

# 光纤通道 SAN 配置指南

ESX Server 3.5、ESX Server 3i 版本 3.5

VirtualCenter 2.5



光纤通道 SAN 配置指南

修订时间：20080410

项目：VI-CHS-Q208-535

我们的网站将提供最新技术文档，网址为：

<http://www.vmware.com/cn/support/>

此外，VMware 网站还提供最新的产品更新。

如果对本文档有任何意见或建议，请将反馈信息提交至以下地址：

[docfeedback@vmware.com](mailto:docfeedback@vmware.com)

© 2008 VMware, Inc. 保留所有权利。受若干项美国专利保护，专利号是 6,397,242、6,496,847、6,704,925、6,711,672、6,725,289、6,735,601、6,785,886、6,789,156、6,795,966、6,880,022、6,944,699、6,961,806、6,961,941、7,069,413、7,082,598、7,089,377、7,111,086、7,111,145、7,117,481、7,149,843、7,155,558 和 7,222,221；以及多项正在申请的专利。

VMware、VMware “箱状”徽标及设计、虚拟 SMP 和 VMotion 都是 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](http://www.vmware.com)

**VMware Global, Inc.**

北京办公室 北京市东城区长安街一号东方广场 W2 办公楼 6 层 601 室  
邮编：100738 电话：+86-10-8520-0148  
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  
邮编：200120 电话：+86-21-6160-1168  
广州办公室 广州市天河北路 233 号中信广场 7401 室  
邮编：510613 电话：+86-20-3877-1938  
<http://www.vmware.com/cn>

# 目录

关于本书	9
<b>1 VMware ESX Server 概述</b>	<b>13</b>
ESX Server 简介	14
系统组件	14
软件和硬件兼容性	15
了解虚拟化	15
CPU、内存及网络虚拟化	16
虚拟 SCSI	16
磁盘配置选项	17
虚拟机文件系统	18
裸设备映射	18
虚拟 SCSI 主机总线适配器	18
与 ESX Server 系统交互	19
VMware Virtual Center	19
ESX Server 3 服务控制台	20
虚拟化概览	21
<b>2 将 ESX Server 与光纤通道 SAN 配合使用</b>	<b>23</b>
存储区域网络概念	24
将 ESX Server 与 SAN 配合使用概述	26
将 ESX Server 与 SAN 配合使用的优点	26
ESX Server 和 SAN 用例	27
查看详细信息	27
将 SAN 阵列与 ESX Server 配合使用的细节	28
跨 ESX Server 共享 VMFS	28
元数据更新	29
LUN 的显示与重新扫描	29
主机类型	30
间接级别	30
数据访问：VMFS 或 RDM	31
第三方管理应用程序	31
区域分配和 ESX Server	31

- 访问控制 (LUN 屏蔽) 和 ESX Server 32
- 了解 VMFS 和 SAN 存储器选择 32
  - 选择较大或较小的 LUN 32
  - 决定 LUN 的大小和数目 33
    - 预测性方案 33
    - 自适应性方案 33
  - 关于决定 LUN 大小和数目的提示 34
- 了解数据访问 34
- 路径管理和故障切换 36
- 选择虚拟机位置 38
- 针对服务器故障的设计 38
  - 使用 VMware HA 39
  - 使用群集服务 39
  - 服务器故障切换和存储器注意事项 40
- 优化资源利用 40
  - 使用 VMotion 迁移虚拟机 40
  - 使用 VMware DRS 迁移虚拟机 41
- 3 要求和安装 43**
  - 常规 ESX Server SAN 要求 44
    - 将 ESX Server 与 SAN 配合使用的限制 44
    - 设置 LUN 分配 45
    - 设置光纤通道 HBA 45
  - 建议 46
  - 从 SAN 引导 ESX Server 的要求 46
  - 安装和设置步骤 47
- 4 设置与 ESX Server 配合使用的 SAN 存储设备 49**
  - 设置概述 50
    - 测试 50
    - 支持的设备 50
  - 常规设置注意事项 51
  - EMC CLARiiON 存储系统 52
    - EMC CLARiiON AX100 与 RDM 52
    - AX100 非活动连接的显示问题 52
    - 将主机配置更改应用于阵列 53
  - EMC Symmetrix 存储系统 53
  - IBM TotalStorage DS4000 存储系统 54
    - 配置用于 DS4000 存储服务器 SAN 故障切换的硬件 54
    - 验证存储处理器端口配置 55

禁用自动卷传输	56
配置存储处理器检测数据	56
IBM TotalStorage DS4000 与路径抖动	57
IBM TotalStorage 8000	58
HP StorageWorks 存储系统	58
HP StorageWorks MSA	58
将 Profile Name 设置为 Linux	58
集线器控制器问题	59
HP StorageWorks EVA	60
HP StorageWorks XP	60
Hitachi Data Systems 存储器	60
Network Appliance 存储器	61
<b>5 对 ESX Server 系统使用从 SAN 引导</b>	<b>63</b>
从 SAN 引导概述	64
从 SAN 引导如何运作	64
从 SAN 引导的优点	65
准备从 SAN 引导	65
开始前	66
从 SAN 引导模式中的 LUN 屏蔽	66
准备 SAN	67
尽可能减少启动器数目	68
为从 SAN 引导设置 FC HBA	68
为从 SAN 引导设置 QLogic FC HBA	68
启用 QLogic HBA BIOS	68
启用可选式引导	69
选择引导 LUN	69
将系统设置为先从 CD-ROM 引导	70
为从 SAN 引导设置 Emulex FC HBA	70
<b>6 管理使用 SAN 存储器的 ESX Server 系统</b>	<b>73</b>
问题与解决方案	74
避免问题发生的准则	75
获取信息	75
查看 HBA 信息	75
查看数据存储信息	76
解决显示问题	77
了解显示屏幕中的 LUN 命名	77
解决有关 LUN 不可见的问题	78
使用重新扫描	78

- 移除数据存储 79
- 高级 LUN 显示配置 79
  - 使用 Disk.MaxLUN 更改扫描的 LUN 的数目 80
  - 使用 Disk.MaskLUNs 屏蔽 LUN 80
  - 使用 Disk.SupportSparseLUN 更改稀疏 LUN 支持 81
- N-Port ID 虚拟化 81
  - 基于 NPIV 的 LUN 访问如何运作 82
  - 使用 NPIV 的要求 82
  - 向虚拟机分配 WWN 83
- 多路径 85
  - 查看当前的多路径状况 85
  - 设置 LUN 多路径策略 88
  - 禁用和启用路径 89
  - 设置固定的路径策略的首选路径 90
  - 路径管理和手动负载平衡 91
- 故障切换 92
  - 对故障切换设置 HBA 超时 93
  - 设置 SCSI 控制器的设备驱动程序选项 94
  - 设置操作系统超时 94
- VMkernel 配置 95
- 共享诊断分区 95
- 避免问题和解决问题 95
- 优化 SAN 存储性能 96
  - 存储阵列性能 96
  - 服务器性能 97
- 解决性能问题 98
  - 监视性能 98
  - 解决路径抖动 98
  - 了解路径抖动 99
  - 平衡虚拟机之间的磁盘访问 100
  - 移除 VMFS-2 驱动程序 100
  - 移除 NFS 驱动程序 100
  - 减少 SCSI 预留 101
  - 设置 HBA 的最大队列深度 101
    - 调整 QLogic HBA 的队列深度 101
    - 调整 Emulex HBA 的队列深度 102
- SAN 存储器备份注意事项 103
  - 快照软件 103
  - 使用第三方备份软件包 104
  - 选择备份解决方案 104
- 分层应用程序 105

基于阵列（第三方）的解决方案	105
基于文件 (VMFS) 的解决方案	105
VMFS 卷重新签名	106
装载原始、快照或副本 VMFS 卷	106
了解重新签名选项	106
状况 1 - EnableResignature=0, DisallowSnapshotLUN=1（默认值）	107
状况 2 - EnableResignature=1（DisallowSnapshotLUN 不相关）	107
状况 3 - EnableResignature=0, DisallowSnapshotLUN=0	107
<b>A 多路径对照表</b>	<b>109</b>
<b>B 实用程序</b>	<b>111</b>
esxstop 和 resxstop 实用程序	112
storageMonitor 实用程序	112
选项	112
示例	113
<b>索引</b>	<b>115</b>



# 关于本书

---

本手册（《光纤通道 SAN 配置指南》）说明如何将 VMware® ESX Server 系统与存储区域网络 (Storage Area Network, SAN) 配合使用。手册将在以下主要主题中讨论背景概念、安装要求和管理信息。

- 了解 ESX Server - 向 SAN 管理员介绍 ESX Server 系统。
- ESX Server 与 SAN 配合使用 - 讨论使用 ESX Server 时在设置 SAN 方面的要求与明显差异，以及如何在两个系统并用的情况下进行管理和故障排除。
- 使 ESX Server 系统可从 SAN 上的 LUN 进行引导 - 讨论从 SAN 进行引导的要求、限制及管理。

---

**注意** 本手册专门介绍通过光纤通道 (Fibre Channel, FC) 实现的 SAN。其中未涉及 iSCSI 或 NFS 存储设备。有关 iSCSI 存储器的信息，请参见《iSCSI SAN 配置指南》。有关其他类型存储器的信息，请参见《ESX Server 3i 配置指南》和《ESX Server 3 配置指南》。

---

《光纤通道 SAN 配置指南》涵盖了 ESX Server 3.5 和 ESX Server 3i 版本 3.5。为方便讲解，本书使用以下产品命名约定：

- 对于特定于 ESX Server 3.5 的主题，本书使用术语“ESX Server 3”。
- 对于特定于 ESX Server 3i 3.5 版本的主题，本书使用术语“ESX Server 3i”。
- 对于上述两款产品的通用主题，本书使用术语“ESX Server”。
- 如果讲解内容需要明确辨识某特定版本，本书将使用带版本号完整名称指代该产品。

- 如果讲解内容适用于 VMware Infrastructure 3 的所有 ESX Server 版本，则本书使用术语“ESX Server 3.x”。

## 目标读者

本手册所示信息的目标读者为经验丰富的 Windows 或 Linux 系统管理员，以及熟悉虚拟机技术和数据中心操作的人员。

## 文档反馈

VMware 欢迎您提出宝贵建议，以便改进我们的文档。如有意见，请将反馈发送到：

[docfeedback@vmware.com](mailto:docfeedback@vmware.com)

## VMware Infrastructure 文档

VMware Infrastructure 文档包括 VMware VirtualCenter 和 ESX Server 文档集。

## 图中使用的缩写

本手册中的图片使用表 1 中列出的缩写形式。

**表 1. 缩写**

缩写	描述
数据库	VirtualCenter 数据库
数据存储	受管主机的存储
dsk#	受管主机的存储磁盘
host <i>n</i>	VirtualCenter 管理的主机
SAN	受管主机之间共享的存储区域网络类型数据存储
tmpl	模板
user#	具有访问权限的用户
VC	VirtualCenter
VM#	受管主机上的虚拟机

## 技术支持和教育资源

下面几节为您介绍提供的技术支持资源。可以通过下列网址访问本手册及其他书籍的最新版本：

<http://www.vmware.com/support/pubs>

## 在线支持和电话支持

通过在线支持可提交技术支持请求、查看产品和合同信息，以及注册您的产品。网址为：<http://www.vmware.com/cn/support>。

具有相应支持合同的客户应通过电话支持获得优先级为 1 的问题的最快响应。网址为：[http://www.vmware.com/cn/support/phone\\_support.html](http://www.vmware.com/cn/support/phone_support.html)。

## 支持服务项目

了解 VMware 支持服务项目如何帮助您满足业务需求。网址为：<http://www.vmware.com/cn/support/services>。

## VMware 教育服务

VMware 课程提供了大量实践操作环境、案例研究示例，以及设计作为作业参考工具的课程材料。有关 VMware 教育服务的详细信息，请访问 <http://mylearn1.vmware.com/mgreg/index.cfm>。



# VMware ESX Server 概述

---

您可以将 ESX Server 与光纤通道存储区域网络 (Storage Area Network, SAN) 配合使用, 后者是使用光纤通道 (Fibre Channel, FC) 协议在您的计算机系统与高性能存储子系统之间传输数据的专用高速网络。将 ESX Server 与 SAN 一同使用可为整合提供额外的存储器, 提高可靠性, 并在灾难恢复方面提供帮助。

要将 ESX Server 与 SAN 有效配合使用, 您必须掌握 ESX Server 系统与 SAN 概念的相关应用知识。本章将对 ESX Server 概念进行概述。这部分内容专为不熟悉 ESX Server 系统的 SAN 管理员提供, 其中包括以下各节:

- “ESX Server 简介” (第 14 页)
- “了解虚拟化” (第 15 页)
- “与 ESX Server 系统交互” (第 19 页)
- “虚拟化概览” (第 21 页)

有关 VMware ESX Server 的详尽信息 (包括文档、硬件兼容性列表、白皮书等等), 请转至 VMware 网站 <http://www.vmware.com/cn>。

## ESX Server 简介

管理员通过 ESX Server 架构可将硬件资源分配给完全隔离的环境中的多个工作负载，这些环境称作**虚拟机**。

### 系统组件

ESX Server 系统具有以下关键组件：

- **虚拟化层** - 该层提供理想化的硬件环境及基础物理资源到虚拟机的虚拟化。其中包括负责监视虚拟化的虚拟机监视器 (Virtual Machine Monitor, VMM) 及 VMkernel。

虚拟化层可调度虚拟机操作系统，如果在 ESX Server 3 主机上运行，还可调度服务控制台。虚拟化层将管理操作系统对物理资源的访问方式。VMkernel 需要使用自身的驱动程序来提供对物理设备的访问。VMkernel 驱动程序是经过修改的 Linux 驱动程序，尽管 VMkernel 并非 Linux 变体。

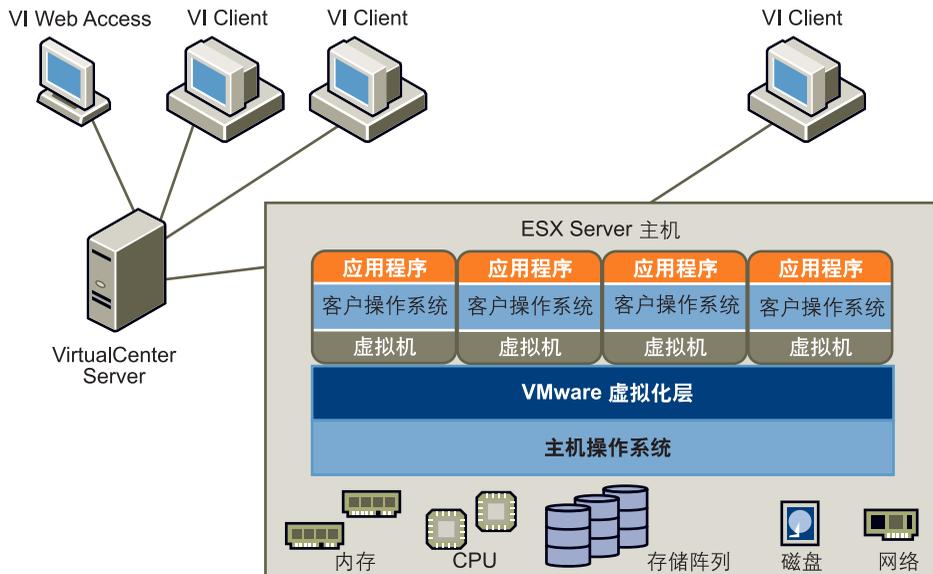
- **硬件界面组件** - 虚拟机使用硬件界面组件与 CPU 或磁盘之类的硬件进行通信。这些组件包括设备驱动程序，可实现特定于硬件的服务交付，同时隐藏与系统其他组成部分的硬件差异。
- **用户界面** - 管理员可通过若干方式查看并管理 ESX Server 主机和虚拟机：
  - VMware Infrastructure Client (VI Client) 可以直接连接 ESX Server 主机。这适合环境中仅有一台主机的情况。
  - VI Client 也可连接 VirtualCenter Server 并与 VirtualCenter Server 管理的所有 ESX Server 主机交互。
  - VI Web Access Client 可让您使用基于浏览器的界面执行多项管理任务。
  - 在极少数情况下，当您需要访问命令行时，可以使用以下选项：
    - 对于 ESX Server 3，使用服务控制台命令行界面。请参见《ESX Server 3 配置指南》中的附录 A。
    - 对于 ESX Server 3i，使用远程命令行界面 (Remote Command-Line Interface, RCLI)。请参见《ESX Server 3i 配置指南》中的附录 A。

图 1-1 显示了组件交互方式。ESX Server 配置了四台虚拟机。每个虚拟机分别运行各自的客户操作系统及应用程序。管理员通过以下方式监视主机和虚拟机：

- 使用 VI Client 直接连接 ESX Server 主机。

- 使用 VI Client 连接 VirtualCenter Management Server。VirtualCenter Server 可管理多台 ESX Server 主机。

图 1-1. Virtual Infrastructure 环境



## 软件和硬件兼容性

在 VMware ESX Server 架构中，虚拟机的操作系统（客户操作系统）仅与虚拟化层呈现的标准 x86 兼容虚拟硬件交互。通过此架构，VMware 产品可支持任何 x86 兼容操作系统。

实际上，VMware 产品支持大部分 x86 兼容操作系统，这些操作系统在整个产品开发周期中经过了测试。VMware 将这些客户操作系统的安装和操作信息编成文档，并为提供相关支持对技术人员进行培训。

大多数应用程序仅与客户操作系统交互，并不与基础硬件交互。因此，只要在虚拟机安装了应用程序所需的操作系统，您便能够在所选硬件上运行应用程序。

## 了解虚拟化

VMware 虚拟化层在各款 VMware 桌面产品（如 VMware Workstation）及服务器产品（如 VMware ESX Server）之间是通用的。该层为应用程序工作负载的开发、测试、交付及支持提供了一致平台，其构成方式如下：

- 每台虚拟机都运行各自的操作系统（客户操作系统）和应用程序。
- 虚拟化层提供映射至特定物理设备份额的虚拟设备。这些设备包括虚拟化的 CPU、内存、I/O 总线、网络接口、存储适配器和存储设备、人机接口设备以及 BIOS。

## CPU、内存及网络虚拟化

VMware 虚拟机可提供完全的硬件虚拟化。虚拟机上运行的客户操作系统和应用程序永远无法直接确定其所访问的物理资源（例如，在多处理器系统中运行了哪个物理 CPU，或者哪个物理内存映射到了其页面）。将发生以下虚拟化过程：

- **CPU 虚拟化** - 每台虚拟机显示为分别运行于各自的 CPU（或一组 CPU）之上，与其他虚拟机完全隔离。各虚拟机的注册表、翻译后备缓冲器及其他控制结构将分开保存。

大多数指令直接在物理 CPU 上执行，使得需占用大量资源的工作负载能以接近本地的速度运行。虚拟化层可安全执行特权指令。

请参见 《资源管理指南》。

- **内存虚拟化** - 各虚拟机可见一块连续的内存空间。但是，所分配的物理内存可能并不连续，而是将非连续的物理页面重新映射并呈现给各虚拟机。对于占用极大内存的负载，服务器内存将处于过量使用状态。在这种情况下，虚拟机的部分物理内存可能映射到共享页面，或者映射到未映射或换出的页面。

ESX Server 无需客户操作系统所具有的信息便可执行此虚拟内存管理，并且不会影响客户操作系统的内存管理子系统。

请参见 《资源管理指南》。

- **网络虚拟化** - 虚拟化层确保每台虚拟机与其他虚拟机相隔离。虚拟机仅可通过用于连接独立物理机的类似网络机制相互通信。

这种隔离可让管理员构建内部防火墙或其他网络隔离环境，从而使一些虚拟机可与外部连接，而另一些虚拟机只能通过虚拟网络与其他虚拟机连接。

请参见 《ESX Server 3 配置指南》或 《ESX Server 3i 配置指南》。

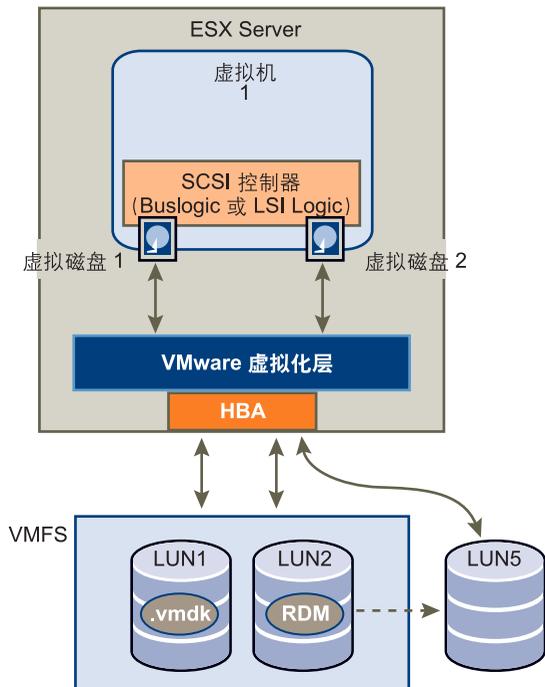
## 虚拟 SCSI

在 ESX Server 环境中，每台虚拟机包括一至四个虚拟 SCSI 主机总线适配器 (Host Bus Adapters, HBA)。这些虚拟适配器显示为 Buslogic 或 LSI Logic SCSI 控制器。虚拟机只能访问这些类型的 SCSI 控制器。

虚拟机可通过上述某一虚拟 SCSI 适配器访问的各虚拟磁盘可驻留在 VMFS 中，也可以是裸磁盘。

图 1-2 概述了存储虚拟化。图中所示为使用 VMFS 的存储器和使用裸设备映射 (Raw Device Mapping, RDM) 的存储器。

图 1-2. SAN 存储器虚拟化



### 磁盘配置选项

您可以配置具有多个虚拟 SCSI 驱动器的虚拟机。有关支持的驱动程序列表，请参见《存储器 ISAN 兼容性指南》，网址为 [www.vmware.com/support/pubs/vi\\_pubs.html](http://www.vmware.com/support/pubs/vi_pubs.html)。客户操作系统可限制 SCSI 驱动器的总数。

虽然所有 SCSI 设备均作为 SCSI 目标呈现，但是存在以下物理实施方案：

- 存储在 VMFS 卷上的虚拟机 .vmdk 文件。请参见“虚拟机文件系统”（第 18 页）。

- 映射至 SAN LUN（逻辑单元号， Logical Unit Number）的设备。请参见“[裸设备映射](#)”（第 18 页）。
- 直接传递到虚拟机的本地 SCSI 设备（例如，本地磁带驱动器）。

从虚拟机的角度而言，每个虚拟磁盘看上去都好像是与 SCSI 适配器连接的 SCSI 驱动器。实际的物理磁盘设备是通过 SCSI、iSCSI、RAID、NFS 还是光纤通道控制器来访问，这对客户操作系统以及虚拟机上运行的应用程序而言是透明的。

## 虚拟机文件系统

在简单配置中，虚拟机磁盘作为虚拟机文件系统 (Virtual Machine File System, VMFS) 内的文件进行存储。当客户操作系统向其虚拟磁盘发出 SCSI 命令时，虚拟化层会将这些命令转换为 VMFS 文件操作。

ESX Server 系统使用 VMFS 存储虚拟机文件。为最大程度减少磁盘 I/O 开销，VMFS 已经过优化，可将多台虚拟机作为一个工作负载来运行。VMFS 还为虚拟机文件提供分布式锁定，因此虚拟机在多台 ESX Server 主机共享一组 LUN 的 SAN 环境中也能安全运作。

VMFS 最初作为 ESX Server 安装的一部分进行配置。创建新的 VMFS-3 卷时，该卷必须为 1200 MB 或更大。请参见《[安装指南](#)》。随后可进行定制，如《[ESX Server 3 配置指南](#)》或《[ESX Server 3i 配置指南](#)》中所述。

一个 VMFS 卷可扩展为 32 个相同存储类型的物理存储扩展。此功能可实现存储池，并可灵活创建虚拟机所需的存储卷。当卷上运行虚拟机时，您可以根据虚拟机的需要向 VMFS 卷添加新的空间来扩展卷。

## 裸设备映射

裸设备映射 (Raw Device Mapping, RDM) 是 VMFS 卷中充当裸设备代理的特殊文件。RDM 可提供 VMFS 文件系统中虚拟磁盘的一些优点，同时保持了直接访问物理设备的部分优势。

如果您使用 Microsoft 群集服务 (Microsoft Cluster Service, MSCS) 或者在虚拟机上运行 SAN 快照或其他分层应用程序，此时可能需要 RDM。RDM 使系统能够更好地使用 SAN 阵列所拥有的硬件功能。有关 RDM 的信息，请参见《[ESX Server 3 配置指南](#)》或《[ESX Server 3i 配置指南](#)》中的“裸设备映射”，有关 MSCS 的信息，请参见《[Microsoft 群集服务的设置](#)》。

## 虚拟 SCSI 主机总线适配器

虚拟机可通过虚拟 SCSI 主机总线适配器 (Host Bus Adapter, HBA) 访问逻辑 SCSI 设备，正如通过物理 HBA 可访问物理存储设备一样。但是，存储管理员（如 SAN 管理

员) 不能通过虚拟 SCSI HBA 访问物理机。您可以将多个虚拟 HBA 隐藏在单个 (或多个) FC HBA 之后。

## 与 ESX Server 系统交互

管理员通过以下方式之一与 ESX Server 系统交互：

- 通过客户端 (VI Client 或 VI Web Access)。客户端可直接连接 ESX Server 主机，您也可以使用 VirtualCenter Management Server 同时管理多台 ESX Server 主机。
- 对于 ESX Server 3，使用服务控制台。在 ESX Server 3.x 中，不必 (也不建议) 使用服务控制台，因为您可以使用 VI Client 或 VI Web Access 执行大多数管理操作。对于脚本式管理，请使用 Virtual Infrastructure SDK。

有关服务控制台的详细信息，请参见“[ESX Server 3 服务控制台](#)” (第 20 页)。

- 对于 ESX Server 3i，使用远程命令行界面 (RCLI)。由于 ESX Server 3i 不包括服务控制台，通常使用 VI Client 配置 ESX Server 3i 主机。但是，如果您想对多台 ESX Server 3i 主机使用相同配置设置，或因其他原因而需访问命令行，此时可以使用 RCLI。

请参见《[ESX Server 3i 配置指南](#)》。

## VMware Virtual Center

您可以通过 VI Client 或 VI Web Access 访问 VirtualCenter Server。

- VirtualCenter Server 可充当在网络上进行连接的 ESX Server 主机的中心管理员。该服务器指导虚拟机和 VMware ESX Server 上的操作。
- VI Client 在 Microsoft Windows 上运行。在多主机环境中，管理员使用 VI Client 向 VirtualCenter Server 发送请求，由此转而影响其虚拟机和主机。在单服务器环境中，VI Client 直接连接 ESX Server 主机。
- 通过 VI Web Access，您可以使用 HTML 浏览器连接 VirtualCenter Server。

[图 1-3](#) 显示 VI Client 显示屏幕的 **[配置 (Configuration)]** 选项卡，其中 **[存储器 (Storage)]** 已选定。所选 ESX Server 主机与 SAN LUN 和本地硬盘连接。显示屏幕仅在设置过程中选定的名称而异。

图 1-3. VI Client 的 [ 配置 (Configuration)] 选项卡中的存储信息

The screenshot shows the 'Configuration' tab in the VI Client interface. It displays a table of storage devices and a detailed view for the 'Shared\_200' device.

标识	设备	容量	可用空间	类型
shared_50	vmhba1:0:3:1	49.75 GB	4.82 GB	vmfs3
Shared_200	vmhba1:1:0:1	199.75 GB	101.96 GB	vmfs3
localStorage	vmhba0:0:0:5	131.25 GB	130.70 GB	vmfs3

**Shared\_200** 199.75 GB 容量

位置: /vmfs/volumes/47067c74...

97.79 GB 已使用  
101.96 GB 可用空间

**路径选择**  
最近使用

**属性**  
卷标: Shared\_200  
数据存储名称: Shared\_200

**扩展**  
vmhba1:1:0:1 199.99...  
总格式化容量 199.75...

**路径**  
总计: 4  
中断: 0  
禁用: 0

**格式化**  
文件系统: VMFS 3.21  
块大小: 1 MB

## ESX Server 3 服务控制台

服务控制台是 ESX Server 3 命令行管理界面。ESX Server 3i 不提供服务控制台。服务控制台支持 ESX Server 3 系统管理功能和界面。其中包括 HTTP、SNMP 和 API 界面，以及身份验证和低性能设备访问等其他支持功能。

由于 VirtualCenter 功能得到了增强，几乎能够提供所有管理操作，现在服务控制台功能受到限制。服务控制台仅在特殊情况下使用。

**注意** 对于脚本式管理，请使用 Virtual Infrastructure SDK。

服务控制台使用经过修改的 Linux 版本得以实施。但是，服务控制台并不直接对应 Linux 命令提示符。

服务控制台中运行以下 ESX Server 3 管理过程及服务：

- **主机守护进程 (hostd)** - 代表服务控制台和 VI Client 执行服务控制台中的操作。
- **身份验证守护进程 (vmauthd)** - 使用用户名和密码数据库对 VI Client 和远程控制台的远程用户进行身份验证。您也可以使用能通过控制台的可插入验证模块

(Pluggable Authentication Module, PAM) 功能访问的其他身份验证存储。如果具有多个密码存储机制，则可以将 VMware ESX Server 与来自 Windows 域控制器、LDAP/RADIUS 服务器或者类似中央身份验证存储的密码配合使用来进行远程访问。

- **SNMP 服务器 (net-snmpd)** - 实施 SNMP 陷阱与数据结构，管理员可使用它们将 ESX Server 系统整合到基于 SNMP 的系统管理工具中。

除上述由 VMware 提供的服务之外，服务控制台还可用于运行其他系统层面或依赖于硬件的管理工具。这些工具可包括特定于硬件的运行状况监视器（如 IBM Director 或 HP Insight Manager）、全系统备份和灾难恢复软件，以及群集和高可用性产品。

---

**注意** 服务控制台不保证可用于通用的 Linux 硬件监控。它并非等同于 Linux shell 程序。

---

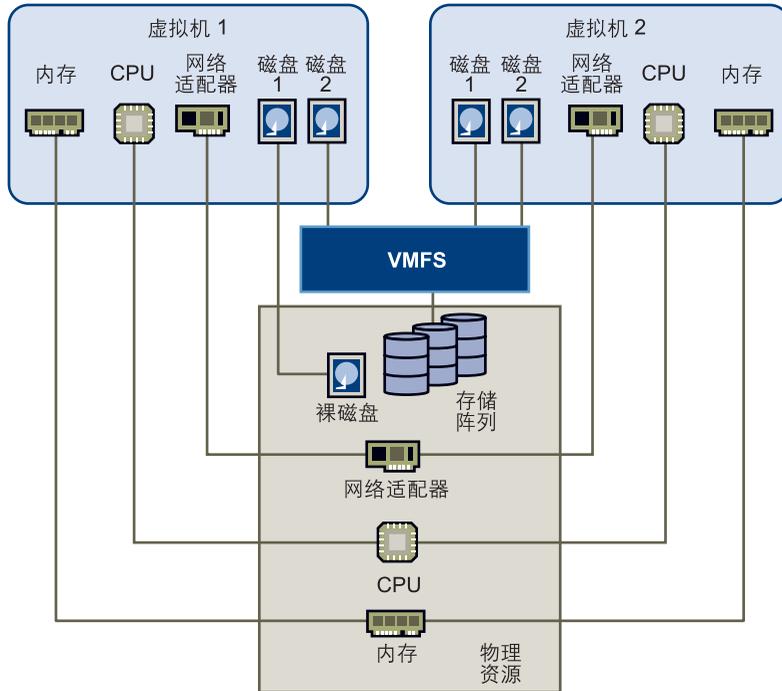
## 虚拟化概览

ESX Server 将物理系统的资源虚拟化，供虚拟机使用。

**图 1-4** 演示了多台虚拟机如何共享物理设备。图中显示两台虚拟机，分别具有如下配置：

- 一个 CPU
- 分配到的内存和一个网络适配器 (NIC)
- 两个虚拟磁盘

图 1-4. 共享物理资源的虚拟机



虚拟机各使用服务器上的某一 CPU 并访问非连续的内存页面，而且其中一个虚拟机的部分内存目前交换到磁盘（未显示）。两个虚拟网络适配器与两个物理网络适配器连接。

磁盘的映射方式如下：

- 虚拟机 1 的磁盘 1 直接映射到裸磁盘。此配置在某些情况下可能较有利。
- 虚拟机 1 的磁盘 2 和虚拟机 2 的两个磁盘驻留在位于 SAN 存储阵列的 VMFS 上。VMFS 可确保相应锁定和安全性始终保持正确无误。

# 将 ESX Server 与光纤通道 SAN 配合使用

---

# 2

将 ESX Server 主机设置为使用 FC SAN 阵列存储时，必须考虑若干特殊注意事项。本章介绍如何将 ESX Server 与 SAN 阵列配合使用，并论述下列主题：

- “[存储区域网络概念](#)”（第 24 页）
- “[将 ESX Server 与 SAN 配合使用概述](#)”（第 26 页）
- “[将 SAN 阵列与 ESX Server 配合使用的细节](#)”（第 28 页）
- “[了解 VMFS 和 SAN 存储器选择](#)”（第 32 页）
- “[了解数据访问](#)”（第 34 页）
- “[路径管理和故障切换](#)”（第 36 页）
- “[选择虚拟机位置](#)”（第 38 页）
- “[针对服务器故障的设计](#)”（第 38 页）
- “[优化资源利用](#)”（第 40 页）

## 存储区域网络概念

如果您是 ESX Server 管理员，并计划设置 ESX Server 主机与 SAN 配合使用，那么您必须掌握 SAN 概念的相关应用知识。在某些出版资料中或 Internet 上可以找到 SAN 的相关信息。两处网络资源如下：

- [www.searchstorage.com](http://www.searchstorage.com)
- [www.snia.org](http://www.snia.org)

这一领域的发展日新月异，请时常查看上述资源。

如果您对 SAN 技术不甚了解，请阅读本节以熟悉《SAN 配置指南》所用到的基本术语。要了解基本的 SAN 概念，请参见《SAN 概念与设计基础》白皮书，网址为：<http://www.vmware.com/support/pubs>。

---

**注意** SAN 管理员可跳过本节，继续阅读本章其余内容。

---

*存储区域网络 (Storage Area Network, SAN)* 是将计算机系统或主机服务器连接到高性能存储子系统的专用高速网络。SAN 组件包括主机服务器中的主机总线适配器 (Host Bus Adapter, HBA)、用于路由存储流量的交换机、线缆、存储处理器 (Storage Processor, SP) 以及存储磁盘阵列。

在网络上至少存在一台交换机的 SAN 拓扑构成了 *SAN 架构*。

为将流量从主机服务器传输到共享存储器，SAN 使用光纤通道 (Fibre Channel, FC) 协议将 SCSI 命令打包到光纤通道帧中。

在本文档中，*端口*是指设备与 SAN 的连接。SAN 中的每个节点（主机、存储设备及架构组件）均有一个或多个端口，用于将其连接到 SAN。端口可通过多种方式进行标识：

- **WWPN**（全球端口名称，World Wide Port Name）- 一种全球唯一的端口标识符，可让特定应用程序访问相应端口。FC 交换机会发现设备或主机的 WWPN，并向该设备分配一个端口地址。

要使用 VI Client 查看 WWPN，请单击主机的 **[配置 (Configuration)]** 选项卡并选择 **[存储适配器 (Storage Adapters)]**。随后便可选择要查看其 WWPN 的存储适配器。

### 详细信息

```

vmhba2
  型号：      QLA231x/2340
  WWPN：     21:01:00:e0:8b:ab:89:a0
  目标：     2
  
```

- **Port\_ID**（或端口地址）- 在 SAN 中，各端口具有唯一的端口 ID，用作端口的 FC 地址。这样就可经由 SAN 将数据路由至相应端口。FC 交换机在设备登录到相应架构时为其分配端口 ID。仅当设备登录后端口 ID 才有效。

使用 N-Port ID 虚拟化 (N-Port ID Virtualization, NPIV) 时，单一 FC HBA 端口 (N-port) 可使用多个 WWPN 向架构注册。这允许一个 N-port 要求使用多个架构地址，每个地址都显示为唯一的实体。在 ESX Server 主机所使用的 SAN 环境中，凭借上述多个唯一标识符，单个虚拟机的 WWN 分配可作为虚拟机配置的一部分。请参见 [“N-Port ID 虚拟化”](#)（第 81 页）。

在主机服务器与存储器之间传输数据时，SAN 将使用多路径技术。使用 *多路径*，您可以通过多条物理路径从 ESX Server 主机到达存储阵列上的 LUN。

如果路径或该路径所经过的任一组件（HBA、线缆、交换机端口或存储处理器）发生故障，服务器会选择另一条可用路径。检测故障路径并切换到另一条路径的过程称为 *路径故障切换*。

存储磁盘阵列可为以下类型：

- *主动 / 主动磁盘阵列*，允许通过所有可用存储处理器同时访问 LUN，但并不会明显降低性能。所有路径始终保持活动状态（除非路径发生故障）。
- *主动 / 被动磁盘阵列*，其中一个存储处理器 (Storage Processor, SP) 主动为给定 LUN 提供服务。另一 SP 将充当该 LUN 的备份并可主动为其他 LUN I/O 提供服

务。只能向活动的处理器发送 I/O。如果主要 SP 发生故障，则某个次要存储处理器将变为活动状态，此操作可自动执行，也可由管理员执行。

为限制服务器访问未分配给该服务器的存储阵列，SAN 将使用[区域分配](#)。通常会为访问一组共享存储设备和 LUN 的各组服务器创建区域。区域定义了哪些 HBA 可连接哪些 SP。某区域以外的设备对该区域之内的设备不可见。

区域分配与 LUN 屏蔽类似，后者常用于权限管理。[LUN 屏蔽](#)是使 LUN 对某些主机可用而对另外一些主机不可用的过程。LUN 屏蔽通常在 SP 或服务器级别执行。

## 将 ESX Server 与 SAN 配合使用概述

支持 FC HBA 使 ESX Server 系统能够连接 SAN 阵列。随后便可使用 SAN 阵列 LUN 存储虚拟机配置信息和应用程序数据。将 ESX Server 与 SAN 配合使用可提高灵活性、效率以及可靠性。这一方式还支持集中式管理，以及故障切换和负载平衡技术。

## 将 ESX Server 与 SAN 配合使用的优点

将 SAN 与 ESX Server 配合使用可提高环境的故障恢复能力：

- 您可以存储冗余数据并配置多个 FC 架构，从而避免出现单一故障点。当一个数据中心不可用时，不会导致整个企业瘫痪。
- ESX Server 系统在默认情况下提供多路径，并自动令所有虚拟机支持这一配置。请参见[“路径管理和故障切换”](#)（第 36 页）。
- 将 SAN 与 ESX Server 系统配合使用可将故障恢复能力扩展到服务器。使用 SAN 存储器时，所有应用程序可在发生主机故障后立即重新启动。请参见[“针对服务器故障的设计”](#)（第 38 页）。

通过专用硬件提供备用服务的成本十分昂贵，相比之下，ESX Server 与 SAN 配合使用让更多应用程序能够负担实现高可用性及自动负载平衡的费用：

- 由于有了共享的中心存储器，因此可以构建使用 MSCS 的虚拟机群集。请参见[“服务器故障切换和存储器注意事项”](#)（第 40 页）。
- 如果将虚拟机用作现有物理服务器的备用系统，那么共享存储器便必不可少，此时 SAN 是最佳解决方案。
- 使用 VMware VMotion 功能将虚拟机从一台主机无缝迁移到另一台主机。
- 将 VMware High Availability (HA) 与 SAN 一同用于冷备用解决方案，该解决方案可保证快速的自动响应。

- 使用 VMware Distributed Resource Scheduler (DRS) 将虚拟机从一台主机迁移到另一台主机以实现负载平衡。由于存储位于 SAN 阵列，应用程序可无缝继续运行。
- 如果使用 VMware DRS 群集，请将 ESX Server 主机置于维护模式，以便系统将所有正在运行虚拟机迁移到其他 ESX Server 主机。随后便可执行升级或其他维护操作。

VMware 虚拟机的可传输性和封装与 SAN 存储器的共享特性相辅相成。虚拟机位于基于 SAN 的存储器时，您可以关闭某服务器上的虚拟机并在另一服务器上将其启动，或在某服务器上将其挂起并在同一网络上的另一服务器上恢复操作，这只需几分钟而已。此功能可实现在迁移计算资源的同时保持一致的共享访问。

## ESX Server 和 SAN 用例

将 ESX Server 系统与 SAN 配合使用对于以下任务非常有效：

- **零停机时间的维护** - 执行维护时，使用 VMware DRS 或 VMotion 将虚拟机迁移到其他服务器。如果共享存储器位于 SAN 上，您无需中断用户服务就可以执行维护。
- **负载平衡** - 使用 VMotion 或 VMware DRS 将虚拟机迁移到其他主机以实现负载平衡。如果共享存储器位于 SAN 上，则无需中断用户连接就可以执行负载平衡。
- **存储整合和存储布局简化** - 如果您使用多台主机，而且每台主机运行多台虚拟机，那么主机的存储便不再够用，需要使用外部存储。选择 SAN 作为外部存储可使系统架构更为简单，同时能享有本节所列的其他优点。先预留一个较大的 LUN，然后根据需要将各部分分配给虚拟机。从存储设备预留及创建 LUN 需一次完成。
- **灾难恢复** - 将所有数据存储于 SAN 上可大大方便数据备份的远程存储。此外，如果某一站点受到损害，可以在远程 ESX Server 主机上重新启动虚拟机进行恢复。

## 查看详细信息

除本文档外，另有许多其他资源可帮助您配置 ESX Server 系统与 SAN 配合使用：

- 对于大多数设置问题，请使用存储阵列供应商的文档。存储阵列供应商可能还会提供介绍在 ESX Server 环境中使用存储阵列的文档。
- VMware 文档网站的网址为 <http://www.vmware.com/support/pubs/>。
- 《iSCSI SAN 配置指南》论述了 ESX Server 与 iSCSI 存储区域网络配合使用的情况。

- 《VMware I/O 兼容性指南》列出了当前已认可的 HBA、HBA 驱动程序以及驱动程序版本。
- 《VMware 存储器/SAN 兼容性指南》列出了当前已认可的存储阵列。
- 《VMware 发行说明》提供了有关已知问题和变通方法的信息。
- 《VMware 知识库》包含有关常见问题和变通方法的信息。

## 将 SAN 阵列与 ESX Server 配合使用的细节

将 SAN 与 ESX Server 主机一起使用在很多方面有别于传统的 SAN 应用，本节将具体论述。

### 跨 ESX Server 共享 VMFS

ESX Server VMFS 专用于从多台物理机进行的并发访问，并在虚拟机文件上执行适当的访问控制。有关 VMFS 的背景信息，请参见“[虚拟机文件系统](#)”（第 18 页）。

VMFS 可以：

- 协调对虚拟磁盘文件的访问 - ESX Server 使用文件级锁定，这由 VMFS 分布式锁定管理器管理。
- 协调对 VMFS 内部文件系统信息（元数据）的访问 - ESX Server 将暂时性 SCSI 预留用作其分布式锁定协议的一部分。SCSI 预留在对 VMFS 卷进行元数据更新的过程中不会保留。

由于虚拟机共享一个公共的 VMFS，因此可能很难得出固定的高峰访问时段或优化性能。请针对高峰时段规划虚拟机存储访问，但是不同的应用程序可能具有不同的高峰访问时段。共享 VMFS 的虚拟机越多，因 I/O 争用所致性能降低的可能也就越大。

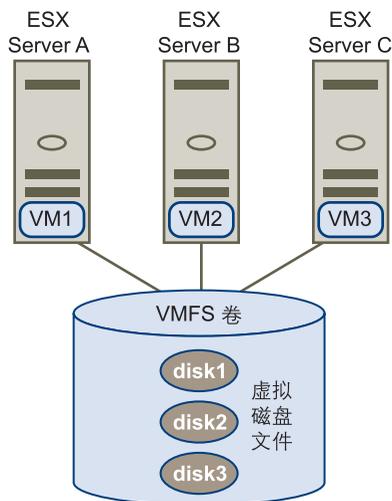
---

**注意** VMware 建议您使服务器、CPU 及存储器上的虚拟机达到负载平衡。在每个给定的服务器和存储器上运行多台虚拟机的组合，以免同时在一个区域面临很高的需求。

---

图 2-1 显示了共享同一 VMFS 卷的若干 ESX Server 系统。

图 2-1. 访问虚拟磁盘文件



## 元数据更新

VMFS 保存文件、目录、符号链接、RDM 等等，以及上述对象的相应元数据。每次访问或修改文件的属性时都会访问元数据。这些操作具体如下，但并不限于此：

- 创建、扩展或锁定文件
- 更改文件的属性
- 启动或关闭虚拟机

## LUN 的显示与重新扫描

SAN 是动态的，某特定主机可使用哪些 LUN 随多种因素而变化，其中包括：

- SAN 存储器阵列上创建的新 LUN
- 对 LUN 屏蔽所作的更改
- SAN 连接性或 SAN 的其他方面发生的变化

VMkernel 在启动时会发现 LUN，这些 LUN 随后即在 VI Client 中可见。如果对 LUN 进行了更改，您必须重新扫描才能看到相应更改。



**小心** 创建了新的 VMFS 数据存储或扩展了现有 VMFS 数据存储之后，您必须从所有可看到该特定数据存储的 ESX Server 主机重新扫描 SAN 存储器。否则，这一共享的数据存储对其中一些主机可能不可见。

## 主机类型

LUN 的行为随访问该 LUN 的主机的类型不同而稍有差别。通常，主机类型分配涉及特定于操作系统的特性或问题。ESX Server 阵列一般配置为 Linux 主机类型，如果可用，也可配置为 ESX 或 VMware 主机类型。

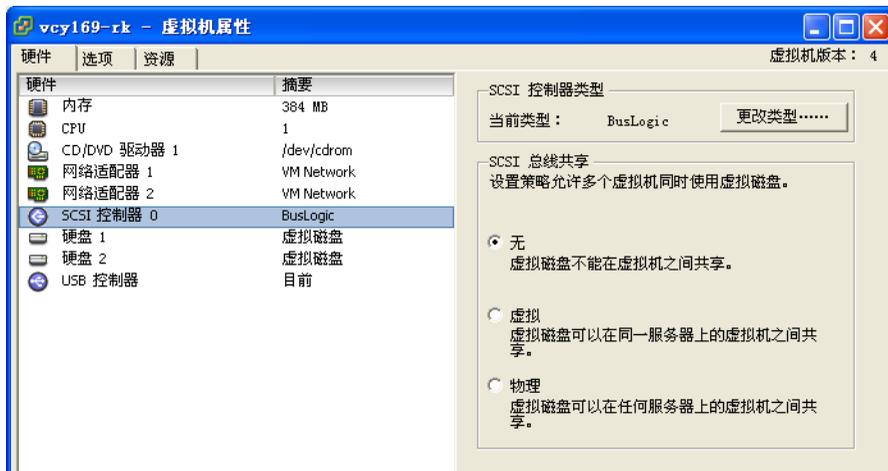
请参见第 6 章，“管理使用 SAN 存储器的 ESX Server 系统”（第 73 页）及 VMware 知识库。

## 间接级别

如果您习惯使用传统的 SAN，则起初可能对间接级别感到困惑。

- 您无法直接访问使用该存储器的虚拟机操作系统。使用传统的工具只能监视 VMware ESX Server 操作系统，而不能监视虚拟机操作系统。可以使用 VI Client 监控虚拟机。
- 默认情况下，各虚拟机在安装过程中将配置一个虚拟硬盘和一个虚拟 SCSI 控制器。可以修改 SCSI 控制器类型及 SCSI 总线共享特性，方法是使用 VI Client 编辑虚拟机设置，如图 2-2 中所示。您也可以将硬盘添加到虚拟机。请参见《基本系统管理》。

图 2-2. 设置 SCSI 控制器类型



- 通过 SAN 管理工具可看到的 HBA 属于 ESX Server 系统，而非虚拟机。
- ESX Server 系统可执行多路径。在虚拟机中不支持多路径软件（如 PowerPath），而且也不需要此种软件。

## 数据访问：VMFS 或 RDM

通常，创建虚拟机时会将虚拟磁盘置于 VMFS 数据存储。当客户操作系统向其虚拟磁盘发出 SCSI 命令时，虚拟化层会将这些命令转换为 VMFS 文件操作。请参见“[虚拟机文件系统](#)”（第 18 页）。

除 VMFS 之外还可使用 RDM。RDM 是 VMFS 卷中充当裸设备代理的特殊文件。RDM 可提供 VMFS 中虚拟磁盘的一些优点，同时保持了直接访问物理设备的部分优势。请参见“[裸设备映射](#)”（第 18 页）。

## 第三管理应用程序

大多数 SAN 硬件都附送 SAN 管理软件。此软件通常在存储阵列或单独一台服务器上运行，独立于使用 SAN 作为存储器的服务器。此第三方管理软件可用于多项任务：

- 存储阵列管理，包括 LUN 创建、阵列缓存管理、LUN 映射以及 LUN 安全。
- 设置复制、检查点、快照或镜像。

如果决定在虚拟机上运行 SAN 管理软件，您可享受有运行虚拟机的一系列优点，包括使用 VMotion 和 VMware HA 的故障切换等等。但是，由于附加的间接级别，管理软件可能无法看到 SAN。使用 RDM 可解决此问题。请参见“[分层应用程序](#)”（第 105 页）。

---

**注意** 虚拟机能否成功运行管理软件取决于存储阵列。

---

## 区域分配和 ESX Server

区域分配可在 SAN 拓扑中提供访问控制。区域分配定义了哪些 HBA 可连接哪些 SP。使用区域分配对 SAN 进行配置后，某区域以外的设备对该分区之内的设备将不可见。

区域分配具有以下作用：

- 减少了对 ESX Server 系统呈现的目标和 LUN 个数。
- 控制和隔离架构中的路径。
- 能够防止非 ESX Server 系统看到特定的存储系统，以及避免 ESX Server VMFS 数据遭到破坏。
- 可用于分隔不同的环境（例如，分隔测试环境与生产环境）。

使用区域分配时，请牢记以下几项：

- 使用共享存储器进行虚拟机故障切换或负载平衡的多台 ESX Server 主机必须属于一个区域。
- 如果进行了大型的部署，您可能需要分别为不同的功能领域创建单独的区域。例如，可以将财务会计与人力资源分隔开来。
- 创建许多较小的区域（例如，两台主机，各有四台虚拟机）并不会达到很好的效果。

---

**注意** 请咨询存储阵列供应商以了解最佳区域分配方案。

---

## 访问控制（LUN 屏蔽）和 ESX Server

使用访问控制可以限制能够看到 LUN 的 ESX Server 主机（或其他主机）数目。访问控制对于以下方面可能很有用：

- 减少对 ESX Server 系统呈现的 LUN 个数。
- 防止非 ESX Server 系统看到 ESX Server LUN，以及避免 VMFS 卷遭到破坏。

## 了解 VMFS 和 SAN 存储器选择

本节论述可用的 VMFS 和 SAN 存储器选择，并提供关于如何进行选择的建议。

### 选择较大或较小的 LUN

设置 ESX Server 系统的存储器时，请选择以下方式之一：

- 设置多个 LUN，每个 LUN 上各有一个 VMFS 卷
- 设置多个 LUN，所有 LUN 属于单个 VMFS 卷

每个 LUN 只能有一个 VMFS 卷。但是，您可以决定使用一个大的 LUN 或多个小的 LUN。

出于以下原因，您可能需要较少、较大的 LUN：

- 更灵活地创建虚拟机，无需向 SAN 管理员要求更多空间。
- 更灵活地调整虚拟磁盘大小、执行快照等等。
- 减少需要标识和管理的 LUN。

出于以下原因，您可能需要较多、较小的 LUN：

- 不同的应用程序可能需要不同的 RAID 特性。
- 增加灵活性（为每个 LUN 设置多路径策略和磁盘份额）。
- 使用 Microsoft 群集服务要求每个群集磁盘资源位于各自的 LUN 上。

## 决定 LUN 的大小和数目

如果没有相关虚拟机的存储描述，请通过以下方式之一来决定 LUN 的大小及要使用多少个 LUN：

- 预测性方案
- 自适应性方案

### 预测性方案

在预测性方案中，将执行以下操作：

- 创建若干具有不同存储特性的 LUN。
- 在每个 LUN 上构建一个 VMFS 卷（依照相应特性标记各个卷）。
- 根据各应用程序的需要将其置于适当的 RAID。
- 使用磁盘份额来区别高优先级虚拟机与低优先级虚拟机。磁盘份额仅表示给定 ESX Server 主机内的比例。分配给某一 ESX Server 主机上虚拟机的份额并不影响其他 ESX Server 主机上的虚拟机。

### 自适应性方案

在自适应性方案中，将执行以下操作：

- 创建一个较大的 LUN (RAID 1+0 或 RAID 5)，同时启用写入缓存。
- 在该 LUN 上构建 VMFS。
- 将多个磁盘置于 VMFS 之上
- 运行应用程序并确定磁盘性能是否可满足需要。
- 如果性能可满足需要，您可以向 VMFS 上再添加虚拟磁盘。如果性能无法满足需要，请新建更大的 LUN（可能采用不同的 RAID 级别）并重复此过程。您可以使用冷迁移，这样便不会在重新创建 LUN 时丢失虚拟机。

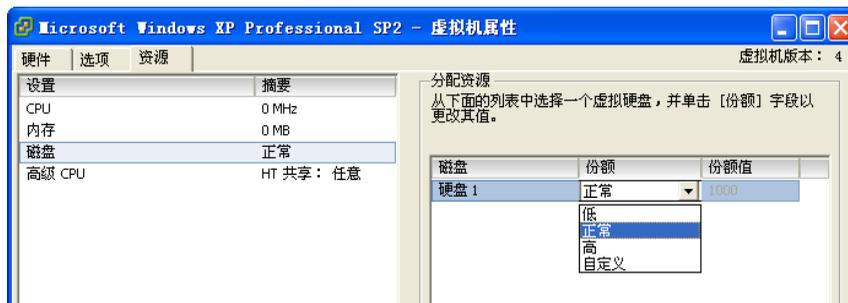
## 关于决定 LUN 大小和数目的提示

决定 LUN 的大小和数目时，请牢记以下几点：

- 每个 LUN 应具有正确的 RAID 级别和存储特性，适用于使用该 LUN 的虚拟机中的应用程序。
- 一个 LUN 必须仅包含一个 VMFS 卷。
- 如果多台虚拟机访问同一个 LUN，请使用磁盘份额区分虚拟机的优先级。

### 使用磁盘份额区分虚拟机的优先级

- 1 启动 VI Client 并连接 VirtualCenter Server。
- 2 从清单中选择虚拟机，右键单击，然后选择 [编辑设置 (Edit Settings)]。
- 3 单击 [资源 (Resources)] 选项卡并单击 [磁盘 (Disk)]。
- 4 右键单击要修改的磁盘的 [份额 (Shares)] 列，然后从下拉菜单中选择所需值。



份额是表示控制所有虚拟机磁盘带宽的相对衡量指标的值。[低 (Low)]、[正常 (Normal)]、[高 (High)] 以及 [自定义 (Custom)] 各值将与服务器上所有虚拟机的总份额相比较，而在 ESX Server 3 主机上，则将与服务控制台上所有虚拟机的总份额进行比较。份额分配象征值可用于将其配置为转换成数字值。

## 了解数据访问

虚拟机通过以下方式之一访问数据：

- **VMFS** - 在简单配置中，虚拟机的磁盘将存储为 ESX Server VMFS 数据存储内的 .vmdk 文件。当客户操作系统向其虚拟磁盘发出 SCSI 命令时，虚拟化层会将这些命令转换为 VMFS 文件操作。

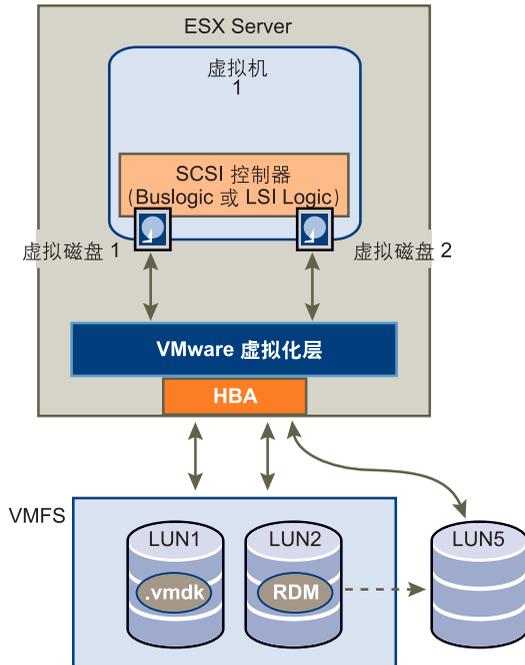
在默认设置中，虚拟机在访问文件时始终经过 VMFS，不管文件位于 SAN 还是主机的本地硬盘。请参见“[虚拟机文件系统](#)”（第 18 页）。

- **RDM** - RDM 是 VMFS 内充当裸设备代理的映射文件。RDM 可赋予客户操作系统访问裸设备的权限。

RDM 适合虚拟机必须直接与 SAN 上的物理磁盘交互的情况。例如，当您要从客户操作系统发出磁盘阵列快照创建命令时，或是数据量很大，您不想将其移动到虚拟磁盘时（较为少见），就会出现这种情况。设置 Microsoft 群集服务时也需要 RDM。请参见 VMware 文档《*Microsoft 群集服务的设置*》。

图 2-3 显示虚拟机如何使用 VMFS 或 RDM 访问数据。

图 2-3. 虚拟机如何访问数据



有关 VMFS 和 RDM 的详细信息，请参见《*ESX Server 3 配置指南*》或《*ESX Server 3i 配置指南*》。

虚拟机与 SAN 交互时，将发生以下过程：

- 1 虚拟机中的客户操作系统需要读写 SCSI 磁盘时，将向虚拟磁盘发出 SCSI 命令。
- 2 虚拟机操作系统中的设备驱动程序将与虚拟 SCSI 控制器进行通信。VMware ESX Server 支持两种类型的虚拟 SCSI 控制器：BusLogic 和 LSILogic。
- 3 虚拟 SCSI 控制器将命令转发至 VMkernel。

- 4 VMkernel:
  - 在 VMFS 卷中查找与客户机虚拟机磁盘对应的文件。
  - 将对虚拟磁盘上块的请求映射到相应物理设备上的块。
  - 将修改后的 I/O 请求从 VMkernel 中的设备驱动程序发送到物理 HBA（主机 HBA）。
- 5 主机 HBA:
  - 将请求从本身的二进制数据形式转换为经光缆传输所需的光学形式。
  - 依据 FC 协议的规则将请求打包。
  - 将请求传输至 SAN。
- 6 根据 HBA 用于连接架构的端口，相应 SAN 交换机会收到请求并将其路由至主机要访问的存储设备。

从主机的角度来看，此存储设备显示为某一特定磁盘，但它可能是与 SAN 上的物理设备相对应的逻辑设备。交换机必须确定哪个物理设备可作为主机的目标逻辑设备。

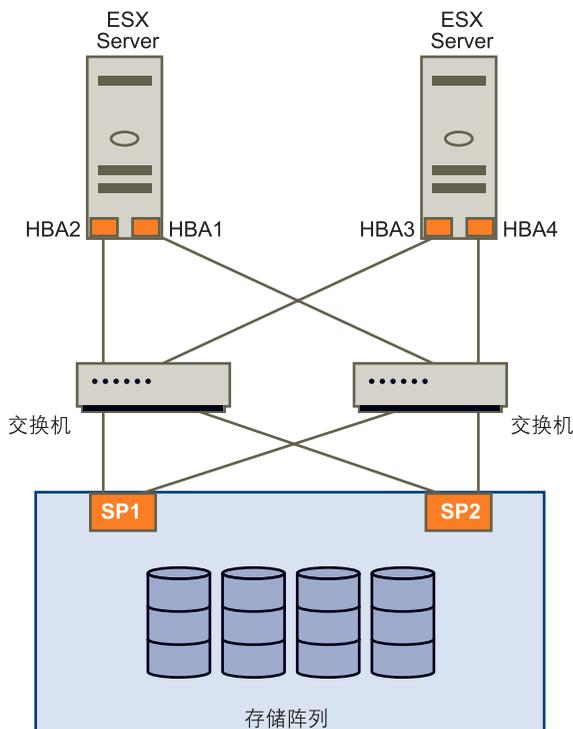
## 路径管理和故障切换

ESX Server 支持多路径，以便在 HBA、交换机、SP 或 FC 线缆发生故障时维持服务器计算机与存储设备之间的持续连接。多路径支持无需特定故障切换驱动程序。

为了支持路径切换，服务器通常具有两个或更多的可用 HBA，使用一个或多个交换机可从这些 HBA 到达存储阵列。或者，设置可以包括一个 HBA 和两个存储处理器，以便 HBA 可以使用不同的路径到达磁盘阵列。

在图 2-4 中，多条路径将每台服务器与存储设备相连。例如，如果 HBA1 或 HBA1 与 FC 交换机之间的链路发生故障，HBA2 会取代 HBA1 并提供服务器与交换机之间的连接。一个 HBA 取代另一个 HBA 的过程称为 *HBA 故障切换*。

图 2-4. 多路径和故障切换



类似地，如果 SP1 发生故障或 SP1 与交换机之间的链路中断，SP2 会取代 SP1 并提供交换机与存储设备之间的连接。此过程称为 *SP 故障切换*。VMware ESX Server 通过多路径功能支持 HBA 和 SP 故障切换。

您可以为系统选择一个多路径策略：**[ 固定的 (Fixed) ]** 或 **[ 最近使用 (Most Recently Used) ]**。如果策略为 **[ 固定的 (Fixed) ]**，您可以指定一条首选路径。对 ESX Server 主机可见的每个 LUN（磁盘）都可以有各自的路径策略。有关如何查看当前多路径状况以及如何设置多路径策略的信息，请参见 [“多路径”](#)（第 85 页）。

发生故障切换时，虚拟机 I/O 最长可能延迟 60 秒，特别是在主动 / 被动阵列上。这一延迟必不可少，以便 SAN 架构能在发生拓扑更改或其他架构事件后稳定自身配置。如果主动 / 被动阵列使用 **[ 固定的 (Fixed) ]** 策略，则路径抖动可能是个问题。请参见 [“解决路径抖动”](#)（第 98 页）。

如果通往存储虚拟机磁盘的存储设备的所有路径都不可用，虚拟机将发生不可预知的故障。

## 选择虚拟机位置

如果您要设法优化虚拟机的性能，存储位置是个重要因素。提供高性能和高可用性但价格昂贵的存储器与性能较低但成本也较低的存储器之间始终存在一个权衡。根据多种因素，存储器可分为不同级别：

- **高端** - 提供高性能和高可用性。可能提供内置快照，便于备份及时间点 (Point-in-Time, PiT) 还原。支持复制、完全 SP 冗余和光纤驱动器。使用高成本心轴。
- **中端** - 提供中等程度的性能、较低可用性、部分 SP 冗余以及 SCSI 驱动器。可能提供快照。使用中等成本的心轴。
- **低端** - 提供低性能及少许内部存储冗余。使用低端 SCSI 驱动器或 SATA（串行低成本心轴）。

并非所有应用程序都需要性能最高、可用性最佳的存储器，至少不是在整个生命周期内一直需要。

如果您需要高端存储器的部分功能（如快照），但不想为此而支付高额成本，那么或许可以在软件中实现某些高性能特性。例如，您可以在软件中创建快照。

决定放置某虚拟机的位置时，请考虑以下问题：

- 该虚拟机的重要程度如何？
- 有哪些性能及可用性要求？
- 有哪些时间点 (Point-in-Time, PiT) 还原要求？
- 有哪些备份要求？
- 有哪些复制要求？

由于重要程度发生改变或技术发展使得目前的较高端功能降为低端功能，虚拟机在其整个生命周期过程中可能变更级别。重要程度是相对的，并且可能由于多种原因而改变，包括组织、操作流程、法规要求及灾难规划等方面的变化。

## 针对服务器故障的设计

SAN 存储器的 RAID 架构本质上是在物理磁盘级别提供故障保护。双重架构拥有所有架构组件的副本，可在发生大多数架构故障的情况下保护 SAN。使整个环境变得耐故障的最后一步即提供针对服务器故障的保护。ESX Server 系统故障切换选项将在随后几节中进行论述。

## 使用 VMware HA

使用 VMware HA 可将虚拟机组织为故障切换组。某一主机发生故障时，其所有虚拟机将在其他主机上立即启动。HA 需要 SAN 存储器。

虚拟机在其他主机上还原后会丢失内存状况，但磁盘状况与主机发生故障时完全一致（崩溃一致故障切换）。HA 需要共享存储器，如 SAN。请参见《资源管理指南》。

---

**注意** 您必须获得许可才能使用 VMware HA。

---

## 使用群集服务

服务器群集是一种使用高速网络连接尝试将两个或更多服务器集合在一起的方法，这样，一组服务器便可作为单个逻辑服务器运行。如果其中一台服务器发生故障，群集内的其他服务器将继续运行，同时接管故障服务器所执行的操作。

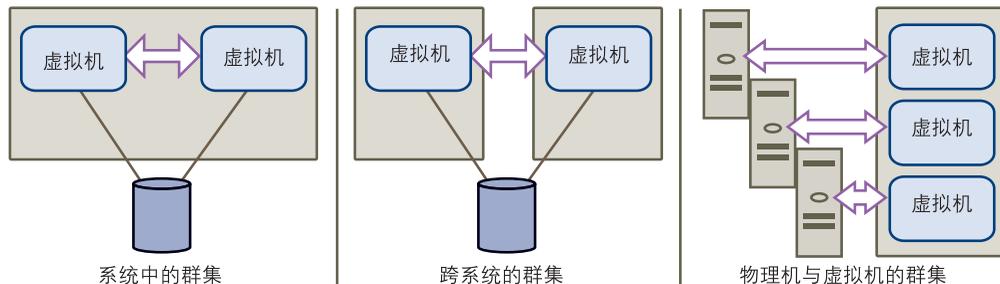
VMware 对 Microsoft 群集服务与 ESX Server 系统配合使用进行了测试，但是其他群集解决方案可能同样可行。多种配置选项可用于通过群集实现故障切换：

- **机箱内群集** - 同一主机上的两台虚拟机可相互作为对方的故障切换服务器。一台虚拟机发生故障后，将由另一台取代。此配置不会对主机故障提供保护。通常在群集应用程序的测试过程中完成此配置。
- **机箱间群集** - 某一 ESX Server 主机上的虚拟机在另一 ESX Server 主机上具有与其相匹配的虚拟机。
- **物理到虚拟群集 (N+1 群集)** - 某一 ESX Server 主机上的虚拟机用作一台物理服务器的故障切换服务器。由于单个主机上运行的虚拟机可用作多台物理服务器的故障切换服务器，因此这一群集方式可提供经济高效的 N+1 解决方案。

请参见《Microsoft 群集服务的设置》。

图 2-5 显示可用于通过群集实现故障切换的多种配置选项。

图 2-5. 使用群集服务的群集



## 服务器故障切换和存储器注意事项

对于各种类型的服务器故障切换，您必须考虑有关存储器的问题：

- 仅当每台服务器都对同一存储器具有访问权限时用于服务器故障切换的方法才可。因为多台服务器需要大量磁盘空间，而且存储阵列故障切换与服务器故障切换相辅相成，所以通常会将 SAN 与服务器故障切换结合使用。
- 规划 SAN 与服务器故障切换协同工作时，由群集虚拟机使用的所有 LUN 必须对所有 ESX Server 主机均可见。这一要求与 SAN 管理员的直觉正好相反，但却适用于使用虚拟机的情况。

虽然主机可以访问 LUN，但并非该主机上的所有虚拟机就一定具有对 LUN 上所有数据的访问权限。虚拟机只能访问为其配置的虚拟磁盘。假设出现配置错误，在虚拟机引导时，虚拟磁盘将被锁定，因此不会发生损坏现象。

---

**注意** 一般来说，使用从 SAN 引导时，只有正在从 LUN 引导的 ESX Server 系统才能看到各引导 LUN。一种例外情况就是当您通过将另一 ESX Server 系统指向同一 LUN 来尝试从故障中恢复时。在这种情况下，上述 SAN LUN 并非真正从 SAN 引导。因 SAN 已损坏，未有任何 ESX Server 系统从其引导。SAN LUN 是对 ESX Server 系统可见的非引导 LUN。

---

## 优化资源利用

使用 VMware Infrastructure 可将虚拟机从过度利用的主机迁移到利用不足的主机，从而优化资源分配。存在以下选项：

- 使用 VMotion 手动迁移虚拟机。
- 使用 VMware DRS 自动迁移虚拟机。

仅当虚拟机位于可由多台服务器访问的共享存储器上时才能使用 VMotion 或 DRS。多数情况下使用 SAN 存储器。有关 VMotion 的其他信息，请参见《基本系统管理》。有关 DRS 的其他信息，请参见《资源管理指南》。

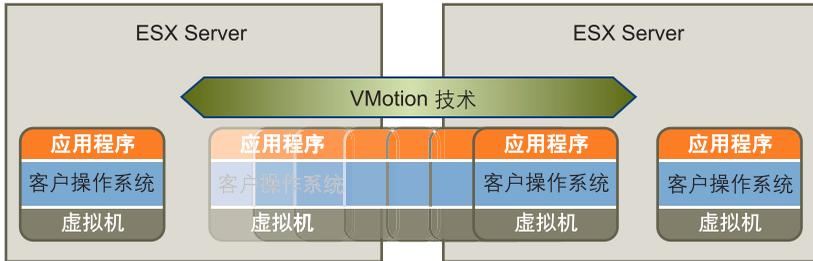
## 使用 VMotion 迁移虚拟机

管理员通过 VMotion 可以手动将虚拟机迁移到不同的主机。管理员可以将正在运行的虚拟机迁移到同一 SAN 所连接的另一台物理服务器，无需中断服务。VMotion 实现了：

- 通过移动虚拟机执行零停机时间维护，使得对基础硬件及存储器的维修不会导致用户会话中断。
- 根据业务需求的变化不断平衡数据中心间的工作负载，实现最有效的资源利用。

图 2-6 显示如何使用 VMotion 迁移虚拟机。

图 2-6. 通过 VMotion 迁移



## 使用 VMware DRS 迁移虚拟机

VMware DRS 有助于改善所有主机和资源池之间的资源分配。DRS 将收集 VMware 群集中所有主机和虚拟机的资源使用情况信息，并在出现以下两种情况之一时给出建议（或迁移虚拟机）：

- **初始放置位置** - 当您在群集中首次启动某个虚拟机时，DRS 将放置该虚拟机或提出建议。
- **负载均衡** - DRS 通过执行虚拟机自动迁移 (VMotion) 或提供虚拟机迁移建议来设法改善群集中的资源利用。

请参见 《资源管理指南》。



## 要求和安装

---

本章将论述有关将 ESX Server 系统与 SAN 存储器配合使用的硬件和系统要求。本章包括以下各节：

- “常规 ESX Server SAN 要求”（第 44 页）
- “从 SAN 引导 ESX Server 的要求”（第 46 页）
- “安装和设置步骤”（第 47 页）

本章仅列出最基本的要求。有关设置系统的详细信息，请阅读第 4 章，“设置与 ESX Server 配合使用的 SAN 存储设备”（第 49 页）。

## 常规 ESX Server SAN 要求

当您准备配置 SAN 并设置 ESX Server 系统以使用 SAN 存储器时，请查看以下要求和建议：

- **硬件和固件。** 仅支持有限的若干 SAN 存储硬件和固件组合与 ESX Server 系统配合使用。有关最新列表，请参见 *《存储器/SAN 兼容性指南》*。
- **每个 LUN 一个 VMFS 卷。** 将系统配置为每个 LUN 只有一个 VMFS 卷。使用 VMFS-3 时无需设置可访问性。
- 除非使用无磁盘服务器，否则请勿在 SAN LUN 上设置诊断分区。  
共享的诊断分区适用于从 SAN 引导无磁盘服务器的情况。请参见 [“共享诊断分区”](#)（第 95 页）。
- VMware 建议从装有 ESX Server 2.5 或更高版本的计算机访问任意裸磁盘时使用 RDM。有关 RDM 的详细信息，请参见 *《ESX Server 3 配置指南》* 或 *《ESX Server 3i 配置指南》*。
- **多路径。** 要使多路径正常工作，每个 LUN 必须对所有 ESX Server 主机呈现相同的 LUN ID 号。
- **队列大小。** 确保客户操作系统中的 BusLogic 或 LSILogic 驱动程序指定一个足够大的队列。您可以在系统设置过程中设置物理 HBA 的队列深度。有关支持的驱动程序，请参见 *《存储器/SAN 兼容性指南》*。
- **SCSI 超时。** 在运行 Microsoft Windows 的虚拟机上，考虑增加 SCSI TimeoutValue 参数值，以使 Windows 能接受因路径故障切换所导致的更长 I/O 延迟。请参见 [“设置操作系统超时”](#)（第 94 页）。

## 将 ESX Server 与 SAN 配合使用的限制

将 ESX Server 与 SAN 配合使用时存在以下限制：

- ESX Server 不支持 FC 连接磁带设备。VMware Consolidated Backup 代理可以管理这些设备。请参见 *《虚拟机备份指南》*。
- 您不能使用虚拟机多路径软件对单个物理 LUN 执行 I/O 负载平衡。
- 您不能使用虚拟机逻辑卷管理器软件对虚拟磁盘进行镜像。Microsoft Windows 虚拟机上的动态磁盘是个例外，但需要进行特殊配置。

## 设置 LUN 分配

设置 LUN 分配时，请注意以下几点：

- **存储器置备。**为确保 ESX Server 系统在启动时识别 LUN，将 SAN 连接到 ESX Server 系统前请将所有 LUN 置备到相应的 HBA。

VMware 建议您同时将所有 LUN 置备到所有 ESX Server HBA。仅当所有 HBA 都看到相同 LUN 时 HBA 故障切换才可行。

- **VMotion 和 VMware DRS。**当您使用 VirtualCenter 以及 VMotion 或 DRS 时，请确保用于虚拟机的 LUN 已置备到所有 ESX Server 主机。这可为移动虚拟机提供最大的自由。
- **主动 / 主动与主动 / 被动阵列比较。**将 VMotion 或 DRS 用于主动 / 被动 SAN 存储设备时，请确保所有 ESX Server 系统通向所有存储处理器的路径均一致。否则在进行 VMotion 迁移时可能导致路径抖动。请参见“[解决路径抖动](#)”（第 98 页）。

对于《[存储器 / SAN 兼容性指南](#)》中未列出的主动 / 被动存储阵列，VMware 不支持存储端口故障切换。在上述情况下，您必须将服务器连接到存储阵列上的主动端口。此配置可确保向 ESX Server 主机呈现 LUN。

## 设置光纤通道 HBA

设置 FC HBA 的过程中，请考虑以下几点：

- **HBA 默认设置：**FC HBA 可在默认配置设置下正常工作。请遵循存储阵列供应商提供的配置准则。

---

**注意** 为达到最佳效果，请在同一服务器中使用相同型号的 HBA。请确保服务器中各 HBA 上的固件级别相同。不支持将同一服务器中的 Emulex 和 QLogic HBA 用于同一目标。

---

- **通过 HBA 的静态负载平衡。**您可以配置某些 ESX Server 系统通过多个 HBA 将流量均衡加载到多个包含特定主动 / 主动阵列的 LUN。

为此，请向 LUN 分配首选路径，以便均衡使用 HBA。例如，如果有两个 LUN（A 和 B）和两个 HBA（X 和 Y），您可以将 HBA X 设置为 LUN A 的首选路径，并将 HBA Y 设置为 LUN B 的首选路径。这么做可以充分利用 HBA。这种情况下的路径策略必须设置为 **[固定的 (Fixed)]**。请参见“[使用 VI Client 设置多路径策略](#)”（第 89 页）。

- **对故障切换设置超时。** 设置用于检测 HBA 驱动程序中的路径出现故障的超时值。VMware 建议将超时设置为 30 秒以确保最佳的性能。要设置此值，请按照“[对故障切换设置 HBA 超时](#)”（第 93 页）中的说明进行操作。
  - **对磁带驱动器使用专用适配器。** 为达到最佳效果，请对 ESX Server 系统所连接的任何磁带驱动器都使用专用的 SCSI 适配器。不支持 FC 连接磁带驱动器。请使用 Consolidated Backup 代理，如《[虚拟机备份指南](#)》中所述。
- 有关从 SAN HBA 设置引导的其他信息，请参见第 5 章，“[对 ESX Server 系统使用从 SAN 引导](#)”（第 63 页）。

## 建议

设置使用 ESX Server 主机和 SAN 的环境时请考虑以下建议：

- 对虚拟机的虚拟磁盘使用 RDM 以利用该磁盘阵列的某些硬件快照功能，或从用于数据 LUN 的冷备用主机配置中的虚拟机以及物理机访问磁盘。
- 对 Microsoft 群集服务设置中的共享磁盘使用 RDM。请参见《[Microsoft 群集服务的设置](#)》。
- 分配一个较大的 LUN 供多台虚拟机使用，并将其设置为 VMFS。随后您便可以动态创建或删除虚拟机，不必在每次添加虚拟机时请求额外的磁盘空间。
- 要使用 VMotion 将虚拟机移至其他主机，那么虚拟机的虚拟磁盘所在的 LUN 必须对所有主机均可见。

有关其他建议以及故障排除信息，请参见第 6 章，“[管理使用 SAN 存储器的 ESX Server 系统](#)”（第 73 页）。

## 从 SAN 引导 ESX Server 的要求

如果已针对 ESX Server 系统对 SAN 存储器进行了配置，则您可以将 ESX Server 引导映像置于该 SAN 之上的某个 LUN。此配置必须满足本节所论述的特定标准。请参见“[对 ESX Server 系统使用从 SAN 引导](#)”（第 63 页）。

要实现从 SAN 引导 ESX Server 系统，请执行以下任务：

- 检查环境是否满足常规要求。请参见“[常规 ESX Server SAN 要求](#)”（第 44 页）。
- 完成表 3-1 中所列的任务。

表 3-1. 从 SAN 引导的要求

要求	描述
ESX Server 系统要求	建议使用 ESX Server 3.x。使用 ESX Server 3.x 系统时，支持 RDM 与从 SAN 启动一起使用。对于 ESX Server 2.5.x 系统，不支持 RDM 与从 SAN 启动一起使用。
HBA 要求	<p>必须启用并正确配置 HBA FC 卡的 HBA BIOS 才能访问启动 LUN。请参见“<a href="#">设置光纤通道 HBA</a>”（第 45 页）。</p> <p>应将 HBA 插入 PCI 总线和插槽号最低的位置。这使得驱动程序能够迅速检测到 HBA，因为驱动程序按 PCI 总线和插槽号的升序来扫描 HBA，与关联的虚拟机 HBA 号无关。</p> <p>有关精确的驱动程序及版本信息，请参见《<a href="#">ESX Server I/O 兼容性指南</a>》。</p>
引导 LUN 注意事项	<ul style="list-style-type: none"> <li>■ 从主动 / 被动存储阵列引导时，与 HBA BIOS 配置中指定的 WWN 相对应的 SP 必须是主动的。如果该 SP 是被动的，则 HBA 无法支持引导过程。</li> <li>■ 为便于 BIOS 配置，请屏蔽各引导 LUN，使其仅对各自的 ESX Server 系统可见。每个 ESX Server 系统应看到自身的引导 LUN，但是不应看到任何其他 ESX Server 系统的引导 LUN。</li> </ul>
SAN 注意事项	<ul style="list-style-type: none"> <li>■ SAN 连接必须经由一个交换机架构拓扑。从 SAN 引导不支持直接连接（即不经过交换机的连接）或 FC 仲裁环路 (Arbitrated Loop) 连接。</li> <li>■ 支持冗余或非冗余配置。在冗余的情况下，ESX Server 会将冗余路径折叠，因此用户仅看到一条 LUN 路径。</li> </ul>
特定于硬件的注意事项	<p>如果正在运行 IBM eServer BladeCenter 并使用从 SAN 引导，您必须禁用刀片服务器上的 IDE 驱动器。</p> <p>有关其他特定于硬件的注意事项，请参见 VMware 知识库文章和 <a href="#">第 4 章</a>，“<a href="#">设置与 ESX Server 配合使用的 SAN 存储设备</a>”（第 49 页）。</p>

## 安装和设置步骤

[表 3-2](#) 概述了安装和设置步骤，并指出了相关信息的参考路径。

表 3-2. 安装和设置步骤

步骤	描述	参考
1	如果尚未配置 SAN，请设计 SAN。多数现有的 SAN 只需略微修改即可用于 ESX Server。	<a href="#">第 2 章</a> ，“ <a href="#">将 ESX Server 与光纤通道 SAN 配合使用</a> ”（第 23 页）。
2	检查所有 SAN 组件是否满足要求。	<a href="#">第 3 章</a> ，“ <a href="#">常规 ESX Server SAN 要求</a> ”（第 44 页）。 <a href="#">《存储器 / SAN 兼容性指南》</a> 。

**表 3-2. 安装和设置步骤 (续)**

步骤	描述	参考
3	设置 ESX Server 主机的 HBA。	有关仅适用于从 SAN 引导的特殊要求, 请参见第 3 章, “从 SAN 引导 ESX Server 的要求” (第 46 页)。另请参见第 5 章, “对 ESX Server 系统使用从 SAN 引导” (第 63 页)。
4	进行任何必要的存储阵列修改。	有关简介请参见第 4 章, “设置与 ESX Server 配合使用的 SAN 存储设备” (第 49 页)。多数供应商会提供各自的特定文档, 介绍如何设置 SAN 与 VMware ESX Server 配合使用。
5	在已连接 SAN 并设置了 HBA 的主机上安装 ESX Server。	《安装指南》。
6	创建虚拟机。	《基本系统管理》。
7	(可选) 为实现 VMware HA 故障切换或使用 Microsoft 群集服务而对系统进行设置。	《资源管理指南》。 《Microsoft 群集服务的设置》。
8	根据需要升级或修改环境。	第 6 章, “管理使用 SAN 存储器的 ESX Server 系统” (第 73 页) 将进行介绍。有关特定于计算机的信息以及最近发布的新闻, 请搜索 VMware 知识库文章。

# 设置与 ESX Server 配合使用的 SAN 存储设备

---

# 4

本章讨论支持与 VMware ESX Server 一起使用的几种存储设备。对于每种设备，本章将列出可能出现的主要已知问题，指示特定供应商的信息（如果有），并包括来自 VMware 知识库文章的信息。

---

**注意** 本文中的信息仅随各版本更新。新的信息或许已可通过其他途径获取。请参考最新版本的《*存储器/SAN 兼容性指南*》，咨询存储阵列供应商，以及查看 VMware 知识库文章。

---

本章将讨论以下主题：

- “[设置概述](#)”（第 50 页）
- “[常规设置注意事项](#)”（第 51 页）
- “[EMC CLARiON 存储系统](#)”（第 52 页）
- “[EMC Symmetrix 存储系统](#)”（第 53 页）
- “[IBM TotalStorage DS4000 存储系统](#)”（第 54 页）
- “[IBM TotalStorage 8000](#)”（第 58 页）
- “[HP StorageWorks 存储系统](#)”（第 58 页）
- “[Hitachi Data Systems 存储器](#)”（第 60 页）
- “[Network Appliance 存储器](#)”（第 61 页）

## 设置概述

VMware ESX Server 支持各种不同配置的 SAN 存储阵列。并非所有存储设备都针对 ESX Server 的所有特性和功能进行了认证，供应商可能就 ESX Server 提供特定形式的支持。有关支持的存储阵列的最新信息，请参见《*存储器/SAN 兼容性指南*》。

## 测试

VMware 会对使用存储阵列的 ESX Server 进行以下配置测试：

- **基本连接** - 测试 ESX Server 能否识别和操作存储阵列。此配置不允许应用多路径或任何类型的故障切换。
- **HBA 故障切换** - 服务器配备了多个 HBA，这些 HBA 与一个或多个 SAN 交换机相连。服务器仅在发生 HBA 和交换机故障时能保证正常运行。
- **存储端口故障切换** - 服务器与多个存储端口相连，并在发生存储端口故障和交换机故障时能正常运行。
- **从 SAN 引导** - ESX Server 主机从 SAN 上配置的 LUN，而非从服务器本身进行引导。
- **直接连接** - 服务器不通过交换机，而仅使用 FC 线缆连接阵列。对于所有其他测试，使用架构连接。FC 仲裁环路 (AL) 不受支持。
- **群集** - 通过在虚拟机中运行 Microsoft 群集服务来测试系统。请参见《*Microsoft 群集服务的设置*》文档。

## 支持的设备

表 4-1 列出 ESX Server 3.x 支持的存储设备，同时指出从何处能找到有关这些设备与 ESX Server 一起使用的详细信息。

**表 4-1. 支持的 SAN 阵列**

制造商	设备	参考
EMC	CLARiiON 存储系统。 也可由 FSC 提供。 Dell, Inc. 同时有售，相应为 Dell/EMC FC RAID 阵列产品系 列。	“ <a href="#">EMC CLARiiON 存储系统</a> ”（第 52 页）。
	Symmetrix 存储系统。	“ <a href="#">EMC Symmetrix 存储系统</a> ”（第 53 页）。

表 4-1. 支持的 SAN 阵列

制造商	设备	参考
IBM	IBM TotalStorage DS 4000 系统 (之前称为 FASTT 存储系统)。 也可由 LSI Eugenio 和 StorageTek 提供。	<a href="#">“IBM TotalStorage DS4000 存储系统”</a> (第 54 页)。
	IBM TotalStorage Enterprise 存储 系统 (即先前的 Shark 存储系 统)。	<a href="#">“IBM TotalStorage 8000”</a> (第 58 页)。
Hewlett Packard	HP StorageWorks (MSA、EVA 和 XP)。	<a href="#">“HP StorageWorks 存储系统”</a> (第 58 页)。
Hitachi	Hitachi Data Systems 存储器。 也可由 Sun 提供, 并且等同于 HP XP。	<a href="#">“Hitachi Data Systems 存储器”</a> (第 60 页)。
Network Appliance	Network Appliance FC SAN 存储 解决方案。	<a href="#">“Network Appliance 存储器”</a> (第 61 页)。

## 常规设置注意事项

对于所有存储阵列, 请确保满足以下要求:

- LUN 对每台主机的每个 HBA 呈现的 LUN ID 号都必须相同。如果使用了不同的号码, 则 ESX Server 主机便识别不出指向同一 LUN 的不同路径。  
  
有关如何配置相同 SAN LUN ID 的说明因供应商而异, 因此请参考存储阵列文档了解详细信息。
- 除非是为本章所讨论的个别存储阵列指定主机类型, 否则请将向 ESX Server 显示的 LUN 的主机类型设置为 `Linux`、`Linux Cluster`, 或者设置为 `vmware` 或 `esx` (如果可用)。
- 如果使用了 VMotion、DRS 或 HA, 则请确保虚拟机的源主机与目标主机都能看到具有相同 LUN ID 的同一 LUN。

出于对可能出现数据错误的担心, SAN 管理员可能认为使多个主机看见相同的 LUN 不合常规。但是, VMFS 可防止多个虚拟机同时对同一文件进行写入, 因此给所有需要的 ESX Server 系统置备 LUN 是可行的。

## EMC CLARiiON 存储系统

EMC CLARiiON 存储系统可与采用 SAN 配置的 ESX Server 计算机配合使用。基本配置步骤包括：

- 1 安装并配置存储设备。
- 2 在交换机级别配置分区。
- 3 创建 RAID 组。
- 4 创建并绑定 LUN。
- 5 注册 SAN 所连接的服务器。
- 6 创建包含服务器和 LUN 的存储组。

使用 EMC 软件执行配置。请参见 EMC 文档。

此阵列是主动 / 被动磁盘阵列，因此适用于以下相关问题。

为避免发生路径抖动的可能，默认多路径策略为 **[最近使用 (Most Recently Used)]**，而非 **[固定的 (Fixed)]**。ESX Server 系统在识别阵列时设置默认策略。请参见 [“解决路径抖动”](#)（第 98 页）。

AX100 存储设备不支持自动对卷重新签名。请参见 [“VMFS 卷重新签名”](#)（第 106 页）。

要使用从 SAN 引导，请确保在 HBA BIOS 中选择主动 SP 作为引导 LUN 的目标。

## EMC CLARiiON AX100 与 RDM

在 EMC CLARiiON AX100 系统上，只有在使用 Navisphere Management Suite 进行 SAN 管理时才支持 RDM。不保证 Navilight 能正常运作。

要成功使用 RDM，给定的 LUN 必须对群集中所有 ESX Server 主机呈现相同的 LUN ID。默认情况下，AX100 不支持此配置。

## AX100 非活动连接的显示问题

使用直接连接 ESX Server 系统的 AX100 FC 存储设备时，必须验证所有连接均可工作，并将任何不再使用的连接注销。否则 ESX Server 将无法发现新的 LUN 或路径。

请考虑以下情况：

- 1 ESX Server 系统直接连接 AX100 存储设备。ESX Server 有两个 FC HBA。其中一个 HBA 先前已在存储阵列注册并且已对其 LUN 进行了配置，但现在连接为非活动状态。
- 2 将 ESX Server 主机上的另一 HBA 连接到 AX100 并进行注册后，ESX Server 主机可正确显示阵列具有一个活动连接。但是，先前配置给 ESX Server 主机的任何 LUN 均不可见，即使反复地重新扫描也是如此。

要解决此问题，请移除非活动的 HBA、注销非活动 HBA 的连接，或使所有非活动连接处于活动状态。这会使得存储组中仅包含活动的 HBA。完成此更改后，通过重新扫描添加已配置的 LUN。

## 将主机配置更改应用于阵列

使用 AX100 存储阵列时，并无主机代理定期检查主机配置并将更改应用于阵列。`axnaviserverutil cli` 实用程序将用于更新所作的更改。此为手动操作并应根据需要执行。

## EMC Symmetrix 存储系统

要在 Symmetrix 网络存储系统上执行 ESX Server 操作，需要以下设置：

- 公共序列号 (Common serial number, C)
- 启用自动协商 (Auto negotiation) (EAN)
- 在此端口上启用 Fibrepath (VCM)
- SCSI 3 (SC3) 集 (启用)
- 唯一的全球名称 (Unique world wide name, UWN)
- 需要 SPC-2 (Decal) (SPC2) SPC-2 标记

使用 EMC 软件配置存储阵列。请参见 EMC 文档。

ESX Server 主机将所有来自 Symmetrix 存储阵列、容量为 50 MB 或更小的所有 LUN 视作管理 LUN。这些 LUN 也称为伪 LUN (pseudo LUN) 或网关 LUN (gatekeeper LUN)。这些 LUN 可显示在 EMC Symmetrix Management Interface 中，并且不应用于保存数据。

## IBM TotalStorage DS4000 存储系统

IBM TotalStorage DS4000 系统曾叫作 IBM FASTT。许多存储阵列供应商（包括 LSI 和 StorageTek）都生产兼容 DS4000 的 SAN 存储阵列。

请参见 IBM 红皮书 *Implementing VMware ESX Server with IBM TotalStorage FASTT* 《实现 VMware ESX Server 与 IBM TotalStorage FASTT 的搭配使用》，网址为：<http://www.redbooks.ibm.com/redbooks/pdfs/sg246434.pdf>。此部分总结了如何配置 IBM TotalStorage 存储系统以使用 SAN 和 Microsoft 群集服务。请参见《Microsoft 群集服务的设置》。

除 IBM TotalStorage 存储系统的正常配置步骤外，您还需执行特定任务。

您必须确保多路径策略设置为 **[最近使用 (Most Recently Used)]**。请参见“[查看当前的多路径状况](#)”（第 85 页）。

### 配置用于 DS4000 存储服务器 SAN 故障切换的硬件

要结合配备了两个存储处理器的 DS4000 存储模块设置具有高可用性的 SAN 故障切换配置，您需要以下硬件组件：

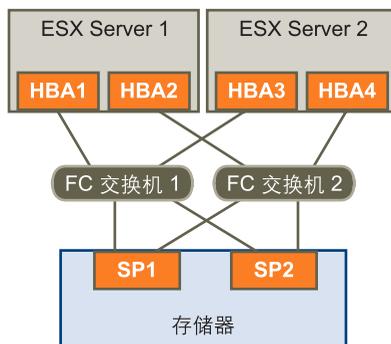
- 每台 ESX Server 计算机上两个 FC HBA，如 QLogic 或 Emulex。
- 两台 FC 交换机，将 HBA 连接到 SAN（例如，FC 交换机 1 和 FC 交换机 2）。
- 两个 SP（例如，SP1 和 SP2）。

每个 SP 必须至少有两个连接 SAN 的端口。

使用以下 ESX Server 主机连接设置，如图 4-1 所示：

- 将每台 ESX Server 计算机上的每个 HBA 各自连接不同交换机。例如，使 HBA1 连接 FC switch 1，HBA2 连接 FC switch 2。
- 在 FC switch 1 上，SP1 连接的交换机端口号低于 SP2，从而确保 SP1 先列出。例如，SP1 连接 FC switch 1 端口 1，SP2 连接 FC switch 1 端口 2。
- 在 FC switch 2 上，SP1 连接的交换机端口号低于 SP2，从而确保 SP1 先列出。例如，SP1 连接 FC switch 2 上的端口 1，SP2 连接 FC switch 2 上的端口 2。

图 4-1. SAN 故障切换



此配置对每个 HBA 提供两条路径，使得每条连接均能在发生故障时切换到冗余路径。此配置中的路径顺序可提供 HBA 和交换机故障切换，无需触发 SP 故障切换。LUN 所有权必须属于首选路径所连的存储处理器。在上述示例配置中，SP1 拥有 LUN 的所有权。

---

**注意** 上述示例假定交换机并非通过一个架构中的交换机内链路 (Inter-Switch Link, ISL) 连接。

---

## 验证存储处理器端口配置

可以通过对比 VI Client 信息和 DS4000 子系统配置文件中的信息来验证 SP 端口配置。

### 验证存储处理器端口配置

- 1 使用 VI Client 连接到 ESX Server 主机。
- 2 选择主机并选择 [ **配置 (Configuration)** ] 选项卡。
- 3 在 [ **硬件 (Hardware)** ] 面板中单击 [ **存储适配器 (Storage Adapters)** ]。

- 4 选择各存储适配器以查看其 WWPN。

**存储适配器** 重新扫描

设备	类型	SAN 标识符
<b>QLA231x/2340</b>		
vmhba1	光纤通道	21:00:00:e0:8b:89:a0
vmhba2	光纤通道	21:01:00:e0:8b:ab:89:ac
<b>Smart Array 6i</b>		
vmhba0	块 SCSI	
<b>iSCSI Software Adapter</b>		

**详细信息**

**vmhba2**

型号: QLA231x/2340

WWPN: 21:01:00:e0:8b:ab:89:ac

目标: 2

**SCSI 目标 0** 隐藏 LUN

路径	规范路径	类型	容量	LUN ID
vmhba2:0:0	vmhba1:1:0	disk	200.00 GB	0
vmhba2:0:1	vmhba1:0:1	disk	100.00 GB	1
vmhba2:0:2	vmhba1:0:2	disk	50.00 GB	2
vmhba2:0:3	vmhba1:0:3	disk	50.00 GB	3

**SCSI 目标 1** 隐藏 LUN

- 5 选择 [ 存储器 (Storage) ] 查看可用数据存储。

对比 WWPN 信息与 DS4000 存储子系统配置文件中列出的信息。

## 禁用自动卷传输

为避免发生路径抖动的可能，请在 SAN 存储处理器上禁用自动卷传输 (Auto Volume Transfer, AVT)。如果启用了 AVT，那么两个存储处理器在特定情况下可交替取得 LUN 的所有权，这会致使性能降低。AVT 也称为 ADT (自动磁盘传输, Auto Disk Transfer)。

请参见“解决路径抖动”(第 98 页)。

要禁用 AVT，在 DS 4000 Storage Manager 中，对于包含一台或多台 ESX Server 计算机的 HBA 的各主机组，请为其中定义的每个端口将主机类型设置为 LNXCL 或 VMware (用于更高版本)。

---

**注意** 更改 AVT 配置后必须重新引导 ESX Server 主机。

---

## 配置存储处理器检测数据

存储处理器可配置为在静态时返回 Unit Attention 或 Not Ready 消息。运行 Windows 做为客户操作系统的 DS4000 SP 在静态时应返回 Not Ready 检测数据 (sense

data)。返回 Unit Attention 可能导致 Windows 客户操作系统在故障切换期间发生故障。

### 配置存储处理器返回 Not Ready 检测数据

- 1 在 shell 程序窗口中，使用以下命令确定 LNXCL 主机类型的索引号：

输入每条命令后均按 Enter。

```
SMcli.exe <ip-addr-for-SPA> show hosttopology; <Enter>
```

```
SMcli.exe <ip-addr-for-SPB> show hosttopology; <Enter>
```

以下命令假定 13 是 NVSRAM 主机类型定义中与 LNXCL 对应的索引号。如果存储处理器的 LNXCL 对应的索引号不同，请将其替换以下命令中的 13。

- 2 对 SPA 执行下述命令使其返回 Not Ready 检测数据。

仅在输入所有命令后按 Enter。

```
SMcli.exe <ip-addr-for-SPA>
```

```
set controller [a] HostNVSARAMBYTE [13,0x12]=0x01;
```

```
set controller [a] HostNVSARAMBYTE [13,0x13]=0x00;
```

```
reset Controller [a];
```

```
<Enter>
```

- 3 对 SPB 执行下述命令使其返回 Not Ready 检测数据。

```
SMcli.exe <ip-addr-for-SPB>
```

```
set controller [b] HostNVSARAMBYTE [13,0x12]=0x01;
```

```
set controller [b] HostNVSARAMBYTE [13,0x13]=0x00;
```

```
reset Controller [b];
```

```
<Enter>
```

---

**注意** 如果使用 DS4000 Storage Manager GUI，请将用于两个存储处理器的配置命令粘贴到单个脚本中，并同时对二者进行配置。如果使用 SMcli.exe，请分别连接每个 SP。

---

## IBM TotalStorage DS4000 与路径抖动

在 DS 4000 或兼容的 SAN 阵列上检测到路径抖动时，以下警告会记录到 vmkernel 日志。

```
FAStT SAN is path thrashing with another system. Check AVT setting.
```

## IBM TotalStorage 8000

IBM TotalStorage 8000 系统使用主动 / 主动阵列，无需特殊配置即可与 VMware ESX Server 一起使用。

要成功使用 RDM，给定的 LUN 需对群集中所有 ESX Server 主机呈现相同的 LUN ID。

在 TotalStorage Configuration Management 工具中，选择 **[对源和目标中的 LUN 使用相同 ID (Use same ID for LUN in source and target)]**。

IBM TotalStorage 8000 系统不支持自动执行重新签名。

---

**注意** 如果要为 ESX Server 主机配置为从 IBM TotalStorage 8000 阵列上的某个 LUN 使用从 SAN 引导，请在安装完成之前禁用相应刀片的内部光纤端口。

---

## HP StorageWorks 存储系统

本节包含用于不同 HP StorageWorks 存储系统的配置信息。

有关其他信息，请参见 HP 网站上有关 VMware ESX Server 的 HP ActiveAnswers 部分。

### HP StorageWorks MSA

本节列出使用 HP StorageWorks MSA 主动 / 被动版本时用户所关心的问题。

#### 将 Profile Name 设置为 Linux

要将 HP StorageWorks MSA 1000 和 MSA 1500 与 ESX Server 系统配合使用，请配置 SAN 阵列与 ESX Server 主机之间的 FC 连接，将 Profile Name 设置为 **Linux**。

#### 设置连接的 Profile Name

- 1 使用 MSA 1000 命令行界面在 MSA 1000 上创建静态连接。  
有关安装和配置命令行界面的信息，请参见 HP StorageWorks MSA 1000 文档。

---

**注意** 无法使用 HP Array Configuration 实用程序创建连接设置。

---

- 2 将 MSA 1000 命令行界面连接至 MSA 1000。
- 3 验证 MSA 1000 与 ESX Server 主机之间的 FC 网络运行正常。
- 4 启动命令行界面，并在提示符处输入以下命令：

**SHOW CONNECTIONS**

输出显示 MSA 1000 所连各个 FC WWNN 和 WWPNN 的连接规范：

```
Connection Name: <unknown>
Host WWNN = 20:02:00:a0:b8:0c:d5:56
Host WWPNN = 20:03:00:a0:b8:0c:d5:57
Profile Name = Default
Unit Offset 0
Controller 1 Port 1 Status = Online
Controller 2 Port 1 Status = Online
```

- 5 确保主机的 WWNN 和 WWPNN 显示 ESX Server 计算机上每个 FC 适配器的正确连接。
- 6 创建静态连接，如下所示：

```
ADD CONNECTION ESX_CONN_1 WWNN=20:02:00:a0:b8:0c:d5:56
WWPN=20:03:00:a0:b8:0c:d5:57 PROFILE=LINUX
```

- 7 验证连接，如下所示：

```
SHOW CONNECTIONS
```

输出显示单个连接，其 WWNN 和 WWPNN 对为 20:02:00:a0:b8:0c:d5:56 和 20:03:00:a0:b8:0c:d5:57，且 Profile Name 设置为 Linux：

```
Connection Name: ESX_CONN_1
Host WWNN = 20:02:00:a0:b8:0c:d5:56
Host WWPNN = 20:03:00:a0:b8:0c:d5:57
Profile Name = Linux
Unit Offset = 0
Controller 1 Port 1 Status = Online
Controller 2 Port 1 Status = Online
```

---

**注意** 确保 WWNN = 20:02:00:a0:b8:0c:d5:56 与 WWPNN = 20:03:00:a0:b8:0c:d5:57 显示单个连接。

---

对于 WWNN= 20:02:00:a0:b8:0c:d5:56 和 WWPNN = 20:03:00:a0:b8:0c:d5:57，此时不应再有 Connection Name 为 unknown 的连接。

- 8 为 ESX Server 主机上的每个 WWNN 和 WWPNN 添加静态连接（使用不同的连接名称值）。

### 集线器控制器问题

ESX Server 系统在使用 MSA 集线器控制器的情况下可能无法正常工作。请改用 2/8 内部交换机或单端口控制器。

## HP StorageWorks EVA

HP StorageWorks EVA 系统分为两种类型，如下所示：EVA\_GL，此为主动 / 被动系统；EVA\_XL，此为主动 / 主动系统。

向 ESX Server 主机显示 LUN 时请将连接类型设置为 Custom。请采用以下值之一：

- 对于 HP EVAgl 3000/5000（主动 / 被动），请使用主机模式类型 000000002200282E。
- 对于 HP EVAgl 固件 4.001（GL 系列的主动 / 主动固件）及更高版本，请使用主机模式类型 VMware。
- 对于所用固件版本低于 5.031 的 EVA4000/6000/8000 主动 / 主动阵列，请使用主机模式类型 000000202200083E。
- 对于使用 5.031 及更高版本固件的 EVA4000/6000/8000 主动 / 主动阵列，则请使用主机模式类型 VMware。

对于其他情况，EVA 系统无需特殊的配置更改即可与 ESX Server 系统配合使用。

请参见“VMware Infrastructure 3 与 HP StorageWorks 最佳配置方案”，网址为：  
[http://h71019.www7.hp.com/ActiveAnswers/downloads/VMware3\\_StorageWorks\\_BestPractice.pdf](http://h71019.www7.hp.com/ActiveAnswers/downloads/VMware3_StorageWorks_BestPractice.pdf)。

## HP StorageWorks XP

对于 HP StorageWorks XP，请将主机模式设置为 Windows（而非 Linux）。Hitachi Data Systems 也提供此系统。

## Hitachi Data Systems 存储器

本节介绍 Hitachi Data Systems 存储器的相关设置。Sun 也可提供此解决方案，且该方案可由 HP XP 存储器实现。

- **LUN 屏蔽** - 要屏蔽 ESX Server 主机上的 LUN，最好使用 HDS Storage Navigator 软件。
- **微码与配置** - 请咨询 HDS 代表获取与 ESX Server 协同操作所需的准确配置及微码级别。如果微码不受支持，则通常不能与 ESX Server 进行交互。
- **模式** - 模式设置具体取决于所使用的机型，例如：
  - 9900 和 9900v 使用 Netware 主机模式。
  - 9500v 系列使用 Hostmode1: standard 及 Hostmode2: SUN Cluster。

有关此处未列出机型的主机模式设置，请咨询 HDS 代表。

## Network Appliance 存储器

配置 Network Appliance 存储设备时，首先为存储阵列设置适当的 LUN 类型及启动器组类型：

- **LUN 类型** - VMware（如果 VMware 类型不可用，请使用 Linux）
- **启动器组类型** - VMware（如果 VMware 类型不可用，请使用 Linux）

随后必须置备存储器。

### 从 Network Appliance 存储设备置备存储器

- 1 如果需要，可使用 CLI 或 FilerView GUI 创建集合 (Aggregate):

```
aggr create <vmware-aggr> <number of disks>
```

- 2 创建灵活卷 (Flexible Volume):

```
vol create <aggregate name> <volume size>
```

- 3 创建 Qtree 存储各个 LUN:

```
qtree create <path>
```

- 4 创建 LUN:

```
lun create -s <size> -t vmware <path>
```

- 5 创建启动器组:

```
igroup create -f -t vmware <igroup name>
```

- 6 将 LUN 映射到刚创建的启动器组:

```
lun map (<path>) <igroup name> <LUN ID>
```

有关 Network Appliance 存储器如何与 VMware 技术配合使用的其他信息，请参见以下 Network Appliance 文档：

- “Network Appliance & VMware ESX Server: Instantaneous Backup & Recovery with NetApp Snapshot Technology（使用 NetApp 快照技术实时备份与恢复）”，网址为：<http://www.netapp.com/library/tr/3428.pdf>。
- “Technical Case Study: Using a Network Appliance SAN with VMware to Facilitate Storage and Server Consolidation（技术案例研究：结合使用 Network Appliance SAN 和 VMware 实现存储器和服务器的整合）”，网址为：<http://www.netapp.com/library/tr/3401.pdf>。



# 对 ESX Server 系统使用从 SAN 引导

---

# 5

本章讨论从 SAN 引导的优点，并介绍将 ESX Server 引导映像存储到 SAN LUN 上时需要执行的任务。

---

**注意** 如果不计划从 SAN 引导 ESX Server 主机，则请跳过本章。

---

本章将讨论以下主题：

- “[从 SAN 引导概述](#)”（第 64 页）
- “[准备从 SAN 引导](#)”（第 65 页）
- “[为从 SAN 引导设置 FC HBA](#)”（第 68 页）

## 从 SAN 引导概述

考虑如何设置系统从 SAN 引导之前，请先判断这对您的环境而言是否可行。

对于以下情况，可使用从 SAN 引导：

- 不想维护本地存储器时。
- 需要对服务控制台进行简单克隆时（仅限 ESX Server 3）。
- 在无磁盘硬件配置中，例如，某些刀片系统上。

对于以下情况，请勿使用从 SAN 引导：

- 正在使用 Microsoft 群集服务时。
- 当服务控制台与 VMkernel 之间可能发生 I/O 冲突时（仅限 ESX Server 3）。

---

**注意** 对于 ESX Server 2.5，从 SAN 引导无法与 RDM 一起使用。而 ESX Server 3.x 中则取消了这一限制。

---

## 从 SAN 引导如何运作

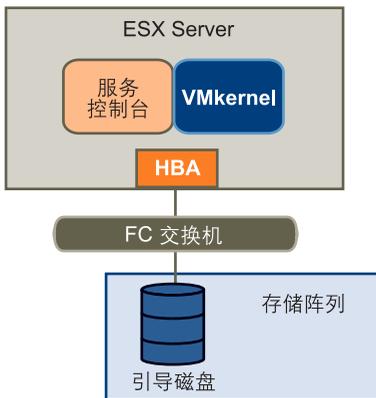
将系统设置为从 SAN 引导时，引导映像并非存储在 ESX Server 系统的本地磁盘上，而是存储于 SAN LUN，如图 5-1 所示。

---

**注意**

---

图 5-1. 从 SAN 引导如何运作



在设置为从 SAN 引导的系统上：

- HBA BIOS 必须指定 FC 卡作为引导控制器。请参见“为从 SAN 引导设置 FC HBA”（第 68 页）。
- 必须配置 FC 卡以启动与目标引导 LUN 的原始连接。

## 从 SAN 引导的优点

在从 SAN 引导的环境中，会在 SAN 阵列中一个或多个 LUN 上安装操作系统。服务器会得到有关引导映像位置的通知。服务器启动之后，即从 SAN 阵列上的 LUN 进行引导。

---

**注意** 对 VMware ESX Server 系统使用从 SAN 引导时，每台服务器必须各自拥有引导 LUN。

---

从 SAN 引导可带来许多好处，其中包括：

- **服务器成本更低** - 无需内部存储器，可更密集地安置服务器，且运行过程散热也更少。
- **服务器更换更方便** - 可以更换服务器并将新服务器指向旧的引导位置。
- **减少了空间浪费。**
- **备份过程更轻松** - 可作为 SAN 整体备份过程的一部分来备份 SAN 中的系统引导映像。
- **改善了管理** - 创建和管理操作系统映像变得更简单且更高效。

## 准备从 SAN 引导

除了常规的 ESX Server 与 SAN 连用相关配置任务，请完成以下任务使得 ESX Server 主机能够从 SAN 引导。

### 实现从 SAN 引导

1 确保配置设置满足从 SAN 引导的基本要求。

2 准备硬件元素。

其中包括 HBA、网络设备和存储系统。请参阅各设备的产品文档。

3 配置 SAN 上的 LUN 屏蔽。

这将确保每个 ESX Server 主机都有专用的 LUN 作引导分区。引导 LUN 必须专用于单台服务器。

#### 4 选择诊断分区的位置。

诊断分区可与引导分区置于相同的 LUN 上。核心转储存储在诊断分区中。

本节剩余部分列出需要完成哪些任务才能从 SAN 成功引导 ESX Server 计算机。

## 开始前

请了解以下内容：

- 所需设置类型的相关建议或样本设置：
  - 指向引导 LUN 的单一路径或冗余路径。
  - FC 交换机架构。
  - 适用于所用存储阵列类型的任何特定建议。
- 限制和要求，其中包括：
  - 从 SAN 引导的限制。
  - 有关要用于从 SAN 引导的存储阵列的供应商建议。
  - 有关从 SAN 引导的服务器的供应商建议。
- 使用以下方法之一查找引导路径 HBA 的 WWN：
  - 在引导刚开始时进入 FC HBA BIOS。
  - 在物理卡上找到 WWN。其与 MAC 地址类似。

## 从 SAN 引导模式中的 LUN 屏蔽

正确的 LUN 屏蔽对从 SAN 引导模式非常重要。

- 每台服务器只能看到其自身的引导 LUN，而看不到其他服务器的引导 LUN。
- 多台服务器可以共享一个诊断分区。这点可以通过使用 LUN 屏蔽来实现。请参见[“共享诊断分区”](#)（第 95 页）。

## 准备 SAN

本节列出的步骤用于准备从 SAN 引导的 SAN 存储阵列。除步骤 1 和 2 外，步骤 3-7 均针对于从 SAN 引导。

### 准备 SAN

- 1 连接 FC 和以太网线缆，请参阅适用于您的设置的任何线缆连接指南。  
检查 FC 交换机接线（如果有）。
- 2 配置存储阵列。
  - a 在 SAN 存储阵列中，使 ESX Server 主机对于 SAN 为可见。这通常称为创建对象。
  - b 在 SAN 存储阵列中，设置 ESX Server 主机，将主机 FC 适配器的 WWPN 作为端口名称或节点名称。
  - c 创建 LUN。
  - d 分配 LUN。
  - e 记录 FC 交换机和存储阵列的 IP 地址。
  - f 记录涉及到的各个 SP 及主机适配器的 WWPN。



**小心** 如果使用脚本式安装在从 SAN 引导模式下安装 ESX Server，需要执行特殊步骤来避免意外丢失数据。请参见 VMware 知识库文章 1540，网址为 [http://www.vmware.com/support/kb/enduser/std\\_adp.php?p\\_faqid=1540](http://www.vmware.com/support/kb/enduser/std_adp.php?p_faqid=1540)。

- 3 为从 SAN 引导配置 HBA BIOS，如以下各节中所述：
  - “为从 SAN 引导设置 QLogic FC HBA”（第 68 页）。
  - “为从 SAN 引导设置 Emulex FC HBA”（第 70 页）
- 4 从 ESX Server 安装 CD 引导 ESX Server 系统。  
请参见《安装指南》。

QLogic BIOS 使用路径 (wwpn:lun) 搜索列表查找引导映像。如果某一 wwpn:lun 路径与一条被动路径相关联（如使用 CLARiiON 或 IBM TotalStorage DS 4000 系统的情况），BIOS 将采用该被动路径且不查找主动路径。如果从 SAN LUN 引导 ESX Server 系统，则引导会在主机尝试访问被动路径时失败。

## 尽可能减少启动器数目

确保分区包含尽量少的主机和存储端口。如果同一分区内包含若干其他启动器而尝试选择引导 LUN，则 Emulex 和 QLogic BIOS 可能变得没有响应。

例如，如果一个分区内有 15 个启动器和 4 个 Symmetrix 端口，则可能因 Emulex 或 QLogic BIOS 没有响应而无法从中选择引导设备。如果将两个主机端口作为一个分区，从这一分区仅看到四个存储端口，那么便可以选择引导 LUN。

## 为从 SAN 引导设置 FC HBA

本节讨论如何设置 HBA。

### 为从 SAN 引导设置 QLogic FC HBA

配置 QLogic HBA BIOS 实现从 SAN 引导 ESX Server 包括以下任务：

---

**注意** 如果正在使用 IBM BladeCenter，则请从服务器断开所有本地磁盘驱动器。

---

#### 启用 QLogic HBA BIOS

配置 QLogic HBA BIOS 实现从 SAN 引导 ESX Server 时，请从启用 QLogic HBA BIOS 开始。

#### 启用 QLogic HBA BIOS

- 1 进入 BIOS Fast!UTIL 配置实用程序：
  - a 引导服务器。
  - b 引导服务器时，按 Ctrl-Q。
- 2 根据 HBA 的数目执行适当的操作。
  - 如果只有一个主机总线适配器 (HBA)，此时将出现 [Fast!UTIL 选项 (Fast!UTIL Options)] 页面。请跳至 [步骤 3](#)。
  - 如果有多个 HBA，请手动选择 HBA：
    - a 在 [选择主机适配器 (Select Host Adapter)] 页面中，使用箭头键将光标定位到适当的 HBA 上。
    - b 按 Enter。
- 3 在 [Fast!UTIL 选项 (Fast!UTIL Options)] 页面中，选择 **[配置设置 (Configuration Settings)]** 并按 Enter。

- 4 在 [ 配置设置 (Configuration Settings)] 页面中, 选择 [ **主机适配器设置 (Host Adapter Settings)**] 并按 Enter。
- 5 设置 BIOS 来搜索 SCSI 设备:
  - a 在 [ 主机适配器设置 (Host Adapter Settings)] 页面中, 选择 [ **主机适配器 BIOS (Host Adapter BIOS)**]。
  - b 按 Enter 将值切换至 [ **已启用 (Enabled)**]。
  - c 按 Esc 退出。

### 启用可选式引导

需要启用可选式引导。

### 启用可选式引导

- 1 选择 [ **可选式引导设置 (Selectable Boot Settings)**] 并按 Enter。
- 2 在 [ 可选式引导设置 (Selectable Boot Settings)] 页面中, 选择 [ **可选式引导 (Selectable Boot)**]。
- 3 按 Enter 将值切换至 [ **已启用 (Enabled)**]。

### 选择引导 LUN

如果使用主动 / 被动存储阵列, 选定的 SP 必须在指向引导 LUN 的首选 (主动) 路径上。如果不确定哪个 SP 在主动路径上, 请使用存储阵列管理软件进行查找。目标 ID 由 BIOS 创建, 并可能随每次重新引导而变化。

### 选择引导 LUN

- 1 使用光标键选择存储处理器列表中的第一个条目。
- 2 按 Enter 打开 [ 选择光纤通道设备 (Select Fibre Channel Device)] 页面。
- 3 使用光标键选择所选 SP 并按 Enter。
  - 如果 SP 仅连接一个 LUN, 则该 LUN 将被选为引导 LUN, 并可以跳至 [步骤 4](#)。
  - 如果 SP 连接多个 LUN, 此时将打开 [ 选择 LUN (Select LUN)] 页面。使用箭头键定位到所选 LUN 并按 Enter。

如果列表中显示任何其他存储处理器, 请定位到相应条目并按 C 清除数据。

- 4 按两次 Esc 退出。
- 5 按 Enter 保存设置。

## 将系统设置为先从 CD-ROM 引导

因为 VMware 安装 CD 已放入 CD-ROM 驱动器，所以请将系统设置为先从 CD-ROM 引导。为此，请在系统 BIOS 设置中更改系统引导顺序。

例如，在 IBM X-Series 345 服务器上，请执行以下操作：

- 1 在系统启动过程中，进入系统 [BIOS 配置 / 设置实用程序 (BIOS Configuration/Setup Utility)]。
- 2 选择 [启动选项 (Startup Options)] 并按 Enter。
- 3 选择 [启动顺序选项 (Startup Sequence Options)] 并按 Enter。
- 4 将 [第一启动设备 (First Startup Device)] 更改为 [CD-ROM]。

现在即可按 《安装指南》中所述安装 ESX Server 系统。

## 为从 SAN 引导设置 Emulex FC HBA

配置 Emulex HBA BIOS 实现从 SAN 引导 ESX Server 包括以下任务：

- 启用 [BootBIOS 提示](#)
- 启用 BIOS

### 启用 BootBIOS 提示

- 1 在 ESX Server 服务控制台或 Linux 命令提示符中，运行 `lputil`。

---

**注意** 考虑到是从装入 Emulex 驱动器的 Linux Administration CD 引导 ESX Server 主机，因而在此处运行 `lputil`。

---

- 2 选择 <3> [固件维护 (Firmware Maintenance)]。
- 3 选择适配器。
- 4 选择 <6> [引导 BIOS 维护 (Boot BIOS Maintenance)]。
- 5 选择 <1> [启用引导 BIOS (Enable Boot BIOS)]。

### 启用 BIOS

- 1 重新引导 ESX Server 计算机。
- 2 在 Emulex 提示处按 <ALT+E>。
  - a 选择适配器（支持 BIOS）。
  - b 选择 <2> [配置适配器参数 (Configure Adapter's Parameters)]。

- c 选择 <I> [**启用或禁用 BIOS (Enable or Disable BIOS)**]。
  - d 选择 <I> 启用 BIOS。
  - e 选择 <X> 退出，选择 <N> 返回主菜单。
- 3 在 Emulex 主菜单中：
- a 选择同一适配器。
  - b 选择 <I> [**配置引导设备 (Configure Boot Devices)**]。
  - c 选择引导条目的位置。
  - d 输入表示引导设备的两位数。
  - e 输入表示启动 LUN 的两位数 (HEX) (例如, 08)。
  - f 选择引导 LUN。
  - g 选择 <I> [**WWPN**]。(使用 WWPN 而非 DID 引导此设备)。
  - h 选择 <X> 退出，选择 <Y> 重新引导。
- 4 引导至系统 BIOS，将 Emulex 移到引导控制器顺序的第一位。
- 5 重新引导并在 SAN LUN 上安装。



# 管理使用 SAN 存储器的 ESX Server 系统

---

# 6

本章可在管理 ESX Server 系统、有效使用 SAN 存储器以及故障排除方面为您提供帮助。本章将讨论以下主题：

- [“问题与解决方案”](#)（第 74 页）
- [“避免问题发生的准则”](#)（第 75 页）
- [“获取信息”](#)（第 75 页）
- [“解决显示问题”](#)（第 77 页）
- [“高级 LUN 显示配置”](#)（第 79 页）
- [“N-Port ID 虚拟化”](#)（第 81 页）
- [“多路径”](#)（第 85 页）
- [“故障切换”](#)（第 92 页）
- [“VMkernel 配置”](#)（第 95 页）
- [“共享诊断分区”](#)（第 95 页）
- [“避免问题和解决问题”](#)（第 95 页）
- [“优化 SAN 存储性能”](#)（第 96 页）
- [“解决性能问题”](#)（第 98 页）
- [“SAN 存储器备份注意事项”](#)（第 103 页）
- [“分层应用程序”](#)（第 105 页）
- [“VMFS 卷重新签名”](#)（第 106 页）

## 问题与解决方案

表 6-1 列出了最常遇到的问题，并说明了如何解决这些问题，或指出了论述相应问题的各节位置。

**表 6-1. 问题与解决方案**

问题	解决方案
LUN 在 VI Client 中不可见。	请参见 <a href="#">“解决显示问题”</a> （第 77 页）。
共享 LUN 及在其上格式化的 VMFS 文件系统对于访问该 LUN 的所有 ESX Server 主机均不可见。	请参见 <a href="#">“解决有关 LUN 不可见的问题”</a> （第 78 页）。
了解路径故障切换如何执行或对路径故障切换的执行方式进行更改。	您可以通过 VI Client 执行相应操作。请参见 <a href="#">“多路径”</a> （第 85 页）。
查看或更改当前的多路径策略或首选路径，或者禁用或启用路径。	您可以通过 VI Client 执行相应操作。请参见 <a href="#">“多路径”</a> （第 85 页）。
增加 Windows 磁盘超时，避免故障切换期间发生中断。	请参见 <a href="#">“设置操作系统超时”</a> （第 94 页）。
定制 QLogic 或 Emulex HBA 的驱动程序选项。	请参见 <a href="#">“设置 SCSI 控制器的设备驱动程序选项”</a> （第 94 页）。
服务器无法访问 LUN，或者访问速度很慢。	问题可能在于路径抖动。请参见 <a href="#">“解决路径抖动”</a> （第 98 页）。
访问速度较慢。	<p>如果您的 VMFS 数据存储数目较多，并且均为 VMFS-3，可在命令行提示符处键入以下命令卸载 VMFS-2 驱动程序：</p> <pre>vmkload_mod -u vmfs2</pre> <p>您将发现刷新数据存储及重新扫描存储适配器等管理操作的执行速度有显著提高。</p> <p>同样，如果未使用 NFS 数据存储，可以键入以下命令卸载 NFS 驱动程序：</p> <pre>vmkload_mod -u nfsclient</pre>
您向存储器添加了新的 LUN 或新路径并希望其在 VI Client 中可见。	您必须重新扫描。请参见 <a href="#">“使用重新扫描”</a> （第 78 页）。

## 避免问题发生的准则

请遵循下述准则避免潜在的问题：

- 每个 LUN 上仅放置一个 VMFS 卷。不建议多个 VMFS 卷位于同一 LUN 上。
- 请勿更改系统预设的路径策略。特别是，使用主动 / 被动阵列的同时将路径策略设置为 **[ 固定的 (Fixed) ]** 可导致路径抖动。

## 获取信息

本节说明如何查找有关 HBA、状态及多路径等项目的信息。如果在执行这些任务时遇到问题，请参见 [“解决显示问题”](#)（第 77 页）。

### 查看 HBA 信息

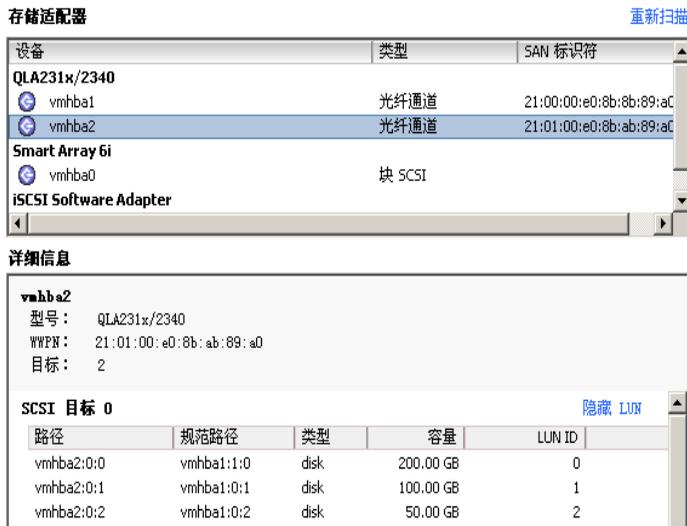
使用 VI Client 显示所有可用存储适配器及其信息。

#### 查看 HBA 类型列表

- 1 选择要查看其 HBA 的主机并单击 **[ 配置 (Configuration) ]** 选项卡。

您可以从 **[ 摘要 (Summary) ]** 选项卡中查看所有存储设备的列表。但从此处看不到详细信息，也无法管理设备。

- 在 [ 硬件 (Hardware)] 面板中，选择 [ 存储适配器 (Storage Adapters)]。此时会显示存储适配器列表。您可以选择每个适配器获取详细信息。



## 查看数据存储信息

使用 VI Client 显示所有格式化的数据存储并查看特定数据存储的详细信息。

### 查看所有存储设备及其相关详细信息

- 选择要查看其存储设备的主机并单击 [ 配置 (Configuration)] 选项卡。
- 在 [ 硬件 (Hardware)] 面板中，选择 [ 存储器 (Storage)]。

此时 [ 存储器 (Storage)] 面板中会显示数据存储列表。

屏幕将显示所选主机的整个 VMFS。只有使用 VMFS 进行格式化的存储器出现在显示屏幕中。

- 要查看有关任一数据存储的详细信息，请将其选中。

[ 详细信息 (Details)] 面板将显示其他信息。其中包括位置和容量、路径数目、路径策略及属性。同时还包括扩展信息。

一个扩展即是一个 VMFS 格式的分区 (LUN 的一块)。例如，vmhba 0:0:14 是一个 LUN，而 vmhba 0:0:14:1 则为一个分区。一个 VMFS 卷可有多个扩展。

**注意** 缩写 vmhba 表示 ESX Server 系统上的物理 HBA (QLogic 或 Emulex), 而非虚拟机所使用的 SCSI 控制器。

**图 6-1. 数据存储详细信息**

The screenshot shows the ESX storage configuration interface. At the top, there are tabs for '配置', '任务与事件', '警报', '权限', and '映射'. Below the tabs is a '存储器' (Storage) section with buttons for '刷新', '移除', and '添加存储器...'. A table lists three storage devices:

标识	设备	容量	可用空间	类型
shared_50	vmhba1:0:3:1	49.75 GB	4.82 GB	vmfs3
Shared_200	vmhba1:1:0:1	199.75 GB	101.96 GB	vmfs3
localStorage	vmhba0:0:0:5	131.25 GB	130.70 GB	vmfs3

Below the table is a '详细信息' (Details) section for the selected 'localStorage' device, with a '属性...' (Properties) link. It shows the device name 'localStorage', capacity '131.25 GB', and location '/vmfs/volumes/480d2417...'. A pie chart indicates '561.00 MB 已使用' (Used) and '130.70 GB 可用空间' (Available space). To the right of the pie chart is an '扩展' (Expand) button.

路径选择	属性	扩展
固定的	卷标: localStor...	vmhba0:0:0:5 131.50..
	数据存储名称: localStor...	总格式化容量 131.25..
<b>路径</b>	<b>格式化</b>	
总计: 1	文件系统: VMFS 3.31	
中断: 0	块大小: 1 MB	
禁用: 0		

- 4 单击 [属性 (Properties)] 查看和更改属性。

## 解决显示问题

如果使用 AX100 存储阵列, 非活动连接可导致显示问题。请参见“AX100 非活动连接的显示问题”(第 52 页)。

## 了解显示屏幕中的 LUN 命名

在 VI Client 中, LUN 显示为三个或四个数字的序列, 中间由冒号分隔:

<SCSI HBA>:<SCSI target>:<SCSI LUN>:<disk partition>

如果最后一个数字为 0 或不显示, 则名称表示整个 LUN。

ESX 设备名称中的前三个数字可能会变化, 但仍表示同一物理设备。例如, vmhba1:2:3 表示 SCSI LUN3, 连接 SCSI 目标 2, 位于 SCSI HBA 1 上。重新引导 ESX Server 系统后, LUN 3 的设备名称可能更改为 vmhba1:1:3。这些数字具有以下含义:

- 第一个数字表示 SCSI HBA，发生变化的条件为：重新引导或重新扫描系统时发生 FC 或 iSCSI 网络中断，需要 ESX 通过其他 SCSI HBA 访问物理设备。
- 第二个数字表示 SCSI 目标，发生变化的条件为：对 ESX Server 主机可见的 FC 或 iSCSI 目标的映射中发生变化。
- 第三个数字表示 SCSI LUN，任何情况下均不会发生变化。

## 解决有关 LUN 不可见的问题

可以使用 VI Client 查看 LUN。

如果显示屏幕（或输出）与预期有所不同，请检查以下几项：

- **线缆连接** - 如果看不到端口，问题可能在于线缆连接或区域分配。请先检查线缆。
- **区域分配** - 限制对特定存储设备的访问，增加安全性，同时降低网络流量。有些存储供应商仅允许单启动器区域。在这种情况下，一个 HBA 可处于仅面向一个目标的多个区域。另一些供应商允许多启动器分区。请参见存储供应商的文档了解区域分配要求。使用 SAN 交换机软件配置和管理区域分配。
- **LUN 屏蔽** - 如果 ESX Server 主机可看到特定存储设备，但在该设备上看不到预期的 LUN，则可能是 LUN 屏蔽的设置不正确。

要从 SAN 引导，请确保每台 ESX Server 主机仅看到所需的 LUN。不要使任何 ESX Server 主机看到不归其所有的任何引导 LUN。使用磁盘阵列软件确保 ESX Server 主机仅能看到应对其可见的 LUN。

确保通过 **[Disk.MaxLUN]** 和 **[Disk.MaskLUNs]** 设置可查看预期为可见的 LUN。请参见“使用 **Disk.MaxLUN** 更改扫描的 LUN 的数目”（第 80 页）和“使用 **Disk.MaskLUNs** 屏蔽 LUN”（第 80 页）。

- **存储处理器** - 如果磁盘阵列有多个 SP，请确保 SAN 交换机连接了要访问的 LUN 所属的 SP。在某些磁盘阵列上，仅有一个 SP 为主动，另一 SP 在发生故障前是被动的。如果连接错误的 SP（对应被动路径的 SP），则您可能看不到预期的 LUN，或可能看到了 LUN，但在尝试进行访问时出错。

## 使用重新扫描

每次完成以下某一操作后请执行重新扫描：

- 针对某一 ESX Server 主机对 SAN 上的新磁盘阵列进行区域分配。
- 在 SAN 磁盘阵列上创建新 LUN。

- 更改 ESX Server 主机磁盘阵列上的 LUN 屏蔽。屏蔽所有指向 LUN 的路径之后，应重新扫描所有提供指向 LUN 的路径的适配器，以便更新配置。
- 重插拔线缆。
- 对群集中的 ESX Server 主机进行更改。
- 对 ESX Server 主机上的数据存储配置进行更改，例如，创建新的数据存储，移除、升级数据存储或对其进行重新签名，或者添加扩展。

---

**注意** 路径发生故障时不要重新扫描。如果一条路径发生故障，将由另一条路径取代，系统所有功能仍继续运作。但是，如果在路径不可用时重新扫描，ESX Server 主机会将该路径从指向设备的路径列表中移除。直到下次当路径处于活动状态时执行重新扫描后，ESX Server 主机才能使用此路径。

---

### 重新执行扫描

- 1 在 VI Client 中，选择一个主机，然后单击 **[配置 (Configuration)]** 选项卡。
- 2 在 **[硬件 (Hardware)]** 面板中，选择 **[存储适配器 (Storage Adapters)]**，然后单击 **[存储适配器 (Storage Adapters)]** 面板上方的 **[重新扫描 (Rescan)]**。

也可选择单个适配器，然后单击 **[重新扫描 (Rescan)]** 仅重新扫描该适配器。

## 移除数据存储

使用 VI Client，您可以从 ESX Server 主机中移除数据存储。移除数据存储之前，请迁移此数据存储上驻留的虚拟机。

### 移除数据存储

- 3 在 **[清单 (Inventory)]** 面板中，选择主机。
- 4 依次单击 **[配置 (Configuration)]** 选项卡和 **[存储器 (Storage)]** 以显示所有存储设备。
- 5 选择要移除的数据存储，然后单击 **[移除 (Remove)]**。
- 6 单击 **[刷新 (Refresh)]** 更新可用存储选项的视图。

## 高级 LUN 显示配置

本节论述多个高级配置选项，其中包括更改 LUN 的数目、屏蔽 LUN 以及更改稀疏 LUN 支持。

## 使用 Disk.MaxLUN 更改扫描的 LUN 的数目

默认情况下，VMkernel 从每一目标的 LUN 0 扫描到 LUN 255（总共 256 个 LUN）。您可以通过更改 **[Disk.MaxLUN]** 参数来更改这一数目。此更改可能会提升发现 LUN 的速度。

---

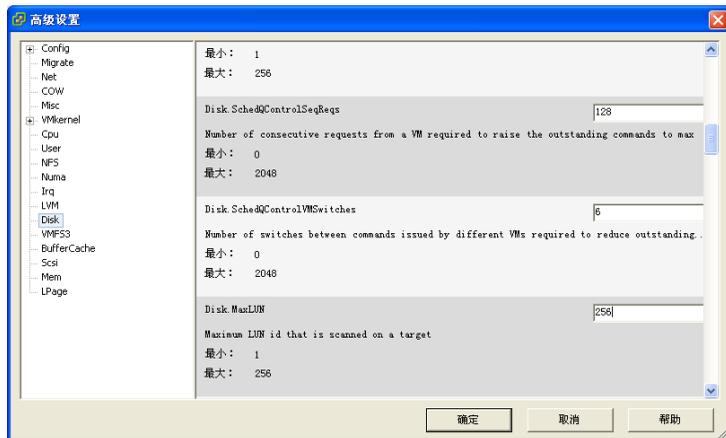
**注意** 无法发现 LUN ID 号大于 255 的 LUN。

---

减小该值可缩短重新扫描时间和引导时间。重新扫描 LUN 所需的时间取决于若干因素，其中包括存储阵列的类型以及是否启用了稀疏 LUN 支持。请参见“[使用 Disk.SupportSparseLUN 更改稀疏 LUN 支持](#)”（第 81 页）。

### 更改 Disk.MaxLUN 的值

- 1 在 VI Client 的 [ 清单 (Inventory) ] 面板中，选择主机。
- 2 单击 [ **配置 (Configuration)** ] 选项卡，然后单击 [ **高级设置 (Advanced Settings)** ]。
- 3 在出现的对话框中选择 [ **磁盘 (Disk)** ]。
- 4 向下滚动至 [ **Disk.MaxLUN** ]，将现有的值更改为所选值，然后单击 [ **确定 (OK)** ]。



## 使用 Disk.MaskLUNs 屏蔽 LUN

可使用 **Disk.MaskLUNs** 参数在特定 HBA 上屏蔽特定 LUN。VMkernel 不会接触或访问已屏蔽的 LUN，即使在初始扫描过程中也是如此。

如果要防止 ESX Server 系统访问某些 FC LUN，但又不想使用 FC 交换机或 FC 设备 LUN 屏蔽机制，您便可以使用此选项。

### 更改 Disk.MaskLUNs 的值

- 1 在 VI Client 的 [ 清单 (Inventory)] 面板中，选择主机。
- 2 单击 [ 配置 (Configuration)] 选项卡，然后单击 [ 高级设置 (Advanced Settings)]。
- 3 在出现的对话框中选择 [ 磁盘 (Disk)]。
- 4 向下滚动至 [Disk.MaskLUNs]，并使用以下格式将现有的值更改为所选值：  
 < 适配器 >:< 目标 >:< 以逗号分隔的 LUN 范围列表 >
- 5 单击 [ 确定 (OK)]。



**小心** 如果目标、LUN 或 vmhba 编号因重新配置服务器或 SAN 而发生变化，那么可能错误地屏蔽或显示 LUN。

### 使用 Disk.SupportSparseLUN 更改稀疏 LUN 支持

对设备上不支持 SCSI-3 标准的 LUN 进行扫描时，VMkernel 使用一种顺序法探测给定 LUN ID 范围内的每个 LUN。默认情况下，VMkernel 配置为支持稀疏 LUN，即并非范围内所有 LUN 都存在的一种配置情况。

如果范围内的所有 LUN 均存在，您可以禁用 **Disk.SupportSparesLUN** 参数。此更改将缩短对 LUN 进行扫描所需的时间。VMkernel 发现范围内某个 LUN 不存在便立即停止对 LUN 的探测。

对于支持 SCSI-3 标准的 LUN，无需更改 **Disk.SupportSparseLUN** 参数。VMkernel 所使用的方法可发现 ESX Server 主机的所有可用 LUN，无需按顺序对 LUN 进行扫描。

#### 禁用稀疏 LUN 支持

- 1 在 VI Client 的 [ 清单 (Inventory)] 面板中，选择主机。
- 2 单击 [ 配置 (Configuration)] 选项卡，然后单击 [ 高级设置 (Advanced Settings)]。
- 3 在 [ 高级设置 (Advanced Settings)] 对话框中，选择 [ 磁盘 (Disk)]。
- 4 向下滚动至 [Disk.SupportSparseLUN]，将值更改为 0，然后单击 [ 确定 (OK)]。

## N-Port ID 虚拟化

N-Port ID 虚拟化 (N-Port ID Virtualization, NPIV) 是一个 ANSI T11 标准，介绍单个光纤通道 HBA 端口如何使用多个全球端口名称 (WorldWide Port Name, WWPN) 向架构注册。这将允许架构所连接的 N-port 要求使用多个架构地址。每个地址在光纤通道架构上都显示为唯一的实体。

## 基于 NPIV 的 LUN 访问如何运作

可向 SAN 对象（如交换机、HBA、存储设备或虚拟机）分配全球名称 (WWN) 标识符。WWN 在光纤通道架构中唯一标识此类对象。如果虚拟机具有 WWN 分配，它们将分配的 WWN 用于所有 RDM 流量，因此虚拟机上任何 RDM 所指向的 LUN 不得针对其 WWN 进行屏蔽。如果虚拟机没有 WWN 分配，它们将使用主机物理 HBA 的 WWN 访问存储 LUN。但是，通过使用 NPIV，SAN 管理员可以按虚拟机监控和路由存储访问。下一节将介绍其运作方式。

NPIV 使得单一 FC HBA 端口可以向架构注册多个唯一的 WWN，其中每个 WWN 都可分配给单个虚拟机。向虚拟机分配了 WWN 后，该虚拟机的配置文件 (.vmx) 将更新为包含一个 WWN 对（由全球端口名称 (WWPN) 和全球节点名称 (World Wide Node Name, WWNN) 组成）。当该虚拟机已启动时，VMkernel 将在物理 HBA 上实例化一个虚拟端口 (Virtual Port, VPORT)，用于访问 LUN。VPORT 是虚拟 HBA，在 FC 架构中显示为物理 HBA，也就是说，它自己唯一的标识符，即分配给虚拟机的 WWN 对。每个 VPORT 均特定于虚拟机，虚拟机关闭后，VPORT 便在主机上消失，并且不会再显示于 FC 架构。

如果启用了 NPIV，在创建时会为每个虚拟机指定四个 WWN 对 (WWPN 和 WWNN)。启动使用 NPIV 的虚拟机后，它会依次使用上述各个 WWN 对来尝试发现指向存储器的访问路径。实例化的 VPORT 数目等于主机上存在的物理 HBA 个数，最大为四。VPORT 创建于建立了物理路径的各物理 HBA 之上。各条物理路径可用于确定将来访问 LUN 的虚拟路径。请注意，NPIV 无法识别的 HBA 在此发现过程中将被跳过，因为其无法实例化 VPORT。

---

**注意** 如果用户将四个物理 HBA 作为指向存储器的路径，所有物理路径必须由 SAN 管理员针对虚拟机进行区域分配。这需要支持多路径，即使一次只有一条路径处于活动状态。

---

## 使用 NPIV 的要求

尝试通过向虚拟机分配 WWN 来实施 NPIV 之前，请注意以下要求和限制：

- NPIV 仅能用于具备 RDM 磁盘的虚拟机。具备常规虚拟磁盘的虚拟机将使用主机物理 HBA 的 WWN。有关 RDM 的详细信息，请参见《ESX Server 3 配置指南》或《ESX Server 3i 配置指南》。
- 对于此 NPIV 实施，ESX Server 主机上的物理 HBA（使用其自身的 WWN）必须对该主机上运行的虚拟机要访问的所有 LUN 具有访问权限。

- ESX Server 主机的物理 HBA 必须支持 NPIV。目前，以下供应商及 HBA 类型提供这一支持：
  - QLogic - 所有 4 GB HBA。
  - Emulex - 具有 NPIV 兼容固件的 4 GB HBA。
- 每台虚拟机仅生成四个 WWN 对。
- 对分配了 WWN 的虚拟机或模板进行克隆后，得到的克隆副本不保留 WWN。
- 使用的交换机必须是 NPIV 可识别的。
- 配置用于在存储器级别进行访问的 NPIV LUN 时，请确保 NPIV LUN 号和 NPIV 目标 ID 与物理 LUN 和目标 ID 相匹配。
- 始终使用 VI Client 对具有 WWN 的虚拟机进行操作。

## 向虚拟机分配 WWN

可在创建带有 RDM 磁盘的新虚拟机时向此虚拟机分配 WWN，或者向可以暂时关闭的现有虚拟机分配 WWN。

### 创建带有 RDM 的虚拟机

- 1 从 VI Client 中，单击导航栏中的 **[ 清单 (Inventory) ]**，并在必要时展开清单。
- 2 在清单列表中，选择要添加新虚拟机的受管主机。
- 3 选择 **[ 文件 (File) ] > [ 新建 (New) ] > [ 虚拟机 (Virtual Machine) ]**。  
此时出现新建虚拟机向导。
- 4 选择 **[ 自定义 (Custom) ]**，然后单击 **[ 下一步 (Next) ]**。
- 5 键入虚拟机名称，然后单击 **[ 下一步 (Next) ]**。
- 6 选择一个文件夹或数据中心的根目录，然后单击 **[ 下一步 (Next) ]**。
- 7 如果资源池选项可用，请展开此树直到找到要运行虚拟机的资源池，使其突出显示，然后单击 **[ 下一步 (Next) ]**。
- 8 选择在其中存储虚拟机文件的数据存储，然后单击 **[ 下一步 (Next) ]**。
- 9 在 **[ 客户操作系统 (Guest operating system) ]** 下，选择操作系统系列（[Microsoft Windows]、[Linux]、[Novell NetWare]、[Solaris] 或 [其他 (Other)]）。
- 10 从下拉菜单中选择版本，然后单击 **[ 下一步 (Next) ]**。
- 11 从下拉列表中选择虚拟机的虚拟处理器数目，然后单击 **[ 下一步 (Next) ]**。

- 12 选择以兆字节为单位的数字配置虚拟机的内存大小，然后单击 [ **下一步 (Next)** ]。
- 13 配置网络连接，然后单击 [ **下一步 (Next)** ]。
- 14 选择希望与虚拟机一起使用的 SCSI 适配器的类型。
- 15 选择 [ **裸设备映射 (Raw Device Mapping)** ]，然后单击 [ **下一步 (Next)** ]。
- 16 在 SAN 磁盘或 LUN 列表中，选择想要虚拟机直接访问的裸 LUN。
- 17 为 RDM 映射文件选择数据存储。

可以将 RDM 文件置于虚拟机文件所驻留的同一数据存储上，也可以选择不同的数据存储。

---

**注意** 如果要对启用了 NPIV 的虚拟机使用 VMotion，请确保 RDM 文件位于虚拟机配置文件所驻留的同一数据存储。启用 NPIV 后，无法在数据存储之间执行 Storage VMotion 或 VMotion。

---

- 18 选择兼容模式：物理或虚拟。
  - 物理兼容性模式允许客户操作系统直接访问硬件。如果正在虚拟机中使用 SAN 感知应用程序，则物理兼容性非常有用。但是，带有物理兼容 RDM 的虚拟机不能克隆，不能制作成模板，或者不能迁移（如果迁移涉及复制磁盘）。
  - 虚拟兼容性允许 RDM 像虚拟磁盘一样工作，因此您可以使用快照和克隆之类的功能。

根据您的选择，后续屏幕将提供不同的选项。

- 19 在 [ 指定高级选项 (Specify Advanced Options) ] 页面中，可以更改虚拟设备节点，然后单击 [ **下一步 (Next)** ]。
- 20 按照 “**分配或修改 WWN**” 中的步骤向虚拟机分配 WWN。在 [ 即将完成新虚拟机 (Ready to Complete New Virtual Machine) ] 页面上，选中 [ 完成前编辑虚拟机设置 (Edit the virtual machine settings before completion) ] 复选框并单击 [ **下一步 (Next)** ]。

创建了带有 RDM 的虚拟机后，您可以向其分配虚拟 WWN。您还可以对带有 RDM 的现有虚拟机修改 WWN 分配。如果要编辑现有的 WWN，请确保关闭相应虚拟机。

### **分配或修改 WWN**

- 1 确保 SAN 管理员已置备了存储 LUN ACL，允许虚拟机的 ESX Server 主机进行访问。
- 2 打开 [ 虚拟机属性 (Virtual Machine Properties) ] 对话框：

- 对于新的虚拟机，按照“[创建带有 RDM 的虚拟机](#)”中所述创建虚拟机后，在 [即将完成新虚拟机 (Ready to Complete New Virtual Machine)] 页面上选中 [在提交创建任务前编辑虚拟机设置 (Edit the virtual machine settings before submitting the creation task)] 复选框，然后单击 [继续 (Continue)]。
  - 对于现有的虚拟机，从清单面板中选择虚拟机，然后单击 [编辑设置 (Edit Settings)] 链接。
- 3 选择 [选项 (Options)] 选项卡。
  - 4 选择 [光纤通道 NPIV (Fibre Channel NPIV)]。
  - 5 在打开的对话框中，选择以下选项之一：
    - [保留不变 (Leave unchanged)] - 保留现有 WWN 分配。此对话框的只读 [WWN 分配 (WWN Assignments)] 部分显示所有现有 WWN 分配的节点和端口值。
    - [生成新的 WWN (Generate New WWNs)] - 生成新的 WWN 并分配给虚拟机，覆盖任何现有 WWN（HBA 自身的 WWN 不受影响）。
    - [移除 WWN 分配 (Remove WWN assignment)] - 移除向虚拟机分配的 WWN，虚拟机将使用 HBA WWN 访问存储 LUN。如果创建新的虚拟机，则此选项不可用。



**小心** 移除或更改虚拟机的现有 WWN 分配可导致其与存储 LUN 的连接丢失。

---

- 6 单击 [确定 (OK)] 保存更改。

## 多路径

有关多路径概念的介绍，请参见“[路径管理和故障切换](#)”（第 36 页）。

---

**注意** SAN 实施中的 LUN 以及指向这些 LUN 的路径较多，可导致 ESX Server 在枚举所有路径之前便耗尽资源。这会使 ESX Server 无法看到所有指向存储器的路径。为避免此情况，请减少指向 LUN 的路径数。

---

## 查看当前的多路径状况

您可以使用 VI Client 查看当前的多路径状况。

### 查看当前的多路径状况

- 1 在 VI Client 的 [ 清单 (Inventory)] 面板中，选择一个主机，然后单击 [ 配置 (Configuration)] 选项卡。
  - 2 在 [ 存储器 (Storage)] 面板中，选择某一数据存储。
- 有关该数据存储的信息将出现在 [ 详细信息 (Details)] 面板中。

**配置** 任务与事件 警报 权限 映射

**存储器** [刷新](#) [移除](#) [添加存储器...](#)

标识	设备	容量	可用空间	类型
shared_50	vmhba1:0:3:1	49.75 GB	4.82 GB	vmfs3
Shared_200	vmhba1:1:0:1	199.75 GB	101.96 GB	vmfs3
localStorage	vmhba0:0:0:5	131.25 GB	130.70 GB	vmfs3

**详细信息** [属性...](#)

**Shared\_200** 199.75 GB 容量

位置: /vmfs/volumes/47067c74...

97.79 GB  已使用  
101.96 GB  可用空间

路径选择	属性	扩展
最近使用	卷标: Shared_200	vmhba1:1:0:1 199.99..
	数据存储名称: Shared_200	总格式化容量 199.75..

路径	格式化
总计: 4	文件系统: VMFS 3.21
中断: 0	块大小: 1 MB
禁用: 0	

- 3 要查看其他信息，或更改多路径策略，请选择 [ 详细信息 (Details)] 面板上方的 [ 属性 (Properties)]。

- 4 如果数据存储有多个扩展，在 [ 扩展 (Extents)] 面板中，选择要查看或更改信息的扩展。

[ 扩展设备 (Extent Device)] 面板将显示设备、此设备上的 VMFS 数据存储、路径选择算法、可用路径以及活动路径的相关信息。



显示屏幕将包括有关各条指向设备的路径的状态信息。其中会出现以下路径信息：

- **[ 活动 (Active)]** - 路径处于工作状态并且是用于传输数据的当前路径。
- **[ 已禁用 (Disabled)]** - 路径已禁用，无法传输数据。
- **[ 备用 (Standby)]** - 路径处于工作状态，但当前并未用于数据传输。
- **[ 中断 (Broken)]** - 软件无法通过此路径连接磁盘。

- 5 如果使用 **[ 固定的 (Fixed) ]** 路径策略并想查看首选路径，请单击 **[ 管理路径 (Manage Paths) ]**。

首选路径的第四列标有一个星号 (\*)。

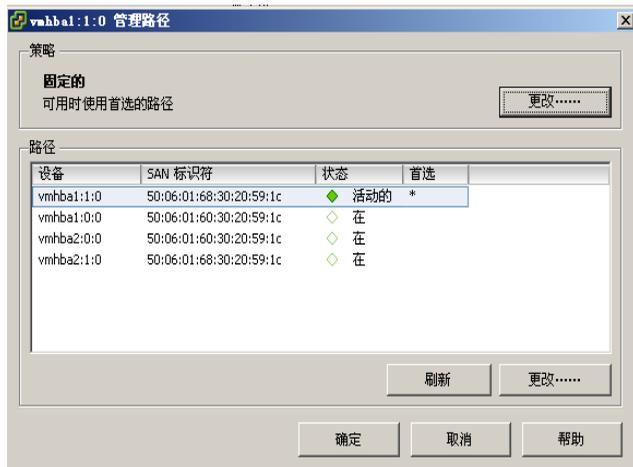


表 6-2 总结了 ESX Server 系统的行为随不同阵列类型和故障切换策略的变化情况。

**表 6-2. 路径策略影响**

策略 / 控制器	主动 / 主动	主动 / 被动
最近使用	发生路径故障后进行故障恢复需要管理员操作。	发生路径故障后进行故障恢复需要管理员操作。
固定的	连接恢复后，VMkernel 继续使用首选路径。	VMkernel 尝试继续使用首选路径。这会导致路径抖动或故障，因为另一 SP 现拥有 LUN 的所有权。请参见 <a href="#">“解决路径抖动”</a> (第 98 页)。

## 设置 LUN 多路径策略

默认情况下，在给定的任意时间里，ESX Server 主机仅使用一条路径（称为活动路径）与特定存储设备进行通信。选择活动路径时，ESX Server 遵循以下多路径策略：

- **[ 固定的 (Fixed) ]** - 当通往磁盘的指定首选路径可用时，ESX Server 主机始终使用此路径。如果主机无法通过首选路径访问磁盘，它会尝试备用路径。**[ 固定的 (Fixed) ]** 是主动 / 主动存储设备的默认策略。

- **[最近使用 (Most Recently Used)]** - ESX Server 主机使用最近使用的通往磁盘的路径，直到此路径变得不可用。也就是说，ESX Server 主机不会自动恢复到首选路径。**[最近使用 (Most Recently Used)]** 是主动 / 被动存储设备的默认策略并且对于这些设备是必需的。
- **[循环 (Round Robin)]** - ESX Server 主机将自动轮流选择所有可用的路径。除了路径故障切换，循环还支持路径间的负载平衡。

---

**注意** 循环负载平衡处于实验阶段，不支持供生产使用。请参见《*循环负载平衡*》白皮书。

---

ESX Server 主机根据其所检测阵列的品牌和型号设置多路径策略。如果不支持检测到的阵列，则将其视为主动 / 主动阵列。有关支持的阵列列表，请参见《*存储器 / SAN 兼容性指南*》。

---

**注意** 不建议手动将 **[最近使用 (Most Recently Used)]** 更改为 **[固定的 (Fixed)]**。系统会为有需要的阵列设置此策略。

---

### 使用 VI Client 设置多路径策略

- 1 在 VI Client 的 [清单 (Inventory)] 面板中，选择主机，然后单击 **[配置 (Configuration)]** 选项卡。
- 2 在 [硬件 (Hardware)] 面板中，选择 **[存储器 (Storage)]**。
- 3 选择要为其更改多路径策略的数据存储，然后单击 [详细信息 (Details)] 面板中的 **[属性 (Properties)]**。
- 4 在 [扩展 (Extent)] 面板中，选择要进行更改的设备，然后单击右侧 [扩展设备 (Extent Device)] 面板中的 **[管理路径 (Manage Paths)]**。  
此时会打开管理路径向导。
- 5 在 [策略 (Policy)] 下，单击 **[更改 (Change)]**。  
此时会打开 [选择策略 (Selection Policy)] 页面。
- 6 选择多路径策略并单击 **[确定 (OK)]**。

## 禁用和启用路径

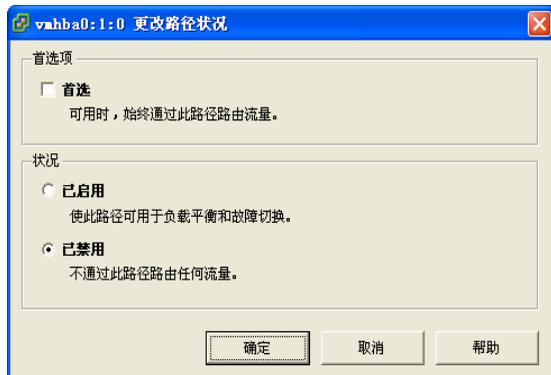
由于维护或其他原因，可以暂时禁用路径。您可以使用 VI Client 完成此操作。

## 禁用路径

- 1 在 VI Client 的 [ 清单 (Inventory)] 面板中，选择主机，然后单击 [ **配置 (Configuration)** ] 选项卡。
- 2 在 [ 硬件 (Hardware)] 面板中，选择 [ **存储器 (Storage)** ]。
- 3 选择要对其禁用路径的设备，然后单击 [ 详细信息 (Details)] 面板中的 [ **属性 (Properties)** ]。
- 4 在 [ 扩展 (Extent)] 面板中，选择要进行更改的设备，然后单击右侧 [ 扩展设备 (Extent Device)] 面板中的 [ **管理路径 (Manage Paths)** ]。

此时会打开管理路径向导。

- 5 在 [ 路径 (Paths)] 下，选择要禁用的路径，然后单击 [ **更改 (Change)** ]。
- 6 选中 [ **已禁用 (Disabled)** ] 单选按钮禁用该路径。



## 启用路径

如果已禁用了路径（例如，为进行维护），您可以按照禁用路径的步骤将其启用，不过要单击 [ **已启用 (Enabled)** ] 单选按钮。

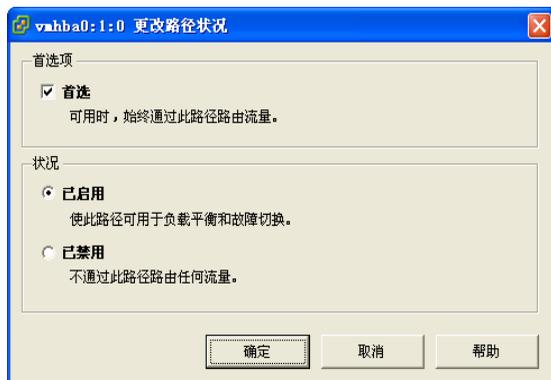
## 设置固定的路径策略的首选路径

如果使用 [ **固定的 (Fixed)** ] 路径策略，请指定服务器在路径可用时所使用的首选路径。

### 设置首选路径

- 1 在 VI Client 的 [ 清单 (Inventory)] 窗格中，选择主机，然后单击 [ **配置 (Configuration)** ] 选项卡。
- 2 在 [ 硬件 (Hardware)] 面板中，选择 [ **存储器 (Storage)** ]。

- 3 选择要对其禁用路径的设备，然后单击 [ 详细信息 (Details) ] 面板中的 [ **属性 (Properties)** ]。
- 4 在 [ 扩展 (Extent) ] 面板中，选择要进行更改的设备，然后单击右侧 [ 扩展设备 (Extent Device) ] 面板中的 [ **管理路径 (Manage Paths)** ]。  
此时会打开管理路径向导。
- 5 在 [ 路径 (Paths) ] 下，选择要作为首选路径的路径，然后单击 [ **更改 (Change)** ]。
- 6 在 [ 首选项 (Preference) ] 窗格中，单击 [ **首选 (Preferred)** ]。



如果 [ **首选 (Preferred)** ] 不可用，请确保路径策略为 [ **固定的 (Fixed)** ]。

- 7 单击 [ **确定 (OK)** ] 两次，以保存设置并退出对话框。

## 路径管理和手动负载均衡

在可用路径间平衡负载将提高性能。通过更改不同 HBA 的首选路径，您可以将系统设置为使用指向不同 LUN 的不同路径。这仅适用于主动 / 主动 SP，并且需要将路径策略设置为 [ **固定的 (Fixed)** ]。

如果某一路径发生故障，其余可用路径将承载所有流量。路径故障切换可能需要一分钟或更长时间，因为架构可能要协同新的拓扑来尝试还原服务。这一延迟必不可少，以便 SAN 架构能在发生拓扑更改或其他架构事件后稳定自身配置。

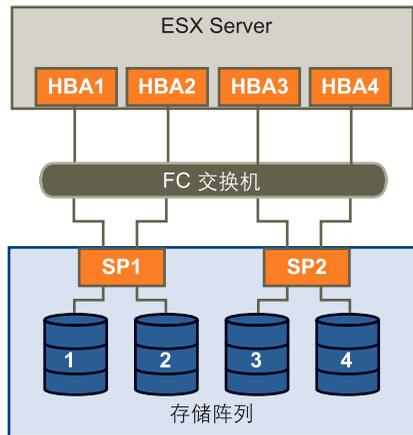
以下示例演示了如何执行手动负载均衡：

使用主动 / 主动阵列时，您可以针对负载均衡对环境进行设置。假定以下设置，如图 6-2 中所示：

- 主动 / 主动 SP

- 一个 ESX Server 系统
- 每台服务器中有四个光纤通道 HBA
- Director 级软件

图 6-2. 手动负载均衡



为实现负载均衡，请按如下所示设置首选路径。

- LUN 1: vmhba1:1:1
- LUN 2: vmhba2:1:2
- LUN 3: vmhba3:2:3
- LUN 4: vmhba4:2:4

请参见“[设置固定的路径策略的首选路径](#)”（第 90 页）了解相关信息。

---

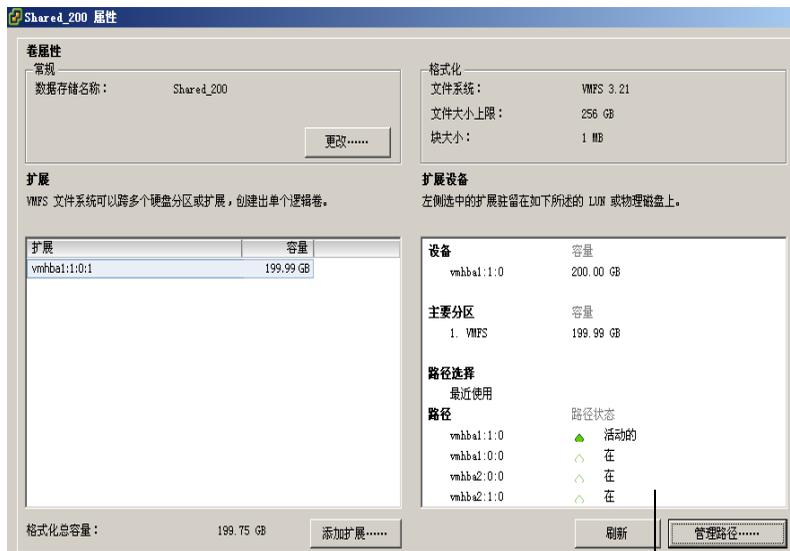
**注意** 尽管此示例使用四个 HBA，其实仅对两个 HBA 即可执行负载均衡。

---

## 故障切换

路径故障切换指的是指向 LUN 的活动路径由某一路径更改为另一路径的情况，通常是因为当前路径沿线的某一 SAN 组件发生故障。服务器通常有一个或两个 HBA，每个 HBA 可看到给定 SAN 阵列上的一个或两个存储处理器。通过查看 LUN 的属性可以确定活动路径，即服务器当前所使用的路径。

图 6-3. 活动路径和备用路径



活动路径和备用路径

拔出 FC 线缆后，I/O 可能会暂停 30-60 秒，直到 FC 驱动程序确定链路无法工作并且故障切换完成为止。因此，虚拟机（其虚拟磁盘安装在 SAN 存储器上）可能变得没有响应。如果尝试显示主机、主机存储设备或主机适配器，操作可能被挂起。故障切换完成后，I/O 恢复正常。

如果发生存在多处故障的灾难性事件，所有与 SAN 存储设备的连接可能都会丢失。如果任何与存储设备的连接均无法工作，某些虚拟机在其虚拟 SCSI 磁盘上可能遇到 I/O 错误。

## 对故障切换设置 HBA 超时

通常在 HBA BIOS 驱动程序中设置用于 I/O 重试操作的超时值。（您可能还想更改操作系统超时，如“设置操作系统超时”（第 94 页）中所述。）VMware 建议将超时值设置为 30 秒。

配置超时值：

- 对于 QLogic HBA，超时值为  $2 * n + 5$  秒，其中  $n$  是 QLogic 卡 BIOS 的 PortDownRetryCount 参数值。通过更改模块参数 qlport\_down\_retry 的值（其默认值来自 BIOS 设置）可以改变路径故障检测时间。此参数的建议设置为 14。

- 对于 Emulex HBA，可以通过更改模块参数 `lpfc_linkdown_tmo`（默认值为 30）和 `lpfc_nodev_tmo`（默认值为 30）的值来修改路径故障检测时间。驱动程序使用两个参数中的最大值来确定路径故障检测时间。二者的建议设置均为默认值。

要更改上述参数，您必须向驱动程序传递额外选项，如 `qlport_down_retry` 或 `lpfc_linkdown_tmo`。下一节介绍如何向驱动程序传递这些选项。

## 设置 SCSI 控制器的设备驱动程序选项

本节将设置 QLogic、Emulex 或其他 SCSI 卡驱动程序的设备驱动程序选项。

### 设置 QLogic、Emulex 或其他 SCSI 卡驱动程序的设备驱动程序选项

- 1 备份 `/etc/vmware/esx.conf` 文件，然后将其打开进行编辑。

文件包含对应每个 SCSI 设备的一节内容，如下例所示：

```
/device/002:02.0/class = "0c0400"
/device/002:02.0/devID = "2312"
/device/002:02.0/irq = "19"
/device/002:02.0/name = "QLogic Corp QLA231x/2340 (rev 02)"
/device/002:02.0/options = ""
/device/002:02.0/owner = "vmkernel"
/device/002:02.0/subsysDevID = "027d"
/device/002:02.0/subsysVendor = "1014"
/device/002:02.0/vendor = "1077"
/device/002:02.0/vmkernelname = "vmhba0"
```

- 2 找到紧邻 `name` 行之下的 `options` 行，然后根据需要进行修改。
- 3 如果需要，对由同一驱动程序控制的所有 SCSI 适配器重复上述操作。

## 设置操作系统超时

您可能想要增加标准磁盘超时值，使得 Windows 客户操作系统不会在故障切换过程中发生大范围中断。

对于 Windows 2000 和 Windows Server 2003 客户操作系统，可以使用注册表设置操作系统超时。

### 设置 Windows 服务器的操作系统超时

- 1 备份 Windows 注册表。
- 2 选择 [开始 (Start)]> [运行 (Run)]，键入 `regedit.exe` 并单击 [确定 (OK)]。
- 3 在左侧面板层次结构视图中，依次双击 [HKEY\_LOCAL\_MACHINE]、[System]、[CurrentControlSet]、[Services] 以及 [Disk]。

- 4 选中 **[TimeOutValue]** 并将数据值设置为 `x03c`（十六进制）或 `60`（十进制）。  
执行此更改后，Windows 将至少等待 60 秒，以便延迟的磁盘操作完成，然后才会生成错误。
- 5 单击 **[确定 (OK)]** 退出注册表编辑器。

## VMkernel 配置

安装 ESX Server 系统时，请决定放置不同存储元素的位置，例如服务控制台的 `/` 和 `/boot` 分区（仅限 ESX Server 3）。《安装指南》中对不同组件进行了更详细的说明。

## 共享诊断分区

如果 ESX Server 主机有一个本地磁盘，则该磁盘最适合用作诊断分区。原因之一就是：如果远程存储器出现问题，引发核心转储，则核心转储将丢失，从而加重解决问题的难度。

但是对于从 SAN 引导的无磁盘服务器，多个 ESX Server 系统可共享 SAN LUN 上的一个诊断分区。如果多个 ESX Server 系统使用一个 LUN 作为诊断分区，必须对该 LUN 进行区域分配以便所有服务器都可对其进行访问。

每台服务器需要 100 MB 空间，因此 LUN 的大小将决定可共享该 LUN 的服务器数目。每个 ESX Server 系统均映射到一个诊断插槽。如果服务器共享诊断分区，VMware 建议磁盘空间至少满足 16 个插槽所需 (1600 MB)。

如果设备上只有一个诊断插槽，所有共享该设备的 ESX Server 系统将映射到同一插槽。这很容易产生问题。如果两个 ESX Server 系统同时执行核心转储，则诊断分区上最后一个插槽上的核心转储会被覆盖。

如果分配足够的内存用于 16 个插槽，即使两个 ESX Server 系统同时执行核心转储，核心转储也不太可能映射到诊断分区上的同一位置。

## 避免问题和解决问题

本节给出一些避免问题和解决问题的提示：

- 将所有信息记录在案。其中包括涉及以下项目的信息：区域分配、访问控制、存储器、交换机、服务器和 FC HBA 配置、软件和硬件版本以及存储器线缆布局。
- 对故障情况进行规划：
  - 对拓扑进行映射并制作多个副本。考虑每一元素发生故障时可能对 SAN 带来的影响。

- 除去不同的链接、交换机、HBA 及其他元素，确保未遗漏设计中的关键故障点。
- 将 ESX Server 主机安装到生产系统时，请在本地安装过程中断开光纤通道 HBA 的连接。



**小心** 安装程序允许擦除所有可访问的磁盘，包括其他服务器正在使用的 SAN LUN。

- 确保根据插槽和总线速度将光纤通道 HBA 安装到 ESX Server 主机中的正确插槽。在服务器中的可用总线之间平衡 PCI 总线负载。
- 在所有可视点（包括 ESX Server 性能图表、FC 交换机统计信息及存储性能统计信息）熟悉存储网络中的不同监视点。

## 优化 SAN 存储性能

有关优化典型 SAN 环境的两个主要因素分别为存储阵列性能及服务器性能。如果环境配置正确，SAN 架构组件（特别是 SAN 交换机）对优化的影响较小，因为这些组件的延迟相对于服务器和存储阵列而言较短。确保经由交换机架构的路径尚未饱和，即交换机架构未以最高吞吐量运行。

### 存储阵列性能

如果存储阵列性能存在问题，务必参考存储阵列供应商文档了解任何相关信息。

分配 LUN 时，请记住，每个 LUN 由多台 ESX Server 主机访问，而且各主机上可运行多台虚拟机。由一台 ESX Server 主机使用的 LUN 可向运行于不同操作系统的多个不同应用程序提供 I/O 服务。由于此工作负载并非恒定不变，ESX Server LUN 所在的 RAID 组不应包括其他主机所使用的 LUN，这些主机上未运行 ESX Server 用于 I/O 密集型应用程序。

确保启用了读 / 写缓存。

负载均衡的过程即是将服务器 I/O 请求分散于所有可用 SP 及其关联的主机服务器路径。目的是针对吞吐量（每秒 I/O 流量、每秒兆字节数或响应时间）实现最佳性能。

需要不断对 SAN 存储阵列进行重新设计和调试，以确保所有存储阵列路径间的 I/O 负载均衡。为满足此要求，请在所有 SP 间分发指向 LUN 的路径以提供最佳负载均衡效果。密切监视可指示何时需要手动重新平衡 LUN 的分发。有关示例请参见“[路径管理和手动负载均衡](#)”（第 91 页）。

调试静态平衡存储阵列即是监视特定性能统计信息（如每秒 I/O 操作数、每秒块数及响应时间）并通过分发 LUN 工作负载将工作负载分散到所有 SP。

---

**注意** ESX Server 目前尚不支持动态负载平衡。

---

## 服务器性能

确保最佳的服务器性能需要考虑多个因素。各服务器应用程序访问其专用存储器时必须满足以下条件：

- 高 I/O 速率（每秒 I/O 操作数）
- 高吞吐量（每秒兆字节数）
- 最小延迟（响应时间）

由于各应用程序的要求不尽相同，您可以选择存储阵列上的适当 RAID 组来实现上述目标。要实现性能目标，请执行以下操作：

- 将每个 LUN 置于提供必要性能级别的 RAID 组。请注意所分配的 RAID 组中其他 LUN 的活动及资源利用率。对于高性能 RAID 组，因有过多应用程序对其执行 I/O 操作，它可能无法满足 ESX Server 主机上运行的应用程序所需的性能目标。
- 确保各服务器具有足够多的 HBA，能满足高峰时段服务器上托管的所有应用程序的最大吞吐量。将 I/O 分散到多个 HBA 可为各应用程序提供更高的吞吐量及更短的延迟。
- 要在 HBA 发生故障时提供冗余，请确保服务器连接一个双冗余架构。
- 为 ESX Server 系统分配 LUN 或 RAID 组时，多个操作系统将使用和共享该资源。因此，当您使用 ESX Server 系统时，存储子系统中各个 LUN 所需的性能要远高于使用物理机的情况。例如，如果计划运行四个 I/O 密集型应用程序，请为 ESX Server LUN 分配四倍大小的性能容量。
- 将多个 ESX Server 系统与 VirtualCenter Server 一起使用时，存储子系统所需的性能将相应增加。
- ESX Server 系统上运行的应用程序所需的待处理 I/O 数目应与 HBA 和存储阵列能处理的 I/O 数目相匹配。

## 解决性能问题

本节论述性能监视以及解决性能问题的可能途径。

为实现最佳性能，请将各虚拟机置于适当级别的存储器。请参见“[选择虚拟机位置](#)”（第 38 页）了解相关信息。

### 监视性能

VI Client 可提供多种工具用于收集性能信息。这些信息随后以图形方式显示在 VI Client 中。请参见《[基本系统管理](#)》了解相关信息。VI Client 可定时更新显示屏幕。

对于 ESX Server 3，您还可以从服务控制台使用 `esxtop` 工具。有关 `esxtop` 的信息，请参见《[资源管理指南](#)》，或从服务控制台查看手册页。可以使用 `esxtop` 实时监视性能。如果使用 ESX Server 3i，则类似功能可由 `resxtop` 工具提供。

### 解决路径抖动

如果服务器无法访问 LUN，或者访问速度很慢，则可能是路径抖动（也称为 LUN 抖动）带来的问题。两台主机通过不同 SP 访问 LUN 时，LUN 从未真正可用，此时可能出现路径抖动。

通常，只有特定的 SAN 配置与以下条件同时出现时可导致路径抖动：

- 使用主动 / 被动阵列。
- 路径策略设置为 **[ 固定的 (Fixed) ]**。
- 两台主机使用相反的路径顺序访问 LUN。例如，主机 A 设置为通过 SP A 访问编号较低的 LUN。主机 B 设置为通过 SP B 访问编号较低的 LUN。

如果主机 A 丢失了某一路径，只能使用指向 SP A 的路径，而主机 B 丢失其他路径，只能使用指向 SP B 的路径，此时也可能出现路径抖动。

直接连接阵列（如 AX100）的一个或多个节点上发生 HBA 故障切换时也可能出现这一问题。

路径抖动问题通常不会出现于其他操作系统：

- 其他常用操作系统均不会对超过两台以上的服务器使用共享 LUN（该设置通常是群集预留的）。
- 对于群集，一次只有一台服务器发出 I/O。路径抖动不会成为问题。

相比而言，多个 ESX Server 系统可能同时向同一 LUN 发出 I/O。

## 解决路径抖动

- 确保共享主动 / 被动阵列上相同 LUN 集合的所有主机同时访问同一存储处理器。
- 更正不同 ESX Server 主机与 SAN 目标间的任何线缆接线不一致问题，以便相同的目标以同一顺序呈现给所有 HBA。
- 确保路径策略设置为 [ **最近使用 (Most Recently Used)** ] (默认值)。

## 了解路径抖动

在所有阵列中，SP 就像对部分共享存储器具有访问权限的独立计算机。如何处理并发的访问将由算法决定。

- 对于主动 / 被动阵列，存储器上组成特定 LUN 的所有扇区每次只能由一个 SP 访问。所有权将在存储处理器之间传递。原因在于存储阵列将使用缓存且 SP A 不得向磁盘写入使 SP B 缓存无效的内容。因为 SP 在完成操作后必须刷新缓存，所以转移所有权需要一些时间。在这期间，任何 SP 均无法处理与 LUN 相关的 I/O。
- 对于主动 / 主动阵列，通过算法可实现对存储器更精细的访问以及将缓存同步。可通过任何 SP 同时访问，无需额外的时间。

阵列使用 AVT 是指主动 / 被动阵列在 I/O 到达时将 LUN 的所有权传递给不同的 SP，试图看起来如同主动 / 主动阵列。此方法适用于群集设置，但如果多个 ESX Server 系统通过不同的 SP 同时访问同一 LUN，则会导致 LUN 抖动。

考虑路径选择如何运作：

- 在主动 / 主动阵列上，系统开始沿新路径发送 I/O。
- 对于主动 / 被动阵列，ESX Server 系统将检查所有备用路径。当前待考虑路径末端的 SP 将向系统发送信息，指示其当前是否拥有 LUN 的所有权。
  - 如果 ESX Server 系统找到拥有 LUN 所有权的 SP，则会选定该路径并沿该路径发送 I/O。
  - 如果 ESX Server 主机找不到此类路径，ESX Server 主机便挑选某一路径并向 SP (位于路径另一端) 发送命令将 LUN 所有权移至该 SP。

如下路径选择可导致路径抖动：如果服务器 A 只能通过一个 SP 到达某个 LUN，而服务器 B 只能通过另一 SP 到达同一个 LUN，那么二者将使得 LUN 的所有权在两个 SP 之间不断转移，就像打乒乓球。由于系统转移所有权的速度很快，存储阵列无法处理任何 I/O (或只能处理非常少的一部分)。因此，依赖该 LUN 的任何服务器都会开始使 I/O 超时。

## 平衡虚拟机之间的磁盘访问

可以在 VI Client 中通过 **[Disk.SchedNumReqOutstanding]** 参数调整待处理的磁盘请求的最大数目。两台或多台虚拟机访问同一 LUN 时，此参数将控制每台虚拟机可向 LUN 发出的待处理请求数目。调整该限制有助于平衡虚拟机之间的磁盘访问。

此限制不适用于 LUN 上只有一个虚拟机处于活动状态的情况。在这种情况下，带宽由存储适配器的队列深度限制。

### 设置待处理磁盘请求的数目

- 1 在 VI Client 的清单面板中选择主机。
- 2 单击 **[配置 (Configuration)]** 选项卡，然后单击 **[高级设置 (Advanced Settings)]**。
- 3 在左侧面板中单击 **[磁盘 (Disk)]** 并向下滚动至 **[Disk.SchedNumReqOutstanding]**。
- 4 将参数值更改为所选的数字并单击 **[确定 (OK)]**。
- 5 重新引导服务器。

此更改可对磁盘带宽调度产生影响，但是实验结果证明此更改对磁盘密集型工作负载有改善作用。

如果在 VMkernel 中调整此值，您可能还需要调整存储适配器中的队列深度。请参见“[设置 HBA 的最大队列深度](#)”（第 101 页）。

## 移除 VMFS-2 驱动程序

如有多个 VMFS 数据存储，并且所有数据存储均为 VMFS-3，则可以通过卸载 VMFS-2 驱动程序潜在提高性能。在命令行提示符下，键入：

```
vmkload_mod -u vmfs2
```

刷新数据存储及重新扫描存储适配器等特定管理操作的速度将得到大幅提升。但是，由于此命令仅对当前引导有效，您需要在每次重新引导后重复这一命令。

## 移除 NFS 驱动程序

如果未使用 NFS 数据存储，您可以在命令行提示符处键入以下命令卸载 NFS 驱动程序。此命令仅对当前引导有效，因此需要在每次重新引导后重复这一命令。

```
vmkload_mod -u nfsclient
```

## 减少 SCSI 预留

需要在 VMFS 中获取文件锁或元数据锁的操作可导致暂时性 SCSI 预留。SCSI 预留将锁定整个 LUN。某一服务器使用过多 SCSI 预留可导致其他服务器访问相同 VMFS 时性能降低。

需要取得文件锁或元数据锁的操作示例包括：

- 启动虚拟机。
- VMotion。
- 虚拟机与虚拟磁盘快照一起运行。
- 需要打开文件或执行元数据更新的文件操作。（请参见“元数据更新”（第 29 页）。）

如果在访问相同 VMFS 的多台服务器上经常执行此类操作则会导致性能降低。例如，不建议在同一 VMFS 上运行来自多台服务器的多台使用虚拟磁盘快照的虚拟机。在 VMFS 上运行多台虚拟机时请限制 VMFS 文件操作的数目。

## 设置 HBA 的最大队列深度

ESX Server 应具有合理的队列深度。但是，如果对 HBA 的性能不甚满意，您可以更改其最大队列深度。

### 调整 QLogic HBA 的队列深度

可以通过以下过程调整 QLogic qla2x00 系列适配器的最大队列深度。

#### 设置 QLogic HBA 的最大队列深度

- 1 以 root 用户身份登录服务控制台。
- 2 备份 /etc/vmware/esx.conf 文件，然后将其打开进行编辑。

文件内容与下例类似：

```
/device/002:02.0/class = "0c0400"
/device/002:02.0/devID = "2312"
/device/002:02.0/irq = "19"
/device/002:02.0/name = "QLogic Corp QLA231x/2340 (rev 02)"
/device/002:02.0/options = ""
/device/002:02.0/owner = "vmkernel"
/device/002:02.0/subsysDevID = "027d"
/device/002:02.0/subsysVendor = "1014"
/device/002:02.0/vendor = "1077"
/device/002:02.0/vmkernelname = "vmhba0"
```

- 3 找到 name 行下方的 options 行，并对其进行修改以指定最大队列深度，如下所示（其中 nn 为队列深度最大值）：

```
/device/001:02.0/options = "ql2xmaxqdepth=nn"
```

---

**注意** ql2xmaxqdepth 中的第二个字符为小写的 “L”。

---

- 4 保存更改并重新引导服务器。

## 调整 Emulex HBA 的队列深度

可以通过以下过程调整 Emulex HBA 的最大队列深度。

### 更改 Emulex HBA 的队列深度

- 1 以 root 用户身份登录服务控制台。
- 2 验证当前已加载的 Emulex HBA 模块：

```
vmkload_mod -l | grep lpfcdd
```

根据 HBA 型号的不同，模块可为以下之一：

- lpfcdd\_7xx
- lpfcdd\_732.o - 此 4 GB 驱动程序随 ESX Server 3.x 一起提供。在某些情况下，可能需要降级为 2 GB 驱动程序。请参见 <http://kb.vmware.com/kb/1560391>。

- 3 对于系统上的单个 Emulex HBA 实例，请运行以下命令。

示例所示为 lpfcdd\_7xx 模块。请使用与 [步骤 2](#) 的输出相对应的模块。

- a esxcfg-module -s lpfc0\_lun\_queue\_depth=16 lpfcdd\_7xx
- b esxcfg-boot -b

在这种情况下，lpfc0 所代表的 HBA 的 LUN 队列深度将设置为 16。

- 4 对于系统上存在的多个 Emulex HBA 实例，请运行以下命令：

- a esxcfg-module -s "lpfc0\_lun\_queue\_depth=16  
lpfc1\_lun\_queue\_depth=16" lpfcdd\_7xx
- b esxcfg-boot -b

在这种情况下，两个 HBA（lpfc0 和 lpfc1）的 LUN 队列深度将设置为 16。

- 5 重新引导。

## SAN 存储器备份注意事项

在 SAN 环境中，备份有两个目的。第一个目的是将联机数据归档至脱机媒体。可对所有联机数据按时间表定期重复执行此过程。第二个目的是提供对脱机数据的访问，用于从故障中恢复。例如，数据库恢复通常需要检索当前未联机的已归档日志文件。

计划备份取决于多种因素：

- 重要应用程序的标识，这些应用程序在给定的一段时间内需要较频繁地备份。
- 恢复点和恢复时间目标。考虑恢复点所需的精确度，以及愿意为此而等待的时间长度。
- 与数据关联的变化率(Rate of Change, RoC)。例如，如果使用同步 / 异步复制，RoC 将影响主存储设备与次要存储设备间所需带宽的大小。
- 对 SAN 环境、存储性能（备份时）以及其他应用程序的总体影响。
- SAN 上高峰流量时段的标识（计划于这些高峰时段执行的备份会降低应用程序和备份过程的运行速度）。
- 计划数据中心内所有备份的时间。
- 备份单个应用程序所需的时间。
- 归档数据的资源可用性；通常为脱机媒体访问（磁带）。

设计备份策略时要包括各应用程序的恢复时间目标。也就是考虑重新置备数据所需的时间和资源。例如，如果计划的备份要存储过多数据，导致恢复需要大量时间，那么请检查已计划的备份。增加执行备份的频率，这可减少每次备份的数据，从而缩短恢复时间。

如果特定的应用程序需要在某一期限内恢复，则备份过程需要提供时间表及特殊数据处理以满足此需求。快速恢复可能需要使用驻留在联机存储器上的恢复卷，从而尽可能避免通过访问速度较慢的脱机媒体来获取缺少的数据组件。

## 快照软件

管理员通过快照软件可以对磁盘子系统内定义的任一虚拟磁盘生成即时副本。快照软件分为不同的可用级别：

- 通过 ESX Server 主机可创建虚拟机的快照。此软件在 ESX Server 基本软件包中附送。
- 第三方备份软件可能允许更全面的备份过程，并可能包含更高级的配置选项。

管理员可出于多种原因而生成快照，其中包括：

- 备份。
- 灾难恢复。
- 多种配置和 / 或版本的可用性。
- 取证 (Forensics) (在系统运行时通过查看快照寻找导致问题的原因)。
- 数据挖掘 (Data mining) (查看数据的副本以降低生产系统的负载)。

## 使用第三方备份软件包

如果使用第三方备份软件，请确保 ESX Server 主机支持该软件。请参见《*备份软件兼容性指南*》。

使用第三方软件具有环境统一的优点。但是，您必须考虑到，随着 SAN 规模的扩大，由第三方快照软件带来的额外开销将变得更高。

如果使用快照来备份数据，请考虑以下几点：

- 有些供应商同时支持 VMFS 和 RDM 的快照。如果二者均支持，您可以生成主机的整个虚拟机文件系统的快照，也可以对单个虚拟机（每磁盘一个）执行快照操作。
- 有些供应商仅对使用 RDM 的设置支持快照。如果仅支持 RDM，您可以对单个虚拟机执行快照。

请参见存储供应商的文档。

---

**注意** ESX Server 系统还包括一个 Consolidated Backup 组件，在《*虚拟机备份指南*》中进行了详细论述。

---

## 选择备份解决方案

选择备份解决方案时，请考虑备份可满足以下一项或全部需求：

- 崩溃一致性
- 文件系统一致性
- 应用程序一致性

VMware 可提供文件系统一致性备份。在多数情况下，使用文件系统一致性备份便可从故障中完全恢复。但是，如果应用程序需要跨文件系统同步或与数据库同步，则 VMware 解决方案可能无法提供足够的一致性。在这类情况下，应详细了解第三方备份解决方案以确定其是否更适合您的需求。

## 分层应用程序

SAN 管理员通常使用基于阵列的专用软件进行备份、灾难恢复、数据挖掘、取证以及配置测试。

存储供应商通常对 LUN 提供两种类型的高级服务，即快照和复制。

- 快照将创建包含 LUN 的高效率副本的空间，这些副本共享公共的数据块。快照通常在主要 LUN 所在的同一阵列上本地执行，用于快速备份、应用程序测试、取证或数据挖掘。
- 复制将创建 LUN 的完整副本。通常对单独的阵列或站点进行副本复制，以防御可使整个阵列或站点变为不可用或遭破坏的主要故障。

将 ESX Server 系统与 SAN 一起使用时，您需要判断基于阵列或基于主机的工具哪一个更适合特定的情形。

## 基于阵列（第三方）的解决方案

考虑基于阵列的解决方案时，请注意以下几点：

---

**注意** ESX Server 系统还包括一个 Consolidated Backup 组件，在《虚拟机备份指南》中进行了详细论述。

---

- 基于阵列的解决方案通常可得到更全面的统计信息。使用 RDM，数据始终采用同一路径，使得性能管理更为简单。
- 使用 RDM 和基于阵列的解决方案时，安全性对于存储管理员而言更为透明，因为使用 RDM 时虚拟机与物理机更为相似。
- 如果使用基于阵列的解决方案，通常会将物理兼容性 RDM 用作虚拟机的存储器。如果不打算使用 RDM，您应查看存储供应商的文档，确认是否支持在带有 VMFS 卷的 LUN 上进行操作。此外，如果在 VMFS LUN 上应用阵列操作，您应仔细阅读重新签名一节。

## 基于文件 (VMFS) 的解决方案

考虑使用 VMware Tools 和 VMFS（而非阵列工具）的基于文件的解决方案时，请注意以下几点：

- 使用 VMware Tools 和 VMFS 更利于置备：分配一个较大 LUN，多个 .vmdk 文件可置于该 LUN。使用 RDM，每台虚拟机都需要一个新 LUN。
- 快照服务随 ESX Server 主机附送，无需额外付费。因此基于文件的解决方案相比基于阵列的解决方案更为经济高效。

- 对于 ESX Server 管理员而言，使用 VMFS 更容易。
- ESX Server 管理员如使用基于文件的解决方案可减轻对 SAN 管理员的依赖性。

## VMFS 卷重新签名

ESX Server 需能区分其 VMFS 卷，为此可使用卷签名。对 VMFS 卷进行复制或生成快照后，得到的 LUN 副本与源 LUN 具有相同的标识。如果对 ESX Server 可见的两个 LUN 具有相同签名，ESX Server 必须对此情形进行处理以防止由于混淆用于访问已注册虚拟机的 LUN 而导致停机时间。ESX Server 3.0 引入重新签名功能来解决此问题。

---

**注意** 当某个 LUN 需要进行重新签名时，vmkernel 日志中将出现警示。如果遇到此类警示，请相应设置重新签名选项，如下节所述。

---

### 装载原始、快照或副本 VMFS 卷

您可以在同一 ESX Server 主机上装载原始、快照或副本 VMFS 卷。

#### 装载原始、快照或副本 VMFS 卷

- 1 执行所需的存储任务：
  - a 生成阵列快照或副本。
  - b 针对 ESX Server 将快照或副本屏蔽或进行区域分配。
- 2 在 VI Client 的清单面板中选择主机。
- 3 单击 [ **配置 (Configuration)** ] 选项卡，然后单击 [ **高级设置 (Advanced Settings)** ]。
- 4 在左侧面板中选择 [ **LVM** ]，然后将 [ **LVM.EnableResignature** ] 选项设置为 [ **1** ]。
- 5 重新扫描以显示任何新的 LUN 或 VMFS 卷。检测到作为快照或副本的卷将被重新签名。

重新扫描后，由复制所得的 VMFS 卷显示为  
/vmfs/volumes/snap-<DIGIT>-<old-label>。

如果任意虚拟机的 .vmx 文件或虚拟机快照的 .vmsd 文件包含 /vmfs/volumes/<标签或 UUID>/ 路径，您必须更改这些项以反映重新签名的卷的路径。

- 6 完成重新签名后，将 [ **LVM.EnableResignature** ] 选项设置为 [ **0** ]。

### 了解重新签名选项

本节论述 [ **EnableResignature** ] 和 [ **DisallowSnapshotLUN** ] 选项的交互方式，并介绍更改上述选项所得的三种状况：

- 状况 1: EnableResignature=0, DisallowSnapshotLUN=1 (ESX Server 3.x 默认值)
- 状况 2: EnableResignature=1 (DisallowSnapshotLUN 不相关)
- 状况 3: EnableResignature=0, DisallowSnapshotLUN=0 (ESX Server 2.x 行为)

### 状况 1 - EnableResignature=0, DisallowSnapshotLUN=1 (默认值)

在此状况中:

- 无法将阵列制作的 VMFS 卷的快照或副本放入 ESX Server 主机, 无论 ESX Server 是否具有原始 LUN 的访问权限。
- 使用 VMFS 格式化的 LUN 对于各 ESX Server 主机必须具有相同的 ID。

### 状况 2 - EnableResignature=1 (DisallowSnapshotLUN 不相关)

在此状况中, 您可以安全地将 VMFS 卷的快照或副本放入原始卷所在的同一服务器, 这些快照或副本将自动重新签名。

### 状况 3 - EnableResignature=0, DisallowSnapshotLUN=0

这类似于 ESX Server 2.x 行为。在此状况中, ESX Server 假定只看到给定 LUN 的一个副本或快照且从不尝试重新签名。这适用于某种 DR 方案, 在其中将 LUN 的一个副本放入 ESX Server 的新群集, 可能位于对源 LUN 没有访问权限的另一站点上。在这种情况下, ESX Server 将副本当作源一样使用。

如果将 LUN 的快照或副本放入对原始 LUN 具有访问权限的服务器, 请不要使用此设置。否则将导致破坏性结果, 包括:

- 如果一次或多次创建 VMFS 卷的快照, 并动态地将其中一个或多个快照放入 ESX Server, 则只有第一个副本可用。可用的副本最有可能是主要副本。重新引导后便不能确定哪个卷 (源或快照之一) 可用。这种非确定的行为非常危险。
- 如果创建跨区 VMFS 卷的快照, ESX Server 主机可能使用属于不同快照的碎片重新集合卷。这可能对文件系统造成损坏。



## 多路径对照表

本附录提供用于不同存储阵列的多路径设置要求对照表。

**表 A-1. 多路径设置要求**

组件	备注
所有存储阵列	没有备用电池时必须禁用写入缓存。
拓扑	一处故障不会导致 HBA 和 SP 同时进行故障切换，尤其是使用主动 - 被动存储阵列时。
IBM TotalStorage DS 4000（之前称为 FastT）	在较高版本中，默认主机类型必须为 LNXCL 或 VMware。 在较高版本中，主机类型必须为 LNXCL 或 VMware。 此主机模式中禁用 AVT（自动卷传输，Auto Volume Transfer）。
HDS 99xx 和 95xxV 系列	HDS 9500V 系列 (Thunder) 需要两种主机模式： <ul style="list-style-type: none"> <li>■ 主机模式 1: Standard。</li> <li>■ 主机模式 2: Sun Cluster</li> </ul> HDS 99xx 系列 (Lightning) 及 HDS Tabma (USP) 需将主机模式设置为 Netware。
EMC Symmetrix	启用 SPC2 和 SC3 设置。请联系 EMC 获取最新设置。
EMC Clariion	所有启动器记录必须包含如下设置： <ul style="list-style-type: none"> <li>■ Failover Mode = 1</li> <li>■ Initiator Type = "Clariion Open"</li> <li>■ Array CommPath = "Enabled" 或 1</li> </ul>
HP MSA	主机类型必须为 Linux。 将每个 HBA 端口的连接类型设置为 Linux。

**表 A-1. 多路径设置要求 (续)**

组件	备注
HP EVA	<p>对于 EVA3000/5000 固件 4.001 及更高版本和 EVA4000/6000/8000 固件 5.031 及更高版本，将主机类型设置为 <b>VMware</b>。</p> <p>否则将主机模式类型设置为 <b>Custom</b>。值为：</p> <ul style="list-style-type: none"> <li>■ EVA3000/5000 固件 3.x: 000000002200282E</li> <li>■ EVA4000/6000/8000: 000000202200083E</li> </ul>
HP XP	<p>对于 XP 128/1024/10000/12000，主机模式应设置为 <b>0C (Windows)</b>，即 <b>zeroC (Windows)</b>。</p>
NetApp	<p>无特定要求</p>
ESX Server 配置	<p>为 ESX Server 主机设置以下高级设置</p> <ul style="list-style-type: none"> <li>■ 将 <b>[Disk.UseLunReset]</b> 设置为 <b>[1]</b></li> <li>■ 将 <b>[Disk.UseDeviceReset]</b> 设置为 <b>[0]</b></li> </ul> <p>对于主动 - 被动阵列，必须将所有托管群集磁盘的 LUN 的多路径策略设置为 <b>[最近使用 (Most Recently Used)]</b>。主动 - 主动阵列上 LUN 的多路径策略可设置为 <b>[最近使用 (Most Recently Used)]</b> 或 <b>[固定的 (Fixed)]</b>。</p> <p>所有 FC HBA 必须为同一型号。</p>

# B

## 实用程序

---

在多数情况下, VI Client 非常适合监控 SAN 存储器所连接的 ESX Server 主机。高级用户有时可能需要使用某些命令行实用程序获取其他详细信息。

本附录提供有关以下实用程序的信息:

- “[esxtop 和 resxtop 实用程序](#)” (第 112 页)
- “[storageMonitor 实用程序](#)” (第 112 页)

## esxtop 和 resxtop 实用程序

esxtop 和 resxtop 命令行工具可提供有关 ESX Server 实时资源利用率的详细显示。有关这两种实用程序的详细信息，请参见《资源管理指南》。

## storageMonitor 实用程序

storageMonitor 实用程序可监视 VMware ESX Server 所连接的存储设备遇到的 SCSI 感知错误。该实用程序通过定期轮询 VMkernel 内运行的 storageMonitor 收集感知错误信息，然后将错误信息发送至标准输出、文件或系统日志。在将错误信息发送至输出之前，该实用程序会对其格式化。例如，它会按照 SCSI-3 规格将感知错误代码转换为相应的文本。

如果未指定配置文件，storageMonitor 将分析默认配置文件 /etc/vmware/storageMonitor.conf 来筛选某些错误并允许显示其他错误。您可以使用 -d 选项实现在交互模式或守护进程模式中运行 storageMonitor。

## 选项

您可以使用以下选项之一从 ESX Server 命令行调用 storageMonitor。

**表 B-1. storageMonitor 命令行选项**

选项	描述
<config-file>	允许您指定一个配置文件。如果未指定此选项，将使用默认值。配置文件可指定 storageMonitor 应允许何种类型的错误，以及在显示错误之前应筛选哪些错误。默认配置文件将说明条目的格式。
-d	指定 storageMonitor 应在守护进程模式中运行。指定了此选项后，所有输出将发送至 syslog 或用户指定的日志文件。如果同时指定了 -s 选项，输出还将写入标准输出。
-h	显示帮助信息。
-l <log_file>	指定了此选项后，程序的输出将写入 <log_file>。仅当同时指定了 -d 选项时此选项才有效。
-p <poll_interval>	允许您指定时间间隔（单位为秒），用于轮询内核驻留存储器以及检索存储设备的状态或错误。如果未指定此选项，则使用 10 秒的默认轮询时间间隔。
-s	指定 storageMonitor 应将输出发送至标准输出。仅当在守护进程模式中启动 storageMonitor（已指定 -d 选项）时此选项才有效。

## 示例

```
storageMonitor -p 60
```

将轮询时间间隔设置为 60 秒。将输出发送至标准输出（因为 `storageMonitor` 不是在守护进程模式中运行）。在发送输出前使用默认配置文件中指定的筛选器。

```
storageMonitor -d -c myconf.conf
```

使用配置文件 `myconf.conf` 在守护进程模式中运行 `storageMonitor`。将输出写入 `syslog`。默认情况下，`syslog` 位于 `/var/log/storageMonitor`。

```
storageMonitor -d -l mylog.log -s
```

使用默认配置文件在守护进程模式中运行 `storageMonitor`。将输出发送至 `mylog.log`，而非 `syslog`。因为指定了 `-s` 选项，同时会将输出写入标准输出。



# 索引

## 符号

- .vmdk 文件 17
- [固定的 (Fixed)] 路径策略 37, 88
  - 路径抖动 99
  - 首选路径 90
- [管理路径 (Manage Paths)] 对话框 90
- [最近使用 (Most Recently Used)] 路径策略 88, 89
  - 路径抖动 99

## A

- AVT 56, 99
- AX100
  - 非活动连接 52
  - 显示问题 52
- 安装
  - 步骤 47
  - SAN 47
  - 准备从 SAN 引导 65

## B

- BIOS
  - 为 BFS 启用 70
  - 为 BFS 启用 Qlogic HBA 68
- BusLogic
  - 队列深度 44
  - SCSI 控制器 16
- 被动磁盘阵列 45, 58, 69, 88
  - 路径抖动 99
- 备份 44
  - 第三方备份软件包 104

- 解决方案 104
- 注意事项 103
- 备用路径状态 87
- 避免问题 95

## C

- CD-ROM, 引导方式 70
- changing disk.maskLuns 81
- changing disk.maxLun 80
- Consolidated Backup 代理 44
- CPU 虚拟化 16
- 参数, lpfc\_nodedev\_tmo 94
- 操作系统超时 94
- 测试 50
- 查看信息 27
- 超时 93, 94
- 磁带设备 46
- 磁盘, 配置选项 17
- 磁盘访问, 平衡 100
- 磁盘份额 34
- 磁盘阵列
  - 针对 ESX Server 对磁盘阵列进行区域分配 78
  - 主动 / 被动 45, 58, 69, 88, 99
  - 主动 / 主动 45, 88, 91
- 从 CD-ROM 引导 70
- 从 SAN 引导
  - Emulex FC HBA 70
  - ESX Server 要求 47
  - 概念性综述 64
  - HBA 要求 47
  - 简介 64

- LUN 屏蔽 **66**
- Qlogic FC HBA **68**
- 启用 Qlogic HBA BIOS **68**
- 要求 **46**
- 引导 LUN 注意事项 **47**
- 优点 **65**
- 诊断分区 **66**
- 准备安装 **65**
- 存储处理器
  - 端口配置 **55**
  - 检测数据 **56**
  - 配置检测数据 **56**
- 存储器选择 **32**
- 存储设备
  - 查看 **76**
  - 详细信息 **76**
- 存储系统
  - EMC CLARiiON **52**
  - EMC Symmetrix **53**
  - Hitachi **60**
  - HP StorageWorks **58**
  - Network Appliance **61**
- 存储阵列
  - 性能 **96**
- 重新签名
  - 选项 **106**
- 重新扫描 **78, 79**
  - LUN 创建 **78**
  - LUN 的显示 **29**
  - LUN 屏蔽 **79**
  - 路径发生故障时 **79**
  - 添加磁盘阵列 **78**
  - 重插拔线缆 **79**

## D

- disallowSnapshotLUN **107**
- disk.maskLuns **80, 81**

- disk.maxLun **80**
- Disk.SchedNumReqOutstanding 参数 **100**
- disk.supportSparseLun **81**
- DRS **41**
- DS4000, 配置用于 SAN 故障切换的硬件 **54**
- 待处理磁盘请求 **100**
- 待处理磁盘请求的数目 **100**
- 当前的多路径状况 **86**
- 低端存储器 **38**
- 第三方备份软件包 **104**
- 第三方管理应用程序 **31**
- 端口, 配置 **55**
- 端口地址 **25**
- 队列深度 **101**
- 多个扩展 **76**
- 多路径 **85**
- 多路径策略 **88**
- 多路径软件 **30**
- 多路径状况 **86**

## E

- EMC CLARiiON **52**
- EMC Symmetrix **53**
  - 伪 LUN **53**
- Emulex FC HBA
  - 从 SAN 引导 **70**
  - lpfc\_linkdown\_tmo **94**
  - NPIV 支持 **83**
  - 设备驱动程序选项 **94**
- enableResignature **107**
- ESX Server
  - 共享 VMFS **28**
  - 简介 **14**
  - 优点 **26**
- ESX Server 3 **14, 64, 95, 98**
- ESX Server 3i **14, 98**
- ESX Server 和 SAN, 要求 **44**
- esxtop 实用程序 **112**

EVA (HP StorageWorks) 60

## F

FC HBA 设置 45

访问

    控制 32

    平衡磁盘访问 100

分布式锁定 18

分层应用程序 105

分配 45

服务控制台 14, 20, 64, 95, 98

    界面 20

    身份验证 20

服务器故障 38

服务器故障切换 40

服务器性能 97

负载均衡 27, 45

    手动 91

## G

高端存储器 38

高级 LUN 显示配置 79

更改 disk.supportSparseLun 81

共享诊断分区 95

故障 38

故障排除 74, 95

故障切换 36, 40, 92

    FASTT 存储器 54

    HBA 93

    I/O 延迟 37

    区域 32

管理应用程序 31

## H

HA 39

HBA

    超时 93

    队列深度 101

Emulex 70, 83

静态负载平衡 45

类型 75

类型列表 75

Qlogic 68, 83

设置 45

    为 BFS 启用 Qlogic HBA BIOS 68

HBA 类型列表 75

Hitachi Data Systems 存储器 60

    微码 60

hostd 20

HP StorageWorks 58

    EVA 60

    MSA 58

    XP 60

活动路径状态 87

获取信息 75

## I

I/O 延迟 37, 44

IBM TotalStorage DS4000 54

    路径抖动 57

IBM TotalStorage Enterprise 存储系统 58

iSCSI 18

ISL 55

## J

集线器控制器问题 59

机箱间群集 39

机箱内群集 39

基于阵列（第三方）的解决方案 105

间接 30

间接级别 30

监视性能 98

交换机内链路 55

解决方案 74

解决问题 95

禁用路径 89, 90

禁用自动卷传输 **56**

卷重新签名 **106**

## **K**

看不到 LUN **77**

可插入验证模块 **20**

可见性问题 **77**

可选式引导, 启用 **69**

跨服务器共享 VMFS **28**

快照 **106**

快照软件 **103**

扩展 **18**

    定义 **76**

扩展数目 **18**

## **L**

LDAP **21**

Linux

    服务控制台 **20**

    Profile Name **58**

    VMkernel **14**

    主机类型 **51**

Linux Cluster 主机类型 **51**

lpfc\_linkdown\_tmo 参数 **94**

lpfc\_nodedev\_tmo 参数 **94**

LSI Logic SCSI 控制器 **16**

LSILogic 队列深度 **44**

LUN

    1 个 VMFS 卷 **44**

    创建, 和重新扫描 **78**

    disk.maskLuns **80**

    多路径策略 **88**

    分配 **45**

    更改扫描的数目 **80**

    基于 NPIV 的访问 **82**

    较少, 较大与较多, 较小 **32**

    决定 **33**

    看不到 **77**

    屏蔽更改和重新扫描 **79**

    扫描数目 **80**

    设置多路径策略 **88**

    稀疏 **81**

    显示配置 **79**

    显示与重新扫描 **29**

    选择引导 LUN **69**

    移除 **79**

    引导 LUN **69**

LUN 不可见

    屏蔽 **78**

    区域分配 **78**

    SP 可见性 **78**

    问题 **78**

    线缆连接 **78**

LUN 发现, VMkernel **29**

LUN 屏蔽 **32, 78**

    从 SAN 引导 **66**

路径

    禁用 **89, 90**

    启用 **89, 90**

    首选 **88, 90**

路径策略

    固定的 **37, 88**

    MRU **88**

    循环 **89**

    最近使用 **89**

路径策略重置

    主动 / 被动磁盘阵列 **75**

路径抖动 **52, 98, 99**

    IBM TotalStorage DS4000 **57**

路径故障切换 **36**

路径故障重新扫描 **79**

路径管理 **36, 91**

路径旁边的 \* **88**

路径旁边的星号 **88**

路径状态 **87**

裸机映射 35  
 裸设备映射 18, 52  
     映射文件 18

## M

maxLun 80  
 Microsoft 群集服务 18, 26, 50  
 MRU 路径策略 88  
 MSA (HP StorageWorks) 58  
 MSCS 26, 50  
 命令, SDK 19

## N

N+1 群集 39  
 net-snmpd 21  
 Netware 主机模式 60  
 Network Appliance 存储器 61  
     置备存储器 61  
 NFS 18  
 N-Port ID 虚拟化 (NPIV) 25, 82  
     要求 82  
 内存虚拟化 16

## P

PAM 20  
 Port\_ID 25  
 PortDownRetryCount 参数 93  
 Profile Name, Linux 58  
 配置  
     存储处理器检测数据 56  
     选项 17  
 配置用于 SAN 故障切换的硬件  
     DS4000 54  
 屏蔽 32  
     LUN 不可见 78  
     使用 disk.maskLuns 80  
 平衡磁盘访问 100

## Q

Qlogic FC HBA  
     从 SAN 引导 68  
     NPIV 支持 83  
     PortDownRetryCount 93  
     设备驱动程序选项 94  
 Qlogic HBA BIOS, 为 BFS 启用 68  
 启用可选式引导 69  
 启用路径 89, 90  
 驱动程序  
     设备驱动程序 94  
     VMFS-2 100  
 区分虚拟机优先级 34  
 区域  
     故障切换 32  
     何时创建 32  
     建议大小 32  
 区域大小 32  
 区域分配  
     和 ESX Server 31  
     LUN 不可见 78  
 全球名称 (WWN) 82  
     全球端口名称 (WWPN) 25, 56, 82  
     全球节点名称 (WWNN) 82  
     向虚拟机分配 83  
 群集服务 39

## R

RADIUS 21  
 RDM 18, 31, 35, 52  
     Microsoft 群集服务 18  
     映射文件 18  
 resxtp 98  
 软件兼容性 15

## S

SAN 32

- 安装注意事项 47
- 备份注意事项 103
- 服务器故障切换 40
- 要求 44
- 硬件故障切换 54
- 阵列 28
- 准备 67
- 准备安装 ESX Server 65
- SAN 存储器, 优点 26
- SAN 存储性能, 优化 96
- SAN 架构 24
- SCSI 控制器, 设备驱动程序选项 94
- SCSI 预留 28
- SCSI 预留, 减少 101
- SDK 19
- SNMP 服务器 21
- SP 可见性, LUN 不可见 78
- storageMonitor 实用程序 112
- 扫描, 更改数目 80
- 设备驱动程序 14
- 设备驱动程序选项 94
  - Emulex 94
  - Qlogic 94
- 设计
  - 针对服务器故障 38
- 设置步骤 47
- 身份验证守护进程 20
- 失效路径 87
- 实用程序
  - esxtop 112
  - storageMonitor 112
- 手动负载平衡 91
- 首选路径 88, 90
- 数据存储, 移除 79
- 数据访问 34
  - RDM 31
  - VMFS 31
- 锁定 18

元数据更新 29

## T

TimeoutValue 参数 44  
提示 34

## V

VI Client 14, 19  
VI 浏览器访问 14, 19  
Virtual Infrastructure SDK 19  
Virtual Machine File System 18  
vmauthd 20  
VMFS 18, 31, 32

- 创建新卷 18
- 大小下限 18
- 跨 ESX Server 共享 28
- 扩展数目 18
- 每个 LUN 1 个卷 44
- SCSI 预留 28
- 锁定 18

VMFS 卷重新签名 106  
VMFS-2 驱动程序 100  
vmhba 77  
VMkernel 14

- LUN 发现 29
- 配置 95

VMM 14  
VMotion 26, 27, 40, 45  
VMware Consolidated Backup 代理 44  
VMware DRS 27, 41  
VMware HA 26, 39  
VMware Infrastructure Client 14

## W

网络虚拟化 16  
为 BFS 启用 BIOS 70  
为 BFS 启用引导 BIOS 提示 70  
维护 27

微码, Hitachi Data Systems 存储器 **60**

问题 **74**

    避免 **95**

    集线器控制器 **59**

    可见性 **77**

    性能 **98**

物理到虚拟群集 **39**

## X

XP (HP StorageWorks) **60**

稀疏 LUN 支持 **81**

线缆

    连接问题 **78**

    重插拔和重新扫描 **79**

显示问题, AX100 **52**

限制 **44**

性能 **96, 97**

    监视 **98**

    SCSI 预留 **28**

    问题 **98**

    移除 VMFS-2 驱动程序 **100**

    优化 **96**

虚拟 SCSI HBA **16, 18**

虚拟端口 (VPORT) **82**

虚拟化 **15, 16**

    概览 **21**

虚拟机

    分配 WWN **83**

    共享设备 **21**

    I/O 延迟 **37**

    默认配置 **30**

    平衡磁盘访问 **100**

    区分优先级 **34**

    位置 **38**

虚拟机的位置 **38**

虚拟机监视器 **14**

循环路径策略 **89**

## Y

要求 **44**

    从 SAN 引导 **46**

移除 LUN **79**

移除 VMFS-2 驱动程序 **100**

移除数据存储 **79**

已禁用路径状态 **87**

引导 BIOS 提示, 为 BFS 启用 **70**

引导 LUN **69**

引导 LUN, 选择 **69**

隐患 **75**

硬件兼容性 **15**

映射文件 **18**

应用程序, 分层 **105**

用例 **27**

优点 **26**

有关存储设备的详细信息 **76**

优化资源利用率 **40**

与 ESX Server 系统交互 **19**

预留

    减少 SCSI 预留 **101**

元数据更新 **29**

## Z

灾难恢复 **27**

诊断分区 **44**

    从 SAN 引导 **66**

    共享 **95**

支持的设备 **50**

直接连接 **50**

中端存储器 **38**

主动 / 被动磁盘阵列 **37, 45, 69, 88**

    从 SAN 引导 **47**

    HP StorageWorks MSA **58**

    路径策略重置 **75**

    路径抖动 **99**

主动 / 主动磁盘阵列 **45, 58, 88, 91**

主机类型 **51**

主机守护进程 **20**

转储分区 **44**

    共享 **95**

自动卷传输 **56**

资源利用率, 优化 **40**

最大 HBA 队列深度 **101**