vSphere 可用性

ESXi 5.0

vCenter Server 5.0

在本文档被更新的版本替代之前，本文档支持列出的每个产品的版本和所有后续版本。要查看本文档的更新版本，请访问 http://www.vmware.com/cn/support/pubs。

ZH_CN-000605-00

vmware®

最新的技术文档可以从 VMware 网站下载：

http://www.vmware.com/cn/support/pubs/

VMware 网站还提供最近的产品更新信息。

您如果对本文档有任何意见或建议，请把反馈信息提交至：

docfeedback@vmware.com

版权所有 © 2009 – 2011 VMware, Inc. 保留所有权利。本产品受美国和国际版权及知识产权法的保护。VMware 产品受一项或多项专利保护，有关专利详情，请访问 http://www.vmware.com/go/patents-cn。

VMware 是 VMware, Inc. 在美国和/或其他法律辖区的注册商标或商标。此处提到的所有其他商标和名称分别是其各自公司的商标。

VMware, Inc.
3401 Hillview Ave.
Palo Alto, CA 94304
www.vmware.com

北京办公室
北京市海淀区科学院南路 2 号
融科资讯中心 C 座南 8 层
www.vmware.com/cn

上海办公室
上海市浦东新区浦东南路 999 号
新梅联合广场 23 楼
www.vmware.com/cn

广州办公室
广州市天河北路 233 号
中信广场 7401 室
www.vmware.com/cn

目录

关于 vSphere 可用性 5

1 业务连续性和最小化停机时间 7
- 减少计划的停机时间 7
- 防止非计划停机时间 8
- vSphere HA 提供快速中断恢复 8
- vSphere Fault Tolerance 提供连续可用性 9

2 创建和使用 vSphere HA 群集 11
- vSphere HA 的工作方式 11
- vSphere HA 接入控制 15
- vSphere HA 对照表 20
- 创建 vSphere HA 群集 21
- 自定义 vSphere HA 行为 25
- vSphere HA 群集的最佳做法 27

3 为虚拟机提供 Fault Tolerance 31
- Fault Tolerance 的工作方式 32
- 将 Fault Tolerance 功能与 DRS 配合使用 33
- Fault Tolerance 用例 33
- Fault Tolerance 对照表 33
- Fault Tolerance 互操作性 35
- 为 Fault Tolerance 准备群集和主机 36
- 为虚拟机提供 Fault Tolerance 39
- 查看有关容错虚拟机的信息 40
- Fault Tolerance 的最佳做法 41
- vSphere Fault Tolerance 配置建议 43

索引 45

关于 vSphere 可用性

《vSphere 可用性》介绍提供专业连续性的解决方案，包括如何建立 vSphere® High Availability (HA) 和 vSphere Fault Tolerance。

目标读者

此信息专供需要通过 vSphere HA 和 Fault Tolerance 解决方案提供业务连续性的用户使用。本书的目标读者为熟悉虚拟机技术和数据中心操作且具有丰富经验的 Windows 或 Linux 系统管理员。

1 业务连续性和最小化停机时间

无论是计划停机时间还是非计划停机时间，都会带来相当大的成本。但是，用于确保更高级别可用性的传统解决方案都需要较大开销，并且难以实施和管理。

VMware 软件可为重要应用程序提供更高级别的可用性，并且操作更简单，成本更低。使用 vSphere，组织可以轻松提高为所有应用程序提供的基准级别，并且以更低成本和更简单的操作来实现更高级别的可用性。使用 vSphere，您可以：

- 独立于硬件、操作系统和应用程序提供更高可用性。
- 减少常见维护操作的计划停机时间。
- 在出现故障时提供自动恢复。

vSphere 可以减少计划的停机时间，防止出现非计划停机，并迅速从断电中恢复。

本章讨论了以下主题：

- 第 7 页，“减少计划的停机时间”
- 第 8 页，“防止非计划停机时间”
- 第 8 页，“vSphere HA 提供快速中断恢复”
- 第 9 页，“vSphere Fault Tolerance 提供连续可用性”

减少计划的停机时间

计划的停机时间通常占数据中心停机时间的 80% 以上。硬件维护、服务器迁移和固件更新均需要将物理服务器停机。为最小化此停机时间的影响，会强制组织延迟维护，直到出现不便且难以调度的停机时间段。

通过 vSphere，组织可以显著减少计划的停机时间。由于 vSphere 环境中的工作负载无需停机或服务中断就可以动态移动到其他物理服务器，所以服务器维护无需应用程序和服务停机就可以执行。通过 vSphere，组织可以执行以下任务：

- 消除常见维护操作的停机时间。
- 消除计划的维护时间段。
- 随时执行维护，无需中断用户和服务。

由于 VMware 环境中的工作负载无需中断服务即可动态移动到不同的物理服务器或基础存储器，所以，通过 vSphere 中的 vSphere vMotion® 和 Storage vMotion 功能，组织可以减少计划的停机时间。管理员可以快速而完整地执行透明的维护操作，无需强制调度不方便的维护时间段。

防止非计划停机时间

在 ESXi 主机为应用程序的运行提供稳定平台时，组织还必须保护自身，避免出现硬件或应用程序故障所导致的非计划停机时间。vSphere 将重要功能构建到数据中心基础结构中，这有助于避免出现非计划停机时间。

这些 vSphere 功能是虚拟基础结构的一部分，因此，对操作系统以及虚拟机中运行的应用程序而言是透明的。这些功能可以进行配置，而且可供物理系统上的所有虚拟机使用，从而降低成本并降低实现高可用性的复杂程度。vSphere 中内置的密钥可用性功能：

- 共享存储器。通过在共享存储器（如光纤通道、iSCSI SAN 或 NAS）上存储虚拟机文件来消除单一故障点。可以使用 SAN 镜像和复制功能将虚拟磁盘的更新副本保留在灾难恢复站点。
- 网络接口绑定。允许单个网卡发生故障。
- 存储多路径。允许存储路径发生故障。

除了这些功能外，vSphere HA 和 Fault Tolerance 功能分别通过提供中断快速恢复和连续可用性来最小化或消除非计划停机时间。

vSphere HA 提供快速中断恢复

vSphere HA 利用配置为群集的多台 ESXi 主机，为虚拟机中运行的应用程序提供快速中断恢复和具有成本效益的高可用性。

vSphere HA 通过以下方式保护应用程序可用性：

- 通过在群集内的其他主机上重新启动虚拟机，防止服务器故障。
- 通过持续监控虚拟机并在检测到故障时对其进行重新设置，防止应用程序故障。

与其他群集解决方案不同，vSphere HA 提供基础架构并使用该基础架构保护所有工作负载：

- 无需在应用程序或虚拟机内安装特殊软件。所有工作负载均受 vSphere HA 保护。配置 vSphere HA 之后，不需要执行操作即可保护新虚拟机。它们会自动受到保护。
- 可以将 vSphere HA 与 vSphere Distributed Resource Scheduler (DRS) 结合使用以防止出现故障，以及在群集内的主机之间提供负载平衡。

与传统的故障切换解决方案相比，vSphere HA 具有多个优势：

最小化设置	设置 vSphere HA 群集之后，群集内的所有虚拟机无需额外配置即可获得故障切换支持。
减少了硬件成本和设置	虚拟机可充当应用程序的移动容器，可在主机之间移动。管理员会避免在多台计算机上进行重复配置。使用 vSphere HA 时，必须拥有足够的资源来对要通过 vSphere HA 保护的主机数进行故障切换。但是，vCenter Server 系统会自动管理资源并配置群集。
提高了应用程序的可用性	虚拟机内运行的任何应用程序的可用性变得更高。虚拟机可以从硬件故障中恢复，提高了在引导周期内启动的所有应用程序的可用性，而且没有额外的计算需求，即使该应用程序本身不是群集应用程序也一样。通过监控和响应 VMware Tools 检测信号并重新启动未响应的虚拟机，可防止客户机操作系统崩溃。
DRS 和 vMotion 集成	如果主机发生了故障，并且在其他主机上重新启动了虚拟机，则 DRS 会提出迁移建议或迁移虚拟机以平衡资源分配。如果迁移的源主机和/或目标主机发生故障，则 vSphere HA 会帮助从该故障中恢复。

vSphere Fault Tolerance 提供连续可用性

vSphere HA 通过在主机出现故障时重新启动虚拟机来为虚拟机提供基本级别的保护。vSphere Fault Tolerance 可提供更高级别的可用性，允许用户对任何虚拟机进行保护以防止主机发生故障时丢失数据、事务或连接。

Fault Tolerance 通过确保主虚拟机和辅助虚拟机的状态在虚拟机的指令执行的任何时间点均相同来提供连续可用性。使用 ESXi 主机平台上的 VMware vLockstep 技术来完成此过程。vLockstep 通过使主虚拟机和辅助虚拟机执行相同顺序的 x86 指令来完成此过程。主虚拟机捕获所有输入和事件（从处理器到虚拟 I/O 设备），并在辅助虚拟机上进行重放。辅助虚拟机执行与主虚拟机相同的指令序列，而仅单个虚拟机映像（主虚拟机）执行工作负载。

如果运行主虚拟机的主机或运行辅助虚拟机的主机发生故障，则会发生即时且透明的故障切换。正常运行的 ESXi 主机将无缝变成主虚拟机的主机，而不会断开网络连接或中断正在处理的事务。使用透明故障切换，不会有数据损失，并且可以维护网络连接。 在进行透明故障切换之后，将重新生成新的辅助虚拟机，并将重新建立冗余。整个过程是透明且全自动的，并且即使 vCenter Server 不可用，也会发生。

创建和使用 vSphere HA 群集 2

vSphere HA 群集允许 ESXi 主机集合作为一个组协同工作，这些主机为虚拟机提供的可用性级别比 ESXi 主机单独提供的级别要高。当规划新 vSphere HA 群集的创建和使用时，您选择的选项会影响群集对主机或虚拟机故障的响应方式。

在创建 vSphere HA 群集之前，应清楚 vSphere HA 标识主机故障和隔离以及响应这些情况的方式。还应了解接入控制的工作方式以便可以选择符合故障切换需要的策略。建立群集之后，不但可以通过高级属性自定义其行为，还可以通过执行建议的最佳做法优化其性能。

注意 尝试使用 vSphere HA 时可能会获得错误消息。有关与 vSphere HA 相关的错误消息的信息，请参见位于 http://kb.vmware.com/kb/1033634 的 VMware 知识库文章。

本章讨论了以下主题：

- 第 11 页，“vSphere HA 的工作方式”
- 第 15 页，“vSphere HA 接入控制”
- 第 20 页，“vSphere HA 对照表”
- 第 21 页，“创建 vSphere HA 群集”
- 第 25 页，“自定义 vSphere HA 行为”
- 第 27 页，“vSphere HA 群集的最佳做法”

vSphere HA 的工作方式

vSphere HA 可以将虚拟机及其所驻留的主机集中在群集内，从而为虚拟机提供高可用性。群集中的主机均会受到监控，如果发生故障，故障主机上的虚拟机将在备用主机上重新启动。

创建 vSphere HA 群集时，您可以选择使用单个主机作为首选主机以与 vCenter Server 通信并监控其他主机、从属主机及其虚拟机的状况。可能会发生不同类型的主机故障，首选主机必须检测并相应地处理故障。首选主机必须可以区分故障主机与处于网络分区中或已与网络隔离的主机。首选主机使用数据存储检测信号来确定故障的类型。

首选主机和从属主机

在将主机添加到 vSphere HA 群集时，代理将上载到主机，并配置为与群集内的其他代理通信。群集中的每台主机作为首选主机或从属主机运行。

如果为群集启用了 vSphere HA，则所有活动主机（未处于待机或维护模式的主机或未断开连接的主机）都将参与选举以选择群集的首选主机。挂载最多数量的数据存储的主机在选举中具有优势。每个群集只存在一台首选主机，其他所有主机都是从属主机。如果首选主机出现故障、关机或从群集中移除，则会进行新的选举。

群集中的首选主机具有很多职责：

- 监控从属主机的状况。如果从属主机发生故障或无法访问，首选主机将确定需要重新启动的虚拟机。
- 监控所有受保护虚拟机的电源状况。如果有一台虚拟机出现故障，首选主机可确保重新启动该虚拟机。使用本地放置引擎，首选主机还可确定执行重新启动的位置。
- 管理群集主机和受保护的虚拟机列表。
- 充当群集的 vCenter Server 管理界面并报告群集健康状况。

从属主机主要通过本地运行虚拟机、监控其运行时状况和向首选主机报告状况更新对群集发挥作用。首选主机也可运行和监控虚拟机。从属主机和首选主机都可实现虚拟机和应用程序监控功能。

首选主机执行的功能之一是虚拟机保护。虚拟机受保护时，vSphere HA 可保证在其出现故障后尝试重新打开电源。首选主机在观察到为响应用户操作，虚拟机的电源状况由关闭电源变为打开电源时，会致力于保护虚拟机。如果发生故障切换，首选主机必须重新启动所负责的受保护的虚拟机。该职责已分配给已在包含虚拟机配置文件的数据存储上以独占方式锁定系统定义的文件的首选主机。

vCenter Server 使用 vSphere HA 主机状况报告主机是首选主机还是从属主机。如果已启用“HA 状况”列，则会在 vSphere Client 中主机的**摘要**选项卡上和群集或数据中心的“主机列表”视图中报告此状况。HA 状况“正在运行 (主机)”表示主机作为 vSphere HA 首选主机。“已连接 (从属)”状况表示主机作为 vSphere HA 从属主机。还提供了其他一些状况，表示正在进行选举或发生错误情况。主机的**摘要**选项卡在主机的 vSphere HA 状况旁提供了一个链接来解释当前状况。有关错误状况的详细信息，请参见 *vSphere 故障排除*文档。

注意 如果断开主机与群集之间的连接，则所有向该主机注册的虚拟机均不受 vSphere HA 保护。

主机故障类型和检测

vSphere HA 群集的首选主机负责检测从属主机的故障。根据检测到的故障类型，在主机上运行的虚拟机可能需要进行故障切换。

在 vSphere HA 群集中，检测三种类型的主机故障：

- 主机停止运行（即发生故障）。
- 主机与网络隔离。
- 主机失去与首选主机的网络连接。

首选主机监控群集中从属主机的活跃度。此通信通过每秒交换一次网络检测信号来完成。当首选主机停止从从属主机接收这些检测信号时，它会在声明该主机已出现故障之前检查主机活跃度。首选主机执行的活跃度检查是要确定从属主机是否在与数据存储之一交换检测信号。请参见第 13 页，“数据存储检测信号”。而且，首选主机还检查主机是否对发送至其管理 IP 地址的 ICMP ping 进行响应。

如果首选主机无法直接与从属主机上的代理进行通信，则该从属主机不会对 ICMP ping 进行响应，并且该代理不会发出被视为已出现故障的检测信号。会在备用主机上重新启动主机的虚拟机。如果此类从属主机与数据存储交换检测信号，则首选主机会假定它处于某个网络分区或隔离网络中，因此会继续监控该主机及其虚拟机。请参见第 13 页，“网络分区”。

当主机仍在运行但无法再监视来自管理网络上 vSphere HA 代理的流量时，会发生主机网络隔离。如果主机停止监视此流量，则它会尝试 ping 群集隔离地址。如果仍然失败，主机将声明自己已与网络隔离。

首选主机监控在独立主机上运行的虚拟机，如果发现虚拟机的电源已关闭，而且该首选主机负责这些虚拟机，则会重新启动这些虚拟机。

注意 如果您确保网络基础结构具有足够的冗余度且至少有一个网络路径始终可用，则主机网络隔离应该在极少数情况下才出现。

网络分区

在 vSphere HA 群集发生管理网络故障时，该群集中的部分主机可能无法通过管理网络与其他主机进行通信。一个群集中可能会出现多个分区。

已分区的群集会导致虚拟机保护和群集管理功能降级。请尽快更正已分区的群集。

- 虚拟机保护。vCenter Server 允许虚拟机打开电源，但仅当虚拟机与负责它的首选主机在相同的分区中运行时，才会对其进行保护。首选主机必须与 vCenter Server 进行通信。如果首选主机以独占方式锁定包含虚拟机配置文件的数据存储上的系统定义的文件，则首选主机将负责虚拟机。
- 群集管理。vCenter Server 只能与群集中的部分主机进行通信，且只能连接到一台首选主机。因此，只有在解决分区之后，配置中影响 vSphere HA 的更改才能生效。此故障可能会导致其中一个分区在旧配置下操作，而另一个分区使用新的设置。

如果 vSphere HA 群集包含 ESXi 5.0 以前版本的主机且出现一个分区，则 vSphere HA 可能会错误地打开被用户关闭电源的虚拟机的电源，或者可能无法重新启动发生故障的虚拟机。

注意 vSphere HA 群集分区后，可以仅将主机添加到 vCenter Server 与之通信的分区。

数据存储检测信号

当 vSphere HA 群集中的首选主机无法通过管理网络与从属主机通信时，首选主机将使用数据存储检测信号来确定从属主机是否出现故障，是否位于网络分区中，或者是否与网络隔离。如果从属主机已停止数据存储检测信号，则认为该从属主机出现故障，并且其虚拟机已在别处重新启动。

vCenter Server 选择一组首选数据存储集用于检测信号。此选择会使可访问检测信号数据存储的主机数最大，并且会将数据存储受同一存储阵列或 NFS 服务器支持的可能性降至最低。要替换所选数据存储，请使用 vSphere Client 的“群集设置”对话框指定检测信号数据存储。通过**数据存储检测信号**选项卡，您可以指定备用数据存储。只有至少由两台主机挂载的数据存储可用。通过查看 HA“群集状态”对话框的**检测信号数据存储**选项卡，您还可以看到 vSphere HA 已选择使用的数据存储。

可以使用高级属性 das.heartbeatdsperhost 更改 vCenter Server 为每个主机选择的检测信号数据存储的数量。默认值为 2，最大有效值为 5。

注意 如果部署聚合网络环境（其中存储和管理网络流量通过同一物理网卡进行传输），请禁用数据存储检测信号。它在该类型的网络中没有任何优点。

vSphere HA 将在用于数据存储检测信号和保留受保护的虚拟机集的每个数据存储的根目录中创建一个目录，目录名称为 `.vSphere-HA`。请勿删除或修改存储在此目录中的文件，因为这可能会对操作产生影响。由于多个群集可能使用一个数据存储，因此将针对每个群集创建该目录的子目录。Root 用户拥有这些目录和文件，并且只有 root 用户可以读写这些目录和文件。vSphere HA 使用的磁盘空间取决于多个因素，包括所用的 VMFS 版本以及将数据存储用于信号检测的主机数。使用 vmfs3 时，最大使用量约为 2 GB，典型使用量约为 3 MB。使用 vmfs5 时，最大使用量和典型使用量约为 3 MB。vSphere HA 使用数据存储增加的开销很小，并且对其他数据存储操作的性能没有影响。

vSphere HA 安全性

多个安全功能增强了 vSphere HA。

选择已打开的防火墙端口 vSphere HA 对代理至代理的通信使用 TCP 和 UDP 端口 8182。防火墙端口将自动打开和关闭，确保仅在需要时打开端口。

使用文件系统权限保护的配置文件 vSphere HA 在本地存储器或 ramdisk（如果没有本地数据存储）上存储配置信息。使用文件系统权限保护这些文件，且仅 root 用户可以访问它们。

详细的日志记录

vSphere HA 放置日志文件的位置取决于主机版本。

- 对于 ESXi 5.x 主机，vSphere HA 默认仅写入 syslog，因此，日志放置在 syslog 所配置的放置位置。vSphere HA 日志文件名前置 `fdm`（fdm 代表故障域管理器，vSphere HA 中的一种服务）。
- 对于旧版 ESXi 4.x 和 3.x 主机，vSphere HA 写入本地磁盘上的 `/var/log/vmware/fdm` 以及 syslog（如果已配置）。
- 对于旧版 ESX 4.x 和 3.x 主机，vSphere HA 写入 `/var/log/vmware/fdm`。

安全 vSphere HA 登录

vSphere HA 使用 vCenter Server 创建的用户帐户 `vpxuser` 登录到 vSphere HA 代理。此帐户与 vCenter Server 用于管理主机的帐户相同。vCenter Server 为此帐户创建随机密码，并定期更改密码。时间段由 vCenter Server `VirtualCenter.VimPasswordExpirationInDays` 设置进行设置。可以使用 vSphere Client 中的“高级设置”控件更改设置。对主机的 root 文件夹具有管理特权的用户也可以登录到代理。

安全通信

vCenter Server 和 vSphere HA 代理之间的所有通信都是通过 SSL 完成的。除选举消息以外（通过 UDP 完成），代理至代理的通信也使用 SSL。选举消息通过 SSL 进行验证，因此，恶意代理只能阻止在其上运行代理的主机被选为首选主机。在这种情况下，将发出群集的配置问题，以便用户了解问题。

需要验证主机 SSL 证书

vSphere HA 要求每个主机都具有一个经过验证的 SSL 证书。每个主机在首次引导时都会生成一个自签署证书。然后，可以重新生成或使用机构颁发的证书替换该证书。如果证书被替换，需要重新配置主机上的 vSphere HA。如果主机在其证书更新后断开与 vCenter Server 的连接，且重新启动 ESXi 或 ESX 主机代理，则主机重新连接到 vCenter Server 时将自动重新配置 vSphere HA。如果此时因禁用 vCenter Server 主机 SSL 证书验证而没有断开连接，请验证新证书并重新配置主机上的 vSphere HA。

结合使用 vSphere HA 和 DRS

将 vSphere HA 和 Distributed Resource Scheduler (DRS) 一起使用，可将自动故障切换与负载平衡相结合。这种结合会在 vSphere HA 将虚拟机移至其他主机后生成一个更平衡的群集。

vSphere HA 执行故障切换并在其他主机上重新启动虚拟机时，其首要的优先级是所有虚拟机的立即可用性。虚拟机重新启动后，其上打开虚拟机电源的主机可能负载很重，而其他主机的负载较轻。vSphere HA 使用虚拟机的 CPU 和内存预留来确定主机是否有足够的空闲容量容纳虚拟机。

在结合使用 DRS 和 vSphere HA 并且启用了接入控制的群集内，可能不会从正在进入维护模式的主机上撤出虚拟机。这种行为的出现是由于用于重新启动虚拟机的预留资源出现了故障。必须使用 vMotion 将虚拟机手动迁出主机。

在某些情况下，vSphere HA 可能由于资源限制而无法对虚拟机进行故障切换。这种情况的出现有多种原因。

- 禁用了 HA 接入控制，但启用了 Distributed Power Management (DPM)。这会导致 DPM 将虚拟机整合到较少数量的主机上，并将空主机置于待机模式，使得没有足够的已打开电源容量来执行故障切换。
- 虚拟机-主机关联性规则（必需）可能会限制可以容纳某些虚拟机的主机。
- 可能有足够多的聚合资源，但这些资源在多台主机上是资源碎片，因此虚拟机无法使用它们进行故障切换。

在这些情况下，vSphere HA 可使用 DRS 尝试调整群集（例如，通过使主机退出待机模式或者迁移虚拟机以整理群集资源碎片），以便 HA 可以执行故障切换。

如果 DPM 处于手动模式，则可能需要确认主机打开电源建议。同样，如果 DRS 处于手动模式，可能需要确认迁移建议。

如果要使用虚拟机-主机关联性规则，请注意不能违反这些规则。如果执行故障切换违反这样的规则，则 vSphere HA 将不会执行故障切换。

有关 DRS 的详细信息，请参见 *vSphere 资源管理*文档。

vSphere HA 接入控制

vCenter Server 使用接入控制来确保群集内具有足够的资源，以便提供故障切换保护并确保考虑虚拟机资源预留。

有三种类型的接入控制可用。

主机	确保主机有足够资源来满足其上运行的所有虚拟机的预留。
资源池	确保资源池有足够资源来满足与其关联的所有虚拟机的预留、份额和限制。
vSphere HA	确保预留了足够的群集资源，以便在主机发生故障时恢复虚拟机。

接入控制对资源使用施加一些限制，违反这些限制的任何操作将不被允许。可能被禁止的操作的示例包括：

- 打开虚拟机电源。
- 将虚拟机迁移到主机、群集或资源池中。
- 增加虚拟机的 CPU 或内存预留。

对于这三种接入控制类型，只有 vSphere HA 接入控制可以被禁用。但是，如果禁用 VMware HA 接入控制，将无法保证有预期数量的虚拟机能够在故障之后重新启动。VMware 建议您不要禁用接入控制，但可能由于以下原因，需要临时将其禁用：

- 当没有足够资源来支持故障切换操作时，您需要违反故障切换限制（例如，如果您打算将主机置于待机模式以测试它们能否与 Distributed Power Management (DPM) 一起使用）。
- 如果自动过程需要执行一些操作，而这些操作可能会暂时违反故障切换限制（例如，在 vSphere Update Manager 执行的升级过程中）。
- 如果需要执行测试或维护操作。

注意 禁用 vSphere HA 接入控制后，即使 DPM 已启用并且可将所有虚拟机整合到单个主机上，vSphere HA 仍可确保群集中至少有两个主机已打开电源。这是为了确保可进行故障切换。

“群集允许的主机故障数目”接入控制策略

可以将 vSphere HA 配置为允许指定的主机故障数目。使用“群集允许的主机故障数目”接入控制策略，vSphere HA 允许指定数目的主机出现故障，同时可以确保群集内留有足够的资源来对这些主机上的虚拟机进行故障切换。

使用“群集允许的主机故障数目”策略，vSphere HA 以下列方式执行接入控制：

1 计算插槽大小。

插槽是内存和 CPU 资源的逻辑表示。默认情况下，会调整插槽的大小来满足群集中任何已打开电源虚拟机的要求。

2 确定群集内每台主机可以拥有的插槽数目。

3 确定群集的当前故障切换容量。

这是可以发生故障并仍然有足够插槽满足所有已打开电源虚拟机的主机的数目。

4 确定“当前故障切换容量”是否小于“配置的故障切换容量”（由用户提供）。

如果是，则接入控制不允许执行此操作。

插槽大小计算

插槽大小由两个组件（CPU 和内存）组成。

- vSphere HA 计算 CPU 组件的方法是先获取每台已打开电源虚拟机的 CPU 预留，然后再选择最大值。如果没有为虚拟机指定 CPU 预留，则系统会为其分配一个默认值 32 MHz。可以使用 das.vmcpuminmhz 高级属性更改此值。
- vSphere HA 计算内存组件的方法是先获取每台已打开电源虚拟机的内存预留和内存开销，然后再选择最大值。内存预留没有默认值。

如果群集内虚拟机的预留值大小不一致，则会影响插槽大小的计算。为避免出现这种情况，可以使用 das.slotcpuinmhz 或 das.slotmeminmb 高级属性分别指定插槽大小的 CPU 或内存组件的上限。请参见第 25 页，“vSphere HA 高级属性”。

使用插槽数目计算当前故障切换容量

计算出插槽大小后，vSphere HA 会确定每台主机中可用于虚拟机的 CPU 和内存资源。这些值包含在主机的根资源池中，而不是主机的总物理资源中。通过使用 vSphere Client 直接连接到主机，然后导航至主机的**资源**选项卡，可以找到由 vSphere HA 使用的主机资源数据。如果群集中的所有主机均相同，则可以用群集级别指数除以主机的数量来获取此数据。不包括用于虚拟化目的的资源。只有处于连接状态、未进入维护模式且没有任何 vSphere HA 错误的主机才列入计算范畴。

然后，即可确定每台主机可以支持的最大插槽数目。为确定此数目，请用主机的 CPU 资源数除以插槽大小的 CPU 组件，然后将结果化整。对主机的内存资源数进行同样的计算。然后，比较这两个数字，较小的那个数字即为主机可以支持的插槽数。

通过确定可以发生故障并仍然有足够插槽满足所有已打开电源虚拟机要求的主机的数目（从最大值开始）来计算当前故障切换容量。

高级运行时信息

如果选择“群集允许的主机故障数目”接入控制策略，**高级运行时信息**链接则会在 vSphere Client 中群集**摘要**选项卡上的 vSphere HA 区域中显示。单击此链接以显示有关群集的下列信息：

- 插槽大小。
- 群集内的插槽总数。群集内正常主机所支持的插槽总数。
- 已使用的插槽数。分配给已打开电源的虚拟机的插槽数目。如果已使用高级选项定义插槽大小的上限，则此数目可以大于已打开电源的虚拟机的数目。这是因为有些虚拟机会占用多个插槽。
- 可用插槽数。可用于打开群集内其他虚拟机的电源的插槽数量。vSphere HA 保留故障切换所需的插槽数量。剩余的插槽可用于打开新虚拟机电源。
- 故障切换插槽数。除已使用的插槽和可用插槽之外的插槽总数。
- 群集中已打开电源虚拟机的总数。
- 群集中的主机总数。
- 群集中的正常主机总数。处于连接状态、未进入维护模式而且没有 vSphere HA 错误的主机数目。

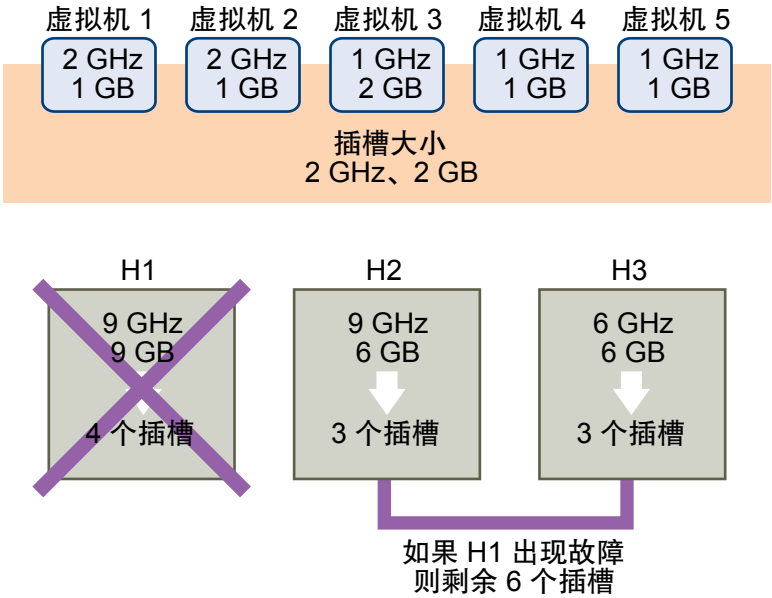
示例：使用“群集允许的主机故障数目”策略的接入控制

示例中展示了使用此接入控制策略计算和使用插槽大小的方式。对群集进行如下假设：

- 群集包括三台主机，每台主机上可用的 CPU 和内存资源数各不相同。第一台主机 (H1) 的可用 CPU 资源和可用内存分别为 9 GHz 和 9 GB，第二台主机 (H2) 为 9 GHz 和 6 GB，而第三台主机 (H3) 则为 6 GHz 和 6 GB。

- 群集内存在五个已打开电源的虚拟机，其 CPU 和内存要求各不相同。VM1 所需的 CPU 资源和内存分别为 2 GHz 和 1 GB，VM2 为 2 GHz 和 1 GB，VM3 为 1 GHz 和 2 GB，VM4 为 1 GHz 和 1 GB，VM5 则为 1 GHz 和 1 GB。
- “群集允许的主机故障数目”设置为 1。

图 2–1 使用“群集允许的主机故障数目”策略的接入控制示例



1 比较虚拟机的 CPU 和内存要求，然后选择最大值，从而计算出插槽大小。

   最大 CPU 要求（由 VM1 和 VM2 共享）为 2 GHz，而最大内存要求（针对 VM3）为 2 GB。根据上述情况，插槽大小为 2 GHz CPU 和 2 GB 内存。

2 由此可确定每台主机可以支持的最大插槽数目。

   H1 可以支持四个插槽。H2 可以支持三个插槽（取 9GHz/2GHz 和 6GB/2GB 中较小的一个），H3 也可以支持三个插槽。

3 计算出当前故障切换容量。

   最大的主机是 H1，如果它发生故障，群集内还有六个插槽，足够供所有五个已打开电源的虚拟机使用。如果 H1 和 H2 都发生故障，群集内将仅剩下三个插槽，这是不够用的。因此，当前故障切换容量为 1。

群集内可用插槽的数目为 1（H2 和 H3 上的六个插槽减去五个已使用的插槽）。

“预留的群集资源的百分比”接入控制策略

可以将 vSphere HA 配置为通过预留特定百分比的群集 CPU 和内存资源来执行接入控制，用于从主机故障中进行恢复。

使用“预留的群集资源的百分比”接入控制策略，vSphere HA 可确保预留 CPU 和内存资源总量的指定百分比以用于故障切换。

使用“预留的群集资源”策略，vSphere HA 可强制执行下列接入控制：

1 计算群集内所有已打开电源虚拟机的总资源要求。

2 计算可用于虚拟机的主机资源总数。

3 计算群集的“当前的 CPU 故障切换容量”和“当前的内存故障切换容量”。

4 确定“当前的 CPU 故障切换容量”或“当前的内存故障切换容量”是否小于对应的“配置的故障切换容量”（由用户提供）。

如果是，则接入控制不允许执行此操作。

vSphere HA 将使用虚拟机的实际预留。如果虚拟机没有预留（即预留量为 0），则将应用默认设置（0 MB 内存和 256 MHz CPU）。

注意 “预留的群集资源的百分比”接入控制策略还会检查群集中是否至少有两个已启用 vSphere HA 的主机（不包括正在进入维护模式的主机）。如果只有一个已启用 vSphere HA 的主机，即使可以使用足够的资源百分比，也不允许执行此操作。进行此次额外检查的原因在于如果群集中只有一个主机，则 vSphere HA 无法进行故障切换。

计算当前故障切换容量

已打开电源的虚拟机的总资源要求由两个组件组成，即 CPU 和内存。vSphere HA 将计算这些值。

- CPU 组件值的计算方法是：加总已打开电源虚拟机的 CPU 预留。如果没有为虚拟机指定 CPU 预留，则系统会为其分配一个默认值 256 MHz（可以使用 das.vmcpuminmhz 高级属性更改此值）。
- 内存组件值的计算方法是：加总每台已打开电源虚拟机的内存预留（以及内存开销）。

计算出主机的 CPU 和内存资源总和，从而得出虚拟机可使用的主机资源总数。这些值包含在主机的根资源池中，而不是主机的总物理资源中。不包括用于虚拟化目的的资源。只有处于连接状态、未进入维护模式而且没有 vSphere HA 错误的主机才列入计算范畴。

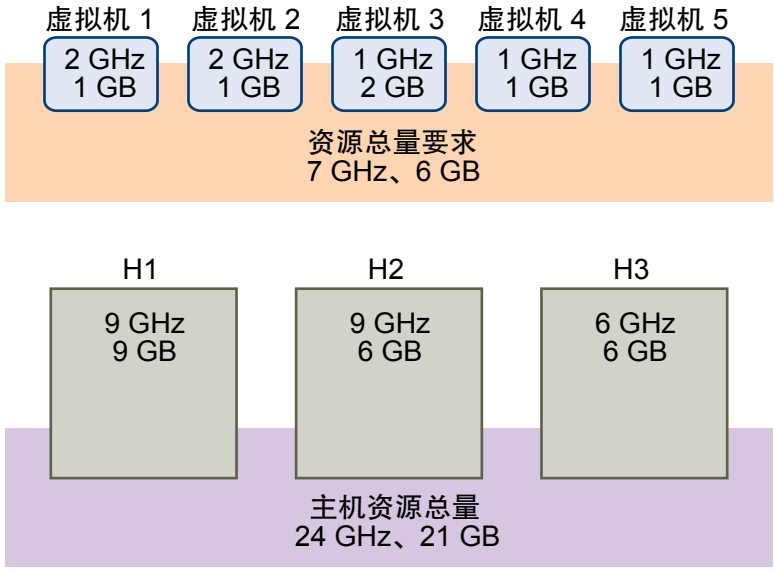
先用主机 CPU 资源总数减去总 CPU 资源要求，然后再用这个结果除以主机 CPU 资源总数，从而计算出“当前的 CPU 故障切换容量”。“当前的内存故障切换容量”的计算方式与之相似。

示例：使用“预留的群集资源的百分比”策略的接入控制

示例中展示了使用此接入控制策略计算和使用“当前故障切换容量”的方式。对群集进行如下假设：

- 群集包括三台主机，每台主机上可用的 CPU 和内存资源数各不相同。第一台主机 (H1) 的可用 CPU 资源和可用内存分别为 9 GHz 和 9 GB，第二台主机 (H2) 为 9 GHz 和 6 GB，而第三台主机 (H3) 则为 6 GHz 和 6 GB。
- 群集内存在五个已打开电源的虚拟机，其 CPU 和内存要求各不相同。VM1 所需的 CPU 资源和内存分别为 2 GHz 和 1 GB，VM2 为 2 GHz 和 1 GB，VM3 为 1 GHz 和 2 GB，VM4 为 1 GHz 和 1 GB，VM5 则为 1 GHz 和 1 GB。
- “配置的故障切换容量”设置为 25%。

图 2–2 使用“预留的群集资源的百分比”策略的接入控制示例



已打开电源的虚拟机的总资源要求为 7 GHz CPU 和 6 GB 内存。可用于虚拟机的主机资源总数为 24 GHz CPU 和 21 GB 内存。根据上述情况，“当前的 CPU 故障切换容量”为 70% ((24GHz - 7GHz)/24GHz)。同样，“当前的内存故障切换容量”为 71% ((21GB-6GB)/21GB)。

由于群集的“配置的故障切换容量”设置为 25%，因此仍然可使用 45% 的群集 CPU 资源总数和 46% 的群集内存资源打开其他虚拟机电源。

“指定故障切换主机”接入控制策略

在配置 vSphere HA 时可以将特定主机指定为故障切换主机。

如果使用“指定故障切换主机”接入控制策略，则在主机发生故障时，vSphere HA 将尝试在指定的故障切换主机之一上重新启动其虚拟机。如果不能使用此方法（例如，故障切换主机发生故障或者资源不足时），则 vSphere HA 会尝试在群集内的其他主机上重新启动这些虚拟机。

为了确保故障切换主机上拥有可用的空闲容量，将阻止您打开虚拟机电源或使用 vMotion 将虚拟机迁移到故障切换主机。而且，为了保持负载平衡，DRS 也不会使用故障切换主机。

注意 如果使用“指定故障切换主机”接入控制策略，并指定多个故障切换主机，则 DRS 不会对故障切换主机进行负载平衡，且不支持虚拟机-虚拟机关联性规则。

vSphere Client 中群集的**摘要**选项卡的 vSphere HA 区域内显示了“当前故障切换主机”。每个主机旁边的状态图标可以是绿色、黄色或红色。

- 绿色。主机处于连接状态、未进入维护模式且没有 vSphere HA 错误。主机上没有任何已打开电源的虚拟机。
- 黄色。主机处于连接状态、未进入维护模式且没有 vSphere HA 错误。但是，主机上驻留了已打开电源的虚拟机。
- 红色。主机已断开连接、处于维护模式或存在 vSphere HA 错误。

选择接入控制策略

应当基于可用性需求和群集的特性选择 vSphere HA 接入控制策略。选择接入控制策略时，应当考虑的因素很多。

避免资源碎片

当总计有足够资源用于虚拟机故障切换时，将出现资源碎片。但是，这些资源位于多个主机上并且不可用，因为虚拟机一次只能在一个 ESXi 主机上运行。通过将插槽定义为虚拟机最大预留值，“群集允许的主机故障数目”策略可避免资源碎片。“群集资源的百分比”策略不解决资源碎片问题。使用“指定故障切换主机”策略不会出现资源碎片，因为该策略会为故障切换预留主机。

故障切换资源预留的灵活性

为故障切换保护预留群集资源时，接入控制策略所提供的控制粒度会有所不同。“群集允许的主机故障数目”策略允许设置多个主机作为故障切换级别。“群集资源的百分比”策略最多允许指定 100% 的群集 CPU 或内存资源用于故障切换。通过“指定故障切换主机”策略可以指定一组故障切换主机。

群集的异构性

从虚拟机资源预留和主机总资源容量方面而言，群集可以异构。在异构群集内，“群集允许的主机故障数目”策略可能过于保守，因为在定义插槽大小时它仅考虑最大虚拟机预留，而在计算当前故障切换容量时也假设最大主机发生故障。其他两个接入控制策略不受群集异构性影响。

注意 vSphere HA 在执行接入控制计算时会包括 Fault Tolerance 辅助虚拟机的资源使用情况。对于“群集允许的主机故障数目”策略，将为辅助虚拟机分配一个插槽；而对于“群集资源的百分比”策略，在计算群集的可用容量时将考虑辅助虚拟机的资源使用情况。

---

vSphere HA 对照表

vSphere HA 对照表包含在创建和使用 vSphere HA 群集之前需要了解的要求。

vSphere HA 群集的要求

在设置 vSphere HA 群集之前，应查看此列表。有关详细信息，请遵循相应的交叉引用或者参见第 21 页，“创建 vSphere HA 群集”。

- 所有主机必须获得 vSphere HA 许可。

  注意 ESX/ESXi 3.5 主机受 vSphere HA 的支持，但必须包含用于解决文件锁问题的修补程序。对于 ESX 3.5 主机，必须应用修补程序 ESX350-201012401-SG；对于 ESXi 3.5，必须应用修补程序 ESXe350-201012401-I-BG。应用这些修补程序之前必须先应用必备修补程序。

- 群集中至少需要有两台主机。
- 需要为所有主机配置静态 IP 地址。如果使用的是 DHCP，必须确保每台主机的地址在重新引导期间保留。
- 所有主机应该至少有一个公共的管理网络，最佳做法则至少需要有两个。所使用的主机版本不同，管理网络也不同。
  - ESX 主机 - 服务控制台网络。
  - 版本 4.0 之前的 ESXi 主机 - VMkernel 网络。
  - 版本 4.0 及更高版本的 ESXi 主机 - 已选中了**管理网络**复选框的 VMkernel 网络。

  请参见第 28 页，“网络连接的最佳做法”。

- 为了确保任何虚拟机都可以在群集内的任何主机上运行，所有主机都应该可以访问相同的虚拟机网络和数据存储。同样，虚拟机必须位于共享而非本地存储器上，否则在主机出现故障时它们将无法进行故障切换。

  注意 vSphere HA 使用数据存储信号检测来区分已分区的主机、已隔离的主机和出现故障的主机。相应地，必须确保为 vSphere HA 预留的数据存储始终立即可用。

- 为了使虚拟机监控工作，必须安装 VMware Tools。请参见第 24 页，“虚拟机和应用程序监控”。
- 应启用主机证书检查功能。请参见第 36 页，“启用主机证书检查”。
- vSphere HA 同时支持 IPv4 和 IPv6。但是，混合使用这两个版本协议的群集更可能导致网络分区。

创建 vSphere HA 群集

vSphere HA 在 ESXi（或旧版 3.5、4.0 和 4.1 ESX）主机群集的环境中运行。必须创建群集，然后用主机填充群集，并配置 vSphere HA 设置，才能建立故障切换保护。

创建 vSphere HA 群集时，必须配置许多可决定功能如何运行的设置。在此之前，请确定群集的节点。这些节点是为支持虚拟机而提供资源，并且将由 vSphere HA 用于故障切换保护的 ESXi 主机。然后应当确定如何互相连接这些节点，以及如何将这些节点连接到虚拟机数据所在的共享存储器。在建立好网络架构后，可以将主机添加到群集并完成 vSphere HA 配置。

将主机节点添加到群集之前，可以启用和配置 vSphere HA。但是，在将主机添加到群集之前，群集的所有功能并非都能运行，部分群集设置不可用。例如，在出现可以指定为故障切换主机的主机之前，“指定故障切换主机”接入控制策略不可用。

注意 为处于（或移入）vSphere HA 群集的主机上驻留的所有虚拟机禁用“虚拟机启动和关机”（自动启动）功能。与 vSphere HA 配合使用时，不支持自动启动。

创建 vSphere HA 群集

可以为群集启用 vSphere HA。启用了 vSphere HA 的群集是 Fault Tolerance 的必备条件。VMware 建议您首先创建空群集。在规划好群集的资源和网络架构之后，可以使用 vSphere Client 将主机添加到群集，并指定群集的 vSphere HA 设置。

使用具有群集管理员权限的帐户将 vSphere Client 连接到 vCenter Server。

前提条件

确认所有虚拟机及其配置文件都驻留在共享存储器上。确认主机配置为具有该共享存储器的访问权限，以便您可以使用群集中的不同主机打开虚拟机电源。

确认主机配置为具有虚拟机网络的访问权限。

注意 VMware 建议为 vSphere HA 设置冗余管理网络连接。有关设置网络冗余的信息，请参见第 29 页，“网络路径冗余”。

步骤

1 选择“主机和群集”视图。

2 右键单击清单树中的数据中心，然后单击**新建群集**。

3 完成新建群集向导。

此时不要启用 vSphere HA（或 DRS）。

4 单击**完成**，关闭向导并创建群集。

此时创建了一个空群集。

5 根据您的群集资源和网络架构计划，使用 vSphere Client 将主机添加到群集。

6 右键单击群集，然后单击**编辑设置**。

在群集的“设置”对话框中，您可以修改群集的 vSphere HA（和其他）设置。

7 在“群集功能”页上，选择**打开 vSphere HA**。

8 根据需要为群集配置 vSphere HA 设置。

- 主机监控状态
- 接入控制
- 虚拟机选项
- 虚拟机监控
- 数据存储检测信号

9 单击**确定**关闭群集的“设置”对话框。

您已配置 vSphere HA 群集，其中包含可用主机。

群集功能

新建群集向导中的第一个面板可用于为群集指定基本选项。

在该面板中，可以指定群集名称并选择一个或两个群集功能。

名称 指定群集的名称。该名称显示在 vSphere Client 清单面板中。必须指定一个名称，才能继续创建群集。

打开 vSphere HA 如果选中此复选框，则在主机出现故障时，虚拟机将在群集内的其他主机上重新启动。要在群集内的任何虚拟机上启用 vSphere Fault Tolerance，必须打开 vSphere HA。

打开 vSphere DRS 如果选中此复选框，则 DRS 将平衡整个群集的虚拟机负载。即使虚拟机受 HA 保护，DRS 也会放置并迁移虚拟机。

以后可更改上面的任何群集功能。

主机监控状态

创建群集后，请启用主机监控以便 vSphere HA 可以监控由群集内每个主机上的 vSphere HA 代理发送的检测信号。

如果选择**启用主机监控**，则会检查群集内的每台主机以确保其正在运行。如果某台主机出现故障，则会在另一台主机上重新启动虚拟机。主机监控还是 vSphere Fault Tolerance 恢复进程正常运行所必需的。

注意 如果需要执行可能会触发主机隔离响应的网络维护，VMware 建议首先禁用主机监控以挂起 vSphere HA。完成维护后，请重新启用“主机监控”。

启用或禁用接入控制

通过新建群集向导，可以为 vSphere HA 群集启用或禁用接入控制，并选择有关其执行方式的策略。

可以为 vSphere HA 群集启用或禁用接入控制。

启用：禁止违反可用性限制的打开虚拟机电源操作	启用接入控制并执行可用性限制，同时保留故障切换容量。不允许在虚拟机上执行会减少群集内的未预留资源并违反可用性限制的任何操作。
禁用：允许违反可用性限制的打开虚拟机电源操作	禁用接入控制。例如，即使打开虚拟机电源会造成故障切换容量不足，仍然可执行该操作。执行该操作时，不会显示任何警告，而且群集不会变为红色。如果群集的故障切换容量不足，vSphere HA 仍可以执行故障切换，并使用“虚拟机重新启动优先级”设置来确定要先打开电源的虚拟机。

如果启用了接入控制，vSphere HA 会提供三个强制接入控制的策略。

- 群集允许的主机故障数量
- 作为故障切换空间容量保留的群集资源的百分比
- 指定故障切换主机

注意 有关 vSphere HA 接入控制如何工作的详细信息，请参见第 20 页，“选择接入控制策略”。

虚拟机选项

默认虚拟机设置控制虚拟机的重新启动顺序（虚拟机重新启动优先级）以及在主机之间失去网络连接时 vSphere HA 的响应方式（主机隔离响应）。

这些设置适用于主机发生故障或主机隔离时群集内的所有虚拟机。此外，也可以为特定虚拟机配置异常。请参见第 27 页，“自定义各个虚拟机的 vSphere HA 行为”。

虚拟机重新启动优先级设置

虚拟机重新启动优先级确定主机发生故障后虚拟机的重新启动相对顺序。这些虚拟机在新主机上按顺序重新启动，首先启动优先级最高的虚拟机，然后是那些低优先级的虚拟机，直到重新启动所有虚拟机或者没有更多的可用群集资源为止。如果主机故障数目超过了接入控制所允许的数目，则系统可能会等到有更多资源可用时再重新启动优先级较低的虚拟机。如果指定了一个故障切换主机，则虚拟机将在该故障切换主机上重新启动。

此设置的值为：已禁用、低、中等（默认）和高。如果选择“已禁用”，则会为虚拟机禁用 vSphere HA，这意味着当其主机出现故障时不会在其他 ESXi 主机上重新启动虚拟机。“已禁用”设置不会对虚拟机监控造成影响，这意味着当正常运行的主机上的某个虚拟机出现故障时，会在同一主机上重置该虚拟机。您可更改各个虚拟机的这种设置。

虚拟机的重新启动优先级设置因用户需求而有所不同。VMware 建议为提供最重要服务的虚拟机分配较高的重新启动优先级。

例如，在多层应用程序中，可以根据虚拟机上所驻留的功能来对分配进行排序。

- 高。将为应用程序提供数据的数据库服务器。
- 中等。使用数据库中的数据并在网页上提供结果的应用程序服务器。
- 低。接收用户请求、将查询传递到应用程序服务器并将结果返回给用户的 Web 服务器。

主机隔离响应设置

主机隔离响应确定当 vSphere HA 群集内的某个主机失去其管理网络连接但仍继续运行时出现的情况。主机隔离响应要求启用“主机监控状态”。如果“主机监控状态”处于禁用状态，则主机隔离响应将同样被挂起。当主机无法与其他主机上运行的代理通信且无法 ping 其隔离地址时，该主机确定其已被隔离。发生这种情况时，主机会执行其隔离响应。响应包括：保持打开电源（默认值）、关闭电源和关机。还可以为各个虚拟机自定义此属性。

要使用“关机”设置，必须在虚拟机的客户机操作系统中安装 VMware Tools。将虚拟机关机的优点在于可以保留其状况。关机操作优于关闭虚拟机电源操作，关闭虚拟机电源不会将最近的更改刷新到磁盘中，也不会提交事务。在关机完成时，正在关机的虚拟机需要更长时间进行故障切换。未在 300 秒内或在高级属性 das.isolationshutdowntimeout 中指定的秒数内关机的虚拟机将被关闭电源。

注意 创建 vSphere HA 群集后，可以替代特定虚拟机的“重新启动优先级”和“隔离响应”的默认群集设置。此替代操作对于用于特殊任务的虚拟机很有帮助。例如，可能需要先打开提供基础架构服务（如 DNS 或 DHCP）的虚拟机的电源，再打开群集内的其他虚拟机的电源。

如果主机禁用其隔离响应（即隔离时使虚拟机处于打开电源状态）且无法访问管理和存储网络，则可能发生“裂脑”情况。在这种情况下，即使虚拟机的原始实例仍在独立主机上运行，独立主机也会丢失磁盘锁且虚拟机会故障切换至另一主机。当主机摆脱隔离时，将会有两个虚拟机副本，尽管最初独立主机上的副本没有权限访问 vmdk 文件进而避免了数据损坏。在 vSphere Client 中，虚拟机显示为在两个主机之间来回切换。

为了从此情况中恢复，ESXi 会针对已丢失硬盘锁的虚拟机生成一个问题（关于主机何时摆脱隔离状态并认识到无法重新获取磁盘锁）。vSphere HA 将自动回答该问题，这就使已丢失磁盘锁的虚拟机实例关闭电源，只留下具有磁盘锁的实例。

虚拟机和应用程序监控

如果在设置的时间内没有收到单个虚拟机的 VMware Tools 检测信号，虚拟机监控将重新启动该虚拟机。同样，如果没有收到虚拟机正在运行的应用程序的检测信号，应用程序监控也可以重新启动该虚拟机。可以启用这些功能，并配置 vSphere HA 监控无响应时的敏感度。

启用虚拟机监控后，虚拟机监控服务（使用 VMware Tools）将通过检查正在客户机内运行的 VMware Tools 进程的常规检测信号和 I/O 活动来评估群集内的每个虚拟机是否正在运行。如果没有收到检测信号或 I/O 活动，则很有可能是客户机操作系统出现故障，或未分配给 VMware Tools 用来完成任务的时间。在这种情况下，虚拟机监控服务会先确定虚拟机已发生故障，然后决定重新引导虚拟机以还原服务。

有时，仍然正常工作的虚拟机或应用程序会停止发送检测信号。为了避免不必要的重置，虚拟机监控服务还监控虚拟机的 I/O 活动。如果在故障时间间隔内未收到任何检测信号，则会检查 I/O 统计间隔（群集级别属性）。I/O 统计间隔确定在前两分钟（120 秒）内是否已发生与虚拟机有关的任何磁盘或网络活动。如果没有，则重置该虚拟机。可以使用高级属性 das.iostatsinterval 更改此默认值（120 秒）。

要启用应用程序监控，必须先获取相应的 SDK（或使用可支持 VMware 应用程序监控的应用程序），然后使用它来设置要监控的应用程序的自定义检测信号。完成此操作后，应用程序监控的工作方式将与虚拟机监控的工作方式大致相同。如果在指定时间内没有收到应用程序的检测信号，将重新启动其虚拟机。

您可以配置监控敏感度的级别。高敏感度监控可以更快得出已发生故障的结论。然而，如果受监控的虚拟机或应用程序实际上仍在运行，但由于资源限制等因素导致未收到检测信号，高敏感度监控可能会错误地认为此虚拟机发生了故障。低敏感度监控会延长实际故障和虚拟机重置之间服务中断的时间。请选择一个有效折衷满足需求的选项。

表 2–1 介绍了监控敏感度的默认设置。也可以通过选中**自定义**复选框来指定监控敏感度和 I/O 统计间隔的自定义值。

表 2–1 虚拟机监控设置

设置	故障时间间隔（秒）	重置期
高	30	1 小时
中	60	24 小时
低	120	7 天

检测到故障后，vSphere HA 会重置虚拟机。重置可确保这些服务仍然可用。为了避免因非瞬态错误而反复重置虚拟机，默认情况下，在某个可配置的时间间隔内将对虚拟机仅重置三次。在对虚拟机执行过三次重置后，指定的时间结束之前，vSphere HA 不会在后续故障出现后进一步尝试重置虚拟机。可以使用**每个虚拟机的最大重置次数**自定义设置来配置重置次数。

自定义 vSphere HA 行为

建立群集后，可以修改会对 vSphere HA 行为方式造成影响的特定属性。还可以更改由单个虚拟机继承的群集默认设置。

检查可用于优化环境中 vSphere HA 群集的高级设置。因为这些属性会影响 vSphere HA 的运行，所以更改时请小心谨慎。

设置高级 vSphere HA 选项

要自定义 vSphere HA 行为，请设置高级 vSphere HA 选项。

前提条件

要修改其设置的 vSphere HA 群集。

群集管理员特权。

步骤

1 在群集的“设置”对话框中，选择 **vSphere HA**。

2 单击**高级选项**按钮，打开“高级选项(HA)”对话框。

3 输入要在**选项**列中的文本框中更改的每个高级属性，并在**值**列中输入一个值。

4 单击**确定**。

群集即可以使用已添加或修改的选项。

vSphere HA 高级属性

可以设置影响 vSphere HA 群集的行为的高级属性。

表 2–2 vSphere HA 高级属性

属性	描述
das.isolationaddress[...]	设置为了确定主机是否与网络隔离而要 ping 的地址。只有当未从群集内的任何其他主机接收到检测信号时才 ping 此地址。如果未指定，则使用管理网络的默认网关。此默认网关必须是可用的可靠地址，以便主机可以确定它是否与网络隔离。可以为群集指定多个隔离地址（最多 10 个）：das.isolationaddressX，其中 X = 1-10。通常每个管理网络应指定一个隔离地址。如果指定的地址太多，则进行隔离检测所需的时间将会较长。
das.usedefaultisolationaddress	默认情况下，vSphere HA 使用控制台网络的默认网关作为隔离地址。此属性指定是否使用此默认设置 (true|false)。

表 2–2 vSphere HA 高级属性（续）

属性	描述
das.isolationshutdowntimeout	关闭虚拟机电源之前，系统等待虚拟机关机的时间段。只有在主机的隔离响应为“关闭虚拟机”时，此选项才适用。默认值为 300 秒。
das.slotmeminmb	定义内存插槽大小的最大限制。如果使用此选项，则插槽大小小于该值，或是小于群集内任何已打开电源虚拟机的最大内存预留以及内存开销。
das.slotcpuinmhz	定义 CPU 插槽大小的最大限制。如果使用此选项，则插槽大小小于该值，或是小于群集内任何已打开电源虚拟机的最大 CPU 预留。
das.vmmemoryminmb	定义在没有指定虚拟机内存预留或者内存预留为零时，分配给虚拟机的默认内存资源值。用于“群集允许的主机故障数目”接入控制策略。如果未指定任何值，则默认值为 0 MB。
das.vmcpuminmhz	定义在没有指定虚拟机 CPU 预留或者内存预留为零时，分配给虚拟机的默认 CPU 资源值。用于“群集允许的主机故障数目”接入控制策略。如果未指定任何值，则默认值为 256 MHz。
das.iostatsinterval	更改虚拟机监控敏感度的默认 I/O 统计间隔。默认值为 120（秒）。可以设置为大于等于 0 的任何值。设置为 0 会禁用检查。
das.ignoreinsufficienthbdatastore	如果主机不具有足够的 vSphere HA 检测信号数据存储，则禁用创建的配置问题。默认值为 false。
das.heartbeatdsperhost	更改所需的检测信号数据存储的数量。有效值范围为 2 至 5，默认值为 2。

注意 如果更改以下任一高级属性的值，则必须先禁用 vSphere HA，再重新启用它，更改才会生效。

- das.isolationaddress[...]
- das.usedefaultisolationaddress
- das.isolationshutdowntimeout

不再受支持的选项

在 vCenter Server 5.0 中，vSphere HA 的许多高级配置选项不再受支持。以下选项不再受支持。

- das.consoleUser
- das.consoleNode
- das.consolePerm
- das.primaryCount
- das.checkVmStateDelay
- das.trace
- das.traceLevel
- das.traceOutput
- das.preferredPrimaries
- das.disableUWSwapRequirement
- das.sensorPollingFreq
- das.bypassNetCompatCheck

- das.defaultfailoverhost
- das.failureDetectionTime
- das.failureDetectionInterval

如果尝试设置某个不受支持的选项，vCenter Server 5.0 会报告该选项无效。另外，如果从定义了其中某些选项的早期版本升级到 vCenter Server 5.0，这些选项将被删除且不再有效。

自定义各个虚拟机的 vSphere HA 行为

vSphere HA 群集中的每个虚拟机均分配了“虚拟机重新启动优先级”、“主机隔离响应”和“虚拟机监控”的群集默认设置。可以通过更改这些默认项来指定每个虚拟机的特定行为。如果虚拟机离开该群集，则将丢弃这些设置。

步骤

1 选择群集并在右键菜单中选择**编辑设置**。

2 选择 vSphere HA 下面的**虚拟机选项**。

3 在“虚拟机设置”窗格中，选择虚拟机，并自定义其**虚拟机重新启动优先级**或**主机隔离响应**设置。

4 在 vSphere HA 下方选择**虚拟机监控**。

5 在“虚拟机设置”窗格中，选择虚拟机，并自定义其**虚拟机监控**设置。

6 单击**确定**。

现在，虚拟机的行为将不同于您更改的每项设置的群集默认设置的行为。

vSphere HA 群集的最佳做法

为确保 vSphere HA 群集性能最佳，VMware 建议您遵循一些最佳做法。在设计和实现群集时，网络配置和冗余非常重要。

将警报设置为监控群集更改

当 vSphere HA 或 Fault Tolerance 执行用于维护可用性的操作时（例如，虚拟机故障切换），可能会向您通知此类更改。将 vCenter Server 中的警报配置为在执行这些操作时触发，并向指定的一组管理员发送警示（如电子邮件）。

提供多个默认的 vSphere HA 警报。

- 故障切换资源不足（群集警报）
- 找不到主机（群集警报）
- 正在进行故障切换（群集警报）
- 主机 HA 状态（主机警报）
- VM 监控错误（虚拟机警报）
- 虚拟机监控操作（虚拟机警报）
- 故障切换失败（虚拟机警报）

注意 默认的警报包括功能名称 vSphere HA。

监控群集有效性

有效群集是尚未违反接入控制策略的群集。

当已打开电源的虚拟机数超过了故障切换需求，即当前故障切换容量小于所配置的故障切换容量，则已启用 vSphere HA 的群集将会变为无效（红色）。如果禁用了接入控制，则群集不会变为无效。

vSphere Client 中的群集的**摘要**选项卡显示群集的配置问题的列表。该列表说明了造成群集无效或过载（黄色）的原因。

如果群集是由于 vSphere HA 问题而变为红色的，则 DRS 行为不会受到影响。

检查群集的操作状态

群集或其主机可能会出现一些配置问题和其他问题，这会对 vSphere HA 的正常运行产生负面影响。可以查看“群集操作状态”屏幕来监控这些错误，能够从 vSphere Client 中群集的**摘要**选项卡的 vSphere HA 部分访问此屏幕。解决此屏幕上列出的问题。

大部分配置问题都有记录的匹配事件。所有 vSphere HA 事件在描述中都包括“vSphere HA”。可以通过搜索此术语来查找相应事件。

vSphere HA 和 Storage vMotion 互操作性

在广泛使用 Storage vMotion 或启用存储 DRS 的群集中，VMware 建议不要部署 vSphere HA。vSphere HA 可通过在某个 ESXi 版本的主机上（不同于出现故障前运行虚拟机的主机版本）重新启动虚拟机来响应主机故障。如果出现故障时虚拟机被卷进 ESXi 5.0 主机上的 Storage vMotion 操作，而且 vSphere HA 在低于 ESXi 5.0 版本的主机上重新启动虚拟机，则会出现问题。虽然虚拟机可能打开电源，但针对快照操作的任何后续尝试都可能会使 vdisk 状态遭到损坏并导致虚拟机不可用。

接入控制最佳做法

以下建议是 vSphere HA 接入控制的最佳做法。

- 选择“预留的群集资源的百分比”接入控制策略。该策略在主机和虚拟机所需空间方面提供了极高的灵活性。大多数情况下，采用 1/N（其中 N 是群集中的节点总数）会得到足够的备用容量。
- 确保设置的所有群集主机大小相等。群集不平衡会导致保留多余的容量用于处理最可能出现故障的节点。
- 尝试使所有已配置的虚拟机的虚拟机空间需求相似。“群集允许的主机故障数目”接入控制策略使用插槽大小来计算需要为每个虚拟机预留的容量。插槽大小取决于任一虚拟机需要的最大预留内存和 CPU。当混用 CPU 和内存要求不同的多个虚拟机时，插槽大小计算将默认为最大可能值（限制整合）。

将 Auto Deploy 与 vSphere HA 配合使用

可以将 vSphere HA 与 Auto Deploy 配合使用来提高虚拟机的可用性。Auto Deploy 可在打开主机电源时置备这些主机，您还可以将其配置为在引导过程中于这些主机上安装 vSphere HA 代理。要使 Auto Deploy 安装 vSphere HA 代理，分配给主机的映像配置文件必须包含 vmware-fdm VIB。有关详细信息，请参见《vSphere 安装和设置》中的 Auto Deploy 文档。

网络连接的最佳做法

VMware 就主机网卡配置和 vSphere HA 的网络拓扑提供最佳做法。最佳做法包括对 ESXi 主机的建议，以及对电缆、交换机、路由器和防火墙的建议。

网络配置和维护

下列网络维护建议可以帮助您避免对由于丢失 vSphere HA 检测信号而发生故障的主机和网络隔离的意外检测。

- 对群集 ESXi 主机所在的网络进行更改时，VMware 建议您挂起主机监控功能。更改网络硬件或网络连接设置会中断 vSphere HA 用于检测主机故障的检测信号，并且这可能导致不必要的虚拟机故障切换尝试。

- 在 ESXi 主机上更改网络配置时（例如，添加端口组或移除 vSwitch），VMware 建议除了挂起主机监控以外，还应将在其中进行了更改的主机置于维护模式。主机退出维护模式后，会重新配置该主机，从而重新检查正在运行的主机的网络信息。如果主机未置于维护模式，则将使用旧的网络配置信息运行 vSphere HA 代理。

注意 由于网络是 vSphere HA 的一个重要组件，因此，如果需要执行网络维护，请通知 vSphere HA 管理员。

用于 vSphere HA 通信的网络

要识别哪些网络操作可能会中断 vSphere HA 的运行，应当了解哪些管理网络用于检测信号和其他 vSphere HA 通信。

- 在群集中的旧版 ESX 主机上，vSphere HA 通信通过被指定为服务控制台网络的所有网络进行传输。这些主机没有将 VMkernel 网络用于 vSphere HA 通信。
- 在群集中的 ESXi 主机上，默认情况下，vSphere HA 通信通过 VMkernel 网络（除了那些标记为用于 vMotion 的通信）进行传输。如果仅有一个 VMkernel 网络，如有必要，vSphere HA 与 vMotion 会共享它。使用 ESXi 4.x 及 ESXi，还必须显式启用“管理网络”复选框以允许 vSphere HA 使用此网络。

注意 VMware 建议您不要在同一子网上使用多个 vmkNIC 配置主机。如果已完成此操作，请注意，vSphere HA 会使用与给定子网关联的任意 pNIC 发送数据包（如果该子网中至少有一个 vNIC 已配置用于管理流量）。

网络隔离地址

网络隔离地址是要 ping 的 IP 地址，以确定主机是否与网络隔离。只有当主机已停止从群集内的任何其他主机接收检测信号时才 ping 此地址。如果主机可以 ping 其网络隔离地址，则说明该主机并未与网络隔离，并且群集内的其他主机已发生故障。但是，如果主机无法 ping 其隔离地址，则可能该主机已与网络隔离，并且不会执行故障切换操作。

默认情况下，网络隔离地址是主机的默认网关。无论已定义多少个管理网络，都只会指定一个默认网关。应当使用 das.isolationaddress[...] 高级属性为其他网络添加隔离地址。请参见第 25 页，“vSphere HA 高级属性”。

其他网络注意事项

在配置支持 vSphere HA 群集的网络时，应考虑其他事项。

- 配置交换机。如果连接服务器的物理网络交换机支持 PortFast（或等效）设置，请将其启用。此设置可防止主机在执行冗长的跨树算法期间误认为网络已被隔离。
- 端口组名称和网络标签。在 VLAN 上使用一致的公用网络端口组名称和网络标签。端口组名称用于重新配置虚拟机对网络的访问。如果在原始服务器和故障切换服务器间使用的名称不一致，虚拟机将在故障切换后中断网络连接。虚拟机使用这些标签在重新启动时重新建立网络连接。
- 配置管理网络，以便群集中主机上的 vSphere HA 代理可以在任何其他主机上使用其中一个管理网络访问这些代理。如果未设置此配置，则在选择首选主机后可能会出现网络分区情况。

网络路径冗余

群集节点之间的网络路径冗余对 vSphere HA 可靠性非常重要。单个管理网络会最终成为单一故障点，并且，尽管只有该网络出现故障，仍可能会导致故障切换。

如果您只有一个管理网络，则主机和群集之间的任何故障都可能导致不必要的（或错误的）故障切换活动。可能的故障包括网卡故障、网络电缆故障、网络电缆移除和交换机重置。考虑主机可能导致故障的上述原因，然后尝试减少这些问题（通常通过提供网络冗余来实现此目的）。

可以使用网卡绑定在网卡级别或在管理网络级别实现网络冗余。在大多数实现中，网卡绑定可以提供足够的冗余，但如果需要，可以使用或增加管理网络冗余。冗余管理网络可以可靠地检测故障并防止出现隔离的情况，因为检测信号可以通过多个网络发送。

在群集内的服务器之间尽量少配置硬件分段，目的是为了限制单一故障点。此外，跃点过多的路由可能会导致检测信号的网络数据包延迟，并增加潜在的故障点数目。

使用网卡绑定的网络冗余

如果用两个连接到不同物理交换机的网卡组成一个网卡组，则可以提高管理网络的可靠性。因为通过两个网卡（并且通过单独的交换机）连接的服务器具有两条独立的路径来发送和接收检测信号，所以群集具有更好的弹性。要为管理网络配置网卡组，请在活动或待机配置的 vSwitch 配置中配置 vNIC。推荐的 vNIC 参数设置如下：

- 默认的负载平衡 = 基于源虚拟端口 ID 的路由
- 故障恢复 = 否

在为 vSphere HA 群集中的一个主机添加网卡之后，必须在该主机上重新配置 vSphere HA。

使用辅助网络的网络冗余

除了使用网卡绑定提供检测信号冗余之外，还可以创建一个辅助管理网络连接，并将其连接到一个单独的虚拟交换机上。主管理网络连接用于网络和管理。辅助管理网络连接创建之后，vSphere HA 会同时通过主管理网络连接和辅助管理网络连接发送检测信号。如果一条路径发生故障，vSphere HA 仍可通过另一条路径发送和接收检测信号。

为虚拟机提供 Fault Tolerance 3

可以为虚拟机启用 vSphere Fault Tolerance，以获得比 vSphere HA 所提供的级别更高的可用性和数据保护，从而确保业务连续性。

Fault Tolerance 是基于 ESXi 主机平台构建的（使用 VMware vLockstep 技术），它通过在单独主机上以虚拟锁步方式运行相同的虚拟机来提供连续可用性。

要获取 Fault Tolerance 的最佳结果，应先熟悉其工作原理、如何为群集和虚拟机启用它及其最佳使用方法。

注意 尝试使用 Fault Tolerance 时，可能会遇到错误消息。有关与 Fault Tolerance 相关的错误消息的信息，请参见 VMware 知识库文章（网址为 http://kb.vmware.com/kb/1033634）。

本章讨论了以下主题：

- 第 32 页，“Fault Tolerance 的工作方式”
- 第 33 页，“将 Fault Tolerance 功能与 DRS 配合使用”
- 第 33 页，“Fault Tolerance 用例”
- 第 33 页，“Fault Tolerance 对照表”
- 第 35 页，“Fault Tolerance 互操作性”
- 第 36 页，“为 Fault Tolerance 准备群集和主机”
- 第 39 页，“为虚拟机提供 Fault Tolerance”
- 第 40 页，“查看有关容错虚拟机的信息”
- 第 41 页，“Fault Tolerance 的最佳做法”
- 第 43 页，“vSphere Fault Tolerance 配置建议”

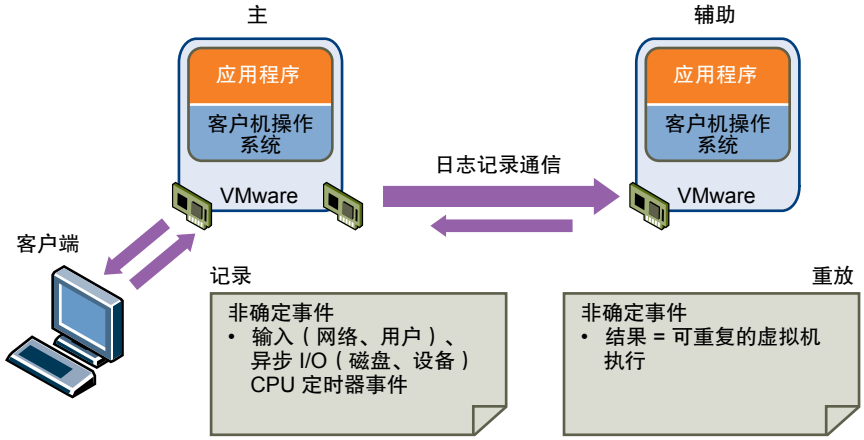
Fault Tolerance 的工作方式

vSphere Fault Tolerance 通过创建和维护与主虚拟机相同，且可在发生故障切换时随时替换主虚拟机的辅助虚拟机，来确保虚拟机的连续可用性。

可以为大多数任务关键虚拟机启用 Fault Tolerance。并会创建一个重复虚拟机（称为辅助虚拟机），该虚拟机会以虚拟锁步方式随主虚拟机一起运行。VMware vLockstep 可捕获主虚拟机上发生的输入和事件，并将这些输入和事件发送到正在另一主机上运行的辅助虚拟机。使用此信息，辅助虚拟机的执行将等同于主虚拟机的执行。因为辅助虚拟机与主虚拟机一起以虚拟锁步方式运行，所以它可以无中断地接管任何点处的执行，从而提供容错保护。

注意 主虚拟机与辅助虚拟机之间的 FT 日志记录通信是未加密的，且包含客户机网络和存储器 I/O 数据以及客户机操作系统的内存内容。此通信可以包含敏感数据，如纯文本格式的密码。为避免这些数据被泄漏，尤其是避免受到“中间人”攻击，请确保此网络是受保护的。例如，可以对 FT 日志记录通信使用专用网络。

图 3–1 Fault Tolerance 对中的主虚拟机和辅助虚拟机



主虚拟机和辅助虚拟机可持续交换检测信号。此交换使得虚拟机对中的虚拟机能够监控彼此的状态，以确保持续提供 Fault Tolerance 保护。如果运行主虚拟机的主机发生故障，系统将会执行透明故障切换，此时会立即启用辅助虚拟机以替换主虚拟机，并将启动新的辅助虚拟机，同时在几秒钟内重新建立 Fault Tolerance 冗余。如果运行辅助虚拟机的主机发生故障，则该主机也会立即被替换。在任一情况下，用户都不会遭遇服务中断和数据丢失的情况。

容错虚拟机及其辅助副本不允许在相同主机上运行。此限制可确保主机故障无法导致两个虚拟机都丢失。也可以使用虚拟机-主机关联性规则来确定要在其上运行指定虚拟机的主机。如果使用这些规则，应了解对于受这种规则影响的任何主虚拟机，其关联的辅助虚拟机也受这些规则影响。有关关联性规则的详细信息，请参见 *vSphere 资源管理*文档。

Fault Tolerance 可避免“裂脑”情况的发生，此情况可能会导致虚拟机在从故障中恢复后存在两个活动副本。共享存储器上锁定的原子文件用于协调故障切换，以便只有一端可作为主虚拟机继续运行，并由系统自动重新生成新辅助虚拟机。

注意 系统会在主虚拟机打开电源后执行反关联性检查。当主虚拟机和辅助虚拟机都处于关闭电源状态时，它们可能会位于相同主机上。但这是正常行为，辅助虚拟机会在主虚拟机打开电源时在其他主机上启动。

将 Fault Tolerance 功能与 DRS 配合使用

启用 Enhanced vMotion Compatibility (EVC) 功能时，可以将 vSphere Fault Tolerance 与 vSphere Distributed Resource Scheduler (DRS) 配合使用。此过程不但可使容错虚拟机受益于更好的初始放置位置，还可以将其纳入群集的负载平衡计算中。

当群集启用了 EVC 时，DRS 将为容错虚拟机提出初始放置位置建议、在重新平衡群集负载期间移动这些虚拟机，并允许您为主虚拟机分配 DRS 自动化级别（辅助虚拟机总是采用与其关联的主虚拟机相同的设置。）

在初始放置或负载平衡期间，DRS 放置在主机上的主虚拟机或辅助虚拟机的数目不会超过一个固定的数目。此限制由高级选项 das.maxftvmsperhost 控制。此选项的默认值为 4。但是，如果将此选项设置为 0，DRS 将忽略此限制。

将 vSphere Fault Tolerance 用于禁用了 EVC 的群集中的虚拟机时，将为容错虚拟机指定 DRS 自动化级别“已禁用”。在这种群集中，每个主虚拟机只在其注册的主机上打开电源，其辅助虚拟机被自动放置，并且不会移动主或辅助容错虚拟机以进行负载平衡。

如果将关联性规则用于一对容错虚拟机，则虚拟机-虚拟机关联性规则仅适用于主虚拟机，而虚拟机-主机关联性规则适用于主虚拟机及其辅助虚拟机。如果为主虚拟机设置了虚拟机-虚拟机关联性规则，则 DRS 会尝试解决在故障切换（即主虚拟机移至新的主机）后出现的任何冲突。

Fault Tolerance 用例

几种典型情况可以受益于 vSphere Fault Tolerance 的使用。

Fault Tolerance 可提供比 vSphere HA 更高级别的业务连续性。当调用辅助虚拟机以替换与其对应的主虚拟机时，辅助虚拟机会立即取代主虚拟机的角色，并会保存其整个状况。应用程序已在运行，并且不需要重新输入或重新加载内存中存储的数据。这不同于 vSphere HA 提供的故障切换，故障切换会重新启动受故障影响的虚拟机。

更高的连续性级别以及增加的状况信息和数据保护功能可在您要部署 Fault Tolerance 时提供方案信息。

- 需要始终保持可用的应用程序，尤其是那些具有长时间客户端连接的应用程序，用户希望在硬件故障期间保持这些连接。
- 不能通过任何其他方式实现群集功能的自定义应用程序。
- 可以通过自定义群集解决方案提供高可用性，但这些解决方案太复杂，很难进行配置和维护的情况。

按需 Fault Tolerance

用 Fault Tolerance 保护虚拟机的另一个关键用例可以描述为按需 Fault Tolerance。在这种情况中，虚拟机在正常操作期间受到 vSphere HA 的充分保护。在某些关键期间，您可能希望增强虚拟机的保护。例如，您可能正在执行季末报告，如果发生中断，则可能会延迟任务关键信息的可用性。使用 vSphere Fault Tolerance，可以在运行此报告之前保护此虚拟机，然后在生成报告之后关闭或禁用 Fault Tolerance。可以在关键时间段使用按需 Fault Tolerance 保护虚拟机，然后在非关键操作期间将资源置回正常状态。

Fault Tolerance 对照表

以下对照表包含在使用 vSphere Fault Tolerance 之前需要了解的群集、主机和虚拟机要求。

在设置 Fault Tolerance 之前，应查看此列表。您也可以使用 VMware SiteSurvey 实用程序（从 http://www.vmware.com/download/shared_utilities.html 下载），更好地了解与要用于 vSphere FT 的群集、主机和虚拟机相关的配置问题。

注意 容错虚拟机的故障切换与 vCenter Server 无关，但必须使用 vCenter Server 来设置 Fault Tolerance 群集。

Fault Tolerance 的群集要求

在使用 Fault Tolerance 之前，必须满足以下群集要求。

- 主机证书检查功能已启用。请参见第 36 页，“启用主机证书检查”。
- 至少有两台通过 FT 认证的主机运行相同的 Fault Tolerance 版本号或主机内部版本号。Fault Tolerance 版本号显示在 vSphere Client 中主机的**摘要**选项卡上。

  注意 对于 ESX/ESXi 4.1 之前的旧版主机，此选项卡列出的是主机内部版本号。修补程序可能会导致 ESX 和 ESXi 安装之间的主机内部版本号不同。为了确保旧版主机与 FT 兼容，请不要在 FT 对中混合使用旧版 ESX 和 ESXi 主机。

- ESXi 主机可以访问相同的虚拟机数据存储和网络。请参见第 41 页，“Fault Tolerance 的最佳做法”。
- 配置了 Fault Tolerance 日志记录和 vMotion 网络。请参见第 36 页，“为主机配置网络”。
- vSphere HA 群集已创建并启用。请参见第 21 页，“创建 vSphere HA 群集”。打开容错虚拟机电源或者将主机添加到已支持容错虚拟机的群集之前，必须启用 vSphere HA。

Fault Tolerance 的主机要求

在使用 Fault Tolerance 之前，必须满足以下主机要求。

- 主机上的处理器必须来自与 FT 兼容的处理器组。此外，强烈建议主机的处理器之间也相互兼容。有关受支持处理器的信息，请参见 VMware 知识库文章，网址为 http://kb.vmware.com/kb/1008027。
- 主机必须获得 Fault Tolerance 的许可。
- 主机必须已通过 Fault Tolerance 认证。请参见 http://www.vmware.com/resources/compatibility/search.php 并选择**按与容错兼容的集合搜索**，确定主机是否已通过认证。
- 在配置每台主机时，都必须在 BIOS 中启用硬件虚拟化 (HV)。

要确认群集内的主机是否兼容，从而判断其是否支持 Fault Tolerance，还可以按 第 38 页，“创建 vSphere HA 群集和检查合规性” 中所说明的那样运行配置文件合规性检查。

注意 当主机无法支持 Fault Tolerance 时，可在 vSphere Client 中主机的**摘要**选项卡上查看原因。单击**已配置 FT 的主机**字段旁边的蓝色标题图标，查看主机不符合的 Fault Tolerance 要求的列表。

Fault Tolerance 的虚拟机要求

在使用 Fault Tolerance 之前，必须满足以下虚拟机要求。

- 没有不受支持的设备连接到虚拟机。请参见第 35 页，“Fault Tolerance 互操作性”。
- 虚拟机必须存储在虚拟 RDM 或厚置备的虚拟机磁盘 (VMDK) 文件中。如果虚拟机存储在精简置备的 VMDK 文件中，则在尝试启用 Fault Tolerance 时将会出现一则指示必须转换 VMDK 文件的消息。要执行该转换，必须关闭虚拟机电源。
- 不兼容的功能一定不能与容错虚拟机一起运行。请参见第 35 页，“Fault Tolerance 互操作性”。
- 虚拟机文件必须存储在共享存储器上。可接受共享的存储解决方案包括光纤通道、（硬件和软件）iSCSI、NFS 和 NAS。
- 只有带有单个 vCPU 的虚拟机才能与 Fault Tolerance 功能兼容。
- 虚拟机必须在一个受支持的客户机操作系统上运行。有关详细信息，请参见 VMware 知识库文章，网址为 http://kb.vmware.com/kb/1008027。

Fault Tolerance 互操作性

在配置 vSphere Fault Tolerance 之前，应当了解 Fault Tolerance 不能与之交互操作的功能和产品。

Fault Tolerance 不支持的 vSphere 功能

容错虚拟机不支持以下 vSphere 功能。

- 快照。在虚拟机上启用 Fault Tolerance 前，必须移除或提交快照。此外，不可能对已启用 Fault Tolerance 的虚拟机执行快照。
- Storage vMotion。不能为已启用 Fault Tolerance 的虚拟机调用 Storage vMotion。要迁移存储器，应当先暂时关闭 Fault Tolerance，然后再执行 Storage vMotion 操作。在完成迁移之后，可以重新打开 Fault Tolerance。
- 链接克隆。不能在为链接克隆的虚拟机上启用 Fault Tolerance，也不能从启用了 Fault Tolerance 的虚拟机创建链接克隆。
- 虚拟机备份。不能使用 Storage API for Data Protection、VMware Data Recovery 或需要使用虚拟机快照（如 ESXi 所执行的那样）的类似备份产品来备份启用了 FT 的虚拟机。要以这种方式备份容错虚拟机，首先必须禁用 FT，然后在执行备份后重新启用 FT。基于存储阵列的快照不影响 FT。

不与 Fault Tolerance 兼容的功能和设备

要使虚拟机与 Fault Tolerance 功能兼容，虚拟机不能使用以下功能或设备。

表 3–1 不与 Fault Tolerance 兼容的功能和设备以及纠正操作

不兼容的功能或设备	纠正操作
对称多处理器 (SMP) 虚拟机。只有带有单个 vCPU 的虚拟机才能与 Fault Tolerance 功能兼容。	将虚拟机重新配置为单个 vCPU。配置为单 vCPU 后，许多工作负载均有较好的性能表现。
物理裸磁盘映射 (RDM)。	请重新配置具有支持物理 RDM 的虚拟设备的虚拟机以使用虚拟 RDM。
由物理或远程设备支持的 CD-ROM 或虚拟软盘设备。	移除 CD-ROM 或虚拟软盘设备，或使用共享存储器上安装的 ISO 重新配置备用功能。
准虚拟化客户机。	如果准虚拟化不是必需的，请重新配置不带 VMI ROM 的虚拟机。
USB 和声音设备。	从虚拟机移除这些设备。
N_Port ID 虚拟化 (NPIV)。	禁用虚拟机的 NPIV 配置。
网卡直通。	Fault Tolerance 不支持此功能，因此必须将其关闭。
vlance 网络驱动程序。	Fault Tolerance 不支持配置了 vlance 虚拟网卡的虚拟机。但是，vmxnet2、vmxnet3 和 e1000 完全受支持。
使用精简置备的存储器或未启用群集功能的厚置备磁盘备份的虚拟磁盘。	当您打开 Fault Tolerance 时，系统会默认执行到适当磁盘格式的转换。必须先关闭虚拟机电源，才能触发此转换。
热插拔设备。	容错虚拟机的热插拔功能将自动禁用。要热插拔设备（添加或移除），必须临时关闭 Fault Tolerance，完成热插拔操作，然后重新启用 Fault Tolerance。 **注意** 使用 Fault Tolerance 时，如果在虚拟机正在运行过程中更改虚拟网卡的设置，该操作即为热插拔操作，因为它要求先拔出网卡，然后重新插入。例如，当正在运行的虚拟机使用虚拟网卡时，如果更改虚拟网卡所连接到的网络，必须首先关闭 FT。
扩展页表/快速虚拟化索引 (EPT/RVI)。	EPT/RVI 在启用了 Fault Tolerance 的虚拟机上将自动禁用。
串行或并行端口	从虚拟机移除这些设备。

表 3–1 不与 Fault Tolerance 兼容的功能和设备以及纠正操作（续）

不兼容的功能或设备	纠正操作
IPv6	将 IPv4 地址与 FT 日志记录网卡配合使用。
启用了 3D 的视频设备。	Fault Tolerance 不支持启用了 3D 的视频设备。

为 Fault Tolerance 准备群集和主机

要为群集启用 vSphere Fault Tolerance，必须满足此功能的必备条件，然后在主机上执行特定的配置步骤。完成这些步骤并创建群集后，还可以检查配置是否符合启用 Fault Tolerance 的要求。

尝试为群集启用 Fault Tolerance 之前，应当完成的任务包括：

- 启用主机证书检查（如果您要从先前版本的 vCenter Server 升级）。
- 为每台主机配置网络。
- 创建 vSphere HA 群集，添加主机，并检查合规性。

在为群集和主机准备好 Fault Tolerance 之后，便可为虚拟机打开 Fault Tolerance。请参见第 40 页，“为虚拟机打开 Fault Tolerance”。

启用主机证书检查

使用主机证书检查，可以将 ESXi 主机配置为相互验证身份，确保维护一个更安全的环境。对于容错虚拟机所在的 ESXi 主机来说，主机证书检查是必需的。

如果安装了 VMware vCenter Server 4.1 版本，系统会自动启用主机证书检查。如果所用版本是从上一版本升级的版本，则必须手动执行此过程。在此过程中，您会看到主机及其待验证的证书的列表。您可以在启用证书检查之前验证主机证书。必须以手动方式对未在此步骤中验证的主机进行验证和重新连接。

步骤

1 将 vSphere Client 连接到 vCenter Server。

2 选择**系统管理**，然后选择 **vCenter Server 设置**。

此时会出现 **vCenter Server 设置**窗口。

3 在左窗格中单击 **SSL 设置**。

4 选中 **vCenter 需要已验证的主机 SSL 证书**框。

5 单击**确定**。

为主机配置网络

在要添加到 vSphere HA 群集的每台主机上，必须配置两个不同的网络交换机，以便让主机也能支持 vSphere Fault Tolerance。

要为主机启用 Fault Tolerance，您必须完成此步骤两次，即为每个端口组选项各进行一次，以确保有足够的带宽可供 Fault Tolerance 日志记录使用。选择一个选项，完成该过程，然后选择另一个端口组选项，再执行一次该过程。

前提条件

需要多个千兆位网络接口卡 (NIC)。对于支持 Fault Tolerance 功能的每台主机，最少需要两个物理千兆位网卡。例如，您需要一个网卡专门用于 Fault Tolerance 日志记录，另一个则专门用于 vMotion。VMware 建议具备三个或更多网卡来确保可用性。

注意 vMotion 和 FT 日志记录网卡必须位于不同的子网上，并且 FT 日志记录网卡不支持 IPv6。

步骤

1 将 vSphere Client 连接到 vCenter Server。

2 在 vCenter Server 清单中，选择主机，然后单击**配置**选项卡。

3 选择**硬件**下的**网络**，然后单击**添加网络**链接。

此时会出现添加网络向导。

4 选择**连接类型**下的 **VMkernel**，然后单击**下一步**。

5 选择**创建虚拟交换机**，然后单击**下一步**。

6 提供交换机的标签。

7 选择**将此端口组用于 vMotion** 或**将此端口组用于 Fault Tolerance 日志记录**，然后单击**下一步**。

8 提供 IP 地址和子网掩码，然后单击**下一步**。

9 单击**完成**。

在创建 vMotion 和 Fault Tolerance 日志记录虚拟交换机后，可以根据需要创建其他虚拟交换机。然后，应该将主机添加到群集，并完成打开 Fault Tolerance 所需的步骤。

下一步

要确认是否在主机上成功启用了 vMotion 和 Fault Tolerance，请在 vSphere Client 中查看该主机的**摘要**选项卡。在“常规”窗格中，**vMotion 已启用**和**已配置 FT 的主机**字段均应显示为“是”。

注意 如果将网络配置为可支持 FT，但随后又禁用了 Fault Tolerance 日志记录端口，则已打开电源的容错虚拟机对将保持打开电源状态。但是，如果出现了故障切换情况，那么，当主虚拟机被其辅助虚拟机替换时，将不会启动新的辅助虚拟机，这会导致新的主虚拟机以“不受保护”状态运行。

主机网络 Fault Tolerance 配置示例

此示例描述了在具有四个千兆位网卡的典型部署中主机网络的 Fault Tolerance 配置。此配置是一种可能的部署，它确保有适当的服务用于示例中标识的每种流量类型。此配置可被认为是最佳做法配置。

在物理主机由于断电、系统应急或类似原因出现故障的过程中，Fault Tolerance 提供充足的正常运行时间。然而，网络或存储路径故障或者不影响主机的正在运行状态的任何其他物理服务器组件可能不会启动 Fault Tolerance 故障切换到辅助虚拟机。因此，强烈鼓励客户使用适当的冗余（如网卡绑定）来减少虚拟机与基础架构组件（如网络或存储阵列）的连接断开的机会。

网卡绑定策略在 vSwitch (vSS) 端口组（或 vDS 的分布式虚拟端口组）上配置，它们用于控制 vSwitch 如何处理和分布流量到虚拟机的物理网卡 (vmnic) 和 vmkernel 端口。通常，对每种流量类型使用一个唯一的端口组，并且通常将每种流量类型分配给一个不同的 VLAN。

主机网络配置准则

您可以按照以下准则配置主机的网络，以便在不同流量类型的组合（如 NFS）和不同数目的物理网卡的情况下支持 Fault Tolerance。

- 将每个网卡绑定分布到两台物理交换机，并确保这两台物理交换机之间的每个 VLAN 的 L2 域连续性。
- 使用确定的绑定策略确保特定流量类型与特定网卡（活动/待机）或网卡集（如源虚拟端口 ID）具有关联性。
- 使用活动/待机策略时，将流量类型配对，以便使两种流量类型共享某个 vmnic 的情况下发生故障切换时所产生的影响最小。
- 使用活动/待机策略时，配置所有活动适配器，以便特定流量类型（如 FT 日志记录）流向相同物理交换机。这样可使网络跃点的数目最少，并降低超额预订交换机到交换机链路的可能性。

4 Gb 网卡的配置示例

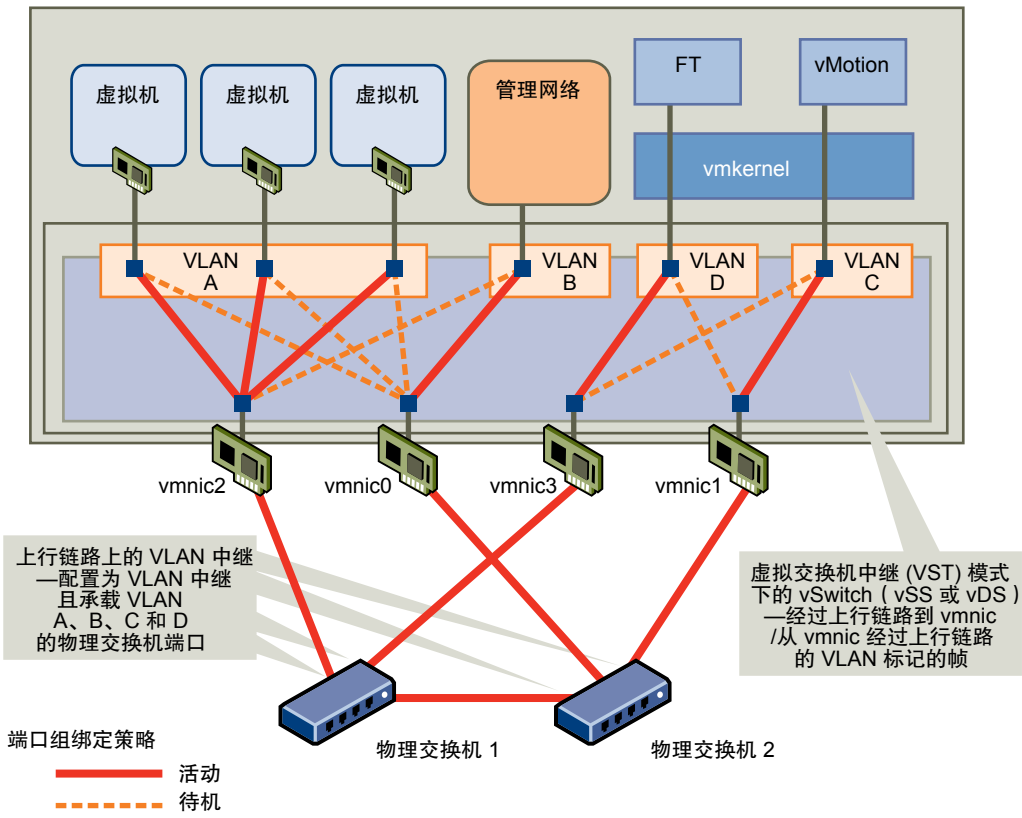
图 3–2 描述了具有四个千兆位网卡以支持 Fault Tolerance 的一台 ESXi 主机的网络配置。FT 群集中的其他主机可类似配置。

此示例使用四个端口组，其配置如下所示：

- VLAN A：虚拟机网络端口组 - 在 vmnic2（连接到物理交换机 #1）上处于活动状态；在 vmnic0（连接到物理交换机 #2）上处于待机状态。
- VLAN B：管理网络端口组 - 在 vmnic0（连接到物理交换机 #2）上处于活动状态；在 vmnic2（连接到物理交换机 #1）上处于待机状态。
- VLAN C:vMotion 端口组 - 在 vmnic1（连接到物理交换机 #2）上处于活动状态；在 vmnic3（连接到物理交换机 #1）上处于待机状态。
- VLAN D：FT 日志记录端口组 - 在 vmnic3（连接到物理交换机 #1）上处于活动状态；在 vmnic1（连接到物理交换机 #2）上处于待机状态。

vMotion 和 FT 日志记录可以共享相同的 VLAN（在这两个端口组中配置相同的 VLAN 号），但要求它们在不同 IP 子网中具有唯一的 IP 地址。但是，如果服务质量 (QoS) 限制在使用基于 VLAN 的 QoS 的物理网络上有效，则使用不同的 VLAN 可能更好。在竞争流量出现的情况下（例如，使用了多个物理交换机跃点，或者当发生故障切换时，多种流量类型争用网络资源），QoS 特别有用。

图 3–2 网络 Fault Tolerance 配置示例



创建 vSphere HA 群集和检查合规性

vSphere Fault Tolerance 在 vSphere HA 群集环境中使用。在每台主机上配置网络之后，创建 vSphere HA 群集并向其中添加主机。可以查看群集配置是否正确以及是否符合成功启用 Fault Tolerance 的要求。

步骤

1 将 vSphere Client 连接到 vCenter Server。

2 在 vCenter Server 清单中，选择群集，然后单击**配置文件合规情况**选项卡。

3 单击**立即检查合规性**运行合规性测试。

要查看运行的测试，请单击**描述**。

合规性测试的结果将显示在屏幕的底部。主机将标记为“合规”或“不合规”。

注意 有关如何创建 vSphere HA 群集的详细论述，请参见第 11 页，第 2 章“创建和使用 vSphere HA 群集”。

为虚拟机提供 Fault Tolerance

在采取了为群集启用 vSphere Fault Tolerance 所需的全部步骤之后，可以为各个虚拟机打开 Fault Tolerance 功能。

如果符合下列任一情况，则用于打开 Fault Tolerance 的选项将不可用并变成灰色：

- 虚拟机所驻留的主机并未获得使用该功能的许可证。
- 虚拟机所驻留的主机处于维护模式或待机模式。
- 虚拟机已断开连接或被孤立（无法访问其 .vmx 文件）。
- 用户没有打开此功能的权限。

如果用于打开 Fault Tolerance 的选项可用，则此任务仍然必须进行验证，并且在未满足某些要求时可能会失败。

打开 Fault Tolerance 时的验证检查

在打开 Fault Tolerance 之前，需要在虚拟机上执行多项验证检查。

- 必须在 vCenter Server 设置中启用 SSL 证书检查。
- 主机必须位于 vSphere HA 群集或包含 vSphere HA 和 DRS 的混合群集内。
- 主机必须安装了 ESX/ESXi 4.0 或更高版本。
- 虚拟机不得有多个 vCPU。
- 虚拟机不得有快照。
- 虚拟机不得是模板。
- 对于虚拟机不得禁用 vSphere HA。
- 虚拟机不得有启用了 3D 的视频设备。

已对已打开电源的虚拟机（或正在打开电源的虚拟机）执行了多项附加验证检查。

- 容错虚拟机所驻留的主机的 BIOS 必须启用了硬件虚拟化 (HV)。
- 支持主虚拟机的主机必须有支持 Fault Tolerance 的处理器。
- 支持辅助虚拟机的主机必须有支持 Fault Tolerance 的处理器，并且该处理器的 CPU 系列和型号与支持主虚拟机的主机相同。
- 您的硬件应认证为与 Fault Tolerance 兼容。为了确认这点，请使用 http://www.vmware.com/resources/compatibility/search.php 中的《VMware 兼容性指南》并选择**按与容错兼容的集合搜索**。
- 虚拟机的客户机操作系统和处理器的组合必须受 Fault Tolerance 支持（例如，基于 AMD 的处理器和 32 位 Solaris 系统的组合当前不受支持）。有关支持的处理器和客户机操作系统组合的信息，请参见 VMware 知识库文章（网址为 http://kb.vmware.com/kb/1008027）。

- 虚拟机的配置必须有效，以便与 Fault Tolerance 功能配合使用（例如，不得包含任何不受支持的设备）。

当为虚拟机打开 Fault Tolerance 这一操作通过验证检查时，将创建辅助虚拟机。辅助虚拟机的放置位置和即时状态取决于在打开 Fault Tolerance 时主虚拟机是已打开电源还是已关闭电源。

如果主虚拟机已打开电源：

- 将复制整个主虚拟机的状况，创建辅助虚拟机，并将其放置在单独的兼容主机上，而且会在通过接入控制时打开电源。
- vSphere Client 中虚拟机的**摘要**选项卡上显示的 Fault Tolerance 状态将为**受保护**。

如果主虚拟机已关闭电源：

- 将立即创建辅助虚拟机并在群集的主机中注册（打开该虚拟机电源时，可能会在更合适的主机上重新进行注册）。
- 辅助虚拟机在主虚拟机打开电源之后打开电源。
- vSphere Client 中虚拟机的**摘要**选项卡上显示的 Fault Tolerance 状态将为**不受保护、虚拟机未运行**。
- 当尝试在打开 Fault Tolerance 之后打开主虚拟机的电源时，将执行上面列出的附加验证检查。要正确打开电源，虚拟机不得使用准虚拟化 (VMI)。

  通过这些检查之后，将打开主虚拟机和辅助虚拟机的电源，并将其分别放置在单独的兼容主机上。vSphere Client 中虚拟机的**摘要**选项卡上显示的 Fault Tolerance 状态将标记为**受保护**。

为虚拟机打开 Fault Tolerance

可以通过 vSphere Client 打开 vSphere Fault Tolerance。

在打开 Fault Tolerance 功能后，vCenter Server 会取消虚拟机的内存限制设置，并将内存预留值设置为虚拟机的内存大小。当 Fault Tolerance 保持打开状态时，不能更改内存预留、大小、限制或份额。在关闭 Fault Tolerance 功能后，已更改的任何参数均不会恢复到其原始值。

使用具有群集管理员权限的帐户将 vSphere Client 连接到 vCenter Server。

步骤

1 选择“主机和群集”视图。

2 右键单击一个虚拟机并选择 **Fault Tolerance > 打开 Fault Tolerance**。

  如果选择多个虚拟机，则 **Fault Tolerance** 菜单处于禁用状态。一次只能为一个虚拟机打开 Fault Tolerance。

特定的虚拟机将指定为主虚拟机，并在另一台主机上建立辅助虚拟机。现在，主虚拟机已启用了容错功能。

查看有关容错虚拟机的信息

可以使用 vSphere Client 查看 vCenter Server 清单中的容错虚拟机。

注意 不能从辅助虚拟机禁用 Fault Tolerance。

主虚拟机的**摘要**选项卡中显示了 vSphere Fault Tolerance 区域（窗格），其中包含有关虚拟机的信息。

Fault Tolerance 状态 表示虚拟机的 Fault Tolerance 状态。

- 受保护。表示主虚拟机和辅助虚拟机均已打开电源并正在按预期运行。

- 不受保护。表示辅助虚拟机未在运行。下表中列出了可能原因。

表 3–2 主虚拟机处于不受保护状态的原因

处于不受保护状态的原因	描述
正在启动	Fault Tolerance 功能正在启动辅助虚拟机。此消息仅在短时间段内可见。
需要辅助虚拟机	主虚拟机正在运行而没有辅助虚拟机，因此主虚拟机当前不受保护。此情况通常在群集内无兼容主机可用于辅助虚拟机时发生。使某台兼容主机联机可解决此问题。如果群集内有联机的兼容主机，则可能需要进一步调查。在某些情况下，禁用 Fault Tolerance 然后重新将其启用可解决此问题。
已禁用	Fault Tolerance 功能当前处于禁用状态（没有辅助虚拟机正在运行）。此情况在用户禁用了 Fault Tolerance 或当 vCenter Server 在无法打开辅助虚拟机电源之后禁用了 Fault Tolerance 时发生。
虚拟机未运行	Fault Tolerance 功能已启用，但虚拟机已关闭电源。打开虚拟机电源，使其达到受保护状态。

辅助虚拟机位置 显示辅助虚拟机所在的 ESXi 主机。

辅助虚拟机的总 CPU 表示辅助虚拟机的 CPU 使用情况，以 MHz 为单位。

辅助虚拟机的总内存 表示辅助虚拟机的内存使用情况，以 MB 为单位。

vLockstep 时间间隔 辅助虚拟机匹配主虚拟机的当前执行状况所需的时间间隔（以秒为单位）。通常，此时间间隔小于半秒。无论 vLockstep 时间间隔值为多大，在故障切换过程中均不会丢失任何状况。

记录带宽 用于将 vSphere Fault Tolerance 日志信息从运行主虚拟机的主机发送到运行辅助虚拟机的主机的网络容量。

对于配置为支持 Fault Tolerance 的每台主机，都可以通过访问 vSphere Client 中主机的**摘要**选项卡来查看有关其容错虚拟机的信息。该屏幕的 **Fault Tolerance** 部分显示了主机上驻留的主虚拟机和辅助虚拟机的总数，以及那些已打开电源虚拟机的数量。如果主机为 ESX/ESXi 4.1 或更高版本，则此部分还将显示主机正运行的 Fault Tolerance 版本。否则，将列出主机内部版本号。如果需要两台主机相互兼容，则必须具有匹配的 FT 版本号或匹配的主机内部版本号。

Fault Tolerance 的最佳做法

为确保获得最佳 Fault Tolerance 结果，VMware 建议您遵循某些最佳做法。

除了以下信息外，您还可以参见白皮书《VMware Fault Tolerance 建议和注意事项》，网址为 http://www.vmware.com/resources/techresources/10040。

主机配置

配置主机时请考虑以下最佳做法。

- 运行主虚拟机和辅助虚拟机的主机应当按照与处理器大致相同的频率运行，否则辅助虚拟机可能会更频繁地重新启动。不依据工作负载进行调整（例如，为省电而执行功率封顶和强制低频率模式）的平台电源管理功能可能会导致处理器频率大范围浮动。如果辅助虚拟机要定期重新启动，请在运行容错虚拟机的主机上禁用所有的电源管理模式，或者确保所有主机以相同电源管理模式运行。

- 对所有主机应用相同的指令集扩展配置（已启用或已禁用）。启用或禁用指令集的过程因 BIOS 而异。有关如何配置指令集，请参见主机的 BIOS 文档。

同类群集

vSphere Fault Tolerance 可以在主机不一致的群集内使用，但在节点兼容的群集内才能起到最好的效果。构建群集时，所有主机都应具有以下配置：

- 来自同一兼容处理器组的处理器。
- 对虚拟机所用数据存储的一般访问权限。
- 相同的虚拟机网络配置。
- 相同的 ESXi 版本。
- 相同的 Fault Tolerance 版本号（或早于 ESX/ESXi 4.1 的主机的内部版本号）。
- 所有主机的相同 BIOS 设置（电源管理和超线程）。

运行**检查合规性**确定不兼容性并进行更正。

性能

要为主虚拟机和辅助虚拟机之间的日志记录通信增加可用带宽，请使用 10 千兆位网卡，并允许使用巨帧。

在共享存储器上存储 ISO 以连续进行访问

将启用了 Fault Tolerance 的虚拟机访问的 ISO 存储在可以供容错虚拟机的两个实例访问的共享存储器上。如果使用此配置，虚拟机中的 CD-ROM 会继续正常工作，即使发生了故障切换也是如此。

对于启用了 Fault Tolerance 的虚拟机，您可能会使用只有主虚拟机才可访问的 ISO 映像。在这种情况下，主虚拟机可以访问 ISO，但是如果出现故障切换，CD-ROM 会报告错误，就好像介质不存在一样。如果 CD-ROM 只是用于临时的非关键性操作（如安装），则此情况是可以接受的。

避免网络分区

当 vSphere HA 群集出现管理网络故障而导致某些主机与 vCenter Server 隔离并且使这些主机彼此隔离时，会出现网络分区。请参见第 13 页，“网络分区”。如果出现分区，则可能降低 Fault Tolerance 保护。

在使用 Fault Tolerance 的已分区 vSphere HA 群集中，主虚拟机（或其辅助虚拟机）可以在由首选主机（不负责管理虚拟机）管理的分区中停止。如果需要故障切换，则仅当主虚拟机位于首选主机（负责管理该主虚拟机）管理的分区中时才重新启动辅助虚拟机。

为确保管理网络尽可能不出现导致网络分区的故障，请遵循第 28 页，“网络连接的最佳做法”中的建议。

升级用于 Fault Tolerance 的主机

当升级包含容错虚拟机的主机时，必须确保主虚拟机和辅助虚拟机继续在具有相同 FT 版本号或主机内部版本号的主机（对于 ESX/ESXi 4.1 之前的主机）上运行。

前提条件

确认您具有群集管理员特权。

确认拥有多组 ESXi 主机，每组由四台或多台主机组成，这些主机托管已打开电源的容错虚拟机。如果虚拟机已关闭电源，则主虚拟机和辅助虚拟机可以重定位到具有不同内部版本的主机。

注意 此升级过程适用于至少包含四个节点的群集。更小的群集也可以遵循此说明，不过不受保护的时间间隔将稍微长一些。

步骤

1 使用 vMotion 从两台主机中迁出容错虚拟机。

2 将这两台已撤出的主机升级到相同的 ESXi 内部版本。

3 在主虚拟机上关闭 Fault Tolerance 功能。

4 使用 vMotion 将已禁用的主虚拟机移动到其中一台已升级的主机上。

5 针对已移动的主虚拟机打开 Fault Tolerance 功能。

6 要在升级的主机上容纳尽可能多的容错虚拟机对，请重复步骤 1 到步骤 5。

7 使用 vMotion 重新分配容错虚拟机。

群集中的所有 ESXi 主机即已升级。

vSphere Fault Tolerance 配置建议

VMware 建议您在配置 Fault Tolerance 时遵循特定准则。

- 除了非容错虚拟机以外，任何一台主机上不得有超过四个的容错虚拟机（主虚拟机或辅助虚拟机）。每台主机上可以安全运行的容错虚拟机数目依据 ESXi 主机和虚拟机的大小以及工作负载确定，且可能会发生变化。
- 如果要使用 NFS 访问共享存储器，请使用至少具有 1 千兆位网卡的专用 NAS 硬件，以获取为了使 Fault Tolerance 功能正常工作所需的网络性能。
- 确保包含容错虚拟机的资源池内存大于虚拟机内存。在打开 Fault Tolerance 功能后，Fault Tolerance 虚拟机的保留内存设置为虚拟机的内存大小。如果资源池中没有额外内存，则可能没有内存可用作开销内存。
- 每个容错虚拟机最多使用 16 个虚拟磁盘。
- 为确保冗余和最大 Fault Tolerance 保护，群集中应至少有三台主机。如果发生故障切换情况，这可确保有主机可容纳所创建的新辅助虚拟机。

索引

A

按需 Fault Tolerance **33**
Auto Deploy **27**

B

必备条件, Fault Tolerance **33**

C

插槽 **15**
插槽大小计算 **15**
创建 vSphere HA 群集 **21**
存储 DRS **27**
存储器
  iSCSI **33**
  NAS **33, 43**
  NFS **33, 43**
错误消息
  Fault Tolerance **31**
  vSphere HA **11**

D

打开 vSphere HA **22**
当前故障切换容量 **15, 17**
当前故障切换主机 **19**
das.heartbeatdsperhost **13, 25**
das.ignoreinsufficienthbdatastore **25**
das.iostatsinterval **24, 25**
das.isolationaddress **25, 28**
das.isolationshutdowntimeout **23, 25**
das.maxftvmsperhost **33**
das.slotcpuinmhz **15, 25**
das.slotmeminmb **15, 25**
das.usedefaultisolationaddress **25**
das.vmcpuminmhz **15, 17, 25**
das.vmmemoryminmb **25**
Distributed Power Management (DPM) **14, 15**
Distributed Resource Scheduler (DRS)
  打开 **22**
  和 Fault Tolerance **35**
  与 vSphere Fault Tolerance 配合使用 **33**
  与 vSphere HA 结合使用 **14**
DNS 查询 **20**
端口组名称 **28**
对称多处理器 (SMP) **35**

E

EVC **33**

F

防火墙端口 **13, 28**
反关联性规则 **32**
Fault Tolerance
  版本 **33**
  必备条件 **33**
  错误消息 **31**
  打开 **40**
  打开限制 **39**
  对照表 **33**
  反关联性规则 **32**
  辅助虚拟机的总 CPU **40**
  辅助虚拟机的总内存 **40**
  辅助虚拟机位置 **40**
  概览 **32**
  合规性检查 **38**
  互操作性 **35**
  记录带宽 **40**
  连续可用性 **9**
  配置建议 **43**
  启用 **36**
  日志记录 **36, 37**
  vLockstep 时间间隔 **40**
  vSphere 配置 **33**
  网络配置 **36, 37**
  验证检查 **39**
  用例 **33**
  最佳做法 **41**
Fault Tolerance 状态
  启动 **40**
  虚拟机未运行 **40**
  需要辅助虚拟机 **40**
  已禁用 **40**
非计划停机时间 **8**
FT 日志记录 **32**
负载平衡 **33**

G

高级属性, vSphere HA **25**
高级运行时信息 **15**
关联性规则 **32, 33**

管理网络 **20, 28**
挂起 vSphere HA **22**
规划 vSphere HA 群集 **11**
故障切换主机 **19**

H
合规性检查, Fault Tolerance **38**
互操作性, Fault Tolerance **35**

I
I/O 统计间隔 **24**
IPv4 **20, 35**
IPv6 **20, 35, 36**
iSCSI SAN **33**
ISO 映像 **41**

J
监控敏感度 **24**
监控 vSphere HA **27**
接入控制
  策略 **23**
  类型 **15**
  启用 **23**
  vSphere HA **15**
接入控制策略
  群集允许的主机故障数目 **15**
  选择 **20**
  预留的群集资源的百分比 **17**
  指定故障切换主机 **19**
计划的停机时间 **7**

K
快速虚拟化索引 (RVI) **35**
快照 **35**
扩展页表 (EPT) **35**

M
每个虚拟机的最大重置次数 **24**
默认网关 **28**
目标读者 **5**

N
N_Port ID 虚拟化 (NPIV) **35**

P
配置的故障切换容量 **15, 17**
配置 vSphere HA 高级选项 **25**
PortFast **28**

Q
群集操作状态 **27**
群集设置 **21**
群集有效性 **27**
群集允许的主机故障数目 **15, 27**

R
RDM **33, 35**
日志文件 **13**

S
升级包含 FT 虚拟机的主机 **42**
事件和警报, 设置 **27**
首选主机选举 **11**
数据存储检测信号 **12, 13**
SSL 证书 **13**
Storage vMotion **7, 27, 35**

T
TCP 端口 **13**
停机时间
  非计划 **8**
  计划的 **7**
透明故障切换 **9, 32**

U
UDP 端口 **13**

V
VLAN **37**
VMDK **33**
VMFS **13, 28**
VMware Tools **24**
VMware vLockstep **9, 31, 32**
vpxuser 用户帐户 **13**
vSphere HA
  从中断中恢复 **8**
  错误消息 **11**
  打开 **22**
  对照表 **20**
  高级属性 **25**
  挂起 **22**
  监控 **27**
  群集设置 **21**
  优点 **8**
  自定义 **25**
vSphere HA 架构 **11**
vSphere HA 群集
  创建 **21, 38**
  从属主机 **11, 12**
  规划 **11**
  接入控制 **15**
  首选主机 **11–13**
  异构性 **20**
  最佳做法 **27**

vSphere HA 网络连接
  路径冗余 **29**
  最佳做法 **28**

W

网卡绑定 **29, 37**
网络标签 **28**
网络分区 **12, 13, 41**
网络隔离地址 **28**
网络配置, Fault Tolerance **36, 37**

X

修改群集设置 **21**
虚拟机-虚拟机关联性规则 **19**
虚拟机保护 **11, 13**
虚拟机重新启动优先级设置 **23**
虚拟机监控 **11, 24**
虚拟机启动和关机功能 **21**
虚拟机替代项 **23, 27**

Y

验证检查 **39**
业务连续性 **7**
硬件虚拟化 (HV) **33, 39**
应用程序监控 **11, 24**
用例, Fault Tolerance **33**
预留的群集资源的百分比 **17, 27**
允许的主机故障数目 **15**

Z

增强型 vMotion 兼容性 **33**
指定故障切换主机 **19**
主机
  网络隔离 **12**
  维护模式 **12, 14**
主机隔离响应设置 **23**
主机监控功能 **22, 28**
主机证书检查 **33, 36**
准虚拟化 **35**
自定义 vSphere HA **25**
资源碎片 **20**
最佳做法
  Fault Tolerance **41**
  vSphere HA 群集 **27**
  vSphere HA 网络连接 **28**
最小化停机时间 **7**