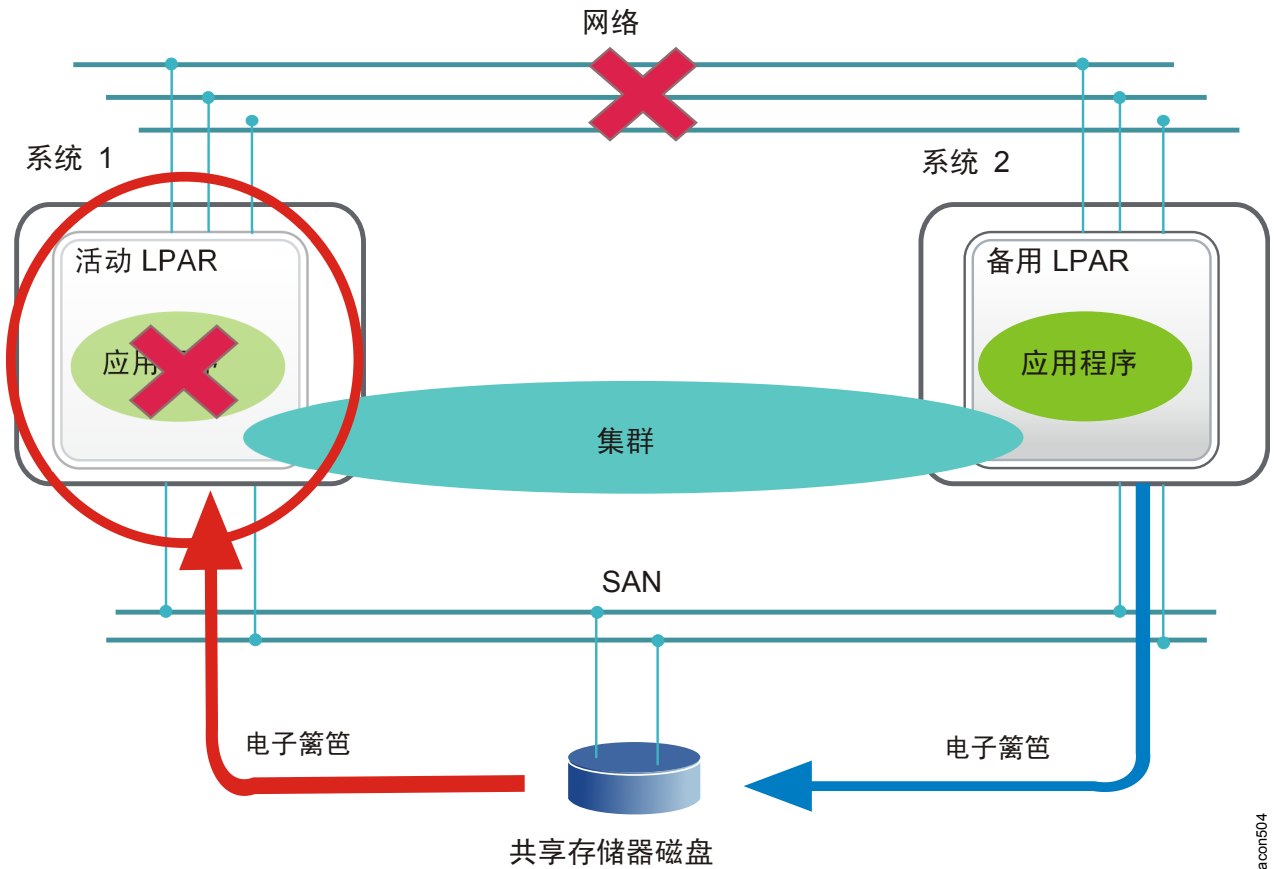
| 注：如果发生错误而导致无法使活动 LPAR 脱机，那么备用 LPAR 将进入错误状态，并需要手动干预以解决问题。

磁盘电子篱笆策略

PowerHA SystemMirror 通过使用磁盘电子篱笆策略来提供对拆分集群的保护。磁盘电子篱笆策略使用 SCSI-3 预留功能将有问题并正在托管工作负载的节点与集群分开（隔离）。在这种情况下，正在问题节点上运行的工作负载将在备用 LPAR 上启动。

电子篱笆过程确保独立节点具有对磁盘的访问权，并确保数据仍然受保护。磁盘电子篱笆策略受主动/被动部署模型支持。磁盘电子篱笆确保工作负载可以在仅对集群中的一个节点具有写访问权的情况下运行。PowerHA SystemMirror 将注册属于任何资源组的所有卷组的磁盘。

下图显示了在使用磁盘电子篱笆策略的情况下，发生集群拆分事件时所发生的情况。在下图中，备用 LPAR 与共享存储器磁盘进行通信，并且撤销了访问活动 LPAR 磁盘的请求。共享存储器磁盘将阻止先前处于活动状态的 LPAR 进行任何写访问，即使重新启动活动 LPAR 也是如此。如果 PowerHA SystemMirror 可以隔离活动 LPAR 中资源组的磁盘，那么备用 LPAR 会使应用程序或资源组联机。如果隔离活动 LPAR 中的磁盘时在备用 LPAR 中发生错误，那么不会使应用程序联机，您必须解决问题并手动使资源组联机。



hacon504

图 4. 磁盘电子篱笆策略

下列关键属性适用于 PowerHA SystemMirror 中的磁盘电子篱笆策略：

- 磁盘电子篱笆仅适用主动/被动集群环境。
- 使用在所有可用节点上联机启动策略的资源组不支持磁盘电子篱笆。
- 磁盘电子篱笆在集群级别受支持。因此，您可以在集群级别启用或禁用磁盘电子篱笆策略。
- 磁盘电子篱笆将管理所有属于包含在资源组中的卷组的磁盘。

- 磁盘电子篱笆在互相接管配置中受支持。如果集群中存在多个资源组，那么必须选择一个资源组作为最关键的资源组。发生集群拆分事件时，关键资源组的相对位置确定哪个站点获胜。获胜的站点是发生集群拆分事件之前未在运行关键资源组的站点。
- PowerHA SystemMirror 使用集群中仅可能多的信息来确定 LPAR 的运行状况。例如，如果活动 LPAR 即将崩溃，那么 PowerHA SystemMirror 会在活动 LPAR 脱机之前，向备用 LPAR 发送消息。这些通知确保备用 LPAR 确定活动 LPAR 已脱机，并确保备用 LPAR 可以使应用程序联机。
- 注：在某些情况下，备用 LPAR 知道活动 LPAR 未在发送脉动信号，但无法确定活动 LPAR 的实际状态。在这种情况下，备用 LPAR 会在等待节点故障检测超时字段中所指定的时间之后，声明活动 LPAR 已发生故障。此时，备用 LPAR 将隔离所有磁盘，然后使资源组联机。只要有一个磁盘未正确隔离，就不会使资源组联机。

PowerHA SystemMirror 多站点解决方案

PowerHA SystemMirror 支持站点和特定于站点的策略的不同类型的定义以实现高可用性灾难恢复 (HADR)。您可以在 PowerHA SystemMirror Standard Edition for AIX 和 PowerHA SystemMirror Enterprise Edition for AIX 中定义多个站点。

您可以使用 PowerHA SystemMirror 管理界面来创建下列多站点解决方案：

延伸集群

包含位于相同地理位置的站点中的节点。延伸集群不支持使用存储器复制管理来实现 HADR。

链接集群

包含位于不同地理位置的站点中的节点。链接集群支持跨站点 LVM 镜像和 HyperSwap。

下表显示了延伸集群与链接集群之间的差别。

表 1. 延伸集群与链接集群之间的差别

功能	延伸集群	链接集群
站点通信	多点广播	单点广播
存储库磁盘	共享	单独
集群通信	- 网络 - 存储区域网络 (SAN) - 磁盘	网络
跨站点逻辑卷管理器镜像	可用	使用 SAN 的情况下可用
HyperSwap	可用	使用 SAN 的情况下可用
并发资源组和 HyperSwap	可用	使用 SAN 的情况下可用

相关概念:

第 3 页的『PowerHA SystemMirror 使用 Cluster Aware AIX』

PowerHA SystemMirror 基于在 AIX 操作系统中受支持的核心集群功能。所有支持 Cluster Aware AIX (CAA) 功能的 AIX 版本都支持 PowerHA SystemMirror。

PowerHA SystemMirror 链接集群

对于每个站点位于不同地理位置的情况，最好是使用链接集群。通常，站点之间相隔足够远，以至于它们无法方便地共享公共存储区网络 (SAN)。每个站点都必须具有它自己的活动存储库磁盘以及任何备份存储库。链接集群始终使用单点广播在站点之间进行通信。链接集群是高可用性灾难恢复 (HADR) 的有用部分。

Cluster Aware AIX (CAA) 可以使用链接集群来支持地理位置相隔遥远（位于不同城市）的多个站点。链接集群将各个 CAA 集群链接至位于不同位置的站点。站点之间的链接用于脉动信号传递和集群通信。

可以对短地理距离通信（例如，同一城市的不同大厦之间进行通信）使用链接集群。

下表显示了拆分策略选项和合并策略选项所支持的链接集群配置。例如，在链接集群中，您可以使用拆分策略值为无且合并策略值为大多数或优先级的配置。但是，不能使用拆分策略值为仲裁且合并策略值为手动的配置。

表 2. 拆分策略和合并策略的链接集群选项

拆分策略选项	合并策略选项			
	大多数	优先级	仲裁	手动
无	X	X		
仲裁			X	
手动				X

下图显示了一个链接集群，其中的站点位于不同城市，并且每个站点都具有自己的存储库磁盘。

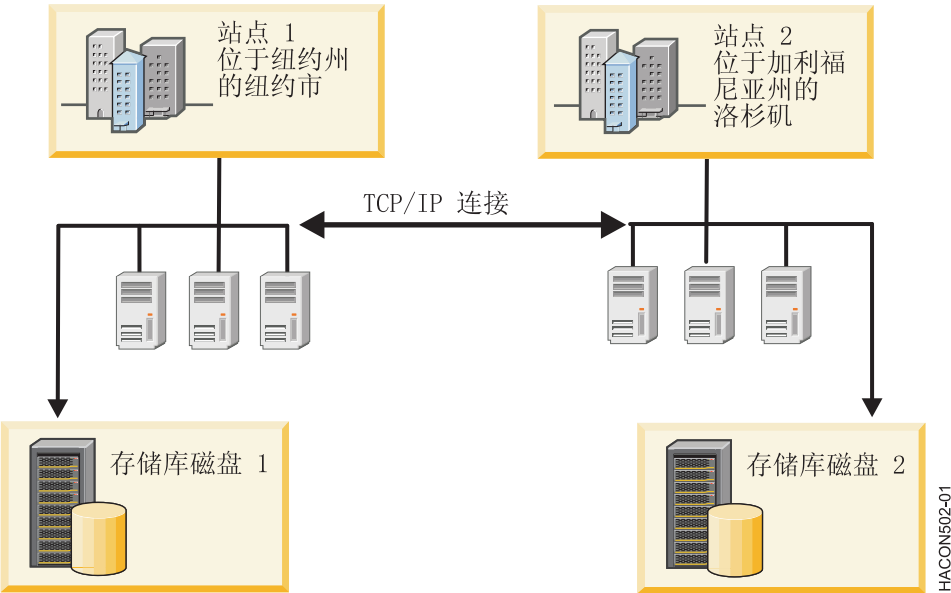


图 5. 链接集群配置

相关概念:

第 3 页的『PowerHA SystemMirror 使用 Cluster Aware AIX』

PowerHA SystemMirror 基于在 AIX 操作系统中受支持的核心集群功能。所有支持 Cluster Aware AIX (CAA) 功能的 AIX 版本都支持 PowerHA SystemMirror。

第 23 页的『PowerHA SystemMirror 延伸集群』

对于每个站点都位于同一地理位置的情况，最好是使用延伸集群。通常，所有节点都在公共存储区网络 (SAN) 中。至少，站点之间将共享活动存储库和任何备份存储库，就像资源组中可以在任一站点运行的任何磁盘一样。延伸集群可以支持 LVM 跨站点镜像和 HyperSwap。

第 11 页的『集群存储库磁盘』

PowerHA SystemMirror 使用共享磁盘作为中央存储库，用来管理集群的配置。集群中的所有节点必须可以使用标准集群部署方法或延伸集群部署方法来访问此磁盘，或者站点中的所有节点必须可以使用链接集群部署方法来访问此磁盘。

PowerHA SystemMirror 延伸集群

对于每个站点都位于同一地理位置的情况，最好是使用延伸集群。通常，所有节点都在公共存储区网络 (SAN) 中。至少，站点之间将共享活动存储库和任何备份存储库，就像资源组中可以在任一站点运行的任何磁盘一样。延伸集群可以支持 LVM 跨站点镜像和 HyperSwap。

当在您所在的环境中使用延伸集群时，会在集群中的站点之间部署 Cluster Aware AIX (CAA) 集群。

如果 IP 通信失败，并且您所在环境使用延伸集群，那么该集群中的所有节点都可以使用磁盘脉动信号功能来进行通信以及使您所在环境保持运作。

延伸集群必须满足下列需求：

- 站点通信可以使用单点广播 IP 地址或多点广播 IP 地址。
- 站点之间至少要共享一个存储库磁盘。

下表显示了拆分策略选项和合并策略选项所支持的延伸集群配置。例如，在延伸集群中，您可以使用拆分策略值为无且合并策略值为大多数的配置。但是，不能使用拆分策略值为仲裁且合并策略值为在大多数的配置。

表 3. 拆分策略和合并策略的延伸集群选项

拆分策略选项	合并策略选项	
	大多数	仲裁
无	X	
仲裁		X

下图是延伸集群的示例，它具有位于同一城市、地理位置相距很近的不同大厦的两个站点，并且这两个站点共享一个存储库磁盘。

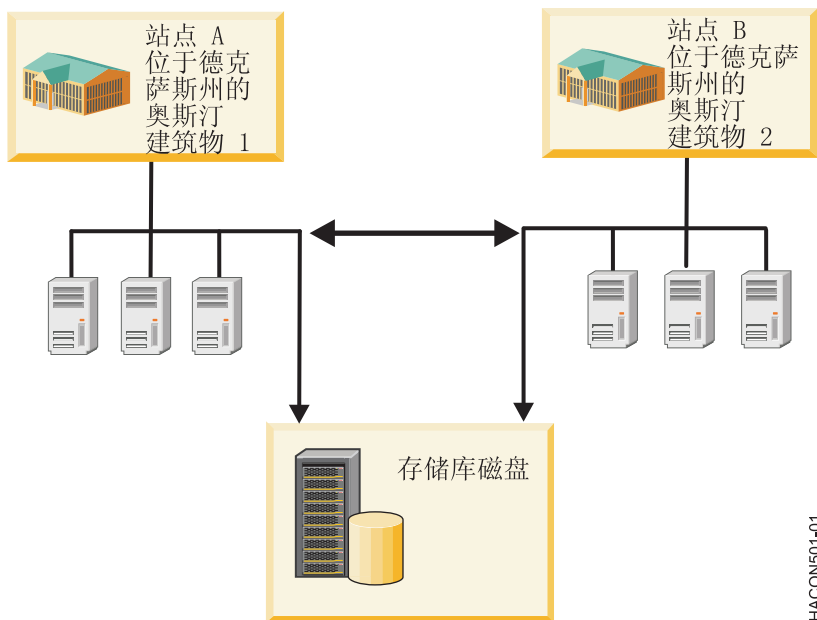


图 6. 延伸集群典型配置

相关概念:

第 21 页的『PowerHA SystemMirror 链接集群』

对于每个站点位于不同地理位置的情况，最好是使用链接集群。通常，站点之间相隔足够远，以至于它们无法方便地共享公共存储区网络 (SAN)。每个站点都必须具有它自己的活动存储库磁盘以及任何备份存储库。链接集群始终使用单点广播在站点之间进行通信。链接集群是高可用性灾难恢复 (HADR) 的有用部分。

第 3 页的『PowerHA SystemMirror 使用 Cluster Aware AIX』

PowerHA SystemMirror 基于在 AIX 操作系统中受支持的核心集群功能。所有支持 Cluster Aware AIX (CAA) 功能的 AIX 版本都支持 PowerHA SystemMirror。

第 11 页的『集群存储库磁盘』

PowerHA SystemMirror 使用共享磁盘作为中央存储库，用来管理集群的配置。集群中的所有节点必须可以使用标准集群部署方法或延伸集群部署方法来访问此磁盘，或者站点中的所有节点必须可以使用链接集群部署方法来访问此磁盘。

集群网络

集群节点通过通信网络互相通信。

如果网络中节点上的一个物理网络接口卡发生故障，那么 PowerHA SystemMirror 会通过将流量转移到同一节点的其他物理网络接口卡来保持到该节点的通信。如果到该节点的"连接"失败，那么 PowerHA SystemMirror 会将资源转移到它可以访问的其他节点。

另外，RSCT 可以通过集群网络在节点间发送脉动信号，以定期检查集群节点本身的运行状况。如果 PowerHA SystemMirror 检测到某节点没有脉动信号，那么该节点将被视为发生故障，且资源会自动转移到其他节点。

我们强烈建议在集群的节点间配置多个通信路径。拥有多个通信网络可以防止集群分区，这种情况下每个分区内的节点均形成它们自己的实体。在分区集群中，每个分区中的节点都可能允许对相同数据进行同时却非同步的访问。这可能导致不同节点中数据的不同视图。

相关概念:

第 28 页的『通过 TCP/IP 和存储区域网络传递脉动信号』

脉动信号是一种在节点间发送的通信数据包类型。脉动信号用于监视节点的运行状况、网络和网络接口，并防止集群分区。

物理网络和逻辑网络

物理网络连接两个或多个物理网络接口。

正如上一节所描述的那样，配置多个基于 TCP/IP 的网络可帮助防止集群分区。必要时，PowerHA SystemMirror 在存储器网络间通信。这一额外的通信路径可通过在基于 TCP/IP 的网络连接变得密集或在集群节点间隔断的情况下，提供额外的通信路径来帮助防止分区集群。

注：如果在您考虑的集群中，物理网络使用外部网络设备从一个网络向另外一个网络发送数据包，那么请考虑以下情况：配置 PowerHA SystemMirror 集群时，PowerHA SystemMirror 验证连通性并访问特定的物理网络上定义的所有接口。但是，PowerHA SystemMirror 不能确定外部网络设备（例如，集群节点间网络路径中的网桥和路由器）的存在。如果网络具有外部网络设备，请确保您使用的设备高度可用而且冗余，这样它们就不会在 PowerHA SystemMirror 集群中创建单点故障。

逻辑网络是连接两个或多个逻辑网络接口或设备的物理网络的一部分。逻辑网络接口或设备是可通过操作系统识别的软件实体。在物理网络接口/设备和逻辑网络接口/设备间存在一对一映射。每个逻辑网络接口都可与相同逻辑网络上的每个逻辑网络接口交换数据包。

如果逻辑网络上的一部分逻辑网络接口在共享相同的物理网络时需要互相通信（但不与其他接口通信），那么将使用子网。子网掩码定义可确定逻辑网络接口是否可向相同逻辑网络上另外一个逻辑网络接口发送数据包的那部分 IP 地址。

PowerHA SystemMirror 中的逻辑网络：

PowerHA SystemMirror 拥有它自己的类似的逻辑网络概念。

PowerHA SystemMirror 网络中的所有逻辑网络接口都可使 PowerHA SystemMirror 数据包互相直接通信。每个逻辑网络都由一个唯一的名称标识。如果针对 PowerHA SystemMirror 集群配置使用自动发现功能，那么 PowerHA SystemMirror 将为它发现的每个 PowerHA SystemMirror 逻辑网络指定一个名称，例如 *net_ether_01*。一个 PowerHA SystemMirror 逻辑网络可以包含一个或多个子网。RSCT 管理逻辑子网间的路由数据包。

相关概念：

第 42 页的『支持 PowerHA SystemMirror 的硬件』

可以使用不同类型的 IBM 硬件与 PowerHA SystemMirror 一起实现基本级别的高度可用环境。

本地和全局网络故障：

当集群网络上发生故障时，PowerHA SystemMirror 将使用网络故障事件管理这种情况。PowerHA SystemMirror 监视并区分两种类型的网络故障事件：本地网络故障事件和全局网络故障事件。

本地网络故障

本地网络故障是个 PowerHA SystemMirror 事件，在该事件中，数据包不能由一个节点通过 PowerHA SystemMirror 逻辑网络发送或接收。例如，如果加入特定 PowerHA SystemMirror 逻辑网络中的节点的所有网络接口卡都发生故障，就会发生该事件。在发生本地网络故障的情况下，网络仍然可由其他节点使用。为了处理本地网络故障，PowerHA SystemMirror 选择性地将（该网络上的）资源从一个节点移到另外一个节点。该操作称为“选择性失败转移”。

全局网络故障

全局网络故障是个 PowerHA SystemMirror 事件，在该事件中，数据包不能由任何节点通过 PowerHA SystemMirror 逻辑网络发送或接收。例如，如果物理网络被损坏，那么就会发生该事件。区分全局网络和全局网络故障事件非常重要。全局网络是 PowerHA SystemMirror 网络的组合。全局网络故障事件指的是一种故障，它影响连接到任何逻辑 PowerHA SystemMirror 网络的所有节点，不一定是全局网络。

PowerHA SystemMirror 通信接口

PowerHA SystemMirror 通信接口是在 PowerHA SystemMirror 中定义的逻辑网络接口、服务 IP 地址和服务 IP 标签的分组。

PowerHA SystemMirror 通信接口组合在一起以创建基于 IP 的网络。PowerHA SystemMirror 通信接口包括：

- 逻辑网络接口是一个名称，AIX 将物理网络接口卡的端口（例如，en0）解析为该名称。
- 服务 IP 地址是一个 IP 地址（例如，129.9.201.1），将通过该 IP 地址提供诸如应用程序等服务，且客户机节点将通过该 IP 地址进行通信。
- 服务 IP 标签是一个映射至服务 IP 地址的标签（例如，**/etc/hosts** 文件中的主机名，或者服务 IP 地址的逻辑等效项，例如 node_A_en_service）。

可以用以下方法使用 PowerHA SystemMirror 中的通信接口：

- 通信接口指的是基于 IP 的网络和网络接口卡 (NIC)。连接到公共物理网络的 NIC 组合为 PowerHA SystemMirror 所使用的逻辑网络。
- 每个 NIC 能够主管多个 TCP/IP 地址。配置集群时，您必须定义 PowerHA SystemMirror 监视的 IP 地址（基本 IP 地址或引导 IP 地址）和 PowerHA SystemMirror 将使其保持高可用性的 IP 地址（服务 IP 地址）。
- PowerHA SystemMirror 中的脉动信号通过通信接口进行传递。PowerHA SystemMirror 使用 Cluster Aware AIX (CAA) 的脉动信号传递设施来监视其网络接口和 IP 地址。CAA 向 RSCT 提供您所创建的网络拓扑，而 RSCT 向 PowerHA SystemMirror 提供故障通知。

PowerHA SystemMirror 非 IP 通信设备

PowerHA SystemMirror 还监视不具备 IP 通信能力的网络设备。这些设备包括存储区域网络 (SAN) 和共享磁盘。

PowerHA SystemMirror 中的子网路由需求

子网路径定义一条路径（已由子网定义），这是为了经过逻辑网络发送数据包到另一个逻辑网络上的地址。

AIX 允许您为内核路由表中的相同目标添加多条路径。如果多条匹配的路径具有相同的条件，那么可使用多个子网路径中的一个替换执行路由。由于以下原因，考虑 PowerHA SystemMirror 中的子网路由是重要的：

- PowerHA SystemMirror 不对共享相同子网路径的逻辑网络接口进行区分。如果一个逻辑网络接口与另一个接口共享路径，那么 PowerHA SystemMirror 无法确定其运行状况。有关网络路径的更多信息，请参阅针对 **route** 命令的 AIX 联机帮助页。
- 网络管理员或 TCP/IP 需求经常将各种约束强加于基于 IP 的网络上。子网和路径也是约束，必须在这些约束的范围内配置 PowerHA SystemMirror 以进行操作。

注：应该将每个通信接口都放置在属于唯一子网的节点上，以使 PowerHA SystemMirror 可以监视每个接口。这不是针对所有情况的严格要求，它取决于多个因素。例如，当每个节点仅有一个网络接口时（例如在虚拟网络环境中），因为存在单个子网路径，所以启动 IP 地址和服务 IP 地址可在相同的子网上。如果需要唯一的子网，那么 PowerHA SystemMirror 集群验证会将其标识为错误。

服务 IP 标签和地址

服务 IP 标签是映射至服务 IP 地址的标签，用来建立客户机节点与服务器节点之间的通信。

可使用通过服务 IP 标签建立的连接来提供服务（例如，数据库应用程序）。服务 IP 标签可作为资源存放到资源组中，允许 PowerHA SystemMirror 监视其运行状况，并在一个节点中使其保持高可用性，或者在已配置 IP 地址接管的情况下，通过在故障事件中将其转移到另一个节点从而在集群节点之间使其保持高可用性。

IP 别名

IP 别名是除网络接口卡 (NIC) 中最初配置的 IP 标签或 IP 地址之外，配置到该网络接口卡中的 IP 标签或 IP 地址。

IP 别名是受 PowerHA SystemMirror 支持的 AIX 功能。AIX 操作系统支持一个 NIC 上使用多个 IP 别名。NIC 上的每个 IP 别名都可在单独的子网上配置。

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

下列主题包含有关 PowerHA SystemMirror 如何根据恢复服务 IP 标签所使用的不同机制将服务 IP 标签绑定到通信接口的信息。

IP 地址接管

IP 地址接管是服务 IP 标签的一种恢复机制 - 当初始网络接口卡 (NIC) 失败时，将服务 IP 标签移至另一个节点上的另一个网络接口卡。

如果一个节点上的物理网络接口卡失败，而在同一网络中的同一节点上没有其他可访问的物理网络接口卡，那么将进行 IPAT。因此，同一节点中的这些 NIC 将无法交换 IP 标签，并且 PowerHA SystemMirror 将通过使用备份节点上的 NIC 并借助 IPAT 来恢复服务 IP 地址。IP 地址接管通过在发生故障之后恢复 IP 地址来使 IP 地址高度可用。PowerHA SystemMirror 使用名为“通过 IP 别名的 IPAT”的方法。

IPAT 和服务 IP 标签

IPAT 控制服务 IP 标签。

当使用“通过 IP 别名的 IPAT”时，将对服务 IP 标签或 IP 地址制作别名或将其作为额外的地址添加到相同的网络接口。也就是说，多个 IP 地址或 IP 标签可同时配置在相同的网络接口上。在该配置中，必须将您定义的所有 IP 地址或标签配置在不同的子网上，除非每个节点上只有单个网络接口。这种方法可节省硬件，但需要额外的子网。

通过 IP 别名进行 IP 地址接管

可以使用 AIX 操作系统的 IP 别名判别网络能力在某些类型的网络中配置 IP 地址接管。

向网络接口定义 IP 别名允许在相同的网络接口上创建多个 IP 标签和地址。“通过 IP 别名的 IPAT”使用许多类型的网络中提供的免费“地址解析协议”(ARP) 功能。

当包含服务 IP 标签的资源组从主节点失败转移到目标节点时，服务 IP 标签将作为别名地址在可用 NIC 的基本 IP 地址基础上添加（和除去）。这使单个 NIC 能够支持多个服务 IP 标签作为别名放置在 NIC 上。因此，同一节点可同时承载多个资源组。

当同一节点上有多个接口连接至同一网络，而这些接口未组合成以太网聚集时，所有引导地址必须全部都在不同的子网中。此外，任何持久地址或服务地址必须与引导地址在不同子网中。

因为 IP 别名判别允许相同网络接口上的多个服务标签共存，所以可在集群中使用较少的物理网络接口卡。失败转移时，PowerHA SystemMirror 将在可用的网络接口卡上平均分配别名。

服务 IP 标签别名的分配首选项

PowerHA SystemMirror 使用"通过 IP 别名进行 IP 地址接管 (IPAT)"方法来保持资源组中的服务 IP 标签的高可用性。

集群启动时，PowerHA SystemMirror（缺省情况下）通过使用最小负载原则，在网络中所有可用的启动接口中分配所有服务 IP 标签别名。PowerHA SystemMirror 为已指定了最少数目的别名或持久 IP 标签的接口指定任何新的服务地址。

但是，在某些情况下，可能希望指定其他类型的分配，或者确保在启动期间和后续的集群事件期间继续以特定方式分配标签。

例如，您可能想要将所有服务 IP 标签别名分配给与当前拥有该节点的持久 IP 标签的接口相同的引导接口。此选项在这样的 VPN 防火墙配置中可能很有用：仅将外部连接授权给一个接口，并且所有 IP 标签（持久 IP 标签别名和服务 IP 标签别名）必须位于同一接口上以启用连接。

可为由 PowerHA SystemMirror 控制的服务 IP 标签的别名配置分配首选项。

服务 IP 标签别名的分配首选项是个网络范围的属性，用于控制集群节点上物理网络接口卡的服务 IP 标签别名的位置。为服务 IP 标签别名配置分配首选项能够：

- 允许您将以前在节点上指定的持久 IP 标签考虑在内，为集群中的服务 IP 标签定制负载均衡。
- 使 PowerHA SystemMirror 根据指定的首选项重新分配别名服务 IP 标签。
- 允许您配置适应 VPN 防火墙外部连接需求的分配首选项类型。
- 尽管服务 IP 标签可能会移到另一个网络接口，PowerHA SystemMirror 仍能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（例如失败转移、回退或同一节点上的接口更改）期间进行维护的。例如，如果将标签指定为映射至相同接口，那么即使最初配置的服务 IP 标签移到另一个节点，标签仍将映射在相同接口上。
- 只要集群中可接受的网络接口可用，就能执行分配首选项。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。

相关概念:

第 91 页的『持久节点 IP 标签』

持久节点 IP 标签是非常有用的管理工具，让您在即使节点上的 PowerHA SystemMirror 集群服务停止的情况下仍能联系该节点。

相关信息:

管理 PowerHA SystemMirror

规划 PowerHA SystemMirror

通过 TCP/IP 和存储区域网络传递脉动信号

脉动信号是一种在节点间发送的通信数据包类型。脉动信号用于监视节点的运行状况、网络和网络接口，并防止集群分区。

为了使 PowerHA SystemMirror 集群识别并响应故障，必须持续地检查集群的运行状况。其中一些检查由脉动信号功能提供。

每个集群节点都将在特定的时间间隔内向其他集群节点发送脉动信号消息，并期望在特定的时间间隔内从节点接收脉动信号消息。如果消息接收停止，那么 PowerHA SystemMirror 将意识到已发生故障。脉动信号可通过以下途径发送：

- TCP/IP 网络
- 存储区域网络
- 集群存储库磁盘

Cluster Aware AIX (CAA) 在所有可用的 TCP/IP 网络和存储区域网络 (SAN) 上都使用脉动信号通信。如果 TCP/IP 网络和 SAN 网络失败，那么 CAA 尝试使用存储库磁盘作为备用脉动信号机制。在 **lscluster** 命令的输出中，备份存储库磁盘的脉动信号路径显示为 **dpcom** 接口。如果 TCP/IP 网络和 SAN 网络正在运行，那么 **lscluster -i** 命令会将 **dpcom** 接口显示为 **restricted**（受限制）。

可将脉动信号功能配置为在节点间使用特定路径。这允许脉动信号监视所有 PowerHA SystemMirror 网络和网络接口以及集群节点本身的运行状况。

由 CAA 自动设置脉动信号路径；您可以选择将点到点和磁盘路径配置为 PowerHA SystemMirror 配置的一部分。

相关概念:

第 24 页的『集群网络』

集群节点通过通信网络互相通信。

通过因特网协议网络传递脉动信号

PowerHA SystemMirror 依赖 Cluster Aware AIX (CAA) 通过因特网协议网络在集群节点之间提供脉动信号传递。缺省情况下，CAA 使用单点广播通信来传递脉动信号。（可选）如果网络配置为支持多点广播，那么可以选择使用多点广播通信。

PowerHA SystemMirror 7.1.2 或更高版本支持因特网协议 V6 (IPv6)。但是，您无法显式指定 IPv6 多点广播地址。CAA 使用从因特网协议 V4 (IPv4) 多点广播地址派生而来的 IPv6 多点广播地址。要确定 IPv6 多点广播地址，使用逻辑运算符 OR 将标准前缀 **0xFF05** 与 IPv4 地址的十六进制等价地址进行组合。例如，IPv4 多点广播地址为 **228.8.16.129** 或 **0xE4081081**。逻辑 OR 运算对标准前缀执行的变换为 **0xFF05:: | 0xE4081081**。因此，获得的 IPv6 多点广播地址为 **0xFF05::E408:1081**。

缺省情况下，CAA 通过所配置的所有 AIX 接口（而不仅仅是在 PowerHA SystemMirror 拓扑中配置的接口）进行脉动信号传递。

注：要停止通过特定网络进行 CAA 脉动信号传递，您可以将该网络定义为 **private**，而不是 **public**。此注释同时适用于 IPv4 接口和 IPv6 接口。

相关概念:

第 26 页的『PowerHA SystemMirror 中的子网路由需求』

子网路由定义一条路径（已由子网定义），这是为了经过逻辑网络发送数据包到另一个逻辑网络上的地址。

相关信息:

用于更改基于 IP 的网络的步骤

通过存储区域网络传递脉动信号

Cluster Aware AIX (CAA) 会通过涵盖集群节点的公共存储区域网络 (SAN) 来自动交换脉动信号。这些连接可为使用基于 TCP/IP 的单个网络的集群提供备用的脉动信号路径。

PowerHA SystemMirror 多点广播

PowerHA SystemMirror 使用作为 AIX 操作系统的一部分嵌入的、称为 Cluster Aware AIX (CAA) 的集群运行状况管理层。通过使用光纤通道适配器以及使用通过中央存储库的基于磁盘的消息传递，CAA 使用 AIX 内核级别的代码通过网络来交换脉动信号。

缺省情况下，CAA 使用单点广播通信来传递脉动信号。（可选）如果网络配置为支持多点广播，那么可以选择使用多点广播通信。对于集群通信，您可以手动配置多点广播地址，或者让 CAA 自动选择多点广播地址。

您最初创建集群时未使用多点广播通信，但是需要使用多点广播通信，集群才能正常运作。您必须测试并验证集群中的所有节点之间可以通过网络进行多点广播通信。

相关信息:

-  Administratively Scoped IP Multicast
-  Multicast address range and assignments
-  Multicast router information

多点广播信息包通信

多点广播是一种寻址方式；在这种方式下，多个节点组成一个组并互相交换消息。

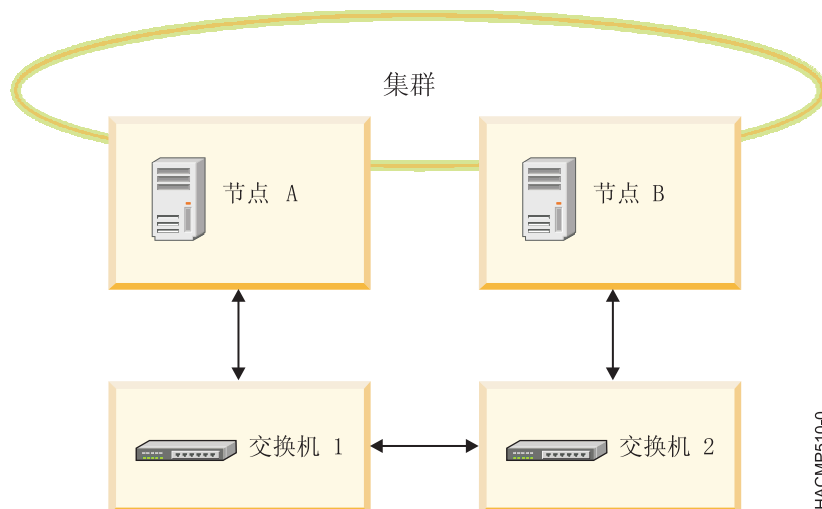
该组中的一个节点所发送的多点广播消息，会被该组中的所有其他节点接收到。通过此功能能够高效进行集群通信。例如，集群中的一个节点可以发送单个多点广播信息包，此信息包可以通知其他节点发生了紧急事件。

多点广播网络交换机

交换机将不同的节点和网络段连接起来，并将网络数据发送至正确的节点。交换机是一个多端口网桥，它在开放式系统互连 (OSI) 模型的数据链路层（第 2 层）处理和传递数据。某些交换机还可以在网路层（第 3 层）处理数据。

通常，在数据中心环境中，多个节点之间通过交换机互连。在交换机级联的情况下，多点广播信息包从一个节点通过交换机发送至另一台交换机，直到它到达集群中的最终目标节点。交换机采用与常规网络通信不同的方式来处理多点广播信息包。因此，如果未正确设置网络，那么可能不会进行交换机与交换机之间的通信以传递多点广播信息包。

下图显示了一个配置，其中，节点 A 包括在集群中，它通过交换机 1 和交换机 2 连接至节点 B。要让多点广播信息包在节点 A 与节点 B 之间流动，必须启用交换机 1 与交换机 2 之间的网络连接以进行多点广播通信。



因特网组管理协议

因特网组管理协议 (IGMP) 是一项通信协议，它使节点（接收方）能够将该节点打算接收特定多点广播流量的意图告知多点广播路由器（IGMP 查询者）。

IGMP 在路由器与允许执行下列操作的节点之间运行：

- 路由器询问节点是否需要特定的多点广播流（IGMP 查询）。
- 节点答复路由器它们是否在寻找特定的多点广播流（IGMP 报告）。

IP 网络中的节点和临近路由器使用 IGMP 通信协议进行交互、建立多点广播通信的基本规则以及建立多点广播组成员资格。

IGMP 探听

IGMP 探听是交换机执行的一项活动，用来跟踪与 IGMP 通信相关的信息包交换，并且适应对多点广播信息包进行过滤。具有 IGMP 探听的交换机通过观察节点与路由器之间的 IGMP 事务来获取有用信息。此功能使交换机在需要时能够将多点广播信息包正确转发至网络路径中的下一个交换机。

交换机将监视 IGMP 流量，并且仅在需要时才发出多点广播信息包。交换机通常会构建一个 IGMP 探听表，此表中具有已请求特定多点广播组的所有端口的列表。IGMP 探听表用来允许多点广播信息包在网络中传递，或者不允许它们在网络中传递。您可以配置交换机以避免进行 IGMP 探听。

多点广播路由

网络设备使用多点广播路由这一过程来决定如何将多点广播信息包传递给多点广播流量的所有预期接收方。

请阅读路由器文档以确定路由器是否实现了多点广播路由。

使用特殊路由算法来转发多点广播信息包的网络实体称为多点广播路由器。节点和其他网络元素实现多点广播路由器，以使多点广播流量能够在网络中传递。

注：某些路由器还支持多点广播软件包路由。

当交换机级联时，您可能需要设置网络交换机以根据需要转发信息包以实现多点广播路由器。如果在您的集群中遇到多点广播通信问题，请检查网络设备是否具有正确的 IGMP 和监听设置。有关设置网络以传递多点广播流量的更多信息，请参阅有关网络设备的文档。

PowerHA SystemMirror 资源和资源组

请阅读本节以了解整个文档以及 PowerHA SystemMirror 用户界面中使用的与资源相关的概念和定义。

识别和保持集群资源可用

PowerHA SystemMirror 软件提供高度可用的环境。

PowerHA SystemMirror 软件通过以下措施实现这一点：

- 标识对于应用程序运作非常重要的集群资源集，并将这些资源组合成资源组。
- 定义可指示 PowerHA SystemMirror 如何管理资源以在集群操作的不同阶段（启动、失败转移和回退）使它们保持高可用性的资源组策略和属性。

通过识别资源和定义资源组策略，PowerHA SystemMirror 软件使大量集群配置成为可能，在定义一个适用于各种需求的集群环境时提供极大的灵活性。

识别集群资源

以下集群资源既可包含硬件资源，也可包含软件资源：

- 磁盘
- 卷组
- 逻辑卷
- 文件系统
- 服务 IP 标签或地址
- 应用程序
- 磁带资源

运行 PowerHA SystemMirror 的处理器拥有用户定义的资源集：磁盘、卷组、文件系统、IP 地址和应用程序。为了使资源保持高可用性，可将互相依赖的资源集配置为资源组。资源组允许您将相关资源组合成单个逻辑实体，以便更容易进行配置和管理。PowerHA SystemMirror 软件将资源组作为一个单元进行处理，因此可使互相依赖的资源全部位于一个节点上并保持它们高度可用。

相关概念：

第 33 页的『卷组』

卷组是一组物理卷，AIX 将其视为可寻址的连续磁盘区域。

第 33 页的『逻辑卷』

逻辑卷是逻辑分区的集合，AIX 使其可用作单个存储单元 - 即，逻辑卷是物理磁盘的逻辑视图。逻辑分区可映射至一个、两个或三个物理分区以实现镜像。

第 33 页的『文件系统』

一个文件系统可写入单个逻辑卷。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。

第 33 页的『应用程序』

高度可用系统的目的是确保用户可以访问关键服务。应用程序要在 PowerHA SystemMirror 环境中运行，通常不需要进行更改。可在异常关闭后成功重新启动的任何应用程序都是 PowerHA SystemMirror 的候选应用程序。

第 34 页的『服务 IP 标签和 IP 地址』

服务 IP 标签用于在客户机节点与服务器节点之间建立通信。可使用通过服务 IP 标签建立的连接来提供服务（例如，数据库应用程序）。

第 35 页的『磁带资源』

可将 SCSI 或光纤通道磁带机配置为非并发资源组中的集群资源，使其对集群中的两个节点高度可用。

集群资源类型

本节提供了可在 PowerHA SystemMirror 中配置且可包含于资源组中以让 PowerHA SystemMirror 保持它们高度可用的资源的简要概述。

卷组

卷组是一组物理卷，AIX 将其视为可寻址的连续磁盘区域。

可对 AIX 操作系统配置卷组，并且卷组可包括在 PowerHA SystemMirror 中的资源组中。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。共享磁盘是那些物理连接到集群节点且逻辑配置在所有集群节点上的磁盘。

逻辑卷

逻辑卷是逻辑分区的集合，AIX 使其可用作单个存储单元 - 即，逻辑卷是物理磁盘的逻辑视图。逻辑分区可映射至一个、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（非并发访问）或原始设备（并发访问）。并发访问不支持文件系统。并发访问环境中的数据库和应用程序必须访问原始逻辑卷（例如，/dev/rsharedlv）。

共享逻辑卷在 PowerHA SystemMirror 集群中必须有唯一的名称。

注：共享卷组不能包含活动的调页空间。

文件系统

一个文件系统可写入单个逻辑卷。通常，可组织一个文件集作为文件系统以方便快速地管理数据。

共享文件系统

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。对于非并发访问，您希望规划共享文件系统，以使它们放置于集群节点共享的外部磁盘上。数据位于这些外部共享磁盘的文件系统中，以使数据高度可用。对于并发访问，您不能使用日志文件系统。相反，请使用原始逻辑卷。

日志文件系统和增强型日志文件系统

相对于日志文件系统 (JFS)，增强型日志文件系统可提供存储更大文件的能力。JFS2 是 64 位内核的缺省文件系统。您可以选择实施为 32 位环境推荐的文件系统 JFS，也可以选择实施提供 64 位功能的 JFS2。JFS2 比 JFS 更加灵活，因为它允许您动态地增加或减少在文件系统中可以拥有的文件数量。JFS2 还可以让您将文件系统日志包含在与数据相同的逻辑卷中，而不是为卷组中所有文件系统的日志分配单独的逻辑卷。

应用程序

高度可用系统的目的是确保用户可以访问关键服务。应用程序要在 PowerHA SystemMirror 环境中运行，通常不需要进行更改。可在异常关闭后成功重新启动的任何应用程序都是 PowerHA SystemMirror 的候选应用程序。

例如，所有商业 DBMS 产品都在某种交易日志中对磁盘状态提供了检查点。在服务器故障事件中，失败转移服务器将重新启动 DBMS，它将重新建立数据库一致性然后恢复处理。

可以使用 AIX Fast Connect 与未将 AIX 作为操作系统来运行的系统共享资源。如果您将 Fast Connect 配置为 SystemMirror 资源，那么 PowerHA SystemMirror 将保持其高度可用，并从节点和网络接口故障进行恢复。PowerHA SystemMirror 还会在进行集群验证期间验证 Fast Connect 的配置。可通过将应用程序作为应用程序控制器资源定义到 PowerHA SystemMirror 来管理应用程序。应用程序控制器包括应用程序启动和停止脚本。当应用程序需要在特定节点联机或脱机时，PowerHA SystemMirror 使用这些脚本保持应用程序高度可用。

注：启动和停止脚本是 PowerHA SystemMirror 对应用程序的控制要点。您指定的脚本能够正常运行以启动和停止应用程序的所有方面是非常重要的。如果脚本未能正确控制应用程序，那么其他部分的应用程序恢复可能会受影响。例如，如果您使用的停止脚本不能完全停止应用程序且进程继续访问磁盘，那么 PowerHA SystemMirror 将无法使发生故障的节点上的卷组处于脱机状态，也无法使其在备份节点上恢复。

仅在已对应用程序启动和停止脚本进行全面测试后，再将应用程序控制器添加到 PowerHA SystemMirror 资源组。

包含应用程序控制器的资源组应该还包含该应用程序所依赖的所有资源，包括服务 IP 地址、卷组和文件系统。一旦创建这样的资源组，PowerHA SystemMirror 将管理整个资源组，并因此将资源组中互相依赖的所有资源作为单个实体进行管理。PowerHA SystemMirror 协调应用程序恢复，并按照确保先激活所有互相依赖的资源再激活其他资源的顺序来管理资源。

另外，PowerHA SystemMirror 包含应用程序监视能力，借助此能力可定义监视器以检测进程的异常终止，或定期轮询应用程序的终止，并根据问题的检测采取自动操作。

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。由于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。或者，可配置一个定制监视器以检查数据库的运行状况，以及一个进程终止监视器以立即检测数据库进程的终止。

还可为应用程序监视器指定方式。它可以跟踪应用程序正在如何运行（运行方式），或者应用程序是否已成功启动（应用程序启动方式）。使用监视器监视应用程序启动对于复杂的集群配置特别有用。

服务 IP 标签和 IP 地址

服务 IP 标签用于在客户机节点与服务器节点之间建立通信。可使用通过服务 IP 标签建立的连接来提供服务（例如，数据库应用程序）。

服务 IP 标签可作为资源放到资源组中，允许 PowerHA SystemMirror 监视其运行状况，并在一个节点中保持其高度可用，或者（如果已配置 IP 地址接管）通过在故障事件中将其转移到另外一个节点从而在集群节点间保持其高度可用。

注：某些子网需求申请将服务 IP 标签配置为不同资源组类型的资源。

相关概念:

第 27 页的『服务 IP 标签和地址』

服务 IP 标签是映射至服务 IP 地址的标签，用来建立客户机节点与服务器节点之间的通信。

相关信息:

规划 PowerHA SystemMirror

磁带资源

可将 SCSI 或光纤通道磁带机配置为非并发资源组中的集群资源，使其对集群中的两个节点高度可用。

可通过以下 PowerHA SystemMirror 功能简化共享磁带机的管理：

- 使用 SMIT 配置工具的磁带机配置
- 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
- 集群资源动态重新配置期间磁带机的受控重新分配。

集群资源组

要让 PowerHA SystemMirror 软件使其高度可用，每个资源都必须包含在资源组中。资源组允许您将相关资源组合到单个逻辑实体中，以便更容易地管理。

受支持的最大资源组数为 64。

参与节点列表

参与节点列表定义可以主管特定资源组的节点的列表。

可在配置资源组时定义节点列表。参与节点列表可以包含集群中的一部分或全部节点。

通常，该列表包含共享相同数据和磁盘的所有节点。

缺省节点优先级

根据特定资源组的节点列表中节点的位置来识别缺省节点优先级。

节点列表中的第一个节点具有最高节点优先级。该节点也称为资源组的主节点。列在某节点前的节点的节点优先级比当前节点高。

根据资源组的回退策略，当（当前正由优先级较低的节点控制的）资源组的较高优先级的节点加入或重新集成至集群时，它将控制该资源组。也就是说，资源组从优先级较低的节点移到优先级较高的节点。

在所给定的任何时间，资源组都具有由参与节点列表所指定的缺省节点优先级。但是，您选择的各种资源组策略可以覆盖缺省节点优先级，并根据特定资源组将要在集群中移到哪个节点来“创建”实际的节点优先级。

动态节点优先级

通过设置动态节点优先级策略，可使用 RSCT 资源变量（例如，最低的 CPU 负载）来为非并发资源组选择接管节点。在启用动态优先级策略的情况下，接管节点列表的顺序将由发生事件时集群的状态确定，正如所选 RSCT 资源变量所度量的那样。可为不同组设置不同策略，或者为多个组设置相同策略。

主节点

主节点（此资源组中具有最高优先级的节点）是列示在非并发资源组的参与节点列表中的第一个节点。

主节点通常是拥有资源组的节点。非并发资源组可能有主节点，也可能没有主节点。非并发资源组是否具有主节点，这取决于资源组的启动、失败转移和回退行为。

由于集群中的不同更改，组可能不总是在主节点上启动。区分资源组的主节点和当前拥有资源组的节点非常重要。

"主节点"术语不用于并发资源组，因为这些资源组归多个节点拥有。

启动、失败转移和回退

当集群中的条件更改时，PowerHA SystemMirror 会通过将资源组从一个节点移动到另一个节点来确保集群资源的可用性。

PowerHA SystemMirror 通过集群启动时在特定节点或多个节点上激活资源组，或者通过集群条件更改时将资源组移到另一个节点，来管理资源组。以下是集群生命周期中影响 PowerHA SystemMirror 如何管理特定资源组的各个阶段：

集群启动

节点启动且根据选择的资源组启动策略在节点间分配资源组。

节点故障

在该节点上活动的资源组失败转移到另一个节点。

节点恢复

节点重新集成至集群且根据选择的资源组策略可重新获取资源组。

资源故障和恢复

资源组可以失败转移到另一个节点，并且在资源变得可用时可重新获取该资源组。

关闭集群

有不同的方法关闭集群，其中之一可确保资源组移到另一个节点。

每个集群阶段期间，PowerHA SystemMirror 中资源组的行为定义如下：

- 哪个或哪些节点在集群启动时激活资源组
- 集群启动期间在节点上允许获取多少个资源组
- 当拥有资源组的节点发生故障及 PowerHA SystemMirror 需要将资源组移到另一个节点时，哪个节点会接管资源组
- 资源组回退至刚刚加入集群的节点还是留在当前拥有它的节点上。

您选择的资源组策略将确定哪个集群节点最初控制资源组及当初始节点放弃控制时哪个集群节点会接管对该资源组的控制。

这些策略的每种组合都允许您指定对哪个或哪些节点控制资源组进行不同程度的控制。

概括地说，PowerHA SystemMirror 对资源组所有权的关注使大量集群配置成为可能，并在定义集群环境以适应应用程序的特定需求上提供了极大的灵活性。启动、失败转移和回退策略的组合总结了前发行版中所有可用的管理策略，而无需指定修改了"预定义"组类型行为的选项集合。

定义资源组行为时，请注意在集群中可由一个或多个节点接管资源组。

启动、失败转移和回退是描述资源组在不同的集群阶段如何运行的特定行为。注意失败转移和回退的区别是重要的。在各种资源组策略的讨论中这些术语会频繁出现。

启动

启动是指在资源组当前所在的一个或多个节点上或者在该资源组的主节点上激活该资源组。集群启动期间或最初获取节点上的组的时候，会发生资源组启动。

失败转移

失败转移是指在当前节点发生故障之后，将资源组从当前拥有该资源组的节点移到另一个活动节点。新所有者不是重新集成的节点或加入节点。

失败转移仅针对非并发资源组才有效。

回退

回退是指将资源组从它当前所在的节点（不是此资源组的主节点）移到加入该集群或重新集成到该集群的节点。

例如，当针对该资源组的优先级更高的节点加入或重新集成至集群时，它将控制该资源组。也就是说，资源组从优先级较低的节点回退到优先级较高的节点。

定义回退行为仅针对非并发资源组才有效。

资源组策略和属性

可配置资源组以使用特定的启动、失败转移和回退策略。本节将描述资源组属性和方案，并帮助您确定哪些资源组适合您的集群需求。

可通过以下方式使用资源组策略：

- 配置资源组以确保在非高峰时段它们可以在重新集成的节点上再次联机。
- 指定包含某应用程序的资源组是唯一一个将给定首选项且可在启动期间在特定节点上获取的资源组。可通过指定节点分配策略来实现。如果可能会在节点上获取多个非并发资源组，但某特定资源组拥有对保持可用更为重要的应用程序，那么这就是适用的。
- 指定在启动、失败转移和回退时，特定资源组共同在同一节点上保持联机，或者分别在不同节点上保持联机。

资源组启动、失败转移和回退

各个资源组都存在若干策略。

这些策略是：

表 4. 资源组策略

资源组	策略
启动	- 仅在主节点上联机。资源组启动期间，资源组仅在其主节点（优先级最高）上联机。这要求具有最高优先级的节点（也就是资源组的节点列表中的第一个节点）可用。 - 在第一个可用节点上联机。资源组在可用的第一个加入节点上联机。 - 在所有可用节点上联机。资源组在所有节点上联机。 - 使用分配策略联机。每个节点上仅一个资源组联机。
失败转移	- 失败转移至列表中的下一个优先级节点。资源组遵循资源组节点列表中指定的缺省节点优先级顺序。 - 使用动态节点优先级失败转移。在选择该选项之前，先从三个预先定义动态节点优先级策略中选择一个。这些都基于 RSCT 变量，例如可用内存最多的节点。配置了站点时，此选项的故障转移不受支持。 - 脱机（仅在错误节点上）。选择该选项可在发生错误的情况下使资源组在节点上脱机。
回退	- 回退至列表中优先级更高的节点。当优先级更高的节点加入集群时，资源组将回退。如果选择该选项，可使用延迟回退计时器。如果不配置延迟回退策略，那么当优先级更高的节点加入集群时，资源组将立即回退。 - 永不回退。当优先级更高的节点加入集群时，资源组不会回退到该节点。

相关信息:

管理 PowerHA SystemMirror

规划 PowerHA SystemMirror

停留时间、动态节点优先级和回退计时器

可为资源组配置一些其他参数，以指示在启动、失败转移和回退时资源组如何运行。

以下是您可以为资源组配置的其他参数：

停留时间

可通过为当前脱机的资源组指定停留时间来配置资源组的启动行为。当未配置停留时间时，资源组将在加入集群的第一个可用的优先级较高的节点上启动。如果已配置停留时间，那么 PowerHA SystemMirror 将在它激活资源组前的停留时间段的持续时间内等待优先级较高的节点以加入集群。当多个节点同时加入时，指定停留时间可使资源组由优先级较高的节点获取。停留时间是集群范围的属性，如果已配置，那么它将影响已为其选择"在第一个可用节点上联机"启动行为的集群中的所有资源组的启动行为。

分发策略

可配置资源组的启动行为以使用基于节点的分发策略。该策略确保在启动期间，一个节点仅获取一个资源组。有关更多信息，请参阅下一节。

动态节点优先级

可配置资源组的失败转移行为以使用三个动态节点优先级策略中的一个。这些策略基于 RSCT 变量，例如，最多的可用内存或者最少使用的 CPU。PowerHA SystemMirror 会在失败转移时选择与策略最佳匹配的节点，以恢复资源组。

注：配置了站点时，不支持使用"动态节点优先级"选项的故障转移。

延迟回退计时器

可通过指定和分配延迟回退计时器来配置资源组的回退行为，使其在预先定义的重现时间之一（每天、每周、每月和每年）或者在特定的日期和时间发生。这很有用，例如安排资源组回退在非高峰业务时间期间发生。

分配策略

集群启动时，可为资源组使用基于节点的分配策略。

如果为多个资源组选择该策略，那么 PowerHA SystemMirror 将尝试在启动时使每个节点仅获取这些资源组中的一个。这允许您在不同的节点上分配 CPU 密集型应用程序。

相关信息:

管理 PowerHA SystemMirror

规划 PowerHA SystemMirror

网络和资源组

所有资源组都支持在别名判别的网络上配置的服务 IP 标签。

服务 IP 标签可包含在任何非并发资源组中 - 该资源组可具有任何允许的启动策略，"在所有可用节点上联机"除外。

资源组依赖关系

PowerHA SystemMirror 支持资源组排序和资源的定制串行处理，来适应在位于不同资源组的应用程序间存在依赖关系的集群配置。

利用定制的串行处理，您可以指定在处理其他资源组之前处理给定资源组。PowerHA SystemMirror 提供了在资源组（和属于它们的应用程序）间配置父代/子代依赖关系的简单方法，以确保集群事件期间的正确处理。

位置依赖关系策略可用于配置资源组，以便不仅在启动集群时而且在失败转移和回退期间用您期望的方式来分配。可配置依赖关系使所指定的资源组在不同节点或相同节点上联机。PowerHA SystemMirror 使用并行处理（可能的话）和串行处理（必要时）来以合理的顺序处理依赖资源组。您不必定制处理。可在资源组间配置不同类型的依赖关系：

- 父代/子代依赖关系
- 位置依赖关系。

资源组间配置的依赖关系是：

- 使用 SMIT 界面明确指定的
- 确定为集群范围的，不仅在本地节点上
- 保证会在集群中被接受的。

子资源组与父资源组的依赖关系

配置资源组父代/子代依赖关系将使您更容易配置集群以及控制具有多层应用程序的集群，在此类集群中，一个应用程序依赖另一个应用程序的成功启动，且两个应用程序都需要使用 PowerHA SystemMirror 以保持高可用性。

以下示例说明了父代/子代依赖关系行为：

- 如果资源组 A 依赖资源组 B，那么在节点启动时，在集群的任何节点获取资源组 A 之前，资源组 B 必须处于联机状态。在失败转移时，顺序则相反：必须先将资源组 A 脱机，然后资源组 B 才能脱机。
- 另外，如果资源组 A 依赖资源组 B，那么在节点启动或节点重新集成期间，资源组 A 不能在资源组 B 联机之前联机。如果将资源组 B 脱机，那么资源组 A 也将脱机，因为它依赖资源组 B。

资源组间的依赖关系为构建具有多层应用程序的集群提供了一种可预测和可靠的方法。有关可使用依赖关系资源组的典型集群环境的更多信息，请参阅“具有多层应用程序的集群配置”。这些术语描述了资源组之间的父代/子代依赖关系：

- 父资源组必须处于联机状态才能启动依赖它的资源组（子代）。
- 子资源组依赖父资源组。仅在已激活父资源组之后，它才能在集群的任何节点上激活。通常，子资源组依赖父资源组提供的一些应用程序服务。

在资源组释放时（例如失败转移或停止集群服务期间），PowerHA SystemMirror 将在父资源组脱机之前，使子资源组脱机。下图说明了资源组之间的父代/子代依赖关系。

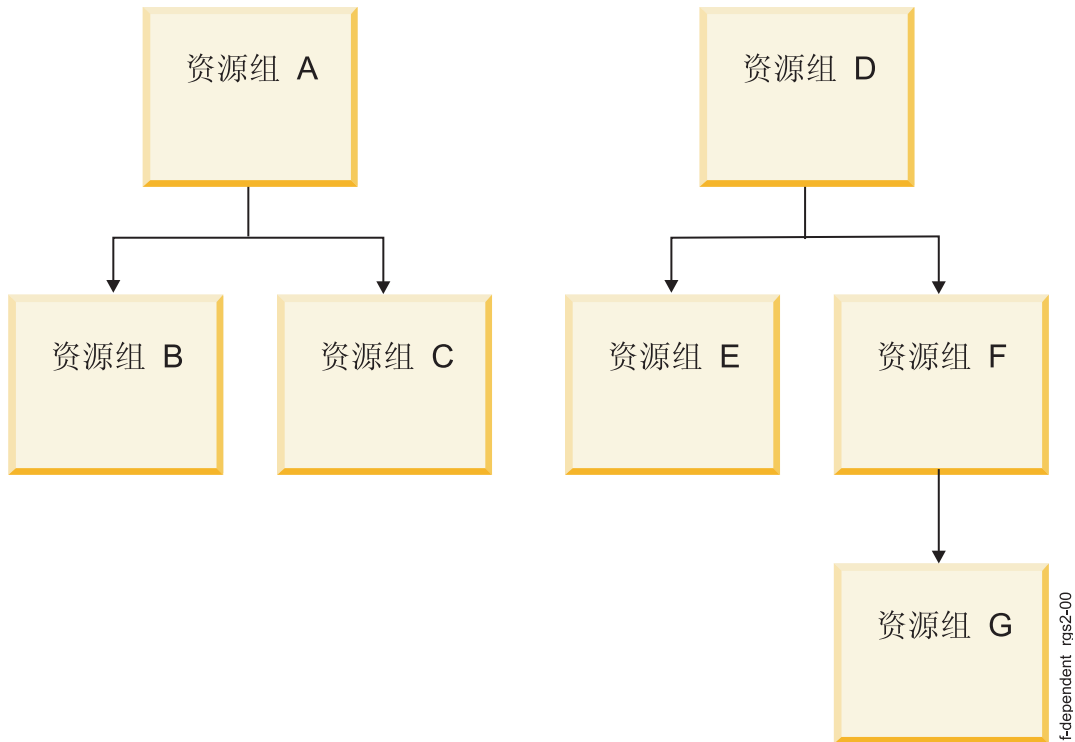


图 7. 资源组间两级和三级依赖关系示例

该示例显示了根据这些指南和限制构造的关系：

- 可配置一种依赖关系类型，其中在子（依赖项）资源组可在节点上激活之前，父资源组必须在集群的任何节点上联机。
- 根据将资源组放置于给定依赖关系链接的哪一端的不同，资源组既可作为父资源组也可作为子资源组。
- 可为资源组指定三级依赖关系。
- 不能在资源组间指定循环依赖关系。

下列准则和局限性也适用于资源组之间的父代/子代依赖关系：

- 可在集群服务运行时，在资源组间添加、更改或删除依赖关系。
- 在删除两个资源组间的依赖关系时，仅这些资源组间的链接会从 PowerHA SystemMirror 配置数据库中除去。资源组不会删除。
- 在父资源组失败转移期间，包含应用程序的子资源组将暂时脱机，然后在任何可用节点上联机。属于子资源组的应用程序也会停止并重新启动。

相关概念：

第 76 页的『具有多层应用程序的集群配置』

可以利用父子依赖资源组的典型集群配置是其中一个应用程序（例如，WebSphere®）依赖于另一个应用程序（例如，DB2®）的环境。

资源组位置依赖关系

除了针对各个资源组以及父代/子代依赖关系的各种策略以外，PowerHA SystemMirror 还提供了处理所有资源组互相依赖关系的策略。PowerHA SystemMirror 可识别这些关系并以合理的顺序处理资源组。

可以配置资源组，以使：

- 所指定的两个或多个资源组在同一节点上始终联机。它们会在同一节点上启动、失败转移和回退。
- 所指定的两个或多个资源组在不同节点上始终联机。它们会在不同节点上启动、失败转移和回退。可为资源组指定优先级，以便在发生失败转移和回退时首先处理最关键的资源组。
- 所指定的两个或多个资源组（具有所复制的资源）在同一节点上始终联机。

在您为各个资源组配置了所给定的位置依赖关系之后，它们将构成一个集合，集群管理器会将其作为一个整体进行处理。当用 **cIRGmove** 命令直接移动资源组时，将应用以下规则：

- 如果某个资源组参与在同一节点上联机的依赖关系集合，那么您只能在同一节点集中的所有其他资源组当前处于联机状态的节点上使其联机。对于集群管理器，这条规则也适用。
- 如果某个资源组参与在同一站点上联机的依赖关系集合，那么您只能在同一站点集的其他资源组当前处于联机状态的节点上使其联机。对于集群管理器，这条规则也适用。
- 如果某个资源组参与在不同节点上联机的依赖关系集合，那么您只能在不主管不同节点依赖关系集中的任何其他资源组的节点上使其联机。（这对集群管理器是相同的规则。）但是，当移动属于该集合的资源组时，将把优先级视为相等的值，而当 PowerHA SystemMirror 使这些组联机时，它将考虑优先级。

样本位置依赖关系模型

一家虚构公司 XYZ Publishing 遵循一个业务连续性模型，该模型涉及到对用于开发 Web 内容的不同平台划分优先级。位置策略用于严格确保一些资源组分别位于不同的节点上而其他资源组则共同位于同一节点上。

下图显示了位置依赖关系模型的一个示例，其中有三个节点，用于三个不同的应用程序及其相关联的数据库。每个应用程序的数据库必须始终与该应用程序在同一节点上运行。例如，生产数据库必须包含生产应用程序。

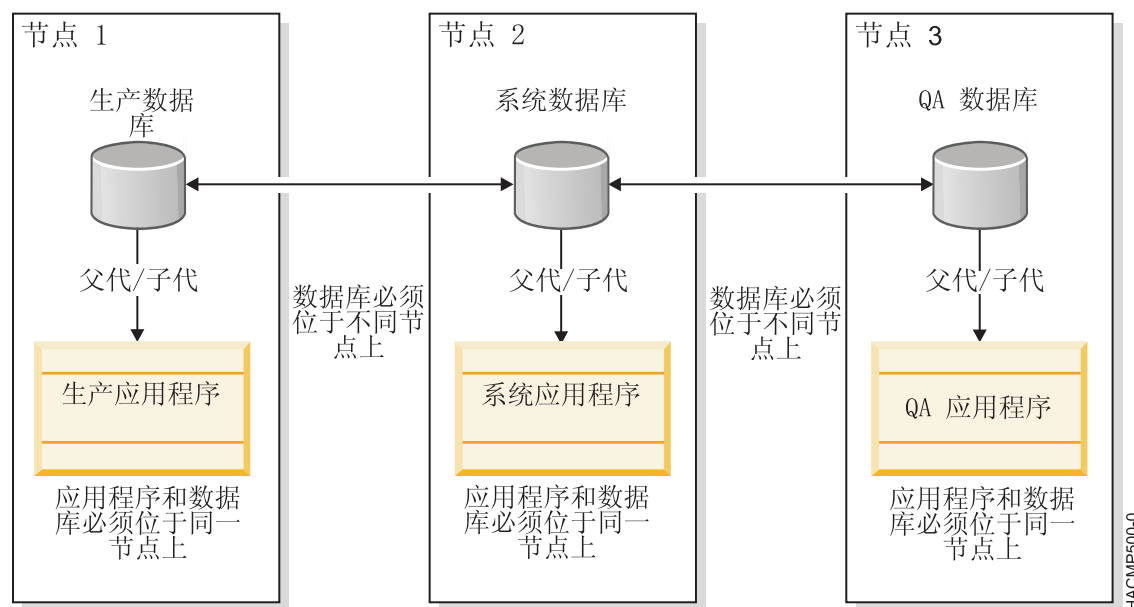


图 8. 分别用于不同数据库的节点 1、2 和 3 的示例

相关信息:

管理 PowerHA SystemMirror

规划 PowerHA SystemMirror

站点和资源组

大多数 PowerHA SystemMirror 配置不包括站点，并且使用“忽略”管理策略（这是缺省情况）。如果您已经安装了 PowerHA SystemMirror Enterprise Edition 组件用于灾难恢复，那么在地理位置上存在一定距离间隔的站点之间分配集群节点，并选择其中一种站点间管理策略。

将您想要复制的资源包括在资源组中。对于所复制资源组的主实例定义启动策略、失败转移策略和回退策略。资源组在主实例中处于联机状态。具有资源组的主实例的节点将激活该组的所有资源。在位于另一个站点上作为备份节点的节点上会激活次要实例（复制）。站点间管理策略与资源组启动策略、失败转移策略和回退策略相结合共同确定主实例首先位于的站点，以及如何处理站点之间的失败转移和回退。

存在下列选项用于配置资源组站点间管理策略：

- 首选主站点
- 在任一站点上联机
- 在两个站点上都联机

如果您为集群定义站点，那么当您为想要复制的每个资源组定义启动策略、失败转移策略和回退策略时，可以对主站点上的节点和另一个站点（辅助站点）上的节点指定资源组。该资源组的主要实例在主站点上运行，次要实例在辅助站点上运行。

如果您具有同时存在的另一个资源组，请将其定义为在所有节点上运行。在这种情况下，您可以对站点间管理策略选择“在两个站点上都联机”选项。然后，这两个站点上的实例都处于活动状态（不存在次要实例）。您还可以选择其他站点间管理策略，以便同时存在的资源组在一个站点的所有节点上都处于联机状态，并且在另一个站点上具有备份实例。

注：当您具有已定义的站点和非并发资源组时，在一个站点节点上使用“未管理”选项来停止集群服务时，会使该资源组在该站点的所有节点上处于“未管理”状态。

您还可以使用 **cIRGmove** 实用程序来移动资源组的主实例并使其跨越站点边界。然后，PowerHA SystemMirror 根据需要重新分发对等次要实例（如果由于配置需求而不允许执行移动操作，那么将对您发出警告）。

当您移动资源组的主要实例，而其次要实例处于脱机状态时，该次要实例将保持脱机。PowerHA SystemMirror 不会自动使该资源组的次要实例联机。

相关信息：

管理指南

规划指南

支持 PowerHA SystemMirror 的硬件

可以使用不同类型的 IBM 硬件与 PowerHA SystemMirror 一起实现基本级别的高度可用环境。

有关 PowerHA SystemMirror 所支持的硬件的摘要，请参阅 IBM Techdocs: PowerHA hardware support matrix (<http://www.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/TD101347>)。

相关信息：

安装 PowerHA SystemMirror

OEM 磁盘、卷组和文件系统调整

PowerHA SystemMirror 集群软件

本节描述了实现高度可用环境的 PowerHA SystemMirror 软件。

PowerHA SystemMirror 节点的软件组件

PowerHA SystemMirror 节点的软件组件具有很多层。

下图显示了 PowerHA SystemMirror 集群中节点上软件的层：



图 9. PowerHA SystemMirror 集群节点的模型

以下列表对每层进行了描述：

应用层

通过由 PowerHA SystemMirror for AIX 提供的服务使任何应用程序高度可用。

PowerHA SystemMirror for AIX 层

识别集群内的更改且协调 AIX 功能部件的使用来为关键数据和应用程序创建高度可用环境的软件。PowerHA SystemMirror 通过提供更多服务来补充较低层，以启用应用程序和相关联的通信，并使数据存储服务高度可用。

RSCT 层

IBM Reliable Scalable Cluster Technology (RSCT) 服务与 AIX 封装在一起。RSCT 层提供了设施，以在分布式环境或集群环境中通过使用可靠消息传递来进行集群活动的同步和协调。RSCT 包括"资源监视与控制"(RMC)、"配置资源管理器"和"组服务"组件。

Cluster Aware AIX 层

AIX 操作系统的此扩展提供了设施来监视节点成员资格、网络接口的运行状况和事件通知。

AIX 层

为 PowerHA SystemMirror 集群提供底层支持的软件。

逻辑卷管理器 (LVM) 子系统层

它在逻辑级别管理数据存储。

TCP/IP 子系统层

它为 PowerHA SystemMirror 集群提供通信支持。

相关概念:

第 45 页的『IBM Reliable Scalable Cluster Technology 可用性服务』

IBM Reliable Scalable Cluster Technology (RSCT) 高可用性服务提供更大的可伸缩性，向已分配的子系统通知软件故障，并在软件堆栈的所有子系统中协调恢复和同步。RSCT 向 PowerHA 集群管理器报告节点和网络故障。

集群管理器

集群状态的更改都可作为集群事件。在每个节点上，集群管理器都为这些事件（例如，应用程序故障事件）监视本地硬件和软件子系统。

集群管理器运行一个或多个事件脚本（例如，重新启动应用程序脚本）以响应这种事件。所有节点上的集群管理器交换消息以协调响应某事件所需的任何操作。

集群管理器是运行在每个集群节点上的守护程序。集群管理器的主要任务是响应计划外事件，例如从软件和硬件故障中恢复，或用户启动的事件，例如加入节点事件。RSCT 子系统将与节点和网络有关的事件通知集群管理器。

集群管理器与其他 PowerHA SystemMirror 守护程序的连接

集群管理器收集与节点和接口的集群状态更改有关的信息。集群信息程序 (Clinfo) 从集群管理器获取该信息，并允许客户机与 Clinfo 通信以了解集群的状态更改。该集群状态信息存储在 PowerHA SystemMirror 管理信息库 (MIB) 中。

相关概念:

第 83 页的『PowerHA SystemMirror 配置过程和设施』

这些主题提供了 PowerHA SystemMirror 集群配置过程以及与 PowerHA SystemMirror 软件一起提供的管理工具的概述。

第 45 页的『IBM Reliable Scalable Cluster Technology 可用性服务』

IBM Reliable Scalable Cluster Technology (RSCT) 高可用性服务提供更大的可伸缩性，向已分配的子系统通知软件故障，并在软件堆栈的所有子系统中协调恢复和同步。RSCT 向 PowerHA 集群管理器报告节点和网络故障。

PowerHA SystemMirror 软件组件

PowerHA SystemMirror 软件具有很多组件。

集群安全通信子系统

PowerHA SystemMirror 具有增加系统间通信安全性的公共通信基础结构。集群实用程序使用在每个节点上都运行的集群通信守护程序在节点间通信。

因为只有一个公共通信路径，所以所有通信都是可靠安全的。尽管大部分组件通过集群通信守护程序通信，但是以下组件使用另一种机制进行节点间通信：

组件	通信方法
集群管理器	RSCT
脉动信号消息传递	Cluster Aware AIX
集群信息程序 (Clinfo)	SNMP

针对需要其他安全性的用户，PowerHA SystemMirror 为在集群节点间发送的消息提供了消息认证和加密。

连接认证

标准安全性方式针对访问列表检查源 IP 地址，检查源端口的值是否在 571 和 1023 之间，并针对远程命令执行使用最小特权原则。标准安全性是缺省安全性方式。对于添加的安全性，可以为节点间连接设置 VPN，以进行 PowerHA SystemMirror 节点间通信。

消息认证和加密

消息认证和消息加密为在集群节点间发送的 PowerHA SystemMirror 信息提供了其他安全性。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转换为初始格式。可以配置安全性选项和使用 SMIT 界面分配加密密钥的选项。

IBM Reliable Scalable Cluster Technology 可用性服务

IBM Reliable Scalable Cluster Technology (RSCT) 高可用性服务提供更大的可伸缩性，向已分配的子系统通知软件故障，并在软件堆栈的所有子系统中协调恢复和同步。RSCT 向 PowerHA 集群管理器报告节点和网络故障。

PowerHA SystemMirror 和 RSCT 软件堆栈在每个集群节点上运行。PowerHA SystemMirror 集群管理器从多个渠道获取可能的故障指示：

- RSCT 监视网络设备的状态
- AIX LVM 监视卷组和磁盘的状态
- 应用程序监视器监视应用程序的状态。

PowerHA SystemMirror 集群管理器驱动组件故障事件中的集群恢复操作。运行在每个节点上的 RSCT 与其对等 RSCT 交换脉动信号，使它能够监视集群中其他节点的可用性。如果脉动信号停止，那么对等系统将驱动恢复过程。对等系统采取必要的操作使关键应用程序运行并确保数据未损坏或丢失。RSCT 服务包含以下组件：

- 资源监视和控制 (PowerHA SystemMirror 的先前版本使用事件管理子系统)。提供一组高可用性服务的已分配的子系统。它通过匹配关于系统资源的状态信息和有关客户机程序关注的资源条件的信息来创建事件。客户机程序可依次使用事件通知来触发从系统故障中的恢复。
- 组服务。对运行在一组节点上的应用程序的状态更改进行协调和监视的系统范围的高度可用的设施。组服务在高度可用应用程序的设计和实现方面，以及多个应用程序的一致性恢复方面提供帮助。它在集成的框架中完成这两项不同的任务。

下图显示了组成 PowerHA SystemMirror 体系结构的主要组件：

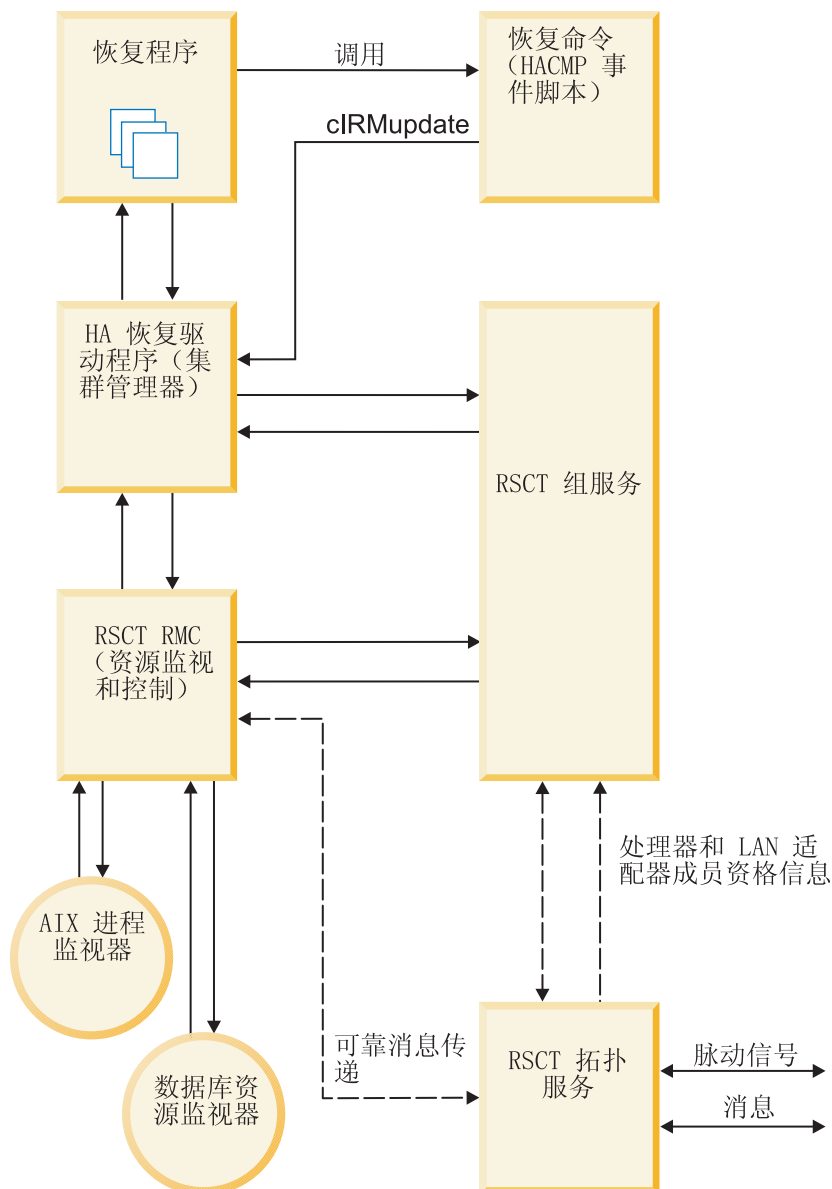


图 10. PowerHA SystemMirror 体系结构的组件

集群管理器和 SNMP 监视程序

PowerHA SystemMirror 集群是动态的，且可以随时间在它的状态中经历各种转换。

例如，节点可以加入或离开集群，或者另一个 IP 标签可以替换物理网络接口卡上的服务 IP 标签。这些更改中的每一个都会影响集群的组合，尤其是当高度可用的客户机和应用程序必须使用由集群节点提供的服务时。

SNMP 支持

集群管理器会提供对客户机应用程序的简单网络管理协议 (SNMP) 支持。SNMP 是用于监视和管理基于 TCP/IP 网络的业界标准规范。SNMP 包括一个协议、一个数据库规范和一组数据对象。这组数据对象构成了管理信息库 (MIB)。SNMP 提供了包括诸如 IP 地址信息和活动 TCP 连接数目信息的标准 MIB。标准 SNMP 代理是 **snmpd** 守护程序。

可以扩展 SNMP 以包括特定于企业的 MIB，此 MIB 包含与离散环境或应用程序相关的信息。在 PowerHA SystemMirror 中，集群管理器会维护在它的 MIB 中定义的对象信息，并将该信息传递到专用的网络监视或网络管理站。

PowerHA SystemMirror MIB

集群管理器在特殊的 PowerHA SystemMirror MIB 中维护集群状态信息。当集群管理器在集群节点上启动时，它将使用 SNMP 守护程序 **snmpd** 注册，然后持续收集集群信息。集群管理器可维护 PowerHA SystemMirror MIB 中集群的更新拓扑，因为它跟踪事件和集群的结果状态。

相关信息:

客户机应用程序编程

集群信息程序

集群信息程序 (Clinfo)，**clinfo** 守护程序，是基于 SNMP 的监视器。

运行在客户机或集群节点上的 Cinfo 可查询 MIB 以获取更新的集群信息。通过 Cinfo，可使 PowerHA SystemMirror 集群、节点和网络的相关状态信息对客户机和应用程序可用。客户机可分为两类：简单型和智能型。

- 简单型客户机将集群机器群看作单个实体。如果集群节点发生故障，那么在未启用 IP 地址接管 (IPAT) 的情况下必须重新启动客户机（或者至少必须重新连接到节点）。
- 另一方面，智能型客户机可以集群感知 - 它可对节点故障作出适当的反应，连接到备用节点，也许还对用户掩饰故障。这样的智能型客户机必须了解集群状态。

PowerHA SystemMirror 软件可通过经由集群管理器和 Cinfo API 功能向客户机提供集群状态更改通知，来向客户机扩展高度可用的服务器、数据和应用程序的优点。

响应集群更改

无论何时发生集群、网络或节点事件，Clinfo 都会调用 **/usr/es/sbin/cluster/etc/clinfo.rc** 脚本。缺省情况下，**clinfo.rc** 脚本会清空系统的 ARP 高速缓存以将更改反映给网络 IP 地址，而且直到另一个地址响应 *ping* 请求时它才会更新高速缓存。通常，如果启用 PowerHA SystemMirror 硬件地址交换设施，那么清空 ARP 高速缓存就是不必要的，因为硬件地址交换将维护 IP 地址和硬件地址间的关系。

在交换式以太网网络中，您可能需要清空 ARP 高速缓存以确保将新的 MAC 地址传递给交换机。

如果希望执行进一步操作，可向 **clinfo.rc** 脚本添加逻辑。

Cinfo API

Cinfo API 可向应用程序开发者提供 C 和 C++ 语言接口，用于访问集群状态信息。PowerHA SystemMirror 软件包含两个版本的 Cinfo API：一个针对单线程应用程序，一个针对多线程应用程序。Clinfo 及其相关联的 API 允许开发者编写识别和响应集群中更改的应用程序。

相关概念:

第 14 页的『PowerHA SystemMirror 集群节点、网络和脉动信号传递概念』

本节介绍了与集群拓扑相关的主要概念，以及在整个文档和 PowerHA SystemMirror 用户界面中使用的定义。

第 50 页的『确保应用程序可用性』

本节描述了 PowerHA SystemMirror 软件如何通过确保集群组件的可用性来确保应用程序可用性。

PowerHA SystemMirror 可消除所有关键系统组件的单点故障，并消除针对大部分例程集群维护任务的计划内停机时间需求。

相关信息:

客户机应用程序编程

高度可用的 NFS 服务器

高度可用的 NFS 服务器功能包含在 PowerHA SystemMirror 产品子系统中。

如果主 NFS 服务器发生故障，高度可用的 NFS 服务器将允许备份处理器恢复当前的 NFS 活动。NFS 服务器特殊功能包括网络文件系统 (NFS) 上高度可用的修改和锁定。您可以执行以下操作：

- 使用可靠的 NFS 服务器能力，它可以保留锁定和副本缓存（如果使用 NFS V2 和 V3，仅双节点集群）。
- 为 NFS 交叉装配指定网络
- 在目录级别定义 NFS 导出和交叉装配
- 为 NFS 导出的目录和文件系统指定导出选项
- 配置两个节点以使用 NFS。

PowerHA SystemMirror 集群最多可以包含 16 个节点。使用 NFS V2 和 V3 的集群最多可以具有两个节点，使用 NFS V4 的集群最多可以具有 16 个节点。

共享外部磁盘访问

PowerHA SystemMirror 软件支持两种共享外部磁盘访问的方法：非并发访问和并发访问。

非并发共享外部磁盘访问：

在非并发环境中，只有一个节点在给定时间内对共享外部磁盘具有访问权。

如果该节点发生故障，那么对等节点中的一个将获取磁盘，装配定义为资源的文件系统，并重新启动应用程序以恢复关键服务。通常，这将耗时 30 - 300 秒，取决于文件系统的数目和大小。

非并发配置可使用：

- SCSI 磁盘
- SCSI 磁盘阵列
- 串行磁盘
- 光纤通道直接连接的磁盘
- 光纤通道 SAN 连接的磁盘

要防止发生故障的磁盘成为单点故障，共享卷组中的每个逻辑卷都应该使用 AIX LVM 设施镜像。如果正在使用 IBM Enterprise Storage System 或其他受支持的 RAID 阵列，请不要使用 LVM 镜像。RAID 设备可提供它们自己的数据冗余。

可以单机方式运行的大部分软件都可无需修改而由 PowerHA SystemMirror 软件管理。

非并发访问通常不需要对服务器程序（例如，数据库管理系统）的任何代码更改，或对应用程序的任何代码更改，以提供高度可用的解决方案。对于最终用户，节点故障看起来像非常快速的机器重启。其中一个剩余节点将拥有发生故障的节点的资源组的所有权，并重新启动高度可用的应用程序。日志文件系统即本机 AIX 文件系统，保证文件系统的完整性。服务器程序保证事务数据完整性。

用户只需登录到继续存在的其中一个节点并重新启动应用程序。登录和应用程序重新启动过程可由 PowerHA SystemMirror 软件驱动。在某些 PowerHA SystemMirror 配置中，用户不必采取任何操作就可继续 - 他们只是会在失败转移期间遭遇延迟。

并发共享的外部磁盘访问：

并发访问功能可增强 PowerHA SystemMirror 集群提供的优点。并发访问允许同时访问已连接到多个（最多 16 个）节点的磁盘子系统上的卷组。

使用并发访问，集群可以提供几乎连续的数据可用性，它可与容错抗衡，但成本要低得多。另外，并发访问提供更高的性能，使应用程序开发更轻松，并允许水平增长。由于并发访问提供了对来自多个节点的数据的同时访问，所以可能需要其他工具来防止多个节点以互相冲突的方式修改相同的数据块。PowerHA SystemMirror 软件提供了 Clinfo 程序，该程序准备了在并发访问环境中运行的应用程序。Clinfo API 提供了一个 API，通过该 API 应用程序可能会变得“集群感知”。在本章前面对 Clinfo 工具进行了描述。

以下列表包括进行并发共享外部磁盘访问的优势：

- 明显的恢复增加可用性。并发访问显著减少了失败转移所需的时间 - 有时只需几秒钟 - 因为对等系统已对共享磁盘具有物理访问权，且正在运行它们自己的应用程序实例。

在并发访问环境中，失败转移基本都涉及到从发生故障的处理器中将进行中的事务撤回。运行在剩余节点上的服务器软件负责恢复崩溃所影响到的任何部分的事务。因为所有节点都可对数据进行并发访问，所以客户机/服务器应用程序可立即在剩余的节点上重试失败的请求，这些剩余节点可继续处理进入的事务。

- 利用多处理器增加吞吐量。应用程序不再受单个处理器吞吐量的限制。相反，多个应用程序实例可同时在多个处理器上运行。需要更多的处理能力时，可将更多系统添加到集群中以增加吞吐量。
- 单个数据库映像使应用程序开发和维护更轻松。在非并发环境中，提高性能的唯一方法是对应用程序及其数据进行分区。将代码和数据分成若干片段使应用程序开发和维护变得更加复杂。

分割数据库需要高级的专业技能以确保数据和工作负载能在若干处理器中平均分配。并发访问环境中并不需要分区代码和数据。要增加吞吐量，运行在不同处理器上的相同应用程序的多个实例可同时对共享外部磁盘上的数据库进行访问。并发配置可以使用：

- SCSI 磁盘
- SCSI 磁盘阵列
- 串行磁盘
- 光纤通道直接连接的磁盘
- 光纤通道 SAN 连接的磁盘

在创建并发访问逻辑卷时，请使用 LVM 镜像来避免磁盘发生单点故障，提供自己镜像的 RAID 磁盘子系统除外。并发访问不支持使用日志文件系统。因此，数据库管理器必须直接写入共享卷组中的原始逻辑卷或 hdisk。应用程序必须使用某种方法仲裁对共享数据的所有请求。大部分商业 UNIX 数据库都提供了使其与 PowerHA SystemMirror 软件兼容的锁定模型。请联系您的数据库供应商以确定特定应用程序是否支持并发访问处理。

相关概念：

『增强的并发方式』

AIX 提供了并发方式的新形式：增强的并发方式。在增强的并发方式中，并发逻辑卷管理器 (CLVM) 实例通过 AIX 中 Reliable Scalable Cluster Technology (RSCT) 设施的"组服务"组件协调节点间的更改。PowerHA SystemMirror 自动将共享卷组创建为增强的并发方式，并将现有的共享卷组转换为增强的并发方式

并发资源管理器

PowerHA SystemMirror 的并发资源管理器提供对高度可用集群中的共享磁盘的并发访问，允许在接管期间采取合适的操作以满足业务需求。

并发资源管理器对所有类型的磁盘上的共享卷组添加增强型并发支持，并对受支持的 RAID 磁盘子系统进行并发共享访问管理。

增强的并发方式：

AIX 提供了并发方式的新形式：增强的并发方式。在增强的并发方式中，并发逻辑卷管理器 (CLVM) 实例通过 AIX 中 Reliable Scalable Cluster Technology (RSCT) 设施的"组服务"组件协调节点间的更改。PowerHA SystemMirror 自动将共享卷组创建为增强的并发方式，并将现有的共享卷组转换为增强的并发方式

"组服务"协议通过集群节点间的通信链路流动。作为到多个节点的附件的受 PowerHA SystemMirror 支持的任何磁盘都可包含在增强的并发方式的卷组中。

补充集群软件

多种其他工具可辅助您在 AIX 环境中有效地构建、管理和扩展高可用性集群。

包括：

- General Parallel File System (GPFS™) for AIX，这是允许用户对跨多个磁盘驱动器和多个节点的文件进行共享访问的集群范围的文件系统。
- Workload Manager for AIX 提供应用程序间的资源平衡。
- 配置 Oracle、DB2、WebSphere 和 PowerHA SystemMirror 集群中的其他应用程序的智能帮助软件。

相关信息：

开发 PowerHA SystemMirror 的 Smart Assist 应用程序

Smart Assist for PowerHA SystemMirror

确保应用程序可用性

本节描述了 PowerHA SystemMirror 软件如何通过确保集群组件的可用性来确保应用程序可用性。PowerHA SystemMirror 可消除所有关键系统组件的单点故障，并消除针对大部分例程集群维护任务的计划内停机时间需求。

概述：应用程序可用性

高度可用集群的关键方面是它能够对那些可能导致所提供的关键服务中断的更改进行检测和响应。即使关键系统组件（例如，网络接口卡）不再可用，PowerHA SystemMirror 软件仍然允许集群继续提供对安装版本至关重要的应用程序服务。

当某组件不可用时，PowerHA SystemMirror 软件能够检测损失，并将工作负载从该组件转移到集群中的其他组件上。在规划高度可用的集群时，请尝试确保关键组件不会成为单点故障。

可使用 PowerHA SystemMirror 的智能帮助来管理高度可用的应用程序，例如 DB2、SAP、Oracle 和 WebSphere。可将这些应用程序看作单个 PowerHA SystemMirror 实体，该实体在 PowerHA SystemMirror 软件中显示为一个应用程序名称。

PowerHA SystemMirror 软件使用称为“动态重新配置”的进程，在执行例程维护任务时允许集群继续提供应用程序服务。在动态重新配置中，可更改正在运行的集群中的组件，例如不必停止和重新启动集群服务而添加或删除节点或网络接口。已更改的配置将以动态形式成为活动配置。还可动态地替换存在故障的磁盘。以下几个章节将概念性地描述如何使用 PowerHA SystemMirror 软件，以：

- 消除集群中的单点故障。
- 利用动态重新配置、资源组管理和集群管理 (C-SPOC) 这几个实用程序使 PowerHA SystemMirror 集群中的计划内停机时间最短。
- 利用快速恢复功能并通过为资源组指定延迟回退计时器策略来使计划外的停机时间最短。
- 使执行磁盘接管所需的时间最短。
- 解释和模拟集群事件。

注：您可能需要在关键组件发生故障且集群继续提供应用程序可用性时监视集群活动。

相关概念:

第 83 页的『PowerHA SystemMirror 配置过程和设施』

这些主题提供了 PowerHA SystemMirror 集群配置过程以及与 PowerHA SystemMirror 软件一起提供的管理工具的概述。

相关信息:

Smart Assit for PowerHA SystemMirror

消除 PowerHA SystemMirror 集群中的单点故障

通过消除单点故障 (SPOF)，PowerHA SystemMirror 软件使您能够构建既高度可用又可扩展的集群。当关键集群功能是由单个组件提供时会存在单点故障。

如果该组件发生故障，那么集群没有其他方法提供该功能且关键服务将不可用。例如，如果关键应用程序的所有数据都位于未镜像的单个磁盘上，且该磁盘发生故障，那么该磁盘将是整个系统的单点故障。客户机节点将无法访问该应用程序直到该磁盘上的数据恢复。

PowerHA SystemMirror 集群中的潜在单点故障

要高度可用，集群必须没有单点故障。

PowerHA SystemMirror 为以下集群组件提供了恢复选项：

- 节点
- 应用程序
- 网络和网络接口
- 磁盘和磁盘适配器。

当目标是消除所有单点故障时，可能需要进行折衷。存在与消除单点故障相关联的成本。例如，冗余硬件会提高成本。消除单点故障的成本应该与该组件发生故障时失去服务的成本比较。PowerHA SystemMirror 软件的目的是提供可不断增长以满足未来处理需求的，划算的高度可用的计算环境。

消除节点作为单点故障

节点离开集群，要么是通过计划内转换（节点关闭或停止节点上的集群服务），要么是由于故障。

当监视邻近节点的节点在定义的时间段内停止接收脉动信号流量时，节点故障就会开始。如果其他集群节点认为该故障是节点故障，发生故障的节点就将从集群中除去，其资源由配置了接管能力的节点接管。例如，活动节点可控制在发生故障的节点上配置的共享磁盘。或者，活动节点可以伪装成发生故障的节点（通过获取其服务 IP 地址）并运行发生故障的节点的进程，同时又保持它自己的进程。因此，客户机应用程序可切换至剩余节点，以使用共享磁盘和处理器服务。PowerHA SystemMirror 软件提供以下设施来处理节点故障：

- 磁盘接管
- 通过 IP 别名进行 IP 地址接管
- 通过替换 IP 来进行 IP 地址接管（具有或不具有硬件地址接管）

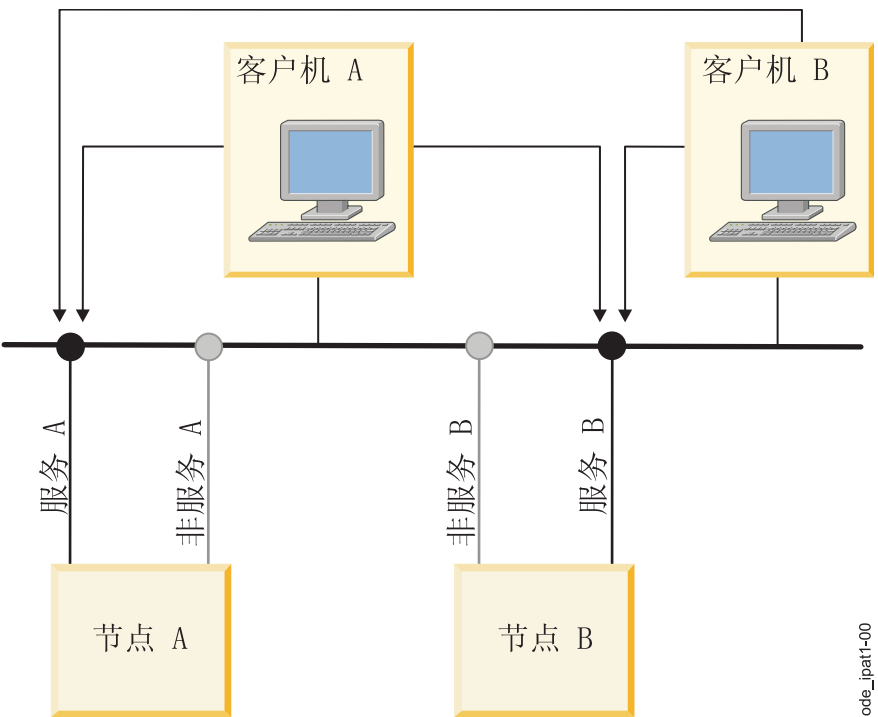
磁盘接管：

在 PowerHA SystemMirror 环境中，共享磁盘物理连接到多个节点。

并发环境中的磁盘接管

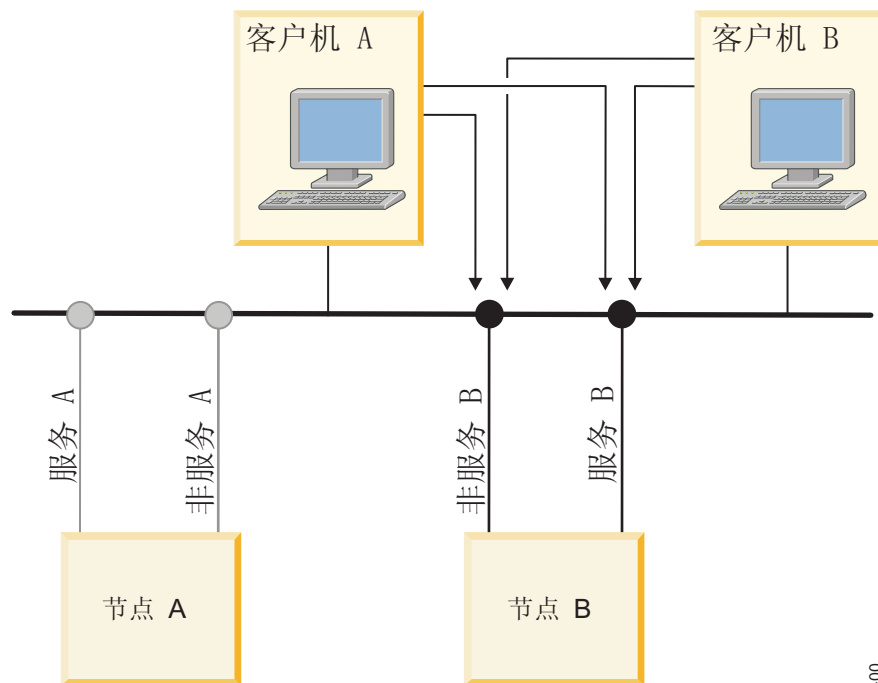
在并发访问配置中，共享磁盘积极地同时连接到多个节点。因此，当节点离开集群时不需要磁盘接管。

下图说明了并发环境中的磁盘接管。



每个节点都提供了单独的网络服务

图 11. 磁盘接管之前的并发访问配置



节点 B 在它的非服务接口上接管节点 A 的 IP 标签，并向此接口上的客户机提供节点 A 的应用程序服务。

c_2node_ipat2-00

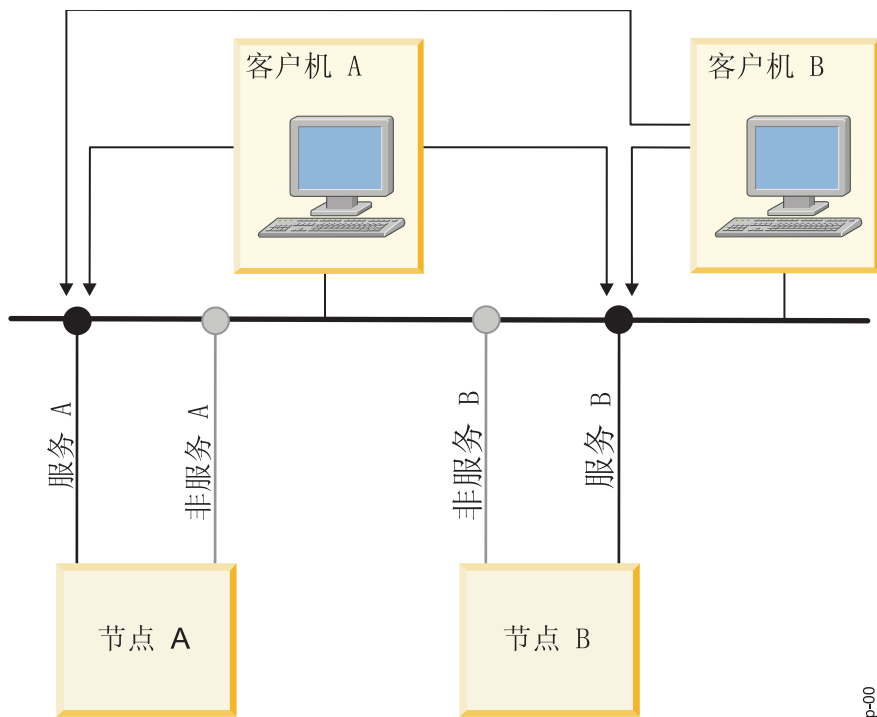
图 12. 磁盘接管之后的并发访问配置

IP 地址接管：

IP 地址接管 (IPAT) 是一种联网能力，它允许节点获取已脱离集群的节点的网络地址。

将提供给客户机的服务绑定到特定的 IP 地址时，即，将向客户机提供服务所经由的服务 IP 标签包含为集群资源组中的资源时，PowerHA SystemMirror 集群中就需要 IP 地址接管。如果某剩余节点简单地进行磁盘和应用程序接管，而不是执行 IPAT，那么客户机将无法在指定的服务器 IP 地址上继续使用应用程序。PowerHA SystemMirror 使用“通过别名的 IPAT”。

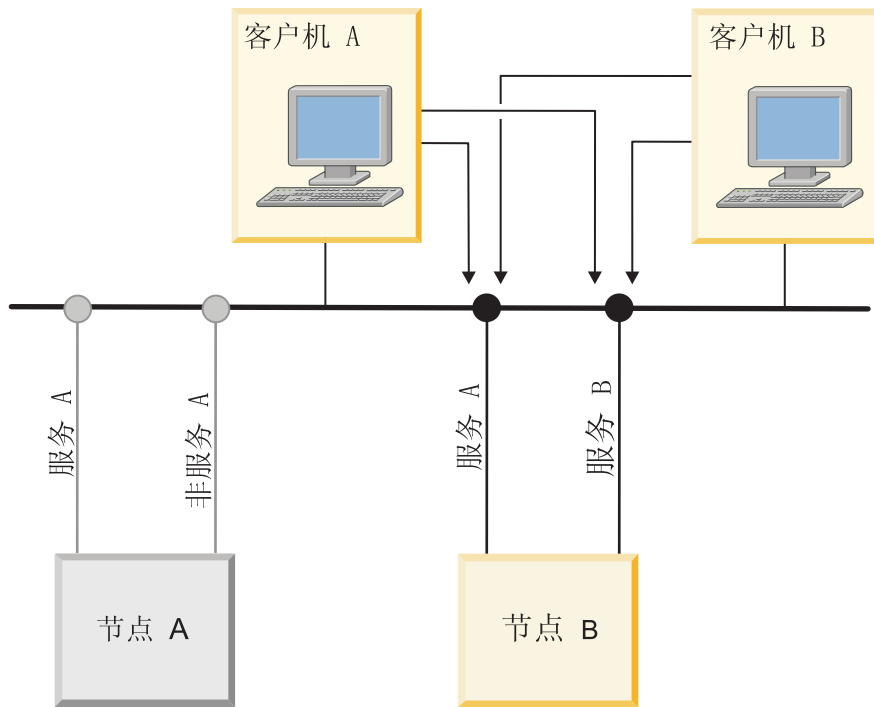
下图说明了 IP 地址接管。



每个节点都提供了单独的网络服务。

c_b4ip-00

图 13. 通过替换 IP 来进行 IP 地址接管之前的配置



节点 B 在其非服务接口上采用节点 A 的地址，并向客户机提供节点 A 的网络服务。

c_atip-00

图 14. 通过替换 IP 来进行 IP 地址接管之后的配置

相关概念:

第 27 页的『通过 IP 别名进行 IP 地址接管』

可以使用 AIX 操作系统的 IP 别名判别网络能力在某些类型的网络中配置 IP 地址接管。

消除应用程序作为单点故障

创建 PowerHA SystemMirror 集群的主要原因是为任务关键型应用程序提供高度可用的环境。

例如，PowerHA SystemMirror 集群可以运行服务客户机应用程序的数据库服务器程序。客户机通过访问存储在共享外部磁盘上的数据库，来向响应客户机请求的服务器程序发送查询请求。

在 PowerHA SystemMirror 集群中，这些关键的应用程序可能是单点故障。要确保这些应用程序的可用性，配置为接管离开集群的节点的资源的节点也应该重新启动应用程序，使它们仍然对客户机进程可用。使应用程序高度可用，可通过使用：

- 应用程序控制器
- 集群控制
- 应用程序监视器
- 应用程序可用性分析工具

要让 PowerHA SystemMirror 控制应用程序，可创建一个应用程序控制器集群资源，该资源将用户定义的服务器名称与用户提供的编写好的用于启动和停止应用程序的脚本名称关联。通过定义应用程序控制器，PowerHA SystemMirror 可在发生失败转移时启动接管节点上的另一个应用程序实例。

可在没有应用程序控制器的情况下使某些应用程序高度可用。您可以通过将这样的应用程序的一个方面配置为资源组的一部分来使集群对其进行控制。例如，快速连接服务可全部作为资源添加至集群资源组，使它们在节点或网络接口故障事件中高度可用。

注：应用程序接管通常与 IP 地址接管关联。如果重新启动应用程序的节点也获取发生故障的节点上的 IP 服务地址，那么客户机只需重新连接到相同的服务器 IP 地址。如果不接管 IP 地址，那么客户机需要连接到新的服务器以继续访问应用程序。

另外，可使用 AIX System Resource Controller (SRC) 来监视应用程序守护程序是否存在并作出相应的响应。

应用程序监视器

还可以配置应用程序监视器来检查进程故障或其他应用程序故障，并自动采取操作以重新启动应用程序。

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。由于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。或者，可配置一个定制监视器以检查数据库的运行状况，以及一个进程终止监视器以立即检测数据库进程的终止。

应用程序可用性分析工具

应用程序可用性分析工具度量任何应用程序可用的确切时长。PowerHA SystemMirror 软件对与您选择要用该工具监视的应用程序有关的大量信息进行收集、标记时间戳记和记录。使用 SMIT，您可选择一个时间段，该工具将显示该时间段内特定应用程序的正常运行时间和停机时间统计数据。

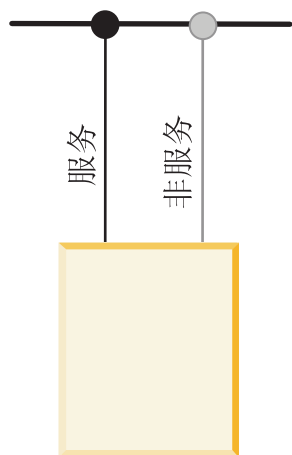
消除通信接口作为单点故障

PowerHA SystemMirror 软件可处理已配置服务 IP 标签的网络接口的故障。

这样的故障有两种类型：

- 在节点上配置的两个网络接口中，具有服务 IP 标签的网络接口发生故障，但是同一节点上的另一个备用网络接口卡仍然可用。在这种情况下，集群管理器将交换该节点上这两个接口卡的角色。这种网络接口故障对您来说是透明的，除了在系统重新配置节点上的网络接口时有一点延迟。
- 在节点上配置的两个网络接口中，另一个备用网络接口发生故障，但已配置服务 IP 标签的网络接口仍然可用。在这种情况下，集群管理器将检测备用网络接口故障、记录事件并向系统控制台发送一条消息。如果希望进行其他处理，可为该事件定制处理。

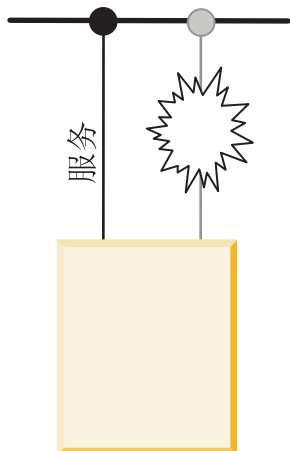
下图说明了发生在同一节点上的网络接口交换：



此时，服务接口提供了与网络的连接。非服务接口应该对应用程序隐藏，仅对于集群管理器可知。

c_adapswap1-00

图 15. 网络适配器交换之前的配置



此时，服务接口已失败，集群管理器指定前一个非服务接口作为新的服务接口。

c_adapswap2-00

图 16. 网络适配器交换之后的配置

消除网络作为单点故障

当 PowerHA SystemMirror 网络对集群中所有节点都发生故障时，将会发生网络故障。当任何集群节点都不能使用为给定 PowerHA SystemMirror 网络配置的任何网络接口卡互相访问时，就会发生这种类型的故障。

下图说明了一个网络故障：

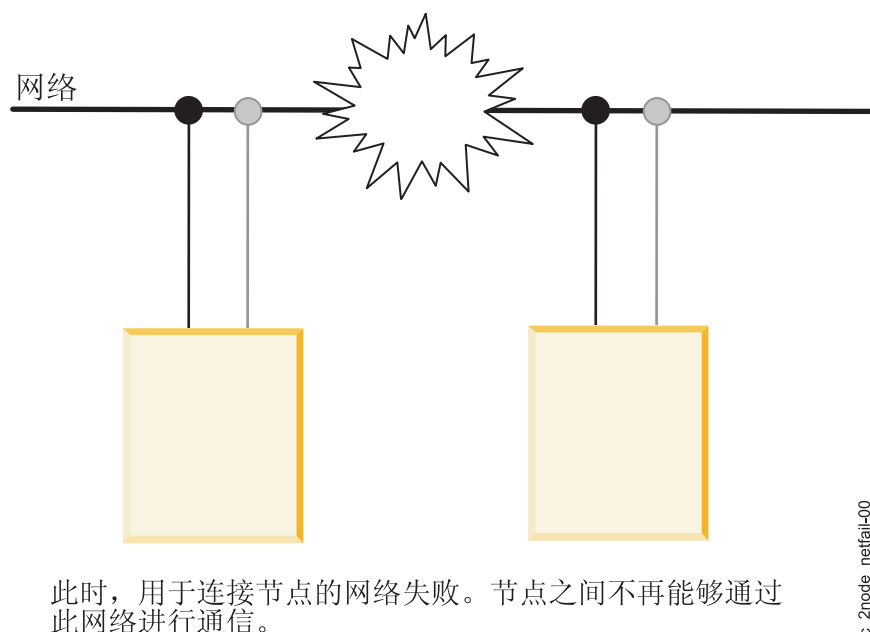


图 17. 网络故障

PowerHA SystemMirror 软件应对网络故障的第一道防线是用多个网络连接集群节点。如果一个网络发生故障，PowerHA SystemMirror 软件将使用仍然可以满足集群流量和监视节点状态的网络。

可以指定其他操作来处理网络故障，例如，通过备用网络重新路由。强烈建议至少用两个网络预防网络故障。

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 IP 标签的资源组采取选择性的恢复操作。集群管理器将尝试只移动受到本地网络故障事件影响的资源组，而不是特定节点上的所有资源组。

节点隔离和已分区集群：

当连接集群的两个或两个以上部分的所有网络失败时，就会进行节点隔离。每个节点组（一个或多个）完全与其他组隔离。分区集群是某些节点组无法与其他节点组通信的集群。

在某个已分区集群的以下图解中，节点 A 和节点 C 在分区的一边，节点 B 和节点 D 在分区的另外一边。

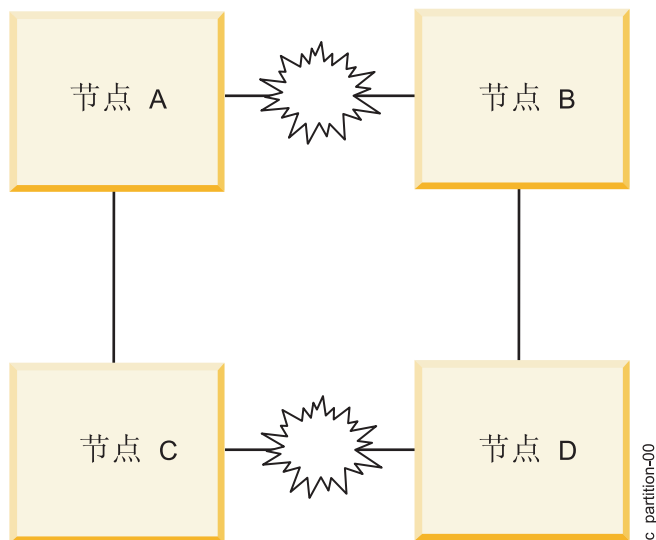


图 18. 已分区集群

已分区集群的问题是，分区一边的节点将分区另外一边的节点的脉动信号的缺失理解为那些节点已发生故障，然后会为那些节点生成节点故障事件。一旦发生这种情况，集群每一边的节点（如果这样配置）都将尝试从仍然活动并因此仍然合理地拥有那些资源的节点那里接管资源。这些接管尝试可导致在集群中发生不可预测的结果，例如，由于磁盘重置造成的数据损坏。

使用存储区域网络防止分区：

为了防止由于失去 TCP/IP 网络而导致分区集群，Cluster Aware AIX 软件将通过跨越集群节点的存储区域网络传递脉动信号。

相关信息：

规划 PowerHA SystemMirror

消除磁盘和磁盘适配器作为单点故障

PowerHA SystemMirror 软件本身不直接处理磁盘和磁盘适配器故障。相反，AIX 通过磁盘上的 LVM 镜像和具有该能力的磁盘上的内部数据冗余来处理这些故障。

例如，通过配置具有多个存储适配器的系统，通过存储区域网络传递脉动信号和 LVM 镜像，磁盘子系统在任何单个组件（适配器、布线和磁盘）发生故障时都不会造成磁盘数据无效。

如果您在使用 IBM 2105 ESS 或其他 RAID 存储子系统，那么磁盘阵列本身将负责提供数据冗余。

AIX 错误通知设施

使用 AIX 错误通知设施，可检测未被 PowerHA SystemMirror 软件特别监视的事件。例如，磁盘适配器无法为对事件的响应编写程序。

磁盘驱动器、控制器或适配器上持久的硬件错误会影响数据的故障恢复。通过错误通知方法监视这些错误使您可以评估故障对集群提供高可用性的能力的影响。错误通知的一个简单实现是向系统管理员发送邮件消息以便进一步研究问题。较复杂的实现可能包括这样的逻辑：分析故障，确定要继续处理、停止处理、还是将故障升级为节点故障并使接管节点让卷组资源对客户机可用。

强烈建议您对影响磁盘子系统的所有错误实施错误通知方法。这样做能确保退化的故障恢复不会不受检测。

AIX 错误通知方法可自动用在 PowerHA SystemMirror 中以监视某些可恢复的 LVM 错误，例如卷组损失错误。

自动错误通知

可使用 SMIT 中的特定选项为某些集群资源自动配置错误通知。如果选择该选项，那么将为特定设备自动开启所有集群节点上的错误通知。

自动错误通知支持某些不可恢复的错误类型，例如磁盘适配器错误。该功能不支持介质错误、已恢复的错误或临时错误。可为自动错误通知支持的所有错误类型指定两种错误通知方法中的一种。

另外，如果向资源组添加卷组，那么 PowerHA SystemMirror 将为它创建 AIX 错误通知方法。在卷组损失定额的情况下，PowerHA SystemMirror 使用该方法选择性地将被影响的资源组移到另一个节点。请勿编辑或更改 PowerHA SystemMirror 生成的错误通知方法。

PowerHA SystemMirror 错误通知设施不支持 MPIO 磁盘。

用 PowerHA SystemMirror 使计划内停机时间最短

PowerHA SystemMirror 软件使您能在活动集群上动态地执行大部分例程维护任务，不必停止然后重新启动集群服务以使已更改的配置成为活动配置。

无需停止应用程序而启动集群服务

可以无需停止应用程序在节点上启动 PowerHA SystemMirror 集群服务。

相关信息:

启动和停止集群服务

动态自动重新配置

当您在活动的集群上进行更改之后对集群配置进行同步时，会触发动态自动重新配置 (DARE)。使用 SMIT 应用集群快照也会触发动态重新配置事件。

例如，要将节点添加到正在运行的集群中，只需简单地将该节点连接至集群，将该节点添加到任何现有集群节点的集群拓扑中，并同步集群。新节点将添加到所有集群节点的集群拓扑定义中，且更改的配置将成为当前的活动配置。动态重新配置事件完成后，可在新节点上启动集群服务。

PowerHA SystemMirror 会在使更改的配置成为当前的活动配置之前对其进行验证，以确保您所做的更改产生了有效的配置。

动态重新配置如何工作：

当在运行中的集群上使用动态自动重新配置 (DARE)，且启动 PowerHA SystemMirror 时，DARE 将创建存储在系统缺省的对象数据模型 (ODM) 中的特定于 PowerHA SystemMirror 的对象类的专用副本。

从现在开始，ODM 称为 PowerHA SystemMirror 配置数据库。两个目录存储配置数据库数据：

- Active Configuration Directory (ACD)，专用目录，用来存储 PowerHA SystemMirror 配置数据库数据，供运行中节点上的所有 PowerHA SystemMirror 守护程序、脚本和实用程序引用。
- Default Configuration Directory (DCD)，系统缺省目录，用来存储 PowerHA SystemMirror 配置数据库和数据。

注：为了保持完整性，此处对 DARE 的操作进行描述。无需手动干预来确保 PowerHA SystemMirror 执行这些操作。PowerHA SystemMirror 可正确管理集群中所有动态重新配置操作。

DCD 是名为 **/etc/objrepos** 的目录。该目录包含缺省系统对象类，例如定制的设备数据库 (CuDv) 和预先定义的设备数据库 (PdDv)，还包含特定的 PowerHA SystemMirror 对象类。ACD 是 **/usr/es/sbin/cluster/etc/objrepos/active**。

注：当配置集群时，可修改存储于 DCD 的 PowerHA SystemMirror 配置数据库数据（不是存储于 ACD 中的数据）。SMIT 和其他 PowerHA SystemMirror 配置实用程序都可修改 DCD 中的 PowerHA SystemMirror 配置数据库数据。另外，显示 PowerHA SystemMirror 配置数据库数据的所有用户命令（例如，**clicsif** 命令）都可从 DCD 读取数据。

下图说明了所有 PowerHA SystemMirror 守护程序、脚本和实用程序如何在访问配置信息时引用 ACD。

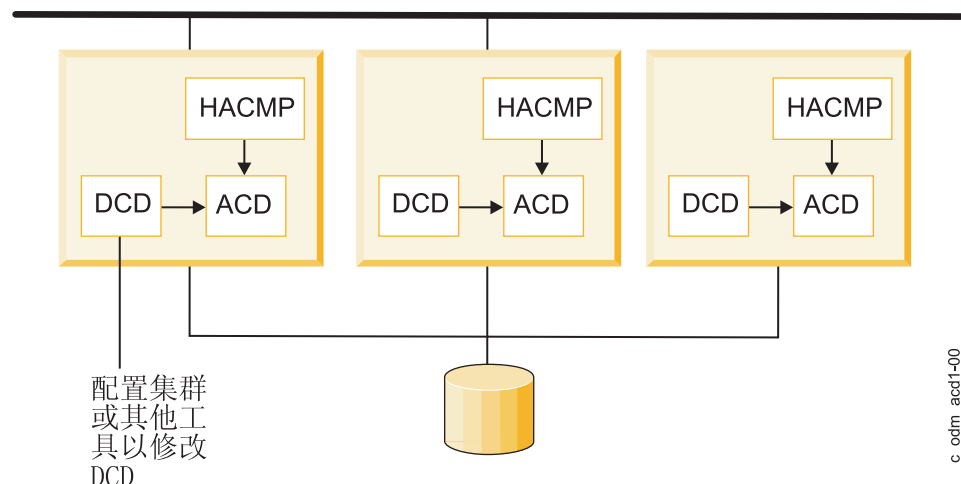


图 19. 当集群启动时，PowerHA SystemMirror 与 ACD 的关系。

动态地重新配置集群：

PowerHA SystemMirror 软件根据某 PowerHA SystemMirror 配置数据库存储库的位置来存储配置数据。

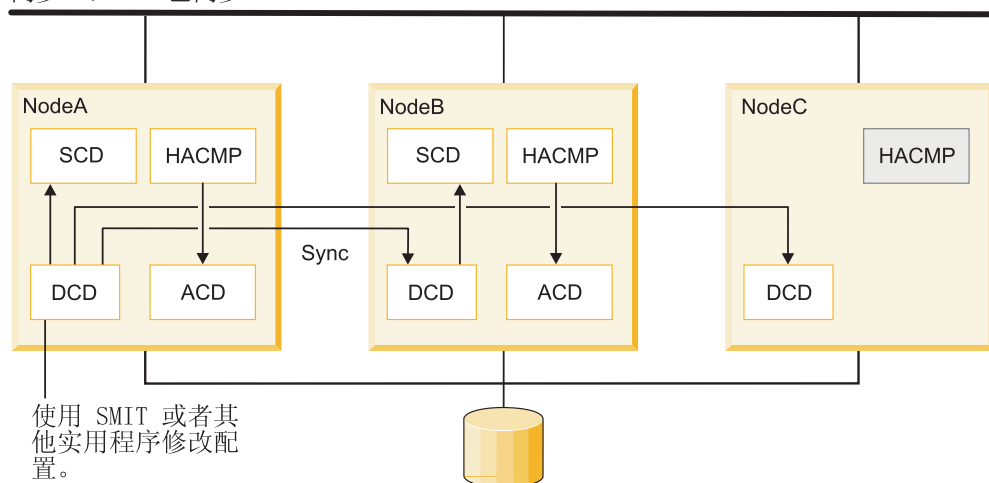
这些存储库的存在与缺失有时可用于确定在集群配置和操作期间采取的步骤。ODMPATH 环境变量允许 PowerHA SystemMirror 配置数据库命令和子例程查询非缺省位置（保留在 ODMDIR 环境变量中）的位置，如果查询的对象不存在于缺省位置中。您可以设置该变量，但一定不能设置为包含 **/etc/objrepos** 目录，否则您将失去 PowerHA SystemMirror 配置信息的完整性。

要更改活动集群的配置，可使用 SMIT 修改存储在缺省配置目录 (DCD) 中的特定 PowerHA SystemMirror 配置数据库类所存储的集群定义。更改活动集群中的集群配置时，可使用相同的 SMIT 路径进行更改，但这些更改不会立即生效。因此，可在一次操作中执行多次更改。在所有集群节点中同步配置时，将发生集群范围的动态重新配置事件。当 PowerHA SystemMirror 处理动态重新配置事件时，它将更新存储在每个集群上的 DCD 中的 PowerHA SystemMirror 配置数据库对象类，并在已协调的集群范围的转换中，用 DCD 中的新的 PowerHA SystemMirror 配置数据库数据替换存储在 ACD 中的 PowerHA SystemMirror 配置数据库数据。它还会刷新集群守护程序，以使它们引用新的配置数据。

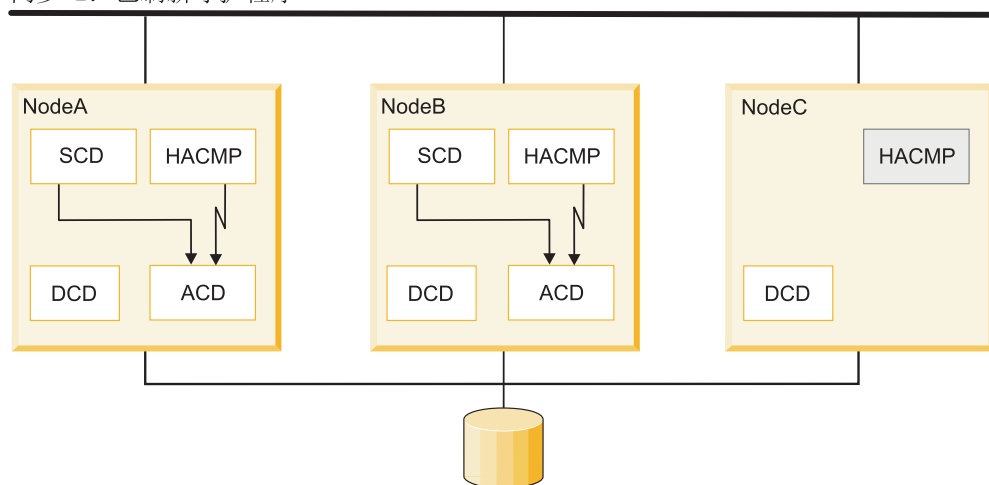
这样处理之后，集群脉动信号将短暂暂挂且集群会处于不稳定状态。已更改的配置将成为活动配置。集群服务在新添加的节点上启动之后，它将自动集成到集群中。

下图说明了涉及使用动态重新配置将节点添加到活动集群的处理。

同步 1: DCD 已同步



同步 2: 已刷新守护程序



同步 3: 完成动态重新配置；在 NodeC 上已启动集群服务

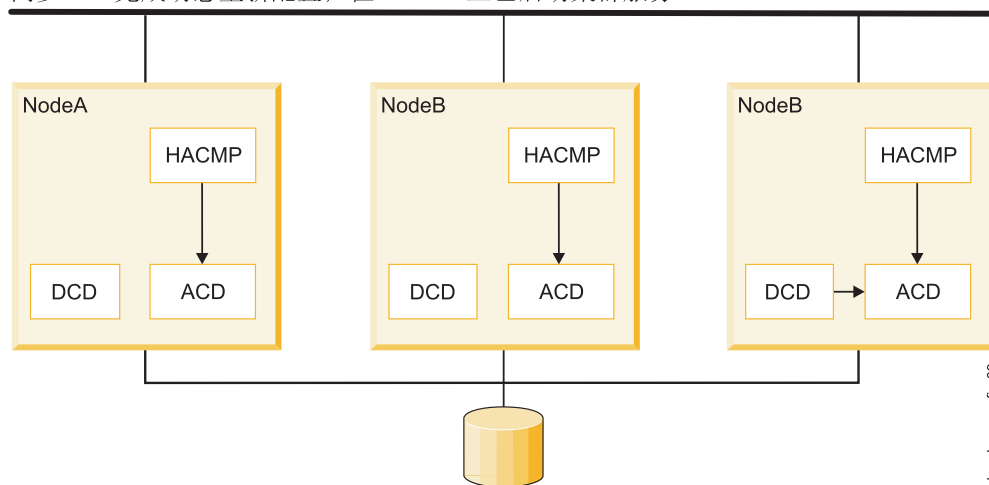


图 20. 动态重新配置处理

要添加的节点连接到运行中的集群，但集群服务在该节点上不活动。配置将在 NodeA 上重新定义。当同步对配置所作的更改时，存储于 NodeA 的 DCD 中的 PowerHA SystemMirror 配置数据库数据将复制到其他集群节点的 DCD 上，且将触发动态重新配置事件。PowerHA SystemMirror 将 DCD 中的新的 PowerHA SystemMirror 配置数据库数据复制到每个节点的临时位置中，称为“登台配置目录 (SCD)”。SCD 的位置是 /usr/es/sbin/cluster/etc/objrepos/stage。通过使用该临时位置，PowerHA SystemMirror 允许您在进行动态重新配置的同时，开始进行其他配置更改。在将 SCD 中的新的 PowerHA SystemMirror 配置数据库数据复制到 ACD 的当前 PowerHA SystemMirror 配置数据库数据之前，PowerHA SystemMirror 将验证新配置。

注：可在进行动态重新配置的同时启动另外的重新配置，但不能对它进行同步。任何集群节点上存在的 SCD 都可作为锁定，防止新的动态重新配置的启动。

资源组管理

您可以使用资源组管理 (clRGmove) 实用程序将资源组移至其他集群节点和站点，或者无需停止集群服务即可使它们联机或脱机。

这在管理资源组及其应用程序方面为您提供了灵活性。也可利用该实用程序释放任何资源组的节点以在特定集群节点上执行系统维护。PowerHA SystemMirror 资源组管理实用程序 (clRGmove) 获得了显著改进，使您更容易移动资源组以进行集群管理。此外，您更容易了解手动移动资源组的结果。例如，您可以预测组是否会保留在它们移到的节点上。PowerHA SystemMirror 遵循以下原则：在所有情况下，当您移动资源组时，它们都将保留在节点上，直到您将其再次移动。需要恢复它们时，PowerHA SystemMirror 会将它们移动。

如果您想要检查资源组当前是否位于目前可用的优先级最高的节点上，那么 PowerHA SystemMirror 将提供针对节点和站点的智能选取列表选项。例如，如果组当前位于一个节点上，并且 PowerHA SystemMirror 发现另一个节点具有更高优先级，那么包含目标节点的 SMIT 选取列表会指出哪个节点具有更高优先级。这样，您就可以总是选择将该组移到该节点。当将组移到其他节点时，将应用这些规则：

- 对于具有“永不回退”这一回退策略的资源组，在将来发生集群事件期间，移动组将不会影响该组的行为。对于采用“在任一站点上联机”站点回退策略的资源组同样如此。
- 对于回退策略不是“永不回退”或者“首选主站点”的资源组，移动某个组将使目标节点成为临时的优先级最高的节点，直到您再次将其移动。移动该节点之后，它就会成为优先级最高的临时节点。

对于具有“回退至优先级最高的节点”策略或者“首选主站点”策略的资源组，此行为有一个重要注意事项。将这样的资源组移到不具备最高优先级的节点或者移到主站点时，它所移到的节点将成为它临时的首选节点，而不会成为它的优先级最高的节点（根据配置）。这种组将留在它们被移到的节点上，直到再次移动。这些组还会回退到这些节点或站点。

相关概念:

第 83 页的『PowerHA SystemMirror 配置过程和设施』

这些主题提供了 PowerHA SystemMirror 集群配置过程以及与 PowerHA SystemMirror 软件一起提供的管理工具的概述。

相关信息:

管理集群中的资源组

规划资源组

用户请求的资源组管理对自动资源组管理:

通常，为了保持应用程序高度可用，PowerHA SystemMirror 会自动管理（并有时移动）其包含的资源组和应用程序。

例如，当需要恢复资源组时，PowerHA SystemMirror 会尝试在失败转移或回退操作期间在另一个节点上自动恢复。而移动组时，PowerHA SystemMirror 会遵循指定的资源组策略和其他设置（例如，可指示 PowerHA SystemMirror 只将组的故障通知给您，而不是在另一个节点上自动恢复发生故障的资源组）。

当您请求 PowerHA SystemMirror 执行资源组管理时，它将使用 *clRGmove* 实用程序，该实用程序通过调用 *rg_move* 事件来移动资源组。

注：当故障诊断记录文件时，区分在某些情况下由 PowerHA SystemMirror 自动触发的 *rg_move* 事件和当您请求 PowerHA SystemMirror 为您管理资源组时发生的 *rg_move* 事件很重要。要识别集群中资源组上执行的操作的原因，请查找 SMIT 中的命令输出和 *hacmp.out* 文件中的信息。

资源组管理操作：

可使用资源组管理来执行大量任务。

使用资源组管理，以：

- 将资源组从一个站点上的节点移到另一个站点上的节点。
- 将资源组从一个节点移到另外一个节点。

在运行中的集群中，可将非并发资源组从它当前所驻留的节点临时移到任何目标节点。您移动的资源组将以您已配置的方式继续持续地运行，也就是说，它们遵循为其指定的启动、失败转移和回退策略。如果您决定将资源组移到其他节点，SMIT 用户界面让您可以清楚地指定和预测该资源组的行为。

如果使用 SMIT 将资源组移到另外一个节点，它将一直停留在新的目标节点上，直到您再次手动移动。PowerHA SystemMirror 可能需要在失败转移期间移动。

- 将资源组移回其最初的最高优先级节点。

资源组可能具有回退策略，也可能不具有回退策略。如果资源组具有“回退至优先级最高的节点”的回退策略，那么当您移动它之后，该组将把“新”节点认为是它目前首选的临时位置，并回退至该节点。要更改该行为，可以总是将组移回其最初的最高优先级的节点。

类似地，如果某资源组具有“永不回退”的回退策略，那么一旦移动该资源组，它将不会回到它被移走的节点，而将一直停留在新的目标节点上，直到再次将它移到其他节点。这样，将确保组总是遵循您为其指定的“永不回退”策略。

- 使资源组在集群中的一个节点或所有节点上联机或脱机。

相关信息：

管理 PowerHA SystemMirror

集群单点控制

利用集群单点控制 (C-SPOC) 实用程序，可从单个集群节点对整个集群进行更改。

无需在每个集群节点上执行管理任务，您可以使用 SMIT 界面在单个节点上发出一次 C-SPOC 命令，然后更改将在所有集群节点中传播。

相关概念：

第 87 页的『使用 C-SPOC 的 PowerHA SystemMirror 系统管理』

为了促进集群管理，PowerHA SystemMirror 提供了一种方法，可从一个节点运行命令然后验证更改并将更改同步到所有其他节点。可使用 PowerHA SystemMirror 系统管理工具和集群单点控制 (C-SPOC)，在无需停止任务关键型作业的情况下自动添加用户、文件和硬件。

动态适配器交换

动态适配器交换功能允许您将活动网络接口卡 (NIC) 的 IP 地址与相同节点和网络上用户指定的活动的可用"备份"网络接口卡的 IP 地址进行交换。

不必停止集群服务来执行交换。

该功能可用于在不关闭节点的情况下，将 IP 地址从运行不稳定的网络接口卡移到另一个 NIC。如果要在节点上替换可热插拔的 NIC，也可以使用该功能。可物理除去并替换可热插拔的 NIC，无需关闭节点电源。从节点中拔出您想要更换的（可热插拔）NIC 时，PowerHA SystemMirror 将使该 NIC 不可作为备份。

可使用 SMIT 来配置适配器交换。服务 IP 地址从它当前的 NIC 移到用户指定的 NIC。然后，该服务 IP 地址会成为可用的"备份"地址。当新卡放到节点中时，将把该 NIC 作为可用的"备份"再次集成到集群中。然后可以将 IP 地址从备份 NIC 交换到初始 NIC。

注：该类型的动态适配器交换仅可在单个节点中执行。不能用该功能将 IP 地址与不同节点上的 IP 地址交换。要将服务 IP 地址移到另一个节点，可使用"资源组管理"实用程序移动其资源组。

自动验证和同步

自动验证和同步可以使向集群添加节点的停机时间最短。

该进程在启动集群服务前运行并检查，以确保加入集群的节点能适当地同步。该进程检查输入活动配置或不活动配置的节点。由自动验证和同步纠正的一些典型的配置不一致性示例是：

- IP 地址在 RSCT 期望的网络接口上配置
- 共享卷组未设置为自动联机
- 文件系统未设置为自动安装。

如果发现任何其他配置错误，那么集群服务将不会在节点上启动，而且详细的错误消息可以使您解决不一致性。

相关信息：

验证和同步 PowerHA SystemMirror 集群

使计划外停机时间最短

PowerHA SystemMirror 的另外一个重要目标是使响应计划外的集群组件故障的计划外停机时间最短。

PowerHA SystemMirror 软件提供以下功能来使计划外停机时间最短：

- 快速恢复以加速大型集群中的失败转移
- 允许定制资源组在指定时间回退的延迟回退计时器
- 服务 IP 标签恢复期间加速处理的通过 IP 别名的 IPAT
- 无论集群节点何时启动，处于 ERROR 状态的资源组的自动恢复。有关更多信息，请参阅下一节。

在节点启动时恢复资源组

集群管理器尝试使加入节点上当前处于 ERROR 状态的资源组变成联机（活动）状态。这将进一步增加使应用程序恢复联机的机会。当节点启动时，如果资源组在集群中的任何节点上处于 ERROR 状态，那么该节点将尝试获取该资源组。

请注意，该节点必须包括在该资源组的节点列表中。

对于非并发和并发资源组，节点启动时的资源组恢复是不同的：

- 如果起始节点未能激活处于"错误"状态的非并发资源组，那么该资源组将继续失败转移到节点列表中的另一个节点（如果有可用节点）。"失败转移"操作将一直持续到已尝试使用节点列表中的所有可用节点为止。
- 如果起始节点未能激活节点上处于"错误"状态的并发资源组，那么该并发资源组在该节点上将保持处于"错误"状态。请注意，该资源组可能仍在其他节点上保持联机。

快速恢复

PowerHA SystemMirror 快速恢复功能可加速大型集群中的失败转移。

快速恢复允许您选择文件系统一致性检查和文件系统恢复方法：

- 如果您配置文件系统以使用一致性检查和恢复方法，那么它将在每个文件系统上运行 `logredo` 而不是运行 `fsck` 来节省时间。如果后续安装失败，然后它将运行完整的 `fsck`。

如果文件系统在故障中遭受损坏，但是仍然可以安装，那么 `logredo` 可能不会成功修复损坏，在数据访问期间会产生错误。

- 另外，它可通过同时（而非依次）获取、释放和回退所有资源组和文件系统来节省时间。

如果您具有共享的嵌套文件系统，请勿将系统设置为同时运行这些命令。这些必须依次恢复。（请注意，集群验证实用程序不会报告文件系统和快速恢复的不一致性。）无论如何设置恢复方法，始终将对卷组同时运行 `varyonvg` 和 `varyoffvg` 命令。

资源组的延迟回退计时器

延迟回退计时器允许资源组在您指定的时间回退至优先级更高的节点。

已配置延迟回退计时器的资源组和当前位于非主节点的资源组，可以在重现时间（每天、每周、每月或每年）或指定日期回退至优先级更高的节点。

相关信息：

规划 PowerHA SystemMirror

使接管时间最短

在集群发生故障的情况下，由于磁盘接管机制得到改善，接管增强的并发卷组的速度比 PowerHA SystemMirror 前发行版更快。

在集群发生故障的情况下，由于磁盘接管机制得到改善，接管增强的并发卷组的速度比 PowerHA SystemMirror 前发行版更快。只要您将增强的并发方式卷组（或将现有卷组转换为增强的并发卷组）包含在非并发资源组中，PowerHA SystemMirror 就将自动检测增强的并发卷组，并确保在发生节点故障事件时启动更快速的卷组接管选项。

该功能对包含大量磁盘的卷组的失败转移特别有用。

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 LVM 信息所需的额外处理。因此，用于增强的并发卷组的 PowerHA SystemMirror 磁盘接管机制比用于非并发资源组中包括的标准卷组的磁盘接管速度更快。

最大限度地进行灾难恢复

PowerHA SystemMirror 可能是贵企业的综合灾难恢复计划的主要部分。

以下是为了可能执行的灾难恢复操作而将数据的备份副本分发至不同站点的几种可以采用的方式：

- 用于地理逻辑卷管理器 (GLVM) 的 PowerHA SystemMirror Enterprise Edition
- 用于高速镜像的 PowerHA SystemMirror Enterprise Edition (与 ESS 和 DS 系统同步的 PPRC)
- 跨站点 LVM 镜像

跨站点 LVM 镜像

您可以设置位于两个不同站点的磁盘使用诸如存储区域网络 (SAN) 进行远程 LVM 镜像。跨站点 LVM 镜像会在每个站点的磁盘子系统之间复制数据以实现灾难恢复。

SAN 是一种高速网络，它允许存储设备在光纤通道支持的距离范围内与处理器（服务器）直接建立连接。因此，位于不同站点的两个或两个以上的服务器（节点）可以通过公共 SAN 来访问相同的物理磁盘，这些磁盘之间可以相隔一定距离。可以使用 AIX 逻辑卷管理器将磁盘组合为一个卷组，并且可以将此卷组导入到位于其他站点的节点。此卷组中的逻辑卷最多可以具有三个镜像。因此，您在每个站点至少可以设置一个镜像。存储在此逻辑卷中的信息保持高可用性，并且在某些故障情况下，另一个站点中的远程镜像仍然具有最新信息，因此，可以在另一个站点中继续执行操作。在发生磁盘或节点故障之后，如果随后进行了重新集成，那么 PowerHA SystemMirror 会自动使镜像同步。

即使其中一个磁盘处于 PVREMOVED 或 PVMISSING 状态，PowerHA SystemMirror 也会处理镜像的自动同步。无法在所有情况下都自动进行同步，但是在发生磁盘或站点故障之后，如果随后进行了重新集成，那么您可以使用 C-SPOC 以手动方式使继续存在的镜像与停滞的镜像的数据同步。

集群事件

本节描述 PowerHA SystemMirror 软件如何响应集群中的更改以维持高可用性。

PowerHA SystemMirror 集群软件监视组成高度可用应用程序的所有组件，包括磁盘、网络接口、节点和应用程序本身。集群管理器使用不同方法监视不同资源：

- RSCT 子系统负责监视网络和节点。
- AIX LVM 子系统针对卷组定额损失产生错误通知。
- 集群管理器本身分派应用程序监视器。

PowerHA SystemMirror 集群环境是事件驱动的。事件是集群管理器识别和处理的集群中的状态更改。集群事件可由影响网络接口卡、网络或节点的更改触发，也可由集群重新配置过程超过其时间限制触发。当集群管理器检测到集群状态的更改时，它将执行被指定为处理事件及其子事件的脚本。

注：为了保持完整性，此处对集群事件的逻辑进行描述。无需手动干预来确保 PowerHA SystemMirror 正确执行集群事件。

以下示例显示了集群管理器识别的一些事件：

- node_up 和 node_up_complete 事件（节点加入集群）
- node_down 和 node_down_complete 事件（节点脱离集群）
- 局部或全局 network_down 事件（网络已失败）
- network_up 事件（已连接网络）
- swap_adapter 事件（网络适配器失败，一个新的网络适配器已取代它）
- 动态重新配置事件
- site_up 和 site_up_complete（站点加入集群）
- site_down 和 site_down_complete（站点脱离集群）

发生集群事件时，集群管理器将为该事件运行相应的事件脚本。处理事件脚本时，可以执行一系列子事件脚本。PowerHA SystemMirror 软件为每个事件和子事件都提供一个脚本。缺省脚本位于 `/usr/es/sbin/cluster/events` 目录中。

缺省情况下，集群管理器将为特定事件调用与 PowerHA SystemMirror 软件一起提供的相应的事件脚本。

如果需要，您可以指定其他处理信息以定制站点的事件处理。

处理集群事件

PowerHA SystemMirror 软件处理的两个主要的集群事件是失败转移和重新集成。

- 失败转移是指集群组件发生故障或节点脱离集群时，PowerHA SystemMirror 软件所执行的操作。
- 重新集成是指先前已脱离集群的组件返回集群时，集群内执行的操作。

事件脚本可控制两种类型的操作。在事件脚本处理期间，“集群感知”的应用程序将集群状态视为是不稳定的。

失败转移

当由于主节点离开集群，资源组从其主节点移到另外一个节点时，就会发生失败转移。

节点离开集群，要么是由于计划内转换（节点关闭或停止节点上的集群服务），要么是由于故障。在前一种情况中，集群管理器将控制现有节点所保留的资源的释放，以及集群中仍然活动的节点对这些资源的获取。必要时，可覆盖资源的释放和获取（例如，执行系统维护）。也可以通过集成节点来推迟资源的获取（方法是为定制资源组设置延迟回退计时器）。

当监视邻近节点的节点在定义的时间段内停止接收保持活动的流量时，节点故障就会开始。如果其他集群节点认为该故障是节点故障，发生故障的节点就将从集群中除去，其资源由配置了接管能力的活动节点接管。

如果其他组件（例如，网络接口卡）发生故障，那么集群管理器将运行事件脚本以将网络流量交换到备份网络接口卡上（如果存在）。

重新集成

当资源组移到刚加入集群的节点时，就会发生重新集成或回退。当节点加入运行中的集群时，集群将变得暂时不稳定。成员节点将协调加入过程的开始阶段，然后运行事件脚本以释放加入节点被配置接管的任何资源。然后加入节点将运行事件脚本以接管这些资源。最后，加入节点将成为集群的成员。此时，集群将再次稳定。

定制事件处理

PowerHA SystemMirror 软件具有一个可以用来调整事件处理的事件定制设施。

集群管理器识别特定序列的事件和子事件的能力实现了一个非常灵活的定制方案。定制事件处理允许您在发生故障时，提供用来处理关键资源的最有效路径。

可以为在更改/显示预先定义的 **PowerHA SystemMirror 事件** SMIT 面板上的选取列表中显示或通过 SMIT 快速路径 `clcscldev.select` 显示的事件列表定义多个事件前和事件后脚本。事件的定制可以包括事件处理前后发送给系统管理员的通知，以及事件处理前后用户定义的命令或脚本，如该列表所示：

- 发送给系统管理员的有关事件要处理的通知
- 事件前脚本或命令
- PowerHA SystemMirror for AIX 事件脚本
- 事件后脚本或命令
- 发送给系统管理员的有关事件处理完成的通知。

针对以下类型的定制使用该工具：

- 事件前处理和事件后处理
- 事件通知
- 事件恢复和重试。

注：在 PowerHA SystemMirror 中，当同步集群资源时，存储于 PowerHA SystemMirror 配置数据库的事件定制信息将在所有集群节点中进行同步。因此，事件前和事件后通知脚本以及恢复事件脚本的名称必须在所有节点上都相同，尽管这些脚本所执行的实际处理可能不同。

集群验证包括借助名为 **cluster_notify** 的新事件自动监视集群配置的功能。可使用该事件来配置 PowerHA SystemMirror 远程通知方法（数字或字母数字页面，或文本消息传递），以便在集群配置中发现错误时向外发送消息。该事件的输出还将记录在运行中集群服务的每个集群节点上的 **hacmp.out** 中。

您还可以通过事件通知脚本向手机发送电子邮件通知；但是，使用远程通知更有优势。如果您负责响应事件通知更改，那么必须手动更改每个事件通知脚本中的地址。可为每个人都定义包含所有事件和节点的远程通知方法，这样您就可以在响应者发生更改时将通知方法作为一个单元进行切换。

定义新事件

在 PowerHA SystemMirror 中，可以定义新事件以及调整现有事件。

事件前处理和事件后处理

要调整事件处理以满足环境要求，请指定在集群管理器生成特定事件前后执行的命令或用户定义的脚本。例如，对于事件前处理，您可能希望在发生某事件时向特定用户发送一条消息，通知他们作好准备。对于事件后处理，您可能希望在特定网络发生故障时禁止某特定用户组登录。

事件通知

可指定一个能够提供通知（例如，邮件）的命令或用户定义的脚本，该通知是关于事件将要发生和事件已刚发生，以及事件的成功或失败。也可以通过 SMIT 界面定义通知方法，以发出定制的远程通知方法来响应集群事件。

事件恢复和重试

可指定尝试从事件命令故障中恢复的命令。如果重试计数大于零且恢复命令成功，那么事件脚本命令将再次运行。还可以指定尝试执行恢复命令的次数。

定制事件持续时间

每当集群事件完成任务所需的时间比指定的超时周期长时，PowerHA SystemMirror 软件都会发出一条系统警告。

使用 SMIT 界面，您可以定制在 PowerHA SystemMirror 发出系统警告前集群事件完成任务所允许的时间段长度。

PowerHA SystemMirror 集群配置

本章提供了由 PowerHA SystemMirror 软件支持的多种集群配置类型的示例。

该列表并不是可使用 PowerHA SystemMirror 软件定义的可能配置的完整目录。相反，可将它们作为考虑最适合于您的环境的集群配置的起点。

备用配置

备用配置是传统的冗余硬件配置，其中一个或多个备用节点闲置，等待服务器节点离开集群。

并发资源组需要所有节点都可同时访问资源组且不可在备用配置中使用。

示例：具有仅在主节点上联机的启动策略的备用配置

该示例显示了具有以下策略的资源组：仅在主节点上联机的启动策略、失败转移至列表中的下一个优先级节点的失败转移策略和回退至列表中优先级更高的节点的回退策略。

在以下备用配置中，资源组具有这些策略：

- 启动策略：仅在主节点上联机
- 失败转移策略：失败转移至列表中的下一个优先级节点
- 回退策略：回退至列表中优先级更高的节点。

在该图中，较小的数字表示较高的优先级：

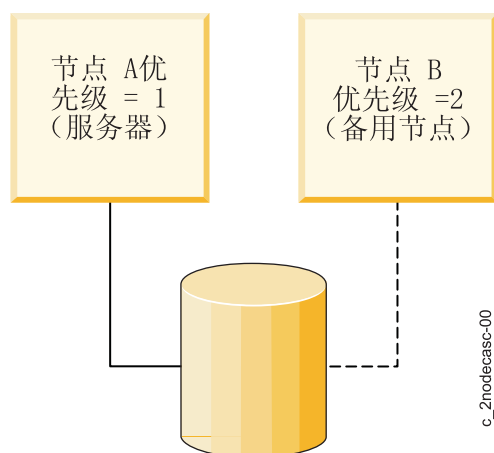


图 21. IP 标签返回至主节点的一对一备用配置

在该设置中，将集群资源定义为单个资源组的一部分。然后定义节点列表包含两个节点。第一个节点（节点 A）指定的接管（所有权）优先级是 1。第二个节点（节点 B）指定的接管优先级是 2。

集群启动时，节点 A（其优先级为 1）具有资源组的所有权。节点 A 是“服务器”节点。节点 B（其优先级为 2）处于空闲状态，在节点 A 发生故障或离开集群时做好准备。实际上，节点 B 是“备用”节点。

如果服务器节点离开集群，那么备用节点将控制由服务器拥有的资源组，启动高度可用的应用程序和服务客户机。备用节点将保持活动状态，直到具有更高接管优先级的节点重新加入集群。此时，备用节点将释放它已接管的资源组，服务器节点则将它们收回。然后备用节点将返回至空闲状态。

扩展备用配置

前面描述的示例中的备用配置可轻易地扩展至更大的集群。该配置的优点是可以更充分地使用硬件。缺点是如果多个服务器节点离开集群，那么集群会遭受严重的性能下降。

下图说明了使用具有这些策略的资源组的三节点备用配置：

- 启动策略：仅在主节点上联机
- 失败转移策略：失败转移至列表中的下一个优先级节点

- 回退策略：回退至列表中优先级更高的节点

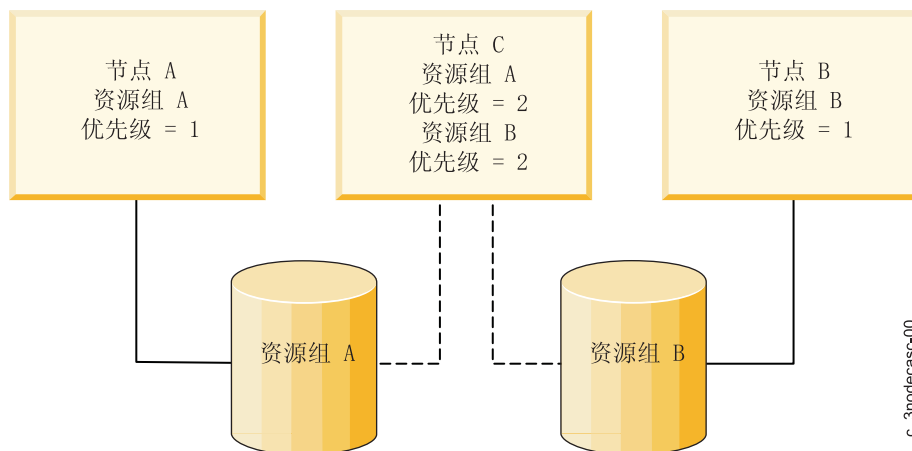


图 22. 具有三个资源组的一对二备用配置

在此配置中，存在两个独立的资源组（A 和 B）和针对每个资源组的独立的节点列表。资源组 A 的节点列表包括节点 A 和节点 C。节点 A 的接管优先级为 1，而节点 C 的接管优先级为 2。资源组 B 的节点列表包括节点 B 和节点 C。节点 B 的接管优先级为 1；而节点 C 的接管优先级为 2。（资源组只能归非并发配置中的单个节点所有。）

由于每个资源组都在其节点列表头处拥有一个不同的节点，所以集群的工作负载可在这两个资源组间进行分割或分区。但是，两个资源组都将同一节点作为它们节点列表中的备用节点。如果其中一个服务器节点离开集群，那么备用节点将控制该服务器节点的资源组并以已离开节点的身份运行。

在该示例中，备用节点有三个网络接口（未显示）以及到每个服务器节点的外部磁盘的单独物理连接。所以，如果需要的话，备用节点可以并行为两个服务器节点接管。但是，在备用节点同时以两个服务器节点的身份运行时，集群性能将很有可能会下降。

示例：具有使用分配策略联机的启动策略的备用配置

此示例说明了您必须对备用配置使用的资源组策略。

在以下备用配置中，资源组具有这些策略：

- 启动策略：使用分配策略联机（基于网络或基于节点）
- 失败转移策略：列表中的下一个优先级节点
- 回退策略：永不回退

该配置与资源组的所有权优先级不固定的备用配置不同。相反，资源组与一个可在节点间轮换的 IP 地址相关联。这使服务器和备用角色不固定，随时间更改。

下图说明了本节中描述的一对一备用配置：

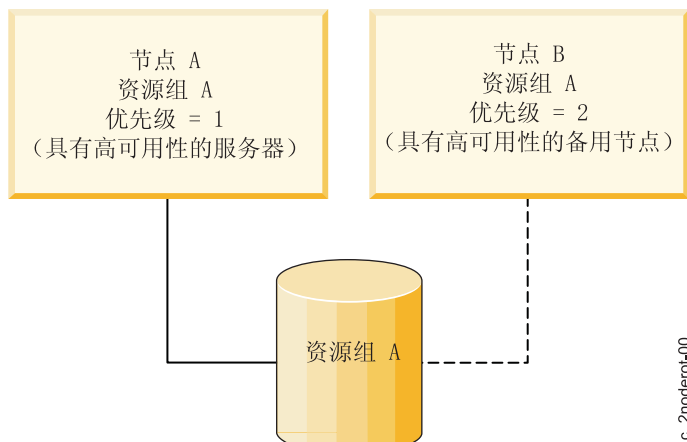


图 23. 具有 IP 标签可切换的资源组的一对一备用配置

系统启动时，资源组将连接至声明共享 IP 地址的节点。只要该节点仍然驻留在集群中，它就会拥有该资源组。如果该节点离开集群，那么对等节点将接受共享 IP 地址并声明该资源组的所有权。现在，只要对等节点仍然驻留在集群中，它就会“拥有”该资源组。

当最初声明资源组的节点重新加入集群时，它不会将资源组收回。相反，只要当前绑定到共享 IP 地址的节点在集群中是活动的，它就将保持空闲状态。仅当对等节点离开集群时，最初“拥有”资源组的节点才会再次声明资源组。因此，资源的所有权可在节点间轮换。

扩展备用配置

正如备用配置的第一个示例一样，该示例中的配置可轻易地扩展至更大的集群。例如，在示例中的这个一对二备用配置中，集群可以拥有两个独立的资源组，每个资源组都包含不同的共享 IP 地址。

当集群启动时，前两个节点都声明一个共享 IP 地址，并拥有与该共享 IP 地址关联的资源组的所有权。第三个节点则仍然处于空闲状态。如果一个活动节点离开集群，那么空闲节点将声明该共享 IP 地址并控制该资源组。

接管配置

在接管配置中，所有集群节点都将执行有用的工作，处理部分集群工作负载。没有备用节点。接管配置会比备用配置更有效地使用硬件资源，因为没有空闲处理器。但是，由于剩余节点负载增加，节点分离后会降低性能。

单边接管

该配置具有主动处理工作的两个节点，但仅一个节点向集群客户机提供高度可用的服务。也就是说，即使在集群中有两个资源集合（例如，处理客户机请求的两个服务器应用程序），也只有一个资源集合需要高度可用。

下图说明了双节点单边接管配置。在该图中，较小的数字表示较高的优先级。

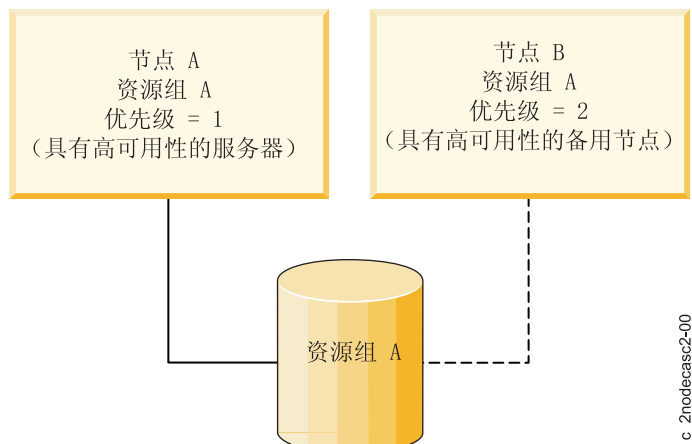


图 24. 资源组中的 IP 标签返回至主节点的单边接管配置

此资源集定义为 PowerHA SystemMirror 资源组，且具有一个包含两个节点的节点列表。第二个资源集未定义为资源组，因此不高度可用。

集群启动时，节点 A（其优先级为 1）具有资源组 A 的所有权。实际上，节点 A 拥有资源组 A。节点 B（对于资源组 A，其优先级为 2）独立处理该资源组中属于它自己的工作负载。

如果节点 A 离开集群，那么节点 B 将控制共享资源。当节点 A 重新加入集群时，节点 B 将释放共享资源。

但是，如果节点 B 离开集群，那么节点 A 将不接管节点 B 的任何资源，因为节点 B 的资源未定义为高度可用的资源组（该节点加入其链）的一部分。

当单个节点能够运行需要对集群客户机高度可用的所有关键应用程序时，该配置是合适的。

相互接管

针对非并发访问配置的相互接管具有多个节点，每个节点都为集群客户机提供唯一的高度可用服务。例如，每个节点可以运行自己的数据库实例并访问自己的磁盘。

另外，每个节点都具有接管能力。如果一个节点离开集群，那么剩下的其中一个节点将接管已离开节点所拥有的资源组。

当集群中的每个节点都在运行需要高度可用的关键应用程序以及每个处理器都能够处理多个节点的负载时，针对非并发访问配置的相互接管是合适的。

下图说明了针对非并发访问的双节点相互接管配置。在该图中，较小的数字表示较高的优先级。

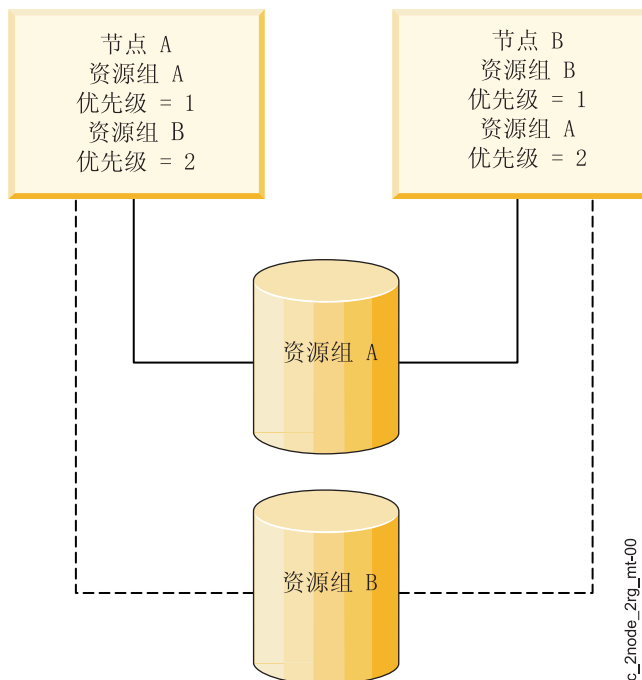


图 25. 针对非并发访问的相互接管配置

该配置的关键功能是将集群的工作负载在节点间分割或分区。存在两个资源组，以及针对每个资源组的独立资源链。加入资源链的节点相同。由这些链中的不同优先级指定该配置为相互接管。

资源组的链由节点 A 和节点 B 组成。对于资源组 A，节点 A 的接管优先级为 1，节点 B 的接管优先级为 2。对于资源组 B，将保留接管优先级。此时，节点 B 的接管优先级是 1，而节点 A 的接管优先级是 2。

集群启动时，节点 A 具有资源组 A 的所有权，而节点 B 具有资源组 B 的所有权。

如果任何一个节点离开集群，那么其对等节点将控制已离开节点的资源组。当该资源组的"所有者"节点重新加入集群，那么接管节点将放弃关联资源；它们将由较高优先级的再整合节点重新获取。

双节点相互接管配置

在该配置中，两个节点可同时访问共享磁盘并拥有相同的磁盘资源。如果某节点离开集群，将不会发生共享磁盘的接管，因为对等节点已将共享卷组联机。

下图说明了针对并发访问的双节点相互接管配置：

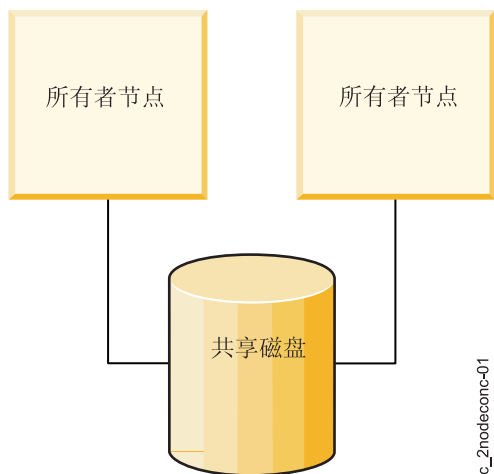


图 26. 针对并发访问的双节点相互接管配置

在该示例中，两个节点都在运行访问共享磁盘上数据库的服务器应用程序实例。该应用程序的专有锁定模型用于仲裁对磁盘资源的应用程序请求。

运行相同服务器应用程序的多个实例允许集群分配处理负载。负载增加时，可添加其他节点以进一步分摊负载。

八节点相互接管配置

在该配置中，所有节点都可同时访问共享磁盘并拥有相同的磁盘资源。

下图说明了针对并发访问的八节点相互接管配置：

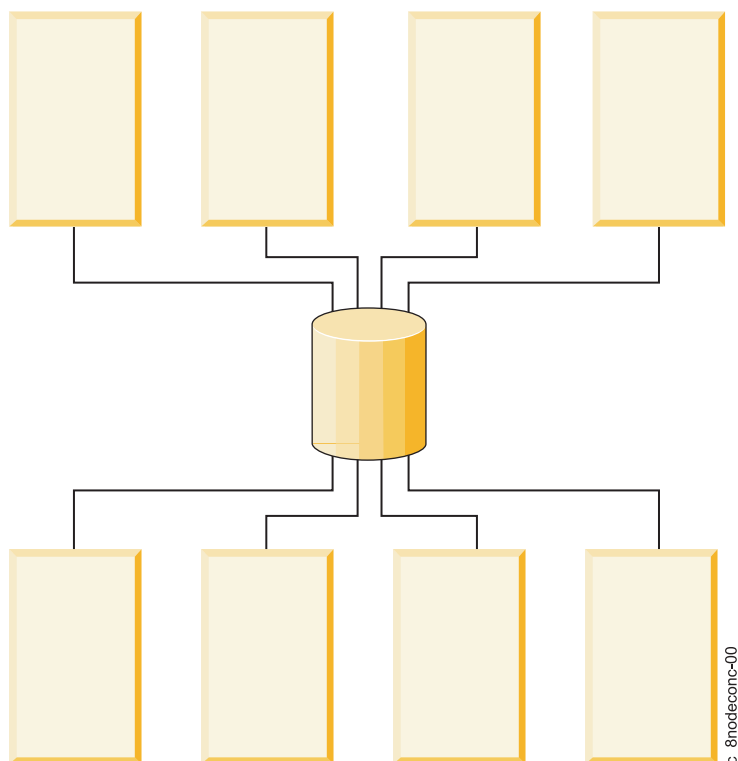


图 27. 八节点相互接管

每个节点都在运行不同的服务器应用程序。客户机在特定的 IP 地址上查询特定的应用程序。因此，必须将每个应用程序控制器及其关联的 IP 地址定义为非并发资源组的一部分，且作为该资源组的潜在所有者的所有节点都必须包含在相应的节点列表中。

在 PowerHA SystemMirror 中，并发访问资源组在拥有多达 16 个节点的集群中受支持。

具有多层应用程序的集群配置

可以利用父子依赖资源组的典型集群配置是其中一个应用程序（例如，WebSphere）依赖于另一个应用程序（例如，DB2）的环境。

为了满足业务需求，必须在两个或多个资源组间定义集群范围的父代/子代依赖关系。

下图说明了利用应用程序间依赖关系的业务方案：

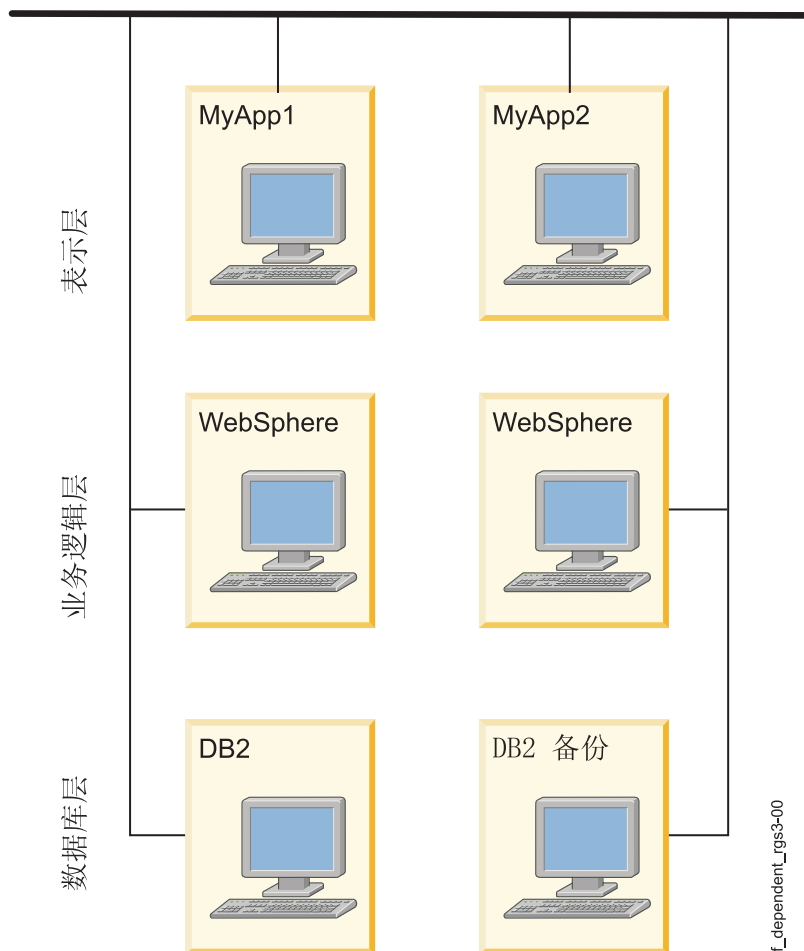


图 28. 具有应用程序之间的依赖关系的典型多层集群环境

多层应用程序

使用分层或多层应用程序的业务配置也可以利用依赖资源组。例如，后端数据库必须在应用程序控制器之前联机。在这种情况下，如果数据库脱机且被移到其他节点，那么包含应用程序控制器的资源组将必须先脱机然后又在集群的任何节点上联机。

环境（例如，SAP）要求无论数据库何时发生故障时，应用程序都能够循环（停止和重新启动）。环境（例如，SAP）提供了很多应用程序服务，且各个应用程序组件经常需要以特定的顺序进行控制。

在资源组间建立互相依赖关系证明有用的另一种情况是当系统服务需要支持应用程序环境时。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。需要更加详细时，可以通过事件前和事件后处理来控制。父/子依赖资源组允许用更简单的方法将系统服务配置为依赖它们所服务的应用程序。

相关概念:

第 39 页的『资源组依赖关系』

PowerHA SystemMirror 支持资源组排序和资源的定制串行处理，来适应在位于不同资源组的应用程序间存在依赖关系的集群配置。

具有资源组位置依赖关系的集群配置

可以对集群进行配置，使某些应用程序不管在启动时还是失败转移和回退事件期间都保留在同一节点或不同节点上。要做到这一点，可将所选资源组配置为位置依赖关系集合的一部分。

具有同一节点和不同节点依赖关系的 Publishing 模型

一家虚构公司 XYZ Publishing 遵循一个业务连续性模型，该模型涉及到分隔用于开发 Web 内容的不同平台。XYZ Publishing 使用位置依赖关系策略来严格确保一些资源组分别位于不同的节点上，而其他资源组共同位于同一节点上。

生产数据库 (PDB) 和生产应用程序 (PApp) 驻留在同一节点上以有利于维护（而且，也许这些资源组的优先级最高的节点拥有最多的内存和更快的处理器）。在它们之间设置父/子关系也是合理的，因为应用程序依赖于数据库。数据库必须联机以便应用程序运行。系统数据库 (SDB) 和系统应用程序 (Sapp)，以及 QA 数据库 (QADB) 和 QA 应用程序 (QAapp) 的情况也是一样的。

由于保持生产数据库和生产应用程序运行的优先级最高，因此配置集群以实现这样的目标才是合理的：使三个数据库资源组放置于不同节点上（使它们属于“在不同节点上联机的依赖关系”集），并为 PDB 资源组指定高优先级。SDB 具有中等优先级，而 QADB 具有低优先级。

将每个数据库及其相关应用程序组都配置为属于“在同一节点上联机的依赖关系”集。

根据配置启动、失败转移和回退策略的方式不同，PowerHA SystemMirror 处理这些组的方式也稍有不同。合理的做法是，使针对每个数据库和应用程序集的参与节点列表不同，以便于将这些资源组保留在首选节点上。

下图显示了三个节点和六个资源组的基本配置。

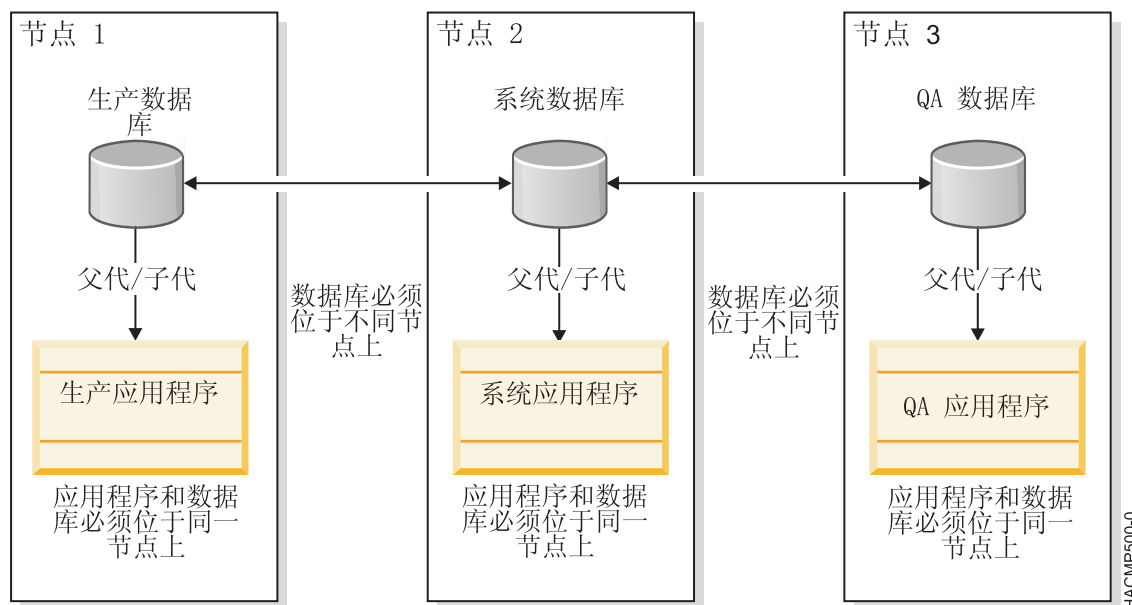


图 29. 具有父依赖关系或子依赖关系和位置依赖关系的 Publishing 模型

资源组策略

为了阐述这种情况，所有六个资源组都具有以下行为策略：

- 启动策略：在第一个可用节点上联机

- 失败转移策略：失败转移至下一个优先级节点
- 回退策略：永不回退

参与节点	位置依赖关系	父/子依赖关系
• PApp: 1、2、3 • PDB: 1、2、3 • SApp: 2、3 • SDB: 2、3 • QAAp: 3 • QADB: 3	在同一节点上联机的依赖关系组： • PApp 和 PDB • SApp 和 SDB • QAAp 和 QADB 在不同节点上联机的依赖关系集：[PDB SDB QADB] 优先级：PDB > SDB > QADB	• PApp（子）依赖于 PDB（父） • SApp（子）依赖于 SDB（父） • QAAp（子）依赖于 QADB（父）

相关信息:
 发生集群事件期间的资源组行为

用于灾难恢复的跨站点 LVM 镜像配置

您可以设置位于两个不同站点的磁盘使用存储区域网络 (SAN) 进行远程 LVM 镜像。

SAN 是一种高速网络，它允许存储设备在光纤通道支持的距离范围内与处理器（服务器）直接建立连接。因此，位于不同站点并且相隔遥远的两个或两个以上的服务器（节点）可以通过公共 SAN 来访问相同的物理磁盘，这些物理磁盘之间可能相隔遥远。这些远程磁盘可以使用 C-SPOC 来组合成卷组。

卷组中的逻辑卷最多可以具有三个镜像或副本，例如，每个站点都有一个镜像。因此，在发生故障的情况下，存储在此逻辑卷中的信息可能会保持高可用性。例如，一个站点中的所有节点（其中包括该站点中的磁盘子系统）都发生了故障，另一个站点中的远程镜像将仍然具有最新信息，并且可以在该站点中继续执行操作。

此功能部件的主要意图是为了支持由两个站点组成的集群，其中，通过 SAN 进行的 LVM 镜像会在每个站点的磁盘子系统之间复制数据以实现灾难恢复。

跨站点 LVM 镜像的另一项优势是，在发生站点/磁盘故障之后，如果随后进行了站点重新集成，那么 PowerHA SystemMirror 会自动尝试使继续存在的磁盘与连接磁盘中的数据同步。将在后台进行同步，不会对重新进行集成所耗用的时间造成重大影响。

下图说明了使用 SAN 的跨站点 LVM 镜像配置：

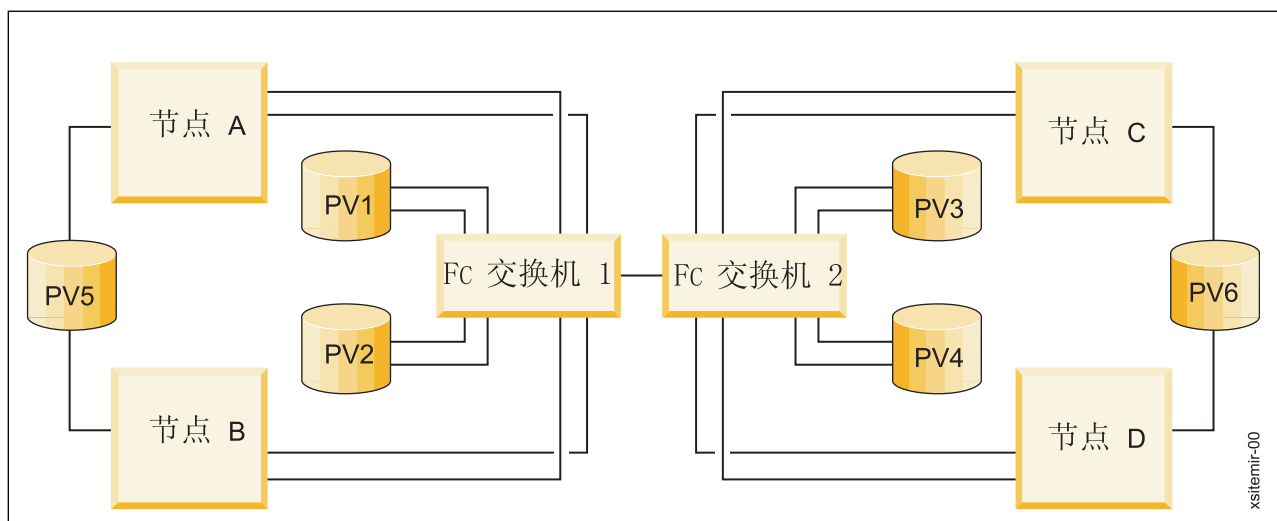


图 30. 使用 SAN 的跨站点 LVM 镜像配置

对于连接至两个站点中每个站点上的至少一个节点的磁盘，可以对其进行镜像。在此示例中，站点 1 上的节点 A 和节点 B 使用光纤通道交换机 1 和光纤通道交换机 2 来查看 PV4，还可以在节点 C 上使用光纤通道交换机 2 来查看 PV4。站点 1 上可以具有 PV4 的镜像。无法跨站点对仅连接至一个站点上的节点的磁盘（PV5 和 PV6）进行镜像。

会将磁盘信息从本地站点复制到远程站点。此数据传输的速度取决于通道的物理特征、距离以及 LVM 镜像性能。

具有动态 LPAR 的集群配置

AIX 的高级分区功能提供了动态分配系统 CPU、内存和 I/O 插槽资源的能力（动态 LPAR）。

结合使用 PowerHA SystemMirror 和 LPAR 允许您：

<Physical Adapter name>_Large_Send

用于指示此适配器的大量发送选项的标志

<Physical Adapter name>_Large_receive

用于指示此适配器的大量接收选项的标志。注意：当设置此标志并且实际适配器支持此标志时，在实际适配器接收到的数据包传递到上一层之前，将聚集这些数据包，从而获得更高性能。

<Physical Adapter name>_DMA_Errors

由于不存在资源错误而被硬件丢弃的传入数据包数。注意：通常由于适配器上的发送缓冲区已用尽而发生此错误。某些适配器可能具有由可配置参数设置的接收缓冲区大小。

<Physical Adapter name>_Speed

指示此适配器的介质速度属性

- 通过系统资源的动态分配执行例程系统升级。与动态 LPAR 一起使用时，PowerHA SystemMirror 可通过将应用程序工作负载从一个逻辑分区自动转换至另外一个逻辑分区来缩短妥善计划的系统升级所需的停机时间，从而使第一个逻辑分区可以在不对应用程序造成风险的情况下升级。
- 有效地对 CPU 和内存资源进行重新分配以管理工作负载。将 PowerHA SystemMirror 与动态 LPAR 组合在一起让您使用定制的应用程序启动和停止脚本来动态地将 CPU 和内存资源重新分配到当前正在执行应

用程序工作负载的逻辑分区，以进一步支持单个框架内的应用程序转换。这样您可以维护支持应用程序所需的处理能力和资源，而使用于升级的资源最少，成为非资源密集型任务。

注：请勿将相同物理服务器内的所有集群节点都配置为 LPAR。该配置可能是个严重的单点故障。以下示例说明了使用三个 LPAR 的集群配置：

- LPAR 1 正在运行后端数据库 (DB2 UDB)
- LPAR 2 正在运行 WebSphere Application Server (WAS)
- LPAR 3 作为 DB2 和 WAS LPAR 的备份（备用）运行。该 LPAR 仅包含最少的 CPU 和内存资源。

当需要将 DB2 或 WAS 应用程序移动到第三个 LPAR 时（例如，由于计划内升级或这些 LPAR 中发生资源故障），可以使用 PowerHA SystemMirror 中的定制应用程序启动和停止脚本，以使 CPU 和内存自动地从主 LPAR 重新动态分配至备用 LPAR。该操作允许第三个 LPAR 获取满足业务性能需求所需的 CPU 和内存资源。当 PowerHA SystemMirror 将包含应用程序的资源组移回其主 LPAR 时，CPU 和内存资源将自动跟随它移动。注：通常，动态 LPAR 允许 CPU、内存和 I/O 插槽资源的动态分配。PowerHA SystemMirror 和动态 LPAR I/O 插槽资源不兼容（尽管可以动态分配 PowerHA SystemMirror 集群以外的 I/O 插槽资源）。

下图说明了该集群环境：

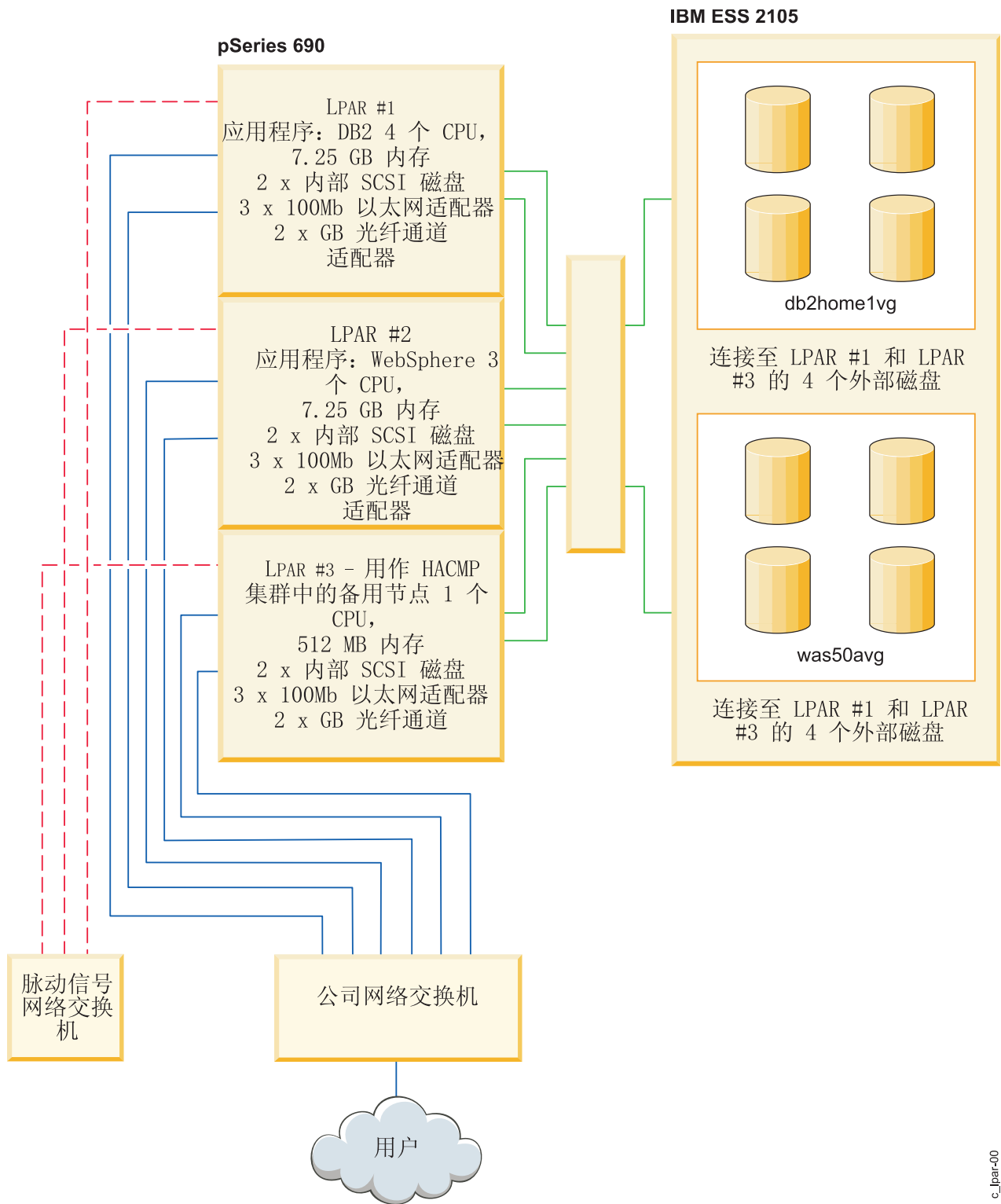


图 31. 具有三个 LPAR 的集群

DLPAR 和 Capacity on Demand

PowerHA SystemMirror 可以在 LPAR 之间移动应用程序资源，可以通过资源优化高可用性 (ROHA) 功能来执行必需的动态资源调整。ROHA 使用 Power Systems 提供的功能来动态管理以下类型的硬件资源：

- Capacity on Demand (CoD) 功能（包括 On/Off CoD 和 Enterprise Pool CoD）在框架 (CEC) 级别管理内存和 CPU 资源。
- DLPAR 功能在逻辑分区级别管理内存和 CPU 资源。

CoD 功能可激活预先安装的非活动资源（CPU 和内存）。当您的环境更改需求资源时，ROHA 使用 CoD 功能激活这些资源。例如，接管过程中，您的环境需要备用框架的额外资源。在本例中，动态提供额外资源，不必添加备用框架中的永久硬件资源。

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。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 配置过程和设施

这些主题提供了 PowerHA SystemMirror 集群配置过程以及与 PowerHA SystemMirror 软件一起提供的管理工具的概述。

您提供给 PowerHA SystemMirror 的信息

配置集群前，请确保已规划和配置构建块，以及对于 PowerHA SystemMirror 已存在到达每个节点的初始通信路径。本节涵盖了配置集群需要执行的基本任务。

关于集群物理配置的信息

集群的物理配置包括多个规划和配置任务。

这些任务包括：

- 确保 TCP/IP 网络支持集群。
- 为集群配置共享磁盘设备。
- 为集群配置共享卷组。
- 考虑使用 PowerHA SystemMirror 所针对的任务关键型应用程序。还要考虑应用程序控制器以及对每个应用程序来说什么类型的资源组管理是最适合的。
- 检查与 PowerHA SystemMirror 客户机有关的问题。
- 通过使用不能中断的多个电路或电源、冗余的物理网络接口卡、连接节点的多个网络和磁盘镜像来确保物理冗余。

相关信息：

规划 PowerHA SystemMirror

AIX 配置信息

集群组件必须在 AIX 级别正确配置。

要完成此任务，请确保：

- 到集群节点的基本通信存在。
- 卷组、逻辑卷、镜像和文件系统已配置并设置。要确保逻辑冗余，请考虑不同类型的资源组，并规划将如何对资源组中的资源进行分组。

相关信息：

安装 PowerHA SystemMirror

由 PowerHA SystemMirror 发现的信息

您只需几个步骤就可以定义基本的集群组件。为了帮助您进行集群配置，PowerHA SystemMirror 可自动从每个节点检索配置所需的信息。

注：为了使集群配置更简单更快速，还可以使用集群配置助手。有关更多信息，请参阅“双节点集群配置助手”。

为使自动发现过程运行，应该在 PowerHA SystemMirror 中满足以下条件：

- 此前已配置物理组件并执行所有必要的 AIX 配置。
- 存在到每个节点的有效通信路径。使用标准配置路径时，该信息将用于自动配置集群 TCP/IP 拓扑。

一旦完成这些任务，PowerHA SystemMirror 将自动发现集群内预先定义的物理组件，并选择缺省行为。另外，如果在配置过程中进行任何更改，PowerHA SystemMirror 将执行集群信息的发现。

运行发现可检索所有集群节点中当前的 AIX 配置信息。该信息显示在选择列表中，以帮助您对现有组件进行准确选择。

PowerHA SystemMirror 自动发现过程简单、快速，且不会为您作为集群管理员带来“等待”负担。

集群配置选项：标准和扩展

本节中，配置过程明显简化了。而配置过程的详细信息将涵盖在 *Administration Guide* 中，本节仅提供配置 PowerHA SystemMirror 集群的两种方法的简要概述。

相关信息：

管理 PowerHA SystemMirror

使用标准配置路径配置 PowerHA SystemMirror 集群

可以用几个步骤将集群的基本组件添加到 PowerHA SystemMirror 配置数据库。标准集群配置路径简化并加速配置过程，因为 PowerHA SystemMirror 会自动启动发现以收集信息并选择缺省行为。

如果使用该路径：

- 缺省情况下，将运行集群信息的自动发现。在启动 PowerHA SystemMirror 配置过程之前，需要配置网络接口或 AIX 中的设备。在 PowerHA SystemMirror 中，可以建立到其他节点的初始通信路径。完成后，PowerHA SystemMirror 将收集该信息并根据物理连接自动配置集群节点和网络。发现的所有网络都将添加到集群配置中。
- IP 别名判别是用于将 IP 标签或地址绑定到网络接口的缺省机制。
- 您可以配置大部分常用的资源类型。但是，资源组失败转移和回退行为的定制会受到限制。

使用扩展配置路径配置 PowerHA SystemMirror 集群

为了配置不常用的集群元素，或者如果未建立到每个集群节点的连接，可以手动输入信息，其方法与 PowerHA SystemMirror 软件的前发行版类似。

使用 PowerHA SystemMirror 的扩展配置 SMIT 路径时，如果有任何组件在远程节点上，那么必须手动开始对集群信息的发现。也就是说，发现是可选的（而不是像使用标准 PowerHA SystemMirror 配置 SMIT 路径时那样是自动的）。

利用扩展配置菜单下的选项，可以将集群的基本组件以及很多其他类型的资源添加到 PowerHA SystemMirror 配置数据库。可使用扩展配置路径为所有组件、策略和标准配置菜单中未包含的选项定制集群。

集群安全性

节点间的所有通信都将通过运行在每个节点上的集群通信守护程序 **clcomd** 发送。

clcomd 守护程序管理节点间的连接认证以及配置的任何消息认证或加密。PowerHA SystemMirror 集群通信守护程序使用可信 **/etc/cluster/rhosts** 文件，并解除对 **/.rhosts** 文件的依赖。该守护程序提供对消息认证和加密的支持。

安装、配置和管理工具

PowerHA SystemMirror 包含了这几节中描述的用于安装、配置和管理集群的工具。

集成特定应用程序和 PowerHA SystemMirror 的智能帮助

给定应用程序的"智能帮助"检查系统上的配置以确定 PowerHA SystemMirror 需要监视的资源（服务 IP 标签、卷组）。然后"智能帮助"配置一个或多个资源组以使应用程序及其资源高度可用。

"智能帮助"采取以下操作：

- 发现应用程序的安装，以及（必要时）当前配置的资源，如服务 IP 地址、文件系统和卷组
- 提供 SMIT 界面，以便获取或更改来用户的包含新的服务 IP 地址的配置信息
- 向 PowerHA SystemMirror 定义应用程序，并为它提供定制启动和停止脚本。
- 为应用程序提供应用程序监视器
- 配置资源组以包含：
 - 如果应用程序解决方案需要，将配置资源组时态和位置依赖关系
 - 指定需要用 PowerHA SystemMirror 文件集合功能同步的文件
 - 必要时，修改先前配置的应用程序
 - 验证配置
 - 测试应用程序的集群配置。

受支持的应用程序

PowerHA SystemMirror 为以下应用程序和配置模型提供了"智能帮助"：

- DB2
- DB2 - 热备用
- DB2 - 相互接管
- WebSphere 6.0
- WebSphere Application Server 6.0
- WebSphere Cluster Transaction Log 恢复
- Deployment Manager
- Tivoli® Directory Server
- IBM HTTP Server
- Oracle 10G

常规应用程序智能帮助

"常规应用程序智能帮助"帮助用户配置已安装而没有它们自己的"智能帮助"的应用程序。

用户提供一些基本信息，例如：

- 主节点 - 缺省情况下是本地节点
- 接管节点 - 缺省情况下是除本地节点以外的所有已配置节点
- 应用程序名称
- 应用程序启动脚本
- 应用程序停止脚本
- 服务 IP 标签

然后"常规智能帮助"用与"双节点集群配置助手"大致相同的方法完成集群配置（但该配置可拥有多于两个节点）。使用"常规应用程序智能帮助"时，用户可以修改、测试或除去应用程序。

智能帮助 API

PowerHA SystemMirror 包含智能帮助开发者指南，以使 OEM 可以开发智能帮助来将他们自己的应用程序与 PowerHA SystemMirror 集成在一起。

相关信息:

开发 PowerHA SystemMirror 的 Smart Assist 应用程序

启动、停止和重新启动集群服务

一旦安装 PowerHA SystemMirror 并配置了集群，您就可以启动集群服务。在 PowerHA SystemMirror 中，启动、停止和重新启动集群服务的选项已得到简化和提高。PowerHA SystemMirror 处理您的请求以在不中断应用程序的情况下启动和停止集群服务，使您可以完全控制。

在 PowerHA SystemMirror 中，您可以：

- 启动和重新启动集群服务。在启动集群服务或在停止后重新启动它们时，缺省情况下，PowerHA SystemMirror 将根据如何定义资源，考虑应用程序依赖关系、应用程序启动和停止脚本、动态属性和其他参数来自动激活资源。也就是说，PowerHA SystemMirror 将自动管理（并在需要时激活）资源组和其中的应用程序。

还可以启动 PowerHA SystemMirror 集群服务，并指示它不要自动为您启动任何资源组（和应用程序）。如果某应用程序已在运行，那么您不再需要在启动集群服务之前停止。

注：PowerHA SystemMirror 依赖应用程序监视器和应用程序启动脚本，来验证它是否需要为您启动应用程序，或者该应用程序是否已在运行（PowerHA SystemMirror 尽量不启动第二个应用程序实例）。PowerHA SystemMirror 依赖已配置的应用程序监视器来检测应用程序故障。必须为 PowerHA SystemMirror 配置应用程序监视器以在启动期间检测正在运行的集群，以使它不会启动重复的应用程序实例。另一种方法是运行可确保不会启动重复的应用程序控制器实例的脚本。

- 关闭集群服务。PowerHA SystemMirror 关闭期间，可以对资源组选择下列三项操作之一：
 - 使其脱机
 - 移到其他节点
 - 使资源组处于"未管理"状态。

集群管理器会记住所有节点的状态并在用户尝试重新启动这些节点时进行适当地响应。

相关信息:

管理 PowerHA SystemMirror

SMIT 界面

可以使用与 PowerHA SystemMirror 软件一起提供的 SMIT 面板来执行多个任务。

这些任务包括：

- 配置集群、节点、网络、资源和事件。
- 捕获及恢复集群配置快照。
- 读取日志文件。
- 诊断集群问题。
- 使用 C-SPOC 实用程序管理集群。
- 执行资源组管理任务。
- 配置自动错误通知。
- 执行动态适配器交换。
- 配置集群性能优化。
- 配置定制磁盘方法。

使用 C-SPOC 的 PowerHA SystemMirror 系统管理

为了促进集群管理，PowerHA SystemMirror 提供了一种方法，可从一个节点运行命令然后验证更改并将更改同步到所有其他节点。可使用 PowerHA SystemMirror 系统管理工具和集群单点控制 (C-SPOC)，在无需停止任务关键型作业的情况下自动添加用户、文件和硬件。

可以使用 C-SPOC 执行以下任务：

- 启动或停止 PowerHA SystemMirror 服务
- PowerHA SystemMirror 通信接口管理
- PowerHA SystemMirror 资源组和应用程序管理
- PowerHA SystemMirror 文件集合管理
- PowerHA SystemMirror 日志查看和管理
- PowerHA SystemMirror 安全性和用户管理
- PowerHA SystemMirror 逻辑卷管理
- PowerHA SystemMirror 并发逻辑卷管理
- PowerHA SystemMirror 物理卷管理
- 在节点上打开 SMIT 会话。

C-SPOC 实用程序简化了维护最多达 32 个节点的集群中的共享 LVM 组件。C-SPOC 命令在集群环境中提供了类似于可用于单个节点的标准 AIX 命令的功能。通过自动执行重复任务，C-SPOC 可消除错误的潜在来源并加快进程。

在没有 C-SPOC 功能的情况下，系统管理员必须分别在每个集群节点上执行管理任务。例如，要添加用户，您通常必须在每个集群节点上执行此任务。利用 C-SPOC 实用程序，在一个节点上执行的命令也会在其他集群节点上执行。因此 C-SPOC 可使管理开销降到最低并减少不一致节点状态的可能性。使用 C-SPOC，您可以在单个节点上发出一次 C-SPOC 命令，用户将被添加到所有指定的集群节点中。

C-SPOC 还可以使逻辑卷组件的管理和集群服务的控制更加有效。可使用 C-SPOC 实用程序从单个节点上启动或停止多个节点上的集群服务。

下图说明了一个双节点配置，以及从单个集群节点上启动集群服务时命令、脚本和节点的交互。注：所有 C-SPOC 命令都以前缀 `cl_` 开头。

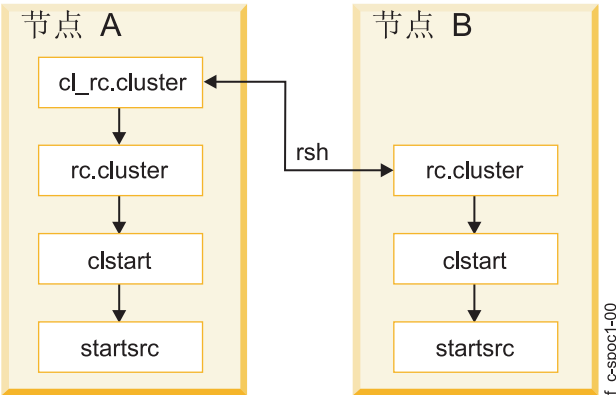


图 32. 集群启动时 C-SPOC 实用程序使用的命令流

C-SPOC 通过它自己的集群管理命令集提供该功能，可通过 SMIT 菜单和面板对其进行访问。要使用 C-SPOC，请从 PowerHA SystemMirror SMIT 菜单中选择集群系统管理选项。

集群快照实用程序

集群快照实用程序允许您保存以后想要恢复的集群配置。

还可以保存有助于诊断系统或集群配置问题的其他系统和集群信息。您可以创建自己的定制快照方法以存储有关集群的其他信息。

集群快照允许在快照中跳过保存日志文件。集群快照可用于记录集群配置信息，但是集群日志仅记录集群的操作而不是配置信息。缺省情况下，当创建集群快照时，PowerHA SystemMirror 不再收集集群日志文件，尽管仍可指定在 SMIT 中收集日志。跳过日志收集会加快快照实用程序的运行时间并减少快照的大小。

定制的事件处理

可定义多个事件前和事件后脚本以使事件处理满足站点的独特需求。

相关信息:

管理 PowerHA SystemMirror

资源组管理实用程序

资源组管理实用程序 `clRGmove` 提供了一种管理集群中的资源组的方法，增强了 PowerHA SystemMirror 的故障恢复能力。

它允许您在无需停止集群服务的情况下使任何类型的资源组（及其资源 - IP 地址、应用程序和磁盘）联机、脱机或将其移到另外一个节点。资源组管理帮助您更有效地管理集群，使您更好地利用集群硬件资源。

资源组管理还允许您在无需重新启动集群或中断操作中的节点的情况下执行选择性的维护。例如，可利用该实用程序释放任何资源组的节点以在特定集群节点上执行系统维护。

使用资源组管理实用程序不会对节点当前所拥有的其他资源组产生影响。当前节点释放它，目标节点获取它，就像节点失败转移期间一样。（如果已在资源组间配置位置依赖关系，那么 PowerHA SystemMirror 将验证并确保依赖关系被接受）。

使用资源组管理来完成下列任务：

- 将非并发资源组从工作集群中的一个节点暂时移到另一个节点（以及从一个站点移到另一个站点）。
- 使资源组在集群中的一个节点或所有节点上联机或脱机。

当移动组时，它将留在被移到的节点上，直到再次移动。如果移动一个具有“回退至优先级最高的节点”的回退策略的组，那么该组将回退或返回至它“新的”临时的优先级最高的节点（如果在后续集群事件期间 PowerHA SystemMirror 必须在其他节点上恢复它）。

如果您希望再次移动组，那么 PowerHA SystemMirror 将智能地通知您（在包含目标节点的选择列表中）它是否发现存在可以主管组的具有更高优先级的节点。您可以总是选择将该组移到该节点。

PowerHA SystemMirror 文件集合管理

像卷组一样，位于每个集群节点上的一些文件需要保持同步，以使 PowerHA SystemMirror（和其他应用程序）正常运行。这种文件包括事件脚本、应用程序脚本和一些 AIX 和 PowerHA SystemMirror 配置文件。PowerHA SystemMirror 文件集合管理提供了请求一组文件在整个集群中保持同步的简单方法。

使用 PowerHA SystemMirror 文件集合，您不必手动将更新的文件复制到每个集群节点，验证文件是否已正确复制，以及确认每个节点都拥有相同的文件版本。

而且，如果这些文件中的一个或多个在一个或多个集群节点上被意外删除或损坏，它会花时间和精力来确认问题。使用 PowerHA SystemMirror 文件集合，这种情况将得到缓解。PowerHA SystemMirror 将检测文件集合中的文件何时被删除或者文件大小是否更改为零，并记录一条消息以通知管理员。缺省情况下会安装两个预先定义的 PowerHA SystemMirror 文件集合：

- **Configuration_Files**。关键的系统文件的容器，例如 **/etc/hosts** 和 **/etc/services**。
- **PowerHA SystemMirror_Files**。PowerHA SystemMirror 配置中所有用户可配置文件的容器。这是一个特殊文件集合，底层的文件集合传播实用程序用它来引用 PowerHA SystemMirror 配置数据库 (ODM) 类中所有的用户可配置文件。

相关信息:

管理 PowerHA SystemMirror

安装 PowerHA SystemMirror

监视工具

PowerHA SystemMirror 提供了多种不同的监视工具。

此处描述的很多实用程序都使用 **clhosts** 文件来启用 PowerHA SystemMirror 集群节点间的通信。

相关概念:

第 93 页的『了解 clhosts 文件』

本主题中描述的很多监视器（包括 Clinfo、HAView 和 clstat 实用程序）都依赖于使用 clhosts 文件。clhosts 文件包含 IP 地址信息，它帮助启用 PowerHA SystemMirror 集群节点间的通信。clhosts 文件位于所有 PowerHA SystemMirror 集群服务器和客户机上。

相关信息:

管理 PowerHA SystemMirror

集群管理器

集群管理器为 SNMP 客户机提供 SNMP 信息和陷阱。

它收集与节点和接口的集群状态更改有关的集群信息。可以使用 SNMP 命令或通过基于 SNMP 的客户机程序来检索集群信息。

相关概念:

第 46 页的『集群管理器和 SNMP 监视程序』

PowerHA SystemMirror 集群是动态的，且可以随时间在它的状态中经历各种转换。

集群信息程序

集群信息程序 (Cinfo) 从 SNMP 收集集群信息，并允许与该程序通信的客户机了解集群的状态更改。

相关概念:

第 47 页的『集群信息程序』

集群信息程序 (Cinfo)，**cinfo** 守护程序，是基于 SNMP 的监视器。

应用程序监视

应用程序监视使您能够为一个应用程序控制器配置多个监视器以监视特定的应用程序和进程；并根据检测到进程的异常终止或其他应用程序故障来定义要采取的操作。

相关概念:

第 55 页的『消除应用程序作为单点故障』

创建 PowerHA SystemMirror 集群的主要原因是为任务关键型应用程序提供高度可用的环境。

clam_nfsv4 应用程序监视器

在资源组中指定了网络文件系统 V4 (NFSV4) 资源时，PowerHA SystemMirror V7.1 会自动添加 clam_nfsv4 应用程序监视器。

缺省情况下，clam_nfsv4 应用程序监视器脚本每分钟运行一次。监视器重新启动计数为 0。发生故障时，clam_nfsv4 应用程序监视器会执行一项“失败转移”操作。进行失败转移时，该监视器记录的任何错误都会导致 NFSV4 应用程序服务器进入脱机状态，并且该应用程序服务器上的资源组将移至下一个具有最高优先级的节点。

您无法除去 clam_nfsv4 应用程序监视器，但是您可以通过将发生故障时执行的操作更改为“通知”来禁用该监视器。

clam_nfsv4 应用程序监视器将确认正确安装了具有交叉安装配置的节点。该监视器还会确认 NFS 服务器可以对通过交叉安装配置发出的客户机请求作出响应。

clam_nfsv4 应用程序监视器将检查下列子系统功能是否正在运行：

- nfsrgyd
- nfsd
- rpc.mountd
- rpc.statd
- rpc.lockd

如果子系统的任何一项功能未运行，clam_nfsv4 应用程序监视器都会将 NFSV4 应用程序服务器标识为已脱机。

您必须注册其中包含资源组的 NFS 节点。您还必须配置 NFS 以正确导出资源组。

相关信息:

将 NFS 与 PowerHA SystemMirror 配合使用

PowerHA SystemMirror 中的 NFS 交叉安装

clam_nfsv4 应用程序监视器故障诊断

显示集群应用程序 SMIT 选项

"显示集群应用程序 SMIT"选项提供了以应用程序为中心的集群配置视图。

该实用程序在"应用程序关闭"类型的视图中显示现有接口和信息。

集群状态实用程序 (clstat)

集群状态实用程序 **/usr/es/sbin/cluster/clstat**，监视集群状态。该实用程序报告关键集群组件的状态：集群本身、集群中的节点、已连接到节点的网络接口和每个节点上的资源组。

它报告集群是正常运行、关机还是不稳定。它还报告节点启动、关闭、加入、离开还是重新配置，以及集群中节点的数目。**clstat** 实用程序提供 ASCII、Motif、X Window System 和 HTML 接口。可从 ASCII SMIT 运行 **clstat**。

对于作为一个整体的集群，**clstat** 指示集群状态和集群节点的数目。对于每个节点，**clstat** 显示连接到节点的每个服务网络接口的 IP 标签和地址，以及该接口开启还是关闭。**clstat** 还显示资源组状态。

可以采用 ASCII 或 X Window System 显示方式或通过 Web 浏览器查看集群状态信息。

如果在 **/usr/es/sbin/cluster/etc/clhosts** 文件中指定了 IPv6 回送地址，但未在集群中使用 IPv6 地址，那么 **clstat** 所显示的集群状态信息可能会不断更改。如果 **clstat** 的状态从 UP 到 DOWN 又从 DOWN 到 UP 反复更改，请从 **/usr/es/sbin/cluster/etc/clhosts** 文件中移除 **::1** 条目。

- | 如果要在 **/usr/es/sbin/cluster/etc/clhosts** 文件中保留 **::1** 条目，请完成以下步骤：
- | 1. 验证 **::1** 条目是否存在于 **/usr/es/sbin/cluster/etc/clhosts** 文件中。
- | 2. 通过将以下行添加到 **/etc/snmpdv3.conf** 文件，将简单网络管理协议 (SNMP) V3 配置为支持 IPv6 地址：
| COMMUNITY public public noAuthNoPriv :: 0 -
- | 3. 重新启动 **snmpd** 守护程序、**clinfoES** 守护程序以及任何从属守护程序。

注：**clstat** 实用程序使用 Clinfo API 检索集群信息。因此，请确保在客户机系统上运行 Clinfo 以查看 **clstat** 显示。

应用程序可用性分析工具

"应用程序可用性分析"工具使用定义到 PowerHA SystemMirror 的应用程序控制器，度量应用程序的正常运行时间统计数据。

PowerHA SystemMirror 软件对与您选择要用该工具监视的应用程序有关的大量信息进行收集、标记时间戳记和记录。使用 SMIT，您可选择一个时间段，该工具将显示该时间段内给定应用程序的正常运行时间和停机时间统计数据。

持久节点 IP 标签

持久节点 IP 标签是非常有用的管理工具，让您在即使节点上的 PowerHA SystemMirror 集群服务停止的情况下仍能联系该节点。

定义持久节点 IP 标签时，PowerHA SystemMirror 将尝试在节点上放置一个 IP 地址。为节点上的网络指定持久节点 IP 标签允许您在用于管理目的的集群网络上拥有节点绑定的 IP 地址，以访问集群中的特定节点。持久节点 IP 标签是可为集群网络上特定节点指定的 IP 别名，以及以下 IP 别名：

- 总是位于同一节点上（是节点绑定）
- 共存于已定义服务 IP 标签的网络接口卡
- 不需要在该节点上安装额外的物理网络接口卡

- 不是任何资源组的部分。

每个节点每个网络上可以有一个持久节点 IP 标签。

PowerHA SystemMirror 的验证和同步

PowerHA SystemMirror 的验证和同步过程将验证对 AIX 系统文件进行的特定 PowerHA SystemMirror 修改是否正确，是否正确配置了集群及其资源，是否正确配置了安全性（如果已设置），所有节点是否都遵循集群拓扑、网络配置及 PowerHA SystemMirror 资源的所有权和接管。

验证还可指示定制集群快照方法是否存在，以及他们是否在每个集群节点上都是可执行的。

无论何时您对集群已进行了配置、重新配置或更新，您之后都应该运行集群验证过程。如果验证成功，那么配置将自动同步。在活动的集群上同步立即生效。

验证实用程序运行后会在每个节点上的 PowerHA SystemMirror 配置数据库中保留信息的详细记录。每个节点的子目录都包含最后一次成功验证 (pass)、倒数第二次成功验证 (pass.prev) 和最后一次不成功验证 (fail) 的信息。

实用程序的消息输出可指示错误出现在哪里（例如，节点、设备、命令等）。

使用自动集群配置监视进行验证：

PowerHA SystemMirror 会提供自动的集群配置监视。缺省情况下，PowerHA SystemMirror 会在每 24 个小时的零时在以字母顺序排列的第一个节点上自动运行一次验证。

如果集群配置已无效，将通知集群管理员。当在选择的集群节点上完成集群验证时，该节点将通知其他集群节点。每个节点都会将有关日期、时间、哪个节点执行了验证和验证的结果等信息存储在 **/var/hacmp/log/clutils.log** 文件中。如果选择的节点不可用或无法完成集群验证，那么可通过 **var/hacmp/log/clutils.log** 文件中报告的缺失来检测到该情况。如果集群验证完成且检测到一些配置错误，那么会将以下潜在问题通知给您：

- 将验证的退出状态与集群验证过程完成的信息一起在集群间发布。
- 在集群间发送广播报文并将其显示在 *stdout* 上。这些报文会将检测到的配置错误通知给您。

使用纠正行动进行验证：

集群验证包括针对各种用户配置的 PowerHA SystemMirror 服务器组件执行的一系列检查。每个检查都尝试检测集群一致性问题或配置错误。

当对 PowerHA SystemMirror 的操作非常重要的信息，但非 PowerHA SystemMirror 软件本身的一部分，未正确传播到所有集群节点时，就会发生一些错误条件。

缺省情况下，针对标准配置和扩展配置，验证会与启用的自动纠正行动方式一起运行。这是运行验证的推荐方式。如果需要的话，可为扩展配置禁用自动纠正行动方式。但是请注意，以自动纠正行动方式运行验证使您能够自动进行很多配置任务，例如创建由很多监视器使用的基于客户机的 **clhosts** 文件。

当验证检测到以下任何条件时，您可在继续错误检查之前对纠正行动进行授权：

- PowerHA SystemMirror 共享的卷组时间戳记不是在所有节点上都匹配。
- 节点上的 **/etc/hosts** 文件没有包含所有的由 PowerHA SystemMirror 管理的 IP 标签和 IP 地址。
- 未在作为资源组一部分的节点上创建文件系统，尽管磁盘可用。
- 磁盘可用，但是未将卷组导入节点。
- 需要的 **/etc/services** 条目在节点上缺失。

- 需要的 PowerHA SystemMirror **snmpd** 条目在节点上缺失。

如果验证期间发现的错误触发任何纠正行动，那么实用程序将在完成第一轮后再次运行所有检查。如果相同的检查再次失败且初始问题是一个错误，那么将记录该错误且验证失败。如果初始条件是一条警告，那么验证成功。

相关概念:

『了解 clhosts 文件』

本主题中描述的很多监视器（包括 Clinfo、HAView 和 clstat 实用程序）都依赖于使用 clhosts 文件。clhosts 文件包含 IP 地址信息，它帮助启用 PowerHA SystemMirror 集群节点间的通信。clhosts 文件位于所有 PowerHA SystemMirror 集群服务器和客户机上。

定制验证方法:

通过 SMIT，还可以添加、更改或除去对集群配置执行特定检查的定制的验证方法。可从命令行或通过 SMIT 界面执行验证，以发出定制的远程通知方法来响应集群事件。

了解 clhosts 文件:

本主题中描述的很多监视器（包括 Clinfo、HAView 和 clstat 实用程序）都依赖于使用 clhosts 文件。clhosts 文件包含 IP 地址信息，它帮助启用 PowerHA SystemMirror 集群节点间的通信。clhosts 文件位于所有 PowerHA SystemMirror 集群服务器和客户机上。

clhosts 文件存在一些差别，视该文件所在的位置而定，下表对此进行了总结：

表 5. clhosts 文件

clhosts 文件	描述
基于服务器的文件	此文件位于所有 PowerHA SystemMirror 服务器节点上的 /usr/es/sbin/cluster/etc/ 目录中。安装 PowerHA SystemMirror 期间，缺省 IP V4 (IPv4) 回送地址 127.0.0.1 将自动添加至该文件。如果在安装时配置了任何 IP V6 (IPv6) 地址，那么还会添加缺省 IPv6 回送地址 ::1。不需要此 IP 地址通常定义的名称 loopback 和别名 localhost。
基于客户机的文件	此文件位于所有 PowerHA SystemMirror 客户机节点上的 /usr/es/sbin/cluster/etc/ 目录中。当您运行验证并且启用了自动更正功能时，将在服务器上自动生成此文件并且名为 clhosts.client 。必须将此文件手动复制到客户机节点上，并将它重命名为 clhosts。此文件包含所有已知的 IP 地址，并且必须从不包含 127.0.0.1、::1、回送地址或本地主机地址。

当监视器守护程序启动时，它将读取本地 /usr/es/sbin/cluster/etc/clhosts 文件以确定哪些节点可用于通信，如下所述：

- 对于在 PowerHA SystemMirror 服务器节点上运行的守护程序，基于服务器的本地 clhosts 文件只需要回送地址（127.0.0.1 和 ::1），当安装 PowerHA SystemMirror 的服务器部分时，该地址将被自动添加到基于服务器的 clhosts 文件中。
- 对于在 PowerHA SystemMirror 客户机节点上运行的守护程序，基于客户机的本地 clhosts 文件应该包含 PowerHA SystemMirror 服务器节点的一个 IP 地址列表。这样，如果特定 PowerHA SystemMirror 服务器节点不可用（例如，断电），那么客户机节点上的守护程序仍可以与其他 PowerHA SystemMirror 服务器节点进行通信。

PowerHA SystemMirror 验证实用程序以如下方式辅助填充基于客户机的 clhosts 文件：查找所有可用的 PowerHA SystemMirror 服务器节点、在服务器节点上创建 /usr/es/sbin/cluster/etc/clhosts.client 文件并用那些 PowerHA SystemMirror 服务器节点的 IP 地址填充该文件。

在集群上完成对 PowerHA SystemMirror 的验证和同步后，必须手动将该 **clhosts.client** 文件复制到每个客户机节点作为 /usr/es/sbin/cluster/etc/clhosts（通过除去 .client 扩展名对它进行重命名）。

相关概念:

第 92 页的『PowerHA SystemMirror 的验证和同步』

PowerHA SystemMirror 的验证和同步过程将验证对 AIX 系统文件进行的特定 PowerHA SystemMirror 修改是否正确，是否正确配置了集群及其资源，是否正确配置了安全性（如果已设置），所有节点是否都遵循集群拓扑、网络配置及 PowerHA SystemMirror 资源的所有权和接管。

故障诊断工具

通常，运行中的 PowerHA SystemMirror 集群需要最少的干预。但是如果发生问题，诊断和恢复技能是必需的。因此，故障诊断需要您能够快速识别问题，并应用您对 PowerHA SystemMirror 软件的理解以将集群恢复到正常运行。

相关概念:

第 96 页的『集群测试工具』

集群测试工具是一个实用程序，它让您测试 PowerHA SystemMirror 集群配置，以评估在一组指定情况下（例如，当节点不可访问、网络不可访问及资源组从一个节点移到另一个节点等情况时）集群如何运行。

日志文件

PowerHA SystemMirror 软件编写它生成到系统控制台和多个日志文件的消息。因为每个日志文件都包含不同级别的详细信息，所以系统管理员可以通过查看不同的日志文件来关注 PowerHA SystemMirror 处理的不同方面。

主要的日志文件包括：

- **/var/hacmp/adm/cluster.log** 文件跟踪集群事件。
- **/var/hacmp/log/hacmp.out** 文件记录配置脚本执行时生成的输出。事件摘要显示在由集群管理器启动的事件的详细输出之后，使扫描 **hacmp.out** 文件以获取重要信息更加容易。另外，事件摘要提供了到 **hacmp.out** 文件中相应事件的 HTML 链接。
- **/var/hacmp/adm/history/cluster.mmddyyyy** 日志文件记录每天的集群历史记录。
- **var/hacmp/clverify/clverify.log** 文件包含验证期间的详细消息输出。集群验证包括针对各种 PowerHA SystemMirror 配置执行的一系列检查。每个检查都尝试检测集群一致性问题或错误。验证实用程序的消息输出可指示错误出现在哪里（例如，节点、设备、命令等）。

PowerHA SystemMirror 允许您查看、重定向、保存并更改日志文件的参数，以使您可以调整它们来满足特定需求。

相关信息:

管理 PowerHA SystemMirror

重置 PowerHA SystemMirror 可调值

配置和测试集群时，可为影响集群性能的其中一个 PowerHA SystemMirror 可调值更改值。或者，您可能希望无需更改任何其他方面的配置而将可调值重置为它们的缺省设置。

可能会要求第三方集群管理员或顾问接管他们没配置过的集群的管理工作，而且可能需要将可调值重置为它们的缺省值。可使用 SMIT 界面来重置集群可调值。PowerHA SystemMirror 将在重置前创建集群快照。在将值重置为缺省值之后，如果您希望返回到定制的集群设置，可以应用集群快照。重置集群可调值将重置集群配置数据库中的信息。重置或除去的信息包括以下几类：

- 由用户提供的信息（例如，事件前和事件后脚本以及像网络掩码这样的网络参数）。注：重置集群可调值不会除去您已配置的事件前和事件后脚本。但是，如果重置可调值，PowerHA SystemMirror 对事件前和事件后脚本的知识将从配置中除去，而且 PowerHA SystemMirror 将不再使用这些脚本管理集群中的资源。如果需要，可重新配置 PowerHA SystemMirror 以再次使用这些脚本。

- 配置和同步期间由 PowerHA SystemMirror 自动生成的信息。这包括节点和网络标识，以及从操作系统中发现的信息，例如网络掩码。通常用户看不到生成的信息。

相关信息:

管理 PowerHA SystemMirror

安装 PowerHA SystemMirror

集群状态信息文件

当使用 PowerHA SystemMirror 集群快照实用程序保存集群配置记录时（正如从每个集群节点看到的），可选择性地使实用程序运行很多标准 AIX 命令和 PowerHA SystemMirror 命令以获取有关集群的状态信息。该信息存储于快照目录中的文件中，该文件可通过 **.info** 扩展名识别。

快照目录由 SNAPSHOTPATH 环境变量的值定义。缺省情况下，集群快照实用程序包含命令（例如，**cllssif**、**cllsnw**、**df**、**ls** 和 **netstat**）的输出。可创建定制快照方法以指定要存储在 **.info** 文件中的其他信息。

集群快照允许在快照中跳过保存日志文件。集群快照可用于记录集群配置信息，但是集群日志仅记录集群的操作而不是配置信息。缺省情况下，当创建集群快照时，PowerHA SystemMirror 不再收集集群日志文件，尽管仍可指定在 SMIT 中收集日志。跳过日志收集会减少快照的大小并加速运行快照实用程序。集群快照的大小取决于配置。例如，基本的双节点配置大约需要 40KB。

自动错误通知

可使用 AIX 错误通知设施来检测未被 PowerHA SystemMirror 软件特别监视的事件。例如，出现磁盘适配器故障，您可以指定事件发生时要产生的响应。

通常，可一个个手动定义错误通知方法。PowerHA SystemMirror 为重要错误提供了一组预先指定的通知方法，您可通过 SMIT 界面在一个步骤中自动"开启"这些方法，无需手动定义每个通知方法从而节省大量时间和精力。

定制远程通知

可通过 SMIT 界面定义通知方法，以发出定制的通知方法来响应集群事件。还可以向包括手机在内的任何地址发送文本消息传递通知，或向电子邮件地址发送邮件。

配置远程通知方法后，可发送一条文本消息以确认配置是否正确。

可以针对不同事件来配置任意数量的具有不同的文本消息和要拨打的电话号码的通知方法。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。

事件前同步信号和摘要

集群事件的详细信息记录在 **hacmp.out** 文件中。该文件的详细输出包含很多行事件信息；可在每个事件的详细信息末尾看到简明摘要。要对集群中最近发生的事件进行快速有效的检查，可通过使用 SMIT 中的显示事件摘要面板来查看当前和以前 **hacmp.out** 文件中仅事件摘要部分的汇编。

还可以选择将编译好的事件摘要保存到您选择的文件中。（可选）事件摘要提供了到 **hacmp.out** 文件中相应事件的 HTML 链接。集群管理器还打印出前同步信号，告知您每个事件中哪些资源组排队等待处理；可以查看将要遵循的处理顺序。

相关信息:

PowerHA SystemMirror 故障诊断

集群测试工具

集群测试工具是一个实用程序，它让您测试 PowerHA SystemMirror 集群配置，以评估在一组指定情况下（例如，当节点不可访问、网络不可访问及资源组从一个节点移到另一个节点等情况时）集群如何运行。

您可以启动测试，让它在无人管理的状态下运行，并在之后返回以评估测试结果。

如果要为拓扑和资源组管理运行基本集群测试的自动套件，那么可从 SMIT 运行自动测试套件。如果您是有经验的 PowerHA SystemMirror 管理员并要调整集群测试以满足环境要求，那么您还可以创建可从 SMIT 运行的定制测试。

建议在以下时段运行该工具：最初配置 PowerHA SystemMirror 之后和将集群放到生产环境之前；集群无法提供服务时进行集群配置更改之后；或者在即使集群看起来正常运行时定期运行。

声明

本信息是为在美国国内供应的产品和服务而编写的。

IBM 可能在其他国家或地区不提供本文档中讨论的产品、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
US*

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*

INTERNATIONAL BUSINESS MACHINES CORPORATION"按现状"提供本出版物，不附有任何种类的（无论是明示的还是默示的）保证，包括但不限于默示的有关非侵权、适销和适用于某种特定用途的保证。某些管辖区域在某些事务中不允许免除明示或默示的保证，因此本声明可能不适用于您。

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。

本信息中对非 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 站点的保证。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不必对您承担任何责任。

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 Director of Licensing
IBM Corporation
North Castle Drive, MD-NC119
Armonk, NY 10504-1785
US*

只要遵守适当的条件和条款，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。

本文档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。

所引用性能数据和客户示例仅供参考。实际性能结果可能会因具体配置和操作条件而有所不同。

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。

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。

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。经销商的价格可与此不同。

本信息仅用于规划目的。在所描述的产品上市之前，此处的信息会有更改。

本信息包括日常业务运作中使用的数据和报告的示例。为了尽可能完整地说明这些示例，示例中可能会包括个人、公司、品牌和产品的名称。所有这些名称均是虚构的，如与实际的人名或商业企业的名称有任何相似之处，纯属巧合。

版权许可：

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，而无需向 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。

这些样本程序的每一个副本或其任何部分或任何衍生产品都必须包括如下的版权声明：

©（贵公司的名称）（年）。

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.（请在此处输入年份）。

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)。

UNIX 是 The Open Group 在美国和其他国家或地区的注册商标。

索引

[A]

安全 85

[B]

并发资源管理器 50

[C]

拆分配置 15

持久节点 IP 标签 91

重置

 可调值 94

磁带 35

磁盘

 消除作为故障 59

磁盘访问

 并发 48

 非并发 48

磁盘设备 10

磁盘适配器

 消除作为故障 59

存储库磁盘

 多路径 12

 故障 13

[D]

定制远程通知 95

动态自动重新配置 60

多点广播 30

 路由 31

 网络交换机 30

 信息包通信 30

 因特网组管理协议 (IGMP) 31

[F]

服务 IP 标签 27, 34

服务 IP 地址 27, 34

[G]

概述

 应用程序可用性 50

故障诊断

 重置可调值 94

 定制远程通知 95

 集群测试工具 96

故障诊断 (续)

 集群状态信息文件 95

 日志文件 94

 事件前同步信号 95

 事件摘要 95

 自动错误通知 95

[H]

合并策略 16

[J]

集群

 安全 85

 测试工具 96

 磁盘设备 10

 多站点解决方案 21

 节点 10, 14

 客户机 13

 快照 88

 链接 22

 软件 42

 事件 67

 网络 10, 24

 物理组件 8

 延伸 23

 硬件 42

 站点 15

 资源组 35

 C-SPOC 64

 IP 地址接管 27

集群服务

 重新启动 86

 开始 86

 停止 86

集群管理器 44

 SNMP 46

集群信息程序 47

集群状态信息文件 95

监视

 持久节点 IP 标签 91

 集群管理器 89

 集群信息程序 90

 集群状态实用程序 91

 应用程序 90

 应用程序可用性分析工具 91

 clhosts 文件 93

节点 10, 14

 消除作为故障 52

卷组 33

[K]

可调值
 重置 94
客户机 13
跨站点 LVM 镜像 66
快速磁盘接管 66
快照 88

[L]

逻辑卷 33
逻辑网络 25

[M]

脉动信号传递 28
 点到点网络 30
 TCP/IP 网络 29

[P]

配置
 备用 70
 示例 70, 71
 标准配置路径 84
 动态 LPAR 80
 多层应用程序 76
 接管 72
 八节点相互 75
 单边 72
 双节点相互 74
 相互 73
 跨站点 LVM 镜像 79
 扩展配置路径 84
 选项 84
 智能帮助 85
 资源组位置依赖关系 78

[R]

软件 42, 43
 补充 50
 组件 45

[S]

示例
 备用配置 70, 71

[T]

通信接口 26
 消除作为故障 56

通信设备 26

[W]

网络 10, 24
 服务 IP 标签 27
 服务 IP 地址 27
 逻辑 25
 脉动信号传递 28
 通信接口 26
 通信设备 26
 物理 25
 消除作为故障 58
 子网 26
 IP 别名 27
网络交换机
 多点广播 30
文件集合 89
文件系统 33
物理网络 25

[X]

消除
 单点故障 51
 磁盘 59
 磁盘适配器 59
 节点 52
 通信接口 56
 网络 58
 应用程序 55

[Y]

验证
 定制方法 93
 使用纠正行动 92
 使用自动集群配置监视 92
应用程序 34
 监视 90
 消除作为故障 55
应用程序监视器 55
 clam_nfsv4 90
应用程序可用性
 概述 50
应用程序可用性分析工具 55, 91
硬件 42

[Z]

灾难恢复 66
站点 15, 42
 解决方案 21
智能帮助 85

资源

- 磁带 35
- 服务 IP 标签 34
- 服务 IP 地址 34
- 卷组 33
- 逻辑卷 33
- 文件系统 33
- 应用程序 34

资源组 35

- 策略 37
- 动态节点优先级 38
- 管理 63
- 管理实用程序 88
- 回退 36, 37
- 回退计时器回退 38
- 启动 36, 37
- 失败转移 36, 37
- 停留时间 38
- 位置依赖关系 40
- 站点 42

子网 26

自动错误通知 95

最大限度

- 灾难恢复 66

最短

- 计划内停机时间 60
- 计划外停机时间 65
- 接管时间 66

C

- clam_nfsv4 90
- clhosts 93
- Clinfo 90
- clinfo 47
- clRGmove 63, 88
- clstat 91
- C-SPOC 64, 87

D

DARE

- 请参阅 动态自动重新配置

H

- hacmp.out 95

I

- IP 别名 27
- IP 地址接管 27

L

- log 94

N

- NFS 48

P

- PowerHA SystemMirror
 - 多点广播 30

S

- SMIT 87
- SNMP 46, 89

[特别字符]

- /.rhosts 85



Printed in China