

# 《vSphere 存储》

VMware vSphere 8.0

VMware ESXi 8.0

vCenter Server 8.0

您可以从 VMware 网站下载最新的技术文档:

<https://docs.vmware.com/cn/>。

**VMware, Inc.**  
3401 Hillview Ave.  
Palo Alto, CA 94304  
[www.vmware.com](http://www.vmware.com)

**威睿信息技术（中国）有  
限公司**  
北京办公室  
北京市  
朝阳区新源南路 8 号  
启皓北京东塔 8 层 801  
[www.vmware.com/cn](http://www.vmware.com/cn)

上海办公室  
上海市  
淮海中路 333 号  
瑞安大厦 804-809 室  
[www.vmware.com/cn](http://www.vmware.com/cn)

广州办公室  
广州市  
天河路 385 号  
太古汇一座 3502 室  
[www.vmware.com/cn](http://www.vmware.com/cn)

版权所有 © 2009-2022 VMware, Inc. 保留所有权利。 [版权和商标信息](#)

# 目录

关于 vSphere 存储 13

## 1 vSphere 环境中的存储简介 14

vSphere 环境中的传统存储虚拟化模型 14

软件定义的存储型号 15

vSphere Storage API 16

## 2 vSphere 环境中的传统存储型号入门 18

ESXi 支持哪些类型的物理存储 18

本地存储 18

联网的存储 19

比较存储类型 23

目标和设备表示形式 23

虚拟机如何访问存储 24

查看可用于 ESXi 主机的存储设备 25

显示 ESXi 主机的存储设备 26

显示适配器的存储设备 27

查看 ESXi 主机上可用的存储适配器 28

ESXi 数据存储的特性 29

显示数据存储信息 30

将持久性内存设备与 ESXi 配合使用 31

监控 PMem 数据存储统计信息 33

将 ESXi 与 SAN 配合使用 34

ESXi 和 SAN 用例 34

将 SAN 存储与 ESXi 配合使用的细节 35

决定 LUN 的大小和数目 35

选择虚拟机位置 36

第三方管理应用程序 37

SAN 存储备份注意事项 37

## 3 将 ESXi 与光纤通道 SAN 配合使用 39

设置 ESXi 光纤通道环境 42

ESXi 光纤通道 SAN 限制 42

设置 LUN 分配 42

设置光纤通道 HBA 43

配置以太网光纤通道 43

设置工作流 43

从光纤通道 SAN 引导 ESXi	44
从 SAN 引导的优点	44
从光纤通道 SAN 引导时的要求和注意事项	44
准备从 SAN 引导 ESXi	45
光纤通道存储最佳做法	46
防止出现光纤通道 SAN 问题	47
禁用自动 ESXi 主机注册	47
优化光纤通道 SAN 存储器性能	48
将光纤通道 NPIV 与 vSphere 虚拟机结合使用	49
基于 NPIV 的 LUN 访问如何运作	49
使用 NPIV 的要求	49
NPIV 功能和限制	50
配置或修改 WWN 分配	50
<b>4 将 ESXi 与 iSCSI SAN 配合使用</b>	<b>52</b>
从 iSCSI SAN 引导	58
关于从 iSCSI SAN 引导的常规建议	59
准备 iSCSI SAN	59
配置独立硬件 iSCSI 适配器进行 SAN 引导	60
iSCSI 存储的最佳做法	61
防止出现 iSCSI SAN 问题	61
优化 iSCSI SAN 存储器性能	62
检查以太网交换机统计信息	65
<b>5 配置 iSCSI 和 iSER 适配器与存储</b>	<b>66</b>
设置独立硬件 iSCSI 适配器	68
查看独立硬件 iSCSI 适配器	69
编辑硬件 iSCSI 的网络设置	69
配置从属硬件 iSCSI 适配器	70
从属硬件 iSCSI 注意事项	71
查看从属硬件 iSCSI 适配器	71
确定 iSCSI 与网络适配器之间的关联	72
配置软件 iSCSI 适配器	72
激活或禁用软件 iSCSI 适配器	73
使用 ESXi 配置 iSER	74
安装和查看支持 RDMA 的网络适配器	75
启用 VMware iSER 适配器	75
修改 iSCSI 或 iSER 适配器的常规属性	77
为 iSCSI 和 iSER 设置网络	78
iSCSI 或 iSER 配置中的多个网络适配器	79
使用软件 iSCSI 配置网络连接的最佳做法	81

- 管理 iSCSI 网络 84
- iSCSI 网络故障排除 85
- 配置 iSCSI 或 iSER 端口绑定 85
  - 为 iSCSI 或 iSER 创建一个 VMkernel 适配器 86
  - 将 iSCSI 或 iSER 适配器绑定到 VMkernel 适配器 87
  - 查看 ESXi 主机上的端口绑定详细信息 88
- 将巨帧与 iSCSI 和 iSER 配合使用 88
  - 为网络连接启用巨帧 89
  - 为独立硬件 iSCSI 启用巨帧 89
- 在 ESXi 主机上配置 iSCSI 和 iSER 的动态或静态发现 90
  - 移除动态或静态 iSCSI 目标 91
- 为 iSCSI 或 iSER 存储适配器配置 CHAP 参数 92
  - 选择 CHAP 身份验证方法 92
  - 为 iSCSI 或 iSER 存储适配器设置 CHAP 93
  - 设置目标的 CHAP 94
- 配置 iSCSI 的高级参数 95
  - 为 ESXi 主机上的 iSCSI 配置高级参数 97
- iSCSI 会话管理 97
  - 查看 iSCSI 会话 98
  - 添加 iSCSI 会话 98
  - 移除 iSCSI 会话 99

## 6 管理存储设备 100

- ESXi 存储设备名称和标识符 100
  - 具有 NGUID 设备标识符的 NVMe 设备 101
  - 重命名存储设备 102
- ESXi 存储的重扫描操作 102
  - 执行存储重新扫描 103
  - 执行适配器重新扫描 103
  - 更改扫描的存储设备的数量 104
- 确定 ESXi 存储连接问题 104
  - 检测 PDL 情况 105
  - 执行计划的存储设备移除 106
  - 从 PDL 情况中恢复 107
  - 处理暂时性 APD 情况 107
  - 验证 ESXi 主机上存储设备的连接状态 109
- 启用或禁用 ESXi 存储设备上的定位符 LED 110
- 擦除存储设备 110
- 更改永久预留设置 111

## 7 对 ESXi 主机使用闪存设备 113

ESXi 主机上闪存设备的最佳做法	114
监控闪存设备	114
估算闪存设备的生命周期	114
保持闪存磁盘无 VMFS	115
标记 ESXi 主机上的存储设备	116
将存储设备标记为闪存设备	116
将存储设备标记为本地	117
关于 vSphere 环境中的虚拟闪存资源	117
虚拟闪存资源的注意事项	117
设置虚拟闪存资源	118
移除虚拟闪存资源	119
设置虚拟闪存使用率警报	119
使用 VMFS 数据存储配置主机缓存	119

## 8 关于 VMware NVMe 存储 121

VMware NVMe 概念	121
基本 VMware NVMe 架构和组件	122
VMware NVMe 存储的要求和限制	125
为 NVMe over RDMA 配置无损以太网	127
为 NVMe over RDMA (RoCE v2) 存储配置适配器	128
查看 RDMA 网络适配器	129
为 RDMA 适配器配置 VMkernel 绑定	129
为 NVMe over TCP 存储配置适配器	136
为 NVMe over TCP 适配器配置 VMkernel 绑定	136
启用 NVMe over RDMA 或 NVMe over TCP 软件适配器	142
为 NVMe over Fabrics 添加控制器	143
移除 NVMe over RDMA 或 NVMe over TCP 软件适配器	144

## 9 使用数据存储 145

了解 VMFS 数据存储	146
VMFS 数据存储的版本	146
作为存储库的 VMFS 数据存储	148
在主机间共享 VMFS 数据存储	149
VMFS 元数据更新	149
VMFS 上的快照格式	150
升级 VMFS 数据存储	151
VMFS 锁定机制	151
显示 VMFS 锁定信息	151
VMFS 的“仅限 ATS”锁定机制	152
将锁定机制更改为 ATS+SCSI	154
了解网络文件系统数据存储	154

- NFS 存储准则和要求 156
- NFS 存储的防火墙配置 159
- 使用第 3 层路由连接访问 NFS 存储 160
- 对 NFS 4.1 使用 Kerberos 161
- 设置 NFS 存储环境 162
- 配置 ESXi 主机进行 Kerberos 身份验证 162
- 收集 NFS 存储的统计信息 165
- 创建数据存储 165
  - 创建 VMFS 数据存储 166
  - 创建 NFS 数据存储 167
  - 创建 Virtual Volumes 数据存储 168
- 管理重复 VMFS 数据存储 169
  - 挂载 VMFS 数据存储副本 170
- 增加 VMFS 数据存储容量 170
- 在 VMFS6 数据存储上启用或禁用对集群虚拟磁盘的支持 172
- 数据存储的管理操作 172
  - 更改数据存储名称 172
  - 卸载数据存储 173
  - 挂载数据存储 174
  - 移除 VMFS 数据存储 174
  - 使用数据存储浏览器 175
  - 关闭存储筛选器 178
- 设置动态磁盘镜像 179
- 在 VMFS 数据存储上收集 ESXi 主机的诊断信息 180
  - 将文件设置为核心转储位置 180
  - 取消激活和删除核心转储文件 182
- 使用 VOMA 检查元数据一致性 183
  - 使用 VOMA 检查元数据一致性 185
- 配置 VMFS 指针块缓存 186
  - 获取 VMFS 指针块缓存的信息 187
  - 更改指针块缓存的大小 187

## 10 了解多路径和故障切换 189

- 光纤通道故障切换 189
- 基于主机的 iSCSI 故障切换 190
- 基于阵列的 iSCSI 故障切换 192
- 路径故障切换和虚拟机 193
  - 在 Windows 客户机操作系统上设置超时 193
- 可插入存储架构和路径管理 194
  - 关于可插入存储架构 195
  - VMware 本机多路径插件 196

- 路径选择插件和策略 198
- VMware SATP 199
- VMware 高性能插件和路径选择方案 200
- 查看和管理路径 207
  - 查看存储设备路径 207
  - 查看数据存储路径 208
  - 更改路径选择策略 209
  - 更改延迟循环的默认参数 209
  - 禁用存储路径 210
- 使用声明规则 211
  - 多路径注意事项 211
  - 列出主机的多路径声明规则 212
  - 添加多路径声明规则 213
  - 删除多路径声明规则 216
  - 屏蔽路径 217
  - 取消路径屏蔽 218
  - 定义 NMP SATP 规则 219
- 虚拟机 I/O 的调度队列 220
  - 在 vSphere Client 中编辑按文件 I/O 调度 220
  - 使用 esxcli 命令启用或禁用按文件 I/O 调度 221

## 11 裸设备映射 222

- 关于裸设备映射 222
  - 裸设备映射的优点 223
  - RDM 注意事项和限制 225
- 裸设备映射特性 225
  - RDM 虚拟兼容模式和物理兼容模式 225
  - 动态名称解析 226
  - 虚拟机集群的裸设备映射 226
  - 比较可用的 SCSI 设备访问模式 226
- 使用 RDM 创建虚拟机 227
- 管理映射的 LUN 的路径 228
- 具有 RDM 的虚拟机必须忽略 SCSI INQUIRY 缓存 229

## 12 基于存储策略的管理 231

- 虚拟机存储策略 232
- 虚拟机存储策略的工作流 232
- 填充“虚拟机存储策略”界面 233
  - 使用存储提供程序填充“虚拟机存储策略”界面 234
  - 向数据存储分配标记 234
- 关于规则和规则集 236



创建和管理虚拟机存储策略	238
为基于主机的数据服务创建虚拟机存储策略	238
为 Virtual Volumes 创建虚拟机存储策略	239
为基于标记的放置创建虚拟机存储策略	241
编辑或克隆虚拟机存储策略	242
关于存储策略组件	243
创建存储策略组件	244
编辑或克隆存储策略组件	245
存储策略和虚拟机	245
将存储策略分配给虚拟机	245
更改虚拟机文件和磁盘的存储策略分配	246
检查虚拟机存储策略的合规性	247
为不合规虚拟机查找兼容存储资源	248
重新应用虚拟机存储策略	249
默认存储策略	249
更改数据存储的默认存储策略	250
<b>13 使用存储提供程序</b>	<b>251</b>
关于存储提供程序	251
存储提供程序和数据呈现形式	252
存储提供程序要求和注意事项	253
注册存储提供程序	253
查看存储提供程序信息	254
管理存储提供程序	254
<b>14 使用 VMware vSphere Virtual Volumes</b>	<b>256</b>
关于 Virtual Volumes	256
Virtual Volumes 概念	257
虚拟卷对象	257
Virtual Volumes 存储提供程序	259
Virtual Volumes 存储容器	259
协议端点	259
绑定和解除绑定 Virtual Volumes	260
Virtual Volumes 数据存储	261
Virtual Volumes 和虚拟机存储策略	261
Virtual Volumes 和存储协议	261
Virtual Volumes 架构	263
Virtual Volumes 和 VMware Certificate Authority	264
虚拟卷快照	265
在启用 Virtual Volumes 之前	265
将 vSphere Storage 环境与网络时间服务器同步	266

- 配置 Virtual Volumes 267
  - 为 Virtual Volumes 注册存储提供程序 267
  - 创建 Virtual Volumes 数据存储 268
  - 查看和管理协议端点 269
  - 更改协议端点的路径选择策略 269
- 配置 NVMe-oF-vVols 270
- 在 Virtual Volumes 数据存储中置备虚拟机 271
- Virtual Volumes 和复制 272
  - Virtual Volumes 复制的要求 272
  - Virtual Volumes 和复制组 273
  - Virtual Volumes 和故障域 273
  - Virtual Volumes 复制工作流 275
  - 复制准则和注意事项 275
- 使用 Virtual Volumes 的最佳做法 276
  - 使用 Virtual Volumes 时的准则和限制 276
  - 存储容器置备最佳做法 278
  - Virtual Volumes 性能最佳做法 279
- 对 Virtual Volumes 进行故障排除 280
  - Virtual Volumes 和 esxcli 命令 280
  - 收集 Virtual Volumes 的统计信息 280
  - Virtual Volumes 数据存储无法访问 281
  - 将虚拟机迁移到 Virtual Volumes 数据存储或将 VM OVF 部署到该数据存储时失败 282

## 15 筛选虚拟机 I/O 283

- 关于 I/O 筛选器 283
  - I/O 筛选器的类型 284
  - I/O 筛选组件 284
  - I/O 筛选器的存储提供程序 285
- 使用闪存设备与缓存 I/O 筛选器 286
- I/O 筛选器系统要求 286
- 在 vSphere 环境中配置 I/O 筛选器 287
  - 在集群中安装 I/O 筛选器 287
  - 查看 I/O 筛选器和存储提供程序 288
- 在虚拟磁盘上启用 I/O 筛选器数据服务 288
  - 将 I/O 筛选器策略分配给虚拟机 289
- 管理 I/O 筛选器 290
  - 从集群卸载 I/O 筛选器 290
  - 升级集群中的 I/O 筛选器 291
- I/O 筛选器准则和最佳实践 291
  - 通过 I/O 筛选器迁移虚拟机 292
- 处理 I/O 筛选器安装故障 292

在单个 ESXi 主机中安装 I/O 筛选器 292

## 16 存储硬件加速 294

硬件加速的优点 294

硬件加速要求 295

硬件加速支持状态 295

块存储设备的硬件加速 295

禁用块存储设备的硬件加速 296

管理块存储设备上的硬件加速 296

NAS 设备上的硬件加速 301

在虚拟机上启用 NAS 本机快照 302

硬件加速注意事项 302

## 17 存储置备和空间回收 304

虚拟磁盘精简置备 304

关于虚拟磁盘置备策略 305

创建精简置备虚拟磁盘 306

查看虚拟机存储资源 306

确定虚拟机的磁盘格式 307

扩充精简虚拟磁盘 307

处理数据存储超额订购 308

ESXi 和阵列精简置备 308

监控空间使用情况 309

识别精简置备的存储设备 309

存储空间回收 310

VMFS 数据存储上的空间回收 311

来自客户机操作系统的空间回收请求 316

## 18 Cloud Native Storage 入门指南 318

Cloud Native Storage 概念和术语 319

Cloud Native Storage 组件 321

使用 vSAN 文件服务置备文件卷 323

Cloud Native Storage 用户 325

适用于 vSphere 管理员的 Cloud Native Storage 325

Cloud Native Storage 的要求和限制 325

Cloud Native Storage 角色和特权 328

为 Kubernetes 创建存储策略 329

配置 Kubernetes 群集虚拟机 331

监控 Kubernetes 集群中的容器卷 331

将加密与云原生存储结合使用 333

## 19 使用 vmkfstools 334

vmkfstools 命令语法 334

vmkfstools 命令选项 335

-v 子选项 335

文件系统选项 336

虚拟磁盘选项 338

存储设备选项 344

# 关于 vSphere 存储

《vSphere 存储》介绍 VMware ESXi™ 和 VMware vCenter Server® 提供的虚拟化和软件定义的存储技术，并说明如何配置和使用这些技术。

VMware 非常重视包容性。为了在客户、合作伙伴和内部社区中促进这一原则，我们采用包容性语言创建内容。

## 目标读者

本信息的目标读者为熟悉虚拟机和存储虚拟化技术、数据中心操作以及 SAN 存储概念且具有丰富经验的系统管理员。

# vSphere 环境中的存储简介

# 1

vSphere 支持传统和软件定义的存储环境中的各种存储选项和功能。vSphere 存储元素和各方面的高级概述可帮助您为虚拟数据中心规划合适的存储策略。

本章讨论了以下主题：

- [vSphere 环境中的传统存储虚拟化模型](#)
- [软件定义的存储型号](#)
- [vSphere Storage API](#)

## vSphere 环境中的传统存储虚拟化模型

通常，存储虚拟化是指从虚拟机及其应用程序中对物理存储资源和容量进行逻辑虚拟化。ESXi 提供主机级别的存储虚拟化。

在 vSphere 环境中，传统型号是围绕以下存储技术以及 ESXi 和 vCenter Server 虚拟化功能构建的。

### 本地存储和联网存储

在传统存储环境中，ESXi 存储管理过程以存储管理员在不同存储系统上预先分配的存储空间开始。

ESXi 支持本地存储和联网存储。

请参见 [ESXi 支持哪些类型的物理存储](#)。

### 存储区域网络

存储区域网络 (SAN) 是一种将计算机系统（或称 ESXi 主机）连接到高性能存储系统的专用高速网络。ESXi 可使用光纤通道或 iSCSI 协议连接到存储系统。

请参见将 [ESXi 与 SAN 配合使用](#)。

### 光纤通道

光纤通道 (FC) 是一种存储协议，SAN 使用该协议将数据流量从 ESXi 主机服务器传输到共享存储。该协议将 SCSI 命令打包到 FC 帧中。要连接到 FC SAN，主机要使用光纤通道主机总线适配器 (HBA)。

请参见第 3 章 将 [ESXi 与光纤通道 SAN 配合使用](#)。

### Internet SCSI

Internet iSCSI (iSCSI) 是一种可在计算机系统（或称 ESXi 主机）和高性能存储系统之间使用以太网连接的 SAN 传输。要连接到存储系统，主机要将硬件 iSCSI 适配器或软件 iSCSI 启动器与标准网络适配器搭配使用。

请参见第 4 章 [将 ESXi 与 iSCSI SAN 配合使用](#)。

## 存储设备或 LUN

在 ESXi 环境中，术语“设备”和“LUN”可互换使用。通常，这两个术语都表示通过块存储系统向主机呈现并可用于进行格式化的存储卷。

请参见[目标和设备表示形式](#) 和第 6 章 [管理存储设备](#)。

## 虚拟磁盘

ESXi 主机上的虚拟机使用虚拟磁盘来存储其操作系统、应用程序文件，以及与其活动关联的其他数据。虚拟磁盘是较大的物理文件或文件集，可以像处理任何其他文件那样复制、移动、存档和备份虚拟磁盘。您可以配置具有多个虚拟磁盘的虚拟机。

要访问虚拟磁盘，虚拟机需使用虚拟 SCSI 控制器。这些虚拟控制器包括 BusLogic 并行、LSI Logic 并行、LSI Logic SAS 和 VMware 准虚拟。虚拟机只能查看和访问以上类型的 SCSI 控制器。

每个虚拟磁盘都位于物理存储上部署的一个数据存储上。从虚拟机的角度而言，每个虚拟磁盘看上去都好像是与 SCSI 控制器连接的 SCSI 驱动器。物理存储是通过主机上的存储适配器还是网络适配器访问，这对于虚拟机客户机操作系统和应用程序而言通常是透明的。

## VMware vSphere® VMFS

块存储设备上部署的数据存储使用本机 vSphere 虚拟机文件系统 (Virtual Machine File System, VMFS) 格式。该格式是一种针对存储虚拟机而优化的特殊高性能文件系统格式。

请参见[了解 VMFS 数据存储](#)。

## NFS

ESXi 中内置的 NFS 客户端使用网络文件系统 (NFS) 协议通过 TCP/IP 访问位于 NAS 服务器上的 NFS 卷。ESXi 主机可以挂载卷，并将其用作 NFS 数据存储。

请参见[了解网络文件系统数据存储](#)。

## 裸设备映射

除虚拟磁盘外，vSphere 还提供称为裸设备映射 (RDM) 的机制。在虚拟机内部的客户机操作系统需要对存储设备的直接访问权限时，RDM 非常有用。有关 RDM 的信息，请参见第 11 章 [裸设备映射](#)。

# 软件定义的存储型号

除了像传统存储模型一样对虚拟机中的底层存储容量进行抽象，软件定义的存储也会对存储功能进行抽象。

通过软件定义的存储型号，虚拟机将成为存储置备单元，可以通过灵活的基于策略的机制进行管理。型号涉及以下 vSphere 技术。

## VMware vSphere® 虚拟卷™ (VVols)

Virtual Volumes 功能可将存储管理范式从管理数据存储内的空间更改为管理存储阵列处理的抽象存储对象。通过 Virtual Volumes，单个虚拟机将成为存储管理单元，而非数据存储。存储硬件可全面控制虚拟磁盘内容、布局和管理。

请参见第 14 章 使用 VMware vSphere Virtual Volumes。

## VMware vSAN

vSAN 是作为管理程序的一部分本机运行的分布式软件层。vSAN 可汇总 ESXi 主机集群的本地或直接连接容量设备，并创建在 vSAN 集群的所有主机之间共享的单个存储池。

请参见《管理 VMware vSAN》。

## 基于存储策略的管理

基于存储策略的管理 (Storage Policy Based Management, SPBM) 是一个框架，跨不同的数据服务和存储解决方案（包括 vSAN 和 Virtual Volumes）提供单一控制面板。该框架使用存储策略使虚拟机的应用程序需求与存储实体提供的功能保持一致。

请参见第 12 章 基于存储策略的管理。

## I/O 筛选器

I/O 筛选器是可以安装到 ESXi 主机上的软件组件，并可以向虚拟机提供其他数据服务。根据实施，服务可能包括复制、加密、缓存等。

请参见第 15 章 筛选虚拟机 I/O。

# vSphere Storage API

Storage API 是指一系列 API，第三方硬件、软件和存储提供程序使用这些 API 来开发可增强多项 vSphere 功能和解决方案的组件。

此存储出版物介绍了多个可用于存储环境的 Storage API。有关此系列中其他 API 的信息，包括 vSphere API - Data Protection，请参见 VMware 网站。

## vSphere APIs for Storage Awareness

这些 API 也称为 VASA。这些 API 由第三方供应商或 VMware 提供，用于实现 vCenter Server 与底层存储之间的通信。通过 VASA，存储实体可以将其配置、功能、存储运行状况和事件通知给 vCenter Server。同时，VASA 可以将 vCenter Server 的虚拟机存储要求传递给存储实体，并确保存储层满足这些要求。

使用 Virtual Volumes、vSAN、vSphere APIs for I/O Filtering (VAIO) 和存储虚拟机策略时，VASA 将会非常重要。请参见第 13 章 使用存储提供程序。

## vSphere APIs for Array Integration

这些 API 也称为 VAAI，包括以下组件：

- 硬件加速 API。帮助阵列与 vSphere 集成，以便 vSphere 可以将某些存储操作卸载给阵列。此集成显著降低了主机上的 CPU 开销。请参见第 16 章 存储硬件加速。



- 阵列精简置备 API。帮助监控精简置备存储阵列上的空间使用情况，以防止出现空间不足的情况，并执行空间回收。请参见 [ESXi 和阵列精简置备](#)。

## vSphere APIs for Multipathing

这些 API 也称为可插入存储架构 (Pluggable Storage Architecture, PSA)。利用这些 API，存储合作伙伴可以创建并提供针对各阵列优化的多路径和负载平衡插件。这些插件与存储阵列通信，并确定最佳路径选择策略，以提高从 ESXi 主机到存储阵列的 I/O 性能和可靠性。有关详细信息，请参见 [可插入存储架构和路径管理](#)。

# vSphere 环境中的传统存储型号入门

## 2

在传统环境中设置 ESXi 存储，包括配置存储系统和设备、启用存储适配器和创建数据存储。

本章讨论了以下主题：

- ESXi 支持哪些类型的物理存储
- 查看可用于 ESXi 主机的存储设备
- 查看 ESXi 主机上可用的存储适配器
- ESXi 数据存储的特性
- 将持久性内存设备与 ESXi 配合使用
- 将 ESXi 与 SAN 配合使用

## ESXi 支持哪些类型的物理存储

在传统存储环境中，ESXi 存储管理过程以存储管理员在不同存储系统上预先分配的存储空间开始。ESXi 支持本地存储和联网存储。

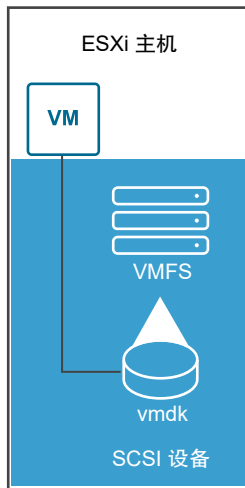
### 本地存储

本地存储可以是位于 ESXi 主机内部的内部硬盘，也可以包括位于主机之外并通过 SAS 或 SATA 等协议直接连接主机的外部存储系统。

本地存储不需要存储网络即可与主机进行通信。您需要一根连接到存储单元的电缆；必要时，主机中需要有一个兼容的 HBA。

下图描述了一台使用本地 SCSI 存储的虚拟机。

图 2-1. 本地存储



在这个本地存储拓扑示例中，ESXi 主机使用的是到存储设备的单一连接。可以在该设备上创建 VMFS 数据存储，以用于存储虚拟机磁盘文件。

虽然可以使用这种存储配置，但它不是最佳做法。如果在存储设备和主机间使用单一连接，则在连接不稳定或出现故障时，会产生将导致中断的单一故障点 (SPOF)。但是，由于大多数本地存储设备不支持多个连接，因此无法使用多个路径访问本地存储。

ESXi 支持各种本地存储设备，包括 SCSI、IDE、SATA、USB、SAS、闪存和 NVMe 设备。

---

**注** 不能使用 IDE/ATA 或 USB 驱动器来存储虚拟机。

---

本地存储不支持在多个主机之间共享。只有一个主机可以访问本地存储设备上的数据存储。因此，虽然可以使用本地存储创建虚拟机，但无法使用需要共享存储的 VMware 功能，如 HA 和 vMotion。

但是，如果您使用的是仅有本地存储设备的主机集群，则可以实施 vSAN。vSAN 可将本地存储资源转变为软件定义的共享存储。有了 vSAN，便可以使用需要共享存储的功能。有关详细信息，请参见《《管理 VMware vSAN》》文档。

## 联网的存储

联网的存储由 ESXi 主机用于远程存储虚拟机文件的外部存储系统组成。通常，主机通过高速存储网络访问这些系统。

网络存储设备将被共享。网络存储设备上的数据存储可同时由多个主机来访问。ESXi 支持多种网络存储技术。

除了本主题中介绍的传统网络存储，VMware 还支持虚拟共享存储（如 vSAN）。vSAN 可将 ESXi 主机的内部存储资源转变为可为虚拟机提供 High Availability 和 vMotion 等功能的共享存储。有关详细信息，请参见《《管理 VMware vSAN》》文档。

---

**注** 同一 LUN 无法通过不同存储协议提供给一个 ESXi 主机或多个主机。要访问 LUN，主机必须始终使用单一协议，例如仅使用光纤通道或仅使用 iSCSI。

---

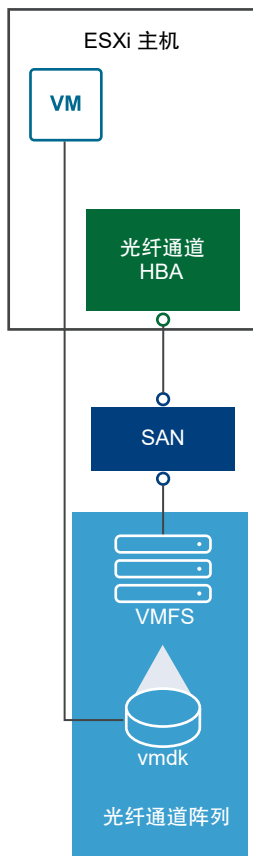
## 光纤通道 (FC)

在 FC 存储区域网络 (SAN) 上远程存储虚拟机文件。FC SAN 是一种将主机连接到高性能存储设备的专用高速网络。该网络使用光纤通道协议，将 SCSI 流量从虚拟机传输到 FC SAN 设备。

要连接到 FC SAN，您的主机应该配有光纤通道总线适配器 (HBA)。除非使用光纤通道直接连接存储，否则需要光纤通道交换机来路由存储流量。如果主机包含 FCoE（以太网光纤通道）适配器，则可以使用以太网网络连接共享光纤通道设备。

光纤通道存储描述了使用光纤通道存储的虚拟机。

图 2-2. 光纤通道存储



在该配置中，主机通过光纤通道适配器连接 SAN 架构（包括光纤通道交换机及存储阵列）。此时，存储阵列的 LUN 变得对于主机可用。您可以访问 LUN 并创建用于满足存储需求的数据存储。数据存储采用 VMFS 格式。

有关设置光纤通道 SAN 的特定信息，请参见第 3 章 [将 ESXi 与光纤通道 SAN 配合使用](#)。

## Internet SCSI (iSCSI)

在远程 iSCSI 存储设备上存储虚拟机文件。iSCSI 将 SCSI 存储流量打包在 TCP/IP 协议中，使其通过标准 TCP/IP 网络（而不是专用 FC 网络）传输。通过 iSCSI 连接，主机可以充当与位于远程 iSCSI 存储系统的目标进行通信的启动器。

ESXi 提供下列 iSCSI 连接类型：

### 硬件 iSCSI

主机通过能够卸载 iSCSI 和网络处理的第三方适配器连接到存储。硬件适配器可以是附属适配器，也可以是独立适配器。

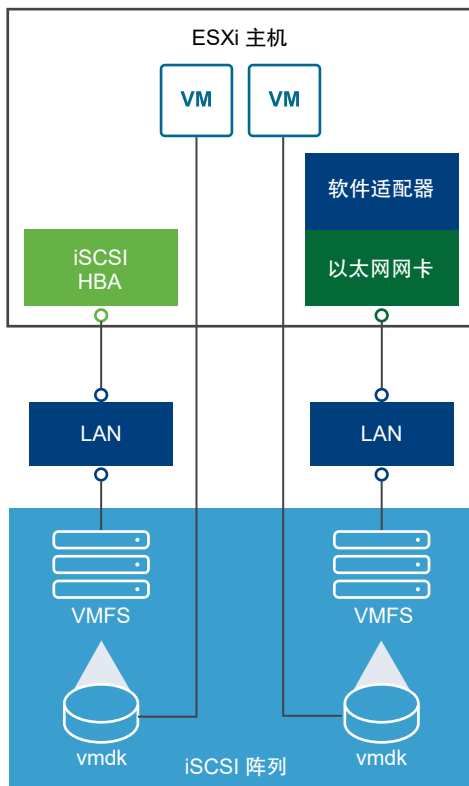
### 软件 iSCSI

主机使用 VMkernel 中基于软件的 iSCSI 启动器连接到存储。通过这种 iSCSI 连接类型，主机只需要一个标准的网络适配器来进行网络连接。

必须配置 iSCSI 启动器以使主机能够访问和显示 iSCSI 存储设备。

iSCSI 存储描述了不同类型的 iSCSI 启动器。

图 2-3. iSCSI 存储



在左侧示例中，主机使用硬件 iSCSI 适配器连接到 iSCSI 存储系统。

在右侧示例中，主机使用软件 iSCSI 适配器和以太网网卡连接到 iSCSI 存储。

此时，存储系统中的 iSCSI 存储设备变得对于主机可用。您可以访问存储设备并创建用于满足存储需求的 VMFS 数据存储。

有关设置 iSCSI SAN 的特定信息，请参见第 4 章 将 ESXi 与 iSCSI SAN 配合使用。

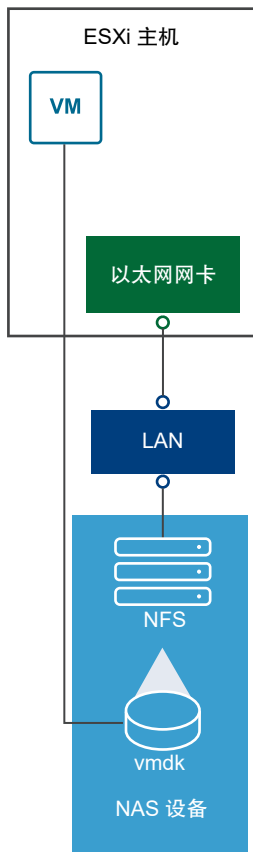
## 网络附加存储 (NAS)

在通过标准 TCP/IP 网络访问的远程文件服务器上存储虚拟机文件。ESXi 中内置的 NFS 客户端使用网络文件系统 (NFS) 协议第 3 版和第 4.1 版来与 NAS/NFS 服务器进行通信。为了进行网络连接，主机需要一个标准的网络适配器。

您可以直接在 ESXi 主机上挂载 NFS 卷。然后，使用 NFS 数据存储来存储和管理虚拟机，这与使用 VMFS 数据存储的方式相同。

NFS 存储描述了使用 NFS 数据存储来存储其文件的虚拟机。在此配置中，主机连接到 NAS 服务器，此服务器通过常规网络适配器存储虚拟磁盘文件。

图 2-4. NFS 存储



有关设置 NFS 存储的特定信息，请参见[了解网络文件系统数据存储](#)。

## 共享串行连接的 SCSI (SAS)

在可向多个主机提供共享访问的直接连接的 SAS 存储系统上存储虚拟机。这种类型的访问允许多个主机访问 LUN 上的同一个 VMFS 数据存储。

## NVMe over Fabrics 存储

VMware NVMe over Fabrics (NVMe-oF) 在主机与共享存储阵列上的目标存储设备之间提供距离连接。VMware 支持 NVMe over RDMA（使用 RoCE v2 技术）、NVMe over Fibre Channel (FC-NVMe) 传输和 NVMe over TCP/IP 技术。有关详细信息，请参见第 8 章 [关于 VMware NVMe 存储](#)。

## 比较存储类型

某些 vSphere 功能是否受支持可能取决于所用存储技术。

下表比较了 ESXi 支持的各种网络存储技术。

表 2-1. ESXi 支持的联网存储

技术	协议	传输	接口
光纤通道	FC/SCSI	数据/LUN 的块访问	FC HBA
以太网光纤通道	FCoE/SCSI	数据/LUN 的块访问	融合网络适配器（硬件 FCoE）
iSCSI	IP/SCSI	数据/LUN 的块访问	<ul style="list-style-type: none"> <li>■ iSCSI HBA 或启用 iSCSI 的网卡（硬件 iSCSI）</li> <li>■ 网络适配器（软件 iSCSI）</li> </ul>
NAS	IP/NFS	文件（无直接 LUN 访问）	网络适配器

下表比较了不同类型存储支持的 vSphere 功能。

表 2-2. 存储支持的 vSphere 功能

存储类型	引导虚拟机	vMotion	数据存储	RDM	虚拟机集群	VMware HA 和 DRS	Storage API - Data Protection
本地存储	是	否	VMFS	否	是	否	是
光纤通道	是	是	VMFS	是	是	是	是
iSCSI	是	是	VMFS	是	是	是	是
NFS 上的 NAS	是	是	NFS 3 和 NFS 4.1	否	否	是	是

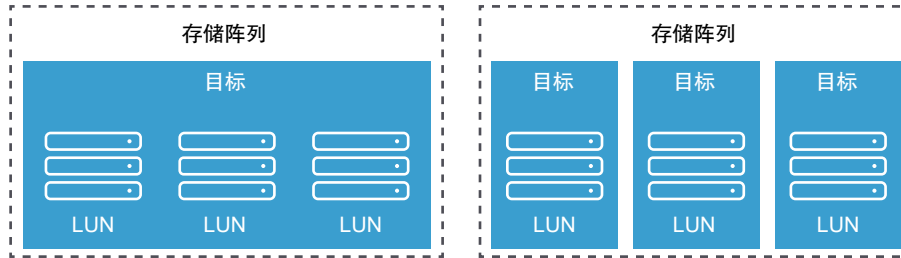
**注** 本地存储支持单个主机上的虚拟机集群（也称为机箱内集群）。需要共享的虚拟磁盘。有关此配置的详细信息，请参见《vSphere 资源管理》文档。

## 目标和设备表示形式

在 ESXi 环境中，“目标”一词标识可以由主机访问的单个存储单元。术语“存储设备”和“LUN”描述代表目标上的存储空间的逻辑卷。在 ESXi 上下文中，这两个术语还表示从存储目标提供给主机并且可以格式化的存储卷。存储设备和 LUN 通常可以互换使用。

不同存储供应商通过不同的方式向 ESXi 主机呈现存储系统。某些供应商在单个目标上呈现多个存储设备或 LUN，而有些供应商则向多个目标各呈现一个 LUN。

图 2-5. 目标和 LUN 表示形式



在此图示中，每种配置都有三个 LUN 可用。在其中一个示例中，主机连接到一个目标，但该目标具有三个可供使用的 LUN。每个 LUN 表示单个存储卷。在另一个示例中，主机检测到三个不同的目标，每个目标都拥有一个 LUN。

通过网络访问的目标都有唯一的名称，该名称由存储系统提供。iSCSI 目标使用 iSCSI 名称。光纤通道目标使用全球名称 (World Wide Name, WWN)。

**注** ESXi 不支持通过不同传输协议（如 iSCSI 和光纤通道）访问相同的 LUN。

设备或 LUN 由其 UUID 名称标识。如果某个 LUN 由多个主机共享，则必须将该 LUN 提供给具有相同 UUID 的所有主机。

## 虚拟机如何访问存储

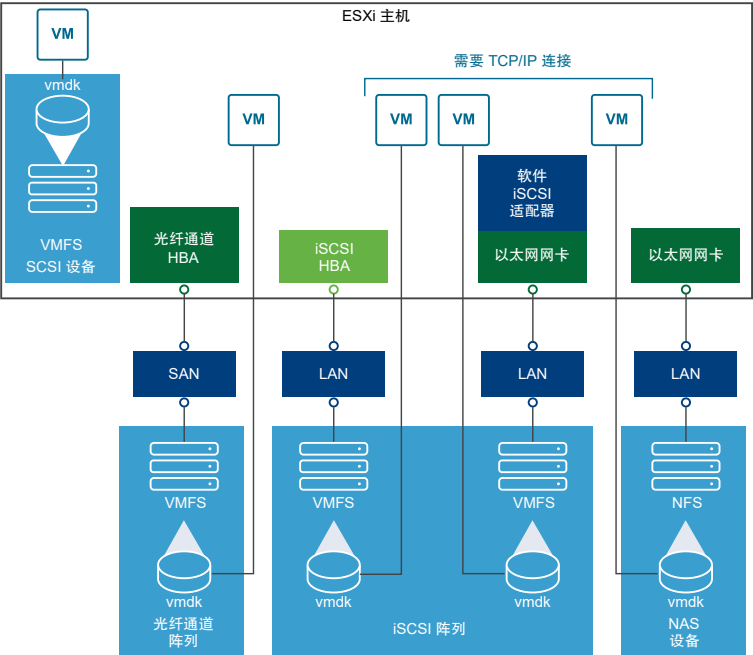
当虚拟机与存储在数据存储上的虚拟磁盘进行通信时，它会发出 SCSI 命令。由于数据存储可以存在于各种类型的物理存储上，因此根据 ESXi 主机用来连接存储设备的协议，这些命令会封装成其他形式。

无论主机使用何种类型的存储设备，虚拟磁盘始终会以挂载的 SCSI 设备形式呈现给虚拟机。虚拟磁盘会向虚拟机操作系统隐藏物理存储层。这样可以在虚拟机内部运行未针对特定存储设备（如 SAN）而认证的操作系统。

下图描述了使用不同存储类型的五个虚拟机，以说明各个类型之间的区别。



图 2-6. 访问不同类型存储的虚拟机



**注** 此图表仅用于展示概念。它并非是推荐的配置。

## 查看可用于 ESXi 主机的存储设备

当 ESXi 主机连接到基于块的存储系统时，支持 ESXi 的 LUN 或存储设备变为对主机可用。

在主机上注册设备后，可以显示所有可用的本地和联网设备并查看其信息。如果使用第三方多路径插件，则通过此类插件可用的存储设备也将出现在列表中。

**注** 如果阵列支持隐式非对称逻辑单元访问 (ALUA)，且只有备用路径，那么该设备的注册将失败。当目标激活备用路径且主机检测到备用路径为活动状态后，可以在主机上注册该设备。高级的系统 `/Disk/FailDiskRegistration` 参数控制主机的此行为。

对于每个存储适配器，可以显示此适配器可用的存储设备的单独列表。

通常，在查看存储设备时，您可看到以下信息。

存储设备信息	描述
名称	也称为显示名称。ESXi 主机根据存储类型和制造商为设备分配的名称。通常，可以根据需要更改此名称。请参见 <a href="#">重命名存储设备</a> 。
标识符	通用唯一标识符是设备的固有名称。请参见 <a href="#">ESXi 存储设备名称和标识符</a> 。
操作状况	指示设备已连接还是已分离。请参见 <a href="#">分离存储设备</a> 。
LUN	SCSI 目标中的逻辑单元号 (LUN)。LUN 号由存储系统提供。如果目标只有一个 LUN，则 LUN 号始终为零 (0)。请参见 <a href="#">目标和设备表示形式</a> 。
类型	设备类型，例如，磁盘或 CD-ROM。

存储设备信息	描述
驱动器类型	有关设备是闪存驱动器还是常规 HDD 驱动器的信息。有关闪存驱动器和 NVMe 设备的信息，请参见第 7 章 <a href="#">对 ESXi 主机使用闪存设备</a> 。
传输	主机用于访问设备的传输协议。协议取决于所使用的存储类型。请参见 <a href="#">ESXi 支持哪些类型的物理存储</a> 。
容量	存储设备的总容量。
所有者	主机用于管理存储设备路径的插件（如 NMP 或第三方插件）。请参见 <a href="#">可插入存储架构和路径管理</a> 。
硬件加速	有关存储设备是否能帮助主机执行虚拟机管理操作的信息。状态可以为“支持”、“不支持”或“未知”。请参见第 16 章 <a href="#">存储硬件加速</a> 。
扇区格式化	指示设备使用的是传统 512n 扇区格式还是高级扇区格式，如 512e 或 4Kn。请参见 <a href="#">设备扇区格式</a> 。
位置	指向 /vmfs/devices/ 目录中存储设备的路径。
分区格式	存储设备采用的分区方案。可以是主引导记录 (MBR) 或 GUID 分区表 (GPT) 格式。GPT 设备可以支持超过 2 TB 的数据存储。请参见 <a href="#">设备扇区格式</a> 。
分区	主分区和逻辑分区，包括 VMFS 数据存储（如果已配置）。请参见第 9 章 <a href="#">使用数据存储</a> 。
多路径策略	主机用于管理存储设备的路径选择策略和存储阵列类型策略。请参见第 10 章 <a href="#">了解多路径和故障切换</a> 。
路径	用于访问存储及其状态的路径。请参见 <a href="#">禁用存储路径</a> 。

## 显示 ESXi 主机的存储设备

显示 ESXi 主机可用的所有存储设备。如果使用任何第三方多路径插件，则通过此类插件可用的存储设备也将出现在列表中。

通过“存储设备”视图，您可列出主机的存储设备、分析其信息并修改属性。

### 步骤

- 1 在 vSphere Client 中，导航到 ESXi 主机。
- 2 单击**配置**选项卡。
- 3 在**存储**下，单击**存储设备**。  
主机可用的所有存储设备都会列在“存储设备”表中。
- 4 要查看特定设备的详细信息，请从列表中选择该设备。
- 5 使用图标执行基本存储管理任务。  
特定图标的可用性取决于设备类型和配置。

- 6 使用以下选项卡访问其他信息，并修改所选设备的属性。

选项卡	描述
属性	查看设备属性和特性。查看和修改设备的多路径策略。请参见 <a href="#">查看和管理路径</a> 。
路径	显示设备的可用路径。停用或启用选定的路径。请参见 <a href="#">禁用存储路径</a> 。
分区详细信息	显示有关分区及其格式的信息。

## 显示适配器的存储设备

显示可通过 ESXi 主机上的特定存储适配器访问的存储设备列表。

### 步骤

- 1 在 vSphere Client 中，导航到 ESXi 主机。

- 2 单击**配置**选项卡。

- 3 在**存储**下，单击**存储适配器**。

主机上安装的所有存储适配器都会列在“存储适配器”表中。

- 4 从列表中选择适配器，然后单击**设备**选项卡。

此时将显示主机可通过该适配器访问的存储设备。

- 5 使用图标执行基本存储管理任务。

特定图标的可用性取决于设备类型和配置。

图标	描述
刷新	刷新有关存储适配器、拓扑和文件系统的信息。
分离存储设备	从主机分离所选设备。
附加存储设备	将所选设备附加到主机。
重命名存储设备	更改所选设备的显示名称。
启用或禁用 ESXi 存储设备上的定位符 LED	打开所选设备的定位器 LED 指示灯。
启用或禁用 ESXi 存储设备上的定位符 LED	关闭所选设备的定位器 LED 指示灯。
标记 ESXi 主机上的存储设备	将所选设备标记为闪存磁盘。
标记 ESXi 主机上的存储设备	将所选设备标记为 HDD 磁盘。
标记 ESXi 主机上的存储设备	将所选设备标记为主机的本地设备。
标记 ESXi 主机上的存储设备	将所选设备标记为主机的远程设备。
擦除存储设备	清除所选设备上的分区。
更改永久预留设置	将所选设备标记为永久保留。
更改永久预留设置	清除所选设备的永久预留。

## 查看 ESXi 主机上可用的存储适配器

存储适配器为 ESXi 主机提供到特定存储单元或网络的连接。ESXi 支持不同的适配器类别，包括 SCSI、iSCSI、RAID、光纤通道、以太网光纤通道 (FCoE) 和以太网。ESXi 通过 VMkernel 中的设备驱动程序直接访问适配器。

根据所使用的存储类型，可能需要在主机上启用和配置存储适配器。

可以显示有关可用存储适配器的详细信息，并查看这些信息。

### 前提条件

必须启用某些适配器（如软件 iSCSI）后，才能查看它们的信息。要配置适配器，请参见 [第 5 章 配置 iSCSI 和 iSER 适配器与存储](#)。

### 步骤

- 1 在 vSphere Client 中，导航到 ESXi 主机。
- 2 单击**配置**选项卡。
- 3 在**存储**下，单击**存储适配器**。
- 4 使用图标执行存储适配器的任务。

特定图标的可用性取决于存储配置。

图标	描述
添加软件适配器	添加存储适配器。适用于软件 iSCSI。
刷新	刷新有关主机上的存储适配器、拓扑和文件系统的信息。
重新扫描存储	重新扫描主机上的所有存储适配器以发现新添加的存储设备或 VMFS 数据存储。
重新扫描适配器	重新扫描选定的适配器以发现新添加的存储设备。

- 5 要查看特定适配器的详细信息，请从列表中选择适配器。
- 6 使用“适配器详细信息”下的选项卡访问其他信息，并修改所选适配器的属性。

选项卡	描述
属性	用于查看一般适配器属性，通常包括适配器的名称和型号，以及根据特定存储标准形成的唯一标识符。对于某些类型的适配器，可以使用此选项卡配置其他属性，例如，身份验证。
设备	用于查看适配器可以访问的存储设备。使用此选项卡可执行常规设备管理任务。请参见 <a href="#">显示适配器的存储设备</a> 。
路径	用于列出和管理适配器用于访问存储设备的所有路径。
目标（光纤通道和 iSCSI）	用于查看和管理通过适配器访问的目标。
网络端口绑定(仅限 iSCSI)	用于配置软件和从属硬件 iSCSI 适配器的端口绑定。
高级选项（仅限 iSCSI）	配置 iSCSI 的高级参数。

## ESXi 数据存储的特性

数据存储是逻辑容器，类似于文件系统，它将各个存储设备的特性隐藏起来，并提供一个统一的模型来存储虚拟机文件。可以显示对 ESXi 主机可用的所有数据存储，并分析其属性。

可通过以下方式将数据存储添加到 vCenter Server 中：

- 您可以使用“新建数据存储”向导创建 VMFS 数据存储、NFS 版本 3 或 4.1 数据存储或 Virtual Volumes 数据存储。启用 vSAN 时，会自动创建 vSAN 数据存储。
- 将 ESXi 主机添加到 vCenter Server 时，主机上的所有数据存储都将添加到 vCenter Server。

下表说明了通过 vSphere Client 查看数据存储时您可以看到的数据存储详细信息。某些特性可能不可用或不适用于所有类型的数据存储。

**表 2-3. 数据存储信息**

数据存储信息	适用的数据存储类型	描述
名称	VMFS NFS vSAN vVol	为数据存储指定的可编辑名称。有关重命名数据存储的信息，请参见 <a href="#">更改数据存储名称</a> 。
类型	VMFS NFS vSAN vVol	数据存储使用的文件系统。有关 VMFS 和 NFS 数据存储以及如何管理它们的信息，请参见第 9 章 <a href="#">使用数据存储</a> 。 有关 vSAN 数据存储的信息，请参见《《管理 VMware vSAN》》文档。 有关 Virtual Volumes 的信息，请参见第 14 章 <a href="#">使用 VMware vSphere Virtual Volumes</a> 。
设备备用	VMFS NFS vSAN	关于底层存储的信息，如部署数据存储所在的存储设备 (VMFS)、服务器和文件夹 (NFS) 或磁盘组 (vSAN)。
协议端点	vVol	有关相应的协议端点的信息。请参见 <a href="#">协议端点</a> 。
数据区	VMFS	数据存储跨越的个别数据区及其容量。
驱动器类型	VMFS	基础存储设备的类型，例如闪存驱动器或常规 HDD 驱动器。有关详细信息，请参见第 7 章 <a href="#">对 ESXi 主机使用闪存设备</a> 。
容量	VMFS NFS vSAN vVol	包括总容量、已置备的空间和可用空间。
挂载点	VMFS NFS vSAN vVol	数据存储在主机的 <code>/vmfs/volumes/</code> 目录中的路径。

表 2-3. 数据存储信息（续）

数据存储信息	适用的数据存储类型	描述
功能集	VMFS	有关基础存储实体提供的存储数据服务的信息。您不能对其进行修改。
	<b>注</b> 多数据区 VMFS 数据存储将仅显示其中一个数据区的功能。	
	NFS	
	vSAN vVol	
Storage I/O Control	VMFS NFS	有关是否启用集群范围存储 I/O 优先化的信息。请参见《vSphere 资源管理》文档。
硬件加速	VMFS	有关基础存储实体是否支持硬件加速的信息。状态可以为“支持”、“不支持”或“未知”。有关详细信息，请参见第 16 章 存储硬件加速。
	NFS	
	vSAN	
	vVol	<b>注</b> NFS 4.1 不支持硬件加速。
标记	VMFS	您定义并将其以标记形式与数据存储关联的数据存储功能。有关信息，请参见向数据存储分配标记。
	NFS	
	vSAN	
	vVol	
与主机的连接	VMFS	挂载了数据存储的主机。
	NFS	
	vVol	
多路径	VMFS	主机用于访问存储的路径选择策略。有关详细信息，请参见第 10 章 了解多路径和故障切换。
	vVol	

## 显示数据存储信息

使用 vSphere Client 导航器访问“数据存储”视图。

使用“数据存储”视图列出 vSphere 基础架构清单中可用的所有数据存储、分析信息和修改属性。

### 步骤

- 1 导航到属于数据存储的有效父对象的任何清单对象，例如主机、集群或数据中心，然后单击**数据存储**选项卡。

清单中的数据存储会显示在中心面板中。

- 2 使用数据存储右键单击菜单中的选项执行所选数据存储的基本任务。

特定选项的可用性取决于数据存储的类型及其配置。

选项	描述
注册虚拟机	注册清单中的一个现有虚拟机。请参见《vSphere 虚拟机管理》文档。
增加数据存储容量	增加 VMFS 数据存储的容量或添加数据区。请参见增加 VMFS 数据存储容量。
浏览文件	导航到数据存储文件浏览器。请参见使用数据存储浏览器。

选项	描述
重命名	更改数据存储名称。请参见 <a href="#">更改数据存储名称</a> 。
挂载数据存储	将数据存储挂载到特定主机。请参见 <a href="#">挂载数据存储</a> 。
卸载数据存储	从特定主机卸载数据存储。请参见 <a href="#">卸载数据存储</a> 。
维护模式	使用数据存储维护模式。请参见《vSphere 资源管理》文档。
配置 Storage I/O Control (VMFS)	为 VMFS 数据存储启用 Storage I/O Control。请参见《vSphere 资源管理》文档。
编辑空间回收 (VMFS)	更改 VMFS 数据存储的空间回收设置。请参见 <a href="#">为现有 VMFS6 数据存储配置固定回收</a> 。
删除数据存储 (VMFS)	移除 VMFS 数据存储。请参见 <a href="#">移除 VMFS 数据存储</a> 。
标记与自定义属性	使用标记对有关数据存储的信息进行编码。请参见 <a href="#">向数据存储分配标记</a> 。

3 要查看特定数据存储详细信息，请单击所选数据存储。

4 使用选项卡访问其他信息，并修改数据存储属性。

选项卡	描述
摘要	查看所选数据存储的统计信息和配置。
监控	查看数据存储的警报、性能数据、资源分配、事件及其他状态信息。
配置	查看和修改数据存储属性。可以查看的菜单项取决于数据存储类型。
权限	分配或编辑选定的数据存储的权限。
文件	导航到数据存储文件浏览器。
主机	查看挂载了数据存储的主机。
虚拟机	查看数据存储上的虚拟机。

## 将持久性内存设备与 ESXi 配合使用

ESXi 支持下一代永久内存设备，也称为非易失性内存 (NVM) 设备。这些设备兼具了内存的性能和速度以及传统存储器的持久性。它们可以在重新引导或发生电源故障后保留已存储的数据。

要求高带宽、低延迟和持久性的虚拟机可以受益于这项技术。示例包括配有加速数据库和分析工作负载的虚拟机。

要与 ESXi 主机搭配使用永久内存，必须熟悉以下概念。

### PMem 数据存储

将永久内存添加到 ESXi 主机之后，主机检测硬件，然后格式化并将其作为本地 PMem 数据存储进行挂载。ESXi 使用 VMFS-L 作为文件系统格式。仅支持每个主机一个本地 PMem 数据存储。

**注** 在管理物理永久内存时，请确保从主机中撤出所有虚拟机，然后将主机置于维护模式。

为缩减管理开销，PMem 数据存储提供了简化的管理模型。传统数据存储任务通常不适用于数据存储，因为主机会自动执行后台所需的所有操作。作为管理员，您无法在 vSphere Client 的“数据存储”视图中显示数据存储或执行其他常规数据存储操作。您可执行的唯一操作是监控 PMem 数据存储的统计信息。

PMem 数据存储用于存储虚拟 NVDIMM 设备和虚拟机的传统虚拟磁盘。带有 vmx 和 vmware.log 文件的虚拟机主目录不能放在 PMem 数据存储上。

## PMem 访问模式

ESXi 将以两种模式向虚拟机公开永久内存。PMem 感知型虚拟机可直接访问永久内存。传统虚拟机可以使用 PMem 数据存储上存储的快速虚拟磁盘。

### 直接访问模式

此模式也称为虚拟 PMem (vPMem) 模式，在此模式下，PMem 区域可以作为虚拟非易失性双列直插式内存模块 (NVDIMM) 提供给虚拟机。虚拟机将 NVDIMM 模块作为可以跨电源周期的标准字节可寻址内存。

置备虚拟机时，可以添加一个或多个 NVDIMM 模块。

虚拟机必须为 ESXi 6.7 或更高版本的硬件版本并具有 PMem 感知型客户机操作系统。NVDIMM 设备可与支持永久内存的最新客户机操作系统（例如 Windows 2016）兼容。

每个 NVDIMM 设备将自动存储在 PMem 数据存储上。

### 虚拟磁盘模式

此模式也称为虚拟 PMem 磁盘 (vPMemDisk) 模式，可用于任何传统的虚拟机，且支持包括所有旧版本在内的任何硬件版本。虚拟机不要求为 PMem 感知型虚拟机。使用此模式时，创建常规 SCSI 虚拟磁盘，并将 PMem 虚拟机存储策略附加到磁盘上。该策略会自动将磁盘置于 PMem 数据存储上。

## PMem 存储策略

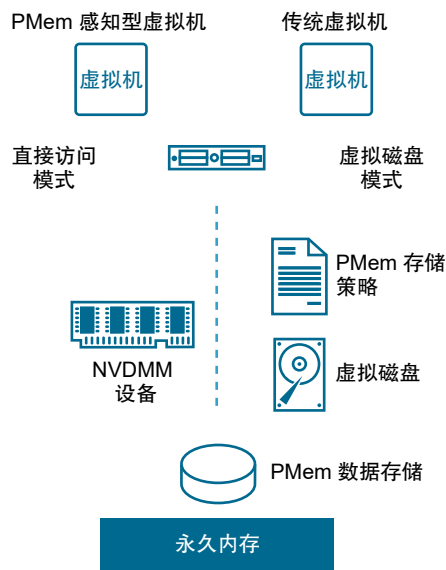
要将虚拟磁盘置于 PMem 数据存储上，必须将主机-本地 PMem 默认存储策略应用到该磁盘。此策略不可编辑。

此策略仅可应用于虚拟磁盘。由于虚拟机主目录未驻留在 PMem 数据存储上，请确保将其置于任何标准数据存储上。

将 PMem 存储策略分配给虚拟磁盘后，无法通过**虚拟机编辑设置**对话框更改策略。要更改策略，请迁移或克隆虚拟机。

下图说明了永久内存组件的交互方式。





有关如何配置和管理具有 NVDIMM 或虚拟永久内存磁盘的虚拟机的信息，请参见《vSphere 资源管理》文档和《vSphere 虚拟机管理》。

## 监控 PMem 数据存储统计信息

您可以使用 vSphere Client 和 `esxcli` 命令来查看 PMem 数据存储的容量和部分其他属性。

但与常规数据存储（例如 VMFS 或 VVol）不同，PMem 数据存储不会出现在 vSphere Client 的“数据存储”视图中。常规数据存储管理任务不会应用于该存储。

### 步骤

- 查看 PMem 数据存储信息。

选项	描述
vSphere Client	a 导航到 ESXi 主机，然后单击摘要。 b 在“硬件”面板中，确认显示了“持久性内存”并查看其容量。
esxcli 命令	使用 <code>esxcli storage filesystem list</code> 列出 PMem 数据存储。

### 示例：查看 PMem 数据存储

当您使用 `esxcli storage filesystem list` 命令列出数据存储时，将显示以下示例输出。

```
# esxcli storage filesystem list
Mount Point          Volume Name          UUID                Mounted  Type      Size
-----
Free
-----
/vmfs/volumes/5xxx...  ds01-102            5xxx...            true     VMFS-6    14227079168
12718178304
/vmfs/volumes/59ex...  ds02-102            59ex...            true     VMFS-6    21206401024
19697500160
```

/vmfs/volumes/59bx...	59bx...	true	vfat	4293591040
4274847744				
/vmfs/volumes/pmem:5ax... PMemDS-56ax...	pmem:5a0x...	true	PMEM	12880707584
11504975872				

## 将 ESXi 与 SAN 配合使用

将 ESXi 与 SAN 配合使用可提高灵活性、效率以及可靠性。将 ESXi 与 SAN 配合使用还支持集中式管理、故障切换和负载均衡技术。

下面是将 ESXi 与 SAN 配合使用的优点：

- 您可安全地存储数据并配置多个指向存储的路径，从而避免出现单一故障点。
- ESXi 主机可访问来自多个存储阵列（包括不同供应商的阵列）的存储设备。
- 将 SAN 与 ESXi 系统配合使用可将故障恢复能力扩展到服务器。使用 SAN 存储时，所有应用程序都可在原始主机出现故障后在其他主机上立即重新启动。
- 可以使用 VMware vMotion 执行虚拟机的实时迁移。
- 如果虚拟机的主机出现故障，可将 VMware High Availability (HA) 与 SAN 结合使用，在另一台服务器上以虚拟机的最后已知状态重新启动虚拟机。
- 可使用 VMware Fault Tolerance (FT) 复制两台不同主机上的受保护虚拟机。如果首选主机出现故障，虚拟机可在辅助主机上继续不间断地运行。
- 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均衡。由于存储位于共享 SAN 阵列，因此应用程序可继续无缝运行。
- 如果使用 VMware DRS 集群，请将 ESXi 主机置于维护模式，以便系统将所有正在运行的虚拟机迁移到其他 ESXi 主机。随后便可在原始主机上执行升级或其他维护操作。

VMware 虚拟机的可移植性和封装与此存储的共享特性相辅相成。当虚拟机位于基于 SAN 的存储上时，您可以快速关闭一台服务器上的虚拟机并在另一台服务器上将其启动，或在一台服务器上将其挂起并在同一网络上的另一台服务器上恢复操作。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。

## ESXi 和 SAN 用例

与 SAN 配合使用时，ESXi 可从 Storage vMotion、Distributed Resource Scheduler (DRS)、High Availability 等多个 vSphere 功能中受益。

将 ESXi 与 SAN 配合使用对于以下任务非常有效：

### 存储整合与存储布局的简化

如果使用多台主机，且每台主机都运行多个虚拟机，那么主机上的存储不再够用。您可能需要使用外部存储。SAN 可提供简单的系统架构和其他优点。

### 零停机时间维护

执行 ESXi 主机或基础架构维护时，使用 vMotion 将虚拟机迁移到其他主机。如果共享存储位于 SAN 上，您无需中断虚拟机的用户服务便可执行维护。虚拟机工作进程在整个迁移过程中继续执行。

## 负载均衡

可以将主机添加到 DRS 集群，该主机的资源将成为集群资源的一部分。将持续监控集群内所有主机和虚拟机的 CPU 和内存资源的分布情况和使用情况。DRS 会将这些衡量指标与理想状态下的资源使用情况进行比较。理想状态下的使用情况会考虑集群内资源池和虚拟机的属性、当前需求以及不均衡目标。如果需要，DRS 会执行或建议执行虚拟机迁移。

## 灾难恢复

您可以使用 VMware High Availability 将多台 ESXi 主机配置为一个集群。集群为虚拟机中运行的应用程序提供快速中断恢复和具有成本效益的高可用性。

## 简化阵列迁移和存储升级

如果购买了新的存储系统，可以使用 Storage vMotion 执行实时迁移，将虚拟机从现有存储迁移到新的目标位置。执行迁移时，可以不中断虚拟机。

## 将 SAN 存储与 ESXi 配合使用的细节

将 SAN 与 ESXi 主机配合使用在很多方面有别于传统的 SAN 使用。

- 您无法使用 SAN 管理工具访问位于该存储上的虚拟机操作系统。使用传统工具，只能监控 VMware ESXi 操作系统。可以使用 vSphere Client 监控虚拟机。
- 通过 SAN 管理工具可看到的 HBA 属于 ESXi 系统（而不是属于虚拟机）。
- 通常，ESXi 系统会为您执行多路径。

## 决定 LUN 的大小和数目

在格式化包含 VMFS 数据存储的 LUN 之前，必须规划如何设置 ESXi 系统的存储。

决定 LUN 的大小和数目，需要考虑以下注意事项：

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

出于以下原因，您可能需要较少、较大的 LUN：

- 在不向存储管理员要求更多空间的情况下，使创建虚拟机的灵活性更大。
- 使调整虚拟磁盘大小、执行快照等等的灵活性更大。
- 使要管理的 VMFS 数据存储变得更少。

出于以下原因，您可能需要较多、较小的 LUN：

- 减少存储空间浪费。
- 不同的应用程序可能需要不同的 RAID 特性。
- 为每个 LUN 设置多路径策略和磁盘份额，从而增加灵活性。

- 使用 Microsoft 集群服务要求每个集群磁盘资源位于各自的 LUN 中。
- 由于对单个卷的争用情况较少，因此性能更佳。

当虚拟机的存储特性不可用时，可能无法轻松确定要置备的 LUN 的数目和大小。可以使用预测性或自适应性方案决定 LUN 大小和数量。

## 使用预测性方案来决定 LUN 的大小和数目

使用预测性方法的试验。

### 步骤

- 1 置备若干具有不同存储特性的 LUN。
- 2 在每个 LUN 上创建一个 VMFS 数据存储，并依照相应特性标记各个数据存储。
- 3 创建虚拟磁盘以包含在 LUN 上创建的 VMFS 数据存储中的虚拟机应用程序数据，这些 LUN 具有满足其上应用程序要求的 RAID 级别。
- 4 使用磁盘份额来区别高优先级虚拟机与低优先级虚拟机。

---

**注** 磁盘份额仅表示给定主机内的比例。分配给某一主机上虚拟机的份额并不影响其他主机上的虚拟机。

---

- 5 运行应用程序以确定虚拟机性能是否可接受。

## 使用自适应性方案来决定 LUN 的大小和数目

可以使用自适应性方案决定 LUN 的大小和数目。

### 步骤

- 1 置备一个较大的 LUN（RAID 1+0 或 RAID 5），同时启用写入缓存。
- 2 在该 LUN 上创建 VMFS。
- 3 在 VMFS 上创建四个或五个虚拟磁盘。
- 4 运行应用程序以确定磁盘性能是否可接受。

### 结果

如果性能可接受，您可以在 VMFS 上再添加虚拟磁盘。如果性能不可接受，请新建大的 LUN（可能采用其他 RAID 级别）并重复此过程。使用迁移，以防止在重新创建 LUN 时丢失虚拟机数据。

## 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。根据您的存储需求，您可以选择具有高性能和高可用性的存储，也可以选择具有低性能的存储。

根据多种因素，存储可分为不同级别：

- 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 (PiT) 还原。支持复制、完全存储处理器冗余和 SAS 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、部分存储处理器冗余和 SCSI 或 SAS 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。

并非所有虚拟机在整个生命周期内都必须位于性能最高和可用性最佳的存储上。

决定某个虚拟机的放置位置时，需要考虑以下注意事项：

- 虚拟机的重要程度
- 性能和可用性要求
- PiT 还原要求
- 备份和复制要求

由于重要程度发生改变或技术发展，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，并且可能由于各种原因（包括组织、操作流程、法规要求及灾难规划等方面的变化）而改变。

## 第三方管理应用程序

可以将第三方管理应用程序与 ESXi 主机配合使用。

大多数 SAN 硬件都附送存储管理软件。在许多情况下，此软件是一个 Web 应用程序，可与连接到网络的任何 Web 浏览器配合使用。在其他情况下，此软件通常在存储系统或单个服务器上运行，独立于使用 SAN 作为存储的服务器。

此第三方管理软件可用于执行下列任务：

- 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。

如果在虚拟机上运行 SAN 管理软件，您可享受虚拟机的一系列优点，包括使用 vMotion 和 VMware HA 进行故障切换。但是，由于附加了间接级别，管理软件可能看不到 SAN。在这种情况下，可以使用 RDM。

---

**注** 虚拟机能否成功运行管理软件取决于特定的存储系统。

---

## SAN 存储备份注意事项

具有正确的备份策略是 SAN 管理的最重要的方面之一。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介质。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存档日志文件。

计划备份取决于多种因素：

- 关键应用程序的标识，这些应用程序在给定时间内需要较频繁地备份。
- 恢复点和恢复时间目标。考虑恢复点必须具有的精确度，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辅助存储设备间所需带宽的大小。
- 对 SAN 环境、存储性能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可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所需的时间。
- 备份单个应用程序所需的时间。
- 存档数据的资源可用性，如脱机介质访问。

设计备份策略时要包括各应用程序的恢复时间目标。也就是考虑执行备份所需的时间和资源。例如，如果已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。

如果应用程序需要在某一时段内恢复，则备份过程必须提供时间表及特定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上的恢复卷。此过程有助于最大限度避免或完全避免通过访问速度较慢的脱机介质来获取丢失的数据组件。

## 使用第三方备份软件包

可以使用第三方备份解决方案来保护虚拟机中的系统、应用程序和用户数据。

VMware 提供的 Storage API - Data Protection 可以与第三方产品一起使用。使用这些 API 时，第三方软件可以执行备份，而无需在 ESXi 主机上加载备份任务的处理。

使用 Storage API - Data Protection 的第三方产品可以执行以下备份任务：

- 对虚拟机执行完整、差异、增量映像备份以及还原。
- 对使用受支持的 Windows 和 Linux 操作系统的虚拟机执行文件级备份。
- 通过对运行受支持的 Microsoft Windows 操作系统的虚拟机使用 Microsoft 卷影复制服务 (Volume Shadow Copy Services, VSS)，确保数据一致性。

因为 Storage API - Data Protection 使用 VMFS 的快照功能，所以备份时不需要停止虚拟机。这些备份是不间断的，可以随时执行，因此不需要更长的备份时段。

有关 Storage API - Data Protection 及与备份产品集成的信息，请参见 [VMware KB 文章 1021175](#)。

# 将 ESXi 与光纤通道 SAN 配合使用

# 3

ESXi 支持光纤通道 (FC)，这是一种存储协议，SAN 使用该协议将数据流量从主机服务器传输到共享存储。本节介绍如何将 ESXi 与光纤通道 SAN 配合使用。有关详细信息，请查看供应商文档。

## 光纤通道 SAN 概念

如果您是 vSphere 管理员，并计划设置主机与 SAN 配合使用，那么您必须掌握 SAN 概念的相关应用知识。在某些出版资料中和 Internet 上可以找到 SAN 的相关信息。这一领域的发展日新月异，请时常查看上述资源。

### 存储区域网络 (SAN)

存储区域网络 (Storage Area Network, SAN) 是将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。

### SAN 架构

在网络上至少存在一台交换机的 SAN 拓扑构成了 SAN 架构。

### 光纤通道 (FC) 协议

为将流量从主机服务器传输到共享存储，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。

### 分区域

为限制服务器访问未分配给该服务器的存储阵列，SAN 将使用分区域。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。

区域分配具有以下作用：

- 减少提供给主机的目标和 LUN 的数目。
- 控制和隔离架构中的路径。
- 可以防止非 ESXi 系统访问特定的存储系统，以及避免 VMFS 数据遭到破坏。
- 可用于分隔不同的环境（例如，将测试环境与生产环境分开）。

对于 ESXi 主机，可使用单启动器区域分配或单启动器单目标区域分配。后者是首选区域分配做法。使用更严格的区域分配可防止出现 SAN 上可能发生的问题和配置错误。

有关详细说明和最佳区域分配做法，请与存储阵列或交换机供应商联系。

## LUN 屏蔽

分区域与 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。

## 多路径

在主机服务器与存储之间传输数据时，SAN 将使用一种叫做多路径的技术。使用多路径，您可以通过多条物理路径从 ESXi 主机到达存储系统上的 LUN。

## 路径故障切换

通常，从主机到 LUN 的单一路径包括 HBA、交换机端口、连接电缆和存储控制器端口。如果路径中的任何组件出现故障，主机将选择另一条可用路径用于 I/O。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。

# 光纤通道 SAN 中的端口

在本文档的上下文中，端口是指设备与 SAN 的连接。SAN 中的每个节点（例如主机、存储设备或架构组件）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）

一种全球唯一的端口标识符，可允许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。

## 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。通过这个唯一 ID 可将数据经由 SAN 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 (NPIV) 时，单一 FC HBA 端口 (N-port) 可使用多个 WWPN 向架构注册。这个方法允许一个 N-port 声明多个架构地址，每个地址都显示为一个唯一的实体。当 ESXi 主机使用 SAN 时，上述多个唯一标识符允许在配置各个虚拟机的过程中为虚拟机分配 WWN。

# 光纤通道存储阵列类型

ESXi 可支持不同的存储系统和阵列。通常分为以下几类。

## 主动-主动存储系统

支持通过所有可用的存储端口同时访问多个 LUN，而不会明显降低性能。所有路径都保持活动状态，除非路径发生故障。

## 主动-被动存储系统



存储处理器在其中主动提供对给定 LUN 的访问权限的系统。其他处理器用作 LUN 的备份，并且可以主动提供对其他 LUN I/O 的访问权限。只能向给定 LUN 的活动端口成功发送 I/O。如果通过主动存储端口访问失败，则其中一个被动存储处理器可通过服务器对其进行访问而激活。

## 非对称存储系统

支持非对称逻辑单元访问 (ALUA)。ALUA 合规存储系统对每个端口提供不同级别的访问。借助 ALUA，主机可以判断目标端口的状态，并区分路径的优先级。主机会将某些活动路径用作主路径，将其他路径用作辅助路径。

## 虚拟机如何访问光纤通道 SAN 上的数据

ESXi 会将虚拟机的磁盘文件存储到位于 SAN 存储设备上的 VMFS 数据存储内。当虚拟机客户机操作系统向其虚拟磁盘发送 SCSI 命令时，SCSI 虚拟化层会将这些命令转换成 VMFS 文件操作。

虚拟机与其在 SAN 上存储的虚拟磁盘交互时，将发生以下过程：

- 1 虚拟机中的客户机操作系统读写 SCSI 磁盘时，将向虚拟磁盘发送 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。
- 3 虚拟 SCSI 控制器将命令转发至 VMkernel。
- 4 VMkernel 将执行以下任务。
  - a 在 VMFS 卷中查找相应的虚拟磁盘文件。
  - b 将对虚拟磁盘上块的请求映射到相应物理设备上的块。
  - c 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。
- 5 物理 HBA 可执行以下任务。
  - a 依据 FC 协议的规则将 I/O 请求打包。
  - b 将请求传输至 SAN。
- 6 根据 HBA 用于连接到架构的端口，其中一个 SAN 交换机会收到请求。交换机将请求路由到相应的存储设备。

本章讨论了以下主题：

- 设置 ESXi 光纤通道环境
- 从光纤通道 SAN 引导 ESXi
- 光纤通道存储最佳做法
- 将光纤通道 NPIV 与 vSphere 虚拟机结合使用

## 设置 ESXi 光纤通道环境

本主题概述了配置光纤通道 SAN 环境以使用 ESXi 时所需遵循的设置步骤。在准备配置光纤通道 SAN 和设置 ESXi 系统时，请查看要求和建议。

### ESXi 光纤通道 SAN 要求

- 请确保 ESXi 系统支持您使用的 SAN 存储硬件和固件组合。如需最新列表，请参见 [VMware 兼容性指南](#)。
- 将系统配置为每个 LUN 只有一个 VMFS 卷。
- 除非使用无磁盘服务器，否则请勿在 SAN LUN 上设置诊断分区。  
如果您使用从 SAN 引导的无磁盘服务器，则共享的诊断分区适用。
- 使用 RDM 访问裸磁盘。有关信息，请参见第 11 章 [裸设备映射](#)。
- 要使多路径正常工作，每个 LUN 必须对所有 ESXi 主机呈现相同的 LUN ID 号。
- 确保存储设备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。
- 在运行 Microsoft Windows 的虚拟机上，将 SCSI TimeoutValue 参数的值增加到 60。由于此值增加，Windows 能够允许路径故障切换导致的延迟 I/O。有关信息，请参见在 [Windows 客户机操作系统上设置超时](#)。

### ESXi 光纤通道 SAN 限制

将 ESXi 与 SAN 配合使用时，会有一些限制。

- ESXi 不支持 FC 连接磁带设备。
- 您不能使用虚拟机内的多路径软件对单个物理 LUN 执行 I/O 负载平衡。但是，此限制不适用于 Microsoft Windows 虚拟机使用动态磁盘的情况。有关配置动态磁盘的信息，请参见 [设置动态磁盘镜像](#)。

### 设置 LUN 分配

此主题提供了有关 ESXi 与 SAN 配合使用时如何分配 LUN 的常规信息。

设置 LUN 分配时，请注意以下几点：

#### 存储置备

为确保 ESXi 系统在启动时识别 LUN，将 SAN 连接到 ESXi 系统前请将所有 LUN 置备到相应的 HBA。

同时将所有 LUN 置备到所有 ESXi HBA。仅当所有 HBA 都看到相同 LUN 时 HBA 故障切换才可行。

对于在多个主机间共享的 LUN，确保 LUN ID 在所有主机间一致。

#### vMotion 和 VMware DRS

当您使用 vCenter Server 以及 vMotion 或 DRS 时，请确保用于虚拟机的 LUN 已置备到所有 ESXi 主机。此操作提供了移动虚拟机的最大能力。

### 主动-主动阵列与主动-被动阵列之间的比较

将 vMotion 或 DRS 用于主动-被动 SAN 存储设备时，请确保所有 ESXi 系统通向所有存储处理器的路径均一致。否则在进行 vMotion 迁移时可能会导致路径抖动。

对于《存储/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 ESXi 主机呈现 LUN。

## 设置光纤通道 HBA

通常，在 ESXi 主机上使用的 FC HBA 可在默认配置设置下正常工作。

应遵循存储阵列供应商提供的配置准则。设置 FC HBA 的过程中，请考虑以下几个问题：

- 不要在单个主机中混合使用不同供应商的 FC HBA。支持拥有相同 HBA 的不同型号，但单个 LUN 无法通过两个不同类型的 HBA 访问，只能通过相同类型的 HBA 访问。
- 请确保各 HBA 上的固件级别相同。
- 设置用于检测故障切换的超时值。要确保获得最佳性能，请不要更改默认值。
- ESXi 支持 32 Gbps 端到端光纤通道连接。

## 配置以太网光纤通道

要访问光纤通道存储，ESXi 主机可以使用以太网光纤通道 (FCoE) 协议。FCoE 协议将光纤通道帧封装到以太网帧中。因此，您的主机不需要特殊的光纤通道链路即可连接到光纤通道存储。主机可以使用 10 Gbit 无损以太网传输光纤通道流量。

### 以太网光纤通道适配器

VMware 支持硬件 FCoE 适配器，这是一种卸载的专用聚合网络适配器 (CNA)，该适配器在同一个卡上包含网络和光纤通道功能。

安装了此类适配器后，主机可检测并使用两个 CNA 组件。在 vSphere Client 中，网络连接组件显示为标准网络适配器 (vmnic)，光纤通道组件显示为 FCoE 适配器 (vmhba)。

有关 VMware FCoE 可以使用的适配器的信息，请参见《VMware 兼容性指南》

## 设置工作流

本主题概述了配置光纤通道 SAN 环境以使用 ESXi 时所需遵循的安装和设置步骤。

请遵循这些步骤来配置 ESXi SAN 环境。

- 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 ESXi。
- 2 检查所有 SAN 组件是否满足要求。
- 3 进行任何必要的存储阵列修改。

多数供应商会提供各自的特定文档，介绍如何设置 SAN 与 VMware ESXi 配合使用。

- 4 为已连接至 SAN 的主机设置 HBA。
- 5 在主机上安装 ESXi。
- 6 创建虚拟机并安装客户机操作系统。
- 7 （可选）为实现 VMware HA 故障切换或使用 Microsoft 集群服务而对系统进行设置。
- 8 根据需要升级或修改环境。

## 从光纤通道 SAN 引导 ESXi

将主机设置为从 SAN 引导时，主机的引导映像存储在 SAN 存储系统中的一个或多个 LUN 上。主机在启动时，将从 SAN 上的 LUN 引导，而不是从其本地磁盘引导。

ESXi 支持通过光纤通道主机总线适配器 (HBA) 或以太网光纤通道 (FCoE) 融合网络适配器 (CNA) 引导。

### 从 SAN 引导的优点

从 SAN 引导可以给 ESXi 环境带来很多好处。但是，在某些情况下，从 SAN 引导与您的主机不兼容。设置系统从 SAN 引导之前，请先判断这是否适合您的环境。

**小心** 对多台 ESXi 主机使用从 SAN 引导时，每台主机都必须拥有自己的引导 LUN。如果将多台主机配置为共享引导 LUN，则 ESXi 映像可能会损坏。

如果使用从 SAN 引导，会和环境带来以下好处：

- 服务器成本更低。无需内部存储，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。没有本地磁盘的服务器通常占用较少空间。
- 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。此外，还可以使用高级阵列功能，如引导映像上的快照。
- 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。
- 可靠性更高。可以通过多条路径访问引导磁盘，从而防止磁盘成为单一故障点。

### 从光纤通道 SAN 引导时的要求和注意事项

ESXi 引导配置必须满足特定要求。

表 3-1. 从 SAN 引导的要求

要求	描述
ESXi 系统要求	遵循供应商针对从 SAN 引导服务器提出的建议。
适配器要求	配置适配器，以便它可以访问引导 LUN。请参见供应商文档。

表 3-1. 从 SAN 引导的要求（续）

要求	描述
访问控制	<ul style="list-style-type: none"> <li>■ 每台主机只能访问其自己的引导 LUN，不能访问其他主机的引导 LUN。使用存储系统软件确保主机只访问指定的 LUN。</li> <li>■ 多台服务器可以共享一个诊断分区。您可以使用针对特定阵列的 LUN 屏蔽来实现此配置。</li> </ul>
多路径支持	由于 BIOS 不支持多路径，并且无法激活备用路径，因此不支持多条路径指向主动/被动阵列上的引导 LUN。
SAN 注意事项	如果阵列未通过直接连接拓扑认证，则必须通过交换的拓扑进行 SAN 连接。如果阵列已通过直接连接拓扑认证，则 SAN 可以直接连接到阵列。交换的拓扑和直接连接拓扑均支持从 SAN 引导。
特定于硬件的注意事项	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。

## 准备从 SAN 引导 ESXi

准备 ESXi 主机以从 SAN 引导时，需要执行多项任务。

此工作流程介绍在机架式安装服务器上从 SAN 引导的一般实现过程。有关在刀片服务器或其他类型的主机上启用“从 SAN 引导”选项的信息，请参阅合作伙伴文档。

有关详细信息，请参见 <https://www.vmware.com/resources/compatibility/search.php>。

### 步骤

#### 1 配置 SAN 组件和存储系统

在将 ESXi 主机设置为从 SAN LUN 引导之前，请先配置 SAN 组件和存储系统。

#### 2 将存储适配器配置为从 SAN 引导

将主机设置为从 SAN 引导时，需要在主机 BIOS 中启用引导适配器。然后，配置该引导适配器以启动与目标引导 LUN 的基本类型连接。

#### 3 将系统设置为从安装介质引导

将主机设置为从 SAN 引导时，首先需要从 VMware 安装介质引导主机。要从安装介质引导，在 BIOS 设置中更改系统引导顺序。

## 配置 SAN 组件和存储系统

在将 ESXi 主机设置为从 SAN LUN 引导之前，请先配置 SAN 组件和存储系统。

由于配置 SAN 组件因供应商而异，因此，请参见各组件的产品文档。

### 步骤

#### 1 请参见适用于您的设置的任何电缆连接指南，连接网络电缆。

检查交换机接线（如果有）。

#### 2 配置存储阵列。

a 在 SAN 存储阵列中，使 ESXi 主机对 SAN 可见。此过程通常称为创建对象。

b 在 SAN 存储阵列中，设置主机以将主机适配器的 WWPN 用作端口名称或节点名称。

- c 创建并分配 LUN。
- d 记录交换机和存储阵列的 IP 地址。
- e 记录各个 SP 的 WWPN。

---

**小心** 如果使用脚本式安装过程在从 SAN 引导模式下安装 ESXi，则执行特殊步骤来避免意外的数据丢失。

---

## 将存储适配器配置为从 SAN 引导

将主机设置为从 SAN 引导时，需要在主机 BIOS 中启用引导适配器。然后，配置该引导适配器以启动与目标引导 LUN 的基本类型连接。

### 前提条件

确定存储适配器的 WWPN。

### 步骤

- ◆ 将存储适配器配置为从 SAN 引导。

由于配置引导适配器因供应商而异，因此，请参阅供应商文档。

## 将系统设置为从安装介质引导

将主机设置为从 SAN 引导时，首先需要从 VMware 安装介质引导主机。要从安装介质引导，在 BIOS 设置中更改系统引导顺序。

由于在 BIOS 中更改引导顺序因供应商而异，因此，请参考供应商文档来了解相关说明。以下过程说明如何在 IBM 主机上更改引导顺序。

### 步骤

- 1 打开系统电源，并进入系统 BIOS Configuration/Setup Utility。
- 2 选择**启动选项**，并按 Enter。
- 3 选择**启动顺序选项**，并按 Enter。
- 4 将**第一启动设备**更改为 **CD-ROM**。

### 结果

现在可以安装 ESXi 了。

## 光纤通道存储最佳做法

将 ESXi 与光纤通道 SAN 配合使用时，请遵照建议以避免性能问题。

vSphere Client 提供多种工具用于收集性能信息。这些信息以图表方式显示并时常更新。

也可以使用 `resxtop` 或 `esxtop` 命令行实用程序。通过这些实用程序，可以详细查看 ESXi 使用资源的情况。有关详细信息，请参见《vSphere 资源管理》文档。

请咨询存储代表，以确定存储系统是否支持 **Storage API - Array Integration** 硬件加速功能。如果支持硬件加速，请参阅供应商文档以在存储系统端启用硬件加速支持。有关详细信息，请参见第 16 章 [存储硬件加速](#)。

## 防止出现光纤通道 SAN 问题

将 ESXi 与光纤通道 SAN 一起使用时，请遵循具体指南以避免 SAN 问题。

为防止出现 SAN 配置问题，请特别注意以下提示：

- 每个 LUN 上仅放置一个 VMFS 数据存储。
- 不要更改系统为您设置的路径策略，除非您了解做出此类更改的影响。
- 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、交换机、服务器和 FC HBA 配置、软件和固件版本以及存储线缆布局。
- 对故障情况进行规划：
  - 制作多个拓扑映射副本。考虑每一元素发生故障时可能对 SAN 带来的影响。
  - 验证不同链接、交换机、HBA 和其他元素，确保未遗漏设计中的关键故障点。
- 确保根据插槽和总线速度将光纤通道 HBA 安装到主机中的正确插槽。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（包括主机的性能图表、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控点。
- 对于包含 ESXi 主机正在使用的 VMFS 数据存储的 LUN，更改其 ID 时，请务必小心操作。如果您更改此 ID，则数据存储变为非活动状态，且其虚拟机失败。对数据存储进行重新签名，使其再次处于活动状态。请参见[管理重复 VMFS 数据存储](#)。

更改 LUN 的 ID 后，重新扫描存储以在主机上重置 ID。有关使用重新扫描的信息，请参见 [ESXi 存储的重扫描操作](#)。

## 禁用自动 ESXi 主机注册

某些存储阵列要求在阵列中注册 ESXi 主机。ESXi 通过将主机的名称和 IP 地址发送到阵列来执行自动主机注册。如果更喜欢使用存储管理软件执行手动注册，请禁用 ESXi 自动注册功能。

### 步骤

- 1 在 vSphere Client 中，导航到 ESXi 主机。
- 2 单击**配置**选项卡。
- 3 在**系统**下，单击**高级系统设置**。
- 4 在“高级系统设置”下，选择 **Disk.EnableNaviReg** 参数，然后单击**编辑**图标。
- 5 将该值设置为 0。

### 结果

此操作将禁用默认情况下启用的自动主机注册。

## 优化光纤通道 SAN 存储器性能

要优化典型 SAN 环境，需要考虑多个因素。

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

存储阵列性能是影响整个 SAN 环境性能的主要因素之一。

如果遇到任何存储阵列性能问题，请参阅存储阵列供应商文档了解任何相关信息。

要提高 vSphere 环境中的阵列性能，请遵循以下常规准则：

- 分配 LUN 时，请注意可能会有许多主机访问 LUN，且每个主机可以运行多台虚拟机。由主机使用的一个 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，ESXi LUN 所在的 RAID 组通常不包括其他服务器所使用的 LUN，这些服务器上未运行 ESXi。
- 请确保读/写缓存可用。
- 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 负载平衡。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平衡效果。密切监控可指示何时需要重新平衡 LUN 的分发。

调整静态平衡的存储阵列就是监控特定性能统计信息，如每秒 I/O 操作数、每秒块数和响应时间。分发 LUN 工作负载以将工作负载分散到所有 SP 也十分重要。

---

**注** 当前，ESXi 不支持动态负载平衡。

---

### 光纤通道的服务器性能

为确保最佳服务器性能，必须考虑几个因素。

各服务器应用程序访问其指定存储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，请遵循以下准则：

- 将每个 LUN 置于提供必要性能级别的 RAID 组。监控分配的 RAID 组中其他 LUN 的活动及资源使用情况。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 ESXi 主机上运行的应用程序所需的性能目标。
- 确保每个主机具有足够 HBA，以在高峰期为主机上的应用程序增加吞吐量。将 I/O 分散到多个 HBA 可为各应用程序提供更快的吞吐量及更短的延迟。
- 要为潜在 HBA 故障提供冗余，请确保主机已连接到双冗余架构。



- 为 ESXi 系统分配 LUN 或 RAID 组时，请注意多个操作系统将使用和共享该资源。ESXi 主机所需的 LUN 性能可能远高于使用常规物理机时所需的 LUN 性能。例如，如果计划运行四个 I/O 密集型应用程序，请为 ESXi LUN 分配四倍大小的性能容量。
- 与 vCenter Server 一起使用多个 ESXi 系统时，存储子系统的性能要求将相应提高。
- ESXi 系统上运行的应用程序所需的待处理 I/O 数必须与 HBA 和存储阵列能够处理的 I/O 数匹配。

## 将光纤通道 NPIV 与 vSphere 虚拟机结合使用

虚拟 N-Port ID 虚拟化 (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 (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。可以将虚拟机配置为使用光纤通道 NPIV 引导。

### 基于 NPIV 的 LUN 访问如何运作

NPIV 使得单一 FC HBA 端口可以向架构注册多个唯一的全球名称 (WWN)，其中每个 WWN 都可分配给单个虚拟机。使用 NPIV 时，SAN 管理员可以根据虚拟机监控和路由存储访问。

仅具有 RDM 的虚拟机可以获得 WWN 分配，并将这些分配用于所有 RDM 流量。

当虚拟机具有分配的 WWN 时，该虚拟机的配置文件 (.vmx) 会更新以包含 WWN 对。WWN 对由全球端口名称 (WWPN) 与全球节点名称 (WWNN) 组成。当该虚拟机已打开电源时，VMkernel 将在物理 HBA 上创建一个虚拟端口 (VPORT)，用于访问 LUN。VPORT 是一个虚拟 HBA，但从 FC 架构角度来看是一个物理 HBA。VPORT 使用已分配给虚拟机的 WWN 对，这是它的唯一标识符。

每个 VPORT 是专用于某个虚拟机的。虚拟机关闭电源后，VPORT 便在主机上消失，并且不会再显示于 FC 架构。当虚拟机从一台主机迁移到另一台主机时，VPORT 会在第一台主机上关闭，然后在目标主机上打开。

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。

### 使用 NPIV 的要求

如果计划在虚拟机上启用 NPIV，则应当了解某些要求。

- NPIV 仅可用于具备 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。
- 主机上的 HBA 必须支持 NPIV。

有关信息，请参见《VMware 兼容性指南》和供应商文档。

- 使用相同类型的 HBA。VMware 不支持访问相同 LUN 的同一主机上的异构 HBA。
- 如果某台主机使用多个物理 HBA 作为指向存储的路径，应将所有物理路径划分到虚拟机。这需要支持多路径，即使一次只有一条路径处于活动状态。
- 确保主机上的物理 HBA 能够检测在该主机上运行且支持 NPIV 的虚拟机将要访问的所有 LUN。
- 该架构中的交换机必须是 NPIV 可识别的。

- 配置用于在存储级别访问 NPIV 的 LUN 时，请确保 NPIV LUN 号和 NPIV 目标 ID 与物理 LUN 和目标 ID 相匹配。
- 为 NPIV WWPN 分区，使其连接到集群主机可以访问的所有存储系统，即使虚拟机不使用存储。如果向包含一个或多个支持 NPIV 的虚拟机的集群中添加任何新的存储系统，则需要添加新区域，以便 NPIV WWPN 可以检测到新的存储系统目标端口。

## NPIV 功能和限制

了解将 NPIV 与 ESXi 配合使用的特定功能和限制。

具有 NPIV 的 ESXi 可支持以下各项：

- NPIV 支持 vMotion。使用 vMotion 迁移虚拟机时，它将保留已分配的 WWN。  
如果将已启用 NPIV 的虚拟机迁移到不支持 NPIV 的主机，则 VMkernel 将恢复为使用物理 HBA 来路由 I/O。
- 如果 FC SAN 环境支持主动-主动阵列中磁盘上的并行 I/O，则连接到两个不同的 NPIV 端口的并行 I/O 也受支持。

将 ESXi 与 NPIV 配合使用时，您必须注意以下限制：

- 由于 NPIV 技术是 FC 协议的扩展，因此，它需要具备 FC 交换机，且不会在直接连接的 FC 磁盘上工作。
- 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。
- NPIV 不支持 Storage vMotion。
- 在虚拟机正在运行时，在 FC 交换机上禁用 NPIV 功能然后重新将其启用可能会导致 FC 链接失败和 I/O 停止。

## 配置或修改 WWN 分配

为虚拟机分配 WWN 设置。可以以后再修改 WWN 分配。

可以创建 1-16 个 WWN 对，它们可以映射到主机上的第 1-16 个物理 FC HBA。

通常不需要更改虚拟机上的现有 WWN 分配。在某些情况下，例如，当手动分配的 WWN 在 SAN 上导致冲突时，可能需要更改或删除 WWN。

### 前提条件

- 配置 WWN 之前，请确保 ESXi 主机可以访问在阵列端配置的存储 LUN 访问控制列表 (Access Control List, ACL)。
- 如果要编辑现有 WWN，请关闭虚拟机电源。

### 步骤

- 1 在清单中右键单击虚拟机，然后选择**编辑设置**。
- 2 单击**虚拟机选项**，然后展开**光纤通道 NPIV**。

### 3 选择以下选项之一，创建或编辑 WWN 分配：

选项	描述
临时禁用此虚拟机的 NPIV	禁用但不移除虚拟机的现有 WWN 分配。
保留不变	保留现有 WWN 分配。只读“WWN 分配”部分显示任何现有 WWN 分配的节点和端口值。
生成新的 WWN	生成新的 WWN，将覆盖所有现有 WWN。HBA 的 WWN 不受影响。指定 WWNN 和 WWPN 的数目。要支持通过 NPIV 进行故障切换，至少需要两个 WWPN。通常只为每个虚拟机创建一个 WWNN。
移除 WWN 分配	移除分配给虚拟机的 WWN。虚拟机使用 HBA WWN 访问存储 LUN。

### 4 单击**确定**保存更改。

#### 后续步骤

在架构中注册新创建的 WWN。

# 将 ESXi 与 iSCSI SAN 配合使用

# 4

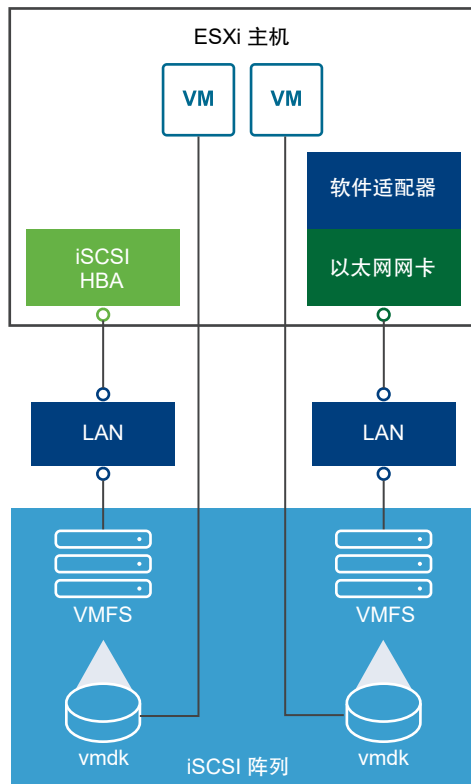
ESXi 可以使用 Internet SCSI (iSCSI) 协议连接到外部 SAN 存储。除了传统的 iSCSI 外，ESXi 还支持 RDMA 的 iSCSI 扩展 (iSER)。

启用 iSER 协议后，主机可以使用相同的 iSCSI 框架，但会将 TCP/IP 传输替换为远程直接内存访问 (RDMA) 传输。

## 关于 iSCSI SAN

iSCSI SAN 在主机和高性能存储子系统之间使用以太网连接。

在主机端，iSCSI SAN 组件包括 iSCSI 主机总线适配器 (host bus adapter, HBA) 或网卡 (NIC)。iSCSI 网络还包括传输存储流量的交换机和路由器、电缆、存储处理器 (Storage Processor, SP) 及存储磁盘系统。



iSCSI SAN 使用客户端-服务器架构。

客户端（也称为 iSCSI 启动器）在 ESXi 主机上操作。它通过发出 SCSI 命令并将封装到 iSCSI 协议中的这些命令传输到 iSCSI 服务器来启动 iSCSI 会话。服务器也称为 iSCSI 目标。通常，iSCSI 目标表示网络上的物理存储系统。

目标也可以是虚拟 iSCSI SAN，例如正在虚拟机中运行的 iSCSI 目标仿真器。iSCSI 目标通过传输必需的 iSCSI 数据响应启动器的命令。

## iSCSI 多路径

在主机服务器与存储之间传输数据时，SAN 将使用一种叫做多路径的技术。通过多路径，ESXi 主机可拥有多个指向存储系统上的 LUN 的物理路径。

通常，从主机到 LUN 的单一路径包括 iSCSI 适配器或网卡、交换机端口、连接电缆和存储控制器端口。如果路径中的任何组件出现故障，主机将选择另一条可用路径用于 I/O。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。

有关多路径的详细信息，请参见 [第 10 章 了解多路径和故障切换](#)。

## iSCSI SAN 节点和端口

iSCSI SAN 上单个可发现的实体（如启动器或目标）表示一个 iSCSI 节点。

每个节点均有一个节点名称。ESXi 使用多种方法来标识节点。

### IP 地址

每个 iSCSI 节点都可具有一个与其相关联的 IP 地址，以便网络上的路由和交换设备可以在主机与存储器之间建立连接。此地址就像为了访问公司的网络或 Internet 而分配给计算机的 IP 地址一样。

### iSCSI 名称

用于标识节点的全球唯一名称。iSCSI 使用 iSCSI 限定名 (IQN) 和扩展唯一标识符 (EUI)。

默认情况下，ESXi 为 iSCSI 启动器生成唯一 iSCSI 名称，例如，  
`iqn.1998-01.com.vmware:iscsitestox-68158ef2`。通常，无需更改默认值；但如果要这样做，请确保输入的新 iSCSI 名称是全球唯一的。

### iSCSI 别名

使用的 iSCSI 设备或端口的更易管理的名称，而不是 iSCSI 名称。iSCSI 别名不是唯一名称，且应为与端口关联的友好名称。

每个节点都有一个或多个将其连接到 SAN 的端口。iSCSI 端口是 iSCSI 会话的端点。

## iSCSI 命名约定

iSCSI 使用一种特殊、唯一的名称来标识 iSCSI 节点（目标或启动器）。

iSCSI 名称通过两种不同方式格式化。最常见的是 IQN 格式。

有关 iSCSI 命名要求和字符串配置文件的更多详细信息，请参见 IETF 网站上的 RFC 3721 和 RFC 3722。

## iSCSI 限定名称格式

iSCSI 限定名称 (IQN) 格式采用 `iqn.yyyy-mm.naming-authority:unique` 名称的形式，其中：

- `yyyy-mm` 是命名机构成立的年份和月份。
- `naming-authority` 是命名机构的 Internet 域名的反向语法。例如，`iscsi.vmware.com` 命名机构可以具有 `iqn.1998-01.com.vmware.iscsi` iSCSI 限定名称形式。该名称表示 `vmware.com` 域名是于 1998 年 1 月注册的，`iscsi` 是由 `vmware.com` 维护的子域。
- `unique name` 是希望使用的任何名称，如主机的名称。命名机构必须确保在冒号后面分配的任何名称都是唯一的，如：
  - `iqn.1998-01.com.vmware.iscsi:name1`
  - `iqn.1998-01.com.vmware.iscsi:name2`
  - `iqn.1998-01.com.vmware.iscsi:name999`

## 企业唯一标识符格式

企业唯一标识符 (EUI) 格式采用 `eui.16_hex_digits` 形式。

例如，`eui.0123456789ABCDEF`。

16 位十六进制数字是 IEEE EUI（扩展唯一标识符）格式的 64 位数的文本表示形式。前 24 位是 IEEE 向特定公司注册的公司 ID。其余 40 位由持有该公司 ID 的实体分配，并且必须是唯一的。

# iSCSI 启动器

要访问 iSCSI 目标，ESXi 主机需要使用 iSCSI 启动器。

启动器是在 ESXi 主机上安装的软件或硬件。iSCSI 启动器启动主机与外部 iSCSI 存储系统之间的通信，并将数据发送到存储系统。

在 ESXi 环境中，在您的主机上配置的 iSCSI 适配器会发挥启动器的作用。ESXi 支持多种类型的 iSCSI 适配器。

有关配置和使用 iSCSI 适配器的信息，请参见 [第 5 章 配置 iSCSI 和 iSER 适配器与存储](#)。

## 软件 iSCSI 适配器

软件 iSCSI 适配器是 VMkernel 中内置的 VMware 代码。使用软件 iSCSI 适配器，您的主机可通过标准网络适配器连接到 iSCSI 存储设备。软件 iSCSI 适配器可在与网络适配器进行通信的过程中负责 iSCSI 处理过程。借助软件 iSCSI 适配器，您无需购买专用硬件便可使用 iSCSI 技术。

## 硬件 iSCSI 适配器

硬件 iSCSI 适配器是第三方适配器，可从主机卸载 iSCSI 和网络处理。硬件 iSCSI 适配器分为几种类别。

- **Dependent Hardware iSCSI Adapter。**基于 VMware 网络以及由 VMware 提供的 iSCSI 配置和管理界面。

这种类型的适配器可以是一张为同一端口提供标准网络适配器和 iSCSI 卸载功能的卡。iSCSI 卸载功能取决于主机获取 IP、MAC 和其他用于 iSCSI 会话的参数的网络配置。iSCSI 许可的 Broadcom 5709 网卡就是一种从属适配器。

- **Independent Hardware iSCSI Adapter。**执行自己的网络及 iSCSI 配置和管理界面。

通常，只提供 iSCSI 卸载功能，或者提供 iSCSI 卸载功能和标准网卡功能的卡就是一种独立硬件 iSCSI 适配器。iSCSI 卸载功能具有独立的配置管理，可分配 IP、MAC 和其他用于 iSCSI 会话的参数。QLogic QLA4052 适配器就是一种独立适配器。

硬件 iSCSI 适配器可能需要获得许可。否则，可能不会出现在客户端或 vSphere CLI 中。有关许可信息，请联系供应商。

## 在 ESXi 中使用 iSER 协议

除了传统的 iSCSI 以外，ESXi 还支持 iSCSI Extensions for RDMA (iSER) 协议。启用 iSER 协议之后，ESXi 主机上的 iSCSI 框架可以使用远程直接内存访问 (RDMA)（而不是 TCP/IP）传输。

传统的 iSCSI 协议通过 TCP/IP 网络在主机上的 iSCSI 启动器和存储设备上的 iSCSI 目标之间传输 SCSI 命令。iSCSI 协议将封装命令，并将相关数据汇编成 TCP/IP 层的数据包。当数据到达时，iSCSI 协议将反汇编 TCP/IP 数据包，以便能够区分 SCSI 命令并将其提供给存储设备。

iSER 与传统的 iSCSI 不同，因为它使用远程直接内存访问 (RDMA) 传输替换 TCP/IP 数据传输模型。通过使用 RDMA 的直接数据放置技术，iSER 协议可以直接在 ESXi 主机和存储设备的内存缓冲区之间传输数据。此方法不仅可以消除不必要的 TCP/IP 处理和数据复制，还可以减少存储设备上的延迟和 CPU 负载。

在 iSER 环境中，iSCSI 的工作方式与以前完全一样，但使用的是底层 RDMA 架构接口（而不是基于 TCP/IP 的接口）。

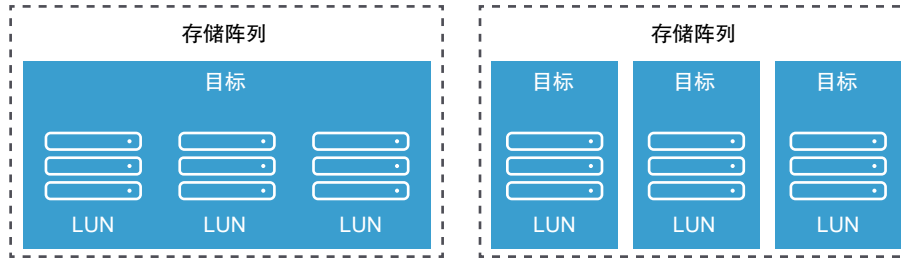
由于 iSER 协议保留了与 iSCSI 基础架构的兼容性，因此在 ESXi 主机上启用 iSER 的过程与 iSCSI 过程类似。请参见[使用 ESXi 配置 iSER](#)。

## 建立 iSCSI 连接

在 ESXi 环境中，“目标”一词标识可以由主机访问的单个存储单元。术语“存储设备”和“LUN”描述代表目标上的存储空间逻辑卷。通常，“设备”和“LUN”等词在 ESXi 环境中表示通过存储器目标向您的主机呈现的 SCSI 卷，对于该卷可以格式化。

不同的 iSCSI 存储供应商通过不同方式向主机呈现存储器。有些供应商在单个目标上呈现多个 LUN，而有些供应商则向多个目标各呈现一个 LUN。

图 4-1. 目标与 LUN 表示形式比较



在这些示例中，每个配置都有三个 LUN 可用。在第一种情况中，主机可以检测到一个目标，但该目标具有三个可供使用的 LUN。每个 LUN 均表示单独的存储卷。在第二种情况中，主机可以检测到三个不同的目标，每个目标都拥有一个 LUN。

基于主机的 iSCSI 启动器建立与各目标的连接。具有包含多个 LUN 的单个目标的存储系统通过一个连接向所有 LUN 输送流量。在具有三个目标（每个均有一个 LUN）的系统中，主机分别连接至三个 LUN。

当您尝试在具有多个 iSCSI 适配器的主机的多个连接上聚合存储流量时，此信息很有用。您可以将一个目标的流量设置为特定适配器，并为另一个目标的流量使用另一个适配器。

## iSCSI 存储系统类型

主机支持的 ESXi 存储类型包括主动-主动、主动-被动和 ALUA 合规。

### 主动-主动存储系统

支持通过所有可用的存储端口同时访问多个 LUN，而不会明显降低性能。所有路径都始终保持活动状态，除非路径发生故障。

### 主动-被动存储系统

存储处理器在其中主动提供对给定 LUN 的访问权限的系统。其他处理器用作 LUN 的备份，并且可以主动提供对其他 LUN I/O 的访问权限。只能向给定 LUN 的活动端口成功发送 I/O。如果通过主动存储端口访问失败，则其中一个被动存储处理器可通过服务器对其进行访问而激活。

### 非对称存储系统

支持非对称逻辑单元访问 (ALUA)。ALUA 合规存储系统对每个端口提供不同级别的访问。借助 ALUA，主机可以判断目标端口的状态，并区分路径的优先级。主机会将某些活动路径用作主路径，其他路径用作辅助路径。

### 虚拟端口存储系统

支持通过一个虚拟端口访问所有可用 LUN。虚拟端口存储系统是主动-主动存储设备，但通过单个端口隐藏其多个连接。默认情况下，ESXi 多路径不会在特定端口与存储之间建立多个连接。某些存储供应商提供会话管理器来建立并管理与其存储的多个连接。这些存储系统以透明的方式处理端口故障切换和连接均衡。此功能通常称为透明故障切换。



## 发现、身份验证和访问控制

可以使用多种机制发现存储并限制对它的访问。

必须配置主机和 iSCSI 存储系统以支持存储访问控制策略。

### 发现

发现会话是 iSCSI 协议的一部分。它返回可在 iSCSI 存储系统上访问的一组目标。ESXi 上提供的两类发现为：动态和静态。动态发现从 iSCSI 存储系统获取可访问目标的列表。静态发现只能通过目标名称和地址访问特定目标。

有关详细信息，请参见在 [ESXi 主机上配置 iSCSI 和 iSER 的动态或静态发现](#)。

### 身份验证

iSCSI 存储系统通过名称和密钥对来对启动器进行身份验证。ESXi 支持 CHAP 身份验证协议。要使用 CHAP 身份验证，ESXi 主机和 iSCSI 存储系统必须启用 CHAP，并具有公共凭据。

有关启用 CHAP 的信息，请参见 [为 iSCSI 或 iSER 存储适配器配置 CHAP 参数](#)。

### 访问控制

访问控制是在 iSCSI 存储系统中设置的策略。大多数实施都支持以下三种访问控制类型的一种或多种：

- 通过启动器名称
- 通过 IP 地址
- 通过 CHAP 协议

仅符合所有规则的启动器才可以访问 iSCSI 卷。

由于 ESXi 主机可以发现所有目标，因此只使用 CHAP 进行访问控制可能会降低重新扫描的速度，但是随后会在身份验证这一步失败。如果主机仅发现其可以进行身份验证的目标，则 iSCSI 重新扫描会进行得更快。

## 虚拟机如何访问 iSCSI SAN 上的数据

ESXi 会将虚拟机的磁盘文件存储到位于 SAN 存储设备上的 VMFS 数据存储内。当虚拟机客户机操作系统向其虚拟磁盘发送 SCSI 命令时，SCSI 虚拟化层会将这些命令转换成 VMFS 文件操作。

虚拟机与其在 SAN 上存储的虚拟磁盘交互时，将发生以下过程：

- 1 虚拟机中的客户机操作系统读写 SCSI 磁盘时，将向虚拟磁盘发送 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。
- 3 虚拟 SCSI 控制器将命令转发至 VMkernel。
- 4 VMkernel 将执行以下任务。
  - a 在 VMFS 卷中查找相应的虚拟磁盘文件。
  - b 将对虚拟磁盘上块请求映射到相应物理设备上的块。

- c 将修改后的 I/O 请求从 VMkernel 中的设备驱动程序发送至 iSCSI 启动器（硬件或软件）。
- 5 如果 iSCSI 启动器是硬件 iSCSI 适配器（独立或从属），则该适配器执行以下任务：
  - a 将 I/O 请求封装到 iSCSI 协议数据单元 (Protocol Data Unit, PDU)。
  - b 将 iSCSI PDU 封装到 TCP/IP 数据包。
  - c 通过以太网将 IP 数据包发送至 iSCSI 存储系统。
- 6 如果 iSCSI 启动器是软件 iSCSI 适配器，则将发生以下事件。
  - a iSCSI 启动器将 I/O 请求封装到 iSCSI PDU 中。
  - b 启动器通过 TCP/IP 连接发送 iSCSI PDU。
  - c VMkernel TCP/IP 堆栈将 TCP/IP 数据包转发到物理网卡。
  - d 物理网卡通过以太网将 IP 数据包发送至 iSCSI 存储系统。
- 7 网络上的以太网交换机和路由器将请求递送至相应的存储设备。

## 错误纠正

为保护 iSCSI 标头和数据完整性，iSCSI 协议定义了错误纠正方法，这些方法称为标头摘要和数据摘要。

这两类参数在 ESXi 主机上默认为禁用，但您可以启用它们。这些摘要分别适用于 iSCSI 启动器和目标之间双向传输的标头和 SCSI 数据。

除其他网络层（例如 TCP 和以太网）提供的完整性检查外，标头和数据摘要还检查非加密数据完整性。标头和数据摘要将检查完整的通信路径，包括可更改网络级流量的所有因素，例如路由器、交换机和代理。

建立 iSCSI 连接之后，将协商摘要的存在和类型。启动器和目标就摘要配置达成一致后，此摘要必须用于它们之间的所有流量。

启用标头和数据摘要要求对启动器和目标进行额外处理，并可影响吞吐量和 CPU 使用性能。

---

**注** 使用 Intel Nehalem 处理器的系统会卸载 iSCSI 摘要计算功能，从而减少对性能的影响。

---

有关启用标头和数据摘要的信息，请参见 [配置 iSCSI 的高级参数](#)。

本章讨论了以下主题：

- [从 iSCSI SAN 引导](#)
- [iSCSI 存储的最佳做法](#)

## 从 iSCSI SAN 引导

将主机设置为从 SAN 引导时，主机的引导映像存储在 SAN 存储系统中的一个或多个 LUN 上。主机在启动时，将从 SAN 上的 LUN 引导，而不是从其本地磁盘引导。

如果您不想处理本地存储的维护，或者具有无磁盘硬件配置（如刀片系统），则可以使用从 SAN 引导。

ESXi 支持从 iSCSI SAN 引导的不同方法。

表 4-1. 从 iSCSI SAN 引导支持

独立硬件 iSCSI	软件 iSCSI
将 iSCSI HBA 配置为从 SAN 引导。有关配置 HBA 的信息，请参见 <a href="#">配置独立硬件 iSCSI 适配器进行 SAN 引导</a> 。	使用软件 iSCSI 适配器和支持 iSCSI 引导固件表 (iBFT) 格式的网络适配器。有关信息，请参见《VMware ESXi 安装和设置》。

## 关于从 iSCSI SAN 引导的常规建议

如果您计划将 iSCSI LUN 设置为主机的引导设备并使用它，请遵循某些常规准则。

下列准则适用于从独立硬件 iSCSI 和 iBFT 引导。

- 查看供应商针对引导配置中使用的硬件提出的任何建议。
- 有关安装必备条件和要求，请查看《vSphere 安装和设置》。
- 使用静态 IP 地址以减少出现 DHCP 冲突的机会。
- 对 VMFS 数据存储和引导分区使用不同的 LUN。
- 在存储系统上配置正确的 ACL。
  - 引导 LUN 必须只对使用该 LUN 的主机可见。不允许 SAN 上的其他任何主机看到该引导 LUN。
  - 如果 LUN 用于 VMFS 数据存储，则多个主机可共享该 LUN。
- 配置诊断分区。
  - 您可以将诊断分区放置在引导 LUN 上（仅适用于独立硬件 iSCSI）。如果配置引导 LUN 中的诊断分区，则此 LUN 不能在多个主机中共享。如果单独的 LUN 用于诊断分区，则多个主机可以共享 LUN。
  - 如果使用 iBFT 从 SAN 进行引导，则不能在 SAN LUN 上设置诊断分区。要收集主机的诊断信息，请在远程服务器上使用 vSphere ESXi Dump Collector。有关 ESXi Dump Collector 的信息，请参见《vCenter Server 安装和设置》和《vSphere 网络连接》。

## 准备 iSCSI SAN

在将主机配置为从 iSCSI LUN 引导之前，需要准备并配置存储区域网络。

**小心** 如果从 SAN 引导时使用脚本式安装来安装 ESXi，则必须执行特殊步骤以避免数据意外丢失。

### 步骤

- 1 请参见适用于您的设置的任何线缆连接指南，连接网络线缆。
- 2 确保存储系统与服务器之间的 IP 连接的正确性。

验证存储网络上任何路由器或交换机的配置。存储系统必须能够 ping 主机中的 iSCSI 适配器。

### 3 配置存储系统。

- a 在存储系统上创建要从中引导主机的卷（或 LUN）。
- b 配置存储系统，以便主机能够访问所分配的 LUN。

此步骤可能涉及到用主机上所使用的 IP 地址、iSCSI 名称和 CHAP 身份验证参数来更新 ACL。在某些存储系统上，除了提供 ESXi 主机的访问信息外，还必须将分配的 LUN 与主机显式关联起来。

- c 确保已将 LUN 正确提供给主机。
- d 确保没有其他系统能够访问配置的 LUN。
- e 记下分配给主机的目标的 iSCSI 名称和 IP 地址。

您必须使用此信息来配置 iSCSI 适配器。

## 配置独立硬件 iSCSI 适配器进行 SAN 引导

如果 ESXi 主机使用独立硬件 iSCSI 适配器（如 QLogic HBA），您可以将适配器配置为从 SAN 引导。

此过程讨论如何启用 QLogic iSCSI HBA 以从 SAN 引导。有关详细信息以及有关 QLogic 适配器配置设置的最新详细信息，请参见 QLogic 网站。

### 步骤

- 1 启动安装媒体并重新引导主机。
- 2 先使用 BIOS 将主机设置为从安装媒体引导。
- 3 在服务器开机自检 (POST) 期间，按 Ctrl+q 进入 QLogic iSCSI HBA 配置菜单。
- 4 选择要配置的 I/O 端口。

默认情况下，适配器引导模式设置为“禁用”。

- 5 配置 HBA。
  - a 在 **Fast!UTIL** 选项菜单中，选择**配置设置 > 主机适配器设置**。
  - b （可选）配置主机适配器的以下设置：启动器 IP 地址、子网掩码、网关、启动器 iSCSI 名称和 CHAP。
- 6 配置 iSCSI 设置。

请参见配置 [iSCSI 引导设置](#)。
- 7 保存更改并重新启动系统。

## 配置 iSCSI 引导设置

配置 iSCSI 引导参数，以便 ESXi 主机可以从 iSCSI LUN 引导。

### 步骤

- 1 在 **Fast!UTIL** 选项菜单中，选择**配置设置 > iSCSI 引导设置**。

- 2 必须先将适配器引导模式设置为**手动**才能设置 SendTargets。
- 3 选择**主要引导设备设置**。
  - a 输入发现**目标 IP**和**目标端口**。
  - b 配置**引导 LUN**和**iSCSI 名称**参数。
    - 如果在目标地址只有一个 iSCSI 目标和一个 LUN 可用，请将**引导 LUN**和**iSCSI 名称**留空。  
您的主机可以访问目标存储系统后，系统会使用适当的信息填充这些文本框。
    - 如果多个 iSCSI 目标和 LUN 可用，请为**引导 LUN**和**iSCSI 名称**提供值。
  - c 保存更改。
- 4 在**iSCSI 引导设置**菜单中，选择主要引导设备。  
此时将自动重新扫描 HBA 以查找新的目标 LUN。
- 5 选择 iSCSI 目标。  
如果目标内存在多个 LUN，则可在找到 iSCSI 设备之后按 **Enter** 来选择特定 LUN ID。
- 6 返回至**主要引导设备设置**菜单。重新扫描后，**引导 LUN**和**iSCSI 名称**将被填充。将**引导 LUN**的值更改为适当的 LUN ID。

## iSCSI 存储的最佳做法

将 ESXi 与 iSCSI SAN 配合使用时，请遵循 VMware 提供的建议以避免出现问题。

请咨询存储代表，以确定存储系统是否支持 Storage API - Array Integration 硬件加速功能。如果支持硬件加速，请参阅供应商文档以在存储系统端启用硬件加速支持。有关详细信息，请参见第 16 章 [存储硬件加速](#)。

## 防止出现 iSCSI SAN 问题

将 ESXi 与 SAN 一起使用时，必须依照特定准则才能避免 SAN 问题。

请遵循以下提示：

- 每个 LUN 上仅放置一个 VMFS 数据存储。
- 不要更改系统为您设置的路径策略，除非您了解做出此类更改的影响。
- 将所有信息记录在案。其中包括涉及以下项目的信息：配置、访问控制、存储、交换机、服务器和 iSCSI HBA 配置、软件和固件版本以及存储线缆布局。
- 对故障情况进行规划：
  - 制作多个拓扑映射副本。考虑每一元素发生故障时可能对 SAN 带来的影响。
  - 除去不同链路、交换机、HBA 和其他元素，确保未遗漏设计中的关键故障点。
- 请确保根据插槽和总线速度将 iSCSI HBA 安装到 ESXi 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。

- 在所有可见点（包括 ESXi 性能图表、以太网交换机统计信息及存储性能统计信息）熟悉存储网络中的各个监控点。
- 仅当 LUN 上部署的 VMFS 数据存储没有正在运行的虚拟机时，才能更改 LUN ID。如果您更改此 ID，则 VMFS 数据存储上运行的虚拟机可能会发生故障。

更改 LUN 的 ID 后，必须重新扫描您的存储以在主机上重置 ID。有关使用重新扫描的信息，请参见 [ESXi 存储的重扫描操作](#)。

- 如果更改 iSCSI 适配器的默认 iSCSI 名称，请确保输入的名称在整个环境中唯一且其格式正确。为了避免存储访问问题，即使在不同的主机上，也绝不将相同的 iSCSI 名称分配给不同的适配器。

## 优化 iSCSI SAN 存储器性能

要优化典型 SAN 环境，需要考虑多个因素。

如果正确配置了网络环境，则 iSCSI 组件将为 iSCSI 启动器和目标提供足够的吞吐量及足够短的延迟。如果网络拥塞，并且链接、交换机或路由器处于饱和状态，则 iSCSI 性能会下降，不能满足 ESXi 环境的需求。

### 存储系统性能

存储系统性能是影响整个 iSCSI 环境性能的主要因素之一。

如果存储系统性能出现问题，请参阅存储系统供应商文档以获取任何相关信息。

分配 LUN 时，请记住，您可以通过多台主机访问每个共享的 LUN，而且每台主机上都可运行多台虚拟机。由 ESXi 主机使用的一个 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，ESXi 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 ESXi 用于 I/O 密集型应用程序。

启用读缓存和写缓存。

负载平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 存储系统进行重新设计和调试，以确保所有存储系统路径间的 I/O 负载平衡。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平衡效果。密切监控可指示何时需要手动重新平衡 LUN 的分发。

调试静态平衡存储系统即是监控特定性能统计信息（例如每秒 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。

### iSCSI 服务器性能

要确保获得最佳 ESXi 主机性能，请考虑以下几个因素。

各服务器应用程序访问其指定存储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，您可以选择存储系统上的适当 RAID 组来实现上述目标。

要实现性能目标，请遵循以下准则：

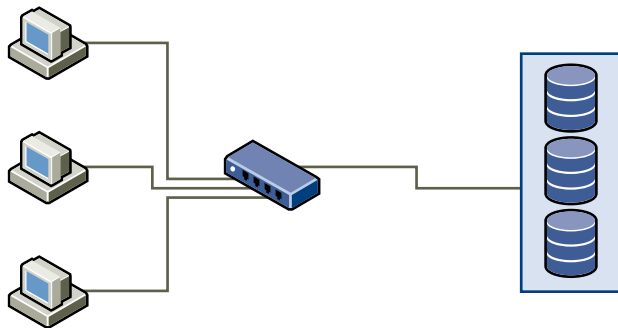
- 将每个 LUN 置于提供必要性能级别的 RAID 组。监控分配的 RAID 组中其他 LUN 的活动及资源使用情况。对于高性能 RAID 组，因有过多应用程序对其执行 I/O 操作，它可能无法满足 ESXi 主机上运行的应用程序所需的性能目标。
- 要在高峰期为主机上的所有应用程序实现最大吞吐量，则安装足够的网络适配器或 iSCSI 硬件适配器。将 I/O 分散到多个端口可为各应用程序提供更快的吞吐量及更短的延迟。
- 要提供软件 iSCSI 的冗余，请确保启动器连接到用于 iSCSI 连接的所有网络适配器。
- 为 ESXi 系统分配 LUN 或 RAID 组时，请注意多个操作系统将使用和共享该资源。ESXi 主机所需的 LUN 性能可能远高于使用常规物理机时所需的 LUN 性能。例如，如果计划运行四个 I/O 密集型应用程序，请为 ESXi LUN 分配四倍大小的性能容量。
- 与 vCenter Server 一起使用多个 ESXi 系统时，存储性能要求会提高。
- ESXi 系统上运行的应用程序所需的未完成 I/O 数必须与 SAN 能够处理的 I/O 数匹配。

## 网络性能

典型的 SAN 包括通过交换机网络连接到一组存储系统的一组计算机。多台计算机通常访问的是同一存储。

下图显示了通过以太网交换机连接到存储系统的多个计算机系统。在此配置中，每个系统都通过单个以太网链路连接到交换机。交换机通过单个以太网链路连接到存储系统。

图 4-2. 单个以太网链路存储的连接

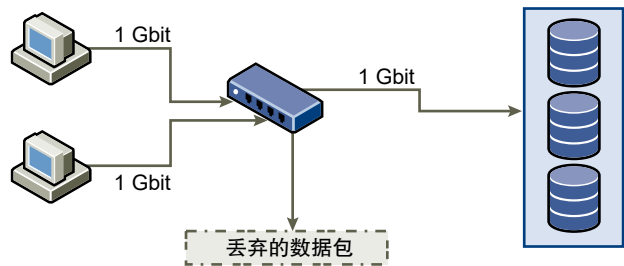


系统从存储读取数据时，存储在响应时会发送足够的填充数据来填充存储系统与以太网交换机之间的链路。任何单个系统或虚拟机都不可能充分利用网络速度。但是，在许多系统共享一个存储设备时可能会发生这种情况。

将数据写入存储时，多个系统或虚拟机可能会尝试填充其链路。因此，系统与存储系统之间的交换机可能会丢弃网络数据包。之所以可能发生丢弃数据的情况，是因为交换机要发送至存储系统的流量超过了单个链路所能承载的流量。交换机可传输的数据量受到交换机与存储系统间的链路速度的限制。



图 4-3. 丢弃的数据包



恢复丢弃的网络数据包会导致性能明显降低。除了花时间确定数据已丢弃，重新传输操作还需使用本可用于当前事务的网络带宽。

iSCSI 流量通过传输控制协议 (Transmission Control Protocol, TCP) 进行网络传输。TCP 是一个可靠的传输协议，可确保重试丢弃的数据包并使其最终到达目标位置。TCP 旨在恢复并快速、无缝地重新传输丢弃的数据包。但是，如果交换机定期丢弃数据包，网络吞吐量将降低。网络会因重新发送数据的请求以及已重新发送的数据包而变得拥堵。传输的数据量少于网络未拥堵时的数据量。

大多数以太网交换机可缓冲或存储数据。此方法为每个尝试发送数据的设备提供到达目标位置的同等机会。缓冲部分传输数据的功能与限制未完成命令数目的许多系统相结合，可以将传输数据缩减为小型突发。多个系统中的突发可以依次发送至存储系统。

如果事务较大，且多台服务器正通过单个交换机端口发送数据，则可能会超出缓冲能力。在这种情况下，交换机会丢弃无法发送的数据，且存储系统必须请求重新传输丢弃的数据包。例如，如果以太网交换机可缓冲 32 KB，但服务器发送到存储设备 256 KB，则会丢弃一些数据。

大多数受管交换机都会提供有关所丢弃数据包的信息，类似以下形式：

```
*: interface is up
IHQ: pkts in input hold queue      IQD: pkts dropped from input queue
OHQ: pkts in output hold queue    OQD: pkts dropped from output queue
RXBS: rx rate (bits/sec)          RXPS: rx rate (pkts/sec)
TXBS: tx rate (bits/sec)          TXPS: tx rate (pkts/sec)
TRTL: throttle count
```

表 4-2. 交换机信息示例

界面	IHQ	IQD	OHQ	OQD	RXBS	RXPS	TXBS	TXPS	TRTL
* GigabitEt hernet0/1	3	9922	0	0	4763030 00	62273	4778400 00	63677	0

在此 Cisco 交换机示例中，使用的带宽为 476303000 位/秒，低于线速的一半。端口缓冲入站数据包，但丢弃了一些数据包。此界面摘要最后一行的 IQD 列表示该端口已丢弃近 10,000 个入站数据包。

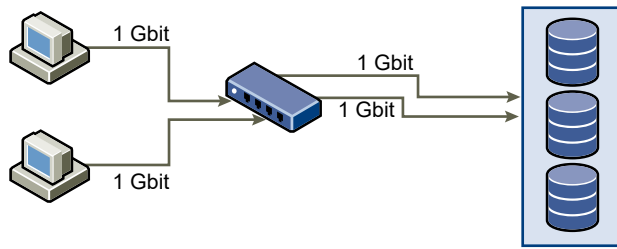
为避免此问题而对配置进行更改时需确保多个输入以太网链路不缩减为一个输出链路，否则会导致链路超额预定。如果在几乎满载时进行传输的多个链路切换为更少数量的链路，则可能发生超额预定情况。

通常，向存储写入大量数据的应用程序或系统必须避免共享存储设备的以太网链路。与存储设备存在多个连接时，这些类型的应用程序可实现最佳性能。



“交换机与存储的多个连接”显示了交换机与存储的多个连接。

图 4-4. 交换机与存储的多个连接



使用 VLAN 或 VPN 不能针对共享配置中的链路超额预定问题提供合适的解决方案。VLAN 和网络的其他虚拟分区提供了按逻辑设计网络的方法。但是，它们不会更改交换机之间链路和中继的物理功能。当存储流量和其他网络流量共享物理连接时，可能会出现超额预定和丢失数据包的现象。此情况也适用于共享交换机间中继的 VLAN。设计 SAN 的性能时，必须考虑网络的物理限制而不是逻辑分配。

## 检查以太网交换机统计信息

许多以太网交换机提供了监控交换机状况的不同方法。

如果端口经常都需要处理接近最大值的吞吐量，则交换机无法提供最佳性能。如果 iSCSI SAN 中的端口需要处理接近最大值的吞吐量，请减少负载。如果端口与 ESXi 系统或 iSCSI 存储器连接，则可使用手动负载平衡来减少负载。

如果端口与多个交换机或路由器连接，请考虑在这些组件间安装更多链接以处理更大的负载。以太网交换机通常还提供有关传输错误、排队的数据包和丢弃的以太网数据包的信息。如果交换机定期报告用于 iSCSI 流量的端口上的上述任何状况，iSCSI SAN 的性能将受影响。

# 配置 iSCSI 和 iSER 适配器与存储

# 5

必须先设置 iSCSI 环境，ESXi 才能与 iSCSI SAN 配合工作。为确保正常工作，您的环境必须遵循特定建议。此外，将 ESXi 与 iSCSI SAN 配合使用时，存在多个限制。

## iSCSI 存储建议

- 验证您的 ESXi 主机是否支持 iSCSI SAN 存储硬件和固件。有关最新的列表，请参见《《VMware 兼容性指南》》。
- 为了确保主机在启动时能够识别 LUN，请配置所有 iSCSI 存储目标，以便主机可以对它们进行访问和使用。请对主机进行配置，使其可以发现所有可用的 iSCSI 目标。
- 除非使用无磁盘服务器，否则请在本地存储上设置诊断分区。如果具有从 iSCSI SAN 引导的无磁盘服务器，请参见[关于从 iSCSI SAN 引导的常规建议](#)获取有关具有 iSCSI 的诊断分区的信息。
- 设置客户机操作系统中的 SCSI 控制器驱动程序以指定一个足够大的队列。
- 在运行 Microsoft Windows 的虚拟机上，增加 SCSI TimeoutValue 参数的值。设置此参数后，Windows 虚拟机能够更好地允许路径故障切换导致的延迟 I/O。有关信息，请参见在[Windows 客户机操作系统上设置超时](#)。
- 将环境配置为每个 LUN 只有一个 VMFS 数据存储。

## iSCSI 存储限制

- ESXi 不支持 iSCSI 连接磁带设备。
- 您不能使用虚拟机多路径软件对单个物理 LUN 执行 I/O 负载均衡。
- 当您将独立硬件适配器与软件适配器或从属硬件适配器组合时，ESXi 不支持多路径。

## 设置 iSCSI ESXi 网络

准备 iSCSI 环境的过程包括以下步骤。

步骤	详细信息
1. 设置 iSCSI 存储	有关信息，请参见存储供应商文档。此外，还请遵循 <a href="#">iSCSI 存储的最佳做法</a> 中的建议。
2. 配置 iSCSI/iSER 适配器	使用相应的工作流配置适配器： <ul style="list-style-type: none"> <li>■ <a href="#">设置独立硬件 iSCSI 适配器</a></li> <li>■ <a href="#">配置从属硬件 iSCSI 适配器</a></li> <li>■ <a href="#">配置软件 iSCSI 适配器</a></li> <li>■ <a href="#">使用 ESXi 配置 iSER</a></li> </ul>
3. 在 iSCSI 存储上创建数据存储	<a href="#">创建数据存储</a>

## iSCSI 网络连接

对于某些类型的 iSCSI 适配器，您必须配置 VMkernel 网络连接。

可以使用 vmkping 实用程序来验证网络配置。

独立硬件 iSCSI 适配器不需要 VMkernel 网络连接。您可以在独立硬件 iSCSI 适配器上配置网络参数，如 IP 地址、子网掩码和默认网关。

所有类型的 iSCSI 适配器都支持 IPv4 和 IPv6 协议。

iSCSI 适配器 (vmhba)	描述	VMkernel 网络连接	适配器网络设置
独立硬件 iSCSI 适配器	从主机卸载 iSCSI 以及网络处理和管理的第三方适配器。	不需要。	有关信息，请参见 <a href="#">编辑硬件 iSCSI 的网络设置</a> 。
软件 iSCSI 适配器	使用标准网卡将主机连接到 IP 网络上的远程 iSCSI 目标。	必需。 有关信息，请参见 <a href="#">为 iSCSI 和 iSER 设置网络</a> 。	不适用
从属硬件 iSCSI 适配器	依赖 VMware 网络以及 iSCSI 配置和管理界面的第三方适配器。	必需 有关信息，请参见 <a href="#">为 iSCSI 和 iSER 设置网络</a> 。	不适用
VMware iSER 适配器	使用具有 RDMA 功能的网络适配器将主机连接到远程 iSCSI 目标。	必需 有关信息，请参见 <a href="#">为 iSCSI 和 iSER 设置网络</a> 。	不适用

## 发现方法

对于所有类型的 iSCSI 适配器，您必须设置动态发现地址或静态发现地址。此外，您必须提供存储系统的目标名称。对于软件 iSCSI 和从属硬件 iSCSI，应该可以使用 vmkping ping 此地址。

请参见在 [ESXi 主机上配置 iSCSI 和 iSER 的动态或静态发现](#)。

## CHAP 身份验证

在启动器和存储系统端启用 CHAP 参数。启用身份验证后，它适用于所有尚未发现的目标。它不适用于已经发现的目标。

请参见为 iSCSI 或 iSER 存储适配器配置 CHAP 参数。

本章讨论了以下主题：

- 设置独立硬件 iSCSI 适配器
- 配置从属硬件 iSCSI 适配器
- 配置软件 iSCSI 适配器
- 使用 ESXi 配置 iSER
- 修改 iSCSI 或 iSER 适配器的常规属性
- 为 iSCSI 和 iSER 设置网络
- 配置 iSCSI 或 iSER 端口绑定
- 将巨帧与 iSCSI 和 iSER 配合使用
- 在 ESXi 主机上配置 iSCSI 和 iSER 的动态或静态发现
- 为 iSCSI 或 iSER 存储适配器配置 CHAP 参数
- 配置 iSCSI 的高级参数
- iSCSI 会话管理

## 设置独立硬件 iSCSI 适配器

独立硬件 iSCSI 适配器是可通过 TCP/IP 访问 iSCSI 存储的专用第三方适配器。此 iSCSI 适配器负责 ESXi 系统的所有 iSCSI 和网络处理及管理。

### 前提条件

- 验证适配器是否必须获得许可。
- 在 ESXi 主机上安装适配器。

有关许可、安装和固件更新的信息，请参见供应商文档。

设置独立硬件 iSCSI 适配器的过程包括以下步骤。

步骤	描述
查看独立硬件 iSCSI 适配器	可以查看独立硬件 iSCSI 适配器并验证它是否已正确安装并准备好进行配置。
修改 iSCSI 或 iSER 适配器的常规属性	如有需要，可以更改分配到 iSCSI 适配器的默认 iSCSI 名称和别名。对于独立硬件 iSCSI 适配器，也可以更改默认 IP 设置。
编辑硬件 iSCSI 的网络设置	更改默认网络设置，以便为 iSCSI SAN 恰当配置适配器。
在 ESXi 主机上配置 iSCSI 和 iSER 的动态或静态发现	设置动态发现。使用动态发现时，启动器每次与指定的 iSCSI 存储系统联系时，都会向该系统发送 SendTargets 请求。iSCSI 系统通过向启动器提供一个可用目标的列表来做出响应。除动态发现方法外，还可以使用静态发现并手动输入目标的信息。
为 iSCSI 或 iSER 存储适配器设置 CHAP	如果 iSCSI 环境使用质询握手身份验证协议 (Challenge Handshake Authentication Protocol, CHAP)，请为配置器配置该协议。
为独立硬件 iSCSI 启用巨帧	如果 iSCSI 环境支持巨型帧，请为适配器启用巨型帧。

## 查看独立硬件 iSCSI 适配器

在 ESXi 主机上，可以查看独立硬件 iSCSI 适配器并验证它是否已正确安装并准备好进行配置。

在主机上安装独立硬件 iSCSI 适配器后，它会显示在可供配置的存储适配器列表上。您可查看其属性。

### 前提条件

所需特权：**主机.配置.存储分区配置**

### 步骤

1 在 vSphere Client 中，导航到 ESXi 主机。

2 单击**配置**选项卡。

3 在**存储**下，单击**存储适配器**。

如果已安装，则硬件 iSCSI 适配器将显示在存储适配器列表上。

4 选择要查看的适配器。

系统将显示适配器的默认详细信息。

适配器信息	描述
型号	适配器的型号。
iSCSI 名称	根据用来标识 iSCSI 适配器的 iSCSI 标准形成的唯一名称。您可以编辑 iSCSI 名称。
iSCSI 别名	用以替代 iSCSI 名称的友好名称。您可以编辑 iSCSI 别名。
IP 地址	分配给 iSCSI HBA 的地址。
目标	通过适配器访问的目标数。
设备	适配器可以访问的所有存储设备或 LUN。
路径	适配器用于访问存储设备的所有路径。

## 编辑硬件 iSCSI 的网络设置

在 ESXi 主机上安装独立硬件 iSCSI 适配器后，您可能需要更改其默认网络设置，以便为 iSCSI SAN 正确配置该适配器。

### 步骤

1 在 vSphere Client 中，导航到 ESXi 主机。

2 单击**配置**选项卡。

3 在**存储**下，单击**存储适配器**，然后选择要配置的适配器 (vmhba#)。

4 单击**网络设置**选项卡，然后单击**编辑**。

- 5 在 IPv4 设置部分中，禁用 IPv6 或选择获取 IP 地址的方法。

**注** 自动 DHCP 选项和静态选项相互排斥。

选项	描述
不使用 IPv4 设置	禁用 IPv4。
自动获取 IPv4 设置	使用 DHCP 获取 IP 设置。
使用静态 IPv4 设置	输入 iSCSI 适配器的 IPv4 IP 地址、子网掩码和默认网关。

- 6 在 IPv6 设置部分中，禁用 IPv6 或选择用于获取 IPv6 地址的相应选项。

**注** 自动选项和静态选项相互排斥。

选项	描述
不使用 IPv6 设置	禁用 IPv6。
启用 IPv6	选择用于获取 IPv6 地址的选项。
通过 DHCP 自动获取 IPv6 地址	使用 DHCP 获取 IPv6 地址。
通过路由器播发自动获取 IPv6 地址	使用路由器播发获取 IPv6 地址。
替代 IPv6 的本地链接地址	通过配置静态 IP 地址，可以替代本地链接 IP 地址。
静态 IPv6 地址	a 单击 <b>添加</b> 以添加新的 IPv6 地址。 b 输入 IPv6 地址和子网前缀长度，然后单击 <b>确定</b> 。

- 7 在 DNS 设置部分中，提供首选 DNS 服务器和备用 DNS 服务器的 IP 地址。

必须同时提供这两个值。

## 配置从属硬件 iSCSI 适配器

从属硬件 iSCSI 适配器是基于 VMware 网络以及由 VMware 提供的 iSCSI 配置和管理界面的第三方适配器。

Broadcom 5709 网卡就是一种从属 iSCSI 适配器。安装在主机上时，将在同一端口上显示两个组件，标准网络适配器和 iSCSI 引擎。iSCSI 引擎作为 iSCSI 适配器 (vmhba) 显示在存储适配器列表中。

默认情况下会启用 iSCSI 适配器。要使其正常工作，必须通过虚拟 VMkernel 适配器 (vmk) 将其连接到与其关联的物理网络适配器 (vmnic)。然后，可以配置此 iSCSI 适配器。

配置从属硬件 iSCSI 适配器后，发现和身份验证数据将通过网络连接传递。iSCSI 流量将绕过网络通过 iSCSI 引擎。

从属硬件 iSCSI 适配器的完整设置和配置过程包括多个步骤。

步骤	描述
查看从属硬件 iSCSI 适配器	查看从属硬件 iSCSI 适配器以验证其加载正确。
修改 iSCSI 或 iSER 适配器的常规属性	如果需要，可以更改分配给适配器的默认 iSCSI 名称和别名。
确定 iSCSI 与网络适配器之间的关联	必须创建网络连接来绑定从属 iSCSI 适配器与物理网络适配器。要正确创建连接，必须确定与从属硬件 iSCSI 适配器关联的物理网卡的名称。
配置 iSCSI 或 iSER 端口绑定	为 iSCSI 组件与物理网络适配器之间的流量配置连接。配置这些连接的过程称为端口绑定。
在 ESXi 主机上配置 iSCSI 和 iSER 的动态或静态发现	设置动态发现。使用动态发现时，启动器每次与指定的 iSCSI 存储系统联系时，都会向该系统发送 SendTargets 请求。iSCSI 系统通过向启动器提供一个可用目标的列表来做出响应。除动态发现方法外，还可以使用静态发现并手动输入目标的信息。
为 iSCSI 或 iSER 存储适配器设置 CHAP	如果 iSCSI 环境使用质询握手身份验证协议 (Challenge Handshake Authentication Protocol, CHAP)，请为配置器配置该协议。
设置目标的 CHAP	还可以为每个发现地址或静态目标配置不同的 CHAP 凭据。
为网络连接启用巨帧	如果 iSCSI 环境支持巨型帧，请为适配器启用巨型帧。

## 从属硬件 iSCSI 注意事项

当将从属硬件 iSCSI 适配器与 ESXi 一起使用时，有一些注意事项。

- 使用任何从属硬件 iSCSI 适配器时，与该适配器关联的网卡的性能报告可能显示很少活动或无活动，即使在 iSCSI 流量很大时也是如此。出现这种行为是因为 iSCSI 流量绕过常规网络堆栈。
- 如果使用第三方虚拟交换机（例如 Cisco Nexus 1000V DVS），请禁用自动固定。而是使用手动固定，确保将 VMkernel 适配器 (vmk) 连接到相应的物理网卡 (vmnic)。有关信息，请参见虚拟交换机供应商文档。
- Broadcom iSCSI 适配器在硬件中执行数据重组，其缓冲区空间有限。在拥堵的网络中或在高负载下使用 Broadcom iSCSI 适配器时，启用流量控制可避免性能下降。

流量控制可管理两个节点之间的数据传输速率，以防止发送方的速度过快而超过接收方。为了获得最佳效果，请在 I/O 路径端点处、主机和 iSCSI 存储系统上启用流量控制。

要为主机启用流量控制，请使用 `esxcli system module parameters` 命令。有关详细信息，请参见 VMware 知识库文章，网址为 <http://kb.vmware.com/kb/1013413>

- 从属硬件适配器支持 IPv4 和 IPv6。

## 查看从属硬件 iSCSI 适配器

在 ESXi 主机上，可以查看从属硬件 iSCSI 适配器以验证其加载是否正确。

如果已经安装，则从属硬件 iSCSI 适配器 (vmhba#) 将显示在以下类别下的存储适配器列表上，如 Broadcom iSCSI 适配器。如果从属硬件适配器未显示在存储适配器列表中，请检查是否需要获取它的许可。请参见供应商文档。

### 步骤

- 1 在 vSphere Client 中，导航到 ESXi 主机。

- 2 单击**配置**选项卡。
- 3 在**存储**下，单击**存储适配器**。
- 4 选择要查看的适配器 (vmhba#)。

将显示适配器的默认详细信息，包括 iSCSI 名称、iSCSI 别名以及状态。

#### 后续步骤

虽然默认情况下从属 iSCSI 适配器已启用，但要使其运行，必须设置适用于 iSCSI 流量的网络，并将此适配器绑定到相应的 VMkernel iSCSI 端口。然后配置发现地址和 CHAP 参数。

## 确定 iSCSI 与网络适配器之间的关联

在 ESXi 主机上，网络连接将从属 iSCSI 与物理网络适配器绑定。必须确定与从属硬件 iSCSI 适配器相关联的物理网卡的名称后，才能正确创建连接。

#### 步骤

- 1 在 vSphere Client 中，导航到 ESXi 主机。
- 2 单击**配置**选项卡。
- 3 在**存储**下，单击**存储适配器**。
- 4 选择 iSCSI 适配器 (vmhba#)，然后在“适配器详细信息”下单击**网络端口绑定**选项卡。
- 5 单击**添加**。

与从属 iSCSI 适配器对应的网络适配器 (vmnic#) 会列在“物理网络适配器”列中。

#### 后续步骤

如果“VMkernel 适配器”列为空，请为物理网络适配器 (vmnic#) 创建 VMkernel 适配器 (vmk#)，然后将其绑定到关联的从属硬件 iSCSI。请参见[为 iSCSI 和 iSER 设置网络](#)。

## 配置软件 iSCSI 适配器

借助基于软件的 iSCSI 实现，可使用标准网卡将主机连接至 IP 网络上的远程 iSCSI 目标。ESXi 中内置的软件 iSCSI 适配器通过利用网络堆栈与物理网卡进行通信，方便了此连接。

如果您使用软件 iSCSI 适配器，请注意以下事项：

- 为 iSCSI 指定一个单独的网络适配器。不要在 100 Mbps 或更慢的适配器上使用 iSCSI。
- 避免在脚本中对软件适配器的名称 vmhbaXX 进行硬编码。名称在不同的 ESXi 版本之间可能会发生更改。如果现有脚本使用硬编码的旧名称，则名称更改可能会导致脚本发生错误。名称更改不影响 iSCSI 软件适配器的行为。

配置软件 iSCSI 适配器的过程包括多个步骤。



步骤	描述
激活或禁用软件 iSCSI 适配器	激活软件 iSCSI 适配器，以便主机可以使用它来访问 iSCSI 存储。
修改 iSCSI 或 iSER 适配器的常规属性	如果需要，可以更改分配给适配器的默认 iSCSI 名称和别名。
配置 iSCSI 或 iSER 端口绑定	为 iSCSI 组件与物理网络适配器之间的流量配置连接。配置这些连接的过程称为端口绑定。
在 ESXi 主机上配置 iSCSI 和 iSER 的动态或静态发现	设置动态发现。使用动态发现时，启动器每次与指定的 iSCSI 存储系统联系时，都会向该系统发送 SendTargets 请求。iSCSI 系统通过向启动器提供一个可用目标的列表来做出响应。除动态发现方法外，还可以使用静态发现并手动输入目标的信息。
为 iSCSI 或 iSER 存储适配器设置 CHAP	如果 iSCSI 环境使用质询握手身份验证协议 (Challenge Handshake Authentication Protocol, CHAP)，请为配置器配置该协议。
设置目标的 CHAP	还可以为每个发现地址或静态目标配置不同的 CHAP 凭据。
为网络连接启用巨帧	如果 iSCSI 环境支持巨型帧，请为适配器启用巨型帧。

## 激活或禁用软件 iSCSI 适配器

必须激活软件 iSCSI 适配器，ESXi 主机才能使用该适配器访问 iSCSI 存储。如果激活后不需要软件 iSCSI 适配器，可将其禁用。

可以仅激活一个软件 iSCSI 适配器。

### 前提条件

所需特权：**主机.配置.存储分区配置**

**注** 如果使用软件 iSCSI 适配器从 iSCSI 进行引导，则在首次引导时会启用该适配器，并创建网络配置。如果禁用了该适配器，它会在您每次引导主机时重新启用。

### 步骤

- 1 在 vSphere Client 中，导航到 ESXi 主机。
- 2 单击**配置**选项卡。

### 3 启用或禁用适配器。

选项	描述
启用软件 iSCSI 适配器	<p>a 在<b>存储</b>下，依次单击<b>存储适配器</b>和<b>添加</b>图标。</p> <p>b 选择<b>软件 iSCSI 适配器</b>，然后确认您要添加该适配器。</p> <p>软件 iSCSI 适配器 (vmhba#) 将启用并显示在存储适配器列表上。启用适配器之后，主机将为其分配默认的 iSCSI 名称。现在，您可以完成适配器配置。</p>
禁用软件 iSCSI 适配器	<p>a 在<b>存储</b>下，单击<b>存储适配器</b>，然后选择要禁用的适配器 (vmhba#)。</p> <p>b 单击<b>属性</b>选项卡。</p> <p>c 单击<b>禁用</b>，并确认您要禁用适配器。</p> <p>状态指示适配器是否已禁用。</p> <p>d 重新引导主机。</p> <p>在重新引导后，适配器不再显示在存储适配器列表中。与该适配器关联的存储设备将变为不可访问。您随后可激活该适配器。</p>

## 使用 ESXi 配置 iSER

除了传统的 iSCSI 以外，ESXi 还支持 iSCSI Extensions for RDMA (iSER) 协议。启用 iSER 协议之后，ESXi 主机上的 iSCSI 框架可以使用远程直接内存访问 (RDMA)（而不是 TCP/IP）传输。您可以在 ESXi 主机上配置 iSER。

有关 iSER 协议的详细信息，请参见在 [ESXi 中使用 iSER 协议](#)。

VMware iSER 的完整设置和配置过程包括多个步骤。

步骤	描述
安装和查看支持 RDMA 的网络适配器	要使用 ESXi 配置 iSER，必须先安装支持 RDMA 的网络适配器，例如，Mellanox Technologies MT27700 Family ConnectX-4。安装此类的适配器后，vSphere Client 将显示它的两个组件：一个 RDMA 适配器和一个物理网络适配器 vmnic#。
启用 VMware iSER 适配器	要想将支持 RDMA 的适配器用于 iSCSI，请使用 <code>esxcli</code> 启用 VMware iSER 存储组件。该组件在 vSphere Client 中显示为 VMware iSCSI over RDMA (iSER) 适配器类别下的 vmhba# 存储适配器。
修改 iSCSI 或 iSER 适配器的常规属性	如果需要，可以更改分配给 iSER 存储适配器 vmhba# 的默认名称和别名。
配置 iSCSI 或 iSER 端口绑定	<p>您必须创建网络连接以绑定 iSER 存储适配器 vmhba# 和支持 RDMA 的网络适配器 vmnic#。配置这些连接的过程称为端口绑定。</p> <p><b>注</b> iSER 不支持网卡绑定。配置端口绑定时，每个 vSwitch 只能使用一个 RDMA 适配器。</p>
在 ESXi 主机上配置 iSCSI 和 iSER 的动态或静态发现	为您的 iSER 存储适配器 vmhba# 设置动态发现或静态发现。使用动态发现时，启动器每次与指定的 iSER 存储系统联系时，都会向该系统发送 <code>SendTargets</code> 请求。iSER 系统通过向启动器提供一个可用目标的列表来做出响应。通过静态发现，您可以手动输入目标的信息。
为 iSCSI 或 iSER 存储适配器设置 CHAP	如果环境使用质询握手身份验证协议 (Challenge Handshake Authentication Protocol, CHAP)，则为 iSER 存储适配器 vmhba# 配置该协议。
设置目标的 CHAP	还可以为每个发现地址或静态目标配置不同的 CHAP 凭据。
为网络连接启用巨帧	如果环境支持巨型帧，则为 iSER 存储适配器 vmhba# 启用巨型帧。

## 安装和查看支持 RDMA 的网络适配器

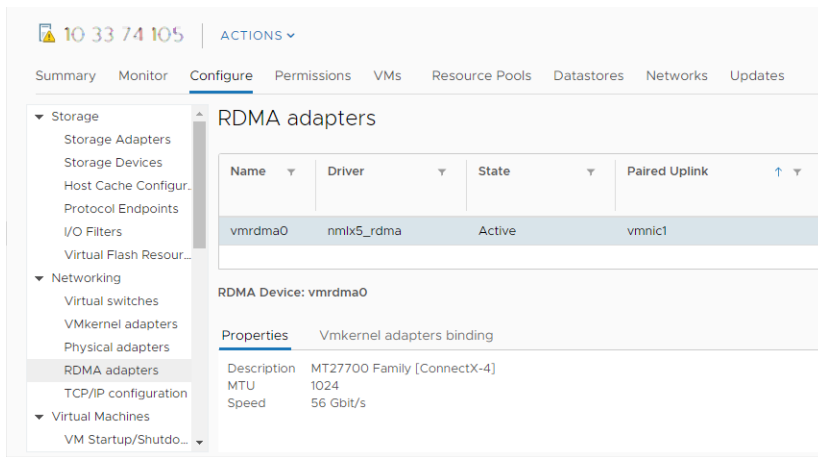
ESXi 支持那些支持 RDMA 功能的网络适配器，例如，Mellanox Technologies MT27700 Family ConnectX-4。在主机上安装此类适配器后，vSphere Client 将显示它的两个组件：一个 RDMA 适配器和一个物理网络适配器。

您可以使用 vSphere Client 查看 RDMA 适配器及其相应的网络适配器。

### 步骤

- 1 在 vSphere Client 中，导航到 ESXi 主机。
- 2 在**网络**下，单击 **RDMA 适配器**。

在此示例中，RDMA 适配器在列表中显示为 vmrdma0。配对上行链路列会将网络组件显示为 vmnic1 物理网络适配器。



- 3 要验证适配器的描述，请从列表中选择 RDMA 适配器，然后单击**属性**选项卡。

### 结果

对于此类存储配置，可以将适配器的 vmnic# 网络组件用作 iSER 或 NVMe over RDMA。有关 iSER 配置步骤，请参见使用 [ESXi 配置 iSER](#)。有关 NVMe over RDMA 的信息，请参见为 [NVMe over RDMA \(RoCE v2\) 存储配置适配器](#)。

## 启用 VMware iSER 适配器

要想将支持 RDMA 的适配器用于 iSCSI，请使用 `esxcli` 启用 VMware iSER 存储组件。启用该组件后，它将在 vSphere Client 中作为 vmhba# 存储适配器显示在“VMware iSCSI over RDMA (iSER) 适配器”类别下。

### 前提条件

- 确保 iSCSI 存储支持 iSER 协议。
- 在 ESXi 主机上安装支持 RDMA 的适配器。有关信息，请参见[安装和查看支持 RDMA 的网络适配器](#)。
- 对于支持聚合以太网 RDMA (RoCE) 且支持 RDMA 的适配器，确定适配器使用的 RoCE 版本。

- 使用支持 RDMA 的交换机。
- 在 ESXi 主机上启用流量控制。要为主机启用流量控制，请使用 `esxcli system module parameters` 命令。有关详细信息，请参见位于 <http://kb.vmware.com/kb/1013413> 的 VMware 知识库文章。
- 确保配置 RDMA 交换机端口以在 iSER 启动器和目标之间创建无损连接。

## 步骤

- 1 使用 ESXi Shell 或 vSphere CLI 启用 VMware iSER 存储适配器并设置其 RoCE 版本。

- a 启用 iSER 存储适配器。

```
esxcli rdma iser add
```

- b 验证是否已添加 iSER 适配器。

```
esxcli iscsi adapter list
```

输出类似以下内容。

```
Adapter Driver State UID Description
-----
vmhba64 iser unbound iscsi.vmhba64 VMware iSCSI over RDMA (iSER) Adapter
```

- c 指定 iSER 用于连接到目标的 RoCE 版本。

使用支持 RDMA 的适配器的 RoCE 版本。输入的命令类似如下：

```
esxcli rdma iser params set -a vmhba64 -r 1
```

命令完成后，VMkernel 日志中将显示类似于以下内容的消息。

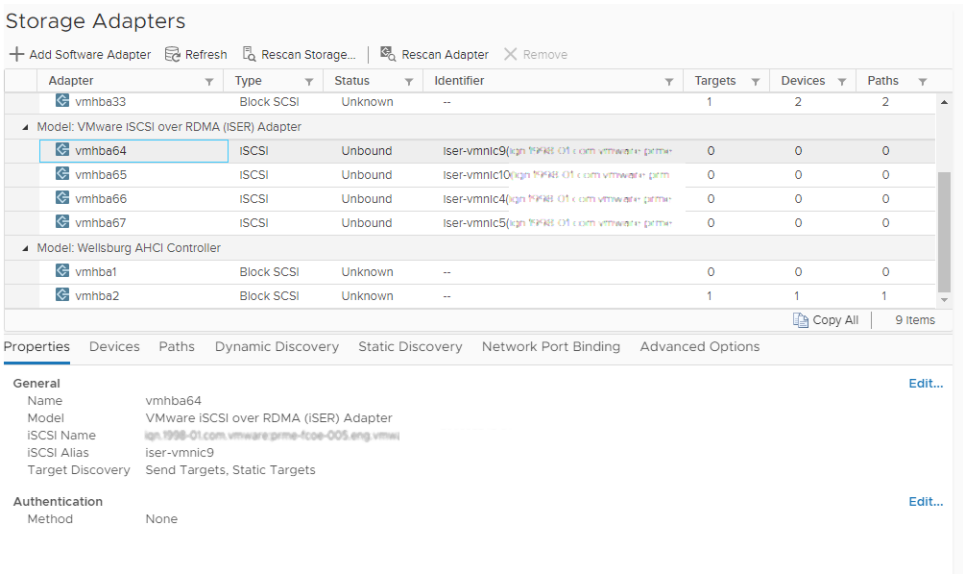
```
vmkernel.0:2020-02-18T18:26:15.949Z cpu6:2100717 opID=45abe37e)iser: iser_set_roce:
Setting roce type: 1 for vmhba: vmhba64
vmkernel.0:2020-02-18T18:26:15.949Z cpu6:2100717 opID=45abe37e)iser: iser_set_roce:
Setting rdma port: 3260 for vmhba: vmhba64
```

如果未指定 RoCE 版本，则主机将默认使用支持 RDMA 的适配器支持的最高 RoCE 版本。

2 使用 vSphere Client 显示 iSER 适配器。

- a 在 vSphere Client 中，导航到 ESXi 主机。
- b 单击配置选项卡。
- c 在存储下，单击存储适配器，然后查看适配器的列表。

如果已启用适配器，它将作为存储 vmhba# 适配器显示在“VMware iSCSI over RDMA (iSER) 适配器”类别下的列表中。



3 选择 iSER 存储 vmhba# 以查看其属性或执行以下任务。

选项	描述
为 iSER 存储适配器配置端口绑定	您必须创建网络连接以绑定 iSER 存储适配器 vmhba# 和支持 RDMA 的网络适配器 vmnic#。配置这些连接的过程称为端口绑定。有关端口绑定的常规信息，请参见为 iSCSI 和 iSER 设置网络。要为 iSER 配置端口绑定，请参见配置 iSCSI 或 iSER 端口绑定。
为 iSER 存储适配器设置动态或静态发现	有关信息，请参见在 ESXi 主机上配置 iSCSI 和 iSER 的动态或静态发现。
为 iSER 存储适配器配置质询握手身份验证协议 (CHAP)	有关信息，请参见为 iSCSI 或 iSER 存储适配器设置 CHAP。

后续步骤

有关详细信息，请参见相应的 VMware 知识库文章，网址为 <https://kb.vmware.com/s/article/79148>。

修改 iSCSI 或 iSER 适配器的常规属性

可以更改 ESXi 主机分配给 iSCSI 或 iSER 存储适配器的默认名称和别名。对于独立硬件 iSCSI 适配器，也可以更改默认 IP 设置。

**重要说明** 在修改适配器的任何默认属性时，请确保使用正确的名称和 IP 地址格式。

前提条件

所需特权: 主机.配置.存储分区配置

步骤

- 1 在 vSphere Client 中，导航到 ESXi 主机。
- 2 单击配置选项卡。
- 3 在存储下，单击存储适配器，然后选择要配置的适配器 (vmhba#)。
- 4 单击属性选项卡，然后在“常规”面板中单击编辑。
- 5 （可选）修改下列常规属性。

选项	描述
iSCSI 名称	根据用来标识 iSCSI 适配器的 iSCSI 标准形成的唯一名称。更改名称时，请确保所输入的名称在整个环境中唯一且其格式正确。否则某些存储设备可能无法识别 iSCSI 适配器。
iSCSI 别名	用以替代 iSCSI 名称的友好名称。

结果

如果更改 iSCSI 名称，该名称将在新的 iSCSI 会话中使用。但对于现有会话，注销并重新登录后才能使用新设置。

后续步骤

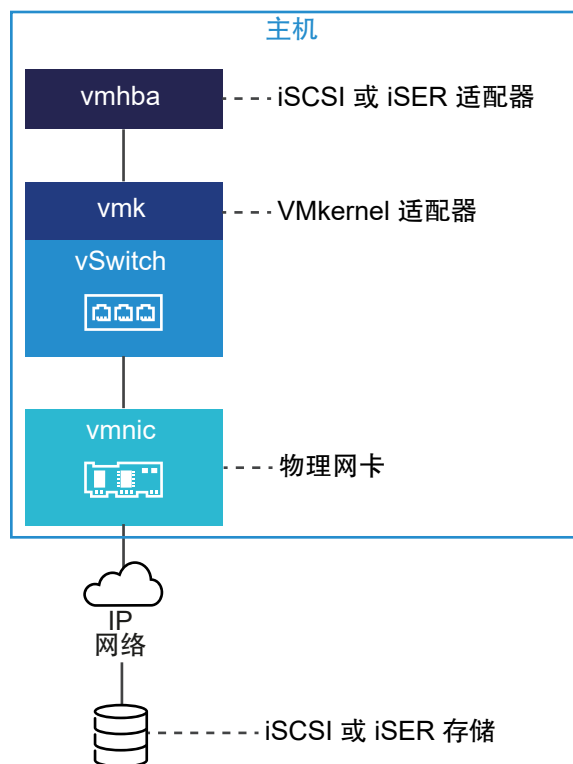
有关可为 iSCSI 或 iSER 存储适配器执行的其他配置步骤，请参见以下主题：

- 设置独立硬件 iSCSI 适配器
- 配置从属硬件 iSCSI 适配器
- 配置软件 iSCSI 适配器
- 使用 ESXi 配置 iSER

## 为 iSCSI 和 iSER 设置网络

某些类型的 iSCSI 适配器取决于 VMkernel 网络连接。这些适配器包括软件或从属硬件 iSCSI 适配器以及 VMware iSCSI over RDMA (iSER) 适配器。如果环境中包括这些适配器中的任何一个，必须为 iSCSI 或 iSER 组件与物理网络适配器之间的流量配置连接。

配置网络连接包括为各个物理网络适配器创建虚拟 VMkernel 适配器。可以在每个虚拟和物理网络适配器之间使用 1:1 映射。然后将 VMkernel 适配器与相应的 iSCSI 或 iSER 适配器关联。此过程称为端口绑定。



配置端口绑定时，请遵循以下规则：

- 可以将软件 iSCSI 适配器与主机上任何可用的物理网卡进行连接。
- 从属 iSCSI 适配器必须只能与其自己的物理网卡进行连接。
- 必须仅将 iSER 适配器连接到支持 RDMA 的网络适配器。

有关配合使用网络连接与软件 iSCSI 的时机和方法的特定注意事项，请参见 VMware 知识库文章，网址为 <http://kb.vmware.com/kb/2038869>。

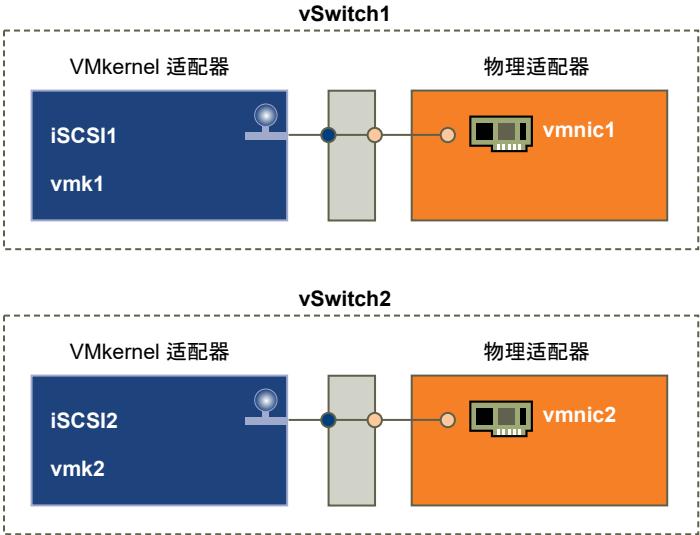
## iSCSI 或 iSER 配置中的多个网络适配器

如果主机有多个物理网络适配器用于 iSCSI 或 iSER，则您可以将这些适配器用于多路径。

可以在单交换机配置或多交换机配置中使用多个物理适配器。

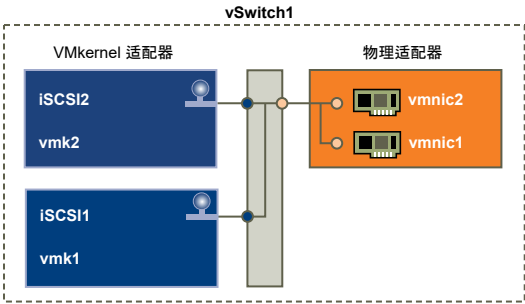
在多交换机配置中，可以为每个虚拟-物理适配器对指定单独的 vSphere 交换机。

图 5-1. 多个单独的 vSphere 标准交换机上的一对一适配器映射



替代方法是所有网卡和 VMkernel 适配器添加到单个 vSphere 交换机。VMkernel 适配器的数量必须与 vSphere 标准交换机上物理适配器的数量相对应。单交换机配置不适用于 iSER，因为 iSER 不支持网卡绑定。

图 5-2. 单个 vSphere 标准交换机上的一对一适配器映射



对于这种配置，必须替代默认网络设置，并确保每个 VMkernel 适配器只映射到一个对应的活动物理适配器，如表中所示。

VMkernel 适配器 (vmk#)	物理网络适配器 (vmnic#)
vmk1 (iSCSI1)	活动适配器
	vmnic1
	未用的适配器
	vmnic2
vmk2 (iSCSI2)	活动适配器
	vmnic2
	未用的适配器
	vmnic1

您也可以使用分布式交换机。有关 vSphere Distributed Switch 以及如何更改默认网络策略的详细信息，请参见《vSphere 网络连接》文档。



使用多个物理适配器时，请注意以下事项：

- 物理网络适配器和它们连接到的存储系统必须位于同一子网上。
- （仅适用于 iSCSI，不适用于 iSER）如果您使用单独的 vSphere 交换机，必须将它们连接到不同的 IP 子网。否则，VMkernel 适配器可能会遇到连接问题，并且主机将无法发现 LUN。
- 单交换机配置不适用于 iSER，因为 iSER 不支持网卡绑定。

如果存在以下任意情况，请不要使用端口绑定：

- 阵列目标 iSCSI 端口处于不同的广播域和 IP 子网中。
- 用于 iSCSI 连接的 VMkernel 适配器位于不同的广播域、IP 子网中，或使用不同的虚拟交换机。

---

**注** 在 iSER 配置中，用于 iSER 连接的 VMkernel 适配器不能用于融合流量。您创建的用于连接具有 iSER 的 ESXi 主机与 iSER 目标的 VMkernel 适配器只能用于 iSER 流量。

---

## 使用软件 iSCSI 配置网络连接的最佳做法

使用软件 iSCSI 配置网络连接时，请考虑以下几种最佳做法。

### 软件 iSCSI 端口绑定

您可以将 ESXi 主机上的软件 iSCSI 启动器绑定到单个或多个 VMkernel 端口，以便 iSCSI 流量仅通过绑定端口。未绑定端口不用于 iSCSI 流量。

配置端口绑定后，iSCSI 启动器将创建从所有绑定端口到所有配置目标门户的 iSCSI 会话。

请参见下面的示例。

VMkernel 端口	目标门户	iSCSI 会话
2 个绑定的 VMkernel 端口	2 个目标门户	4 个会话 (2 x 2)
4 个绑定的 VMkernel 端口	1 个目标门户	4 个会话 (4 x 1)
2 个绑定的 VMkernel 端口	4 个目标门户	8 个会话 (2 x 4)

---

**注** 如果使用端口绑定，请确保所有目标门户可从所有 VMkernel 端口进行访问。否则，可能无法创建 iSCSI 会话。因此，重新扫描操作需要的时间可能比预期更长。

---

### 无端口绑定

如果不使用端口绑定，ESXi 网络连接层会根据其路由表选择最佳 VMkernel 端口。主机通过该端口使用目标门户创建 iSCSI 会话。若无端口绑定，只能为每个目标门户创建一个会话。

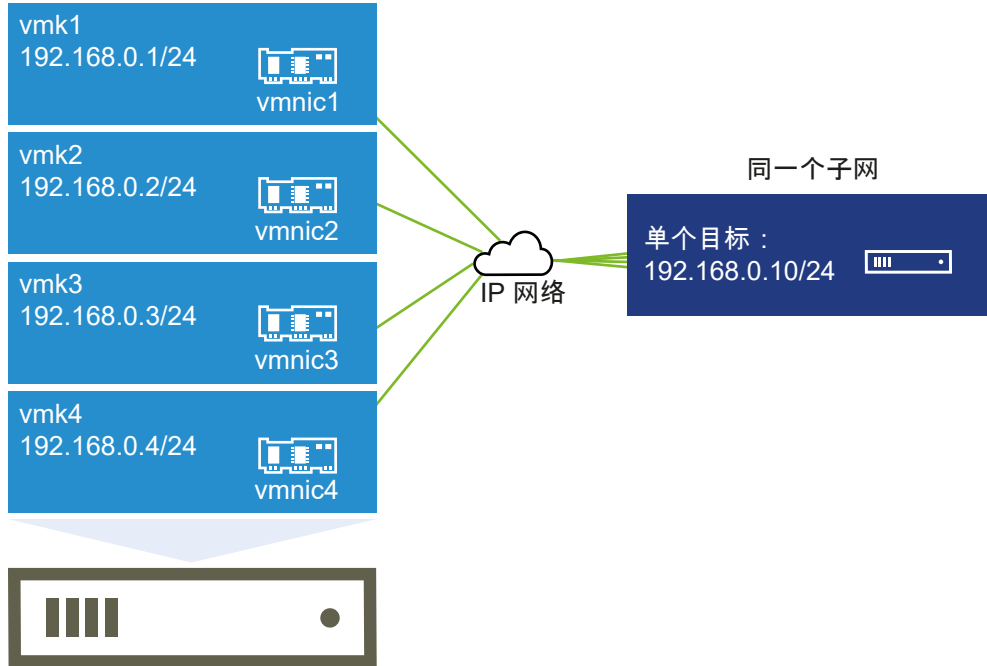
请参见下面的示例。

VMkernel 端口	目标门户	iSCSI 会话
2 个未绑定的 VMkernel 端口	2 个目标门户	2 个会话
4 个未绑定的 VMkernel 端口	1 个目标门户	1 个会话
2 个未绑定的 VMkernel 端口	4 个目标门户	4 个会话

## 软件 iSCSI 多路径

### 示例 1. 含单个网络门户的 iSCSI 目标的多个路径

如果您的目标只有一个网络门户，则可通过在 ESXi 主机上添加多个 VMkernel 端口并将其绑定到 iSCSI 启动器，创建多个指向该目标的路径。

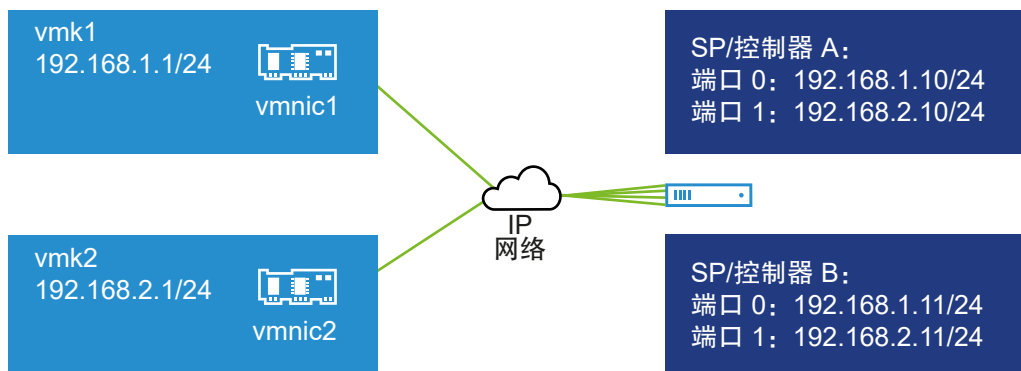


在此示例中，所有启动器端口和目标门户均在同一子网中配置。该目标可通过所有绑定端口访问。您有四个 VMkernel 端口和一个目标门户，因此总共创建了四个路径。

若无端口绑定，只会创建一个路径。

### 示例 2. 含不同子网中的 VMkernel 端口的多个路径

通过在不同的 IP 子网上配置多个端口和目标门户，即可创建多个路径。通过将启动器和目标端口保留在不同的子网中，您可以强制 ESXi 创建经过特定端口的路径。在此配置中，并不使用端口绑定，因为端口绑定要求所有启动器和目标端口位于同一子网上。



ESXi 在连接到控制器 A 和控制器 B 的端口 0 时选择 vmk1，因为所有这三个端口均位于同一子网上。同样，在连接到控制器 A 和控制器 B 的端口 1 时选择 vmk2。可以在此配置中使用网卡绑定。

总共创建了四个路径。

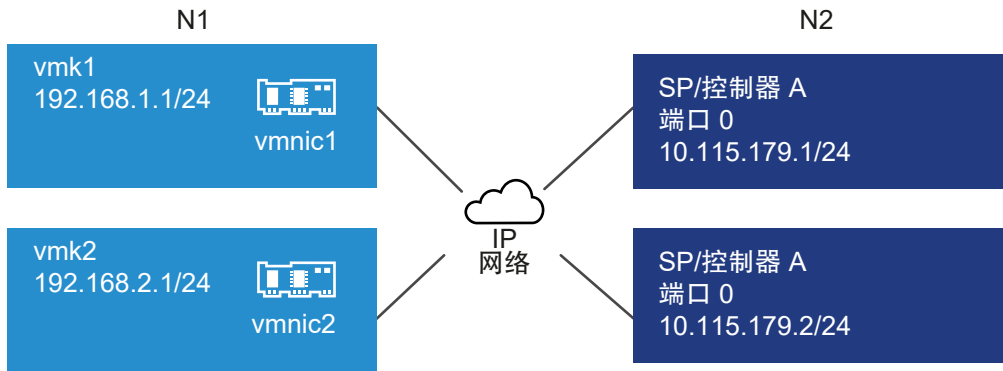
路径	描述
路径 1	vmk1 和控制器 A 的端口 0
路径 2	vmk1 和控制器 B 的端口 0
路径 3	vmk2 和控制器 A 的端口 1
路径 4	vmk2 和控制器 B 的端口 1

使用软件 iSCSI 路由

您可以使用 `esxcli` 命令为 iSCSI 流量添加静态路由。配置静态路由后，不同子网中的启动器和目标端口可以相互通信。

示例 1.使用具有端口绑定的静态路由

在此示例中，您将所有绑定的 VMkernel 端口保留在一个子网 (N1) 中，并在另一个子网 (N2) 中配置所有目标门户。然后，可以为目标子网 (N2) 添加静态路由。

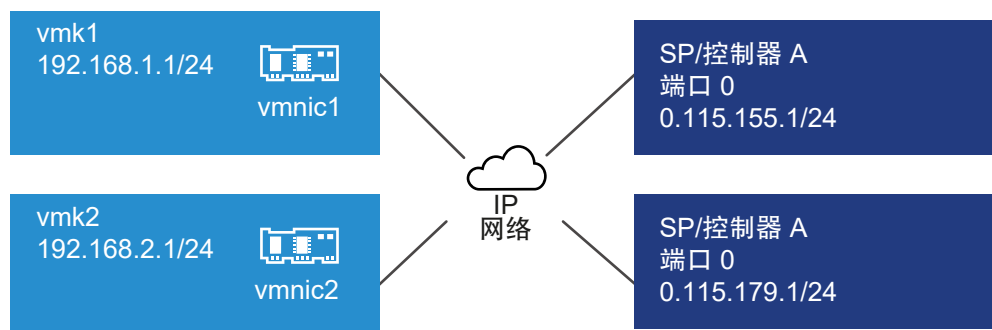


使用以下命令：

```
# esxcli network ip route ipv4 add -gateway 192.168.1.253 -network 10.115.179.0/24
```

示例 2.使用静态路由创建多个路径

在此示例中，使用不同的子网时将使用静态路由。不能在此配置中使用端口绑定。



您在独立子网 192.168.1.0 和 192.168.2.0 中配置 vmk1 和 vmk2。此外，目标门户也位于独立子网 10.115.155.0 和 10.115.179.0 中。

可以从 vmk1 为 10.115.155.0 添加静态路由。请确保网关可从 vmk1 进行访问。

```
# esxcli network ip route ipv4 add -gateway 192.168.1.253 -network 10.115.155.0/24
```

然后从 vmk2 为 10.115.179.0 添加静态路由。请确保网关可从 vmk2 进行访问。

```
# esxcli network ip route ipv4 add -gateway 192.168.2.253 -network 10.115.179.0/24
```

使用控制器 A 的端口 0 进行连接时，将使用 vmk1。

使用控制器 B 的端口 0 进行连接时，将使用 vmk2。

示例 3.每个 VMkernel 端口使用一个单独的网关进行路由

从 vSphere 6.5 开始，您可以为每个 VMkernel 端口配置一个单独网关。如果使用 DHCP 获取 VMkernel 端口的 IP 配置，还可使用 DHCP 获取网关信息。

要查看每个 VMkernel 端口的网关信息，请使用以下命令：

```
# esxcli network ip interface ipv4 address list
```

Name	IPv4 Address	IPv4 Netmask	IPv4 Broadcast	Address Type	Gateway	DHCP	DNS
vmk0	10.115.155.122	255.255.252.0	10.115.155.255	DHCP	10.115.155.253	true	
vmk1	10.115.179.209	255.255.252.0	10.115.179.255	DHCP	10.115.179.253	true	
vmk2	10.115.179.146	255.255.252.0	10.115.179.255	DHCP	10.115.179.253	true	

通过每个 VMkernel 端口的单独网关，您可使用端口绑定访问不同子网中的目标。

## 管理 iSCSI 网络

特殊注意事项适用于与 iSCSI 适配器关联的网络适配器，包括物理网络适配器和 VMkernel 网络适配器。

为 iSCSI 创建网络连接后，iSCSI 指示器在 vSphere Client 中变为启用。此指示器显示特定虚拟网络适配器或物理网络适配器是 iSCSI 绑定的。为避免 iSCSI 流量中断，请在管理 iSCSI 绑定的虚拟网络适配器和物理网络适配器时遵循以下准则和注意事项：

- 确保 VMkernel 网络适配器已在与其所连接的 iSCSI 存储门户相同的子网上分配了地址。

- 使用 VMkernel 适配器的 iSCSI 适配器无法连接到不同子网上的 iSCSI 端口，即使 iSCSI 适配器发现了这些端口。
- 使用单独的 vSphere 交换机连接物理网络适配器和 VMkernel 适配器时，请确保 vSphere 交换机连接到不同的 IP 子网。
- 如果 VMkernel 适配器位于同一子网上，则它们必须连接到一个 vSwitch。
- 如果将 VMkernel 适配器迁移到其他 vSphere 交换机，请移动关联的物理适配器。
- 请勿对 iSCSI 绑定的 VMkernel 适配器或物理网络适配器进行配置更改。
- 请勿进行可能中断 VMkernel 适配器与物理网络适配器的关联的更改。移除其中一个适配器或移除连接适配器的 vSphere 交换机时，可能会中断关联。或者更改其连接的 1:1 网络策略时，也可能会中断关联。

## iSCSI 网络故障排除

警告标记表示 iSCSI 绑定的 VMkernel 适配器的不合规端口组策略。

### 问题

VMkernel 适配器的端口组策略在以下情况下被认为不合规：

- VMkernel 适配器未连接到活动物理网络适配器。
- VMkernel 适配器连接到多个物理网络适配器。
- VMkernel 适配器连接到一个或多个待机物理适配器。
- 活动物理适配器已更改。

### 解决方案

为 iSCSI 绑定的 VMkernel 适配器设置正确的网络策略。请参见 [配置 iSCSI 或 iSER 端口绑定](#)。

## 配置 iSCSI 或 iSER 端口绑定

端口绑定将为特定类型的 iSCSI 和 iSER 适配器与物理网络适配器之间的通信创建连接。

以下类型的适配器需要端口绑定：

- 软件 iSCSI 适配器
- 从属硬件 iSCSI 适配器
- VMware iSCSI over RDMA (iSER) 适配器

以下任务讨论了包含 vSphere 标准交换机和单个物理网络适配器的网络配置。如果您有多个网络适配器，请参见 [iSCSI 或 iSER 配置中的多个网络适配器](#)。

---

**注** iSER 不支持网卡绑定。为 iSER 配置端口绑定时，每个 vSwitch 仅使用一个启用了 RDMA 的物理适配器 (vmnic#) 和一个 VMkernel 适配器 (vmk#)。

---

您还可以在端口绑定配置中使用 VMware vSphere<sup>®</sup> Distributed Switch<sup>™</sup> 和 VMware NSX<sup>®</sup> Virtual Switch<sup>™</sup>。有关 NSX Virtual Switch 的信息，请参见 VMware NSX Data Center for vSphere 文档。

如果将具有多个上行链路端口的 vSphere Distributed Switch 用于端口绑定，则需要为每个物理网卡创建一个单独的分布式端口组。然后，应设置绑定策略，以便每个分布式端口组只有一个活动上行链路端口。有关分布式交换机的详细信息，请参见《vSphere 网络连接》文档。

## 为 iSCSI 或 iSER 创建一个 VMkernel 适配器

将运行 iSCSI 存储服务的 VMkernel 连接到 ESXi 主机上的物理网络适配器。然后，您可以在具有 iSCSI 或 iSER 适配器的端口绑定配置中使用已创建的 VMkernel 适配器。

### 前提条件

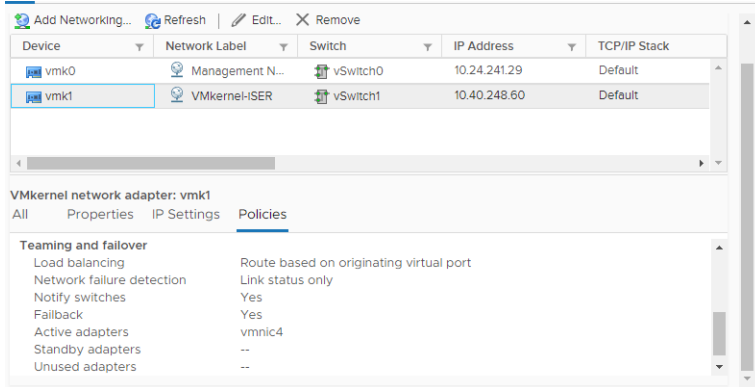
- 如果要为从属硬件 iSCSI 创建 VMkernel 适配器，必须使用与 iSCSI 组件对应的物理网络适配器 (vmnic#)。请参见[确定 iSCSI 与网络适配器之间的关联](#)。
- 对于 iSER 适配器，请务必使用支持 RDMA 的适当 vmnic#。请参见[安装和查看支持 RDMA 的网络适配器](#)。

### 步骤

- 1 在 vSphere Client 中，导航到 ESXi 主机。
- 2 从右键单击菜单中选择**添加网络**。
- 3 选择 **VMkernel 网络适配器**，然后单击**下一步**。
- 4 选择**新建标准交换机**以创建 vSphere 标准交换机。
- 5 单击**添加适配器**图标，然后选择要用于 iSCSI 的适当网络适配器 (vmnic#)。确保将该适配器分配给活动适配器。
- 6 输入网络标签。  
网络标签是用于标识您所创建的 VMkernel 适配器的友好名称，如 iSCSI 或 iSER。
- 7 指定 IP 设置。
- 8 检查信息，然后单击**完成**。  
此时即为主机上的物理网络适配器 (vmnic#) 创建了虚拟 VMkernel 适配器 (vmk#)。

## 9 验证您的配置。

- 在**网络连接**下，选择 **VMkernel 适配器**，然后从列表中选择 VMkernel 适配器 (vmk#)。
- 单击**策略**选项卡，然后确认对应的物理网络适配器 (vmnic#) 在**绑定和故障切换**下显示为活动适配器。



### 后续步骤

如果主机具有一个用于 iSCSI 流量的物理网络适配器，请将所创建的 VMkernel 适配器绑定到 iSCSI 或 iSER vmhba 适配器。

如果具有多个网络适配器，可以创建附加的 VMkernel 适配器，然后执行 iSCSI 绑定。虚拟适配器的数量必须与主机上物理适配器的数量相对应。有关信息，请参见 [iSCSI 或 iSER 配置中的多个网络适配器](#)。

## 将 iSCSI 或 iSER 适配器绑定到 VMkernel 适配器

在 ESXi 主机上，将 iSCSI 或 iSER 适配器与 VMkernel 适配器绑定。

### 前提条件

为主机上的每个物理网络适配器创建虚拟 VMkernel 适配器。如果使用多个 VMkernel 适配器，请设置正确的网络策略。

所需特权：**主机.配置.存储分区配置**

### 步骤

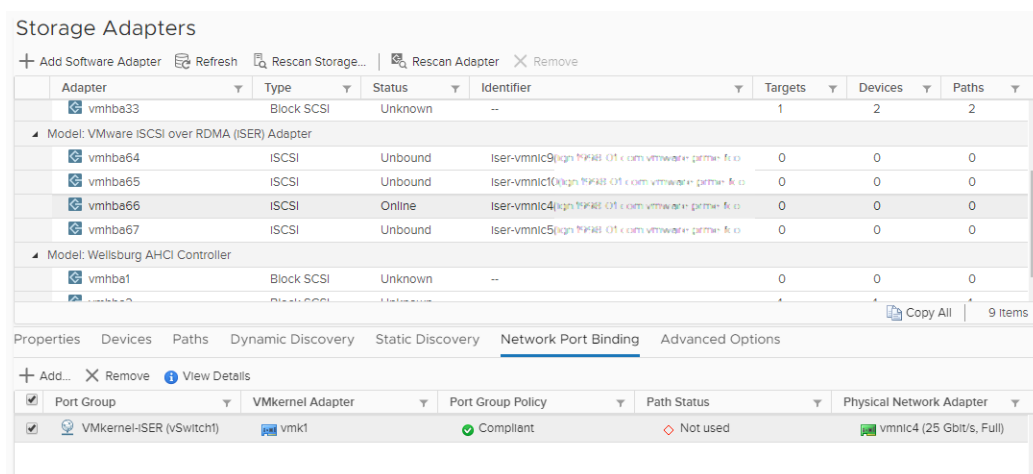
- 在 vSphere Client 中，导航到 ESXi 主机。
- 单击**配置**选项卡。
- 在**存储**下，单击**存储适配器**，然后从列表中选择相应的 iSCSI 或 iSER 适配器 (vmhba#)。
- 单击**网络端口绑定**选项卡，然后单击**添加**图标。
- 选择要与 iSCSI 或 iSER 适配器绑定的 VMkernel 适配器。

**注** 确保 VMkernel 适配器的网络策略符合绑定要求。

可以将软件 iSCSI 适配器绑定到一个或多个 VMkernel 适配器。对于从属硬件 iSCSI 适配器或 iSER 适配器，只能使用一个与正确的物理网卡关联的 VMkernel 适配器。

## 6 单击确定。

iSCSI 或 iSER 适配器的网络端口绑定列表上将显示此网络连接。



## 查看 ESXi 主机上的端口绑定详细信息

查看绑定到 iSCSI 或 iSER vmhba 适配器的 VMkernel 适配器的网络连接详细信息。

### 步骤

- 1 在 vSphere Client 中，导航到 ESXi 主机。
- 2 单击配置选项卡。
- 3 在存储下，单击存储适配器，然后从列表中选择相应的 iSCSI 或 iSER 适配器。
- 4 单击网络端口绑定选项卡，然后从列表中选择 VMkernel 适配器。
- 5 单击查看详细信息图标。
- 6 通过切换可用选项卡来查看 VMkernel 适配器和物理适配器信息。

### 后续步骤

有关可为 iSCSI 或 iSER 存储适配器执行的其他配置步骤，请参见以下主题：

- 配置从属硬件 iSCSI 适配器
- 配置软件 iSCSI 适配器
- 使用 ESXi 配置 iSER

## 将巨帧与 iSCSI 和 iSER 配合使用

ESXi 支持将巨帧与 iSCSI/iSER 配合使用。

巨帧是大小超过 1500 字节的以太网帧。通常情况下，最大传输单元 (MTU) 参数用于测量巨帧的大小。

将巨帧用于 iSCSI 流量时，需要考虑以下注意事项：

- 所有网络组件都必须支持巨帧。



- 请咨询供应商以确保您的物理网卡和 iSCSI 适配器支持巨帧。
- 要为巨帧设置和验证物理网络交换机，请参见供应商文档。

下表说明了 ESXi 对巨帧的支持级别。

表 5-1. 对巨帧的支持

iSCSI 适配器的类型	巨帧支持
软件 iSCSI	支持
从属硬件 iSCSI	受支持。咨询供应商。
独立硬件 iSCSI	受支持。咨询供应商。
VMware iSER	受支持。咨询供应商。

## 为网络连接启用巨帧

可以为使用 VMkernel 网络传输流量的 ESXi 存储适配器启用巨帧。这些适配器包括软件 iSCSI 适配器、从属硬件 iSCSI 适配器和 VMware iSER 适配器。

要启用巨帧，请更改最大传输单元 (MTU) 参数的默认值。您要用于 iSCSI 流量的 vSphere 交换机更改 MTU 参数。有关详细信息，请参见《vSphere 网络连接》文档。

### 步骤

- 1 在 vSphere Client 中，导航到 ESXi 主机。
- 2 单击**配置**选项卡。
- 3 在**网络**下，单击**虚拟交换机**，然后从列表中选择要修改的 vSphere 交换机。
- 4 单击**编辑设置**图标。
- 5 在“属性”页面上，更改 MTU 参数。

此步骤为该标准交换机上的所有物理网卡设置 MTU。将 MTU 值设置为连接到标准交换机的所有网卡中最大的 MTU 大小。ESXi 支持 MTU 最多包含 9000 个字节。

## 为独立硬件 iSCSI 启用巨帧

要为 ESXi 主机上的独立硬件 iSCSI 适配器启用巨型帧，请更改最大传输单元 (MTU) 参数的默认值。

使用“高级选项”设置更改 iSCSI HBA 的 MTU 参数。

### 步骤

- 1 在 vSphere Client 中，导航到 ESXi 主机。
- 2 单击**配置**选项卡。
- 3 在**存储**下单击**存储适配器**，然后从适配器列表中选择独立硬件 iSCSI 适配器。
- 4 单击**高级选项**选项卡，然后单击**编辑**。

## 5 更改 MTU 参数的值。

ESXi 支持 MTU 最多包含 9000 个字节。

# 在 ESXi 主机上配置 iSCSI 和 iSER 的动态或静态发现

需要设置目标发现地址，以便 iSCSI 或 iSER 存储适配器确定网络上可供访问的存储资源。

ESXi 系统支持以下发现方法：

## 动态发现

也称为 **SendTargets** 发现。启动器每次与指定的 iSCSI 服务器联系时，都会向该服务器发送 **SendTargets** 请求。服务器通过向启动器提供一个可用目标的列表来做出响应。这些目标的名称和 IP 地址显示在**静态发现**选项卡上。如果移除了通过动态发现添加的静态目标，则该目标会在下次进行重新扫描、重置存储适配器或重新引导主机时恢复到列表中。

---

**注** 使用软件和从属硬件 iSCSI，ESXi 可基于指定的 iSCSI 服务器地址的 IP 系列筛选目标地址。如果地址是 IPv4，则会筛选出可能进入来自 iSCSI 服务器的 **SendTargets** 响应的 IPv6 地址。当 DNS 名称用于指定 iSCSI 服务器，或来自 iSCSI 服务器的 **SendTargets** 响应包含 DNS 名称时，ESXi 将依赖于 DNS 查询中第一个解析条目的 IP 系列。

---

## 静态发现

除动态发现方法外，还可以使用静态发现并手动输入目标的信息。iSCSI 或 iSER 适配器使用您提供的目标列表与 iSCSI 服务器进行连接和通信。

设置静态或动态发现时，只能添加新的 iSCSI 目标。您可以更改现有目标的任何参数。要进行更改，请移除现有目标，然后添加一个新目标。

## 前提条件

所需特权：**主机.配置.存储分区配置**

## 步骤

- 1 在 vSphere Client 中，导航到 ESXi 主机。
- 2 单击**配置**选项卡。
- 3 在**存储**下，单击**存储适配器**，然后选择要配置的适配器 (vmhba#)。

4 配置发现方法。

发现方法	描述
动态发现	<div><div>a 依次单击<b>动态发现</b>和<b>添加</b>。</div><div>b 输入存储系统的 IP 地址或 DNS 名称，然后单击<b>确定</b>。</div><div>c 重新扫描 iSCSI 适配器。</div></div> <div>与 iSCSI 系统建立 SendTargets 会话后，主机会以新发现的所有目标填充“静态发现”列表。</div> <div><b>注</b> 即使将动态发现的目标从阵列端移除后，该目标仍会保留在列表中。</div>
静态发现	<div><div>a 依次单击<b>静态发现</b>和<b>添加</b>。</div><div>b 输入目标信息，然后单击<b>确定</b></div><div>c 重新扫描 iSCSI 适配器。</div></div>

后续步骤

有关可为 iSCSI 或 iSER 存储适配器执行的其他配置步骤，请参见以下主题：

- 设置独立硬件 iSCSI 适配器
- 配置从属硬件 iSCSI 适配器
- 配置软件 iSCSI 适配器
- 使用 ESXi 配置 iSER

移除动态或静态 iSCSI 目标

移除连接到 ESXi 主机的 iSCSI 服务器。

步骤

- 1 在 vSphere Client 中，导航到 ESXi 主机。
- 2 单击**配置**选项卡。
- 3 在**存储**下，单击**存储适配器**，然后从列表中选择要修改的 iSCSI 适配器。
- 4 在**动态发现**和**静态发现**之间切换。
- 5 选择要移除的 iSCSI 服务器，然后单击**移除**。
- 6 重新扫描 iSCSI 适配器。

如果要移除动态发现的静态目标，则需要执行重新扫描前将其从存储系统中移除。否则，重新扫描适配器时，主机自动发现该目标并将其添加到静态目标列表中。

## 为 iSCSI 或 iSER 存储适配器配置 CHAP 参数

由于 iSCSI 技术用来连接远程目标的 IP 网络不保护其传输的数据，因此必须确保连接的安全。质询握手身份验证协议 (CHAP) 是 iSCSI 实现的协议之一，该协议会验证访问网络上目标的启动器的合法性。

在主机和目标建立连接时，CHAP 使用三路握手算法验证主机和 iSCSI 目标（如果适用的话）的身份。系统根据启动器和目标共享的预定义的专用值或 CHAP 密钥进行验证。

ESXi 支持适配器级别的 CHAP 身份验证。在这种情况下，所有目标从 iSCSI 启动器接收相同的 CHAP 名称和密钥。对于软件和从属硬件 iSCSI 适配器以及 iSER 适配器，ESXi 还支持每个目标的 CHAP 身份验证，此身份验证使您能够为每个目标配置不同凭据以实现更高级别的安全性。

### 选择 CHAP 身份验证方法

ESXi 支持为所有类型的 iSCSI/iSER 启动器设置单向 CHAP，以及为软件和从属硬件 iSCSI 和 iSER 设置双向 CHAP。

在配置 CHAP 之前，请检查是否在 iSCSI 存储系统中启用了 CHAP。此外，还请获取有关系统支持的 CHAP 身份验证方法的信息。如果已启用 CHAP，请为启动器配置 CHAP，并确保 CHAP 身份验证凭据与 iSCSI 存储上的身份验证凭据相匹配。

ESXi 支持下列 CHAP 身份验证方法：

#### 单向 CHAP

在单向 CHAP 身份验证中，目标需验证启动器，但启动器无需验证目标。

#### 双向 CHAP

双向 CHAP 身份验证额外增加了一层安全保护。通过此方法，启动器也可以对目标进行身份验证。  
VMware 仅对软件和从属硬件 iSCSI 适配器以及 iSER 适配器支持此方法。

对于软件和从属硬件 iSCSI 适配器以及 iSER 适配器，可以为每个适配器或在目标级别设置单向 CHAP 和双向 CHAP。独立硬件 iSCSI 仅支持适配器级别的 CHAP。

设置 CHAP 参数时，请指定 CHAP 的安全级别。

**注** 指定 CHAP 安全级别时，存储阵列的响应方式取决于阵列的 CHAP 实施，且因供应商而异。有关不同的启动器和目标配置中的 CHAP 身份验证行为的信息，请参阅阵列文档。

表 5-2. CHAP 安全级别

CHAP 安全级别	描述	支持的存储适配器
无	主机不使用 CHAP 身份验证。如果启用了身份验证，则使用此选项会将其禁用。	独立硬件 iSCSI 软件 iSCSI 从属硬件 iSCSI iSER
使用单向 CHAP (如果目标需要)	主机首选非 CHAP 连接，但如果目标要求可以使用 CHAP 连接。	软件 iSCSI 从属硬件 iSCSI iSER

表 5-2. CHAP 安全级别（续）

CHAP 安全级别	描述	支持的存储适配器
使用单向 CHAP (除非目标禁止)	主机首选 CHAP，但如果目标不支持 CHAP，可以使用非 CHAP 连接。	独立硬件 iSCSI 软件 iSCSI 从属硬件 iSCSI iSER
使用单向 CHAP	主机需要成功的 CHAP 身份验证。如果 CHAP 协商失败，则连接失败。	独立硬件 iSCSI 软件 iSCSI 从属硬件 iSCSI iSER
使用双向 CHAP	主机和目标支持双向 CHAP。	软件 iSCSI 从属硬件 iSCSI iSER

## 为 iSCSI 或 iSER 存储适配器设置 CHAP

在 iSCSI/iSER 适配器级别设置 CHAP 名称和密钥时，所有目标都从适配器接收相同的参数。默认情况下，所有发现地址或静态目标都继承在适配器级别设置的 CHAP 参数。

CHAP 名称不得超过 511 个字母数字字符，CHAP 密钥不得超过 255 个字母数字字符。有些适配器（例如 QLogic 适配器）的限值更低，CHAP 名称不得超过 255 个字母数字字符，CHAP 密钥不得超过 100 个字母数字字符。

### 前提条件

- 为软件或从属硬件 iSCSI 设置 CHAP 参数前，要确定是配置单向还是双向 CHAP。独立硬件 iSCSI 适配器不支持双向 CHAP。
- 验证在存储端配置的 CHAP 参数。您配置的参数必须与在存储端配置的相符。
- 所需特权：**主机.配置.存储分区配置**

### 步骤

- 1 导航到 iSCSI 或 iSER 存储适配器。
  - a 在 vSphere Client 中，导航到 ESXi 主机。
  - b 单击**配置**选项卡。
  - c 在**存储**下，单击**存储适配器**，然后选择要配置的适配器 (vmhba#)。
- 2 单击**属性**选项卡，然后在**身份验证**面板中单击**编辑**。
- 3 指定身份验证方法。
  - 无
  - 使用单向 CHAP (如果目标需要)
  - 使用单向 CHAP (除非目标禁止)

- **使用单向 CHAP**
- **使用双向 CHAP。**要配置双向 CHAP，必须选择该选项。

#### 4 指定出站 CHAP 名称。

确保指定的名称与在存储端配置的名称相匹配。

- 要将 CHAP 名称设置为 iSCSI 适配器名称，请选中**使用启动器名称**。
- 要将 CHAP 名称设置为除 iSCSI 启动器名称之外的任何其他名称，请取消选中**使用启动器名称**，然后在**名称**文本框中输入名称。

#### 5 输入要在身份验证过程中使用的出站 CHAP 密钥。使用在存储端输入的同一密钥。

#### 6 如果配置双向 CHAP，请指定入站 CHAP 凭据。

确保对出站和入站 CHAP 使用不同的密钥。

#### 7 单击**确定**。

#### 8 重新扫描 iSCSI 适配器。

### 结果

如果更改了 CHAP 参数，则它们会用于新的 iSCSI 会话。但对于现有会话，注销并重新登录后才能使用新设置。

### 后续步骤

有关可为 iSCSI 或 iSER 存储适配器执行的其他配置步骤，请参见以下主题：

- [设置独立硬件 iSCSI 适配器](#)
- [配置从属硬件 iSCSI 适配器](#)
- [配置软件 iSCSI 适配器](#)
- [使用 ESXi 配置 iSER](#)

## 设置目标的 CHAP

如果使用软件和从属硬件 iSCSI 适配器或 iSER 存储适配器，可为每个发现地址或静态目标配置不同的 CHAP 凭据。

CHAP 名称不得超过 511 个字母数字字符，CHAP 密钥不得超过 255 个字母数字字符。

### 前提条件

- 在设置 CHAP 参数之前，先确定是要配置单向 CHAP 还是双向 CHAP。
- 验证在存储端配置的 CHAP 参数。您配置的参数必须与在存储端配置的相符。
- 所需特权：**主机.配置.存储分区配置**

## 步骤

- 1 导航到 iSCSI 或 iSER 存储适配器。
  - a 在 vSphere Client 中，导航到 ESXi 主机。
  - b 单击**配置**选项卡。
  - c 在**存储**下，单击**存储适配器**，然后选择要配置的适配器 (vmhba#)。
- 2 单击**动态发现**或**静态发现**。
- 3 从现有目标列表中，选择要配置的目标，然后单击**身份验证**。
- 4 取消选中**从父项继承设置**，然后指定身份验证方法。
  - 无
  - 使用单向 CHAP (如果目标需要)
  - 使用单向 CHAP (除非目标禁止)
  - 使用单向 CHAP
  - 使用双向 CHAP。要配置双向 CHAP，必须选择该选项。
- 5 指定出站 CHAP 名称。

确保指定的名称与在存储端配置的名称相匹配。

  - 要将 CHAP 名称设置为 iSCSI 适配器名称，请选中**使用启动器名称**。
  - 要将 CHAP 名称设置为除 iSCSI 启动器名称之外的任何其他名称，请取消选中**使用启动器名称**，然后在**名称**文本框中输入名称。
- 6 输入要在身份验证过程中使用的出站 CHAP 密钥。使用在存储端输入的同一密钥。
- 7 如果配置双向 CHAP，请指定入站 CHAP 凭据。

确保对出站和入站 CHAP 使用不同的密钥。
- 8 单击**确定**。
- 9 重新扫描存储适配器。

## 结果

如果更改了 CHAP 参数，则它们会用于新的 iSCSI 会话。但对于现有会话，直到注销并重新登录之后才能使用新设置。

## 配置 iSCSI 的高级参数

可能需要为 ESXi 主机上的 iSCSI 启动器配置其他参数。例如，有些 iSCSI 存储系统要求 ARP（地址解析协议）重定向，以在端口间动态移动 iSCSI 流量。在这种情况下，必须在主机上激活 ARP 重定向。

下表列出了可以使用 vSphere Client 进行配置的高级 iSCSI 参数。此外，可以使用 vSphere CLI 命令配置部分高级参数。有关信息，请参见《ESXCLI 入门》文档。

根据适配器的类型，某些参数可能不可用。

**重要说明** 请勿更改 iSCSI 高级设置，除非 VMware 技术支持或存储供应商指导您进行更改。

表 5-3. iSCSI 启动器的其他参数

高级参数	描述
头摘要	增加数据完整性。启用头摘要参数后，系统会对 iSCSI 协议数据单元 (PDU) 的每个标头部分计算校验和。系统会使用 CRC32C 算法验证数据。
数据摘要	增加数据完整性。启用数据摘要参数后，系统会对 PDU 的每个数据部分计算校验和。系统会使用 CRC32C 算法验证数据。  <b>注</b> 使用 Intel Nehalem 处理器的系统会卸载软件 iSCSI 的 iSCSI 摘要计算。该卸载有助于减少对性能的影响。
ErrorRecoveryLevel	主机上的 iSCSI 启动器在登录期间协商的 iSCSI 错误恢复级别 (ERL) 值。
LoginRetryMax	ESXi iSCSI 启动器在结束尝试之前尝试登录到目标的最大次数。
MaxOutstandingR2T	定义在收到确认 PDU 前可转换的 R2T（即将传输）PDU。
FirstBurstLength	指定在执行单个 SCSI 命令期间 iSCSI 启动器可以发送到目标的未经请求的数据的最大数量，以字节为单位。
MaxBurstLength	传入数据或请求的传出数据 iSCSI 序列中的最大 SCSI 数据负载，以字节为单位。
MaxRecvDataSegLength	在 iSCSI PDU 中可以接收的最大数据段长度，以字节为单位。
MaxCommands	可以在 iSCSI 适配器中排队的最大 SCSI 命令数。
DefaultTimeToWait	当连接意外终止或重置后，在尝试注销或重新分配活动任务之前需等待的最小秒数。
DefaultTimeToRetain	当连接终止或重置后，仍然可以重新分配活动任务的最大秒数。
LoginTimeout	启动器等待登录响应完成的时间，以秒为单位。
LogoutTimeout	启动器等待获取注销请求 PDU 响应的的时间，以秒为单位。
RecoveryTimeout	指定执行会话恢复的时间限制（秒）。如果该超时时间超过限制，则 iSCSI 启动器将结束会话。
无操作时间间隔	指定在从 iSCSI 启动器发送到 iSCSI 目标的 NOP-Out 请求之间的时间间隔（秒）。NOP-Out 请求充当验证 iSCSI 启动器和 iSCSI 目标之间的连接是否处于活动状态的 ping 机制。
无操作超时	指定主机收到 NOP-In 消息之前的时间限制（秒）。iSCSI 目标会发送该消息以响应 NOP-Out 请求。无操作超时时间超过限制时，启动器将结束当前会话并启动新的会话。
ARP 重定向	启用此参数后，存储系统可以将 iSCSI 流量从一个端口动态移动到另一个端口。存储系统执行基于阵列的故障切换需要有 ARP 参数。
延迟的 ACK	启用此参数后，存储系统可以延迟对已接收数据包的确认。



## 为 ESXi 主机上的 iSCSI 配置高级参数

高级 iSCSI 设置控制如标头、数据摘要、ARP 重定向、延迟的 ACK 等参数。

**小心** 除非在与 VMware 支持团队进行合作，或拥有为设置所提供值的全面信息，否则不要对高级 iSCSI 设置进行任何更改。

### 前提条件

所需特权：**主机.配置.存储分区配置**

### 步骤

- 1 在 vSphere Client 中，导航到 ESXi 主机。
- 2 单击**配置**选项卡。
- 3 在**存储**下，单击**存储适配器**，然后选择要配置的适配器 (vmhba#)。
- 4 配置高级参数。

选项	描述
在适配器级别	单击 <b>高级选项</b> 选项卡，然后单击 <b>编辑</b> 。
在目标级别	<ol style="list-style-type: none"> <li>a 单击<b>动态发现</b>或<b>静态发现</b>。</li> <li>b 从可用目标的列表中，选择要配置的目标，然后单击<b>高级</b>。</li> </ol>

- 5 为您要修改的高级参数输入任何所需的值。

## iSCSI 会话管理

要相互通信，iSCSI 启动器和目标需要建立 iSCSI 会话。您可以使用 vSphere CLI 查看和管理 iSCSI 会话。

默认情况下，软件 iSCSI 和从属硬件 iSCSI 启动器会在每个启动器端口与每个目标端口之间启动一个 iSCSI 会话。如果 iSCSI 启动器或目标有多个端口，则主机可以建立多个会话。每个目标的默认会话数等于 iSCSI 适配器上的端口数乘以目标端口数。

使用 vSphere CLI，您可以显示所有当前会话，以便对它们进行分析和调试。要创建多条指向存储系统的路径，可以通过复制 iSCSI 适配器与目标端口之间的现有会话来增加默认会话数。

此外，也可以建立与特定目标端口的会话。如果主机连接到仅向启动器呈现一个目标端口的单端口存储系统，则此功能非常有用。然后，系统可将更多会话重定向到不同的目标端口。在 iSCSI 启动器与另一个目标端口之间建立新会话，将创建一条指向存储系统的其他路径。

以下注意事项适用于 iSCSI 会话管理：

- 某些存储系统不支持源自同一启动器名称或端点的多个会话。尝试创建多个与此类目标的会话可能会导致 iSCSI 环境的行为不可预知。
- 存储供应商可提供自动会话管理器。使用自动会话管理器添加或删除会话不能保证持续的结果，并且可能影响存储性能。

## 查看 iSCSI 会话

使用 vCLI 命令可显示 iSCSI 适配器与存储系统之间的 iSCSI 会话。

### 前提条件

安装 ESXCLI。请参见《ESXCLI 入门》。要进行故障排除，请在 ESXi Shell 中运行 `esxcli` 命令。

### 步骤

- ◆ 要列出 iSCSI 会话，请运行以下命令：

```
esxcli iscsi session list
```

该命令采用以下选项：

选项	描述
<b>-A --adapter=</b> <i>str</i>	iSCSI 适配器名称，如 vmhba34。
<b>-s --isid=</b> <i>str</i>	iSCSI 会话标识符。
<b>-n --name=</b> <i>str</i>	iSCSI 目标名称，如 iqn.X。

## 添加 iSCSI 会话

使用 vCLI 为指定的目标添加 iSCSI 会话或复制现有会话。通过复制会话，可增加默认会话数并创建指向存储系统的其他路径。

### 前提条件

安装 ESXCLI。请参见《ESXCLI 入门》。要进行故障排除，请在 ESXi Shell 中运行 `esxcli` 命令。

### 步骤

- ◆ 要添加或复制 iSCSI 会话，请运行以下命令：

```
esxcli iscsi session add
```

该命令采用以下选项：

选项	描述
<b>-A --adapter=</b> <i>str</i>	iSCSI 适配器名称，如 vmhba34。此为必需选项。
<b>-s --isid=</b> <i>str</i>	要复制的会话的 ISID。您可以通过列出所有会话来进行查找。
<b>-n --name=</b> <i>str</i>	iSCSI 目标名称，如 iqn.X。

### 后续步骤

重新扫描 iSCSI 适配器。

## 移除 iSCSI 会话

使用 vCLI 命令可移除 iSCSI 适配器与目标之间的 iSCSI 会话。

### 前提条件

安装 ESXCLI。请参见《ESXCLI 入门》。要进行故障排除，请在 ESXi Shell 中运行 `esxcli` 命令。

### 步骤

- ◆ 要移除会话，请运行以下命令：

```
esxcli iscsi session remove
```

该命令采用以下选项：

选项	描述
<code>-A --adapter=str</code>	iSCSI 适配器名称，如 <code>vmhba34</code> 。此为必需选项。
<code>-s --isid=str</code>	要移除的会话的 ISID。您可以通过列出所有会话来查找。
<code>-n --name=str</code>	iSCSI 目标名称，如 <code>iqn.X</code> 。

### 后续步骤

重新扫描 iSCSI 适配器。

# 管理存储设备

# 6

管理 ESXi 主机可访问的本地和已联网的存储设备。

有关 ESXi 支持的存储设备的信息，请参见 [ESXi 支持哪些类型的物理存储和目标和设备表示形式](#)。

本章讨论了以下主题：

- [ESXi 存储设备名称和标识符](#)
- [ESXi 存储的重扫描操作](#)
- [确定 ESXi 存储连接问题](#)
- [启用或禁用 ESXi 存储设备上的定位符 LED](#)
- [擦除存储设备](#)
- [更改永久预留设置](#)

## ESXi 存储设备名称和标识符

在 ESXi 环境中，每个存储设备由多个名称进行标识。

### 设备标识符

ESXi 主机使用不同的算法和约定为每个存储设备生成标识符，具体取决于存储类型。

#### 存储提供的标识符

ESXi 主机查询目标存储设备的设备名称。主机从返回的元数据中提取或生成设备的唯一标识符。该标识符基于特定存储标准，在所有主机之间具有唯一和持久性，且采用以下格式之一：

- `naa.xxx`
- `eui.xxx`
- `t10.xxx`

#### 基于路径的标识符

如果设备未提供标识符，主机将生成 `mpx.path` 名称，其中 *path* 代表设备的第一个路径，例如 `mpx.vmhba1:C0:T1:L3`。此标识符的使用方法可以与存储提供的标识符相同。

假设本地设备的路径名称唯一时，才会为其创建 `mpx.path` 标识符。但是，此标识符不是唯一的也不是永久的，并且每次系统重新启动后都会发生变化。

设备路径通常采用以下格式：

*vmhbaAdapter:CChannel:TTarget:LLUN*

- *vmhbaAdapter* 是存储适配器的名称。此名称指的是主机上的物理适配器，而不是由虚拟机使用的 SCSI 控制器。

- *CChannel* 是存储通道号。

软件 iSCSI 适配器和从属硬件适配器使用通道号来显示到同一目标的多个路径。

- *TTarget* 为目标号。目标编号由主机确定，对主机可见的目标的映射更改时，编号也可能更改。由不同主机共享的目标可能没有相同的目标号。

- *LLUN* 是显示目标中 LUN 位置的 LUN 号。LUN 号由存储系统提供。如果目标只有一个 LUN，则 LUN 号始终为零 (0)。

例如，*vmhba1:C0:T3:L1* 表示通过存储适配器 *vmhba1* 和通道 0 访问的目标 3 上的 LUN 1。

## 旧标识符

除了设备提供的标识符或 *mpx.Path* 标识符，ESXi 还会为每个设备生成一个备用的旧名称。标识符具有以下格式：

*vml.number*

旧标识符包含一系列对于设备唯一的数字。可以从通过 SCSI INQUIRY 命令获取的元数据部分派生出标识符。对于未提供 SCSI INQUIRY 标识符的非本地设备，使用 *vml.number* 标识符作为唯一可用的标识符。

## 示例：在 vSphere CLI 中显示设备名称

您可以在 vSphere CLI 中使用 `esxcli storage core device list` 命令显示所有设备名称。输出与下例类似：

```
# esxcli storage core device list
naa.XXX
    Display Name: DGC Fibre Channel Disk(naa.XXX)
    ...
    Other UUIDs: vml.000XXX
mpx.vmhba1:C0:T0:L0
    Display Name: Local VMware Disk (mpx.vmhba1:C0:T0:L0)
    ...
    Other UUIDs: vml.0000000000XYZ
```

## 具有 NGUID 设备标识符的 NVMe 设备

对于 NVMe 设备，ESXi 会根据从设备检索的信息生成设备标识符。通常情况下，NVMe 设备支持 EUI64 或 NGUID 格式的标识符，或者使用这两种格式。NGUID 是指命名空间全局唯一标识符，使用 EUI64 16 字节指示符格式。

对于仅支持 NGUID 格式的设备，主机会创建两个标识符：*eui.xxx (NGUID)*（作为主标识符）和 *t10.xxx\_controller\_serial\_number*（作为备用主标识符）。

设备支持的 ID 格式		主机生成的设备标识符
EUI64 ID 格式	NGUID ID 格式	ESXi 8.0
是	是	t10.xxx_EUI64
是	否	t10.xxx_EUI64
否	是	eui.xxx (NGUID) (作为主 ID) t10.xxx_controller_serial_number (作为备用主 ID)

验证主设备标识符和备用设备标识符之间的映射

使用 `esxcli storage core device uidmap list` 命令验证设备标识符。输出类似以下内容：

```
esxcli storage core device uidmap list
eui.0000xyz.....
  Primary UID: eui.0000xyz.....
  Alternative Primary UIDs: t10.0000abc....
  Legacy UID: vml.00000000000766d68....
  Alternative Legacy UIDs: vml.0000000000080906....
```

重命名存储设备

ESXi 主机基于存储类型和制造商为存储设备分配显示名称。您可更改设备的显示名称。

无法重命名某些类型的本地设备。

步骤

- 1 在 vSphere Client 中，导航到 ESXi 主机。
- 2 单击配置选项卡。
- 3 在存储下，单击存储设备。
- 4 选择要重命名的设备，然后单击重命名。
- 5 将设备名称更改为友好名称。

ESXi 存储的重扫描操作

当执行存储管理任务或进行 SAN 配置更改时，可能需要重新扫描存储。

执行 VMFS 数据存储管理操作时，例如创建 VMFS 数据存储或 RDM、添加数据区和增加或删除 VMFS 数据存储，主机或 vCenter Server 将自动重新扫描和更新存储。可以通过关闭“主机重新扫描筛选器”来禁用自动重新扫描功能。请参见[关闭存储筛选器](#)。

在某些情况下，需要执行手动重新扫描。您可以重新扫描您的主机或文件夹中的所有主机、集群和数据中心可用的所有存储。

如果进行的更改只针对通过特定适配器连接的存储，则需要重新扫描此适配器。

每次进行以下更改之一后请手动执行重新扫描：

- 对 SAN 上的新磁盘阵列进行区域分配。
- 在 SAN 上创建新 LUN。
- 更改主机上的路径屏蔽。
- 重新连接线缆。
- 更改 CHAP 设置（只适用于 iSCSI）。
- 添加或移除发现或静态地址（只适用于 iSCSI）。
- 在 vCenter Server 中编辑或移除由 vCenter Server 主机和单台主机共享的数据存储之后，向 vCenter Server 中添加该单台主机。

**重要说明** 如果在某条路径不可用时重新扫描，则主机将从到设备的路径的列表中移除该路径。当该路径再次变得可用并且开始工作时，会立即重新显示在列表中。

## 执行存储重新扫描

在 SAN 配置中进行更改后，可能需要重新扫描存储。可以重新扫描 ESXi 主机、集群或数据中心可用的所有存储。如果进行的更改只针对通过特定主机访问的存储，则只重新扫描此主机。

### 步骤

- 1 在 vSphere Client 对象导航器中，浏览到主机、集群、数据中心或者包含主机的文件夹。
- 2 从右键单击菜单中选择**存储 > 重新扫描存储**。
- 3 指定重新扫描的数据区。

选项	描述
扫描新的存储设备	重新扫描所有适配器以发现新的存储设备。新发现的设备将显示在设备列表上。
扫描新的 VMFS 卷	重新扫描所有存储设备以发现上次扫描后添加的新数据存储。任何新数据存储均显示在数据存储列表中。

## 执行适配器重新扫描

在 SAN 配置中进行更改并且这些更改被隔离到通过 ESXi 主机上的特定适配器访问的存储中时，请只针对该适配器执行重新扫描。

### 步骤

- 1 在 vSphere Client 中，导航到 ESXi 主机。
- 2 单击**配置**选项卡。
- 3 在**存储**下，单击**存储适配器**，然后从列表中选择要重新扫描的适配器。
- 4 单击**重新扫描适配器**图标。

## 更改扫描的存储设备的数量

ESXi 主机的已扫描 LUN ID 的范围是 0 到 16,383。ESXi 会忽略大于 16,383 的 LUN ID。可配置的 `Disk.MaxLUN` 参数控制扫描 LUN ID 的范围。该参数的默认值为 1024。

当 SCSI 目标不支持使用 `REPORT_LUNS` 进行直接发现时，`Disk.MaxLUN` 参数还决定 SCSI 扫描代码尝试使用单个 `INQUIRY` 命令发现的 LUN 的数量。

您可以视需要修改 `Disk.MaxLUN` 参数。例如，如果环境中的存储设备数较少，LUN ID 介于 1 到 100 之间，请将值设置为 101。因此，您可以提高不支持 `REPORT_LUNS` 的目标上的设备发现速度。减小该值可缩短重新扫描时间和引导时间。但是，重新扫描存储设备所用的时间可能还取决于其他因素，包括存储系统的类型和存储系统上的负载。

在其他情况下，如果环境使用的 LUN ID 大于 1023，可能需要增大该值。

### 步骤

- 1 在 vSphere Client 中，导航到 ESXi 主机。
- 2 单击配置选项卡。
- 3 在系统下，单击高级系统设置。
- 4 在“高级系统设置”表中，选择 **Disk.MaxLUN**，然后单击编辑图标。
- 5 将现有值更改为所选的值，并单击确定。

所输入的值指定的是要发现的最后一个 LUN ID 后面的 LUN ID。

例如，要发现从 1 到 100 的 LUN ID，请将 **Disk.MaxLUN** 设置为 101。

## 确定 ESXi 存储连接问题

如果 ESXi 主机在连接到存储设备时遇到问题，则主机将根据某些因素将该问题视为永久或临时性问题。

存储连接问题由多种原因造成。尽管 ESXi 不能始终确定存储设备或其路径不可用的原因，但是主机对设备的永久设备丢失 (PDL) 状况和存储的暂时性全部路径异常 (APD) 状况能够加以区分。

### 永久设备丢失 (PDL)

是当存储设备出现永久性故障或以管理方式被移除或排除时所出现的一种情况。预计该设备将来不再可用。当设备永久不可用时，ESXi 会从存储阵列收到相应的感知代码或登录拒绝，并能够确认此设备已永久丢失。

### 全部路径异常 (APD)

是当主机无法访问存储设备且没有可用的设备路径时所出现的一种情况。ESXi 会将其视为一种暂时性状况，因为设备中出现的问题通常为临时性的，因而设备有望再次可用。



## 连接问题和 vSphere High Availability

当设备进入 PDL 或 APD 状态时，vSphere High Availability (HA) 会检测到连接问题，并为 ESXi 主机上受影响的虚拟机提供自动恢复。vSphere HA 使用虚拟机组件保护 (VMCP) 来防止在 vSphere HA 集群中主机上运行的虚拟机发生可访问性故障。有关 VMCP 以及如何配置在出现 APD 或 PDL 情况时数据存储和虚拟机的响应的详细信息，请参见《vSphere 可用性》文档。

### 检测 PDL 情况

如果 ESXi 主机永久无法使用某个存储设备，则会将其视为处于永久设备丢失 (PDL) 状态。

通常，如果无意移除了某个设备、其唯一 ID 发生更改或该设备发生不可恢复的硬件错误，就会出现 PDL 情况。

如果存储阵列确定设备永久无法使用，会向 ESXi 主机发送 SCSI 感知代码。收到感知代码后，您的主机将该设备识别为失败，并将设备的状态注册为 PDL。对于要视为永久丢失的设备，必须在其所有路径上都收到感知代码。

注册设备的 PDL 状态后，主机将停止尝试与设备重新建立连接或向其发送命令。

vSphere Client 会显示关于设备的以下信息：

- 该设备的操作状态更改为 Lost Communication。
- 全部路径显示为 Dead。
- 设备上的数据存储不可用。

如果设备不存在打开的连接或者在最后一个连接关闭之后，主机将移除 PDL 设备和该设备的所有路径。通过将高级主机参数 `Disk.AutoremoveOnPDL` 设置为 0 可禁用路径的自动删除功能。

如果设备通过 PDL 条件返回，则主机可以发现该设备，但会将其视为新设备。不保证已恢复设备上虚拟机的数据一致性。

---

**注** 当设备出现故障而未发送相应的 SCSI 感知代码或 iSCSI 登录拒绝时，主机无法检测 PDL 状况。在这种情况下，主机将继续将设备连接问题视为 APD，即使设备出现永久故障也是如此。

---

### 永久设备丢失和 SCSI 感知代码

以下是 VMkernel 日志中 SCSI 感知代码的示例，表明设备处于 PDL 状态。

```
H:0x0 D:0x2 P:0x0 Valid sense data: 0x5 0x25 0x0 or Logical Unit Not Supported
```

### 永久设备丢失和 iSCSI

在每个目标对应一个 LUN 的 iSCSI 阵列中，通过 iSCSI 登录失败检测 PDL。iSCSI 存储阵列会拒绝主机尝试启动 iSCSI 会话，原因为 Target Unavailable。与感知代码一样，也必须在所有路径上都收到该响应，才能将设备视为永久丢失。

## 永久设备丢失和虚拟机

注册设备的 PDL 状态后，主机会关闭虚拟机的所有 I/O。vSphere HA 可检测 PDL 并重新启动发生故障的虚拟机。

## 执行计划的存储设备移除

当存储设备出现故障时，您可以取消永久设备丢失 (PDL) 条件或全部路径异常 (APD) 条件。对存储设备执行有计划的移除和重新连接。

计划的设备移除是指有意断开存储设备的连接。您也可以出于多种原因计划移除设备，如升级硬件或重新配置存储设备。当按顺序执行移除和重新连接存储设备时，您需要完成一些任务。

任务	描述
从计划分离的设备迁移虚拟机。	《vCenter Server 和主机管理》
卸载设备上部署的数据存储。	请参见 <a href="#">卸载数据存储</a> 。
分离存储设备。	请参见 <a href="#">分离存储设备</a> 。
对于每目标具有一个 LUN 的 iSCSI 设备，请从具有该存储设备路径的每个 iSCSI HBA 中删除该静态目标项。	请参见 <a href="#">移除动态或静态 iSCSI 目标</a> 。
通过使用阵列控制台执行任何必要的存储设备重新配置。	请参见 <a href="#">供应商文档</a> 。
重新附加存储设备。	请参见 <a href="#">附加存储设备</a> 。
挂载数据存储并重新启动虚拟机。	请参见 <a href="#">挂载数据存储</a> 。

## 分离存储设备

可以安全地从 ESXi 主机分离存储设备。

您可能需要分离设备以使主机无法对其进行访问，例如，在存储端执行硬件升级时。

### 前提条件

- 该设备不包含任何数据存储。
- 该设备未被任何虚拟机作为 RDM 磁盘使用。
- 该设备不包含诊断分区或暂存分区。

### 步骤

- 1 在 vSphere Client 中，导航到 ESXi 主机。
- 2 单击**配置**选项卡。
- 3 在**存储**下，单击**存储设备**。
- 4 选择要分离的设备，然后单击**分离**图标。

### 结果

此时该设备无法进行访问。该设备的操作状态更改为“已卸载”。

后续步骤

如果多个主机共用该设备，请从每个主机分别分离该设备。

附加存储设备

重新附加先前与 ESXi 主机分离的存储设备。

步骤

- 1 在 vSphere Client 中，导航到 ESXi 主机。
- 2 单击配置选项卡。
- 3 在存储下，单击存储设备。
- 4 选择分离的存储设备，然后单击附加按钮。

结果

此时该设备可以进行访问。

从 PDL 情况中恢复

当存储设备变为永久不可用且未从 ESXi 主机正常分离时，会出现非计划的永久设备丢失 (PDL) 情况。

vSphere Client 中的以下各项表明设备处于 PDL 状况：

- 设备上部署的数据存储不再可用。
- 设备的操作状态更改为“通信中断”。
- 全部路径显示为“不活动”。
- VMkernel 日志文件中将出现有关设备永久不可用的警告消息。

要从非计划的 PDL 情况中恢复，并从主机中移除不可用的设备，请执行以下任务。

任务	描述
将受 PDL 情况影响的数据存储上运行的所有虚拟机关闭电源并取消注册。	请参见《vSphere 虚拟机管理》。
卸载数据存储。	请参见卸载数据存储。
重新扫描已访问该设备的所有 ESXi 主机。	请参见执行存储重新扫描。

**注** 如果重新扫描失败并且主机继续列出设备，则设备的某些挂起 I/O 或活动引用可能仍然存在。检查是否仍存在对设备或数据存储的引用处于活动状态的任何项目。这类项目包括虚拟机、模板、ISO 映像、裸设备映射等。

处理暂时性 APD 情况

如果存储设备在未指定的时间段内对 ESXi 主机不可用，则被视为处于全部路径异常 (APD) 状况。

例如，APD 状况的原因可能会是交换机发生故障，或者存储电缆断开连接。

与永久设备丢失 (PDL) 状况相反，主机将 APD 状况视为瞬间现象，并预计设备会再次可用。

主机继续重新尝试发出的命令，以便重新建立与设备之间的连接。如果主机命令在较长一段时间内无法重试，则主机可能存在具有性能问题的风险。主机及其虚拟机也可能会无法响应。

为了避免出现这些问题，您的主机使用默认的 APD 处理功能。当设备进入 APD 状态时，主机会打开定时器。在定时器打开的情况下，主机会仅在有限的时间段内继续重试非虚拟机命令。

默认情况下，APD 超时时间设置为 140 秒。通常，此值大于大多数设备从连接丢失中恢复所需的时间。如果设备在此时间内可用，主机及其虚拟机则会继续运行，而不会遭遇任何问题。

如果设备未恢复并且超时时间结束，则主机停止重新尝试，并停止任何非虚拟机 I/O。虚拟机 I/O 会继续重试。vSphere Client 对 APD 超时过期的设备显示以下信息：

- 该设备的操作状态更改为 Dead or Error。
- 全部路径显示为 Dead。
- 设备中的数据存储显示为灰色。

即便设备和数据存储不可用，虚拟机仍处于响应状态。您可以关闭虚拟机电源，或者将其迁移到不同的数据存储或主机。

如果稍后设备路径处于运行状态，则主机可以将 I/O 恢复到设备并结束特殊 APD 处理。

## 禁用存储 APD 处理

默认情况下，ESXi 主机上的存储全部路径异常 (APD) 处理功能处于启用状态。如果启用了该功能，主机会在有限的一段时间内持续向处于 APD 状态的存储设备重试非虚拟机 I/O 命令。时限到期后，主机将停止重试尝试，并终止所有非虚拟机 I/O。您可禁用主机上的 APD 处理功能。

如果禁用了 APD 处理功能，主机将无限期持续重试发出命令，尝试重新连接到 APD 设备。该行为可能导致主机上的虚拟机超过其内部 I/O 超时值而无响应或发生故障。主机可能与 vCenter Server 断开连接。

### 步骤

- 1 在 vSphere Client 中，导航到 ESXi 主机。
- 2 单击配置选项卡。
- 3 在系统下，单击高级系统设置。
- 4 在“高级系统设置”表中，选择 Misc.APDHandlingEnable 参数，然后单击 Edit 图标。
- 5 将该值设置为 0。

### 结果

如果禁用了 APD 处理功能，您可在某个设备进入 APD 状态时将其重新启用并将其值设置为 1。内部 APD 处理功能立即启动，定时器从处于 APD 状态的每个设备当前的超时值开始。

## 更改存储 APD 的超时限值

超时参数控制 ESXi 主机必须在多少秒内对出现全部路径异常 (APD) 状况的存储设备重试 I/O 命令。您可以更改默认超时值。

在设备进入 APD 状况后，超时期限立即开始。超时结束后，主机将 APD 设备标记为无法访问。主机将停止尝试重试并非来自虚拟机的任何 I/O。主机将继续重试虚拟机 I/O。

默认情况下，主机上的超时参数设置为 140 秒。在某些情况下，例如连接到您的 ESXi 主机的存储设备从断开连接状态恢复所需的时间超过 140 秒，您可以增加该超时值。

---

**注** 如果您在设备变得不可用之后更改超时参数，则更改对于该特定的 APD 事件不会生效。

---

### 步骤

- 1 在 vSphere Client 中，导航到 ESXi 主机。
- 2 单击配置选项卡。
- 3 在系统下，单击高级系统设置。
- 4 在“高级系统设置”表中，选择 **Misc.APDTimeout** 参数，然后单击 Edit 图标。
- 5 更改默认值。

您可以输入介于 20 和 99999 秒之间的值。

## 验证 ESXi 主机上存储设备的连接状态

使用 `esxcli` 命令可验证特定存储设备的连接状态。

### 前提条件

安装 ESXCLI。请参见《ESXCLI 入门》。要进行故障排除，请在 ESXi Shell 中运行 `esxcli` 命令。

### 步骤

- 1 运行 `esxcli storage core device list -d=device_ID` 命令。
- 2 查看 Status: 区域中的连接状态。
  - on- 设备已连接。
  - dead- 设备已进入 APD 状态。APD 定时器已启动。
  - dead timeout- APD 超时已过期。
  - not connected- 设备处于 PDL 状态。

## 启用或禁用 ESXi 存储设备上的定位符 LED

使用定位符 LED 识别特定存储设备，以便可从其他设备中对其进行定位。您可以打开或关闭定位符 LED。

### 步骤

- 1 在 vSphere Client 中，导航到 ESXi 主机。
- 2 单击**配置**选项卡。
- 3 在**存储**下，单击**存储设备**。
- 4 从存储设备列表中选择一个或多个磁盘，并启用或禁用定位符 LED 指示器。

选项	描述
启用	单击 <b>打开 LED</b> 图标。
禁用	单击 <b>关闭 LED</b> 图标。

## 擦除存储设备

某些功能（例如，vSAN 或虚拟闪存资源）要求 ESXi 主机使用干净存储设备。您可以清除 HDD 或闪存设备，并移除所有现有数据。

### 前提条件

- 确保主机处于连接状态。
- 确认计划擦除的设备未被使用。
- 所需特权：**主机.配置.存储**

### 步骤

- 1 在 vSphere Client 中，导航到 ESXi 主机。
- 2 单击**配置**选项卡。
- 3 在**存储**下，单击**存储设备**。
- 4 选择一个或多个设备，然后单击**清除分区**图标。
- 5 确认将要擦除的分区信息不是重要信息。
- 6 单击**确定**以确认您的更改。

## 更改永久预留设置

可以调整在 Windows Server 故障切换集群 (WSFC) 配置中用作物理裸设备映射 (RDM) 的存储设备上的永久预留设置。

分布于多个 ESXi 主机上的 WSFC 集群节点需要物理 RDM。RDM 在运行集群节点的所有主机之间共享。具有主动节点的主机在所有共享 RDM 设备上持有永久性 SCSI-3 预留。当主动节点正在运行且设备处于锁定状态时，任何其他主机都无法写入设备。如果主动节点锁定设备时其他参与主机引导，则引导可能需要很长时间，因为主机尝试连接锁定的设备失败。此同一问题还可能会影响重新扫描操作。

为防止出现此问题，请对具有 RDM 的辅助 WSFC 节点所在 ESXi 主机上的所有设备激活永久预留。此设置会通知 ESXi 主机设备上的永久性 SCSI 预留，以便主机可在引导或存储重新扫描过程中跳过设备。

如果稍后将标记的设备重新用作 VMFS 数据存储，请移除预留以避免出现不可预知的数据存储行为。

有关 WSFC 集群的信息，请参见《Windows Server 故障切换集群设置》文档。

### 前提条件

在将设备标记为永久保留之前，请确保该设备不包含 VMFS 数据存储。

### 步骤

- 1 在 vSphere Client 中，导航到 ESXi 主机。
- 2 单击**配置**选项卡。
- 3 在**存储**下，单击**存储设备**。
- 4 从存储设备列表中选择设备，然后单击以下图标之一。

选项	描述
标记为永久保留	将所选设备标记为永久保留。  <b>注</b> 对加入 WSFC 集群的每个 RDM 设备重复此过程。
取消标记为永久保留	清除之前标记的设备的永久预留。

### 结果

该配置将随 ESXi 主机一起永久存储，不受重新启动的影响。

### 示例

此外，还可以使用 `esxcli` 命令标记加入 WSFC 集群的设备。

- 1 将设备标记为永久保留。

```
esxcli storage core device setconfig -d naa.id --perennially-reserved=true
```

- 2 验证设备是否已永久保留。

```
esxcli storage core device list -d naa.id
```

在 `esxcli` 命令的输出中，搜索条目 `Is Perennially Reserved: true`。

- 3 要移除永久保留标记，请运行以下命令。

```
esxcli storage core device setconfig -d naa.id --perennially-reserved=false
```



# 对 ESXi 主机使用闪存设备

# 7

除了常规存储硬盘驱动器 (Hard Disk Drive, HDD) 以外, ESXi 还支持闪存设备。

与属于机电设备（包含运动零件）的常规 HDD 不同, 闪存设备将半导体用作其存储介质且没有运动零件。通常, 闪存设备具有很好的弹性, 并可提供对数据的快速访问。

要检测闪存设备, ESXi 可使用基于 T10 标准的查询机制。请咨询供应商, 以确认您的存储阵列是否支持 ESXi 闪存设备检测机制。

主机检测到闪存设备后, 您可以将其用于多项任务和功能。

如果使用 NVMe 存储, 请启用高性能插件 (HPP) 以提高您的存储性能。请参见 [VMware 高性能插件和路径选择方案](#)。

有关将 NVMe 存储与 ESXi 配合使用的详细信息, 请参见 [第 8 章 关于 VMware NVMe 存储](#)。

表 7-1. 对 ESXi 使用闪存设备

功能互操作	描述
vSAN	vSAN 需要闪存设备。有关详细信息, 请参见《管理 VMware vSAN》文档。
VMFS 数据存储	在闪存设备上创建 VMFS 数据存储。数据存储可用作下列用途: <ul style="list-style-type: none"><li>■ 存储虚拟机。某些客户机操作系统可以将存储在这些数据存储上的虚拟磁盘标识为闪存虚拟磁盘。</li><li>■ 为 ESXi 主机交换缓存分配数据存储空间。请参见 <a href="#">使用 VMFS 数据存储配置主机缓存</a>。</li></ul>
虚拟闪存资源 (VFFS)	如果供应商要求, 请设置虚拟闪存资源并对 I/O 缓存筛选器使用该资源。请参见 <a href="#">第 15 章 筛选虚拟机 I/O</a> 。

## 闪存设备和虚拟机

客户机操作系统可以将驻留在闪存数据存储上的虚拟磁盘标识为闪存虚拟磁盘。

客户机操作系统可使用标准查询命令, 例如对 SCSI 设备运行 SCSI VPD Page (B1h), 对 IDE 设备运行 ATA IDENTIFY DEVICE (Word 217)。

对于链接克隆、本机快照和增量磁盘, 查询命令将报告基础磁盘的虚拟闪存状态。

在下列情况下, 操作系统可以检测虚拟磁盘是否为闪存磁盘:

- 支持在虚拟硬件版本 8 或更高版本的虚拟机上检测闪存虚拟磁盘。

- 在所有主机上，必须将支持共享 VMFS 数据存储的设备标记为闪存。
- 如果 VMFS 数据存储包含多个设备数据区，则所有底层物理数据区都必须基于闪存。

本章讨论了以下主题：

- [ESXi 主机上闪存设备的最佳做法](#)
- [标记 ESXi 主机上的存储设备](#)
- [关于 vSphere 环境中的虚拟闪存资源](#)
- [使用 VMFS 数据存储配置主机缓存](#)

## ESXi 主机上闪存设备的最佳做法

在 vSphere 环境中使用闪存设备时，请遵循以下最佳做法。

- 使用《VMware 兼容性指南》批准的闪存设备。
- 确保将最新的固件与闪存设备配合使用。时常向存储供应商确认更新内容。
- 严密监控使用闪存设备的频率并计算其预计生命周期。预计生命周期取决于您继续使用闪存设备的频率。请参见[估算闪存设备的使用寿命](#)。
- 如果使用 NVMe 设备进行存储，请启用高性能插件 (HPP) 以提高您的存储性能。有关使用 NVMe 设备的详细信息，请参见 [VMware 高性能插件和路径选择方案](#)

## 监控闪存设备

您可以从 ESXi 主机监控某些关键闪存设备参数，包括 Media Wearout Indicator、Temperature 和 Reallocated Sector Count。

使用 `esxcli` 命令可以监控闪存设备。

### 前提条件

安装 ESXCLI。请参见《ESXCLI 入门》。要进行故障排除，请在 ESXi Shell 中运行 `esxcli` 命令。

### 步骤

- ◆ 通过运行以下命令显示闪存设备统计信息：

```
esxcli storage core device smart get -d=flash device_ID
```

## 估算闪存设备的使用寿命

在使用闪存设备时，可监控闪存设备的使用频率并估算其生命周期。

通常，存储供应商将可靠地估算闪存设备在理想情况下的生命周期。例如，供应商可能会保证在每天写入 20 GB 的情况下生命周期为 5 年。然而，在实际使用中，设备的预期生命周期将取决于 ESXi 主机每天实际生成的写入量。可按照下列步骤来计算闪存设备的使用寿命。

## 前提条件

请注意距上次重新引导 ESXi 主机已过去的天数。例如，十天。

## 步骤

- 1 获取自上次重新引导后写入闪存设备的块的总数量。

运行 **esxcli storage core device stats get -d=device\_ID** 命令。例如：

```
~ # esxcli storage core device stats get -d t10.000000000000000000
Device: t10.000000000000000000
Successful Commands: xxxxxxxx
Blocks Read: xxxxxxxx
Blocks Written: 629145600
Read Operations: xxxxxxxx
```

输出中的“写入的块”项显示自上次重新引导后写入设备的块的数量。在此示例中，该值为 629,145,600。每次重新引导后，该值会重置为 0。

- 2 计算写入的总量，并转换为 GB。

一个块是 512 字节。要计算写入的总量，请将“写入的块”值乘以 512，然后将得到的值转换为 GB。在此示例中，自上次重新引导后写入的总量大约为 322 GB。

- 3 估算每天平均写入量（以 GB 单位）。

用距上次重新引导后写入的总量除以距上次重新引导的天数。

如果上次重新引导是在十天之前，则每天的写入量为 32 GB。您可以算出在一段时间内的每天平均写入量。

- 4 使用以下公式估算设备的生命周期：

*供应商提供的每天写入量乘以供应商提供的生命周期除以每天实际平均写入量*

例如，如果供应商保证在每天写入 20 GB 的情况下生命周期为 5 年，而每天实际写入量为 30 GB，则闪存设备的生命周期约为 3.3 年。

## 保持闪存磁盘无 VMFS

如果您在安装或自动部署 ESXi 时使用自动分区引导选项，自动分区选项将在主机的本地存储中创建一个 VMFS 数据存储。在某些情况下，您需要保持本地存储闪存磁盘未格式化。

## 问题

默认情况下，自动分区将在主机上所有未使用的本地存储磁盘（包括闪存磁盘）中部署 VMFS 文件系统。

但是，使用 VMFS 格式化的闪存磁盘将不可用于虚拟闪存和 vSAN 等功能。这两种功能都要求使用未格式化的闪存磁盘，并且都不能与任何其他文件系统共享磁盘。

## 解决方案

要确保自动分区不使用 VMFS 格式化闪存磁盘，请在首次安装 ESXi 或引导 ESXi 主机时使用以下引导选项：

- **autoPartition=TRUE**
- **skipPartitioningSsds=TRUE**

如果使用 Auto Deploy，请在引用 ESXi 主机上设置这些参数。

- 1 在 vSphere Client 中，导航到要用作引用主机的主机，然后单击**配置**选项卡。
- 2 单击**系统**打开系统选项，然后单击**高级系统设置**。
- 3 设置以下各项。

参数	值
VMkernel.Boot.autoPartition	True
VMkernel.Boot.skipPartitioningSsds	True

- 4 重新引导主机。

如果您计划用于虚拟闪存资源和 vSAN 的闪存磁盘已具有 VMFS 数据存储，请移除这些数据存储。

## 标记 ESXi 主机上的存储设备

配置 vSAN 或设置虚拟闪存资源时，存储环境必须包含本地闪存设备。但是，当设备供应商不支持自动闪存设备检测时，ESXi 可能无法将某些存储设备识别为闪存设备。在其他情况下，某些设备不会检测为本地设备，ESXi 会将它们标记为远程设备。当设备未被识别为本地闪存设备时，针对 vSAN 或虚拟闪存资源提供的设备列表中不包含这些设备。将这些设备标记为本地闪存可使其可用于 vSAN 和虚拟闪存资源。

### 将存储设备标记为闪存设备

如果 ESXi 未将其设备识别为闪存，请将其标记为闪存设备。

当设备供应商不支持自动闪存磁盘检测时，ESXi 不会将某些设备识别为闪存。设备的“驱动器类型”列显示其类型为 HDD。

**小心** 将 HDD 设备标记为闪存会使数据存储以及使用这些数据存储的服务的性能降低。仅当您确定设备为闪存设备时标记它们。

#### 前提条件

确认设备不在使用中。

#### 步骤

- 1 在 vSphere Client 中，导航到 ESXi 主机。
- 2 单击**配置**选项卡。
- 3 在**存储**下，单击**存储设备**。

- 4 从存储设备列表中，选择一个或多个 HDD 设备，然后单击**标记为闪存磁盘 (F)** 图标。
- 5 单击**是**以保存所做的更改。

#### 结果

设备的类型更改为闪存。

#### 后续步骤

如果要标记的闪存设备可以在多个主机之间共享，请确保从共享此设备的所有主机标记设备。

## 将存储设备标记为本地

使用 ESXi 可以将设备标记为本地设备。当 ESXi 无法确定某些设备是否为本地设备时，此操作非常有用。

#### 前提条件

- 确保设备未共享。
- 关闭设备上的虚拟机电源，然后卸载关联的数据存储。

#### 步骤

- 1 在 vSphere Client 中，导航到 ESXi 主机。
- 2 单击**配置**选项卡。
- 3 在**存储**下，单击**存储设备**。
- 4 从存储设备列表中，选择一个或多个远程设备，然后单击**标记为本地**图标。
- 5 单击**是**以保存所做的更改。

## 关于 vSphere 环境中的虚拟闪存资源

可以将 ESXi 主机上的本地闪存设备汇总成单个虚拟化缓存层（称为虚拟闪存资源）。设置虚拟闪存资源时，将创建新文件系统，即虚拟闪存文件系统 (VFFS)。VFFS 是 VMFS 的派生物，后者可针对闪存设备进行优化并用于将物理闪存设备分组到单个缓存资源池。作为非持久资源，无法用于存储虚拟机。

设置虚拟闪存资源后，可以将其用于 I/O 缓存筛选器。请参见第 15 章 [筛选虚拟机 I/O](#)。

## 虚拟闪存资源的注意事项

配置虚拟闪存资源时，需要考虑几个注意事项。

- 一台 ESXi 主机上只能有一个虚拟闪存资源。在主机级别管理虚拟闪存资源。
- 不能使用虚拟闪存资源来存储虚拟机。虚拟闪存资源只是缓存层。
- 对于虚拟闪存资源，只能使用本地闪存设备。
- 可以从混合闪存设备中创建虚拟闪存资源。所有设备类型都将一视同仁，在进行 SAS、SATA 或 PCI Express 连接时不会区别对待。从混合闪存设备中创建资源时，请确保将执行方式类似的设备组合在一起，以实现最佳性能。

- 不能对虚拟闪存资源和 vSAN 使用相同的闪存设备。两者均需要各自独占和专用的闪存设备。

## 设置虚拟闪存资源

可以设置虚拟闪存资源或为现有虚拟闪存资源添加容量。

要设置虚拟闪存资源，请使用连接到主机或主机集群的本地闪存设备。要增加虚拟闪存资源的容量，您可以添加多个设备，最多不超过最高配置文档中指示的最大数目。单个闪存设备必须以独占方式分配给虚拟闪存资源。任何其他 vSphere 功能（例如 vSAN 或 VMFS）都不能与虚拟闪存资源共享此设备。

### 步骤

- 1 在 vSphere Client 中，导航到 ESXi 主机。
- 2 单击配置选项卡。
- 3 在虚拟闪存下，单击虚拟闪存资源管理。
- 4 单击下列选项之一。

选项	描述
添加容量	如果要在单个主机上创建虚拟闪存资源。
在集群上添加容量	如果要在集群上创建虚拟闪存资源。

- 5 从可用实体列表中，选择要用于虚拟闪存资源的一个或多个实体，然后单击确定。

如果闪存设备未显示在列表中，请参见[标记 ESXi 主机上的存储设备](#)。

选项	描述
本地 VMware 磁盘	选择空闲闪存设备的任意组合。 ESXi 将在其中一个设备上创建 VFFS 卷，然后在其余设备上扩展该卷。系统会在整个 VFFS 卷上配置虚拟闪存资源。 在主机上存在 VFFS 卷时，如果不首先选择现有 VFFS 卷，将无法选择任何空闲设备。
volume ID - 使用现有 VFFS 卷数据区进行配置	如果以前已使用 vmkfstools 命令在主机的一个闪存设备上创建 VFFS 卷，该卷也会显示在符合条件的实体列表中。仅为虚拟闪存资源选择此卷即可。也可以将它与空闲设备合并使用。ESXi 使用现有 VFFS 卷将其扩展到其他设备。

### 结果

将创建虚拟闪存资源。“设备备份”区域列出了所有用于虚拟闪存资源的设备。

### 后续步骤

对通过 vSphere APIs for I/O Filtering 开发的 I/O 缓存筛选器使用虚拟闪存资源。请参见[使用闪存设备与缓存 I/O 筛选器](#)。

可以通过向虚拟闪存资源中添加多个闪存设备来增加其容量。

## 移除虚拟闪存资源

您可能需要移除连接到 ESXi 主机的本地闪存设备上已部署的虚拟闪存资源。移除虚拟闪存资源将释放设备供其他服务使用。

### 前提条件

- 确认虚拟闪存资源未用于 I/O 筛选器。

### 步骤

- 1 在 vSphere Client 中，导航到 ESXi 主机。
- 2 单击**配置**选项卡。
- 3 在**虚拟闪存**下，单击**虚拟闪存资源管理**，然后单击**移除全部**。

### 结果

移除虚拟闪存资源并清除闪存设备后，该设备将可用于其他操作。

## 设置虚拟闪存使用率警报

设置警报以在 ESXi 主机上的虚拟闪存资源使用率超出指定的阈值时发出指示。

### 步骤

- 1 在 vSphere Client 中，导航到 ESXi 主机。
- 2 单击**配置**选项卡。
- 3 在**系统**下，单击**高级系统设置**。
- 4 选择要更改的设置，然后单击**编辑**按钮。

参数	描述
<b>VFLASH.ResourceUsageThreshold</b>	当虚拟闪存资源使用率超过阈值时，系统会触发主机 vFlash 资源使用情况警报。默认阈值为 80%。您可以将阈值更改为适当的值。当虚拟闪存资源使用率降低至阈值时，系统会清除该警报。

- 5 单击**确定**。

## 使用 VMFS 数据存储配置主机缓存

允许 ESXi 主机交换到主机缓存。此外，还可以更改为主机缓存分配的空间量。

ESXi 主机可以将闪存支持的存储实体的一部分用作所有虚拟机共享的交换缓存。

主机级别的缓存由短延迟时间磁盘上的文件组成，ESXi 将该磁盘用作虚拟机交换文件的回写缓存。主机上运行的所有虚拟机共享该缓存。虚拟机页面的主机级别交换充分利用可能受限的闪存设备空间。

### 前提条件

使用闪存设备作为支持设备创建 VMFS 数据存储。请参见[创建 VMFS 数据存储](#)。

## 步骤

- 1 在 vSphere Client 中，导航到 ESXi 主机。
- 2 单击**配置**选项卡。
- 3 在**存储**下，单击**主机缓存配置**。
- 4 在列表中选择闪存数据存储，然后单击**编辑**图标。
- 5 为主机缓存分配适当空间
- 6 单击**确定**。



# 关于 VMware NVMe 存储

# 8

在数据中心内，使用永久内存的非易失性内存 (NVM) 存储设备变得越来越普遍。NVM Express (NVMe) 是一种标准化协议，专门为与 NVM 设备进行高性能多队列通信而设计。ESXi 支持 NVMe 协议以连接到本地和网络存储设备。

本章讨论了以下主题：

- [VMware NVMe 概念](#)
- [VMware NVMe 存储的要求和限制](#)
- [为 NVMe over RDMA \(RoCE v2\) 存储配置适配器](#)
- [为 NVMe over TCP 存储配置适配器](#)
- [启用 NVMe over RDMA 或 NVMe over TCP 软件适配器](#)
- [为 NVMe over Fabrics 添加控制器](#)
- [移除 NVMe over RDMA 或 NVMe over TCP 软件适配器](#)

## VMware NVMe 概念

在 ESXi 环境中开始使用 NVMe 存储之前，可以让自己先熟悉基本 NVMe 概念。

### NVM Express (NVMe)

NVMe 是一种在主机和目标存储系统之间连接和传输数据的方法。NVMe 专为使用具有非易失性内存的更快存储介质（如闪存设备）而设计。这种类型的存储可实现低延迟、低 CPU 使用率和高性能，通常可以替代 SCSI 存储。

### NVMe 传输

NVMe 存储可以使用 PCIe 接口直接连接到主机，也可以通过不同的架构传输间接连接到主机。VMware NVMe over Fabrics (NVMe-oF) 在主机与共享存储阵列上的目标存储设备之间提供距离连接。

当前存在以下类型的 NVMe 传输。有关详细信息，请参见 [VMware NVMe 存储的要求和限制](#)。

NVMe 传输	ESXi 支持
NVMe over PCIe	本地存储。
NVMe over RDMA	共享 NVMe-oF 存储。采用 RoCE v2 技术。

NVMe 传输	ESXi 支持
NVMe over Fibre Channel (FC-NVMe)	共享 NVMe-oF 存储。
NVMe over TCP	共享 NVMe-oF 存储。

## NVMe 命名空间

在 NVMe 存储阵列中，命名空间是一定数量的非易失性内存支持的一个存储卷。在 ESXi 的上下文中，命名空间类似于存储设备或 LUN。在 ESXi 主机发现 NVMe 命名空间后，将在 vSphere Client 的存储设备列表中显示表示该命名空间的闪存设备。您可以使用此设备创建 VMFS 数据存储并存储虚拟机。

## NVMe 控制器

控制器与一个或多个 NVMe 命名空间关联，并提供 ESXi 主机与存储阵列中的命名空间之间的访问路径。要访问控制器，主机可以使用两种机制，即控制器发现和控制器连接。有关信息，请参见[NVMe over Fabrics 添加控制器](#)。

### 控制器发现

通过此机制，ESXi 主机首先连接发现控制器。发现控制器将返回可用控制器列表。选择供主机访问的控制器后，您的主机将可以使用与此控制器关联的所有命名空间。

### 控制器连接

ESXi 主机将连接到您指定的控制器。您的主机将可以使用与此控制器关联的所有命名空间。

## NVMe 子系统

通常，NVMe 子系统是一个存储阵列，其中可能包含多个 NVMe 控制器、多个命名空间、一个非易失性内存存储介质以及控制器与非易失性内存存储介质之间的一个接口。该子系统由子系统 NVMe 限定名称 (NQN) 标识。

## VMware 高性能插件 (HPP)

默认情况下，ESXi 主机使用 HPP 声明 NVMe-oF 目标。为 I/O 请求选择物理路径时，HPP 将应用相应的路径选择方案 (PSS)。有关 HPP 的信息，请参见[VMware 高性能插件和路径选择方案](#)。要更改默认路径选择机制，请参见[更改路径选择策略](#)。

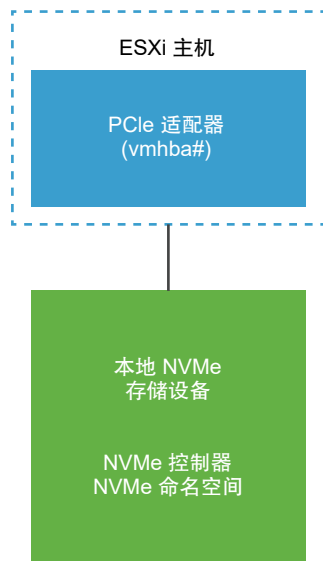
## 基本 VMware NVMe 架构和组件

ESXi 支持本地 NVMe over PCIe 存储和共享 NVMe-oF 存储，例如 NVMe over Fibre Channel、NVMe over RDMA (RoCE v2) 和 NVMe over TCP。

在 NVMe-oF 环境中，目标可以在主动/主动或非对称访问模式下向主机显示与 SCSI 中的 LUN 等效的命名空间。ESXi 能够发现和使用以任一方式显示的命名空间。ESXi 在内部将 NVMe-oF 目标模拟为 SCSI 目标，并将它们显示为主动/主动 SCSI 目标或隐式 ALUA SCSI 目标。

## VMware NVMe over PCIe

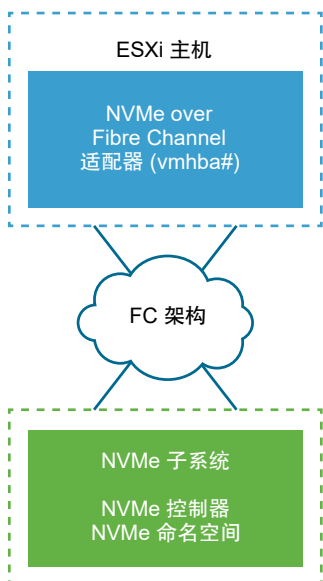
在此配置中，ESXi 主机使用 PCIe 存储适配器访问一个或多个本地 NVMe 存储设备。在主机上安装适配器后，主机发现可用 NVMe 设备，然后它们将显示在 vSphere Client 中的存储设备列表中。



## VMware NVMe over FC

此技术将 NVMe 映射到光纤通道协议，以便在主机计算机和目标存储设备之间传输数据和命令。此传输可以使用已升级的现有光纤通道基础架构来支持 NVMe。

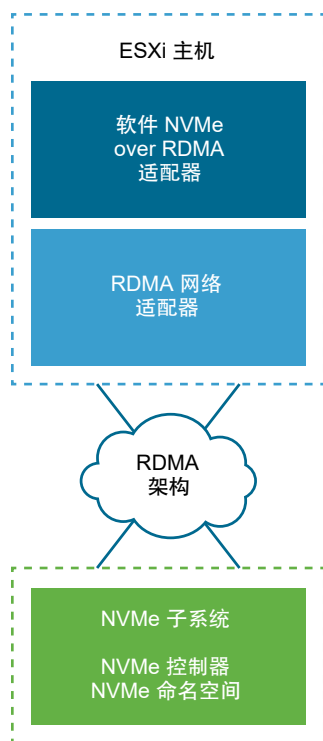
要访问 NVMe over Fibre Channel 存储，请在 ESXi 主机上安装支持 NVMe 的光纤通道存储适配器。无需配置该适配器。它会自动连接到相应的 NVMe 子系统，并发现它可以访问的所有共享 NVMe 存储设备。您可以稍后重新配置适配器并断开其控制器连接，或者连接在主机引导期间不可用的其他控制器。有关详细信息，请参见为 [NVMe over Fabrics 添加控制器](#)。



## NVMe over RDMA (RoCE v2)

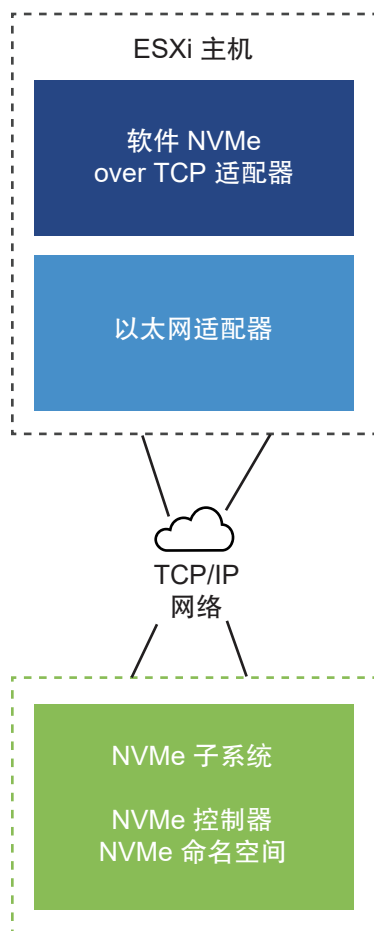
此技术在网络上的两个系统之间使用远程直接内存访问 (RDMA) 传输。该传输使得能够在主内存中进行数据交换，从而绕过操作系统或任一系统的处理器。ESXi 支持聚合以太网 RDMA v2 (RoCE v2) 技术，这使得可以通过以太网进行远程直接内存访问。

要访问存储，ESXi 主机使用主机上安装的 RDMA 网络适配器和软件 NVMe over RDMA 存储适配器。必须配置这两个适配器以使用它们进行存储发现。有关详细信息，请参见为 [NVMe over RDMA \(RoCE v2\) 存储配置适配器](#)。



## NVMe over TCP

此技术在两个系统之间使用以太网连接。要访问存储，ESXi 主机使用您的主机上安装的网络适配器和软件 NVMe over TCP 存储适配器。必须配置这两个适配器以使用它们进行存储发现。有关详细信息，请参见为 [NVMe over TCP 存储配置适配器](#)。



## VMware NVMe 存储的要求和限制

将 NVMe 技术与 VMware 结合使用时，请遵循特定准则和要求。

### NVMe over PCIe 的要求

ESXi 存储环境必须包含以下组件：

- 本地 NVMe 存储设备。
- 兼容的 ESXi 主机。
- 硬件 NVMe over PCIe 适配器。安装该适配器后，ESXi 主机会检测到该适配器并将其作为存储适配器 (vmhba) 显示在 vSphere Client 中，且将协议指示为 PCIe。无需配置该适配器。

### NVMe over RDMA (RoCE v2) 的要求

- 支持 NVMe over RDMA (RoCE v2) 传输的 NVMe 存储阵列。
- 兼容的 ESXi 主机。
- 支持无损网络的以太网交换机。

- 支持聚合以太网 RDMA (RoCE v2) 的网络适配器。要配置该适配器，请参见[查看 RDMA 网络适配器](#)。
- 软件 NVMe over RDMA 适配器。必须在 ESXi 主机上启用该软件组件，并将其连接到相应的网络 RDMA 适配器。有关信息，请参见[启用 NVMe over RDMA 或 NVMe over TCP 软件适配器](#)。
- NVMe 控制器。配置软件 NVMe over RDMA 适配器后，必须添加控制器。请参见[NVMe over Fabrics 添加控制器](#)。
- 无损以太网。

## NVMe over Fibre Channel 的要求

- 支持 NVMe 的光纤通道存储阵列。有关信息，请参见[第 3 章 将 ESXi 与光纤通道 SAN 配合使用](#)。
- 兼容的 ESXi 主机。
- 硬件 NVMe 适配器。通常，它是支持 NVMe 的光纤通道 HBA。安装该适配器后，ESXi 主机检测到该适配器并将其作为标准光纤通道适配器 (vmhba) 显示在 vSphere Client 中，且将存储协议指示为 NVMe。您无需配置硬件 NVMe 适配器即可进行使用。
- NVMe 控制器。无需配置该控制器。安装所需的硬件 NVMe 适配器后，它会自动连接到当前可访问的所有目标 and 控制器。您可以稍后断开这些控制器的连接，或者连接在主机引导期间不可用的其他控制器。请参见[NVMe over Fabrics 添加控制器](#)。

## NVMe over TCP 的要求

- 支持 NVMe over TCP 传输的 NVMe 存储阵列。
- 兼容的 ESXi 主机。
- 以太网适配器。
- 软件 NVMe over TCP 适配器。必须在 ESXi 主机上启用该软件组件，并将其连接到相应的网络适配器。有关详细信息，请参见[启用 NVMe over RDMA 或 NVMe over TCP 软件适配器](#)。
- NVMe 控制器。配置软件 NVMe over TCP 适配器后，必须添加控制器。请参见[NVMe over Fabrics 添加控制器](#)。

## VMware NVMe over Fabrics 共享存储支持

在 ESXi 环境中，NVMe 存储设备的显示类似于 SCSI 存储设备，并且可用作共享存储。使用 NVMe-oF 存储时，请遵循以下规则。

- 请勿混合使用传输类型来访问同一个命名空间。
- 确保将活动路径提供给主机。在发现活动路径之前，无法注册命名空间。

共享存储功能	SCSI over Fabric 存储	NVMe over Fabric 存储
RDM	支持	不受支持
核心转储	支持	不受支持
SCSI-2 预留	支持	不受支持

共享存储功能	SCSI over Fabric 存储	NVMe over Fabric 存储
集群 VMDK	支持	不受支持
带多写入者标志的共享 VMDK	支持	支持 (vSphere 7.0 Update 1 及更高版本)。 有关详细信息, 请参见 <a href="#">知识库文章</a> 。
Virtual Volumes	支持	不受支持
具有 VAAI 插件的硬件加速	支持	不受支持
默认 MPP	NMP	HPP (NMP 无法声明 NVMe-oF 目标)
限制	LUN=1024, 路径=4096	命名空间=32, 路径=128 (主机中每个命名空间最多 4 条路径)

## 为 NVMe over RDMA 配置无损以太网

ESXi 中的 NVMe over RDMA 需要使用无损以太网网络。

要建立无损网络, 您可以选择一个可用的 QoS 设置。

### 启用全局暂停流量控制

在此网络配置中, 请确保在网络交换机端口上启用了全局暂停流控制。此外, 请确保主机中支持 RDMA 的网卡可自动协商以使用正确的流控制。

要检查流量控制, 请运行以下命令:

```
#esxcli network nic get -n vmnicX
  Pause RX: true
  Pause TX: true
```

如果以上命令选项未设置为 true, 请运行以下命令。

```
#esxcli network nic pauseParams set -r true -t true -n vmnicX
```

### 启用优先级流控制

要确保 RoCE 流量无损, 必须在物理交换机和主机中将 PFC 优先级值配置为 3。可以通过两种方式在 ESXi 主机中配置 PFC:

- 自动配置。如果 RNIC 驱动程序支持 DCB 和 DCBx, 则在主机 RNIC 上自动应用 DCB PFC 配置。

您可以通过运行以下命令验证当前的 DCB 设置:

```
#esxcli network nic dcb status get -n vmnicX
```

- 手动配置。在某些情况下, RNIC 驱动程序提供了一种使用驱动程序特定的参数手动配置 DCB PFC 的方法。要使用此方法, 请参阅供应商特定的驱动程序文档。例如, 在 Mellanox ConnectX-4/5 驱动程序中, 可以通过运行以下命令, 然后重新引导主机, 将 PFC 优先级值设置为 3。

```
#esxcli system module parameters set -m nmlx5_core -p "pfctx=0x08 pfcrx=0x08"
```

## 启用基于 DSCP 的 PFC

基于 DSCP 的 PFC 是配置无损网络的另一种方式。在物理交换机和主机中，必须将 DSCP 值设置为 26。要使用此选项，请参阅供应商特定的驱动程序文档。例如，在 Mellanox ConnectX-4/5 驱动程序中，可以通过运行以下命令将 DSCP 标记值设置为 26。

1 启用 PFC 和 DSCP 信任模式。

```
#esxcli system module parameters set -m nmlx5_core -p "pfctx=0x08 pfcrx=0x08 trust_state=2"
```

2 将 DSCP 值设置为 26。

```
#esxcli system module parameters set -m nmlx5_rdma -p "dscp_force=26"
```

3 验证要检查的参数，以确认设置是否正确并且已设置。

```
esxcli system module parameters list -m nmlx5_core | grep 'trust_state\|pfctx\|pfcrx'
```

4 重新引导主机。

## 为 NVMe over RDMA (RoCE v2) 存储配置适配器

ESXi 主机上的适配器配置过程涉及到为 RDMA 网络适配器设置 VMkernel 绑定，然后启用软件 NVMe over RDMA 适配器。

整个配置过程包括以下操作。

操作	描述
<a href="#">查看 RDMA 网络适配器</a>	在 ESXi 主机上安装支持 RDMA (RoCE v2) 的网络适配器后，请使用 vSphere Client 查看 RDMA 适配器和物理网络适配器。
<a href="#">为 RDMA 适配器配置 VMkernel 绑定</a>	NVMe over RDMA 的端口绑定包括创建交换机以及将物理网络适配器和 VMkernel 适配器连接到交换机。RDMA 适配器将通过此连接绑定到 VMkernel 适配器。在配置中，可以使用 vSphere 标准交换机或 vSphere Distributed Switch。
<a href="#">启用 NVMe over RDMA 或 NVMe over TCP 软件适配器</a>	使用 vSphere Client 启用 NVMe over RDMA 软件存储适配器。
<a href="#">为 NVMe over Fabrics 添加控制器</a>	使用 vSphere Client 添加 NVMe 控制器。添加控制器后，与该控制器关联的 NVMe 命名空间将可用于您的 ESXi 主机。表示 ESXi 环境中命名空间的 NVMe 存储设备将显示在存储设备列表中。

以下视频将指导您完成配置 NVMe over RDMA 适配器的步骤。





## 查看 RDMA 网络适配器

在 ESXi 主机上安装支持 RDMA (RoCE v2) 的网络适配器后，请使用 vSphere Client 查看 RDMA 适配器和物理网络适配器。

### 步骤

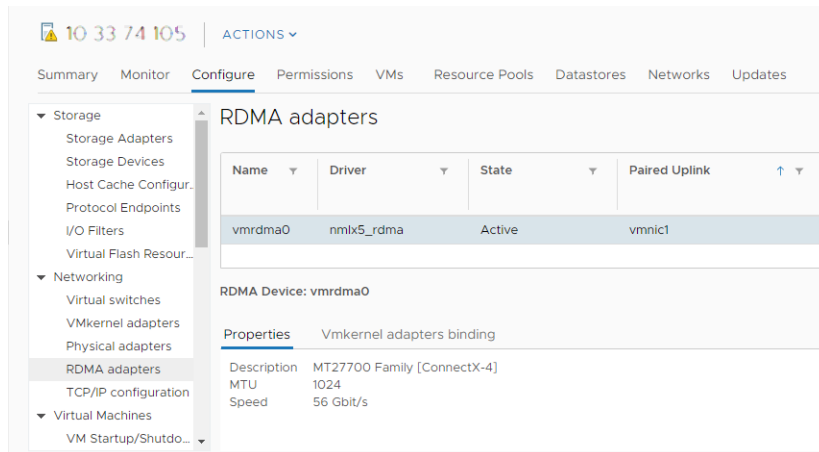
- 1 在 ESXi 主机上，安装支持 RDMA (RoCE v2) 的适配器，例如，Mellanox Technologies MT27700 系列 ConnectX-4。

主机将发现此适配器，并且 vSphere Client 将显示它的两个组件：一个 RDMA 适配器和一个物理网络适配器。

- 2 在 vSphere Client 中，确认您的主机已发现 RDMA 适配器。

- a 导航到主机。
- b 单击**配置**选项卡。
- c 在**网络**下，单击 **RDMA 适配器**。

在此示例中，RDMA 适配器在列表中显示为 vmrdma0。**配对上行链路**列会将网络组件显示为 vmnic1 物理网络适配器。



- d 要验证适配器的描述，请从列表中选择 RDMA 适配器，然后单击**属性**选项卡。

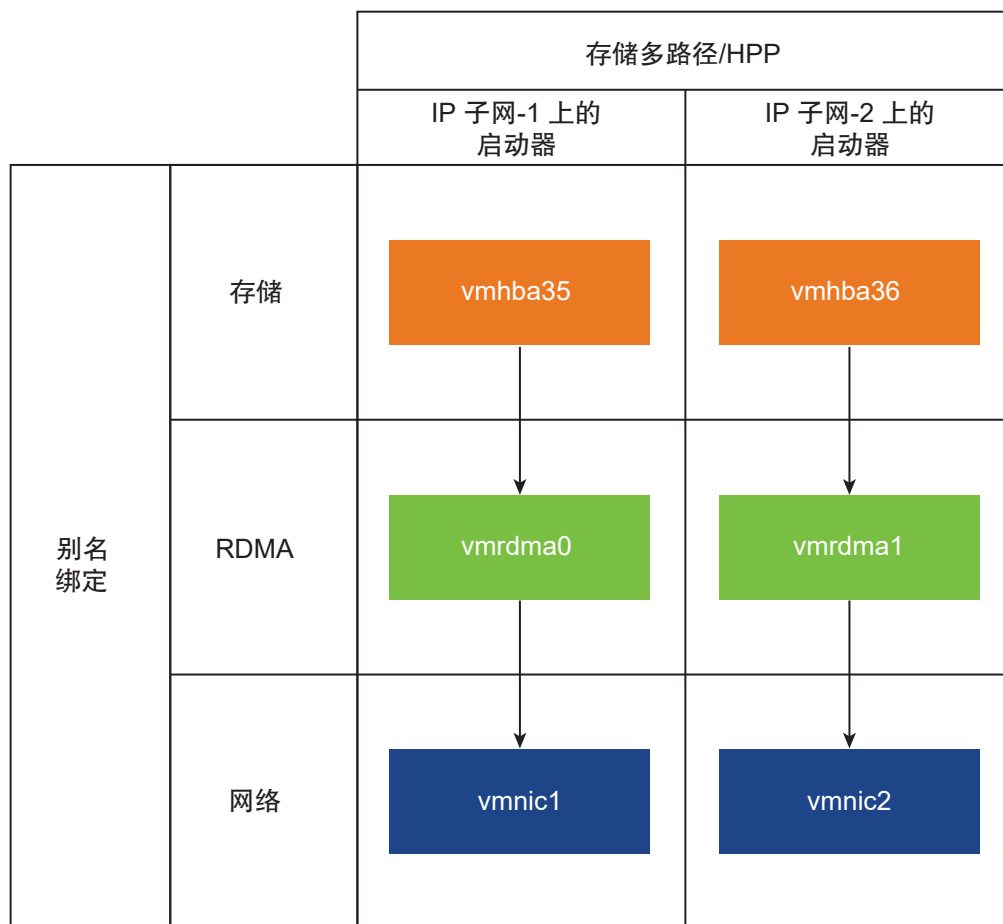
### 后续步骤

现在，您可以创建软件 NVMe over RDMA 适配器。

## 为 RDMA 适配器配置 VMkernel 绑定

NVMe over RDMA 的端口绑定包括创建交换机以及将物理网络适配器和 VMkernel 适配器连接到交换机。RDMA 适配器将通过此连接绑定到 VMkernel 适配器。在配置中，可以使用 vSphere 标准交换机或 vSphere Distributed Switch。

下图显示了 NVMe over RDMA 适配器的端口绑定。

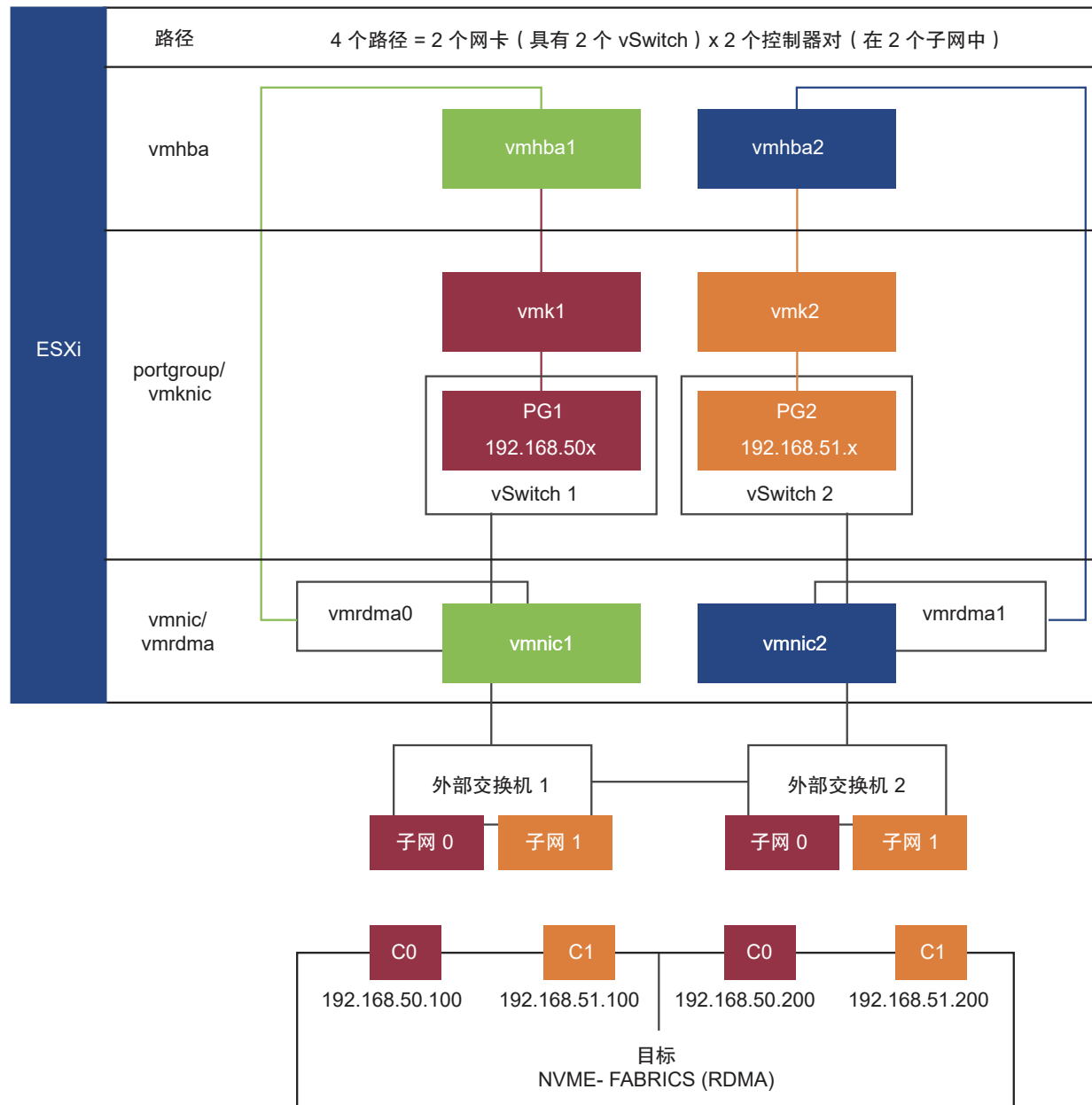


有关创建交换机的详细信息，请参见《vSphere 网络连接》文档中的“创建 vSphere 标准交换机”或“创建 vSphere Distributed Switch”。

## 具有 NVMe over RDMA 的网络拓扑示例

在此示例中，两个 vSphere 标准交换机和两个上行链路（支持 RDMA 的网卡）提供了高可用性。它们连接到两个子网中的两个控制器对。

通过多个 vSwitch 和多个上行链路 (RNIC) 实现 HA



## 使用 vSphere 标准交换机配置 VMkernel 绑定

可以使用 vSphere 标准交换机和每个交换机一个上行链路为 RDMA 适配器配置 VMkernel 端口绑定。配置网络连接包括为各个物理网络适配器创建虚拟 VMkernel 适配器。可以在每个虚拟和物理网络适配器之间使用 1:1 映射。

### 步骤

- 1 创建具有 VMkernel 适配器和网络组件的 vSphere 标准交换机。
  - a 在 vSphere Client 中，选择您的主机，然后单击**网络**选项卡。
  - b 单击**操作 > 添加网络**。
  - c 选择 **VMkernel 网络适配器**，然后单击**下一步**。
  - d 选择**新建标准交换机**，然后单击**下一步**。
  - e 在**分配的适配器**下，单击 **+**。  
将显示可用物理适配器的列表。
  - f 选择所需的物理适配器 vmnic，然后单击**确定**。

---

**注** 确保选择与 RDMA 适配器对应的物理网络适配器。要查看 RDMA 适配器 vmrdma 与物理网络适配器 vmnic 之间的关联，请参见[查看 RDMA 网络适配器](#)。

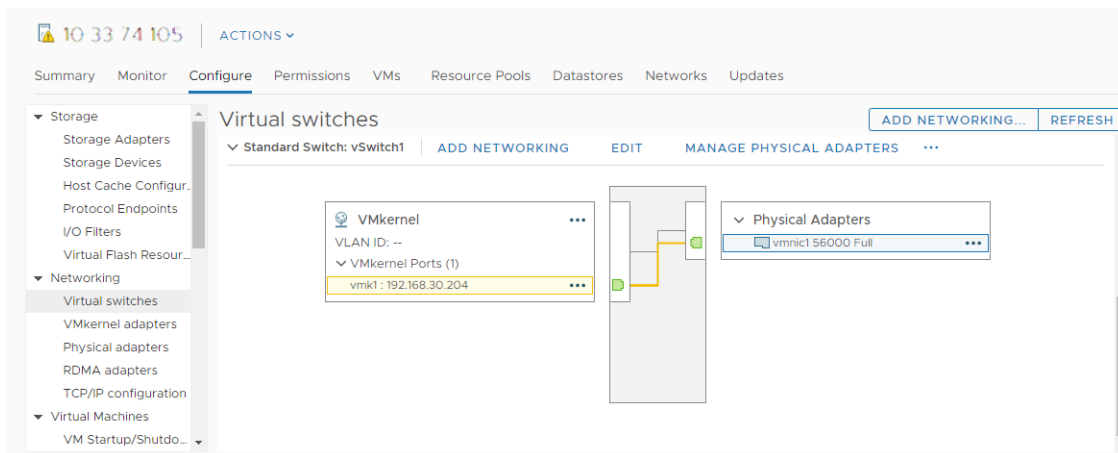
---

- g 在 **VMkernel 端口设置**下，输入所需的值。  
如果要对存储路径使用 VLAN，请输入 VLAN ID。
  - h 在 **IP 设置**列表中，输入 VMkernel IPv4 设置。
  - i 在“可用服务”下，选择 **NVMe over RDMA**。

## 2 验证交换机是否已正确配置。

- a 在**配置**选项卡上，选择**网络**下的**虚拟交换机**。
- b 展开交换机并验证其配置。

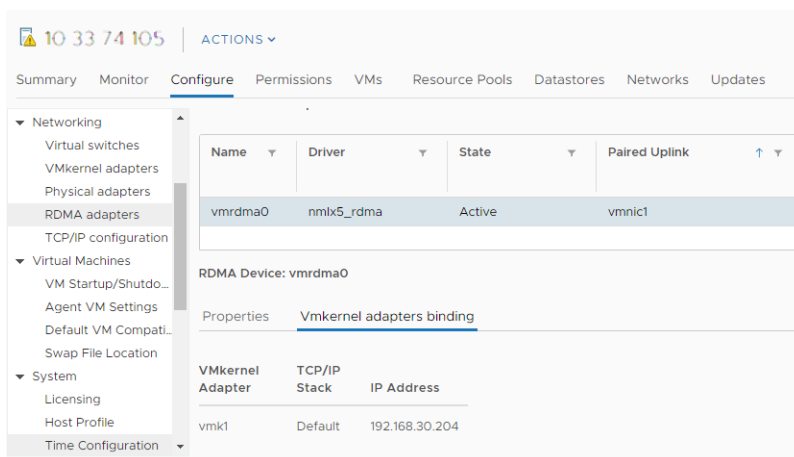
下图显示物理网络适配器和 VMkernel 适配器连接到了 vSphere 标准交换机。通过此连接，RDMA 适配器已绑定到 VMkernel 适配器。



## 3 验证 RDMA 适配器的 VMkernel 绑定配置。

- a 在**网络**列表下，单击 **RDMA 适配器**，然后从列表中选择 RDMA 适配器。
- b 单击 **VMkernel 适配器绑定** 选项卡，然后验证关联的 VMkernel 适配器是否显示在页面上。

在此示例中，vmrdma0 RDMA 适配器与 vmnic1 网络适配器进行配对并连接到 vmk1 VMkernel 适配器。



## 使用 vSphere 标准交换机和网卡绑定配置 VMkernel 绑定

可以使用具有网卡绑定配置的 vSphere 标准交换机为 RDMA 适配器配置 VMkernel 端口绑定。您可以使用网卡绑定来实现网络冗余。您可以将两个或多个网络适配器（网卡）配置为一个组来提供高可用性和负载均衡。

### 步骤

- 1 创建具有 VMkernel 适配器和采用网卡绑定配置的网络组件的 vSphere 标准交换机。
  - a 在 vSphere Client 中，选择您的主机，然后单击**网络**选项卡。
  - b 单击**操作 > 添加网络**。
  - c 选择 **VMkernel 网络适配器**，然后单击**下一步**。
  - d 选择**新建标准交换机**，然后单击**下一步**。
  - e 在**分配的适配器**下，单击 **+**。  
将显示可用物理适配器的列表。
  - f 选择所需的物理适配器 vmnic，然后将其添加到**活动适配器**下。
  - g 选择另一个物理适配器 vmnic，然后将其添加到**未用的适配器**下。
  - h 在 **VMkernel 端口设置**下，输入所需的值。  
如果要对存储路径使用 VLAN，请输入 VLAN ID。
  - i 在 **IP 设置**列表中，指定 VMkernel IPv4 设置。
  - j 在“可用服务”下，选择 **NVMe over RDMA**。

重复步骤 1 以配置现有标准交换机。
- 2 为交换机配置网卡绑定配置。
  - a 单击**配置**选项卡，然后选择**网络**下的**虚拟交换机**。
  - b 选择相应的 VMkernel 适配器。
  - c 从右键单击菜单中，单击**编辑设置**。
  - d 选择**绑定和故障切换**。
  - e 在**活动适配器**下，移动所需物理适配器 vmnic。
  - f 在**备用适配器 > 故障切换顺序**下，移动其他物理适配器。
  - g 设置合适的负载平衡和其他属性。
  - h 重复上述步骤以配置其他 VMkernel 适配器。
- 3 重复步骤 1 和 2，添加并配置一组额外的绑定 rnic。要验证适配器是否已配置，请单击**配置**选项卡，然后选择 **VMkernel 适配器**

## 使用 vSphere Distributed Switch 配置 VMkernel 绑定

可以使用 vSphere Distributed Switch 按每个交换机分配一个上行链路的方式为 RDMA 适配器配置 VMkernel 端口绑定。配置网络连接包括为各个物理网络适配器创建虚拟 VMkernel 适配器。可以在每个虚拟和物理网络适配器之间使用 1:1 映射。

### 步骤

- 1 创建具有 VMkernel 适配器和网络组件的 vSphere Distributed Switch。
  - a 在 vSphere Client 中，选择**数据中心**，然后单击**网络**选项卡。
  - b 单击**操作**，然后选择 **Distributed Switch > 新建 Distributed Switch**。
  - c 选择交换机的名称。  
确保主机中存在数据中心的位置，然后单击**下一步**。
  - d 选择兼容的 ESXi 版本，然后单击**下一步**。
  - e 输入所需的上行链路数，然后单击**完成**。
- 2 将一个或多个主机添加到分布式虚拟交换机。
  - a 在 vSphere Client 中，选择**数据中心**，然后单击 **Distributed Switch**。  
此时将显示可用 DSwitch 的列表。
  - b 右键单击 DSwitch，然后从菜单中选择**添加和管理主机**。
  - c 选择**添加主机**，然后单击**下一步**。
  - d 选择您的主机，然后单击**下一步**。
  - e 选择**分配上行链路**。
  - f 输入相关上行链路以分配 vmnic。
  - g 分配 VMkernel 适配器，然后单击**下一步**。
  - h 在 vSphere Client 中，选择 DSwitch，然后单击**端口**选项卡。  
您可以在此处查看为交换机创建的上行链路。
- 3 为 NVMe over RDMA 存储路径创建分布式端口组。
  - a 在 vSphere Client 中，选择所需的 DSwitch。
  - b 单击**操作**，然后选择**分布式端口组 > 新建分布式端口组**。
  - c 在**配置设置**下，输入端口组的常规属性。  
如果配置了特定的 VLAN，请将其添加到 VLAN ID 中。

---

**注** 如果未正确配置 VLAN，则可能会出现网络连接问题。

---

- 4 配置 VMkernel 适配器。
  - a 在 vSphere Client 中，展开 **DSwitch** 列表，然后选择分布式端口组。
  - b 单击**操作 > 添加 VMkernel 适配器**。
  - c 在**选择成员主机**对话框中，选择您的主机，然后单击**确定**。
  - d 在**配置 VMkernel 适配器**对话框中，确保 MTU 与交换机 MTU 匹配。
  - e 在**可用服务**下，选择 **NVMe over RDMA** 以进行适当的标记。
  - f 单击**完成**。
  - g 重复步骤 b 和步骤 c，添加多个支持 RDMA 的网卡。
- 5 为分布式端口组设置网卡绑定策略。
  - a 在**分布式端口组**中，单击**操作 > 编辑设置**。
  - b 单击**绑定和故障切换**，然后验证活动上行链路。
  - c 将一个上行链路分配为端口组的**活动**上行链路，将另一个上行链路分配为**未使用**。  
对创建的每个端口组重复步骤 c。

后续步骤

完成配置后，单击**配置**，然后验证主机上的物理适配器选项卡是否列出所选网卡的 DVSwitch。

## 为 NVMe over TCP 存储配置适配器

ESXi 主机上的适配器配置过程涉及到为 TCP 网络适配器设置 VMkernel 绑定，然后启用 NVMe over TCP 软件适配器。

整个配置过程包括以下操作。

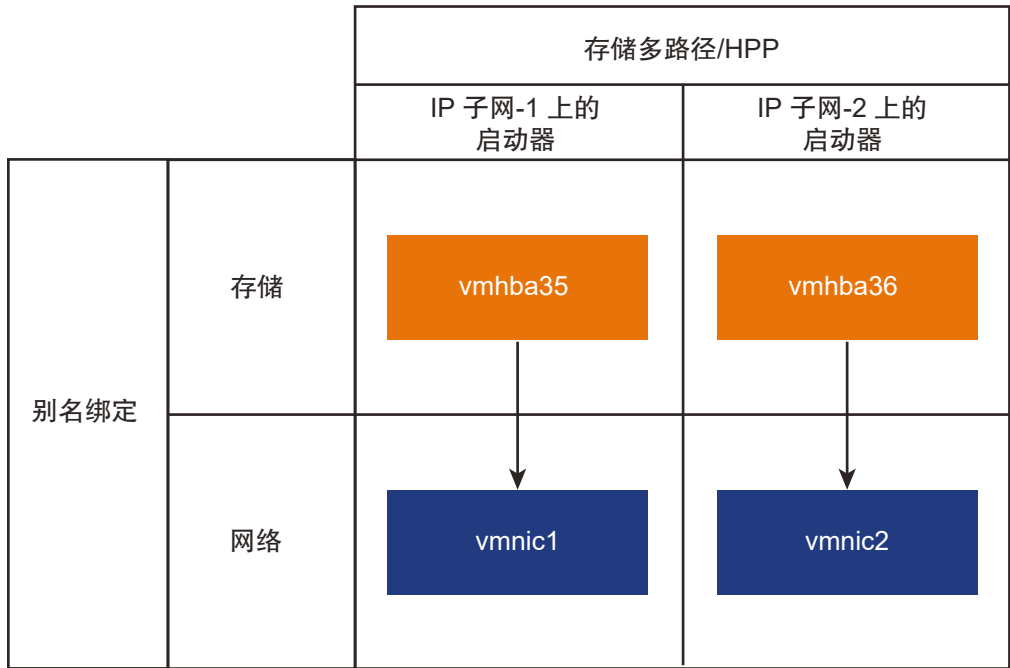
操作	描述
为 NVMe over TCP 适配器配置 VMkernel 绑定	NVMe over TCP 的端口绑定涉及创建虚拟交换机以及将物理网络适配器和 VMkernel 适配器连接到虚拟交换机。通过此连接，TCP 适配器将绑定到 VMkernel 适配器。在配置中，可以使用 vSphere 标准交换机或 vSphere Distributed Switch。
启用 NVMe over RDMA 或 NVMe over TCP 软件适配器	使用 vSphere Client 启用 NVMe over TCP 软件存储适配器。
为 NVMe over Fabrics 添加控制器	使用 vSphere Client 添加 NVMe 控制器。添加控制器后，与该控制器关联的 NVMe 命名空间将可用于您的 ESXi 主机。表示 ESXi 环境中命名空间的 NVMe 存储设备将显示在存储设备列表中。

## 为 NVMe over TCP 适配器配置 VMkernel 绑定

NVMe over TCP 的端口绑定涉及创建虚拟交换机以及将物理网络适配器和 VMkernel 适配器连接到虚拟交换机。通过此连接，TCP 适配器将绑定到 VMkernel 适配器。在配置中，可以使用 vSphere 标准交换机或 vSphere Distributed Switch。



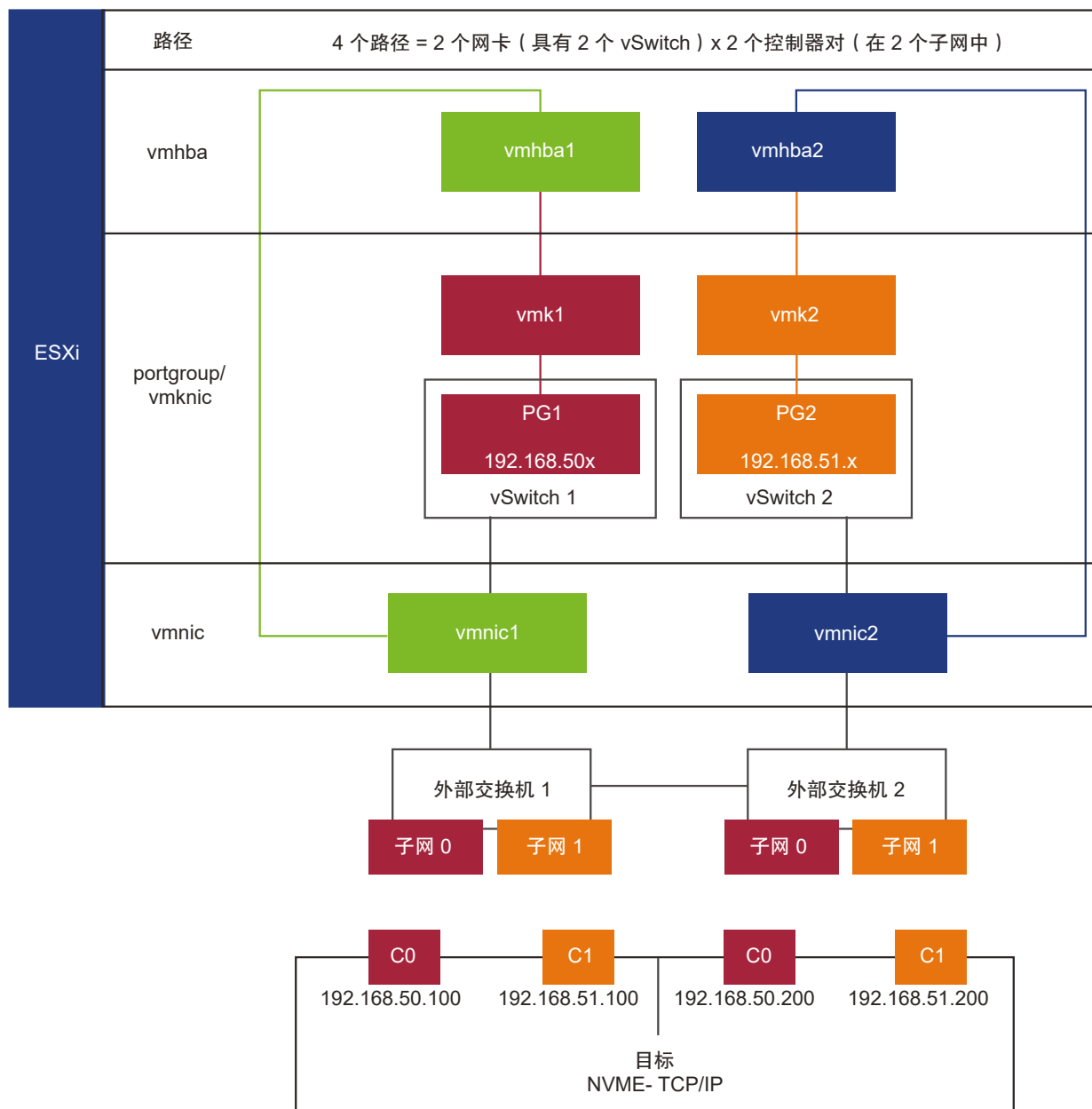
下图显示了 NVMe over TCP 适配器的端口绑定。



有关创建交换机的详细信息，请参见《vSphere 网络连接》文档中的“创建 vSphere 标准交换机”或“创建 vSphere Distributed Switch”。

## 具有 NVMe over TCP 的网络拓扑示例

在此示例中，两个 vSphere 标准交换机和两个上行链路提供高可用性。它们连接到两个子网中的两个控制器对。



## 为具有 vSphere Standard Switch 的 TCP 适配器配置 VMkernel 绑定

可以使用 vSphere 标准交换机按每个交换机分配一个上行链路的方式为 TCP 适配器配置 VMkernel 端口绑定。配置网络连接包括为各个物理网络适配器创建虚拟 VMkernel 适配器。可以在每个虚拟和物理网络适配器之间使用 1:1 映射。

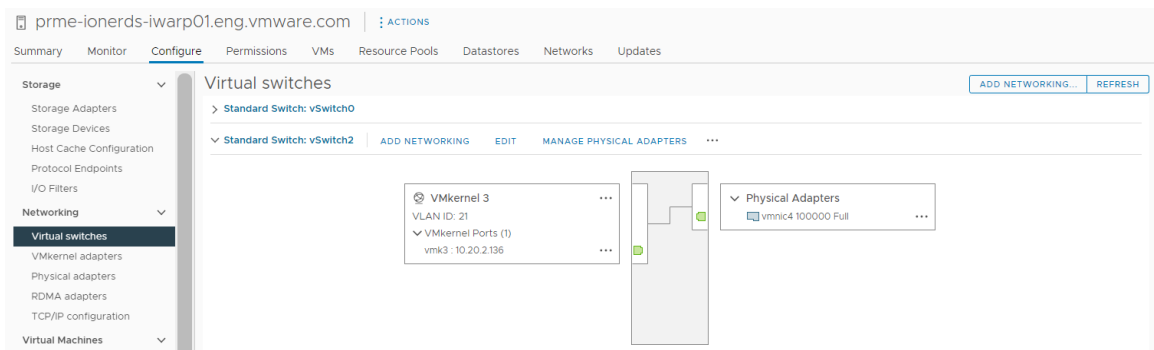
### 步骤

- 1 创建具有 VMkernel 适配器和网络组件的 vSphere 标准交换机。
  - a 在 vSphere Client 中，选择您的主机，然后单击**网络**选项卡。
  - b 单击**操作 > 添加网络**。
  - c 选择 **VMkernel 网络适配器**，然后单击**下一步**。
  - d 选择**新建标准交换机**，然后单击**下一步**。
  - e 在**分配的适配器**下，单击 **+**。  
将显示可用物理适配器的列表。
  - f 选择所需的物理适配器 vmnic，然后单击**确定**。

**注** 确保选择与 TCP/IP 适配器对应的物理网络适配器。

- g 在 **VMkernel 端口设置**下，输入所需的值。  
如果要对存储路径使用 VLAN，请输入 VLAN ID。
- h 在 **IP 设置**列表中，输入 VMkernel IPv4 设置。
- i 在**可用服务**下，选择 **NVMe over TCP** 以进行相应的标记。
- 2 验证交换机是否已正确配置。
  - a 在**配置**选项卡上，选择**网络**下的**虚拟交换机**。
  - b 展开交换机并验证其配置。

下图显示物理网络适配器和 VMkernel 适配器连接到了 vSphere 标准交换机。通过此连接，TCP 适配器将绑定到 VMkernel 适配器。



## 使用 vSphere 标准交换机和网卡绑定为 TCP 适配器配置 VMkernel 绑定

可以使用具有网卡绑定配置的 vSphere 标准交换机为 TCP 适配器配置 VMkernel 端口绑定。您可以使用网卡绑定来实现网络冗余。您可以将两个或多个网络适配器（网卡）配置为一个组来实现高可用性。

### 步骤

#### 1 创建具有 VMkernel 适配器和采用网卡绑定配置的网络组件的 vSphere 标准交换机。

- a 在 vSphere Client 中，选择您的主机，然后单击**网络**选项卡。
  - b 单击**操作 > 添加网络**。
  - c 选择 **VMkernel 网络适配器**，然后单击**下一步**。
  - d 选择**新建标准交换机**，然后单击**下一步**。
  - e 在**分配的适配器**下，单击 **+**。  
将显示可用物理适配器的列表。
  - f 选择所需的物理适配器 vmnic，然后将其添加到**活动适配器**下。
  - g 选择另一个物理适配器 vmnic，然后将其添加到**未用的适配器**下。
  - h 在 **VMkernel 端口设置**下，输入所需的值。  
如果要对存储路径使用 VLAN，请输入 VLAN ID。
  - i 在 **IP 设置**列表中，指定 VMkernel IPv4 设置。
  - j 在**可用服务**下，选择 **NVMe over TCP** 以进行相应的标记。
- 重复步骤 1 以配置现有标准交换机。

#### 2 为交换机配置网卡绑定配置。

- a 单击**配置**选项卡，然后选择**网络**下的**虚拟交换机**。
- b 选择相应的 VMkernel 适配器。
- c 从右键单击菜单中，单击**编辑设置**。
- d 选择**绑定和故障切换**。
- e 在**活动适配器**下，移动所需物理适配器 vmnic。
- f 在**备用适配器 > 故障切换顺序**下，移动其他物理适配器。
- g 设置合适的负载平衡和其他属性。
- h 重复上述步骤以配置其他 VMkernel 适配器。

#### 3 重复步骤 1 和 2，添加并配置一组额外的绑定 rnic。要验证适配器是否已配置，请单击**配置**选项卡，然后选择 **VMkernel 适配器**

## 为具有 vSphere Distributed Switch 的 TCP 适配器配置 VMkernel 绑定

可以使用 vSphere Distributed Switch 按每个交换机分配一个上行链路的方式为 TCP 适配器配置 VMkernel 端口绑定。配置网络连接包括为各个物理网络适配器创建虚拟 VMkernel 适配器。可以在每个虚拟和物理网络适配器之间使用 1:1 映射。

### 步骤

#### 1 创建具有 VMkernel 适配器和网络组件的 vSphere Distributed Switch。

- a 在 vSphere Client 中，选择**数据中心**，然后单击**网络**选项卡。
- b 单击**操作**，然后选择 **Distributed Switch > 新建 Distributed Switch**。
- c 选择交换机的名称。  
确保主机中存在数据中心的位置，然后单击**下一步**。
- d 选择兼容的 ESXi 版本，然后单击**下一步**。
- e 输入所需的上行链路数，然后单击**完成**。

#### 2 将一个或多个主机添加到分布式虚拟交换机。

- a 在 vSphere Client 中，选择**数据中心**，然后单击 **Distributed Switch**。  
此时将显示可用 DSwitch 的列表。
- b 右键单击 DSwitch，然后从菜单中选择**添加和管理主机**。
- c 选择**添加主机**，然后单击**下一步**。
- d 选择您的主机，然后单击**下一步**。
- e 选择**分配上行链路**。
- f 输入相关上行链路以分配 vmnic。
- g 分配 VMkernel 适配器，然后单击**下一步**。
- h 在 vSphere Client 中，选择 DSwitch，然后单击**端口**选项卡。  
您可以在此处查看为交换机创建的上行链路。

#### 3 为 NVMe over RDMA 存储路径创建分布式端口组。

- a 在 vSphere Client 中，选择所需的 DSwitch。
- b 单击**操作**，然后选择**分布式端口组 > 新建分布式端口组**。
- c 在**配置设置**下，输入端口组的常规属性。  
如果配置了特定的 VLAN，请将其添加到 VLAN ID 中。

---

**注** 如果未正确配置 VLAN，则可能会出现网络连接问题。

---

#### 4 配置 VMkernel 适配器。

- a 在 vSphere Client 中，展开 **DSwitch** 列表，然后选择分布式端口组。
- b 单击**操作 > 添加 VMkernel 适配器**。
- c 在**选择成员主机**对话框中，选择您的主机，然后单击**确定**。
- d 在**配置 VMkernel 适配器**对话框中，确保 MTU 与交换机 MTU 匹配。
- e 单击**完成**。
- f 重复步骤 b 和步骤 c，添加多个支持 RDMA 的网卡。

#### 5 为分布式端口组设置网卡绑定策略。

- a 在分布式端口组中，单击**操作 > 编辑设置**。
- b 单击**绑定和故障切换**，然后验证活动上行链路。
- c 将一个上行链路分配为端口组的**活动**上行链路，将另一个上行链路分配为**未使用**。

对创建的每个端口组重复步骤 c。

#### 后续步骤

完成配置后，单击**配置**，然后验证主机上的物理适配器选项卡是否列出所选网卡的 DVSwitch。

## 启用 NVMe over RDMA 或 NVMe over TCP 软件适配器

ESXi 支持 NVMe over RDMA 和 NVMe over TCP 软件适配器。可以使用 vSphere Client 启用 NVMe over RDMA 和 NVMe over TCP 软件存储适配器。

#### 前提条件

- 在 ESXi 主机上，安装支持以下存储类型的适配器。
  - NVMe over RDMA 适配器。例如，Mellanox Technologies MT27700 Family ConnectX-4。
  - NVMe over TCP 适配器。例如，i40en。
- 为适配器配置 VMkernel 绑定。
  - 对于 NVMe over RDMA，请参见为 [RDMA 适配器配置 VMkernel 绑定](#)。
  - 对于 NVMe over TCP，请参见为 [NVMe over TCP 适配器配置 VMkernel 绑定](#)。

#### 步骤

- 1 在 vSphere Client 中，导航到 ESXi 主机。
- 2 单击**配置**选项卡。
- 3 在**存储**下，依次单击**存储适配器**和**添加软件适配器**图标。
- 4 根据需要选择适配器类型。
  - **NVMe over RDMA 适配器**

## ■ NVMe over TCP 适配器

- 5 根据第 4 步中的选择，从下拉菜单中选择相应的 RDMA 适配器或 TCP 网络适配器 (vmnic)。

**注** 如果收到错误消息阻止您创建软件 NVMe over RDMA 适配器，请确保已正确配置 RDMA 适配器的 VMkernel 绑定。有关详细信息，请参见[为 RDMA 适配器配置 VMkernel 绑定](#)。

### 结果

NVMe over RDMA 和 NVMe over TCP 软件适配器将作为 vmhba 存储适配器显示在列表中。如果需要释放底层 RDMA 和 TCP 网络适配器用于其他目的，则可以移除这些适配器。

## 为 NVMe over Fabrics 添加控制器

使用 vSphere Client 添加 NVMe 控制器。添加控制器后，与该控制器关联的 NVMe 命名空间将可用于您的 ESXi 主机。表示 ESXi 环境中命名空间的 NVMe 存储设备将显示在存储设备列表中。

如果使用的是 NVMe over RDMA (RoCE v2) 存储，则必须在配置软件 NVMe over RDMA 适配器后添加控制器。如果使用的是 NVMe over TCP 存储，则必须在配置软件 NVMe over TCP 适配器后添加控制器。对于 FC-NVMe 存储，在安装所需的适配器后，它会自动连接到当前可访问的所有目标。您可以稍后重新配置适配器并断开其控制器连接，或者连接在主机引导期间不可用的其他控制器。

### 前提条件

确保 ESXi 主机已具有适用于您的存储类型的相应适配器。请参见 [VMware NVMe 存储的要求和限制](#)。

### 步骤

- 1 在 vSphere Client 中，导航到 ESXi 主机。
- 2 单击**配置**选项卡。
- 3 在**存储**下，单击**存储适配器**，然后选择要配置的适配器 (vmhba#)。
- 4 单击**控制器**选项卡，然后单击**添加控制器**。

5 要添加控制器，请选择以下选项之一，然后单击**添加**。

选项	描述
自动发现控制器	<p>此方法表示您的主机可以接受与任何可用控制器的连接。</p> <ol style="list-style-type: none"> <li>为发现控制器指定以下参数。 <ul style="list-style-type: none"> <li>对于 NVMe over RDMA (RoCE v2)，指定 IP 地址和传输端口号。</li> <li>对于 NVMe over TCP，指定 IP 地址、传输端口号和摘要参数。</li> </ul> </li> <li>单击<b>发现控制器</b>。</li> <li>从控制器列表中，选择要使用的控制器。</li> </ol>
手动输入控制器详细信息	<p>使用此方法，主机将使用以下参数请求连接到特定控制器：</p> <ul style="list-style-type: none"> <li>子系统 NQN</li> <li>目标端口标识。对于 NVMe over RDMA (RoCE v2)，指定 IP 地址和传输端口号（可选）。对于 FC-NVMe，指定 WorldWideNodeName 和 WorldWidePortName。</li> <li>对于 NVMe over TCP，指定 IP 地址、传输端口号（可选）和摘要参数（可选）。</li> <li>管理队列大小。指定控制器的管理队列大小的可选参数。默认值为 16。</li> <li>保持活动超时。一个可选参数，用于指定适配器与控制器的保持活动状态超时（以秒为单位）。默认超时值为 60 秒。</li> <li>IO 队列大小和 IO 队列编号。只能通过 esxcli 设置的可选参数。</li> </ul>

## 结果

控制器将显示在控制器列表中。现在，您的主机可以发现与该控制器关联的 NVMe 命名空间。表示 ESXi 环境中命名空间的 NVMe 存储设备将显示在 vSphere Client 中的存储设备列表上。

## 移除 NVMe over RDMA 或 NVMe over TCP 软件适配器

使用 vSphere Client 移除 NVMe over RDMA 或 NVMe over TCP 软件适配器。如果出于其他目的需要释放底层 RDMA 适配器或以太网适配器，可以移除该适配器。

无法移除 NVMe over PCIe 和 FC-NVMe 适配器。

## 步骤

- 在 vSphere Client 中，导航到 ESXi 主机。
- 单击**配置**选项卡。
- 在**存储**下，单击**存储适配器**，然后选择要移除的适配器 (vmhba#)。
- 移除连接到适配器的 NVMe 控制器。
  - 单击**控制器**选项卡。
  - 选择控制器，然后单击**移除**。  
NVMe 控制器现在将断开连接并从列表中消失。
- 单击**移除**图标（移除主机的存储适配器）以移除 NVMe over RDMA 适配器或 NVMe over TCP 适配器。



# 使用数据存储

## 9

数据存储是一种逻辑容器，类似于文件系统，它将物理存储的特性隐藏起来，提供一个统一模型来存储虚拟机文件。数据存储还可以用来存储 ISO 映像、虚拟机模板和软盘映像。

vCenter Server 和 ESXi 支持以下类型的数据存储。

表 9-1. 数据存储的类型

数据存储类型	描述
VMFS（版本 5 和 6）	块存储设备上部署的数据存储使用 vSphere 虚拟机文件系统 (Virtual Machine File System, VMFS) 格式。VMFS 是一种针对存储虚拟机而优化的特殊高性能文件系统格式。请参见 <a href="#">了解 VMFS 数据存储</a> 。
NFS（版本 3 和 4.1）	ESXi 中内置的 NFS 客户端使用网络文件系统 (NFS) 协议通过 TCP/IP 访问指定 NFS 卷。卷位于 NAS 服务器中。ESXi 主机将卷作为 NFS 数据存储挂载，并将其用于存储需求。ESXi 支持 3 和 4.1 版本的 NFS 协议。请参见 <a href="#">了解网络文件系统数据存储</a> 。
vSAN	vSAN 会将主机上所有可用的本地容量设备聚合到由 vSAN 集群中的所有主机共享的单个数据存储中。请参见《管理 VMware vSAN》文档。
vVol	Virtual Volumes 数据存储表示 vCenter Server 和 vSphere Client 中的存储容器。请参见 <a href="#">第 14 章 使用 VMware vSphere Virtual Volumes</a> 。

根据您的存储类型，以下某些任务可用于数据存储。

- 创建数据存储。可以使用 vSphere Client 创建某些类型的数据存储。
- 在数据存储上执行管理操作。某些操作（如重命名数据存储）适用于所有类型的数据存储。其他操作适用于特定类型的数据存储。
- 组织数据存储。例如，可以根据业务实践将其分组到文件夹。对数据存储分组后，可以一次性在组中的数据存储上分配相同的权限和警报。
- 将数据存储添加到数据存储集群。数据存储集群是具有共享资源和共享管理接口的数据存储的集合。创建数据存储集群后，可以使用 Storage DRS 管理存储资源。有关数据存储集群的信息，请参见《vSphere 资源管理》文档。

本章讨论了以下主题：

- 了解 VMFS 数据存储
- VMFS 锁定机制
- 了解网络文件系统数据存储
- 创建数据存储
- 管理重复 VMFS 数据存储
- 增加 VMFS 数据存储容量
- 在 VMFS6 数据存储上启用或禁用对集群虚拟磁盘的支持
- 数据存储的管理操作
- 设置动态磁盘镜像
- 在 VMFS 数据存储上收集 ESXi 主机的诊断信息
- 使用 VOMA 检查元数据一致性
- 配置 VMFS 指针块缓存

## 了解 VMFS 数据存储

要存储虚拟磁盘，ESXi 可使用数据存储。数据存储是逻辑容器，它对虚拟机隐藏物理存储的特性，并提供一个统一的模型来存储虚拟机文件。块存储设备上部署的数据存储使用本机 vSphere 虚拟机文件系统 (Virtual Machine File System, VMFS) 格式。该格式是一种针对存储虚拟机而优化的特殊高性能文件系统格式。

使用 vSphere Client，可以在 ESXi 主机发现的基于块的存储设备上预先设置 VMFS 数据存储。VMFS 数据存储可扩展为包括 SAN LUN 和本地存储的多个物理存储设备。使用此功能可以将存储放在存储池中，并灵活地创建虚拟机所需的数据存储。

当虚拟机在数据存储上运行时，您可以增加数据存储的容量。这一功能可让您在虚拟机需要新空间时为 VMFS 数据存储添加新空间。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相应的访问控制。

## VMFS 数据存储的版本

VMFS 文件系统自推出以来已发行多个版本。目前，ESXi 支持 VMFS5 和 VMFS6。

ESXi 对所有受支持的 VMFS 版本提供完全读取和写入支持。您可以在受支持的 VMFS 数据存储上创建虚拟机，并打开其电源。

表 9-2. 主机对 VMFS 版本的访问

VMFS	ESXi
VMFS6	读取和写入
VMFS5	读取和写入

下表比较了 VMFS5 和 VMFS6 的主要特性。有关其他信息，请参见《最高配置》。

表 9-3. 比较 VMFS5 和 VMFS6

特性和功能	VMFS5	VMFS6
对 ESXi 主机 6.5 和更高版本的访问	是	是
对 ESXi 主机 6.0 和更早版本的访问	是	否
每个主机的数据存储	512	512
512n 存储设备	是	是（默认）
512e 存储设备	是。本地 512e 设备上不支持。	是（默认）
4Kn 存储设备	否	是
自动空间回收	否	是
通过 <code>esxcli</code> 命令进行手动空间回收。请参见 <a href="#">手动回收 VMFS5 上累积的存储空间</a> 。	是	是
从客户机操作系统回收空间	受限制	是
GPT 存储设备分区	是	是
MBR 存储设备分区	是 适用于之前从 VMFS3 升级的 VMFS5 数据存储。	否
每个 VMFS 数据区的存储设备大于 2 TB	是	是
支持具有大容量虚拟磁盘或磁盘大于 2 TB 的虚拟机	是	是
支持 1 KB 的小文件	是	是
支持 ATS 的存储设备默认使用“仅限 ATS”锁定机制。请参见 <a href="#">VMFS 锁定机制</a> 。	是	是
块大小	标准 1 MB	标准 1 MB
默认快照	虚拟磁盘的 VMFSsparse 小于 2 TB。 虚拟磁盘的 SEsparse 大于 2 TB。	SEsparse
虚拟磁盘模拟类型	512n	512n
vMotion	是	是
跨不同数据存储类型的 Storage vMotion	是	是
High Availability 和 Fault Tolerance	是	是
DRS 和 Storage DRS	是	是
RDM	是	是

使用 VMFS 数据存储时，请注意以下内容：

- 数据存储数据区。跨区的 VMFS 数据存储必须仅使用同类存储设备，512n、512e 或 4Kn。跨区的数据存储无法在不同格式的设备之间扩展。
- 块大小。VMFS 数据存储上的块大小定义最大文件大小和文件占用的空间量。VMFS5 和 VMFS6 数据存储支持 1 MB 的块大小。
- Storage vMotion。Storage vMotion 支持跨 VMFS、vSAN 和 Virtual Volumes 数据存储执行迁移。vCenter Server 执行兼容性检查以验证跨不同类型的数据存储的 Storage vMotion。
- Storage DRS。VMFS5 和 VMFS6 可以在同一数据存储集群中共存。然而，集群中的所有数据存储必须使用同类存储设备。请勿在同一数据存储集群中混用不同格式的设备。
- 设备分区格式。任何新 VMFS5 或 VMFS6 数据存储都使用 GUID 分区表 (GUID Partition Table, GPT) 来格式化存储设备。利用 GPT 格式，可以创建大于 2 TB 的数据存储。如果 VMFS5 数据存储之前是从 VMFS3 升级而来，则它继续使用主引导记录 (MBR) 分区格式，这种格式是 VMFS3 的特征。仅当将数据存储扩展至大于 2TB 之后，才会转换为 GPT。

## 作为存储库的 VMFS 数据存储

ESXi 可以将基于 SCSI 的存储设备格式化为 VMFS 数据存储。VMFS 数据存储主要充当虚拟机的存储库。

---

**注** 每个 LUN 始终只具有一个 VMFS 数据存储。

---

可以在同一个 VMFS 数据存储上存储多个虚拟机。封装在一组文件中的各个虚拟机都会占用单独的一个目录。对于虚拟机内的操作系统，VMFS 会保留内部文件系统语义，这样可以确保正确的应用程序行为以及在虚拟机中运行的应用程序的数据完整性。

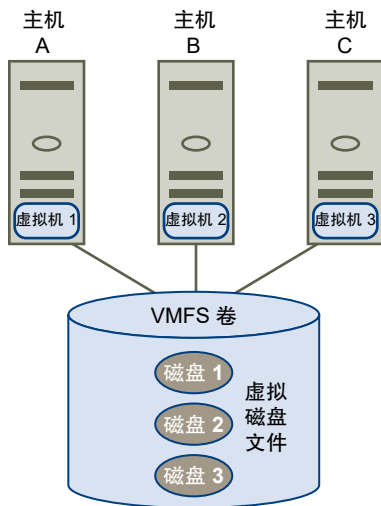
当运行多个虚拟机时，VMFS 针对虚拟机文件提供特定的锁定机制。因此，在多个 ESXi 主机共享同一个 VMFS 数据存储的 SAN 环境中，虚拟机可以安全地操作。

除了虚拟机之外，VMFS 数据存储也可以存储其他文件，如虚拟机模板和 ISO 映像。

## 在主机间共享 VMFS 数据存储

作为一个集群文件系统，VMFS 允许多个 ESXi 主机同时访问同一个 VMFS 数据存储。

图 9-1. 在主机间共享 VMFS 数据存储



有关可连接到单个 VMFS 数据存储的最大主机数目的信息，请参见《最高配置》文档。

为了确保多个主机不会同时访问同一个虚拟机，VMFS 提供了磁盘锁定。

在多个主机之间共享 VMFS 卷可提供一些优势，例如以下优势：

- 可以使用 VMware Distributed Resource Scheduling (DRS) 和 VMware High Availability (HA)。

可以跨越不同的物理服务器分配虚拟机。这意味着，每个服务器上会运行一组虚拟机，这样一来，所有服务器就不会同时在同一个区域面临很高的需求。如果某台服务器发生故障，可以在另一台物理服务器上重新启动虚拟机。如果发生故障，每个虚拟机的磁盘锁定会被释放。有关 VMware DRS 的详细信息，请参见《vSphere 资源管理》文档。有关 VMware HA 的信息，请参见《vSphere 可用性》文档。

- 可以使用 vMotion 将正在运行的虚拟机从一台物理服务器移动到另一台物理服务器。有关迁移虚拟机的信息，请参见《vCenter Server 和主机管理》文档。

要创建共享数据存储，可将数据存储挂载到要求数据存储访问的 ESXi 主机。请参见 [挂载数据存储](#)。

## VMFS 元数据更新

VMFS 数据存储保存虚拟机文件、目录、符号链接、RDM 描述符文件等。数据存储还维护关于这些对象的所有映射信息的一致视图。这些映射信息称为元数据。

每次执行数据存储或虚拟机管理操作时，都会更新元数据。需要更新元数据的操作示例包括：

- 创建、增加或锁定虚拟机文件
- 更改文件的属性
- 打开或关闭虚拟机电源
- 创建或删除 VMFS 数据存储

- 展开 VMFS 数据存储
- 创建模板
- 从模板部署虚拟机
- 通过 vMotion 迁移虚拟机

在共享存储环境中执行元数据更改时，VMFS 使用特定的锁定机制保护其数据，并阻止多个主机同时向元数据执行写入操作。

## VMFS 上的快照格式

生成快照时，将保留虚拟磁盘的状况，从而阻止客户机操作系统在该磁盘写入，并会创建增量磁盘或子磁盘。增量表示虚拟机磁盘的当前状况和上次生成快照时存在的状况之间的差异。在 VMFS 数据存储中，增量磁盘为稀疏磁盘。

稀疏磁盘使用写入时复制机制，在这种机制下，虚拟磁盘不包含任何数据，直到通过写入操作将数据复制到磁盘。此优化可节省存储空间。

根据数据存储的类型，增量磁盘使用不同的稀疏格式。

快照格式	VMFS5	VMFS6
VMFSsparse	适用于小于 2 TB 的虚拟磁盘。	不适用
SEsparse	适用于大于 2 TB 的虚拟磁盘。	适用于所有磁盘。

### VMFSsparse

对于小于 2 TB 的虚拟磁盘，VMFS5 使用 VMFSsparse 格式。

VMFSsparse 在 VMFS 上实现。VMFSsparse 层处理向快照虚拟机发出的 I/O。从技术方面来讲，VMFSsparse 是一个重做日志，在刚生成虚拟机快照后最初是空的。在生成虚拟机快照后对整个 vmdk 重新写入新数据时，该重做日志将扩展至其基础 vmdk 的大小。该重做日志是 VMFS 数据存储中的一个文件。在创建快照之后，连接到虚拟机的基础 vmdk 会更改为新创建的稀疏 vmdk。

### SEsparse

SEsparse 是 VMFS6 数据存储中所有增量磁盘的默认格式。在 VMFS5 中，SEsparse 用于大小为 2 TB 以及更大的虚拟磁盘。

SEsparse 格式与 VMFSsparse 类似，但具有一些增强功能。这种格式可以节省空间，并支持空间回收技术。通过空间回收，可标记客户机操作系统删除的块。系统将命令发送到管理程序中的 SEsparse 层以对这些块取消映射。取消映射可帮助在客户机操作系统删除该数据后回收 SEsparse 分配的空间。有关空间回收的详细信息，请参见[存储空间回收](#)。

## 快照迁移

利用快照可在不同的数据存储之间迁移虚拟机。下列注意事项适用：

- 如果利用 VMFSsparse 快照将虚拟机迁移到 VMFS6，则快照格式会更改为 SEsparse。
- 在将 vmdk 小于 2 TB 的虚拟机迁移到 VMFS5 时，快照格式会更改为 VMFSsparse。

- 在同一层次结构中，不能混用 SEsparse 重做日志和 VMFSsparse 重做日志。

## 升级 VMFS 数据存储

您无法将 VMFS5 数据存储升级到 VMFS6。如果您的环境是 VMFS5 数据存储，请创建 VMFS6 数据存储，并将虚拟机从 VMFS5 数据存储迁移到 VMFS6。

## VMFS 锁定机制

在共享存储环境中，当多台主机访问同一 VMFS 数据存储时，将使用特定的锁定机制。这些锁定机制可防止多台主机同时写入元数据并确保不会发生数据损坏。

根据 VMFS 数据存储的配置和底层存储的类型，VMFS 数据存储可能会使用不同类型的锁定机制。它可能会以独占方式使用原子测试并设置锁定机制（仅限 ATS），也可能会结合使用 ATS 和 SCSI 预留 (ATS+SCSI)。

### “仅限 ATS” 机制

对于支持基于 T10 标准的 VAAI 规格的存储设备，VMFS 可提供 ATS 锁定，也称为硬件辅助锁定。ATS 算法支持按磁盘扇区进行离散锁定。如果基础存储支持，则所有新格式化的 VMFS5 和 VMFS6 数据存储都将使用“仅限 ATS”机制，绝不会使用 SCSI 预留。

在创建使用 ATS 的多数据区的数据存储时，vCenter Server 会筛选出非 ATS 设备。通过此筛选，您可以仅使用支持 ATS 原语的设备。

在某些情况下，您可能需要更改 VMFS5 或 VMFS6 数据存储的默认锁定机制。有关信息，请参见[将锁定机制更改为 ATS+SCSI](#)。

---

**注** 如果运行的是 VMware vSAN 环境或具有仅限 ATS 的 VMFS 卷，请勿取消激活 ATS。取消激活 ATS 可能会导致中断，因为没有可用的锁定机制。有关详细信息，请参见 [VMware 知识库文章](#)。

---

### ATS+SCSI 机制

支持 ATS+SCSI 机制的 VMFS 数据存储将配置为使用 ATS 并尽可能尝试使用该机制。如果 ATS 失败，则 VMFS 数据存储将恢复为 SCSI 预留。与 ATS 锁定不同，SCSI 预留在执行需要元数据保护的操作时锁定整个存储设备。操作完成后，VMFS 会释放预留，并且可继续执行其他操作。

使用 ATS+SCSI 机制的数据存储包括从 VMFS3 升级的 VMFS5 数据存储。此外，存储设备上不支持 ATS 的新 VMFS5 或 VMFS6 数据存储将使用 ATS+SCSI 机制。

如果 VMFS 数据存储恢复为 SCSI 预留，则您可能会注意到因过多 SCSI 预留而导致的性能降低。

## 显示 VMFS 锁定信息

使用 `esxcli` 命令可获取有关 VMFS 数据存储使用的锁定机制的信息。

### 前提条件

安装 ESXCLI。请参见《ESXCLI 入门》。要进行故障排除，请在 ESXi Shell 中运行 `esxcli` 命令。

步骤

- ◆ 要显示与 VMFS 锁定机制相关的信息，请运行以下命令：

```
esxcli storage vmfs lockmode list
```

结果

下表列出了该命令的输出可能包括的项目。

表 9-4. VMFS 锁定信息

字段	值	描述
锁定模式		表示数据存储的锁定配置。
	ATS-only	数据存储配置为使用“仅限 ATS”锁定模式。
	ATS+SCSI	数据存储配置为使用 ATS 模式。如果 ATS 失败或不受支持，数据存储可以恢复至 SCSI。
	ATS upgrade pending	数据存储正在联机升级到“仅限 ATS”模式。
	ATS downgrade pending	数据存储正在联机降级到 ATS+SCSI 模式。
ATS 兼容		表示数据存储是否可以配置为“仅限 ATS”模式。
ATS 升级模式		表示数据存储支持的升级类型。
	None	数据存储与“仅限 ATS”不兼容。
	Online	数据存储在其升级到“仅限 ATS”模式的过程中可以使用。
	Offline	数据存储在其升级到“仅限 ATS”模式的过程中无法使用。
ATS 不兼容原因		如果数据存储与“仅限 ATS”不兼容，该项会指明不兼容的原因。

VMFS 的“仅限 ATS”锁定机制

如果 VMFS 数据存储使用 ATS+SCSI 锁定机制，则可以更改为“仅限 ATS”锁定。

通常，此前从 VMFS3 升级的 VMFS5 数据存储继续使用 ATS+SCSI 锁定机制。如果在启用 ATS 的硬件上部署数据存储，则这些数据存储符合升级到“仅限 ATS”锁定的条件。根据您的 vSphere 环境，可以使用以下升级模式之一：

- 联机升级到“仅限 ATS”机制适合于大多数单数据区 VMFS5 数据存储。当您在其中一个主机上执行联机升级时，其他主机可以继续使用数据存储。
- 脱机升级到“仅限 ATS”必须用于跨越多个物理数据区的 VMFS5 数据存储。由多个数据区组成的数据存储不符合联机升级的条件。这些数据存储要求在收到升级请求时没有主机主动使用数据存储。



## 准备升级到“仅限 ATS”锁定

要为联机或脱机升级到“仅限 ATS”锁定做好环境准备，您必须执行多个步骤。

### 步骤

- 1 将访问 VMFS5 数据存储的所有主机升级到最新版本的 vSphere。
- 2 通过运行 `esxcli storage vmfs lockmode list` 命令来确定数据存储是否符合升级其当前锁定机制的条件。

以下示例输出表明数据存储符合升级条件，同时还显示当前锁定机制及适用于数据存储的升级模式。

Locking Mode	ATS Compatible	ATS Upgrade Modes
-----	-----	-----
ATS+SCSI	true	Online or Offline

- 3 根据适用于数据存储的升级模式，执行以下操作之一：

升级模式	操作
联机	验证所有主机与 VMFS 数据存储的存储连接是否一致。
脱机	验证所有主机是否均未使用数据存储。

## 将锁定机制升级到“仅限 ATS”类型

如果 VMFS 数据存储为“仅限 ATS”兼容，则可以将其锁定机制从 ATS+SCSI 升级至“仅限 ATS”。  
未跨越多个数据区的大多数数据存储符合联机升级的条件。在其中一个 ESXi 主机上执行联机升级时，其他主机可以继续使用数据存储。只有在所有主机已关闭数据存储后，才能完成联机升级。

### 前提条件

如果打算通过将数据存储置于维护模式来完成锁定机制升级，请禁用 Storage DRS。此必备条件仅适用于联机升级。

### 步骤

- 1 通过运行以下命令执行锁定机制升级：  
  
`esxcli storage vmfs lockmode set -a|--ats -l|--volume-label= VMFS label -u|--volume-uuid= VMFS UUID。`

## 2 对于联机升级，请执行额外的步骤。

- a 关闭对数据存储具有访问权限的所有主机上的数据存储，以便主机可以识别更改。

可使用下列方法之一：

- 卸载并挂载数据存储。
- 将数据存储置于维护模式并退出维护模式。

- b 要验证数据存储的锁定模式状态是否已更改为“仅限 ATS”，请运行以下命令：

```
esxcli storage vmfs lockmode list
```

- c 如果锁定模式显示任何其他状态（例如 ATS UPGRADE PENDING），请通过运行以下命令检查尚未处理升级的主机：

```
esxcli storage vmfs host list
```

## 将锁定机制更改为 ATS+SCSI

在支持原子测试和设置 (ATS) 锁定的设备上创建 VMFS5 数据存储时，该数据存储会使用“仅限 ATS”锁定机制。在某些情况下，可能需要将“仅限 ATS”锁定降级到 ATS+SCSI。

您可能需要切换到 ATS + SCSI 锁定机制，例如，在存储设备已降级，或者固件更新失败且设备不再支持 ATS 等情况下。

降级过程与“仅限 ATS”升级过程很相似。与升级一样，根据您的存储配置，可以在联机或脱机模式下执行降级。

---

**注** 如果运行的是 VMware vSAN 环境或具有仅限 ATS 的 VMFS 卷，请勿取消激活 ATS。取消激活 ATS 可能会导致中断，因为没有可用的锁定机制。有关详细信息，请参见 [VMware 知识库文章](#)。

---

### 步骤

- 1 通过运行以下命令将锁定机制更改为 ATS+SCSI：

```
esxcli storage vmfs lockmode set -s|--scsi -l|--volume-label= VMFS label
-u|--volume-uuid= VMFS UUID。
```

- 2 对于联机模式，在对数据存储具有访问权限的所有主机上关闭数据存储，以便主机可以识别更改。

## 了解网络文件系统数据存储

ESXi 中内置的 NFS 客户端使用网络文件系统 (NFS) 协议通过 TCP/IP 访问位于 NAS 服务器上的指定 NFS 卷。ESXi 主机可以挂载该卷，并用该卷来满足其存储需求。vSphere 支持 NFS 协议版本 3 和 4.1。

通常，NFS 卷或目录由存储管理员创建并从 NFS 服务器中导出。您无需使用本地文件系统（如 VMFS）将 NFS 卷格式化。相反，您可以直接在 ESXi 主机上挂载卷，并将其用于存储和引导虚拟机，使用方式与 VMFS 数据存储相同。

除了在 NFS 数据存储上存储虚拟磁盘以外，还可以使用 NFS 作为 ISO 映像、虚拟机模板等的中央存储库。如果使用适用于 ISO 映像的数据存储，则可将虚拟机的 CD-ROM 设备连接到数据存储上的 ISO 文件。然后，可从 ISO 文件安装客户机操作系统。

ESXi 支持 NFS 协议版本 3 和 4.1。为了同时支持这个版本，ESXi 使用了两个不同的 NFS 客户端。

## NFS 客户端各版本的比较

下表列出了 NFS 版本 3 和 4.1 支持的功能。

特性	NFS 版本 3	NFS 版本 4.1
安全机制	AUTH_SYS	AUTH_SYS 和 Kerberos (krb5 和 krb5i)
使用 Kerberos 的加密算法	不适用	AES256-CTS-HMAC-SHA1-96 和 AES128-CTS-HMAC-SHA1-96
多路径	不受支持	通过会话中继支持
锁定机制	专有客户端锁定	服务器端锁定
硬件加速	支持	支持
厚虚拟磁盘	支持	支持
IPv6	支持	AUTH_SYS 和 Kerberos 支持
对于虚拟机显示为 CD-ROM 的 ISO 映像	支持	支持
虚拟机快照	支持	支持
虚拟磁盘大于 2 TB 的虚拟机	支持	支持

## NFS 协议和 vSphere 解决方案

下表列出了 NFS 版本支持的主要 vSphere 解决方案。

vSphere 功能	NFS 版本 3	NFS 版本 4.1
vMotion 和 Storage vMotion	是	是
High Availability (HA)	是	是
Fault Tolerance (FT)	是	是
Distributed Resource Scheduler, DRS	是	是
主机配置文件	是	是
Storage DRS	是	否
Storage I/O Control	是	否
Site Recovery Manager	是	Site Recovery Manager 不支持 NFS 4.1 数据存储进行基于阵列的复制和 Virtual Volumes 复制。您可以将 Site Recovery Manager 与 NFS v 4.1 数据存储一起用于 vSphere Replication。
Virtual Volumes	是	是

vSphere 功能	NFS 版本 3	NFS 版本 4.1
vSphere Replication	是	是
vRealize Operations Manager	是	是

## NFS 4.1 和 Fault Tolerance

NFS 4.1 上的虚拟机支持 vSphere 6.0 中引入的新 Fault Tolerance 机制。该机制可容纳最多具有四个 vCPU 的对称多处理器 (SMP) 虚拟机。

NFS 4.1 虚拟机不支持旧版 Fault Tolerance 机制。

## NFS 升级

从低于 6.5 的版本升级 ESXi 时，现有 NFS 4.1 数据存储会自动开始支持之前 ESXi 版本中不可用的功能。这些功能包括 Virtual Volumes、硬件加速等。

ESXi 不支持从 NFS 版本 3 到 NFS 4.1 的自动数据存储转换。

如果要升级 NFS 3 数据存储，可以使用以下选项：

- 创建 NFS 4.1 数据存储，然后使用 Storage vMotion 将虚拟机从旧数据存储迁移到新数据存储。
- 使用 NFS 存储服务器提供的转换方法。有关详细信息，请与存储供应商联系。
- 卸载 NFS 3 数据存储，然后作为 NFS 4.1 存储进行挂载。

**小心** 如果使用此选项，请确保从有权访问该数据存储的所有主机中卸载数据存储。无法同时使用上述两种协议挂载数据存储。

## NFS 存储准则和要求

在使用 NFS 存储时，请遵循与 NFS 服务器配置、网络连接、NFS 数据存储等相关的特定准则。

### ■ NFS 服务器配置

在将 NFS 服务器配置为与 ESXi 结合使用时，请遵循存储供应商的建议。除了这些一般性建议，还请使用适用于 vSphere 环境中 NFS 的特定指导准则。

### ■ NFS 网络

ESXi 主机使用 TCP/IP 网络连接来访问远程 NAS 服务器。使用 NFS 存储配置网络时，需遵循一定的准则和最佳实践。

### ■ NFS 文件锁定

使用文件锁定机制，可以将对存储在服务器上的数据的访问限制为每次仅一个用户或一个进程。两个 NFS 版本的锁定机制不兼容。NFS 3 使用专用锁定，NFS 4.1 则使用本机协议指定的锁定。

### ■ NFS 安全

借助 NFS 3 和 NFS 4.1，ESXi 可支持 AUTH\_SYS 安全。此外，对于 NFS 4.1，还支持 Kerberos 安全机制。

## ■ NFS 多路径

根据协议规范，NFS 4.1 支持多路径。对于 NFS 3，多路径并不适用。

## ■ NFS 和硬件加速

默认情况下，在 NFS 数据存储上创建的虚拟磁盘为精简置备的虚拟磁盘。为了能够创建厚置备虚拟磁盘，必须使用支持“预留空间”操作的硬件加速。

## ■ NFS 数据存储

在创建 NFS 数据存储时，请务必遵循特定的准则。

# NFS 服务器配置

在将 NFS 服务器配置为与 ESXi 结合使用时，请遵循存储供应商的建议。除了这些一般性建议，还请使用适用于 vSphere 环境中 NFS 的特定指导准则。

准则包括以下项。

- 确保使用 VMware HCL 中列出的 NAS 服务器。使用正确的服务器固件版本。
- 确保使用 NFS 通过 TCP 导出 NFS 卷。
- 确保 NAS 服务器将特定共享导出为 NFS 3 或 NFS 4.1。NAS 服务器不得为同一共享提供两种协议版本。由于 ESXi 不会阻止通过不同 NFS 版本挂载相同的共享，因此 NAS 服务器必须强制执行此策略。
- NFS 3 和非 Kerberos (AUTH\_SYS) NFS 4.1 不支持借助于非 root 凭据启用对 NFS 卷访问权的委派用户功能。如果使用 NFS 3 或非 Kerberos NFS 4.1，请确保每台主机对卷具有根访问权限。不同的存储供应商采用不同的方法启用此功能，但通常 NAS 服务器使用 no\_root\_squash 选项。如果 NAS 服务器未授予根访问权限，您仍然可以在主机上挂载 NFS 数据存储。但是，无法在数据存储上创建任何虚拟机。
- 如果底层 NFS 卷是只读的，则应确保该卷由 NFS 服务器导出为只读共享，或在 ESXi 主机上将它挂载为只读数据存储。否则，主机将该数据存储视为可以读写，并可能无法打开文件。

# NFS 网络

ESXi 主机使用 TCP/IP 网络连接来访问远程 NAS 服务器。使用 NFS 存储配置网络时，需遵循一定的准则和最佳实践。

有关详细信息，请参见《《vSphere 网络连接》》文档。

- 对于网络连接，请在 ESXi 主机中使用标准网络适配器。
- ESXi 支持第 2 层和第 3 层网络交换机。如果使用第 3 层交换机，则 ESXi 主机和 NFS 存储阵列必须位于不同的子网上，并且网络交换机必须处理路由信息。
- 为 NFS 存储配置 VMkernel 端口组。您可以在现有虚拟交换机 (vSwitch) 或在新 vSwitch 上为 IP 存储创建 VMkernel 端口组。vSwitch 可以是 vSphere 标准交换机 (VSS) 或 vSphere Distributed Switch (VDS)。
- 如果为 NFS 流量使用多个端口，请确保正确配置虚拟交换机和物理交换机。
- NFS 3 和 NFS 4.1 支持 IPv6。

## NFS 文件锁定

使用文件锁定机制，可以将对存储在服务器上的数据的访问限制为每次仅一个用户或一个进程。两个 NFS 版本的锁定机制不兼容。NFS 3 使用专用锁定，NFS 4.1 则使用本机协议指定的锁定。

ESXi 上的 NFS 3 锁定不会使用网络锁定管理器 (NLM) 协议。相反，VMware 提供自带的锁定协议。通过在 NFS 服务器上创建锁定文件来实施 NFS 3 锁定。锁定文件的名称为 `.lck-file_id`。

NFS 4.1 使用共享预留作为锁定机制。

因为 NFS 3 和 NFS 4.1 客户端使用的锁定协议并不相同，因此您无法使用不同的 NFS 版本在多个主机上挂载相同的数据存储。从两个不兼容的客户端访问同一虚拟磁盘可能导致不正确的行为，并导致数据损坏。

## NFS 安全

借助 NFS 3 和 NFS 4.1，ESXi 可支持 AUTH\_SYS 安全。此外，对于 NFS 4.1，还支持 Kerberos 安全机制。

NFS 3 支持 AUTH\_SYS 安全机制。使用该机制时，存储流量将以未加密格式通过 LAN 进行传输。由于此安全性有限，请仅使用可信网络上的 NFS 存储并隔离不同物理交换机上的流量。您还可以使用专用 VLAN。

NFS 4.1 支持 Kerberos 身份验证协议以确保与 NFS 服务器的通信安全。使用 Kerberos 时，非 root 用户可以访问文件。有关详细信息，请参见[对 NFS 4.1 使用 Kerberos](#)。

除了 Kerberos，NFS 4.1 还支持采用 AUTH\_SYS 安全的传统非 Kerberos 挂载。在此情况下，对 NFS 版本 3 使用根访问准则。

---

**注** 不能对多个主机共享的同一个 NFS 4.1 数据存储使用两个安全机制：AUTH\_SYS 和 Kerberos。

---

## NFS 多路径

根据协议规范，NFS 4.1 支持多路径。对于 NFS 3，多路径并不适用。

NFS 3 将一个 TCP 连接用于 I/O。因此，ESXi 仅支持 NFS 服务器的一个 IP 地址或主机名上的 I/O，且不支持多路径。根据您的网络基础架构和配置，可以使用网络堆栈配置存储目标的多个连接。在此情况下，必须拥有多个数据存储，且每个数据存储在主机与存储之间使用单独的网络连接。

NFS 4.1 为支持会话中继的服务器提供多路径。中继可用时，可以使用多个 IP 地址访问单个 NFS 卷。不支持客户端 ID 中继。

## NFS 和硬件加速

默认情况下，在 NFS 数据存储上创建的虚拟磁盘为精简置备的虚拟磁盘。为了能够创建厚置备虚拟磁盘，必须使用支持“预留空间”操作的硬件加速。

NFS 3 和 NFS 4.1 支持硬件加速，从而允许主机与 NAS 设备相集成并使用 NAS 存储提供的多个硬件操作。有关详细信息，请参见[NAS 设备上的硬件加速](#)。

## NFS 数据存储

在创建 NFS 数据存储时，请务必遵循特定的准则。

NFS 数据存储的准则和最佳做法包括以下项：

- 不能使用不同的 NFS 版本在不同主机上挂载同一数据存储。NFS 3 和 NFS 4.1 客户端是不兼容的，并且使用不同的锁定协议。因此，从两个不兼容的客户端访问同一虚拟磁盘可能导致不正确的行为，并导致数据损坏。
- NFS 3 和 NFS 4.1 数据存储可以在同一主机上共存。
- ESXi 无法自动将 NFS 版本 3 自动升级至版本 4.1，但您可以使用其他转换方法。有关信息，请参见[了解网络文件系统数据存储](#)。
- 当在不同主机上挂载相同 NFS 3 卷时，确保各主机之间的服务器名称和文件夹名称相同。如果名称不匹配，则主机会将同一 NFS 版本 3 卷视为两个不同的数据存储。该错误可能导致诸如 vMotion 之类的功能运行失败。例如，如果在一台主机上输入 `filer` 作为服务器名称，而在另一台主机上输入 `filer.domain.com` 作为服务器名称，就出现了这种名称不匹配的情况。此准则不适用于 NFS 版本 4.1。
- 如果使用非 ASCII 字符命名数据存储和虚拟机，请确保基础 NFS 服务器提供了国际化支持。如果该服务器不支持国际字符，请仅使用 ASCII 字符，否则可能会出现不可预知的故障。

## NFS 存储的防火墙配置

ESXi 在管理接口和网络之间设有防火墙。该防火墙在默认情况下启用。安装时，ESXi 防火墙配置为阻止除 NFS 等默认服务的流量之外的入站和出站流量。

在 ESXi 防火墙目录 `/etc/vmware/firewall/` 中的规则集配置文件中对受支持的服务（包括 NFS）进行了说明。该文件包含防火墙规则以及它们与端口和协议的关系。

NFS 客户端规则集 (`nfsClient`) 的行为与其他规则集不同。

有关防火墙配置的详细信息，请参见《《vSphere 安全性》》文档。

### NFS 客户端防火墙行为

NFS 客户端防火墙规则集的行为方式与其他 ESXi 防火墙规则集不同。挂载或卸载 NFS 数据存储时，ESXi 将配置 NFS 客户端设置。对于不同版本的 NFS，行为有所不同。

添加、挂载或卸载 NFS 数据存储时，产生的行为取决于 NFS 版本。

#### NFS v3 防火墙行为

添加或挂载 NFS v3 数据存储时，ESXi 将检查 NFS 客户端 (`nfsClient`) 防火墙规则集的状态。

- 如果停用了 `nfsClient` 规则集，则 ESXi 将激活规则集，并通过将 `allowedAll` 标记设置为 `FALSE` 来停用“允许所有 IP 地址”策略。NFS 服务器的 IP 地址将会添加到允许的出站 IP 地址的列表中。



- 如果激活了 nfsClient 规则集，则规则集状态和允许的 IP 地址策略将不会更改。NFS 服务器的 IP 地址将会添加到允许的出站 IP 地址的列表中。

**注** 如果手动激活 nfsClient 规则集或手动设置“允许所有 IP 地址”策略，则将 NFS v3 数据存储添加到系统之前或之后，卸载最新 NFS v3 数据存储时将替代您的设置。卸载所有 NFS v3 数据存储时，将停用 nfsClient 规则集。

移除或卸载 NFS v3 数据存储时，ESXi 会执行以下操作之一。

- 如果未从已卸载数据存储的服务器挂载任何剩余的 NFS v3 数据存储，则 ESXi 将从出站 IP 地址列表中移除该服务器的 IP 地址。
- 如果执行卸载操作后没有剩余任何挂载的 NFS v3 数据存储，则 ESXi 将停用 nfsClient 防火墙规则集。

### NFS v4.1 防火墙行为

挂载第一个 NFS v4.1 数据存储时，ESXi 将激活 nfs41client 规则集并将其 allowedAll 标记设置为 TRUE。此操作将打开所有 IP 地址的端口 2049。卸载 NFS v4.1 数据存储不会影响防火墙状态。也就是说，第一个 NFS v4.1 挂载将打开端口 2049，除非明确关闭该端口，否则该端口将保持激活状态。

### 验证 NFS 客户端的防火墙端口

要启用对 NFS 存储的访问，则在挂载 NFS 数据存储时，ESXi 将自动打开 NFS 客户端的防火墙端口。出于故障排除原因，您可能需要验证相关端口是否已打开。

#### 步骤

- 1 在 vSphere Client 中，导航到 ESXi 主机。
- 2 单击**配置**选项卡。
- 3 在**系统**下，单击**防火墙**，然后单击**编辑**。
- 4 向下滚动到相应版本的 NFS 以确保端口已打开。

## 使用第 3 层路由连接访问 NFS 存储

使用第 3 层 (L3) 路由连接访问 NFS 存储时，请考虑某些要求和限制。

确保您的环境满足以下要求：

- 在 IP 路由器中使用 Cisco 的热备份路由器协议 (HSRP)。如果使用的是非 Cisco 路由器，请使用虚拟路由器冗余协议 (VRRP)。
- 要区分具有带宽限制的网络上 NFS L3 流量的优先级，或区分拥堵网络上的 NFS L3 流量的优先级，请使用服务质量 (QoS)。有关详细信息，请参见路由器文档。
- 遵循存储供应商提供的路由 NFS L3 建议。有关详细信息，请联系您的存储供应商。
- 禁用网络 I/O 资源管理 (NetIORM)。
- 如果计划使用具有机架置顶式交换机或独立于交换机的 I/O 设备分区的系统，请联系您的系统供应商获取兼容性和支持。



在 L3 环境中，以下限制适用：

- 该环境不支持 VMware Site Recovery Manager。
- 该环境仅支持 NFS 协议。在同一物理网络内不使用其他存储协议（例如 FCoE）。
- 该环境中的 NFS 流量不支持 IPv6。
- 该环境中的 NFS 流量仅可以通过 LAN 路由。不支持其他环境（例如 WAN）。

## 对 NFS 4.1 使用 Kerberos

使用 NFS 版本 4.1 时，ESXi 支持 Kerberos 身份验证机制。

RPCSEC\_GSS Kerberos 机制是一种身份验证服务。它允许 ESXi 上安装 NFS 4.1 客户端在挂载 NFS 共享之前向 NFS 服务器证明其身份。Kerberos 安全在不安全的网络连接中使用加密进行工作。

ESXi 针对 NFS 4.1 实施 Kerberos 可提供两种安全模型：krb5 和 krb5i，分别提供不同的安全级别。

- 仅用于身份验证的 Kerberos (krb5) 支持身份认证。
- 用于身份验证和数据完整性的 Kerberos (krb5i) 除了提供身份认证，还提供数据完整性服务。这些服务通过检查潜在的数据包修改操作，帮助保护 NFS 流量免受篡改。

Kerberos 支持加密算法，可防止未经授权的用户访问 NFS 流量。ESXi 上的 NFS 4.1 客户端尝试使用 AES256-CTS-HMAC-SHA1-96 或 AES128-CTS-HMAC-SHA1-96 算法访问 NAS 服务器上的共享。使用 NFS 4.1 数据存储之前，确保在 NAS 服务器上启用 AES256-CTS-HMAC-SHA1-96 或 AES128-CTS-HMAC-SHA1-96。

下表比较了 ESXi 支持的 Kerberos 安全级别。

**表 9-5. Kerberos 安全类型**

		ESXi 6.0	ESXi 6.5 及更高版本
仅用于身份验证的 Kerberos (krb5)	RPC 标头的完整性校验和	是，使用 DES	是，使用 AES
	RPC 数据的完整性校验和	否	否
用于身份验证和数据完整性的 Kerberos (krb5i)	RPC 标头的完整性校验和	无 krb5i	是，使用 AES
	RPC 数据的完整性校验和		是，使用 AES

使用 Kerberos 身份验证时，需要考虑以下注意事项：

- ESXi 使用 Kerberos 与 Active Directory 域。
- 作为 vSphere 管理员，您可以指定 Active Directory 凭据以向 NFS 用户提供 NFS 4.1 Kerberos 数据存储的访问权限。一组凭据可用于访问在该主机上挂载的所有 Kerberos 数据存储。
- 多个 ESXi 主机共享 NFS 4.1 数据存储时，必须对访问共享数据存储的所有主机使用相同的 Active Directory 凭据。要自动执行分配过程，请在主机配置文件中设置用户并将配置文件应用于所有 ESXi 主机。
- 不能对多个主机共享的同一个 NFS 4.1 数据存储使用两个安全机制：AUTH\_SYS 和 Kerberos。

有关分步说明，请参见《《vSphere 存储》》文档。

## 设置 NFS 存储环境

您必须先执行若干配置步骤，然后再在 vSphere 中挂载 NFS 数据存储。

### 前提条件

- 熟悉 [NFS 存储准则和要求](#)中的准则。
- 有关配置 NFS 存储的详细信息，请参阅存储供应商文档。
- 如果使用 Kerberos，请确保在 NAS 服务器上启用 AES256-CTS-HMAC-SHA1-96 或 AES128-CTS-HMAC-SHA1-96。

### 步骤

- 1 在 NFS 服务器上，配置 NFS 卷并将其导出以挂载在 ESXi 主机上。
  - a 请记下 NFS 服务器的 IP 地址或 DNS 名称以及 NFS 共享的完整路径或文件夹名称。

对于 NFS 4.1，您可以收集多个 IP 地址或 DNS 名称以利用 NFS 4.1 数据存储提供的多路径支持。
  - b 如果计划将 Kerberos 身份验证与 NFS 4.1 搭配使用，请指定 ESXi 用于进行身份验证的 Kerberos 凭据。
- 2 在每台 ESXi 主机上，配置 NFS 流量的 VMkernel 网络端口。

有关详细信息，请参见《《vSphere 网络连接》》文档。
- 3 如果计划将 Kerberos 身份验证与 NFS 4.1 数据存储搭配使用，请配置 ESXi 主机进行 Kerberos 身份验证。

请参见[配置 ESXi 主机进行 Kerberos 身份验证](#)。

### 后续步骤

现在，您可以在 ESXi 主机上创建 NFS 数据存储。

## 配置 ESXi 主机进行 Kerberos 身份验证

如果将 NFS 4.1 与 Kerberos 配合使用，则必须执行多项任务以设置主机进行 Kerberos 身份验证。

多个 ESXi 主机共享 NFS 4.1 数据存储时，必须对访问共享数据存储的所有主机使用相同的 Active Directory 凭据。您可以通过设置主机配置文件中的用户并将该配置文件应用到所有 ESXi 主机来自动执行分配流程。

### 前提条件

- 确保已将 Microsoft Active Directory (AD) 和 NFS 服务器配置为使用 Kerberos。
- 在 AD 上启用 AES256-CTS-HMAC-SHA1-96 或 AES128-CTS-HMAC-SHA1-96 加密模式。NFS 4.1 客户端不支持 DES-CBC-MD5 加密模式。

- 确保已将 NFS 服务器导出配置为授予 Kerberos 用户完全访问权限。

步骤

1 为使用 Kerberos 的 NFS 4.1 配置 DNS

如果将 NFS 4.1 与 Kerberos 结合使用，则必须更改 ESXi 主机上的 DNS 设置。设置必须指向配置为为 Kerberos 密钥分发中心 (KDC) 分发 DNS 记录的 DNS 服务器。例如，如果使用 AD 作为 DNS 服务器，则使用 Active Directory 服务器地址。

2 为使用 Kerberos 的 NFS 4.1 配置网络时间协议

如果将 NFS 4.1 与 Kerberos 配合使用，则 ESXi 主机、NFS 服务器和 Active Domain 服务器需要同步时间。通常，在设置中，Active Domain 服务器用作网络时间协议 (NTP) 服务器。

3 在 Active Directory 中启用 Kerberos 身份验证

如果将 NFS 4.1 存储与 Kerberos 配合使用，则必须将每个 ESXi 主机添加到 Active Directory 域并启用 Kerberos 身份验证。Kerberos 可与 Active Directory 集成以实现单点登录，并且在不安全的网络连接中使用时可提供额外的安全层。

为使用 Kerberos 的 NFS 4.1 配置 DNS

如果将 NFS 4.1 与 Kerberos 结合使用，则必须更改 ESXi 主机上的 DNS 设置。设置必须指向配置为为 Kerberos 密钥分发中心 (KDC) 分发 DNS 记录的 DNS 服务器。例如，如果使用 AD 作为 DNS 服务器，则使用 Active Directory 服务器地址。

步骤

- 1 在 vSphere Client 中，导航到 ESXi 主机。
- 2 单击配置选项卡。
- 3 在网络下，单击 TCP/IP 配置。
- 4 选择默认，然后单击编辑图标。
- 5 手动输入 DNS 设置。

选项	描述
域	AD 域名
首选 DNS 服务器	AD 服务器 IP
搜索域	AD 域名

为使用 Kerberos 的 NFS 4.1 配置网络时间协议

如果将 NFS 4.1 与 Kerberos 配合使用，则 ESXi 主机、NFS 服务器和 Active Domain 服务器需要同步时间。通常，在设置中，Active Domain 服务器用作网络时间协议 (NTP) 服务器。

以下任务介绍了如何同步 ESXi 主机和 NTP 服务器。

最佳做法是使用 Active Domain 服务器作为 NTP 服务器。

## 步骤

- 1 在 vSphere Client 中，导航到 ESXi 主机。
- 2 单击**配置**选项卡。
- 3 在**系统**下，选择**时间配置**。
- 4 单击**编辑**并设置 NTP 服务器。
  - a 选择**使用网络时间协议 (启用 NTP 客户端)**。
  - b 要与 NTP 服务器同步，请输入其 IP 地址。
  - c 选择**启动 NTP 服务**。
  - d 设置 NTP 服务启动策略。
- 5 单击**确定**。

此时，主机将与 NTP 服务器同步。

## 在 Active Directory 中启用 Kerberos 身份验证

如果将 NFS 4.1 存储与 Kerberos 配合使用，则必须将每个 ESXi 主机添加到 Active Directory 域并启用 Kerberos 身份验证。Kerberos 可与 Active Directory 集成以实现单点登录，并且在网络不安全连接中使用时可提供额外的安全层。

## 前提条件

设置 AD 域和有权将主机添加到域的域管理员帐户。

## 步骤

- 1 在 vSphere Client 中，导航到 ESXi 主机。
- 2 单击**配置**选项卡。
- 3 在**系统**下，单击**身份验证服务**。
- 4 将 ESXi 主机添加到 Active Directory 域。
  - a 在“身份验证服务”窗格中，单击**加入域**。
  - b 提供域设置，然后单击**确定**。
- 5 配置或编辑 NFS Kerberos 用户的凭据。
  - a 在“NFS Kerberos 凭据”窗格中，单击**编辑**。
  - b 输入用户名和密码 (Windows session credentials cannot be used to log into this server. Enter a user name and password)。

使用这些凭据可访问存储在所有 Kerberos 数据存储中的文件。

NFS Kerberos 凭据的状态将更改为“已启用”。

后续步骤

为主机配置 Kerberos 后，可以创建启用 Kerberos 的 NFS 4.1 数据存储。

收集 NFS 存储的统计信息

可以在 ESXi 主机中使用 `nfsStats` 工具显示有关 NFS 调用和远程过程调用 (RPC) 的统计信息。此命令显示 ESXi 主机上的 NFS 3 和 NFS 4.1 挂载的统计信息。

通常，`nfsStats` 工具将执行以下任务。

- 收集 NFS 统计信息，以调查在 NFS 环境中部署新配置（例如新的 NFS 服务器或网络）时遇到的问题。
- 提供有关 NFS 操作成功和失败的统计信息。
- 发布有关 NFS 操作成功和失败的延迟统计信息。
- 对 NFS 性能问题进行故障排除。

命令语法为 `nfsStats options`。

可以使用以下命令选项。

表 9-6. `nfsStats` 命令

命令选项	描述
No option	获取所有 NFS 数据存储的 NFS 统计信息和 RPC 统计信息。
-3	仅显示 NFS 3 统计信息。
-4	仅显示 NFS 4.1 统计信息。
-n	仅显示 NFS 3 和 NFS 4.1 统计信息。
-r	显示 RPC 统计信息。
-i interval	以等于指定值（以秒为单位）的时间间隔显示 NFS 和 RPC 统计信息。例如，如果输入的值为 10，则统计信息将每 10 秒刷新一次。
-j	以 json 格式显示统计信息。
-e	显示 NFS 3、NFS 4.1 和 RPC 协议的错误统计信息。
-d, --datastore-name DSNAME1, DSNAME2, ...	显示指定 NFS 数据存储的 NFS 和 RPC 统计信息。将此选项与 NFS 数据存储的类型（例如，-3 或 -4）一起使用。

创建数据存储

您可以使用“新建数据存储”向导创建数据存储。根据存储类型和存储需求，可创建 VMFS、NFS 或 Virtual Volumes 数据存储。

启用 vSAN 时，会自动创建 vSAN 数据存储。有关信息，请参见《管理 VMware vSAN》文档。

也可以使用“新建数据存储”向导管理 VMFS 数据存储副本。

■ [创建 VMFS 数据存储](#)

VMFS 数据存储充当虚拟机的存储库。可以在主机发现的基于 SCSI 的任何存储设备（包括光纤通道、iSCSI 和本地存储设备）上设置 VMFS 数据存储。

■ [创建 NFS 数据存储](#)

您可使用[新建数据存储](#)向导挂载 NFS 卷。

■ [创建 Virtual Volumes 数据存储](#)

您可以使用[新建数据存储](#)向导创建 Virtual Volumes 数据存储。

## 创建 VMFS 数据存储

VMFS 数据存储充当虚拟机的存储库。可以在主机发现的基于 SCSI 的任何存储设备（包括光纤通道、iSCSI 和本地存储设备）上设置 VMFS 数据存储。

### 前提条件

- 1 安装并配置存储所需的所有适配器。
- 2 要发现新添加的存储设备，请执行重新扫描。请参见[ESXi 存储的重扫描操作](#)。
- 3 验证您计划用于数据存储的存储设备是否可用。请参见[查看可用于 ESXi 主机的存储设备](#)。

### 步骤

- 1 在 vSphere Client 对象导航器中，浏览到主机、集群或数据中心。
- 2 从右键单击菜单中，选择**存储 > 新建数据存储**。
- 3 选择“VMFS”作为数据存储类型。
- 4 输入数据存储名称，并选择该数据存储的放置位置（如有必要）。  
对于数据存储名称，系统强制执行 42 个字符的限制。
- 5 选择要用于数据存储的设备。

---

**重要说明** 所选设备不得具有任何在“快照卷”列中显示的值。如果存在某值，则设备包含现有 VMFS 数据存储的副本。有关管理数据存储副本的信息，请参见[管理重复 VMFS 数据存储](#)。

---

- 6 指定数据存储版本。

选项	描述
VMFS6	支持 VMFS6 的所有主机上的默认格式。版本 6.0 或更早版本的 ESXi 主机无法识别 VMFS6 数据存储。
VMFS5	VMFS5 数据存储支持版本 6.7 或更早版本的 ESXi 主机进行访问。

## 7 为数据存储定义配置详细信息。

**注** VMFS6 数据存储所需的最小大小为 2 GB。

### a 指定分区配置。

选项	描述
使用所有可用分区	将整个磁盘专用于单个 VMFS 数据存储。如果选择此选项，则当前存储在此设备上的所有文件系统和数据将损坏。
使用可用空间	在剩余的可用磁盘空间中部署 VMFS 数据存储。

- b 如果为数据存储分配的空间超过您的计划，请调整“数据存储大小”字段中的容量值。  
默认情况下，将分配存储设备上的全部可用空间。
- c 对于 VMFS6，请指定块大小并定义空间回收参数。请参见 [VMFS 数据存储上的空间回收](#)。

## 8 在“即将完成”页面，检查数据存储配置信息，然后单击**完成**。

### 结果

即会在基于 SCSI 的存储设备上创建数据存储。该数据存储可用于有权访问此设备的所有主机。

### 后续步骤

创建 VMFS 数据存储以后，可以执行以下任务：

- 更改数据存储的容量。请参见[增加 VMFS 数据存储容量](#)。
- 编辑空间回收设置。请参见[为现有 VMFS6 数据存储配置固定回收](#)。
- 启用共享 VMDK 支持。请参见[在 VMFS6 数据存储上启用或禁用对集群虚拟磁盘的支持](#)。

## 创建 NFS 数据存储

您可使用[新建数据存储](#)向导挂载 NFS 卷。

### 前提条件

- 设置 NFS 存储环境。
- 如果计划将 Kerberos 身份验证与 NFS 4.1 数据存储搭配使用，请确保配置 ESXi 主机进行 Kerberos 身份验证。

### 步骤

- 1 在 vSphere Client 对象导航器中，浏览到主机、集群或数据中心。
- 2 从右键单击菜单中，选择**存储 > 新建数据存储**。
- 3 选择 NFS 作为数据存储类型，并指定 NFS 版本。
  - NFS 3

## ■ NFS 4.1

**重要说明** 如果多个主机访问同一数据存储，则所有主机上必须使用相同的协议。

## 4 输入数据存储参数。

选项	描述
数据存储名称	对于数据存储名称，系统强制执行 42 个字符的限制。
文件夹	挂载点文件夹名称
Server	服务器名称或 IP 地址。您可以使用 IPv6 或 IPv4 格式。 使用 NFS 4.1 时，如果 NFS 服务器支持中继，则可以添加多个 IP 地址或服务器名称。ESXi 主机将使用这些值实现 NFS 服务器挂载点的多路径。

5 如果 NFS 服务器将卷作为只读存储导出，则选择**挂载只读 NFS**。

## 6 要将 Kerberos 安全功能与 NFS 4.1 结合使用，请启用 Kerberos 并选择适当的 Kerberos 模型。

选项	描述
仅将 Kerberos 用于身份验证 (krb5)	支持身份验证
将 Kerberos 用于身份验证和数据完整性 (krb5i)	除了身份验证，还提供数据完整性服务。这些服务通过检查潜在的数据包修改操作，帮助保护 NFS 流量免受篡改。

如果您不启用 Kerberos，则数据存储将使用默认的 AUTH\_SYS 安全功能。

## 7 如果在数据中心或集群级别创建数据存储，请选择挂载数据存储的主机。

8 检查配置选项，然后单击**完成**。

## 创建 Virtual Volumes 数据存储

您可以使用**新建数据存储**向导创建 Virtual Volumes 数据存储。

### 步骤

## 1 在 vSphere Client 对象导航器中，浏览到主机、集群或数据中心。

2 从右键单击菜单中，选择**存储 > 新建数据存储**。3 选择 **VVOL** 作为数据存储类型。

## 4 输入数据存储名称，并从存储容器列表中选择备用存储容器。

确保使用的名称不会与数据中心环境中的其他数据存储名称重复。

如果将同一 Virtual Volumes 数据存储挂载到多台主机上，则该数据存储的名称必须在所有主机中保持一致。

## 5 选择需要访问数据存储的主机。

6 检查配置选项，然后单击**完成**。



## 后续步骤

创建 Virtual Volumes 数据存储后，可以执行重命名数据存储、浏览数据存储文件、卸载数据存储等数据存储操作。

无法将 Virtual Volumes 数据存储添加到数据存储集群中。

## 管理重复 VMFS 数据存储

当存储设备包含 VMFS 数据存储副本时，您可以利用现有签名挂载此数据存储或为其分配新签名。

在存储设备中创建的每个 VMFS 数据存储都有一个唯一签名，也称为 UUID，存储在文件系统超级块中。对存储设备进行复制或在线端生成快照后，生成的设备副本的每个字节都与原始设备完全相同。例如，如果原始存储设备包含具有 UUIDX 的 VMFS 数据存储，则副本会显示为包含具有相同 UUIDX 的数据存储副本。

除了 LUN 快照和复制，某些设备操作，如 LUN ID 变更，可能会生成一份原始数据存储。

ESXi 可以检测 VMFS 数据存储副本。使用数据存储副本的原始 UUID 对其进行挂载或更改 UUID。更改 UUID 的过程称为数据存储重新签名。

您是选择重新签名，还是选择挂载而不重新签名，这取决于如何在存储环境中对 LUN 进行掩码。如果您的主机能够查看 LUN 的两个副本，则重新签名是最佳方法。

## 保留现有的数据存储签名

如果不需要对 VMFS 数据存储副本进行再签名，则无需更改其签名即可挂载。

例如，如果作为灾难恢复计划的一部分，在辅助站点上维护虚拟机的同步副本，则可以保留签名。在主站点发生灾难时，可以在辅助站点上挂载数据存储副本并打开虚拟机电源。

## 对 VMFS 数据存储副本进行重新签名

如果要保留 VMFS 数据存储副本上所存储的数据，请使用数据存储再签名。

对 VMFS 副本进行重新签名时，ESXi 会为副本分配新的签名 (UUID)，并将副本装载为与原始数据存储明显不同的数据存储。对虚拟机配置文件中原始签名的所有引用都会进行更新。

在执行数据存储再签名时，请考虑以下几点：

- 数据存储再签名不可逆。
- 重新签名之后，不再将包含 VMFS 副本的存储设备副本视为副本。
- 仅当跨区数据存储的所有数据区联机时，才可对其进行再签名。
- 再签名过程具有容错性。如果过程中断，可以稍后恢复。
- 您可挂载新的 VMFS 数据存储，而无需承担其 UUID 与设备快照层次结构中的任何其他任何数据存储 UUID 相冲突的风险。

## 挂载 VMFS 数据存储副本

如果要保留 VMFS 数据存储副本上所存储的数据，请使用数据存储再签名。如果不需要对 VMFS 数据存储副本进行重新签名，则无需更改其签名即可挂载。

### 前提条件

- 在主机上执行存储重新扫描，以便更新存储设备在主机上呈现的视图。
- 卸载与您计划挂载的副本具有相同 UUID 的原始 VMFS 数据存储。仅当与原始 VMFS 数据存储不冲突时，才可以挂载 VMFS 数据存储副本。

### 步骤

- 1 在 vSphere Client 对象导航器中，浏览到主机、集群或数据中心。
- 2 从右键单击菜单中，选择**存储 > 新建数据存储**。
- 3 选择“VMFS”作为数据存储类型。
- 4 输入数据存储名称，并选择该数据存储的放置位置（如有必要）。
- 5 从存储设备列表中，选择具有显示在“快照卷”列中的特殊值的设备。  
“快照卷”列中显示的值表示设备是包含现有 VMFS 数据存储的副本。
- 6 挂载数据存储。

选项	描述
重新签名后进行挂载	在 <b>挂载选项</b> 下，选择 <b>分配新签名</b> ，然后单击 <b>下一步</b> 。
无需重新签名而进行挂载	在“挂载选项”下面，选择 <b>保留现有的签名</b> 。

- 7 检查数据存储配置信息，然后单击**完成**。

## 增加 VMFS 数据存储容量

您可以增加 VMFS 数据存储的容量。当您将虚拟机添加到数据存储，或者在数据存储上运行的虚拟机需要更多空间时，可能需要更多容量。

如果共享数据存储有已打开电源的虚拟机并被 100% 占用，您可以增加数据存储的容量。只能从已打开电源的虚拟机的注册主机执行此操作。

根据您的存储配置，您可以使用以下方法之一增加数据存储容量。在使用增加数据存储容量的任一方法时，不需要关闭虚拟机的电源。

### 扩展现有数据存储

增加可扩展数据存储的大小。如果支持存储设备在紧邻数据存储数据区之后具有可用空间，则该数据存储会被视为可扩展。

### 添加数据区

通过将新存储设备添加到数据存储来增加现有 VMFS 数据存储的容量。数据存储可以跨越多个存储设备，但仍显示为单个卷。

跨区的 VMFS 数据存储可以随时使用其任何或所有数据区。使用下一个数据区之前，不需要填充特定数据区。

**注** 仅支持硬件辅助锁定的数据存储（也称为原子测试和设置 (Atomic Test and Set, ATS) 机制）无法跨非 ATS 设备。有关详细信息，请参见 [VMFS 锁定机制](#)。

### 前提条件

如果主机存储符合以下条件之一，您可以增加数据存储的容量：

- 现有数据存储的支持设备有足够的可用空间。
- 您已向主机添加新的存储设备。

### 步骤

- 1 在 vSphere Client 中，导航到数据存储。
- 2 从数据存储右键单击菜单中选择**增加数据存储容量**。
- 3 从存储设备列表选择一个设备。

您的选择取决于是否有可用的可扩展存储设备。

选项	描述
扩展现有数据存储数据区	选择对应的“可扩展”列为“是”的设备。
添加数据区	选择“可扩展”列为“否”的设备。

- 4 检查**分区布局**以查看可用的配置。
- 5 从底部面板选择配置选项。

根据磁盘的当前布局和以前的选择，您看到的菜单项可能有所不同。

菜单项	描述
使用可用空间扩展数据存储	将现有数据区扩展到所需容量。
使用可用空间	在剩余的可用磁盘空间中部署数据区。此菜单项仅在添加数据区时可用。
使用所有可用分区	将整个磁盘专用于单个数据区。此菜单项仅在添加数据区并且所格式化的磁盘非空白时可用。磁盘将被重新格式化，其中所包含的数据存储和任何数据将被擦除。

- 6 设置数据区的容量。

数据区最小大小为 1.3 GB。默认情况下，存储设备上的全部可用空间均可供使用。

- 7 单击**下一步**。
- 8 检查推荐的数据区布局和数据存储的新配置，然后单击**完成**。

## 在 VMFS6 数据存储上启用或禁用对集群虚拟磁盘的支持

如果计划在 Windows Server 故障切换集群 (WSFC) 配置中使用虚拟磁盘，则 VMFS6 数据存储必须支持集群虚拟磁盘。使用 vSphere Client 启用集群磁盘支持。

有关在虚拟机集群中使用集群虚拟磁盘的信息，请参见《Windows Server 故障切换集群设置》文档。

### 前提条件

对集群虚拟磁盘使用数据存储时，请遵循以下准则：

- 存储阵列必须支持 ATS、独占写入 - 所有注册者 (WEAR) SCSI-3 预留类型。
- ESXi 仅对此类型的配置支持光纤通道阵列。
- 仅 VMFS6 数据存储支持集群磁盘。您使用的数据存储无法展开或跨多个数据区。
- 存储设备必须由 NMP 声明。在集群虚拟磁盘配置中，ESXi 不支持第三方插件 (MPP)。
- 确保用于集群的虚拟磁盘采用厚置备快速置零格式。

### 步骤

- 1 在 vSphere Client 中，导航到数据存储。
- 2 依次单击**配置**选项卡和**常规**。
- 3 在**数据存储功能**下，单击**集群 VMDK** 项旁边的以下选项之一。

选项	描述
启用	在数据存储上启用对集群虚拟磁盘的支持。启用支持后，便可以将集群虚拟磁盘置于此 VMFS 数据存储上。
禁用	禁用支持。在禁用之前，请务必关闭具有集群虚拟磁盘的所有虚拟机的电源。

- 4 确认配置。

## 数据存储的管理操作

创建数据存储后，可以对数据存储执行多个管理操作。某些操作（如重命名数据存储）适用于所有类型的数据存储。其他操作适用于特定类型的数据存储。

### 更改数据存储名称

使用 vSphere Client 可更改现有数据存储的名称。您可以重命名有虚拟机运行的数据存储，不会带来任何负面影响。

**注** 如果主机由 vCenter Server 管理，则无法通过直接从 VMware Host Client 访问主机来重命名数据存储。必须从 vCenter Server 重命名数据存储。

### 步骤

- 1 在 vSphere Client 中，导航到数据存储。

- 2 右键单击要重命名的数据存储，然后选择**重命名**。
- 3 输入新的数据存储名称。

对于数据存储名称，系统强制执行 42 个字符的限制。

#### 结果

在具有该数据存储访问权限的所有主机上都将显示新名称。

## 卸载数据存储

卸载数据存储时，它会保持原样，但是在指定的主机上再也看不到该存储。该数据存储会继续显示在其他主机上并在这些主机上保持挂载状态。

正在进行卸载时，请勿对数据存储执行任何可能会导致 I/O 的配置操作。

---

**注** 确保 vSphere HA Heartbeating 未使用数据存储，因此 vSphere HA Heartbeating 不会阻止您卸载数据存储。但是，如果数据存储用于检测信号，卸载它可能会导致主机发生故障并重新启动任何活动虚拟机。

---

#### 前提条件

适用时，在卸载数据存储之前，请确保符合以下必备条件：

- 数据存储上不存在任何虚拟机。
- Storage DRS 不会管理数据存储。
- 已为该数据存储禁用 Storage I/O Control。

#### 步骤

- 1 在 vSphere Client 中，导航到数据存储。
- 2 右键单击数据存储，然后选择**卸载据存储**。
- 3 如果数据存储已共享，选择要从中卸载数据存储的主机。
- 4 确认要卸载数据存储。

#### 结果

将某个 VMFS 数据存储从所有主机上卸载之后，该数据存储将被标记为非活动。如果从所有主机上卸载 NFS 或 Virtual Volumes 数据存储，该数据存储将从清单中消失。您可以挂载已卸载的 VMFS 数据存储。要挂载已从清单中移除的 NFS 或 Virtual Volumes 数据存储，请使用“新建数据存储”向导。

#### 后续步骤

如果是在存储移除过程中卸载的 VMFS 数据存储，则现在可分离支持该数据存储的存储设备。请参见[分离存储设备](#)。

## 挂载数据存储

您可以挂载之前已卸载的数据存储。您可以将数据存储挂载在其他主机上，让其成为共享存储。

已从所有主机上卸载的 VMFS 数据存储仍在清单中，但会被标记为不可访问。请参见[卸载数据存储](#)。

可以使用此任务将 VMFS 数据存储挂载到指定的主机或多个主机。

如果已从所有主机上卸载 NFS 或 Virtual Volumes 数据存储，该数据存储将从清单中消失。要挂载已从清单中移除的 NFS 或 Virtual Volumes 数据存储，请使用“新建数据存储”向导。

从部分主机卸载同时又挂载在其他主机上的任何类型的数据存储在清单中显示为处于活动状态。

### 步骤

- 1 在 vSphere Client 中，导航到数据存储。
- 2 右键单击要挂载的数据存储，然后选择以下选项之一：
  - **挂载数据存储**
  - **在其他主机上挂载数据存储**是否可以查看一个或另一个选项取决于所使用的数据存储类型。
- 3 选择应访问数据存储的主机，然后单击**确定**。
- 4 要列出共享数据存储的所有主机，请导航到数据存储，然后单击**主机**选项卡。

## 移除 VMFS 数据存储

可以删除任何类型的 VMFS 数据存储（包括已挂载但未再签名的副本）。删除数据存储时，会对其造成损坏，而且它将从具有数据存储访问权限的所有主机中消失。

---

**注** 数据存储的删除操作会永久删除与数据存储上的虚拟机关联的所有文件。尽管您可以不进行卸载便删除数据存储，但您最好先卸载数据存储。

---

### 前提条件

- 从数据存储中移除或迁移所有虚拟机。
- 从所有主机卸载数据存储。
- 为数据存储禁用 Storage DRS。
- 为数据存储禁用 Storage I/O Control。
- 确保数据存储未用于 vSphere HA 检测信号。

### 步骤

- 1 在 vSphere Client 中，导航到数据存储。
- 2 右键单击要移除的数据存储，然后选择**删除数据存储**。
- 3 确认要移除数据存储。

## 使用数据存储浏览器

使用数据存储文件浏览器可管理数据存储的内容。您可以浏览存储在数据存储中的文件夹和文件。您还可以使用该浏览器上传文件并对文件夹和文件执行管理任务。

### 步骤

- 1 打开数据存储浏览器。
  - a 显示清单中的数据存储。
  - b 右键单击数据存储，然后选择**浏览文件**。
- 2 导航到现有文件夹和文件，了解数据存储的内容。
- 3 利用图标和选项执行管理任务。

图标和选项	描述
上传文件	将文件上传到数据存储。
上传文件夹（仅在 vSphere Client 中可用）	将文件夹上传到数据存储。
下载	从数据存储下载。
新建文件夹	在数据存储上创建一个文件夹。
复制到	将所选文件夹或文件复制到一个新的位置（该位置可以在同一数据存储上，也可在其他数据存储上）。
移至	将所选文件夹或文件移到一个新的位置（该位置可以在同一数据存储上，也可在其他数据存储上）。
重命名为	重命名所选文件。
删除	删除所选文件夹或文件。
扩充	将所选精简虚拟磁盘转换为厚虚拟磁盘。该选项仅适用于精简置备磁盘。

## 将文件或文件夹上传至数据存储

使用数据存储文件浏览器可将文件上传至 ESXi 主机上的数据存储。如果使用 vSphere Client，也可以上传文件夹。

除了用作虚拟机文件存储这一传统用法外，数据存储还可用于存储与虚拟机有关的数据或文件。例如，可以将操作系统的 ISO 映像从本地计算机上传至主机上的数据存储，然后使用这些映像在新虚拟机上安装客户机操作系统。

**注** 您无法直接将文件上传到 Virtual Volumes 数据存储。您必须首先在 Virtual Volumes 数据存储中创建一个文件夹，然后将文件上传到此文件夹中。在 Virtual Volumes 数据存储中为块存储创建的文件夹的存储容量空间有限 (4GB)。Virtual Volumes 数据存储支持直接上传文件夹。

### 前提条件

所需特权：**数据存储.浏览数据存储**

**步骤**

- 1 打开数据存储浏览器。
  - a 显示清单中的数据存储。
  - b 右键单击数据存储，然后选择**浏览文件**。
- 2 （可选） 创建一个文件夹用于存储文件或文件夹。
- 3 上载文件或文件夹。

选项	描述
上载文件	<ol style="list-style-type: none"> <li>a 选择目标文件夹，然后单击<b>上载文件</b>。</li> <li>b 在本地计算机中找到要上载的项目，然后单击<b>打开</b>。</li> </ol>
上载文件夹（仅在 vSphere Client 中可用）	<ol style="list-style-type: none"> <li>a 选择数据存储或目标文件夹，然后单击<b>上载文件夹</b>。</li> <li>b 在本地计算机中找到要上载的项目，然后单击<b>确定</b>。</li> </ol>

- 4 刷新数据存储文件浏览器，查看列表中已上载的文件或文件夹。

**后续步骤**

在部署先导出、然后再上载到数据存储的 OVF 模板时可能会出现问题。有关详细信息和解决办法，请参见 VMware 知识库文章 [2117310](#)。

**从数据存储下载文件**

使用数据存储文件浏览器从您的 ESXi 主机上可用的数据存储下载文件到您的本地计算机上。

**前提条件**

所需特权：**数据存储.浏览数据存储**

**步骤**

- 1 打开数据存储浏览器。
  - a 显示清单中的数据存储。
  - b 右键单击数据存储，然后选择**浏览文件**。
- 2 导航到要下载的文件，然后单击**下载**。
- 3 按照提示将文件保存至您的本地计算机。

**移动或复制数据存储文件夹或文件**

可以使用数据存储浏览器将文件夹或文件移动或复制到一个新的位置（该位置可以在同一数据存储上，也可在其他数据存储上）。

**注** 虚拟磁盘文件无需进行格式转换即可移动或复制。如果您将虚拟磁盘移到的数据存储并非属于源主机，您可能需要转换虚拟磁盘。否则，您可能无法使用该磁盘。

您无法跨 vCenter Server 复制虚拟机文件。



### 前提条件

所需特权：**数据存储.浏览数据存储**

### 步骤

- 1 打开数据存储浏览器。
  - a 显示清单中的数据存储。
  - b 右键单击数据存储，然后选择**浏览文件**。
- 2 浏览到要移动或复制的对象（文件夹或文件）。
- 3 选择对象并单击**移至或复制到**。
- 4 指定目标位置。
- 5 （可选）选择**覆盖目标位置上具有匹配名称的文件和文件夹**。
- 6 单击**确定**。

## 重命名数据存储文件

使用数据存储浏览器重命名文件。

### 前提条件

所需特权：**数据存储.浏览数据存储**

### 步骤

- 1 打开数据存储浏览器。
  - a 显示清单中的数据存储。
  - b 右键单击数据存储，然后选择**浏览文件**。
- 2 浏览到想要重命名的文件。
- 3 选择文件，然后单击**重命名为**。
- 4 指定新名称，然后单击**确定**。

## 扩充精简虚拟磁盘

如果创建的是精简格式的虚拟磁盘，可以将其格式更改为厚磁盘。

使用数据存储浏览器扩充精简虚拟磁盘。

### 前提条件

- 请确保虚拟机驻留的数据存储具有足够的空间。
- 请确保虚拟磁盘为精简磁盘。
- 移除快照。
- 关闭虚拟机电源。

步骤

- 1 在 vSphere Client 中，导航到要扩充的虚拟磁盘的文件夹。
  - a 导航到虚拟机。
  - b 单击**数据存储**选项卡。  
此时将列出存储虚拟机文件的数据存储。
  - c 右键单击数据存储，然后选择**浏览文件**。  
数据存储浏览器将显示该数据存储的内容。
- 2 扩展虚拟机文件夹，浏览到要转换的虚拟磁盘文件。  
该文件使用 .vmdk 扩展名，并带有虚拟磁盘图标标记 (🗄️)。
- 3 选择虚拟磁盘文件，然后单击**扩充**。

**注** 如果虚拟磁盘为厚磁盘或当虚拟机正在运行时，该选项可能不可用。

结果

扩充的虚拟磁盘将占据最初为其置备的整个数据存储空间。

关闭存储筛选器

在执行 VMFS 数据存储管理操作时，vCenter Server 使用默认存储保护筛选器。通过仅检索可用于特定操作的存储设备，这些筛选器可帮助您避免存储设备损坏。不适合的设备不会显示出来供选择。可以关闭筛选器来查看所有设备。

前提条件

在更改设备筛选器之前，请咨询 VMware 技术支持。

步骤

- 1 浏览到 vCenter Server 实例。
- 2 单击**配置**选项卡。
- 3 在**设置**下，单击**高级设置**，然后单击**编辑设置**。
- 4 指定要关闭的筛选器。  
在页面底部的**名称**和**值**文本框中，输入相应的信息。

名称	值
config.vpxd.filter.vmfsFilter	无效
config.vpxd.filter.rdmFilter	无效

名称	值
config.vpxd.filter.sameHostsAndTransportFilter	无效
config.vpxd.filter.hostRescanFilter	无效
<b>注</b> 如果关闭此筛选器，则每次将一个新的 LUN 提供给主机或集群时，主机仍继续执行重新扫描。	

5 单击**添加**，然后单击**保存**以保存更改。

无需重新启动 vCenter Server 系统。

## 存储筛选

vCenter Server 提供存储筛选器来帮助避免发生因使用不受支持的存储设备而可能导致的存储设备损坏或性能下降。默认情况下，这些筛选器是可用的。

表 9-7. 存储筛选器

筛选器名称	描述
config.vpxd.filter.vmfsFilter (VMFS 筛选器)	筛选出 vCenter Server 管理的任何主机上的 VMFS 数据存储已使用的存储设备或 LUN。这些 LUN 不会显示为要使用其他 VMFS 数据存储进行格式化或用作 RDM 的候选对象。
config.vpxd.filter.rdmFilter (RDM 筛选器)	筛选出 vCenter Server 管理的任何主机上的 RDM 已引用的 LUN。这些 LUN 不会显示为要使用 VMFS 进行格式化或要用于其他 RDM 的候选对象。 要使虚拟机可以访问相同的 LUN，虚拟机必须共享相同的 RDM 映射文件。有关此配置类型的信息，请参见《vSphere 资源管理》文档。
config.vpxd.filter.sameHostsAndTransportFilter (相同主机和传输筛选器)	筛选出因为主机或存储类型不兼容而不适合用作 VMFS 数据存储数据区的 LUN。防止将以下 LUN 添加为数据区： <ul style="list-style-type: none"> <li>■ 不向共享原始 VMFS 数据存储的所有主机公开的 LUN。</li> <li>■ 使用的存储类型与原始 VMFS 数据存储使用的存储类型不同的 LUN。例如，无法将光纤通道数据区添加到本地存储设备上的 VMFS 数据存储。</li> </ul>
config.vpxd.filter.hostRescanFilter (主机重新扫描筛选器)	在执行数据存储管理操作之后，会自动重新扫描和更新 VMFS 数据存储。该筛选器可帮助提供 vCenter Server 管理的所有主机上的所有 VMFS 数据存储的一致视图。 <b>注</b> 如果将新的 LUN 提供给主机或集群，无论主机重新扫描筛选器是打开还是关闭，主机都会自动执行重新扫描。

## 设置动态磁盘镜像

通常，您不能使用虚拟机上的 LUN 管理器软件对虚拟磁盘进行镜像。但是，如果您的 Microsoft Windows 虚拟机支持动态磁盘，则可以跨两个 SAN LUN 对虚拟磁盘进行镜像。镜像可帮助您防止虚拟机发生存储设备意外丢失。

### 前提条件

- 使用支持动态磁盘的 Windows 虚拟机。
- 所需特权：**虚拟机.配置.设置**

**步骤**

- 1 创建具有两个虚拟磁盘的虚拟机。

将这两个磁盘置于不同的数据存储上。

- 2 登录到虚拟机并将磁盘配置为动态镜像磁盘。

请参见 [Microsoft 文档](#)。

- 3 同步磁盘后，关闭虚拟机的电源。

- 4 更改虚拟机设置以启用动态磁盘镜像。

- a 右键单击虚拟机，然后选择**编辑设置**。
- b 单击**虚拟机选项**选项卡，然后展开**高级**菜单。
- c 单击“配置参数”旁边的**编辑配置**。
- d 单击**添加配置参数**，然后添加以下参数：

名称	值
<code>scsi#.returnNoConnectDuringAPD</code>	True
<code>scsi#.returnBusyOnNoConnectStatus</code>	无效

- e 如果使用的是 ESXi6.7 或更高版本，请为参与软件 RAID-1 配置的每个虚拟磁盘提供一个额外参数。

该参数用于防止在存储设备出现故障时出现客户机操作系统 I/O 故障。

名称	值
<code>scsi#:1.passthruTransientErrors</code>	True
<code>scsi#:2.passthruTransientErrors</code>	True

- f 单击**确定**。

## 在 VMFS 数据存储上收集 ESXi 主机的诊断信息

在主机故障过程中，ESXi 必须能够将诊断信息保存到预配置的位置，以供诊断和技术支持之用。

通常，在 ESXi 安装期间会在本地存储设备上创建收集诊断信息的分区（也称为核心转储）。您还可以配置 ESXi Dump Collector，以在网络服务器上保留核心转储。有关设置 ESXi Dump Collector 的信息，请参见《VMware ESXi 安装和设置》文档。

另一个选项是使用 VMFS 数据存储上的文件收集诊断信息。

### 将文件设置为核心转储位置

如果可用核心转储分区的大小不足，则可以配置 ESXi 以使用 VMFS 数据存储中的文件获取诊断信息。

---

**注** 软件 iSCSI 上的 VMFS 数据存储不支持核心转储文件。

---

## 前提条件

安装 ESXCLI。请参见《ESXCLI 入门》。要进行故障排除，请在 ESXi Shell 中运行 `esxcli` 命令。

## 步骤

- 1 通过运行以下命令创建 VMFS 数据存储核心转储文件：

```
esxcli system coredump file add
```

该命令可以使用以下选项，但这些选项并不是必需的，可以省略：

选项	描述
<b>--auto   -a</b>	如果未找到文件，将自动创建一个文件。
<b>--datastore   -d</b> <i>datastore_UUID 或</i> <i>datastore_name</i>	指定转储文件的数据存储。如果不提供此选项，系统将选择一个具有足够大小的数据存储。
<b>--enable   -e</b>	创建诊断文件后启用该文件。
<b>--file   -f file_name</b>	指定转储文件的文件名。如果未提供，系统将为该文件创建一个唯一的名称。
<b>--size   -s file_size_MB</b>	设置转储文件的大小（以 MB 为单位）。如果不提供此选项，系统将根据主机上安装的内存创建一个大小合适的文件。

- 2 验证文件是否已创建：

```
esxcli system coredump file list
```

您可能会看到类似如下的输出内容：

Path	Active	Configured	Size
-----	-----	-----	-----
/vmfs/volumes/52b021c3-.../vmkdump/test.dumpfile	false	false	104857600

- 3 激活主机的核心转储文件：

```
esxcli system coredump file set
```

该命令采用以下选项：

选项	描述
<b>--enable   -e</b>	启用或禁用转储文件。取消配置转储文件时，无法指定此选项。
<b>--path   -p</b>	要使用的核心转储文件的路径。文件必须是预先分配的。
<b>--smart   -s</b>	此标记只能与 <b>--enable   -e=true</b> 一起使用。该选项会使用智能选择算法来选择文件。 例如， <b>esxcli system coredump file set --smart --enable true</b>
<b>--unconfigure   -u</b>	取消配置当前的 VMFS 转储文件。

- 4 验证核心转储文件是否处于活动状态且已配置：

```
esxcli system coredump file list
```

类似如下的输出内容表明核心转储文件处于活动状态且已配置：

Path	Active	Configured	Size
/vmfs/volumes/52b021c3-.../vmkdump/test.dumpfile	True	True	104857600

后续步骤

有关可用于管理核心转储文件的其他命令的信息，请参见《《ESXCLI 参考指南》》文档。

取消激活和删除核心转储文件

取消激活配置的核心转储文件，并根据需要将其从 VMFS 数据存储中移除。

您可以临时取消激活核心转储文件。如果不计划使用已取消激活的文件，可以将其从 VMFS 数据存储中移除。要移除尚未取消激活的文件，您可以使用带有 `--force | -F` 参数的 `esxcli system coredump file remove` 命令。

前提条件

安装 ESXCLI。请参见《ESXCLI 入门》。要进行故障排除，请在 ESXi Shell 中运行 `esxcli` 命令。

步骤

1 列出核心转储文件：

```
esxcli system coredump file list
```

2 运行以下命令取消激活核心转储文件：

```
esxcli system coredump file set --unconfigure | -u
```

3 从 VMFS 数据存储中移除该文件：

```
esxcli system coredump file remove --file | -f file_name
```

该命令采用以下选项：

选项	描述
<code>--file   -f</code>	输入要移除的转储文件的名称。如果不输入名称，该命令会移除默认配置的核心转储文件。
<code>--force   -F</code>	取消激活并取消配置所移除的转储文件。如果文件之前尚未取消激活且正处于活动状态，则此选项为必需项。

结果

核心转储文件将被禁用并从 VMFS 数据存储中移除。

## 使用 VOMA 检查元数据一致性

使用 vSphere On-disk Metadata Analyzer (VOMA) 可确定并修复影响文件系统或底层逻辑卷的元数据损坏事件。

### 问题

在遇到与 VMFS 数据存储或虚拟闪存资源有关的问题时，您可以检查元数据一致性。例如，发生下列任一情况时，请执行元数据检查：

- 出现存储故障。
- 重新构建 RAID 或执行磁盘更换后。
- 您在 vmkernel.log 文件中看到了如下所示的元数据错误：

```
cpu11:268057)WARNING: HBX: 599: Volume 50fd60a3-3aae1ae2-3347-0017a4770402
("<Datastore_name>") may be damaged on disk. Corrupt heartbeat detected at offset 3305472:
[HB state 0 offset 6052837899185946624 gen 15439450 stampUS 5 $
```

- 无法在 VMFS 上访问文件。
- 您在 vCenter Server 的事件选项卡中看到正在报告数据存储的损坏问题。

### 解决方案

要检查元数据一致性，请从 ESXi 主机上的 CLI 中运行 VOMA。VOMA 可用于检查并修复 VMFS 数据存储或支持 VMFS 数据存储的逻辑卷的轻微不一致问题。

VOMA 可检查并修复以下各项。

表 9-8. VOMA 功能

VOMA 功能	描述
元数据检查和修复	元数据检查和修复的示例包括但不限于以下内容： <ul style="list-style-type: none"><li>■ 验证 VMFS 卷标头以实现基本元数据一致性。</li><li>■ 检查 VMFS 资源文件（系统文件）的一致性。</li><li>■ 检查所有文件的路径名和连接。</li></ul>
关联性元数据检查和修复	要为 VMFS6 启用关联性检查，请使用 -a --affinityChk 选项。 关联性元数据检查和修复示例如下所示： <ul style="list-style-type: none"><li>■ 资源类型和 FS3_ResFileMetadata 中的关联性标记。</li><li>■ 验证 SFB RC 元中的关联性标记 (FS3_ResourceClusterMDVMFS6)。</li><li>■ 验证 RC rcMeta 的 affinityInfo 条目中的所有条目，包括溢出键，以确保不存在无效项。检查缺少的条目。</li></ul>

表 9-8. VOMA 功能（续）

VOMA 功能	描述
目录验证	<p>VOMA 可以检测并更正以下错误：</p> <ul style="list-style-type: none"><li>■ 目录哈希块损坏。</li><li>■ 分配映射损坏。</li><li>■ 链接块损坏。</li><li>■ 目录项块损坏。</li></ul> <p>VOMA 会根据损坏的性质只修复损坏的条目或完全重新构建哈希块、分配映射块和链接块。</p>
丢失和找到的文件	<p>在文件系统检查过程中，VOMA 可以查找文件系统的任意位置均未引用的文件。这些孤立的文件是有效、完整的，但在系统上没有名称或目录项。</p> <p>如果 VOMA 在扫描过程中遇到孤立的文件，它将在卷的根目录中创建一个名为 <code>lost+found</code> 的目录，用于存储孤立的文件。文件的名称使用 <code>Filesequence-number</code> 格式。</p>

VOMA 工具附带的命令选项包括以下各项。

表 9-9. VOMA 命令选项

命令选项	描述
<code>-m --module</code>	要运行的模块，包括：
	<code>vmfs</code> 如果您未指定模块名称，则默认情况下使用此选项。 您可以检查 VMFS 文件系统，以及为虚拟闪存资源提供支持的文件系统。如果指定了此模块，也会对 LVM 执行最少量的检查。
	<code>lvm</code> 检查为 VMFS 数据存储提供支持的逻辑卷。
	<code>ptbl</code> 检查并验证 VMFS 分区，例如 MBR 或 GPT。如果不存在分区，请确定是否应存在分区。
<code>-f --func</code>	要执行的功能，包括：
	<code>query</code> 列出模块所支持的功能。
	<code>check</code> 检查错误。
	<code>fix</code> 检查并修复错误。
	<code>dump</code> 收集元数据转储。
<code>-a --affinityChk</code>	为 VMFS6 包括关联性相关的检查和修复。
<code>-d --device</code>	<p>表示要检查的设备或磁盘。请务必提供为 VMFS 数据存储提供后备支持的设备分区的绝对路径。如果数据存储跨多个设备，请提供主数据区的 UUID。</p> <p>例如，<code>voma -m vmfs -f check -d /vmfs/devices/disks/naa.xxxx:x</code></p> <p>如果使用 <code>-x --extractDump</code> 命令，请输入多个设备路径，带分区限定符，以逗号分隔。输入的设备路径数等于跨区设备的数量。</p>
<code>-b --blockSize</code>	表示磁盘块大小。



表 9-9. VOMA 命令选项（续）

命令选项	描述
-s --logfile	指定日志文件的路径以输出结果。
-x --extractDump	使用 VOMA 提取收集的转储。
-D --dumpfile	指出转储文件以保存收集的元数据转储。
-v --version	显示 VOMA 的版本。
-h --help	显示 VOMA 命令的帮助消息。
-Y	指示在运行 VOMA 时没有使用 PE 表进行地址解析。
-Z --file	指示对已提取的设备文件运行 VOMA。

示例

从跨区卷收集元数据转储：

```
voma -m vmfs -f dump -d head_extent -D dump_filename
```

将收集的转储提取回跨区卷的设备：

```
voma -x dump_filename -d head_extent,extent_2,extent_3...extent_n
```

使用 VOMA 检查元数据一致性

该任务演示了如何使用 VOMA 检查 VMFS 元数据一致性。VOMA 可用于检查并修复 VMFS 数据存储或虚拟闪存资源的微小不一致问题。从 ESXi 主机上的 CLI 中运行 VOMA。

前提条件

关闭所有正在运行的虚拟机的电源，或者将这些虚拟机迁移至另一数据存储。

步骤

- 1 获取为想要检查的 VMFS 数据存储提供后备支持的设备的名称和分区编号。

```
#esxcli storage vmfs extent list
```

输出中的“Device Name”列和“Partition”列用于标识该设备。例如：

Volume Name	.....	Device Name	Partition
1TB_VMFS6	.....	naa.xxxx	3

- 2 检查 VMFS 错误。

提供为 VMFS 数据存储提供后备支持的设备分区的绝对路径，并为分区编号提供设备名称。例如：

```
# voma -m vmfs -f check -d /vmfs/devices/disks/naa.xxxx:x
```

输出将列出可能的错误。例如，以下输出表示检测信号地址无效。

```
XXXXXXXXXXXXXXXXXXXXXXXXX
Phase 2: Checking VMFS heartbeat region
  ON-DISK ERROR: Invalid HB address
Phase 3: Checking all file descriptors.
Phase 4: Checking pathname and connectivity.
Phase 5: Checking resource reference counts.

Total Errors Found:          1
```

## 配置 VMFS 指针块缓存

指针块也称为间接块，是包含 VMFS 文件块的文件系统资源。在 ESXi 主机上打开 vmdk 文件时，与该文件相关的指针块会存储在指针块缓存中。指针块缓存的大小是可配置参数。

指针块缓存是主机范围的缓存，与 VMFS 无关。从同一个 ESXi 主机访问的所有数据存储共享该缓存。

指针块缓存的大小由 /VMFS3/MinAddressableSpaceTB 和 /VMFS3/MaxAddressableSpaceTB 控制。可以配置每个 ESXi 主机上的最小大小和最大大小。

### /VMFS3/MinAddressableSpaceTB

最小值是系统向指针块缓存保证的最小内存量。例如，1 TB 的打开文件空间需要大约 4 MB 的内存。默认值为 10 TB。

### /VMFS3/MaxAddressableSpaceTB

此参数定义可在内存中缓存的指针块最大限制。默认值为 32 TB。最大值为 128 TB。通常，/VMFS3/MaxAddressableSpaceTB 参数的默认值足以满足需求。

但是，随着打开 vmdk 文件大小的增加，与这些文件相关的指针块数量也将随之增加。如果增加幅度会导致性能降低，您可以将此参数调整为其最大值，以便为指针块缓存提供更多空间。使指针块缓存的最大大小基于工作集或所需的活动指针块。

### 指针块逐出

/VMFS3/MaxAddressableSpaceTB 参数还控制指针块缓存的增长。当指针块缓存的大小接近配置的最大大小时，将启动指针块逐出过程。该机制将保留活动的指针块，而从缓存中移除非活动的或活动较少的块，以便可以重用空间。

要更改指针块缓存的值，请使用 vSphere Client 的**高级系统设置**对话框或 `esxcli system settings advanced set -o` 命令。

您可以使用 `esxcli storage vmfs pbcache` 命令获取有关指针块缓存大小的信息以及其他统计信息。这些信息可帮助您调整指针块缓存的最小和最大大小，以便您可以获得最佳性能。

## 获取 VMFS 指针块缓存的信息

您可以获取有关 VMFS 指针块缓存使用情况的信息。该信息可帮助您了解指针块缓存占用的空间大小。您还可以确定是否必须要调整指针块缓存的最小和最大大小。

### 前提条件

安装 ESXCLI。请参见《ESXCLI 入门》。要进行故障排除，请在 ESXi Shell 中运行 `esxcli` 命令。

### 步骤

- ◆ 要获取或重置指针块缓存统计信息，请使用以下命令：

```
esxcli storage vmfs pbcache
```

选项	描述
<code>get</code>	获取 VMFS 指针块缓存统计信息。
<code>reset</code>	重置 VMFS 指针块缓存统计信息。

### 示例： 获取指针块缓存统计信息

```
#esxcli storage vmfs pbcache get
Cache Capacity Miss Ratio: 0 %
Cache Size: 0 MiB
Cache Size Max: 132 MiB
Cache Usage: 0 %
Cache Working Set: 0 TiB
Cache Working Set Max: 32 TiB
Vmfs Heap Overhead: 0 KiB
Vmfs Heap Size: 23 MiB
Vmfs Heap Size Max: 256 MiB
```

## 更改指针块缓存的大小

可以调整指针块缓存的最小大小和最大大小。

**小心** 更改高级选项将被视为不受支持。通常，使用默认设置即可获得最佳结果。仅当 VMware 技术支持或知识库文章提供了具体指示时，才能更改高级选项。

### 步骤

- 1 浏览到主机。
- 2 单击配置选项卡。
- 3 在系统下，单击高级系统设置。

- 在“高级系统设置”中，选择相应项。

选项	描述
<code>VMFS3.MinAddressableSpaceTB</code>	VMFS 缓存保证支持的所有打开文件的最小大小。
<code>VMFS3.MaxAddressableSpaceTB</code>	VMFS 缓存在逐出开始前支持的所有打开文件的最大大小。

- 单击 **编辑** 按钮并更改值。
- 单击 **确定**。

### 示例：使用 `esxcli` 命令更改指针块缓存

还可以使用 `esxcli system settings advanced set -o` 来修改指针块缓存的大小。以下示例介绍如何将大小设置为最大值 128 TB。

- 要将 `/VMFS3/MaxAddressableSpaceTB` 的值更改为 128 TB，请输入以下命令：

```
# esxcli system settings advanced set -i 128 -o /VMFS3/MaxAddressableSpaceTB
```

- 要确认值是否设置正确，请输入以下命令：

```
# esxcli system settings advanced list -o /VMFS3/MaxAddressableSpaceTB
```

# 了解多路径和故障切换

# 10

要维持主机及其存储之间的持续连接，ESXi 必须支持多路径。借助多路径，您可以使用多条物理路径在主机和外部存储设备之间传输数据。

如果 SAN 网络中的任何元素（如适配器、交换机或电缆）发生故障，则 ESXi 可切换到另一个可行的物理路径。这种避免使用故障组件的路径切换过程称为路径故障切换。

除路径故障切换外，多路径还提供负载平衡。负载平衡是在多个物理路径中分配 I/O 负载的过程。负载平衡可以减少或消除潜在的瓶颈。

---

**注** 在路径故障切换发生时，虚拟机 I/O 最多能延迟 60 秒。通过这些延迟，SAN 可以在拓扑更改后稳定其配置。通常，主动-被动阵列上的 I/O 延迟时间可能会更长，而在主动-主动阵列上则更短。

---

本章讨论了以下主题：

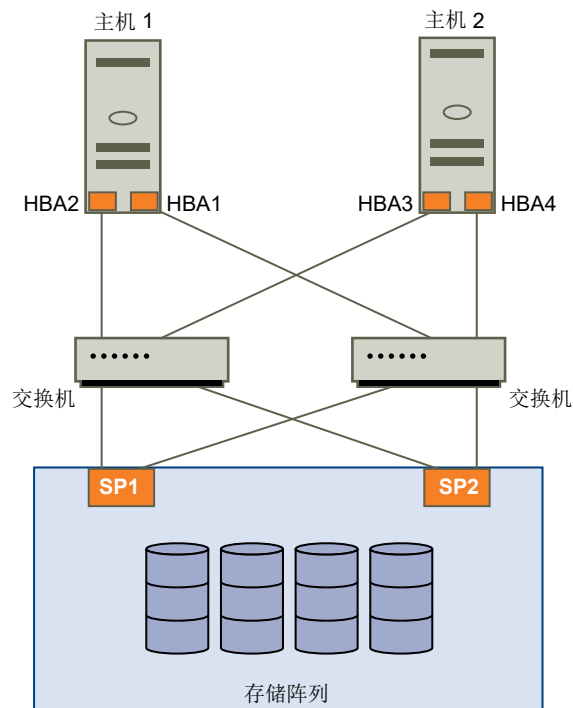
- 光纤通道故障切换
- 基于主机的 iSCSI 故障切换
- 基于阵列的 iSCSI 故障切换
- 路径故障切换和虚拟机
- 可插入存储架构和路径管理
- 查看和管理路径
- 使用声明规则
- 虚拟机 I/O 的调度队列

## 光纤通道故障切换

要支持多路径，主机通常具有两个或更多个可用的 HBA。此配置对 SAN 多路径配置进行补充。通常，SAN 多路径为 SAN 架构提供一个或多个交换机，并为存储阵列设备本身提供一个或多个存储处理器。

在下图中，可通过多个物理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 会接管并提供连接。一个 HBA 取代另一个 HBA 的过程称为 HBA 故障切换。

图 10-1. 多路径和光纤通道故障切换



同样，如果 SP1 发生故障或 SP1 与交换机之间的链路中断，SP2 会接管。SP2 会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 ESXi 支持 HBA 和 SP 故障切换。

## 基于主机的 iSCSI 故障切换

在设置 ESXi 主机的多路径和故障切换功能时，可以使用多个 iSCSI HBA，或者将多个网卡与软件 iSCSI 适配器结合使用。

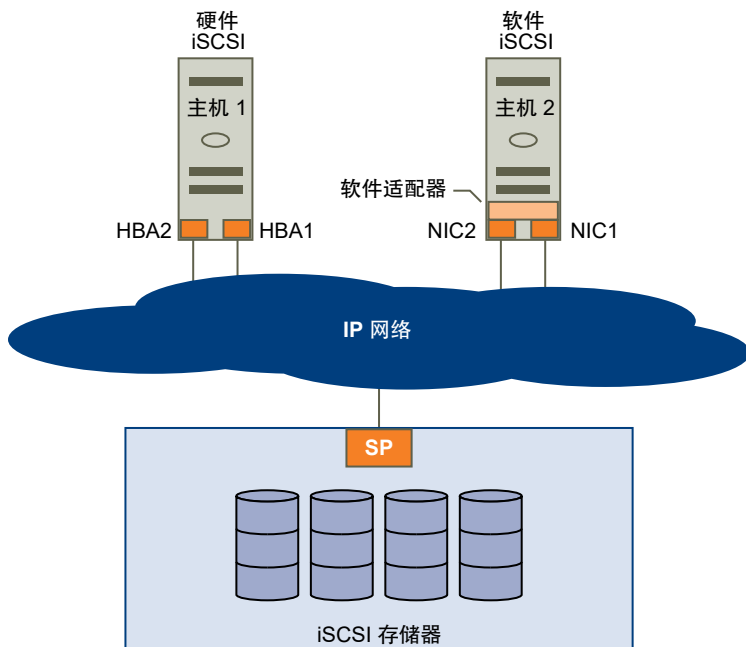
有关不同类型的 iSCSI 适配器的信息，请参见 [iSCSI 启动器](#)。

使用多路径时，需要注意特定事项。

- 将独立硬件适配器与同一主机中的软件 iSCSI 适配器或从属 iSCSI 适配器组合使用时，ESXi 不支持多路径。
- 支持在同一主机内的软件和从属适配器之间使用多路径。
- 在不同的主机上，可以混合使用从属适配器和独立适配器。

下图显示了对于不同类型的 iSCSI 启动器可能的多路径设置。

图 10-2. 基于主机的路径故障切换



## 硬件 iSCSI 和故障切换

对于硬件 iSCSI，主机通常具有两个或更多的硬件 iSCSI 适配器。主机使用适配器通过一个或多个交换机访问存储系统。或者，设置可能包括一个适配器和两个存储处理器，以便适配器可以使用不同的路径访问存储系统。

如图所示，主机 1 有两个硬件 iSCSI 适配器（HBA1 和 HBA2），这提供了两条到存储系统的物理路径。主机上的多路径插件，不论是 VMkernel NMP 还是任何第三方 MPP，默认情况下都可以访问这些路径。这些插件可以监控每条物理路径的运行状况。例如，如果 HBA1 或 HBA1 与网络之间的链路发生故障，多路径插件可以将路径切换到 HBA2。

## 软件 iSCSI 和故障切换

如图中的主机 2 所示，通过软件 iSCSI，您可以使用多张网卡为 iSCSI 连接提供故障切换和负载均衡功能。

多路径插件不能直接访问主机上的物理网卡。因此，对于此设置，您首先必须将每个物理网卡连接至单独的 VMkernel 端口。然后使用端口绑定技术将所有的 VMkernel 端口与软件 iSCSI 启动器相关联。连接到单独网卡的每个 VMkernel 端口将成为 iSCSI 存储堆栈及其存储感知多路径插件可使用的另一条路径。

有关为软件 iSCSI 配置多路径的信息，请参见[为 iSCSI 和 iSER 设置网络](#)。

## 基于阵列的 iSCSI 故障切换

有些 iSCSI 存储系统以透明的方式自动管理其 ESXi 端口的路径使用。

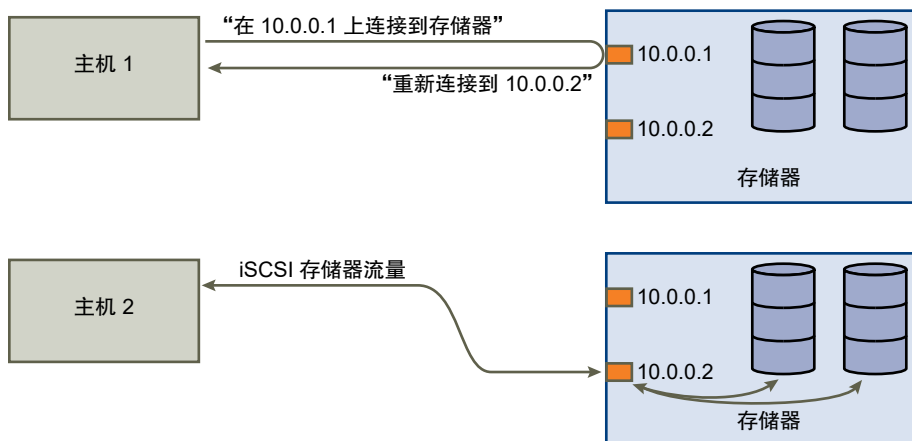
使用此类存储系统时，主机无法识别存储上的多个端口，且无法选择要连接的存储端口。这些系统提供了一个虚拟端口地址，供您主机用于初始通信。在初始通信期间，存储系统可以重定向主机，使其与存储系统上的另一端口进行通信。主机中的 iSCSI 启动器遵守此重新连接请求，并连接到系统上的其他端口。存储系统使用此技术将负载分散在多个可用端口中。

如果 ESXi 主机与其中某一端口断开连接，则会自动尝试重新连接存储系统的虚拟端口，且应重定向到可用的活动端口。此重新连接和重定向将会快速完成，且通常不会中断正在运行的虚拟机。这些存储系统也可请求 iSCSI 启动器重新连接到系统，以更改其连接的存储端口。这样可以最有效地使用多个端口。

“端口重定向”图显示了端口重定向示例。主机尝试连接到 10.0.0.1 虚拟端口。存储系统将此请求重定向到 10.0.0.2。主机与 10.0.0.2 连接并使用此端口进行 I/O 通信。

**注** 存储系统并不会始终重定向连接。10.0.0.1 的端口也可以用于通信。

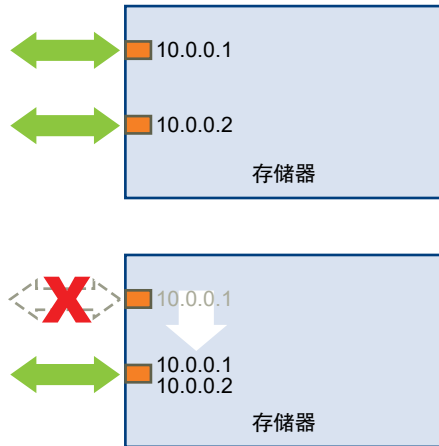
图 10-3. 端口重定向



如果存储系统上充当虚拟端口的端口变为不可用，则存储系统会将虚拟端口的地址重新分配给系统上的其他端口。“端口重新分配”显示了此类端口重新分配的示例。在这种情况下，虚拟端口 10.0.0.1 变为不可用，且存储系统会将虚拟端口 IP 地址重新分配给其他端口。第二个端口将响应两个地址。



图 10-4. 端口重新分配



借助这种基于阵列的故障切换形式，仅当使用 ESXi 主机上的多个端口时，才能拥有多个存储路径。这些路径为主动-主动路径。有关其他信息，请参见 [iSCSI 会话管理](#)。

## 路径故障切换和虚拟机

当 LUN 的活动路径从一个路径更改为另一个路径时，会发生路径故障切换。通常，发生路径故障切换是由于当前路径发生 SAN 组件故障。

路径发生故障后，存储 I/O 可能会暂停 30-60 秒，直到主机确定链路不可用并且执行故障切换为止。如果尝试显示主机、主机存储设备或适配器，操作可能会停止。虚拟机（其磁盘安装在 SAN 上）可能变得没有响应。故障切换后，I/O 恢复正常，且虚拟机继续运行。

当故障切换需要很长时间时，Windows 虚拟机可能会中断 I/O 并最终出现故障。为了避免出现此故障，请将 Windows 虚拟机的磁盘超时值至少设置为 60 秒。

## 在 Windows 客户机操作系统上设置超时

要在路径故障切换过程中避免中断，增加 Windows 客户机操作系统上的标准磁盘超时值。

此过程说明了如何使用 Windows 注册表更改超时值。

### 前提条件

备份 Windows 注册表。

### 步骤

- 1 选择开始 > 运行。
- 2 键入 `regedit.exe`，然后单击确定。
- 3 在左侧面板层次结构视图中，双击 `HKEY_LOCAL_MACHINE > 系统 > CurrentControlSet > 服务 > 磁盘`。
- 4 双击 `TimeOutValue`。

- 5 将数据值设置为 0x3c（十六进制）或 60（十进制），然后单击**确定**。

进行此更改后，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。

- 6 重新引导客户机操作系统以使更改生效。

## 可插入存储架构和路径管理

本主题介绍了 ESXi 存储多路径背后的关键概念。

### 可插入存储架构 (PSA)

为管理多路径，ESXi 使用特殊的 VMkernel 层（即，可插入存储架构 (PSA)）。PSA 是开放式模块框架，用于协调负责多路径操作的各种软件模块。这些模块包括 VMware 提供的通用多路径模块、NMP 和 HPP，以及第三方 MPP。

### 本机多路径插件 (NMP)

NMP 是 ESXi 默认提供的 VMkernel 多路径模块。NMP 将物理路径与特定存储设备相关联，并基于阵列类型提供了默认路径选择算法。NMP 是可扩展的，可管理其他子模块，即名为路径选择策略 (PSP) 和存储阵列类型策略 (SATP) 的子模块。PSP 和 SATP 可由 VMware 或第三方提供。

### 路径选择插件 (PSP)

PSP 是 VMware NMP 的子模块。PSP 负责为 I/O 请求选择物理路径。

### 存储阵列类型插件 (SATP)

SATP 是 VMware NMP 的子模块。SATP 负责特定于阵列的操作。SATP 可判断某个特定于阵列的路径的状态、执行路径激活，并检测任何路径错误。

### 多路径插件 (MPP)

PSA 提供了第三方用于创建自己的多路径插件 (MPP) 的 VMkernel API 集合。模块为特定存储阵列提供特定的负载平衡和故障切换功能。MPP 可以安装在 ESXi 主机上。它们可以作为 VMware 本机模块的补充插件运行，或作为其替代插件运行。

### VMware 高性能插件 (HPP)

HPP 替换了高速设备（例如 NVMe）的 NMP。HPP 可以提高安装在 ESXi 主机本地的超快速闪存设备的性能，并且是声明 NVMe 目标的默认插件。

为了支持多路径，HPP 使用路径选择方案 (PSS)。一个特定的 PSS 负责选择 I/O 请求的物理路径。

有关信息，请参见 [VMware 高性能插件和路径选择方案](#)。

### 声明规则

PSA 使用声明规则确定哪个插件拥有通往特定存储设备的路径。

表 10-1. 多路径缩写形式

缩写形式	定义
PSA	可插入存储架构
NMP	本机多路径插件。使用 SCSI 存储设备的通用 VMware 多路径模块。
PSP	路径选择插件。处理 SCSI 存储设备的路径选择。
SATP	存储阵列类型插件。处理给定 SCSI 存储阵列的路径故障切换。
MPP（第三方）	多路径插件。由第三方开发和提供的多路径模块。
HPP	由 VMware 提供的本机高性能插件。它与超快速本地和联网闪存设备（如 NVMe）配合使用。
PSS	路径选择方案。处理 NVMe 存储设备的多路径。

## 关于可插入存储架构

可插入存储架构 (PSA) 是一个开放式模块框架，用于协调负责多路径操作的各种软件模块。

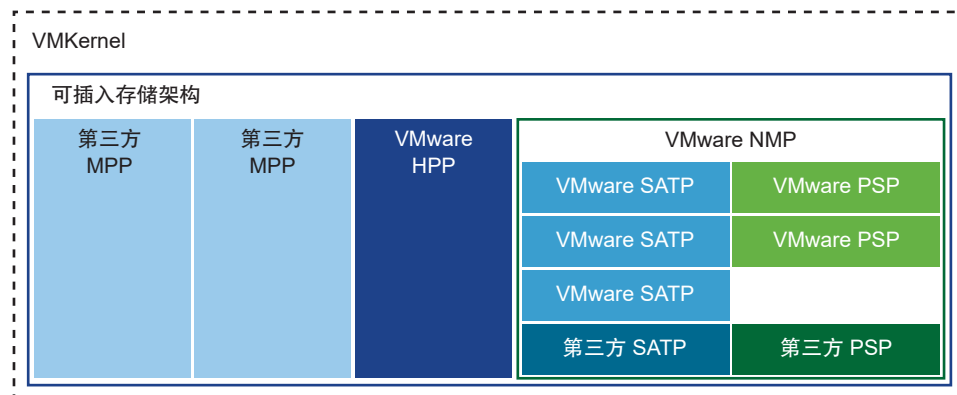
VMware 提供了通用本机多路径模块：VMware NMP 和 VMware HPP。此外，PSA 还提供了一组第三方开发人员可以使用的 VMkernel API。软件开发人员可以为特定的存储阵列创建自己的负载平衡和故障切换模块。这些第三方多路径模块 (MPP) 可以安装在 ESXi 主机上，并独立于 VMware 本机模块运行，或作为其替换项运行。

当协调 VMware 本机模块和安装的任何第三方 MPP 时，PSA 将执行以下任务：

- 加载和卸载多路径插件。
- 从特定插件隐藏虚拟机细节。
- 将特定逻辑设备的 I/O 请求路由到管理该设备的 MPP。
- 处理逻辑设备的 I/O 排队操作。
- 在虚拟机之间实现逻辑设备带宽共享。
- 处理物理存储 HBA 的 I/O 排队操作。
- 处理物理路径发现和移除。
- 提供逻辑设备和物理路径 I/O 统计信息。

如“可插入存储架构”图所示，多个第三方 MPP 可以与 VMware NMP 或 HPP 并行运行。安装时，第三方 MPP 可以替换本机模块的行为。MPP 可以控制指定存储设备的路径故障切换和负载平衡操作。

图 10-5. 可插入存储架构



## VMware 本机多路径插件

默认情况下，ESXi 提供名为本机多路径插件 (NMP) 的可扩展多路径模块。

一般来说，VMware NMP 支持 VMware 存储 HCL 上列出的所有存储阵列，并基于阵列类型提供默认的路径选择算法。NMP 将一组物理路径与特定存储设备或 LUN 关联。

对于其他多路径操作，NMP 将使用被称为 SATP 和 PSP 的子模块。NMP 向 SATP 委派为设备处理路径故障切换的特定详细信息。PSP 为设备处理路径选择。

通常情况下，NMP 会执行以下操作：

- 管理物理路径声明和取消声明。
- 注册和取消注册逻辑设备。
- 将物理路径与逻辑设备关联。
- 支持路径故障检测和修复。
- 处理逻辑设备的 I/O 请求：
  - 为请求选择最佳物理路径。
  - 执行处理路径故障和 I/O 命令重试的必要操作。
- 支持管理任务，如重置逻辑设备。

ESXi 自动为您使用的阵列安装相应的 SATP。您不需要获取或下载任何 SATP。

## VMware NMP I/O 流

当虚拟机向 NMP 管理的存储设备发出 I/O 请求时，将发生以下过程。

- 1 NMP 调用分配给此存储设备的 PSP。
- 2 PSP 将选择要通过其发出 I/O 的相应物理路径。
- 3 NMP 在 PSP 选择的 path 上发出 I/O 请求。
- 4 如果 I/O 操作成功，则 NMP 报告其完成。
- 5 如果 I/O 操作报告错误，则 NMP 调用适当的 SATP。

6 SATP 解释 I/O 命令错误，并在适当时激活非活动路径。

7 此时将调用 PSP 以选择要通过其发出 I/O 的新路径。

## 显示多路径模块

使用 `esxcli` 命令列出已加载到系统中的所有多路径模块。多路径模块管理将主机与存储相连的物理路径。这些模块包含 VMware 本机 NMP 和 HPP，以及任何第三方 MPP。

### 前提条件

安装 ESXCLI。请参见《ESXCLI 入门》。要进行故障排除，请在 ESXi Shell 中运行 `esxcli` 命令。

### 步骤

- ◆ 要列出多路径模块，请运行以下命令：

```
esxcli storage core plugin list --plugin-class=MP
```

### 结果

此命令通常显示 NMP，如果已加载，则显示 HPP 和 MASK\_PATH 模块。如果已加载任何第三方 MPP，也会将它们一并列出。

Plugin name	Plugin class
-----	-----
NMP	MP

有关此命令的详细信息，请参见《《ESXCLI 概念和示例》》和《《ESXCLI 参考指南》》文档。

## 显示 NMP 存储设备

使用 `esxcli` 命令列出由 VMware NMP 控制的所有存储设备，并显示与每个设备关联的 SATP 和 PSP 信息。

### 前提条件

安装 ESXCLI。请参见《ESXCLI 入门》。要进行故障排除，请在 ESXi Shell 中运行 `esxcli` 命令。

### 步骤

- ◆ 要列出所有存储设备，请运行以下命令：

```
esxcli storage nmp device list
```

使用 `--device` | `-d=device_ID` 参数筛选此命令的输出，以显示单个设备。

### 示例：显示 NMP 存储设备

```
# esxcli storage nmp device list
mpx.vmhbal:C0:T2:L0
  Device Display Name: Local VMware Disk (mpx.vmhbal:C0:T2:L0)
  Storage Array Type: VMW_SATP_LOCAL
  Storage Array Type Device Config: SATP VMW_SATP_LOCAL does not support device
configuration.
```

```

Path Selection Policy: VMW_PSP_FIXED
Path Selection Policy Device Config: {preferred=vmhba1:C0:T2:L0;current=vmhba1:C0:T2:L0}
Path Selection Policy Device Custom Config:
Working Paths: vmhba1:C0:T2:L0
Is USB: false

.....

eui.6238666462643332
Device Display Name: SCST_BIO iSCSI Disk (eui.6238666462643332)
Storage Array Type: VMW_SATP_DEFAULT_AA
Storage Array Type Device Config: {action_OnRetryErrors=off}
Path Selection Policy: VMW_PSP_FIXED
Path Selection Policy Device Config: {preferred=vmhba65:C0:T0:L0;current=vmhba65:C0:T0:L0}
Path Selection Policy Device Custom Config:
Working Paths: vmhba65:C0:T0:L0
Is USB: false

```

有关此命令的详细信息，请参见《《ESXCLI 概念和示例》》和《《ESXCLI 参考指南》》文档。

## 路径选择插件和策略

VMware 路径选择插件 (PSP) 负责选择 I/O 请求的物理路径。

插件是 VMware NMP 的子模块。NMP 根据设备类型为每个逻辑设备分配默认 PSP。可以替代默认 PSP。有关详细信息，请参见[更改路径选择策略](#)。

每个 PSP 启用并执行相应的路径选择策略。

### VMW\_PSP\_MRU - 最近使用 (VMware)

“最近使用 (VMware)”策略由 VMW\_PSP\_MRU 执行。它将选择在系统引导时间发现的第一个工作路径。当路径不可用时，主机会选择一个替代路径。当该路径可用时，主机不会恢复到原始路径。“最近使用”策略不会使用首选路径设置。这是大多数主动-被动存储设备的默认策略。

VMW\_PSP\_MRU 支持路径排名。要将排名设置为单个路径，请使用 `esxcli storage nmp psp generic pathconfig set` 命令。有关详细信息，请参见 VMware 知识库文章（网址为 <http://kb.vmware.com/kb/2003468>）和《ESXCLI 参考指南》文档。

### VMW\_PSP\_FIXED - 固定 (VMware)

此固定 (VMware) 策略由 VMW\_PSP\_FIXED 强制执行。该策略会使用指定的首选路径。如果没有分配首选路径，该策略将会选择在系统引导时间发现的第一个工作路径。如果首选路径不可用，主机将选择替代的可用路径。首选路径再次可用时，主机将返回到之前定义的首选路径。

“固定”是大多数主动-主动存储设备的默认策略。

### VMW\_PSP\_RR - 循环 (VMware)

VMW\_PSP\_RR 启用“循环 (VMware)”策略。“循环”是许多阵列的默认策略。它使用自动路径选择算法轮流选择配置的路径。

主动-主动和主动-被动阵列使用策略来实现不同 LUN 的路径之间的负载平衡。通过主动-被动阵列，该策略会使用活动路径。通过主动-主动阵列，该策略会使用可用路径。

默认情况下为该策略激活的延迟机制使其更具自适应性。为了实现更好的负载平衡结果，该机制将通过考虑以下路径特性动态选择最佳路径：

- I/O 带宽
- 路径延迟

要更改自适应延迟循环策略的默认参数或禁用延迟机制，请参见[更改延迟循环的默认参数](#)。

要为 VMW\_PSP\_RR 设置其他可配置参数，请使用 `esxcli storage nmp psp roundrobin` 命令。有关详细信息，请参见《ESXCLI 参考指南》文档。

## VMware SATP

存储阵列类型插件 (SATP) 负责阵列特定的操作。SATP 是 VMware NMP 的子模块。

ESXi 为 VMware 支持的各种类型阵列提供 SATP。ESXi 还支持非特定主动-主动、主动-被动、ALUA 和本地设备的默认 SATP。

每个 SATP 均满足特定存储阵列类的特殊特性。SATP 可以执行检测路径状态和激活非活动路径所需的阵列特定的操作。因此，NMP 模块本身可以使用多个存储阵列，而无需了解存储设备的特性。

通常，NMP 确定将哪个 SATP 用于特定存储设备，并将此 SATP 与该存储设备的物理路径相关联。SATP 实现的任务包括：

- 监控每个物理路径的运行状况。
- 报告每个物理路径的状态变化。
- 执行存储器故障切换所需的阵列特定的操作。例如，对于主动-被动设备，它可以激活被动路径。

ESXi 包括存储阵列的多个通用 SATP 模块。

### VMW\_SATP\_LOCAL

适用于直接连接的本地设备的 SATP。

从 vSphere 6.5 Update 2 版本开始，VMW\_SATP\_LOCAL 为本地设备（4K 本机格式的设备除外）提供多路径支持。无需早期 vSphere 版本中使用其他 SATP 这一过程，即可声明本地设备的多个路径。

VMW\_SATP\_LOCAL 支持 VMW\_PSP\_MRU 和 VMW\_PSP\_FIXED 路径选择插件，但不支持 VMW\_PSP\_RR。

### VMW\_SATP\_DEFAULT\_AA

适用于主动-主动阵列的通用 SATP。

### VMW\_SATP\_DEFAULT\_AP

适用于主动-被动阵列的通用 SATP。

### VMW\_SATP\_ALUA

适用于 ALUA 合规阵列的 SATP。

有关详细信息，请参见《VMware 兼容性指南》和《ESXCLI 参考指南》文档。

## 显示主机的 SATP

使用 `esxcli` 命令列出已加载到系统中的 VMware NMP SATP。显示关于 SATP 的信息。

### 前提条件

安装 ESXCLI。请参见《ESXCLI 入门》。要进行故障排除，请在 ESXi Shell 中运行 `esxcli` 命令。

### 步骤

- ◆ 要列出 VMware SATP，请运行以下命令：

```
esxcli storage nmp satp list
```

### 结果

对于每个 SATP，该输出会显示表明 SATP 支持的存储阵列或系统类型的信息。该输出还显示使用此 SATP 的任何 LUN 的默认 PSP。“描述”列中的 Placeholder (plugin not loaded) 表示未加载 SATP。

### 示例：显示主机的 SATP

```
# esxcli storage nmp satp list
Name                Default PSP        Description
VMW_SATP_MSA        VMW_PSP_MRU        Placeholder (plugin not loaded)
VMW_SATP_ALUA        VMW_PSP_MRU        Placeholder (plugin not loaded)
VMW_SATP_DEFAULT_AP VMW_PSP_MRU        Placeholder (plugin not loaded)
VMW_SATP_SVC        VMW_PSP_FIXED      Placeholder (plugin not loaded)
VMW_SATP_EQL        VMW_PSP_FIXED      Placeholder (plugin not loaded)
VMW_SATP_INV        VMW_PSP_FIXED      Placeholder (plugin not loaded)
VMW_SATP_EVA        VMW_PSP_FIXED      Placeholder (plugin not loaded)
VMW_SATP_ALUA_CX    VMW_PSP_RR         Placeholder (plugin not loaded)
VMW_SATP_SYMM        VMW_PSP_RR         Placeholder (plugin not loaded)
VMW_SATP_CX          VMW_PSP_MRU        Placeholder (plugin not loaded)
VMW_SATP_LSI        VMW_PSP_MRU        Placeholder (plugin not loaded)
VMW_SATP_DEFAULT_AA VMW_PSP_FIXED      Supports non-specific active/active arrays
VMW_SATP_LOCAL       VMW_PSP_FIXED      Supports direct attached devices
```

有关此命令的详细信息，请参见《《ESXCLI 概念和示例》》和《《ESXCLI 参考指南》》文档。

## VMware 高性能插件和路径选择方案

VMware 提供高性能插件 (HPP) 来提高 ESXi 主机上的存储设备的性能。

HPP 替换高速设备（例如 NVMe）的 NMP。HPP 是声明 NVMe-oF 目标的默认插件。在 ESXi 中，将模拟 NVMe-oF 目标并将其作为 SCSI 目标提供给用户。HPP 仅支持主动/主动和隐式 ALUA 目标。

从 vSphere 7.0 Update 2 开始，HPP 将成为本地 NVMe 和 SCSI 设备的默认插件，但可以将其替换为 NMP。



HPP 支持	vSphere 7.0 Update 2 及更高版本
存储设备	本地 NVMe 和 SCSI 共享 NVMe（仅限主动/主动和隐式 ALUA 目标）
多路径	是
第二级插件	否
SCSI-3 永久预留	否
带软件模拟的 4Kn 设备	是

## 路径选择方案

为了支持多路径，HPP 在为 I/O 请求选择物理路径时使用路径选择方案 (PSS)。

可以使用 vSphere Client 或 `esxcli` 命令更改默认路径选择机制。

有关在 vSphere Client 中配置路径机制的信息，请参见[更改路径选择策略](#)。要使用 `esxcli` 命令进行配置，请参见 [ESXi esxcli HPP 命令](#)。

ESXi 支持以下路径选择机制。

### 固定

在此方案中，为 I/O 请求使用指定的首选路径。如果没有分配首选路径，主机将选择在引导时发现的第一个工作路径。如果首选路径不可用，主机将选择替代的可用路径。首选路径再次可用时，主机将返回到之前定义的首选路径。

将 **FIXED** 配置为路径选择机制时，请选择首选路径。

### LB-RR（负载均衡 - 循环）

这是 HPP 声明的设备的默认方案。在当前路径上传输指定数量的字节或 I/O 后，该方案将使用循环算法选择路径。

要配置 **LB-RR** 路径选择机制，请指定以下属性：

- **IOPS** 指示路径上要用作设备路径切换条件的 I/O 计数。
- **字节** 指示路径上要用作设备路径切换条件的字节计数。

### LB-IOPS（负载均衡 - IOPS）

在当前路径上传输指定数量的 I/O（默认值为 1000）后，系统将选择具有最少未完成 I/O 数量的最佳路径。

配置此机制时，请指定 **IOPS** 参数，以指示路径上要用作设备路径切换条件的 I/O 计数。

### LB-BYTES（负载均衡 - 字节）

在当前路径上传输指定数量的字节（默认值为 10 MB）后，系统将选择具有最少未完成字节数量的最佳路径。

要配置此机制，请使用**字节**参数指示路径上要用作设备路径切换条件的字节计数。

### 负载均衡 - 延迟 (LB-Latency)

为了实现更好的负载均衡结果，该机制将通过考虑以下路径特性动态选择最佳路径：

- **延迟评估时间**参数指示必须以怎样的时间间隔评估路径延迟（以毫秒为单位）。
- **每个路径的采样 I/O** 参数控制为了计算路径延迟必须在每个路径上发出多少采样 I/O。

## HPP 最佳做法

要在高速存储设备上获得最快的吞吐量，请遵循这些建议。

- 使用支持 HPP 的 vSphere 版本。
- 将 HPP 用于本地 NVMe 和 SCSI 设备以及 NVMe-oF 设备。
- 如果使用 NVMe over Fibre Channel 设备，请遵循对光纤通道存储的常规建议。请参见第 3 章 [将 ESXi 与光纤通道 SAN 配合使用](#)。
- 如果使用 NVMe-oF，请勿混合使用传输类型来访问同一个命名空间。
- 使用 NVMe-oF 命名空间时，请确保将活动路径提供给主机。在发现活动路径之前，无法注册命名空间。
- 配置您的虚拟机以使用 VMware 准虚拟控制器。请参见《vSphere 虚拟机管理》文档。
- 设置延迟敏感阈值。
- 如果设备的 I/O 工作负载很大一部分由单台虚拟机承担，请考虑将 I/O 分散到多个虚拟磁盘。将磁盘附加到虚拟机中单独的虚拟控制器。

否则，I/O 吞吐量可能会由于负责在特定虚拟存储控制器上处理 I/O 的 CPU 内核饱和而受到限制。

有关仅支持 NGUID ID 格式的 NVMe 设备的设备标识符信息，请参见具有 [NGUID 设备标识符的 NVMe 设备](#)。

## 启用高性能插件和路径选择方案

高性能插件 (HPP) 是声明本地 NVMe 和 SCSI 设备以及 NVMe-oF 目标的默认插件。如有必要，可以将其替换为 NMP。在 vSphere 版本 7.0 Update 1 及更低版本中，NMP 仍为本地 NVMe 和 SCSI 设备的默认插件，但您可以将其替换为 HPP。

使用 `esxcli storage core claimrule add` 命令在 ESXi 主机上启用 HPP 或 NMP。

要运行 `esxcli storage core claimrule add`，您可以使用 ESXi Shell 或 vSphere CLI。有关更多信息，请参见《《ESXCLI 入门》》和《《ESXCLI 参考指南》》。

本主题中的示例说明了如何启用 HPP 以及设置路径选择方案 (PSS)。

---

**注** PXE 引导的 ESXi 主机上不支持启用 HPP。

---

### 前提条件

设置 VMware NVMe 存储环境。有关详细信息，请参见第 8 章 [关于 VMware NVMe 存储](#)。

## 步骤

- 1 通过运行 `esxcli storage core claimrule add` 命令创建 HPP 声明规则。

使用以下方法之一添加声明规则。

方法	描述
基于 NVMe 控制器型号	<pre>esxcli storage core claimrule add --type vendor --nvme-controller-model</pre> <p>例如，</p> <pre>esxcli storage core claimrule add --rule 429 --type vendor --nvme-controller-model "ABCD*" --plugin HPP</pre>
基于 PCI 供应商 ID 和子供应商 ID	<pre>esxcli storage core claimrule add --type vendor --pci-vendor-id --pci-sub-vendor-id</pre> <p>例如，</p> <pre>esxcli storage core claimrule add --rule 429 --type vendor --pci-vendor-id 8086 --pci-sub-vendor-id 8086 --plugin HPP.</pre>

- 2 配置 PSS。

使用下列方法之一。

方法	描述
根据设备 ID 设置 PSS	<pre>esxcli storage hpp device set</pre> <p>例如，</p> <pre>esxcli storage hpp device set --device=device --pss=FIXED --path=preferred path</pre>
根据供应商/型号设置 PSS	<p>将 <code>--config-string</code> 选项与 <code>esxcli storage core claimrule add</code> 命令一起使用。</p> <p>例如，</p> <pre>esxcli storage core claimrule add -r 914 -t vendor -V vendor -M model -P HPP --config-string "pss=LB-Latency,latency-eval-time=40000"</pre>

- 3 重新引导主机以使更改生效。

## 设置延迟敏感阈值

在为存储设备使用 HPP 时，设置该设备的延迟敏感阈值，使 I/O 可以绕过 I/O 调度程序。

默认情况下，ESXi 通过 I/O 调度程序传递每个 I/O。但是，使用调度程序可能会创建内部队列，这不是利用高速存储设备的有效方法。

您可以配置延迟敏感阈值并启用直接提交机制，帮助 I/O 绕过调度程序。若启用此机制，则 I/O 会通过 HPP 直接从 PSA 传递到设备驱动程序。

要顺利执行直接提交，观察到的平均 I/O 延迟必须低于您指定的延迟阈值。如果 I/O 延迟超过延迟阈值，系统会停止直接提交，并暂时恢复到使用 I/O 调度程序。当平均 I/O 延迟再次下降至低于延迟阈值时会恢复直接提交。

可以为一系列 HPP 声明的设备设置延迟阈值。使用供应商和型号对、控制器型号或者 PCIe 供应商 ID 和子供应商 ID 对，设置延迟阈值。

步骤

1 通过运行以下命令，设置设备的延迟敏感阈值：

**esxcli storage core device latencythreshold set -t value in milliseconds**

使用下列选项之一。

选项	示例
供应商/型号	为具有指定供应商和型号的所有设备设置延迟敏感阈值参数： <b>esxcli storage core device latencythreshold set -v 'vendor1' -m 'model1' -t 10</b>
NVMe 控制器型号	为具有指定控制器型号的所有 NVMe 设备设置延迟敏感阈值： <b>esxcli storage core device latencythreshold set -c 'controller_model1' -t 10</b>
PCIe 供应商/子供应商 ID	为 PCIe 供应商 ID 为 0x8086 和 PCIe 子供应商 ID 为 0x8086 的设备设置延迟敏感阈值。 <b>esxcli storage core device latencythreshold set -p '8086' -s '8086' -t 10</b>

2 确认延迟阈值已设置：

**esxcli storage core device latencythreshold list**

Device	Latency Sensitive Threshold
-----	-----
naa.55cd2e404c1728aa	0 milliseconds
naa.500056b34036cdfd	0 milliseconds
naa.55cd2e404c172bd6	50 milliseconds

3 监控延迟敏感阈值的状态。检查以下条目的 VMkernel 日志：

- Latency Sensitive Gatekeeper turned on for device 设备. Threshold of XX msec is larger than max completion time of YYY msec
- Latency Sensitive Gatekeeper turned off for device 设备. Threshold of XX msec is exceeded by command completed in YYY msec

ESXi esxcli HPP 命令

您可以使用 ESXi Shell 或 vSphere CLI 命令配置和监控高性能插件。

有关 esxcli 命令概况信息，请参见《ESXCLI 入门》；有关该命令的使用详细信息，请参见《ESXCLI 参考指南》。

命令	描述	选项
<code>esxcli storage hpp path list</code>	列出当前由高性能插件声明的路径。	<code>-d --device=device</code> 显示特定设备的信息。 <code>-p --path=path</code> 将输出限制到特定路径。
<code>esxcli storage hpp device list</code>	列出当前受高性能插件控制的设备。	<code>-d --device=device</code> 显示特定设备。

命令	描述	选项
esxcli storage hpp device set	为 HPP 设备配置设置。	<p>-B --bytes=<i>long</i> 路径上的最大字节数，超过此值后，将切换该路径。</p> <p>--cfg-file 使用新设置更新配置文件和运行时。如果设备已由另一个 PSS 声明，则在应用到运行时配置时，请忽略任何错误。</p> <p>-d --device=<i>device</i> 要在上面运行的 HPP 设备。使用设备报告的任何 UID。必需。</p> <p>-I --iops=<i>long</i> 路径上的最大 IOPS 数，超过此值后，将切换该路径。</p> <p>-T --latency-eval-time=<i>long</i> 控制必须以怎样的时间间隔（以毫秒为单位）评估路径的延迟。</p> <p>-L --mark-device-local=<i>Bool</i> 将 HPP 设置为是否将设备视为本地设备。</p> <p>-M --mark-device-ssd=<i>bool</i> 指定 HPP 是否将设备视为 SSD。</p> <p>-p --path=<i>str</i> 要设置为设备首选路径的路径。</p> <p>-P --pss=<i>pss_name</i> 要分配给设备的路径选择方案。如果未指定值，系统将选择使用默认值。有关路径选择方案的说明，请参见 <a href="#">VMware 高性能插件和路径选择方案</a>。选项包括：</p> <ul style="list-style-type: none"> <li>■ <b>FIXED</b></li> <li>使用 -p --path=<i>str</i> 子选项设置首选路径。</li> <li>■ <b>LB-Bytes</b></li> <li>使用 -B --bytes=<i>long</i> 子选项指定输入。</li> <li>■ <b>LB-IOPS</b></li> <li>使用 -I --iops=<i>long</i> 子选项指定输入。</li> <li>■ <b>LB-Latency</b></li> <li>子选项包括： <ul style="list-style-type: none"> <li>-T --latency-eval-time=<i>long</i></li> <li>-S --sampling-ios-per-path=<i>long</i></li> </ul> </li> <li>■ <b>LB-RR 默认值</b></li> <li>子选项包括： <ul style="list-style-type: none"> <li>-B --bytes=<i>long</i></li> <li>-I --iops=<i>long</i></li> <li>-S --sampling-ios-per-path=<i>long</i> 控制为了计算路径延迟必须在每个路径上发出多少采样 I/O。</li> </ul> </li> </ul>

命令	描述	选项
		-U --use-ano=bool 要在用于在此设备上发出 I/O 的一组活动路径中包含未优化路径，请将该选项设置为 true。否则，请将该选项设置为 false。
esxcli storage hpp device usermarkedssd list	列出已由用户标记或取消标记为 SSD 的设备。	-d --device=device 将输出限制到特定设备。

## 查看和管理路径

启动 ESXi 主机或重新扫描存储适配器时，主机会发现它可以使用的存储设备的所有物理路径。基于上一组声明规则，主机将确定 NMP、HPP 或 MPP 中的哪个多路径模块拥有通往特定设备的路径。

拥有设备的模块需要负责管理该设备的多路径支持。默认情况下，主机将每隔五分钟执行一次周期性路径评估，并将尚未声明的路径分配给相应的模块。

对于由 NMP 模块管理的路径，将使用第二组声明规则。这些规则将 SATP 和 PSP 模块分配给每个存储设备，并确定要应用哪个存储阵列类型策略和路径选择策略。

使用 vSphere Client 查看分配给特定存储设备的存储阵列类型策略和路径选择策略。还可以检查此存储设备的所有可用路径的状态。如果需要，可以使用客户端更改默认的路径选择策略。

要更改默认多路径模块或 SATP，请使用 vSphere CLI 修改声明规则。

可以在 [使用声明规则](#) 中找到有关修改声明规则的部分信息。

## 查看存储设备路径

查看主机用于特定存储设备的多路径策略，以及该存储设备的所有可用路径的状态。

### 步骤

- 1 在 vSphere Client 中，导航到 ESXi 主机。
- 2 单击 **配置** 选项卡。
- 3 在 **存储** 下，单击 **存储设备**。
- 4 选择要查看其路径的存储设备。
- 5 单击 **属性** 选项卡，然后查看拥有设备的模块，例如 NMP 或 HPP。

在“多路径策略”下，还可以查看分配给设备的“路径选择策略”和“存储阵列类型策略”（如果适用）。

- 6 单击**路径**选项卡，可查看存储设备的所有可用路径以及每个路径的状态。其中会显示以下路径状态信息：

状态	描述
活动 (I/O)	工作路径或当前传输数据的多个路径。
备用	处于非活动状态的路径。如果活动路径失败，他们可以变成工作状态并开始传输 I/O。
已禁用	管理员禁用的路径。
不活动	不可再用于处理 I/O 的路径。物理中级别故障或阵列配置错误可能会导致此状态。

如果正在使用**固定**路径策略，就可以看到哪一条路径是首选路径。首选路径的“首选”列标有一个星号(\*)。

## 查看数据存储路径

检查连接到支持您的 VMFS 数据存储的存储设备的路径。

### 步骤

- 1 在 vSphere Client 中，导航到数据存储。
- 2 单击**配置**选项卡。
- 3 单击**连接和多路径**。
- 4 选择一个主机来查看其设备的多路径详细信息。
- 5 在多路径策略下，查看拥有设备的模块，例如 NMP。您也可以查看分配给设备的路径选择策略和存储阵列类型策略。

例如，您可能会看到以下内容：

路径选择策略	首选路径
存储阵列类型策略	VMW_SATP_LOCAL
所有者插件	NMP

- 6 在“路径”下，查看设备路径和每个路径的状态。其中会显示以下路径状态信息：

状态	描述
活动 (I/O)	工作路径或当前传输数据的多个路径。
备用	处于非活动状态的路径。如果活动路径失败，他们可以变成工作状态并开始传输 I/O。
已禁用	管理员禁用的路径。
不活动	不可再用于处理 I/O 的路径。物理中级别故障或阵列配置错误可能会导致此状态。

如果正在使用**固定**路径策略，就可以看到哪一条路径是首选路径。首选路径的“首选”列标有一个星号(\*)。

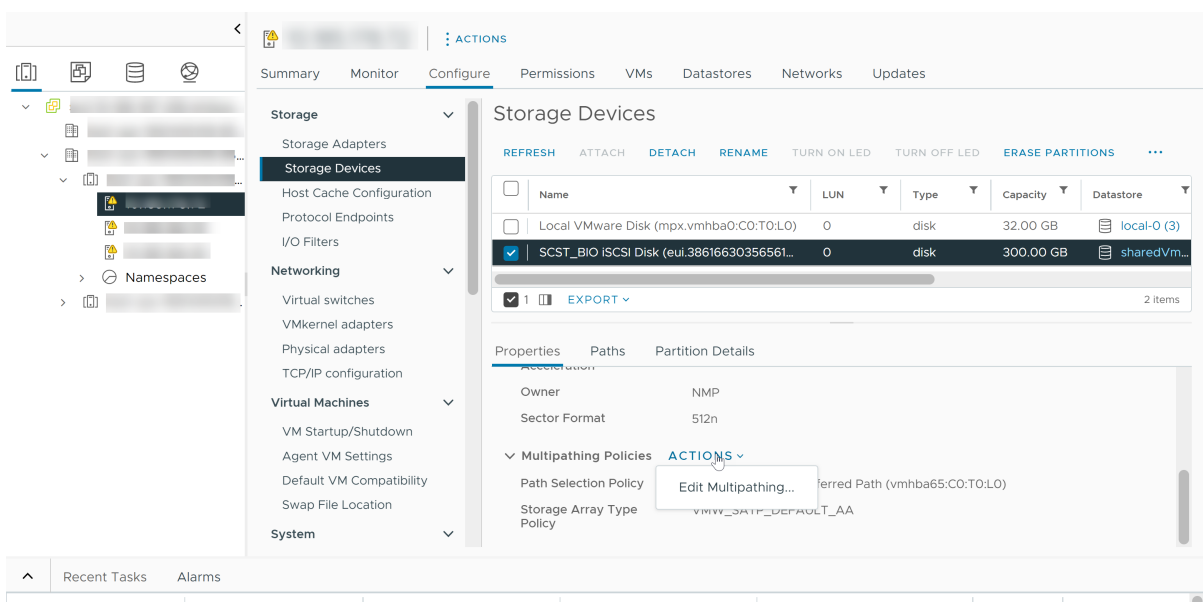


## 更改路径选择策略

通常，不需要更改 ESXi 主机用于特定存储设备的默认多路径设置。如果要进行任何更改，可以使用**编辑多路径策略**对话框修改路径选择策略。您也可以使用该对话框更改基于 SCSI 的协议端点的多路径。

### 步骤

- 1 在 vSphere Client 中，导航到 ESXi 主机。
- 2 单击**配置**选项卡。
- 3 在**存储**下，单击**存储设备**或**协议端点**。
- 4 选择要更改其路径的项目，然后单击**属性**选项卡。
- 5 在“多路径策略”下，从**操作**菜单中选择**编辑多路径**。



- 6 选择路径策略并配置其设置。这些选项会发生更改，具体取决于所使用存储设备的类型。
  - 有关 SCSI 设备的路径策略的信息，请参见[路径选择插件和策略](#)。
  - 有关 NVMe 设备的路径机制的信息，请参见 [VMware 高性能插件和路径选择方案](#)。
- 7 要保存设置并退出对话框，请单击**确定**。

## 更改延迟循环的默认参数

在 ESXi 主机上，默认情况下，会为循环路径选择策略激活延迟机制。该机制会考虑 I/O 带宽和路径延迟，以选择最佳的 I/O 路径。使用延迟机制时，循环策略可以动态选择最佳路径，并实现更好的负载均衡效果。

可以使用 `esxcli` 命令更改延迟机制的默认参数，也可以禁用该机制。

### 前提条件

将路径选择策略设置为“循环”。请参见[更改路径选择策略](#)。

步骤

- 1 使用以下命令配置延迟机制。

```
esxcli storage nmp psp roundrobin deviceconfig set --type=latency --device=device ID
```

该命令采用以下参数：

参数	描述
-S --num-sampling-cycles= <i>sampling value</i>	当 --type 设置为 latency 时，此参数控制要用于计算每个路径平均延迟的 I/O 数。此参数的默认值为 16。
-T --latency-eval-time= <i>time in ms</i>	当 --type 设置为 latency 时，此参数控制路径延迟的更新频率。默认值为 3 分钟。

- 2 验证延迟循环及其参数是否正确配置。

```
esxcli storage nmp psp roundrobin deviceconfig get --device=device ID
```

或

```
esxcli storage nmp device list --device=device ID
```

以下示例输出显示了路径的配置：

```
Path Selection Policy: VMW_PSP_RR
Path Selection Policy Device Config:
{policy=latency,latencyEvalTime=180000,samplingCycles=16,curSamplingCycle=16,useANO=0;
CurrentPath=vmhba1:C0:T0:L0: NumIOsPending=0,latency=0}
```

后续步骤

要禁用延迟机制，请在主机的“高级系统设置”中，将 Misc.EnablePSPLatencyPolicy 参数更改为 0。

禁用存储路径

由于维护或其他原因，可以暂时禁用路径。

使用“路径”面板禁用路径。您有以下几种方法访问路径面板，即从数据存储、存储设备、适配器或者 Virtual Volumes 协议端点视图访问。

步骤

- 1 在 vSphere Client 中，导航到 ESXi 主机。
- 2 单击配置选项卡。
- 3 在存储下，请单击以下项目之一：
  - 存储适配器
  - 存储设备

## ■ 协议端点

- 4 在右侧窗格中，选择您想要禁用其路径的项目、适配器、存储设备或协议端点，然后单击**路径**选项卡。
- 5 选择要禁用的路径，然后单击**禁用**。

该路径的状态将更改为“已禁用”。

## 使用声明规则

声明规则会确定哪个多路径模块拥有通往特定存储设备的路径。他们也定义了主机向设备提供的多路径支持的类型。

声明规则在主机的 `/etc/vmware/esx.conf` 文件中列出。

这些规则分为以下几类：

### 核心声明规则

这些声明规则会确定 NMP、HPP 或第三方 MPP 中的哪个多路径模块为特定设备提供声明。

### SATP 声明规则

根据设备类型，这些规则会分配特定 SATP 子模块，为设备提供特定于供应商的多路径管理。

您可以使用 `esxcli` 命令添加或更改核心和 SATP 声明规则。通常情况下，您要添加声明规则以加载第三方 MPP，或从主机中隐藏 LUN。如果特定设备的默认设置不足，则可能需要更改声明规则。

有关可用于管理 PSA 声明规则的命令的详细信息，请参见《ESXCLI 入门》。

有关存储阵列和相应的 SATP 和 PSP 的列表，请参见以下文档的“存储/SAN”一节：vSphere Compatibility Guide。

## 多路径注意事项

管理存储多路径插件和声明规则时，需考虑一些特定的注意事项。

以下注意事项可帮助您使用多路径：

- 如果声明规则未将任何 SATP 分配给设备，则 iSCSI 或 FC 设备的默认 SATP 将是 `VMW_SATP_DEFAULT_AA`。默认 PSP 是 `VMW_PSP_FIXED`。
- 当系统搜索 SATP 规则以查找给定设备的 SATP 时，会首先搜索驱动程序规则。如果没有匹配项，则搜索供应商/型号规则，最后搜索传输规则。如果未出现匹配项，NMP 将选择设备的默认 SATP。
- 如果已将 `VMW_SATP_ALUA` 分配给某特定存储设备，但该设备不能识别 ALUA，则该设备将没有任何匹配的声明规则。设备将由默认 SATP 根据设备的传输类型进行声明。
- 由 `VMW_SATP_ALUA` 声明的所有设备的默认 PSP 是 `VMW_PSP_MRU`。`VMW_PSP_MRU` 选择由 `VMW_SATP_ALUA` 报告的主动/优化的路径，如果没有主动/优化的路径，则选择主动/未优化的路径。在有更好的路径可用 (MRU) 之前，将一直使用此路径。例如，如果 `VMW_PSP_MRU` 当前正在使用主动/未优化的路径，那么，当主动/优化的路径变为可用时，`VMW_PSP_MRU` 会将当前路径切换到主动/优化的路径。

- 尽管默认情况下通常为 ALUA 阵列选择 VMW\_PSP\_MRU，但某些 ALUA 存储阵列仍需使用 VMW\_PSP\_FIXED。要查看存储阵列是否需要使用 VMW\_PSP\_FIXED，请参阅《VMware 兼容性指南》或联系存储供应商。如果将 VMW\_PSP\_FIXED 用于 ALUA 阵列，则除非明确指定首选路径，否则 ESXi 主机将选择最优的工作路径，并将其指定为默认首选路径。如果主机选择的路径不可用，主机将选择替代的可用路径。但如果您明确指定首选路径，则无论其为何状态，都将始终为首选路径。
- 默认情况下，PSA 声明规则 101 会屏蔽 Dell 阵列伪设备。除非您要取消屏蔽这些设备，否则请不要删除此规则。

## 列出主机的多路径声明规则

使用 `esxcli` 命令可列出可用的多路径声明规则。

声明规则指出是 NMP、HPP 还是第三方 MPP 管理给定的物理路径。每个声明规则基于以下参数标识一组路径：

- 供应商/型号字符串
- 传输，如 SATA、IDE、光纤通道
- 适配器、目标或 LUN 位置
- 设备驱动程序（如 Mega-RAID）

### 步骤

- ◆ 通过运行 `esxcli storage core claimrule list --claimrule-class=MP` 命令来列出多路径声明规则。

如果不使用 `claimrule-class` 选项，则隐含 MP 规则类。

### 示例： `esxcli storage core claimrule list` 命令的示例输出

Rule	Class	Rule	Class	Type	Plugin	Matches
MP		10	runtime	vendor	HPP	vendor=NVMe model=*
MP		10	file	vendor	HPP	vendor=NVMe model=*
MP		50	runtime	transport	NMP	transport=usb
MP		51	runtime	transport	NMP	transport=sata
MP		52	runtime	transport	NMP	transport=ide
MP		53	runtime	transport	NMP	transport=block
MP		54	runtime	transport	NMP	transport=unknown
MP		101	runtime	vendor	MASK_PATH	vendor=DELL model=Universal Xport
MP		101	file	vendor	MASK_PATH	vendor=DELL model=Universal Xport
MP		200	runtime	vendor	MPP_1	vendor=NewVend model=*
MP		200	file	vendor	MPP_1	vendor=NewVend model=*
MP		201	runtime	location	MPP_2	adapter=vmhba41 channel=* target=* lun=*
MP		201	file	location	MPP_2	adapter=vmhba41 channel=* target=* lun=*
MP		202	runtime	driver	MPP_3	driver=megaraid
MP		202	file	driver	MPP_3	driver=megaraid
MP		65535	runtime	vendor	NMP	vendor=* model=*

此示例表示以下内容：

- NMP 声明与使用 USB、SATA、IDE 和块 SCSI 传输的存储设备相连的所有路径。
- 已添加 HPP、MPP\_1、MPP\_2 和 MPP\_3 规则，以便模块可声明指定设备。例如，HPP 声明供应商为 NVMe 的所有设备。无论实际供应商如何，所有由内置的 nvme 驱动程序处理的设备均已声明。MPP\_1 模块声明与任何型号 NewVend 存储阵列相连的所有路径。
- 可以使用 MASK\_PATH 模块隐藏主机中未使用的设备。默认情况下，PSA 声明规则 101 会屏蔽供应商字符串为 DELL 且型号字符串为 Universal Xport 的 Dell 阵列伪设备。
- 输出中的“Rule Class”列说明声明规则的类别。类别可以是 MP（多路径插件）、Filter 或 VAAI。
- “Class”列显示已定义的规则和已加载的规则。“Class”列中的 file 参数表示规则已定义。runtime 参数表示规则已被加载到系统中。为了激活用户定义的声明规则，必须存在两个具有相同规则编号的行，一行是具有 file 参数的规则，另一行是具有 runtime 的规则。几个默认系统定义的声明规则只有一行，其“Class”为 runtime。您无法修改这些规则。
- 默认规则 65535 将所有空闲路径分配给 NMP。请勿删除此规则。

## 添加多路径声明规则

使用 `esxcli` 命令将多路径 PSA 声明规则添加到系统上的一组声明规则。为了激活新声明规则，请先定义规则，然后将其加载到系统中。

下面是一些添加 PSA 声明规则的示例：

- 加载新的第三方 MPP 并且必须定义此模块声明的路径。
- 您必须启用本机 HPP。

---

**警告** 您不能在两个不同的插件声明路径指向同一设备的情况下创建规则。尝试创建这些声明规则失败，`vmkernel.log` 中会显示警告。

---

### 前提条件

安装 ESXCLI。请参见《ESXCLI 入门》。要进行故障排除，请在 ESXi Shell 中运行 `esxcli` 命令。

### 步骤

- 1 要定义新声明规则，请使用以下命令：

```
esxcli storage core claimrule add
```

该命令采用以下选项：

选项	描述
<b>-A --adapter=&lt;adapter&gt;</b>	要使用的路径适配器。仅当 <code>--type</code> 为 <code>location</code> 时才有效。
<b>-u --autoassign</b>	根据声明特性添加声明规则。规则编号并非必填项。
<b>-C --channel=&lt;channel&gt;</b>	要使用的路径通道。仅当 <code>--type</code> 为 <code>location</code> 时才有效。

选项	描述
<b>-c --claimrule-class=&lt;cl&gt;</b>	此操作中要使用的声明规则类。您可以指定 MP（默认值）、Filter 或 VAAI。 要为新的阵列配置硬件加速，请添加两个声明规则：一个用于 VAAI 筛选器，另一个用于 VAAI 插件。有关详细说明，请参见 <a href="#">添加硬件加速声明规则</a> 。
<b>-d --device=&lt;device_uid&gt;</b>	设备 UID。仅当 --type 为 device 时才有效。
<b>-D --driver=&lt;driver&gt;</b>	要使用的路径 HBA 的驱动程序。仅当 --type 为 driver 时才有效。
<b>-f --force</b>	强制声明规则在任何情况下均忽略有效性检查并安装规则。
<b>--force-reserved</b>	替代对预留规则 ID 范围的保护。 预留声明规则是指 ID 小于 100 的规则。您可以使用它们将本地设备重新分配给特定的插件，例如将 NVMe 设备重新分配给 HPP。
<b>--if-unset=&lt;str&gt;</b>	如果此高级用户变量未设置为 1，请运行该命令。
<b>-i --iqn=&lt;iscsi_name&gt;</b>	目标的 iSCSI 限定名称。仅当 --type 为 target 时才有效。
<b>-L --lun=&lt;lun_id&gt;</b>	路径的 LUN。仅当 --type 为 location 时才有效。 LUN ID 不得高于高级配置选项 /Disk/MaxLUN 的值。
<b>-M --model=&lt;model&gt;</b>	要使用的路径型号。仅当 --type 为 vendor 时才有效。 有效值是 SCSI 查询字符串中的型号字符串的值。在每台设备上运行 <code>vicfg-scsidevs &lt;conn_options&gt; -l</code> 以查看型号字符串值。
<b>-P --plugin=&lt;plugin&gt;</b>	要使用的 PSA 插件。值为 NMP、MASK_PATH 或 HPP。第三方还可以提供自己的 PSA 插件。必需。
<b>-r --rule=&lt;rule_ID&gt;</b>	要使用的规则 ID。规则 ID 指示声明规则的评估顺序。用户定义的声明规则按照从 101 开始的数字顺序进行评估。 您可以运行 <code>esxcli storage core claimrule list</code> 以确定哪些规则 ID 可用。
<b>-T --target=&lt;target&gt;</b>	要使用的路径目标。仅当 --type 为 location 时才有效。
<b>-R --transport=&lt;transport&gt;</b>	要使用的路径传输。仅当 --type 为 transport 时才有效。支持以下值。 <ul style="list-style-type: none"> <li>■ block—块存储</li> <li>■ fc—光纤通道</li> <li>■ iscsivendor—iSCSI</li> <li>■ iscsi—当前未使用</li> <li>■ ide—IDE 存储</li> <li>■ sas—SAS 存储</li> <li>■ sata—SATA 存储</li> <li>■ usb—USB 存储</li> <li>■ parallel—并行</li> <li>■ fcoe—FCoE</li> <li>■ unknown</li> </ul>

选项	描述
<b>-t --type=&lt;type&gt;</b>	用于操作的匹配类型。以下为有效值。必需。 <ul style="list-style-type: none"> <li>■ vendor</li> <li>■ location</li> <li>■ driver</li> <li>■ transport</li> <li>■ device</li> <li>■ target</li> </ul>
<b>-V --vendor=&lt;vendor&gt;</b>	要使用的路径供应商。仅当 --type 为 vendor 时才有效。 有效值是 SCSI 查询字符串中的供应商字符串的值。在每台设备上运行 <code>vicfg-scsidevs &lt;conn_options&gt; -l</code> 以查看供应商字符串值。
<b>--wwnn=&lt;wwnn&gt;</b>	目标的全球节点编号。
<b>--wwpn=&lt;wwpn&gt;</b>	目标的全球端口编号。
<b>-a --xcopy-use-array-values</b>	使用阵列报告的值来构建要发送到存储阵列的 XCOPY 命令。这仅适用于 VAAI 声明规则。
<b>-s --xcopy-use-multi-segs</b>	当发出 XCOPY 请求时，使用多个分段。仅当指定 --xcopy-use-array-values 时才有效。
<b>-m --xcopy-max-transfer-size</b>	使用不同于阵列所报告的传输大小时的最大数据传输大小 (MB)。仅当指定 --xcopy-use-array-values 时才有效。
<b>-k --xcopy-max-transfer-size-kib</b>	使用不同于阵列所报告的传输大小时 XCOPY 命令的最大传输大小 (KiB)。仅当指定 --xcopy-use-array-values 时才有效。

## 2 要将新声明规则加载到系统中，请使用以下命令：

```
esxcli storage core claimrule load
```

此命令会将 `esx.conf` 配置文件中所有新创建的多路径声明规则加载到 VMkernel 中。该命令没有选项。

## 3 要应用加载的声明规则，请使用以下命令：

```
esxcli storage core claimrule run
```

该命令采用以下选项：

选项	描述
<b>-A --adapter=&lt;adapter&gt;</b>	如果 --type 为 location，表示在其上运行声明规则的路径的 HBA 名称。要从所有适配器的路径上运行声明规则，请忽略此选项。
<b>-C --channel=&lt;channel&gt;</b>	如果 --type 为 location，表示在其上运行声明规则的路径的 SCSI 通道号的值。要在含任何通道号的路径上运行声明规则，请忽略此选项。
<b>-c --claimrule-class=&lt;cl&gt;</b>	此操作中要使用的声明规则类。
<b>-d --device=&lt;device_uid&gt;</b>	设备 UID。
<b>-L --lun=&lt;lun_id&gt;</b>	如果 --type 为 location，表示在其上运行声明规则的路径的 SCSI LUN 值。要在含任何 LUN 的路径上运行声明规则，请忽略此选项。

选项	描述
<b>-p --path=&lt;path_uid&gt;</b>	如果 <b>--type</b> 为 <b>path</b> ，则此选项指示唯一路径标识符 (UID) 或在其上运行声明规则的路径的运行名称。
<b>-T --target=&lt;target&gt;</b>	如果 <b>--type</b> 为 <b>location</b> ，表示在其上运行声明规则的路径的 SCSI 目标编号值。要在含任何目标编号的路径上运行声明规则，请忽略此选项。
<b>-t --type=&lt;location path all&gt;</b>	要执行的声明类型。默认情况下，将使用 <b>all</b> ，这意味着在不限于特定路径或 SCSI 地址的情况下运行声明规则。有效值为 <b>location</b> 、 <b>path</b> 和 <b>all</b> 。
<b>-w --wait</b>	<p>仅当还使用了 <b>--type all</b> 时才可以使用此选项。</p> <p>如果包括此选项，则声明将等待路径进入稳定状态后，再运行声明操作。在这种情况下，只有在开始声明过程之前系统上的所有路径都出现后，系统才会启动声明过程。</p> <p>声明过程开始后，该命令在设备注册完成后返回。</p> <p>如果在声明或发现过程中添加或移除路径，则此选项可能无法正常工作。</p>

## 示例：定义多路径声明规则

在以下示例中，将添加并加载编号为 500 的规则。此规则为 NMP 插件声明型号字符串为 NewMod 以及供应商字符串为 NewVend 的所有路径。

```
# esxcli storage core claimrule add -r 500 -t vendor -V NewVend -M NewMod -P NMP
```

```
# esxcli storage core claimrule load
```

运行 **esxcli storage core claimrule list** 命令之后，可以看到新的声明规则显示在列表中。

以下输出表明编号为 500 的声明规则已加载到系统中，并处于活动状态。

Rule	Class	Rule	Class	Type	Plugin	Matches
...	...	...	...	...	...	...
MP		500	runtime	vendor	NMP	vendor=NewVend model=NewMod
MP		500	file	vendor	NMP	vendor=NewVend model=NewMod

## 删除多路径声明规则

使用 **esxcli** 命令将多路径 PSA 声明规则从系统上的一组声明规则中移除。

### 前提条件

安装 ESXCLI。请参见《ESXCLI 入门》。要进行故障排除，请在 ESXi Shell 中运行 **esxcli** 命令。

### 步骤

- 1 从声明规则集中删除声明规则。

```
esxcli storage core claimrule remove
```

**注** 默认情况下，PSA 声明规则 101 会屏蔽 Dell 阵列伪设备。除非您要取消屏蔽这些设备，否则请不要删除此规则。



该命令采用以下选项：

选项	描述
<code>-c --claimrule-class=&lt;str&gt;</code>	表示声明规则类（MP、筛选器和 VAAI）。
<code>-P --plugin=&lt;str&gt;</code>	表示插件。
<code>-r --rule=&lt;long&gt;</code>	表示规则 ID。

此步骤将从“文件”类中移除声明规则。

## 2 从系统移除声明规则。

```
esxcli storage core claimrule load
```

此步骤将从“运行时”类中移除声明规则。

## 屏蔽路径

可以阻止主机访问存储设备或 LUN，或阻止其使用指向某个 LUN 的单个路径。使用 `esxcli` 命令可屏蔽路径。屏蔽路径时，请创建声明规则，该规则将 MASK\_PATH 插件分配给指定路径。

### 前提条件

安装 ESXCLI。请参见《ESXCLI 入门》。要进行故障排除，请在 ESXi Shell 中运行 `esxcli` 命令。

### 步骤

#### 1 检查下一个可用的规则 ID 是多少。

```
esxcli storage core claimrule list
```

用于屏蔽路径的声明规则的规则 ID 在 101 - 200 范围内。如果此命令显示规则 101 和 102 已经存在，您可以指定 103 来添加规则。

#### 2 通过为 MASK\_PATH 插件创建新声明规则，将该插件分配给某个路径。

```
esxcli storage core claimrule add -P MASK_PATH
```

#### 3 将 MASK\_PATH 声明规则加载到系统中。

```
esxcli storage core claimrule load
```

#### 4 验证是否正确添加了 MASK\_PATH 声明规则。

```
esxcli storage core claimrule list
```

#### 5 如果存在已屏蔽路径的声明规则，则将规则移除。

```
esxcli storage core claiming unclaim
```

#### 6 运行路径声明规则。

```
esxcli storage core claimrule run
```

## 结果

在将 MASK\_PATH 插件分配给路径之后，路径状态变为不相关，且不再由主机进行维护。因此，显示屏蔽路径信息的命令可能将路径状态显示为失效。

## 示例：屏蔽 LUN

在此示例中，屏蔽目标 T1 和 T2 上通过存储适配器 vmhba2 和 vmhba3 访问的 LUN 20。

```
1 #esxcli storage core claimrule list

2 #esxcli storage core claimrule add -P MASK_PATH -r 109 -t location -A vmhba2 -C 0 -T 1 -L 20
  #esxcli storage core claimrule add -P MASK_PATH -r 110 -t location -A vmhba3 -C 0 -T 1 -L 20
  #esxcli storage core claimrule add -P MASK_PATH -r 111 -t location -A vmhba2 -C 0 -T 2 -L 20
  #esxcli storage core claimrule add -P MASK_PATH -r 112 -t location -A vmhba3 -C 0 -T 2 -L 20

3 #esxcli storage core claimrule load

4 #esxcli storage core claimrule list

5 #esxcli storage core claiming unclaim -t location -A vmhba2
  #esxcli storage core claiming unclaim -t location -A vmhba3

6 #esxcli storage core claimrule run
```

## 取消路径屏蔽

当需要主机访问被屏蔽的存储设备时，请取消对该设备路径的屏蔽。

**注** 使用设备属性（例如设备 ID 或供应商）运行取消声明操作时，不会取消声明 MASK\_PATH 插件声明的路径。MASK\_PATH 插件不会跟踪其声明的路径的任何设备属性。

### 前提条件

安装 ESXCLI。请参见《ESXCLI 入门》。要进行故障排除，请在 ESXi Shell 中运行 `esxcli` 命令。

### 步骤

- 1 删除 MASK\_PATH 声明规则。  
`esxcli storage core claimrule remove -r rule#`
- 2 验证是否正确删除了此声明规则。  
`esxcli storage core claimrule list`

- 3 将配置文件中的路径声明规则重新加载到 VMkernel 中。

```
esxcli storage core claimrule load
```

- 4 对于每个指向被屏蔽的存储设备的路径，运行 **esxcli storage core claiming unclaim** 命令。

例如：

```
esxcli storage core claiming unclaim -t location -A vmhba0 -C 0 -T 0 -L 149
```

- 5 运行路径声明规则。

```
esxcli storage core claimrule run
```

## 结果

您的主机即可访问之前被屏蔽的存储设备。

## 定义 NMP SATP 规则

NMP SATP 声明规则定义存储设备由哪些 SATP 来管理。通常情况下，您可以使用为存储设备提供的默认 SATP。如果默认设置不充足，请使用 **esxcli** 命令更改特定设备的 SATP。

当安装特定存储阵列的第三方 SATP 时，可能需要创建一条 SATP 规则。

### 前提条件

安装 ESXCLI。请参见《ESXCLI 入门》。要进行故障排除，请在 ESXi Shell 中运行 **esxcli** 命令。

### 步骤

- 1 要添加特定 SATP 的声明规则，请运行 **esxcli storage nmp satp rule add** 命令。该命令采用以下选项。

选项	描述
<b>-b --boot</b>	此规则是在引导时添加的系统默认规则。请勿修改 <code>esx.conf</code> 或添加到主机配置文件。
<b>-c --claim-option=string</b>	添加 SATP 声明规则时设置声明选项字符串。
<b>-e --description=string</b>	添加 SATP 声明规则时设置声明规则描述。
<b>-d --device=string</b>	添加 SATP 声明规则时设置设备。设备规则与供应商/型号规则和驱动程序规则互斥。
<b>-D --driver=string</b>	添加 SATP 声明规则时设置驱动程序字符串。驱动程序规则与供应商/型号规则互斥。
<b>-f --force</b>	强制声明规则在任何情况下均忽略有效性检查并安装规则。
<b>-h --help</b>	显示帮助消息。
<b>-M --model=string</b>	添加 SATP 声明规则时设置型号字符串。供应商/型号规则与驱动程序规则互斥。
<b>-o --option=string</b>	添加 SATP 声明规则时设置选项字符串。

选项	描述
<code>-P --psp=string</code>	设置 SATP 声明规则的默认 PSP。
<code>-O --psp-option=string</code>	设置 SATP 声明规则的 PSP 选项。
<code>-s --satp=string</code>	为其添加新规则的 SATP。
<code>-R --transport=string</code>	添加 SATP 声明规则时设置声明传输类型字符串。
<code>-t --type=string</code>	添加 SATP 声明规则时设置声明类型。
<code>-V --vendor=string</code>	添加 SATP 声明规则时设置供应商字符串。供应商/型号规则与驱动程序规则互斥。

**注** 当搜索 SATP 规则以查找给定设备的 SATP 时，NMP 会首先搜索驱动程序规则。如果没有匹配项，则搜索供应商/型号规则，最后搜索传输规则。如果仍然没有匹配项，NMP 将选择设备的默认 SATP。

## 2 重新引导主机。

### 示例：定义 NMP SATP 规则

以下示例命令将分配 VMW\_SATP\_INV 插件以管理供应商字符串为 NewVend 和型号字符串为 NewMod 的存储阵列。

```
# esxcli storage nmp satp rule add -V NewVend -M NewMod -s VMW_SATP_INV
```

运行 `esxcli storage nmp satp list -s VMW_SATP_INV` 命令时，您可以看到新规则已位于 VMW\_SATP\_INV 规则列表中。

## 虚拟机 I/O 的调度队列

vSphere 在默认情况下提供了一种机制，可为每个虚拟机文件创建调度队列。每个文件（例如 .vmdk）可获得自己的带宽控制。

此机制可确保特定虚拟机文件的 I/O 进入其自己单独的队列中，且可避免干扰其他文件的 I/O。

默认情况下会启用此功能。可以使用 vSphere Client 或 `esxcli` 命令禁用或重新启用此功能。

### 在 vSphere Client 中编辑按文件 I/O 调度

`VMkernel.Boot.isPerFileSchedModelActive` 高级参数控制 VMFS 和 NFS 3 数据存储上的按文件 I/O 调度机制。在 ESXi 主机上，该机制默认处于启用状态。可以使用 **高级系统设置** 对话框禁用该机制。

如果您关闭每个文件 I/O 调度模型，主机将恢复到旧版调度机制。旧版调度机制只为每个虚拟机和存储设备对维护一个 I/O 队列。虚拟机及其虚拟磁盘之间的所有 I/O 都将移到此队列。因此，来自不同虚拟磁盘的 I/O 在共享带宽时可能会相互干扰并影响彼此的性能。

**注** 如果您为高速本地设备配置了 HPP 插件和延迟敏感阈值参数，请不要禁用每个文件调度。禁用每个文件调度可能会导致出现不可预测的行为。

**步骤**

- 1 在 vSphere Client 中，导航到 ESXi 主机。
- 2 单击配置选项卡。
- 3 在系统下，单击高级系统设置。
- 4 编辑 **VMkernel.Boot.isPerFileSchedModelActive** 参数的值。

选项	描述
无效	禁用每个文件调度机制。
True (默认)	重新启用每个文件调度机制。

- 5 重新引导主机以使更改生效。

## 使用 esxcli 命令启用或禁用按文件 I/O 调度

可以使用 esxcli 命令更改 I/O 调度功能。默认情况下会启用此功能。

**前提条件**

安装 ESXCLI。请参见《ESXCLI 入门》。要进行故障排除，请在 ESXi Shell 中运行 esxcli 命令。

**步骤**

- ◆ 要启用或禁用按文件 I/O 调度，请运行以下命令：

选项	描述
<b>esxcli system settings kernel set -s isPerFileSchedModelActive -v FALSE</b>	禁用按文件 I/O 调度
<b>esxcli system settings kernel set -s isPerFileSchedModelActive -v TRUE</b>	启用按文件 I/O 调度

裸设备映射 (RDM) 为虚拟机提供了一种机制，用于直接访问物理存储子系统上的 LUN。

以下主题包含 RDM 的相关信息，并且说明如何创建和管理 RDM。

本章讨论了以下主题：

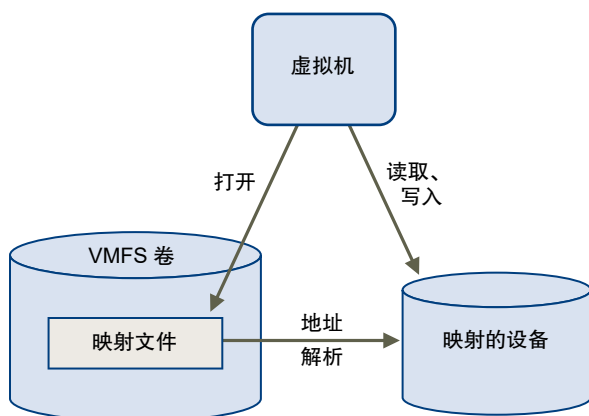
- 关于裸设备映射
- 裸设备映射特性
- 使用 RDM 创建虚拟机
- 管理映射的 LUN 的路径
- 具有 RDM 的虚拟机必须忽略 SCSI INQUIRY 缓存

## 关于裸设备映射

RDM 是独立 VMFS 卷中的映射文件，它可充当裸物理存储设备的代理。通过 RDM，虚拟机可以直接访问和使用存储设备。RDM 包含用于管理和重定向对物理设备进行磁盘访问的元数据。

该文件具有直接访问物理设备的一些优点，但保留了 VMFS 中虚拟磁盘的一些优点。因此，它可以将 VMFS 易管理性结合到裸设备访问中。

图 11-1. 裸设备映射



通常，针对大多数虚拟磁盘存储都使用 VMFS 数据存储。在某些情况下，您可能会使用裸 LUN 或者位于 SAN 中的逻辑磁盘。

例如，在以下情况下，您可能会使用裸 LUN 处理 RDM：

- 当在虚拟机中运行 SAN 快照或其他分层应用程序时。RDM 通过使用 SAN 特有的功能来启用备份卸载系统。
- 在任何跨物理主机的 MSCS 集群情况下，例如虚拟到虚拟集群以及物理到虚拟集群。在此情况下，集群数据和仲裁磁盘配置为 RDM 而非共享 VMFS 上的虚拟磁盘。

将 RDM 视为从 VMFS 卷到裸 LUN 的符号链接。映射使 LUN 显示为 VMFS 卷中的文件。在虚拟机配置中引用 RDM 而非裸 LUN。RDM 包含对裸 LUN 的引用。

RDM 有两种可用兼容模式：

- 在虚拟兼容模式下，RDM 的行为与虚拟磁盘文件类似。RDM 可以使用快照。
- 在物理兼容模式下，通过 RDM，那些需要较低级别控制的应用程序可以直接访问 SCSI 设备。

## 裸设备映射的优点

RDM 具有许多优点，但并非在每种情况下都适用。通常，对于易管理性而言，虚拟磁盘文件优于 RDM。但是，当需要裸设备时，必须使用 RDM。

RDM 提供几个好处。

### 用户友好的持久名称

为所映射的设备提供用户友好的名称。使用 RDM 时，不必通过设备名称引用设备。可以根据映射文件的名称来引用设备，例如：

```
/vmfs/volumes/myVolume/myVMDirectory/myRawDisk.vmdk
```

### 动态名称解析

为各个映射设备存储唯一的标识信息。VMFS 将每个 RDM 与其当前的 SCSI 设备相关联，而不考虑由于适配器硬件更改、路径更改、设备重定位等所引起的服务器物理配置的变化。

### 分布式文件锁定

使为裸 SCSI 设备使用 VMFS 分布式锁定成为可能。当位于不同服务器上的两个虚拟机试图访问同一 LUN 时，RDM 上的分布式锁定使其能够安全使用共享裸 LUN 而不会丢失数据。

### 文件权限

使文件权限成为可能。在文件打开时，强制执行映射文件权限，以保护映射的卷。

### 文件系统操作

通过将映射文件作为代理，可以实现使用文件系统实用程序处理映射的卷。对普通文件有效的大部分操作都可应用于映射文件，并且可重定向在映射设备上进行操作。

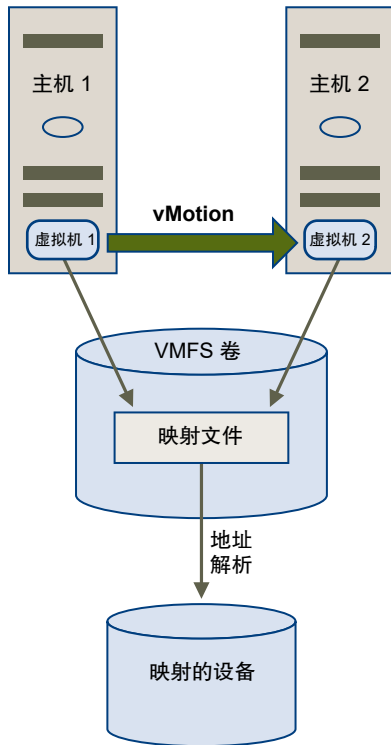
### 快照

使在映射的卷上使用虚拟机快照成为可能。在物理兼容模式下使用 RDM 时，快照不可用。

### vMotion

允许通过 vMotion 迁移虚拟机。映射文件可充当代理，允许 vCenter Server 使用与迁移虚拟磁盘文件相同的机制迁移虚拟机。

图 11-2. 使用裸设备映射的虚拟机的 vMotion



## SAN 管理代理

使在虚拟机内运行某些 SAN 管理代理成为可能。与此相似，可以在虚拟机内运行需要使用硬件特定 SCSI 命令访问设备的任何软件。这种软件称为基于 SCSI 目标的软件。使用 SAN 管理代理时，需要为 RDM 选择物理兼容模式。

## N-Port ID 虚拟化 (NPIV)

令使用 NPIV 技术成为可能，通过该技术，单一光纤通道 HBA 端口可使用多个全球端口名称 (WWPN) 向光纤通道架构注册。通过此功能，HBA 端口可显示为多个虚拟端口，每个端口均有其自身的 ID 和虚拟端口名称。因此，虚拟机就可声明其中每个虚拟端口，并将其用于所有 RDM 流量。

**注** 只能将 NPIV 用于具备 RDM 磁盘的虚拟机。

VMware 与存储器管理软件的供应商合作，确保他们的软件能够在包括 ESXi 的环境下正常工作。下面是一些这种类型的应用程序：

- SAN 管理软件
- 存储资源管理 (SRM) 软件
- 快照软件
- 复制软件



此类软件将物理兼容模式用于 RDM，以便能够直接访问 SCSI 设备。

各种管理产品都可以完美地集中运行（而不是在 ESXi 计算机上运行），而其他产品则可以在虚拟机中良好运行。VMware 未正式认可这些应用程序，也未提供兼容性列表。要了解在 ESXi 环境中是否支持某个 SAN 管理应用程序，请与该 SAN 管理软件的提供商联系。

## RDM 注意事项和限制

使用 RDM 时，有一些注意事项和限制。

- RDM 不可用于直接连接的块设备或某些 RAID 设备。RDM 使用 SCSI 序列号标识映射设备。由于块设备和某些直连 RAID 设备不能导出序列号，因此不能将其用于 RDM。
- 如果在物理兼容模式下使用 RDM，则不能使用磁盘快照。物理兼容模式允许虚拟机管理自己基于存储的快照或镜像操作。  
虚拟机快照可用于虚拟兼容模式的 RDM。
- 无法映射到磁盘分区。RDM 需将映射设备作为整个 LUN。
- 如果使用 vMotion 迁移带有 RDM 的虚拟机，应确保使 RDM 的 LUN ID 在所有参与的 ESXi 主机间保持一致。

## 裸设备映射特性

RDM 是 VMFS 卷中管理映射设备元数据的一种特殊映射文件。管理软件将映射文件视作普通磁盘文件，可用于常规文件系统操作。对于虚拟机，存储器虚拟化层将映射设备显示为虚拟 SCSI 设备。

映射文件中元数据的主要内容包括映射设备的位置（名称解析）、映射设备的锁定状况和权限，等等。

## RDM 虚拟兼容模式和物理兼容模式

可以在虚拟兼容或物理兼容模式中使用 RDM。虚拟模式指定映射设备的完整虚拟化。物理模式指定映射设备的最小 SCSI 虚拟化，实现了 SAN 管理软件的最大灵活性。

在虚拟模式下，VMkernel 只将 READ 和 WRITE 发送到映射设备。映射的设备对于客户机操作系统而言正如虚拟磁盘文件在 VMFS 卷中一样。隐藏真实的硬件特性。如果在虚拟模式中使用裸磁盘，您能够认识到 VMFS 的诸多优点，例如，用于保护数据的高级文件锁定和可简化开发流程的快照等。虚拟模式比物理模式在存储硬件上的移植性更强，表现行为如同虚拟磁盘文件。

在物理模式下，VMkernel 将所有 SCSI 命令传递给设备。唯一例外是对 REPORT LUN 命令进行了虚拟化，以便 VMkernel 可将 LUN 隔离到所属的虚拟机中。其他情况下，底层硬件的所有物理特性都将显现。物理模式对于在虚拟机中运行 SAN 管理代理或其他基于 SCSI 目标的软件非常有用。物理模式还允许虚拟到物理的集群构建，更具成本效益和高可用性。

VMFS5 和 VMFS6 支持虚拟和物理模式下大小超过 2 TB 的 RDM 磁盘。

## 动态名称解析

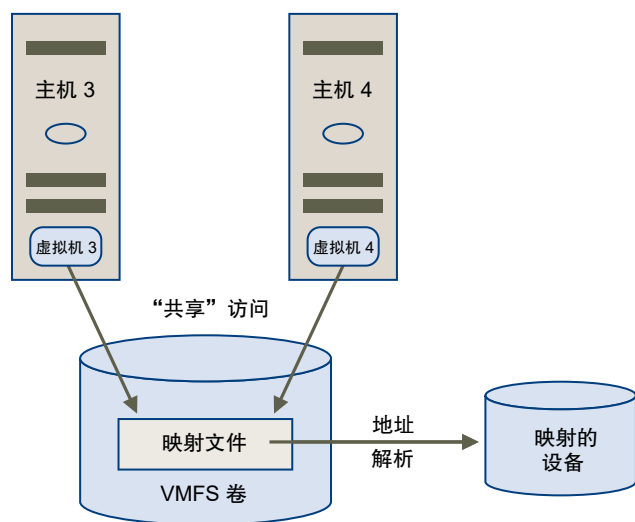
RDM 文件支持动态名称解析（裸设备的路径发生更改时）。

所有映射的存储设备都由 VMFS 进行唯一标识，并且标识存储在其内部数据结构中。裸设备路径中的任何更改（如光纤通道交换机发生故障或添加新的 HBA）都可以造成设备名称发生变化。动态名称解析可解析这些更改，并将原始设备与其新名称自动关联起来。

## 虚拟机集群的裸设备映射

对需要访问同一裸 LUN 以实施故障切换场景的虚拟机集群执行 RDM。其设置与访问同一虚拟磁盘文件的虚拟机集群的设置相同，但 RDM 会替换虚拟磁盘文件。

图 11-3. 从集群虚拟机进行访问



## 比较可用的 SCSI 设备访问模式

访问基于 SCSI 的存储设备的方法包括 VMFS 数据存储上的虚拟磁盘文件、虚拟模式 RDM 和物理模式 RDM。

下表提供了对不同模式下可用功能的比较。

表 11-1. 虚拟磁盘和原始设备映射的可用功能

ESXi 功能	虚拟磁盘文件	虚拟模式 RDM	物理模式 RDM
SCSI 命令已传递	否	否	是 REPORT LUNs 未传递
vCenter Server 支持	是	是	是
快照	是	是	否
分布式锁定	是	是	是

表 11-1. 虚拟磁盘和原始设备映射的可用功能（续）

ESXi 功能	虚拟磁盘文件	虚拟模式 RDM	物理模式 RDM
集群	仅限机箱内集群	机箱内集群 机箱间集群	物理到虚拟集群 机箱间集群
基于 SCSI 目标的软件	否	否	是

为集群的机箱内集群类型使用虚拟磁盘文件。如果计划将机箱内集群重新配置为机箱间集群，请为机箱内集群采用虚拟模式 RDM。

## 使用 RDM 创建虚拟机

授予虚拟机对原始 SAN LUN 的直接访问权限时，会创建位于 VMFS 数据存储并指向 LUN 的 RDM 磁盘。您可创建 RDM 作为新虚拟机的初始磁盘，或将其添加到现有虚拟机中。创建 RDM 时，可以指定要映射的 LUN 以及要用来放置 RDM 的数据存储。

尽管 RDM 磁盘文件与常规虚拟磁盘文件的扩展名均为 .vmdk，但 RDM 仅包含映射信息。实际虚拟磁盘数据直接存储在 LUN 上。

该过程假设正在创建新虚拟机。有关信息，请参见《vSphere 虚拟机管理》文档。

### 步骤

- 1 创建虚拟机。
  - a 右键单击属于虚拟机的有效父对象的任何清单对象，例如数据中心、文件夹、集群、资源池或主机，然后选择**新建虚拟机**。
  - b 选择**创建新的虚拟机**，然后单击**下一步**。
  - c 遵循创建虚拟机所需的步骤。
- 2 在“自定义硬件”页面上，单击**虚拟硬件**选项卡。
- 3 （可选）要删除系统为您的虚拟机创建的默认虚拟硬盘，请将光标移至该磁盘上，然后单击**移除**图标。
- 4 添加 RDM 磁盘。
  - a 单击**添加新设备**，然后从列表中选择 **RDM 磁盘**。
  - b 从 LUN 列表中，选择目标裸 LUN，然后单击**确定**。  
系统即会创建将虚拟机映射到目标 LUN 的 RDM 磁盘。RDM 磁盘在虚拟设备列表中显示为新硬盘。

## 5 配置 RDM 磁盘。

- a 单击**新硬盘**三角形符号展开 RDM 磁盘属性。
- b 为该 RDM 选择位置。

您可将 RDM 置于虚拟机配置文件所在的同一数据存储上，也可选择不同的数据存储。

**注** 要将 vMotion 用于启用了 NPIV 的虚拟机，请确保 RDM 文件和虚拟机文件位于同一数据存储上。启用 NPIV 后，无法执行 Storage vMotion。

- c 选择兼容模式。

选项	描述
物理	允许客户机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，则物理兼容模式非常有用。但是，带有物理兼容 RDM 的虚拟机不能克隆，不能制作成模板，也不能迁移（如果迁移涉及复制磁盘）。
虚拟	允许 RDM 像虚拟磁盘一样工作，因此您可以使用执行快照和克隆之类的功能。当克隆磁盘或将其制作成模板时，LUN 的内容将复制到 .vmdk 虚拟磁盘文件中。当迁移虚拟兼容模式 RDM 时，可以迁移映射文件，或将 LUN 的内容复制到虚拟磁盘中。

- d 如果已选择虚拟兼容模式，请选择磁盘模式。

磁盘模式对使用物理兼容模式的 RDM 磁盘不可用。

选项	描述
从属	快照中包含从属磁盘。
独立 - 持久	持久模式磁盘的行为与物理机上常规磁盘的行为相似。写入持久模式磁盘的所有数据都会永久性地写入磁盘。
独立 - 非持久	关闭虚拟机电源或重置虚拟机时，对非持久模式磁盘的更改将丢失。使用非持久模式，您可以每次使用相同的虚拟磁盘状态重新启动虚拟机。对磁盘的更改会写入重做日志文件并从中读取，重做日志文件会在关闭虚拟机电源或重置虚拟机时被删除。

## 6 完成虚拟机配置。

## 管理映射的 LUN 的路径

当在虚拟机中使用 RDM 时，您可以管理映射的裸 LUN 的路径。

### 步骤

- 1 右键单击虚拟机，然后选择**编辑设置**。
- 2 依次单击**虚拟硬件**选项卡和**硬盘**以展开磁盘选项菜单。
- 3 单击显示在**物理 LUN**旁边的设备 ID 以打开**编辑多路径策略**对话框。

- 4 使用**编辑多路径策略**对话框启用或禁用路径，设置多路径策略并指定首选路径。
- 有关管理路径的信息，请参见第 10 章 [了解多路径和故障切换](#)。

## 具有 RDM 的虚拟机必须忽略 SCSI INQUIRY 缓存

具有 RDM 的某些虚拟机必须从 LUN 获取 SCSI INQUIRY 信息，而不是使用 ESXi 缓存的 SCSI INQUIRY 数据。

### 问题

在具有 RDM 的虚拟机中运行的某些客户机操作系统或应用程序会显示不可预知的行为。

### 原因

该行为可能是由于影响特定客户机操作系统和应用程序的缓存 SCSI INQUIRY 数据所致。

当 ESXi 主机首次连接到目标存储设备时，此主机会发出 SCSI INQUIRY 命令以从设备获取基本标识数据。默认情况下，ESXi 会缓存收到的 SCSI INQUIRY 数据（标准，第 80 页和第 83 页），且这些数据之后保持不变。从缓存返回对后续 SCSI INQUIRY 命令的响应。

但是，在具有 RDM 的虚拟机中运行的特定客户机操作系统必须查询 LUN，而不是使用 ESXi 缓存的 SCSI INQUIRY 数据。在这些情况下，您可以配置虚拟机忽略 SCSI INQUIRY 缓存。

### 解决方案

- ◆ 使用下列方法之一。

仅当存储供应商推荐做出更改时才这样做。

选项	描述
修改具有 RDM 的虚拟机的 .vmx 文件	<p>为具有硬件版本 8 或更高版本的虚拟机使用此方法。</p> <p>a 将以下参数添加到该文件：</p> <pre>scsix:y.ignoreDeviceInquiryCache = "true"</pre> <p>其中，<i>x</i> 是 SCSI 控制器编号，<i>y</i> 是 RDM 的 SCSI 目标编号。</p> <p>b 重新引导虚拟机。</p>
使用 esxcli 命令	<p>由于是在主机级别配置此设置的，因此任何虚拟机硬件版本限制均不适用。</p> <pre>esxcli storage core device inquirycache set --device device id --ignore true</pre> <p>无需重新引导虚拟机。</p>

无论使用哪种方法将 SCSI INQUIRY 缓存参数设置为 true，虚拟机均会直接开始为 SCSI INQUIRY 数据连接 LUN。

忽略 vmx 中的 DeviceInquiryCache 参数	忽略 esxcli 中的 inquirycache 参数	查询请求的出处
True	True	LUN
False（参数不存在时的默认值）	True	LUN

忽略 vmx 中的 DeviceInquiryCache 参数	忽略 esxcli 中的 inquirycache 参数	查询请求的出处
True	False	LUN
False（参数不存在时的默认值）	False	缓存

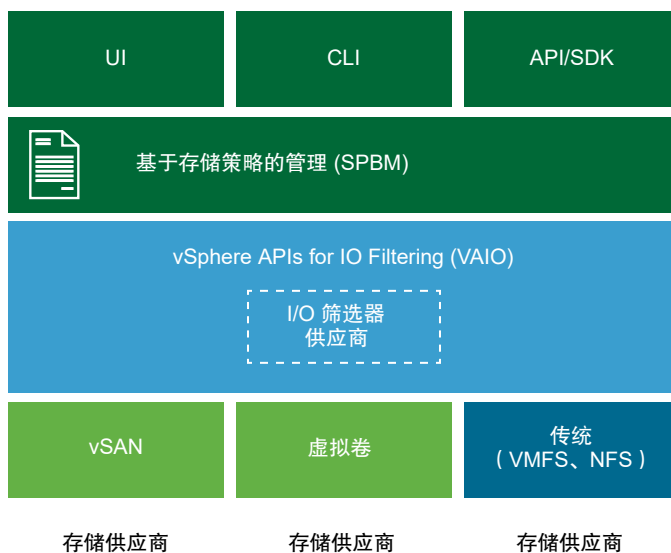
# 基于存储策略的管理

# 12

在软件定义的数据中心内，基于存储策略的管理 ( Storage Policy Based Management, SPBM) 起着重要作用，有助于根据虚拟机的应用程序需求调整存储。它提供一个存储策略框架，可充当跨多种数据服务和存储解决方案的统一控制面板。

作为一个抽象层，SPBM 将 Virtual Volumes、vSAN、I/O 筛选器或其他存储实体提供的存储服务进行抽象。

SPBM 提供适用于多种不同存储实体类型的通用框架，而不是与每种类型的存储和数据服务集成。



SPBM 提供以下机制：

- 存储阵列和诸如 I/O 筛选器、服务等其他实体的存储功能和数据服务的播发。
- 一方的 ESXi 和 vCenter Server 与另一方的存储阵列和实体之间的双向通信。
- 基于虚拟机存储策略的虚拟机置备。

本章讨论了以下主题：

- 虚拟机存储策略
- 虚拟机存储策略的工作流
- 填充“虚拟机存储策略”界面

- 关于规则和规则集
- 创建和管理虚拟机存储策略
- 关于存储策略组件
- 存储策略和虚拟机
- 默认存储策略

## 虚拟机存储策略

要通过 SPBM 置备虚拟机，虚拟机存储策略必不可少。策略控制为虚拟机提供哪种类型的存储，以及如何将虚拟机放置在存储中。它们还确定虚拟机可以使用的数据服务。

vSphere 提供了默认存储策略。此外，您可以定义策略并将它们分配给虚拟机。

使用“虚拟机存储策略”界面创建存储策略。定义策略时，可为虚拟机上运行的应用程序指定各种存储要求。您也可以使用存储策略来为虚拟磁盘请求缓存或复制等特定的数据服务。

您可以在创建、克隆或迁移虚拟机时应用该存储策略。应用存储策略后，SPBM 机制会帮助您将虚拟机放置在匹配的数据存储中。在某些存储环境中，SPBM 确定如何置备和分配存储资源内的虚拟机存储对象，以保证达到要求的服务级别。SPBM 还为虚拟机启用请求的数据服务并帮助您监控策略合规性。

## 虚拟机存储策略的工作流

创建和管理存储策略的整个过程通常包括多个步骤。

是否必须执行特定步骤可能取决于您的环境所提供的存储或数据服务的类型。

步骤	描述
使用合适的数据填充“虚拟机存储策略”界面。	<p>“虚拟机存储策略”界面填充有关存储环境中可用的数据存储和数据服务的信息。此信息从存储提供程序和数据存储标记获取。</p> <ul style="list-style-type: none"><li>■ 对于由存储提供程序呈现的实体，验证是否已注册相应的提供程序。</li></ul> <p>使用存储提供程序的实体包括 vSAN、Virtual Volumes 和 I/O 筛选器。根据存储实体的类型，有些提供程序可以自行注册。其他提供程序必须手动注册。</p> <p>请参见使用存储提供程序填充“虚拟机存储策略”界面和为 Virtual Volumes 注册存储提供程序。</p> <ul style="list-style-type: none"><li>■ 标记不是由存储提供程序呈现的数据存储。您还可以使用标记指出并非通过存储提供程序通信的属性，如地理位置或管理组。</li></ul> <p>请参见向数据存储分配标记。</p>
创建预定义的存储策略组件。	<p>存储策略组件描述必须为虚拟机提供的单个数据服务，例如复制。您可以提前定义组件并将其与多个虚拟机存储策略关联。组件可重复使用并且可互换。</p> <p>请参见创建存储策略组件。</p>
创建虚拟机存储策略。	<p>为虚拟机定义存储策略时，可指定在虚拟机上运行的应用程序的存储要求。</p> <p>请参见创建和管理虚拟机存储策略。</p>



步骤	描述
将虚拟机存储策略应用于虚拟机。	部署虚拟机或配置其虚拟磁盘时，可以应用存储策略。 请参见 <a href="#">将存储策略分配给虚拟机</a> 。
检查虚拟机存储策略的合规性。	确认虚拟机使用符合分配的存储策略的数据存储。 请参见 <a href="#">检查虚拟机存储策略的合规性</a> 。

要创建和管理您的存储策略，可以使用 vSphere Client 的“虚拟机存储策略”界面。

## 填充“虚拟机存储策略”界面

开始创建虚拟机存储策略之前，必须在“虚拟机存储策略”界面中填充有关存储环境中可用的存储实体和数据服务的信息。

该信息可从存储提供程序（也称为 VASA 提供程序）获得，也可从数据存储标记获得。

### 存储功能与服务

某些数据存储（例如，Virtual Volumes 和 vSAN）由存储提供程序呈现。通过存储提供程序，数据存储可在“虚拟机存储策略”界面中播发其功能。这些数据存储功能、数据服务及其他带有值范围的特性会填充到“虚拟机存储策略”界面。

可以使用这些特性为存储策略定义基于数据存储的放置规则和服务规则。

### 数据服务

主机上的 I/O 筛选器也由存储提供程序表示。存储提供程序可将有关筛选器数据服务的信息填充到“虚拟机存储策略”界面。可以使用这些信息为基于主机的数据服务定义规则（也称为常用规则）。与特定于数据存储的规则不同，常用规则不定义虚拟机的存储放置和存储要求，而是会为虚拟机激活请求的 I/O 筛选器数据服务。

### 标记

通常，VMFS 和 NFS 数据存储不由存储提供程序表现。它们不在“虚拟机存储策略”界面中显示其功能和数据服务。您可以使用标记将有关这些数据存储的信息进行编码。例如，您可将 VMFS 数据存储标记为 VMFS-Gold 和 VMFS-Silver 以表示不同的服务级别。

对于 Virtual Volumes 和 vSAN 数据存储，可以使用标记对不由存储提供程序播发的信息进行编码，例如地理位置（帕洛阿尔托）或管理组（会计）。

与存储功能和特性类似，与数据存储关联的所有标记都显示在“虚拟机存储策略”界面中。您可在定义基于标记的放置规则时使用标记。

## 使用存储提供程序填充“虚拟机存储策略”界面

对于由存储 (VASA) 提供程序呈现的实体，验证是否已注册相应的提供程序。注册存储提供程序后，“虚拟机存储策略”界面将填充提供程序呈现的数据存储和数据服务的有关信息。

使用存储提供程序的实体包括 vSAN、Virtual Volumes 和 I/O 筛选器。根据实体的类型，有些提供程序可以自行注册。其他提供程序必须手动注册，如 Virtual Volumes 存储提供程序。存储提供程序注册后，将向“虚拟机存储策略”界面提供以下数据：

- Virtual Volumes 和 vSAN 等数据存储的存储功能和特性。
- I/O 筛选器提供的数据服务。

### 前提条件

注册需要手动注册的存储提供程序。有关详细信息，请参见相应的文档：

- 《管理 VMware vSAN》
- [第 14 章 使用 VMware vSphere Virtual Volumes](#)
- [第 15 章 筛选虚拟机 I/O](#)

### 步骤

- 1 浏览到 vCenter Server 实例。
- 2 依次单击**配置**选项卡和**存储提供程序**。
- 3 在“存储提供程序”列表中，查看向 vCenter Server 注册的存储提供程序。

此列表显示一般信息，包括存储提供程序的名称、其 URL 和状态、提供程序呈现的存储实体等。

- 4 要显示更多详细信息，请从列表中选择特定的存储提供程序或其组件。

## 向数据存储分配标记

可使用标记对有关数据存储的信息进行编码。当数据存储并非由存储提供程序呈现并且未在“虚拟机存储策略”界面中播发其服务时，标记非常有用。您还可以使用标记指明不通过存储提供程序传递的属性，例如地理位置或管理组。

可以将包含常规存储信息的新标记应用于数据存储。有关标记、其类别以及如何管理标记的更多详细信息，请参见《vCenter Server 和主机管理》文档。

### 前提条件

所需特权：

- 在根 vCenter Server 实例上的 **vSphere 标记.创建 vSphere 标记** 特权
- 在根 vCenter Server 实例上的 **vSphere 标记.创建 vSphere 标记类别** 特权
- 在根 vCenter Server 实例上的 **vSphere 标记.分配或取消分配 vSphere 标记** 特权

步骤

- 1 在 vSphere Client 中，为存储标记创建类别。
  - a 在“主页”菜单中，单击**标记与自定义属性**。
  - b 依次单击**标记**选项卡和**类别**。
  - c 单击**添加类别**图标。
  - d 指定类别属性。请参见下面的示例。

类别属性	示例
类别名称	存储位置
描述	与存储位置相关的标记类别
每个对象标记数	多个标记
可关联的对象类型	数据存储和 <b>数据存储集群</b>

- e 单击**确定**。
- 2 创建存储标记。
  - a 在**标记**选项卡上，单击**标记**。
  - b 单击**添加标记**图标。
  - c 指定标记的属性。请参见下面的示例。

标记属性	示例
名称	德克萨斯
描述	位于德克萨斯的数据存储
类别	存储位置

- d 单击**确定**。
- 3 将标记应用于数据存储。
  - a 导航到数据存储。
  - b 右键单击数据存储，然后选择**标记与自定义属性 > 分配标记**。
  - c 从标记列表中选择适当的标记（例如“存储位置”类别中的“德克萨斯”），然后单击**分配**。

结果

新标记将分配给数据存储，并显示在**标记**窗格中的数据存储**摘要**选项卡上。

后续步骤

创建虚拟机存储策略时，您可以引用标记以将标记的数据存储包含在兼容存储资源列表中。请参见[为基于标记的放置创建虚拟机存储策略](#)。

或者，您可以从虚拟机存储策略中排除标记的数据存储。例如，虚拟机存储策略可以包含位于德克萨斯和加利福尼亚的 **Virtual Volumes** 数据存储，但排除位于内华达的数据存储。

要了解有关如何在虚拟机存储策略中使用标记的更多信息，请观看以下视频。



(在存储策略中使用标记)

## 关于规则和规则集

在“虚拟机存储策略”界面中填充适当的数据后，您可以开始创建存储策略。创建策略涉及到定义用于配置数据服务的特定存储放置规则和规则集。

### 规则

规则是虚拟机存储策略的基本要素。每个规则都是描述一项虚拟机存储和数据服务要求的语句。

### 规则集

在存储策略中，各个规则被组织到不同的规则集合或规则集中。通常，规则集可分为两类：基于主机的服务的规则以及数据存储特定规则。

### 数据存储特定规则集

每个规则集都必须包含描述虚拟机存储资源要求的放置规则。单个规则集中的所有放置规则代表一个存储实体。这些规则可以基于存储功能或标记。

此外，数据存储特定规则集可以包含可选规则或存储策略组件，它们描述要为虚拟机提供的数据服务。通常，这些规则会请求存储系统提供缓存、复制等服务。

要定义存储策略，必须提供一个数据存储特定规则集。其他规则集则可选择性提供。单个策略可以使用多个规则集来定义备用存储放置参数，这些参数通常来自多个存储提供程序。

### 放置规则：基于功能

放置规则指定虚拟机的特定存储要求并使 **SPBM** 可以从清单内的所有数据存储中区分兼容的数据存储。这些规则还描述如何分配数据存储内的虚拟机存储对象，以达到要求的服务级别。例如，规则可以将 **Virtual Volumes** 作为目标列出并为 **Virtual Volumes** 对象定义最大恢复点目标 (RPO)。

当您置备虚拟机时，这些规则将指导 **SPBM** 做出有关虚拟机放置的决定。**SPBM** 查找与规则匹配并满足虚拟机存储要求的 **Virtual Volumes** 数据存储。请参见 [为 Virtual Volumes 创建虚拟机存储策略](#)。

### 放置规则：基于标记

基于标记的规则引用数据存储标记。这些规则可以定义虚拟机放置，例如请求将虚拟机放置到具有 **VMFS-Gold** 标记的所有数据存储。您还可以使用基于标记的规则进一步精确调整您的虚拟机放置请求。例如，从 **Virtual Volumes** 数据存储列表中排除具有 **Palo Alto** 标记的数据存储。请参见 [为基于标记的放置创建虚拟机存储策略](#)。

### 基于主机的服务的规则

使用此规则集可激活主机提供的数据服务。基于主机的服务的规则集可以包含描述特定数据服务（如加密或复制）的规则或存储策略组件。

与数据存储特定规则不同，此规则集不包含放置规则。基于主机的服务的规则对于所有存储类型是通用的，不依赖于数据存储。请参见[为基于主机的数据服务创建虚拟机存储策略](#)。

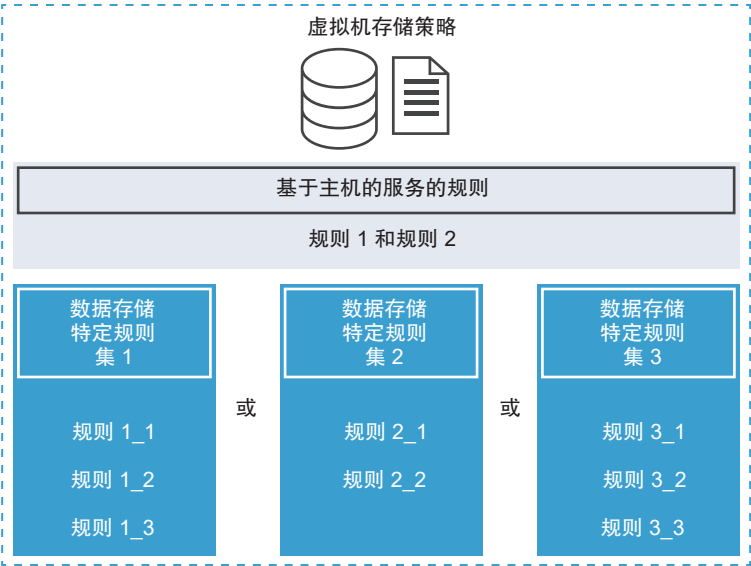
表 12-1. 虚拟机存储策略的结构

基于主机的服务的规则	数据存储特定规则集
激活安装在 ESXi 主机上的数据服务的规则或预定义存储策略组件。例如，I/O 筛选器进行的复制。	描述虚拟机存储资源要求的基于功能或基于标记的放置规则。例如，Virtual Volumes 放置。
	激活存储提供的数据服务的规则或预定义存储策略组件。例如，按 Virtual Volumes 进行缓存。

规则和规则集之间的关系

布尔运算符 OR 定义策略中不同数据存储特定规则集之间的关系。AND 运算符定义单个规则集中所有规则之间的关系。策略可以只包含基于主机的服务的规则集，或者只包含数据存储特定规则集，也可以同时包含这两者。

如果不存在基于主机的服务的规则集，则满足单个数据存储特定规则集中的所有规则即满足整个策略的要求。如果存在基于主机的服务的规则集，则策略要匹配的数据存储需同时满足主机服务规则以及一个数据存储特定规则集中的所有规则。



## 创建和管理虚拟机存储策略

要为虚拟机创建和管理存储策略，可以使用“虚拟机存储策略”界面。

### 为基于主机的数据服务创建虚拟机存储策略

要在 vSphere Client 中定义虚拟机存储策略，请使用[创建虚拟机存储策略](#)向导。在此任务中，您可以为 ESXi 主机提供的数据服务创建规则。包含这些规则的虚拟机存储策略为虚拟机激活指定的数据服务。

可用的数据服务包括加密、I/O 控制和缓存等。某些数据服务（例如加密）由 VMware 提供。主机上安装的第三方 I/O 筛选器可以提供其他数据服务。

数据服务通常对所有存储类型是通用的，不依赖于数据存储。向存储策略添加特定于数据存储的规则是可选操作。

如果您添加特定于数据存储的规则并且主机和存储上的 I/O 筛选器提供同一种类型的服务（例如加密），则您的策略可以从这两个提供程序请求该服务。因此，虚拟机数据由 I/O 筛选器和存储加密两次。然而，Virtual Volumes 提供的复制和 I/O 筛选器提供的复制不能共存于同一存储策略中。

#### 前提条件

- 有关加密虚拟机的信息，请参见《vSphere 安全性》文档。
- 有关 I/O 筛选器的信息，请参见[第 15 章 筛选虚拟机 I/O](#)。
- 有关存储策略组件的信息，请参见[关于存储策略组件](#)。
- 所需特权：[虚拟机存储策略.更新](#)和[虚拟机存储策略.查看](#)。

#### 步骤

- 1 打开[创建虚拟机存储策略](#)向导。
  - a 单击[菜单 > 策略和配置文件](#)。
  - b 在[策略和配置文件](#)下，单击[虚拟机存储策略](#)
  - c 单击[创建](#)。
- 2 输入策略名称和描述，然后单击[下一步](#)。

选项	操作
vCenter Server	选择 vCenter Server 实例。
名称	输入存储策略的名称。
描述	输入存储策略的描述。

- 3 在[策略结构](#)页面上的[基于主机的服务](#)下，启用基于主机的规则。

#### 4 在[基于主机的服务](#)页面上，定义规则以启用并配置由主机提供的数据服务。

- a 单击数据服务类别对应的选项卡，例如[复制](#)。
- b 为数据服务类别定义自定义规则，或者使用预定义的组件。

选项	描述
已禁用	默认情况下，基于主机的服务处于禁用状态。
使用存储策略组件	从下拉菜单中选择存储策略组件。如果数据库中有预定义的组件，则此选项可用。
自定义	通过为规则指定相应的提供程序和价值来定义数据服务类别的自定义规则。

**注** 可以启用多个数据服务。如果将加密与其他数据服务结合使用，请将在[加密前允许 I/O 筛选器](#)参数设置为 **True**，以便其他服务（例如复制）可以在加密明文数据之前对其进行分析。

#### 5 在[存储兼容性](#)页面上，检查与此策略匹配的数据存储列表。

要与基于主机的服务的策略兼容，数据存储必须连接到提供这些服务的主机。如果向策略添加特定于数据存储的规则集，则兼容的数据存储还必须满足该策略的存储要求。

#### 6 在[查看并完成](#)页面上，查看存储策略设置，然后单击[完成](#)。

要更改任何设置，请单击[后退](#)以转至相关页面。

### 结果

基于主机的数据服务的新虚拟机存储策略会显示在列表中。

## 为 Virtual Volumes 创建虚拟机存储策略

要在 vSphere Client 中定义虚拟机存储策略，请使用[创建虚拟机存储策略](#)向导。在此任务中，将创建与 Virtual Volumes 兼容的自定义存储策略。为 Virtual Volumes 定义虚拟机存储策略时，可以创建规则来配置 Virtual Volumes 数据存储提供的存储和数据服务。将虚拟机放置在 Virtual Volumes 数据存储中时应用这些规则。自定义存储策略可以取代 VMware 提供的 Virtual Volumes 的默认“无要求”存储策略。

该过程假设您要为 Virtual Volumes 创建虚拟机存储策略。有关 vSAN 存储策略的信息，请参见《管理 VMware vSAN》文档。

### 前提条件

- 确认 Virtual Volumes 存储提供程序可用且处于活动状态。请参见[为 Virtual Volumes 注册存储提供程序](#)。
- 确保“虚拟机存储策略”界面填充了有关存储环境中可用的存储实体和数据服务的信息。请参见[填充“虚拟机存储策略”界面](#)。
- 定义相应的存储策略组件。请参见[创建存储策略组件](#)。
- 所需特权：[虚拟机存储策略.更新](#)和[虚拟机存储策略.查看](#)。



步骤

- 1 打开**创建虚拟机存储策略**向导。
  - a 单击**菜单 > 策略和配置文件**。
  - b 在**策略和配置文件**下，单击**虚拟机存储策略**。
  - c 单击**创建**。
- 2 输入策略名称和描述，然后单击**下一步**。

选项	操作
vCenter Server	选择 vCenter Server 实例。
名称	输入存储策略的名称，例如 Virtual Volumes 存储策略。
描述	输入存储策略的描述。

- 3 在**策略结构**页面上的“数据存储特定规则”下，为目标存储实体（例如 Virtual Volumes 存储）启用规则。

可以为多个数据存储启用规则。多个规则集允许使用一个策略来定义备用存储放置参数，这些参数通常来自多个存储提供程序。
- 4 在 *Virtual Volumes* 规则页面上，为目标 Virtual Volumes 数据存储定义存储放置规则。
  - a 单击**放置**选项卡，然后单击**添加规则**。
  - b 从“添加规则”下拉菜单中，选择可用功能并指定其值。

例如，您可以指定 Virtual Volumes 对象每秒的读取操作数。

您可以根据需要为所选的存储实体包含任意数量的规则。验证您提供的值是否位于 Virtual Volumes 数据存储播发的值范围内。
  - c 要进一步精确调整放置请求，请单击**标记**选项卡并添加基于标记的规则。

基于标记的规则可以通过包含或排除特定放置标准来筛选数据存储。例如，虚拟机存储策略可以包含位于德克萨斯和加利福尼亚的 Virtual Volumes 数据存储，但排除位于内华达的数据存储。
- 5 （可选）定义规则以配置特定于数据存储的服务。

加密、缓存或复制等数据服务由存储提供。将虚拟机放置在 Virtual Volumes 数据存储中时，引用数据服务的虚拟机存储策略会为虚拟机请求这些服务。

  - a 单击数据服务类别对应的选项卡，例如**复制**。
  - b 为数据服务类别定义自定义规则，或者使用预定义的组件。

选项	描述
已禁用	默认情况下，特定于数据存储的服务处于禁用状态。
使用存储策略组件	从下拉菜单中选择存储策略组件。如果数据库中有预定义的组件，则此选项可用。
自定义	通过为规则指定相应的提供程序和价值来定义数据服务类别的自定义规则。



- 6 在**存储兼容性**页面上，检查与此策略匹配的数据存储列表。
- 如果策略包含多个规则集，则数据存储至少必须满足一个规则集以及该集中的所有规则。
- 7 在**查看并完成**页面上，查看存储策略设置，然后单击**完成**。
- 要更改任何设置，请单击**后退**以转至相关页面。

结果

与 Virtual Volumes 兼容的新虚拟机存储策略将显示在列表中。

后续步骤

现在即可将此策略与虚拟机关联，或将该策略指定为默认策略。

为基于标记的放置创建虚拟机存储策略

基于标记的规则引用分配给数据存储的标记，并可以筛选要用于放置虚拟机的数据存储。要在 vSphere Client 中定义基于标记的放置，请使用**创建虚拟机存储策略**向导。

前提条件

- 确保“虚拟机存储策略”界面填充了有关存储环境中可用的存储实体和数据服务的信息。请参见[填充“虚拟机存储策略”界面](#)。
- 所需特权：[虚拟机存储策略.更新](#)和[虚拟机存储策略.查看](#)。

步骤

- 1 打开**创建虚拟机存储策略**向导。
- a 单击**菜单 > 策略和配置文件**。
  - b 在**策略和配置文件**下，单击**虚拟机存储策略**。
  - c 单击**创建**。
- 2 输入策略名称和描述，然后单击**下一步**。

选项	操作
vCenter Server	选择 vCenter Server 实例。
名称	输入存储策略的名称。
描述	输入存储策略的描述。

- 3 在**策略结构**页面的“数据存储特定规则”下，启用基于标记的放置规则。

- 4 在**基于标记放置**页面上，创建标记规则。
  - a 单击**添加标记规则**并定义基于标记的放置标准。可参考以下示例：

选项	示例
标记类别	服务级别
使用情况选项	使用具有以下标记的存储
标记	金级

具有金级标记的所有数据存储将与存储放置目标兼容。

- b （可选）添加更多基于标记的规则。
- 5 在**存储兼容性**页面上，检查与此策略匹配的数据存储列表。
- 6 在**查看并完成**页面上，查看存储策略设置，然后单击**完成**。

要更改任何设置，请单击**后退**以转至相关页面。

## 结果

与标记的数据存储兼容的新虚拟机存储策略将显示在列表中。

## 编辑或克隆虚拟机存储策略

如果虚拟机和虚拟磁盘的存储要求已更改，则可以修改现有的存储策略。还可以通过克隆来创建现有虚拟机存储策略的副本。克隆时，可以选择自定义原始存储策略。

### 前提条件

所需特权：**StorageProfile.View**

### 步骤

- 1 在 vSphere Client 中，导航到存储策略。
  - a 单击**菜单 > 策略和配置文件**。
  - b 在**策略和配置文件**下，单击**虚拟机存储策略**。
- 2 选择存储策略，然后单击以下图标之一：
  - **编辑**
  - **克隆**
- 3 （可选）修改该策略，然后单击**确定**。

4 如果编辑由虚拟机使用的存储策略，请将该策略重新应用于虚拟机。

选项	描述
稍后手动应用	如果选择此选项，则与存储策略关联的所有虚拟磁盘和虚拟机 <b>home</b> 对象的合规性状态将更改为“已过期”。要更新配置和合规性，请手动将存储策略重新应用于所有关联的条目。请参见 <a href="#">重新应用虚拟机存储策略</a> 。
现在	编辑存储策略后立即更新虚拟机和合规性状态。

## 关于存储策略组件

虚拟机存储策略可以包含一个或多个可重复使用和可互换的构造块，称为存储策略组件。每个组件描述要为虚拟机提供的一项特定数据服务。您可以提前定义策略组件并将它们与多个虚拟机存储策略关联。

您无法将预定义的组件直接分配给虚拟机或虚拟磁盘。必须将组件添加到虚拟机存储策略，然后将策略分配给虚拟机。

组件描述来自一个服务提供程序的一种服务类型。根据您的提供程序，服务可能有所不同，但通常属于以下类别之一。

- 压缩
- 缓存
- 加密
- 复制

创建存储策略组件时，可以为一个特定类型和级别的服务定义规则。

以下示例说明虚拟机 **VM1** 和 **VM2** 具有相同的放置要求，但必须具有不同级别的复制服务。您可以创建具有不同复制参数的存储策略组件并将这些组件添加到相关存储策略中。

表 12-2. 存储策略组件

虚拟机	放置规则	存储策略组件
VM1 要求每隔 2 小时复制一次	Virtual Volumes 数据存储	间隔 2 小时的复制
VM2 要求每隔 4 小时复制一次	Virtual Volumes 数据存储	间隔 4 小时的复制

服务提供程序可以是存储系统、I/O 筛选器或其他实体。如果组件引用 I/O 筛选器，则该组件会添加到存储策略中基于主机的规则。如果组件引用除 I/O 筛选器以外的实体（例如存储系统），则该组件会添加到特定于数据存储的规则集中。

处理组件时，请遵循以下准则：

- 每个组件只能包含一个规则集。此规则集中的所有特性都属于一个数据服务提供程序。
- 如果组件在虚拟机存储策略中引用，则无法删除组件。在删除组件之前，您必须将其从存储策略中移除或者删除存储策略。
- 将组件添加到策略时，您只能针对每个规则集使用同一类别（例如缓存）中的一个组件。

## 创建存储策略组件

存储策略组件描述必须为虚拟机提供的单个数据服务，例如复制。您可以提前定义组件并将其与多个虚拟机存储策略关联。组件可重复使用并且可互换。

### 步骤

- 1 在 vSphere Client 中，打开**新建存储策略组件**对话框。
  - a 单击**菜单 > 策略和配置文件**。
  - b 在**策略和配置文件**下，单击**存储策略组件**。
- 2 单击**创建存储策略组件**。
- 3 选择 vCenter Server 实例。
- 4 输入策略组件的名称（例如间隔 4 小时的复制）和描述。

确保名称不与其他组件或存储策略的名称相冲突。
- 5 选择服务类别，例如**复制**。
- 6 选择服务提供程序。
- 7 为所选类别定义规则。

例如，如果您正在配置间隔 4 小时的复制，请将恢复点目标 (RPO) 值设置为 4。

对于基于 I/O 筛选器的加密，请设置在**加密前允许 I/O 筛选器**参数。存储提供的加密不需要此参数。

选项	描述
False（默认）	不允许在加密筛选器前使用其他 I/O 筛选器。
True	允许在加密筛选器前使用其他 I/O 筛选器。其他筛选器（如复制）可以在加密明文数据之前对其进行分析。

- 8 单击**确定**。

### 结果

新组件将显示在存储策略组件列表中。

### 后续步骤

您可以将组件添加到虚拟机存储策略。如果组件引用的数据服务由 I/O 筛选器提供，您可以将该组件添加到存储策略中基于主机的规则。如果组件引用除 I/O 筛选器以外的实体（例如存储系统），则该组件会添加到特定于数据存储的规则集中。

## 编辑或克隆存储策略组件

您可以修改现有的存储策略组件。还可以通过克隆来创建现有组件的副本。

### 步骤

- 1 在 vSphere Client 中，导航到要编辑或克隆的存储策略组件。
  - a 单击 **菜单 > 策略和配置文件**。
  - b 在 **策略和配置文件** 下，单击 **存储策略组件**。
- 2 选择组件，然后单击以下图标之一。

选项	描述
编辑设置	编辑时，您不能更改数据服务和提供程序的类别。例如，如果原始组件引用 I/O 筛选器提供的复制，则这些设置必须保持不变。
克隆	克隆时，您可以自定义原始组件的任何设置。

- 3 修改相应的值，然后单击 **确定**。
- 4 如果分配至虚拟机的虚拟机存储策略引用了您编辑的策略组件，则将该存储策略重新应用于虚拟机。

菜单项	描述
稍后手动应用	如果选择此选项，则与存储策略关联的所有虚拟磁盘和虚拟机 <b>home</b> 对象的合规性状态将更改为“已过期”。要更新配置和合规性，请手动将存储策略重新应用于所有关联的条目。请参见 <a href="#">重新应用虚拟机存储策略</a> 。
现在	编辑存储策略后立即更新虚拟机和合规性状态。

## 存储策略和虚拟机

定义虚拟机存储策略后，即可将其应用于虚拟机。置备虚拟机或配置其虚拟磁盘时，可以应用存储策略。根据其类型和配置，策略可用于不同的目的。该策略可为虚拟机选择最合适的数据存储并强制执行要求的服务水平。或者，它可以为虚拟机及其磁盘启用特定数据服务。

如果未指定存储策略，则系统将使用与数据存储关联的默认存储策略。如果虚拟机上的应用程序的存储要求更改，则可以修改最初应用于虚拟机的存储策略。

## 将存储策略分配给虚拟机

在虚拟机初始部署期间或在执行其他虚拟机操作（如克隆或迁移）时，可以分配虚拟机存储策略。

本主题介绍如何在创建虚拟机时分配虚拟机存储策略。有关克隆和从模板部署等其他部署方法的信息，请参见《《vSphere 虚拟机管理》》文档。

您可以将同一存储策略应用于虚拟机配置文件及其所有虚拟磁盘。如果虚拟磁盘和配置文件的要求不同，则可以将不同的存储策略与虚拟机配置文件及选定的虚拟磁盘关联。

步骤

- 1 启动虚拟机置备流程并按照相应的步骤操作。
- 2 将同一存储策略分配给所有虚拟机文件和磁盘。
  - a 在**选择存储**页面上，从**虚拟机存储策略**下拉菜单中选择存储策略。

存储策略会根据其配置，将所有数据存储分成兼容数据存储和不兼容数据存储。如果策略引用了某一特定存储实体（如 **Virtual Volumes**）提供的数据服务，则兼容列表将包含仅表示该存储类型的数据存储。
  - b 从兼容数据存储列表中选择相应的数据存储。

该数据存储将成为虚拟机配置文件及所有虚拟磁盘的目标存储资源。
  - c 如果将复制服务与 **Virtual Volumes** 结合使用，请指定复制组。

复制组指明需要将哪些虚拟机和虚拟磁盘一起复制到目标站点。

选项	描述
预配置复制组	在存储端预先配置的复制组。vCenter Server 和 ESXi 会发现复制组，但不管理其生命周期。
自动复制组	Virtual Volumes 会创建一个复制组并将所有虚拟机对象分配到此组。

- 3 更改虚拟磁盘的虚拟机存储策略。

如果虚拟磁盘的存储放置的要求不同，请使用此选项。您还可以使用此选项来为虚拟磁盘启用 I/O 筛选器服务，例如缓存和复制。

  - a 在**自定义硬件**页面上，展开**新硬盘**窗格。
  - b 从**虚拟机存储策略**下拉菜单中，选择要分配给虚拟磁盘的存储策略。
  - c （可选）更改虚拟磁盘的存储位置。

使用此选项可将虚拟磁盘存储在除虚拟机配置文件所在的数据存储以外的数据存储上。
- 4 完成虚拟机置备流程。

结果

创建虚拟机后，**摘要**选项卡将显示已分配的存储策略及其合规性状态。

后续步骤

如果配置文件或虚拟磁盘的存储放置要求更改，则可随后修改虚拟策略分配。

更改虚拟机文件和磁盘的存储策略分配

如果虚拟机上应用程序的存储要求发生更改，则可以编辑最初应用于该虚拟机的存储策略。  
可以为关闭电源或打开电源的虚拟机编辑存储策略。

在更改虚拟机存储策略分配时，您可以将同一存储策略应用于虚拟机配置文件及其所有虚拟磁盘。您还可以将不同存储策略与虚拟机配置文件和虚拟磁盘关联。您可能会应用不同策略，例如，当您的虚拟磁盘和配置文件的存储要求不同时。

步骤

- 1 在 vSphere Client 中，浏览到虚拟机。
  - a 单击**菜单 > 策略和配置文件**。
  - b 在**策略和配置文件**下，单击**虚拟机存储策略**。
  - c 单击要更改的存储策略，然后单击**虚拟机合规性**。  
您可以看到使用此存储策略的虚拟机列表。
  - d 单击要修改其策略的虚拟机。
- 2 依次单击**配置**选项卡和**策略**。
- 3 单击**编辑虚拟机存储策略**。
- 4 为虚拟机指定虚拟机存储策略。

选项	操作
将同一存储策略应用于所有虚拟机对象	从 <b>虚拟机存储策略</b> 下拉菜单中选择策略。
将不同的存储策略应用于虚拟机主对象和虚拟磁盘	<ol style="list-style-type: none"><li>a 打开<b>按磁盘配置</b>选项。</li><li>b 选择对象，例如，虚拟机主页。</li><li>c 在“虚拟机存储策略”列中，从下拉菜单中选择策略。</li></ol>

- 5 如果您使用 Virtual Volumes 策略进行复制，请配置复制组。  
复制组指明需要将哪些虚拟机和虚拟磁盘一起复制到目标站点。  
虚拟机的所有存储对象应属于同一复制组。无法将不同的复制组分配给虚拟机的不同存储对象。
- 6 单击**确定**以保存虚拟机存储策略更改。

结果

存储策略即分配给虚拟机及其磁盘。

## 检查虚拟机存储策略的合规性

您可以检查虚拟机使用的数据存储是否与该虚拟机存储策略中指定的存储要求兼容。

前提条件

验证虚拟机是否包含与其关联的存储策略。

步骤

- 1 在 vSphere Client 中，导航到相关虚拟机。
- 2 依次单击**配置**选项卡和**策略**。

### 3 单击检查虚拟机存储策略的合规性。

系统会验证合规性。

### 4 查看合规性状态。

合规性状态	描述
合规	虚拟机或虚拟磁盘使用的数据存储具有与策略要求兼容的存储功能。
不合规	虚拟机或虚拟磁盘使用的数据存储没有与策略要求兼容的存储功能。可以将虚拟机文件和虚拟磁盘迁移到合规的数据存储中。
已过期	该状态表示策略已被编辑，但新的要求尚未传递至虚拟机对象所在的数据存储。要传递更改，请将策略重新应用到已过期的对象。
不适用	此存储策略引用不受虚拟机所在的数据存储支持的数据存储功能。

#### 后续步骤

当无法将不合规的数据存储变为合规时，请将文件或虚拟磁盘迁移到兼容的数据存储。请参见[为不合规虚拟机查找兼容存储资源](#)。

如果状态为“已过期”，请将策略重新应用于对象。请参见[重新应用虚拟机存储策略](#)。

## 为不合规虚拟机查找兼容存储资源

确定与虚拟机关联的存储策略兼容的数据存储。

有时，分配给虚拟机的存储策略可能处于不合规状态。此状态表示虚拟机或其磁盘使用的数据存储与策略不兼容。可以将虚拟机文件和虚拟磁盘迁移到兼容的数据存储中。

使用此任务确定满足策略要求的数据存储。

#### 步骤

#### 1 验证虚拟机的存储策略是否处于不合规状态。

- a 在 vSphere Client 中，导航到相关虚拟机。
- b 单击**摘要**选项卡。

“虚拟机存储策略”窗格上的“虚拟机存储策略合规性”面板将显示“不合规”状态。

#### 2 导航到不合规的存储策略。

- a 单击**菜单 > 策略和配置文件**。
- b 在**策略和配置文件**下，单击**虚拟机存储策略**。

#### 3 显示不合规的存储策略的兼容数据存储的列表。

- a 单击该存储策略。
- b 单击**存储兼容性**。

此时将显示与策略的要求相匹配的数据存储列表。



## 后续步骤

可将虚拟机或其磁盘迁移到列表中的其中一个数据存储中。

## 重新应用虚拟机存储策略

编辑已与虚拟机对象关联的存储策略后，必须重新应用该策略。通过重新应用该策略，可将新的存储要求传递到虚拟机对象所在的数据存储。

### 前提条件

虚拟机的合规性状态为“已过期”。该状态表示策略已被编辑，但新的要求尚未传递到数据存储。

### 步骤

- 1 在 vSphere Client 中，导航到相关虚拟机。
- 2 依次单击**配置**选项卡和**策略**。
- 3 验证合规性状态是否为“已过期”。
- 4 单击**重新应用虚拟机存储策略**。
- 5 检查合规性状态。

合规性状态	描述
合规	虚拟机或虚拟磁盘使用的数据存储具有策略所需的存储功能。
不合规	<p>数据存储支持指定的存储要求，但目前无法符合存储策略。例如，当数据存储的物理资源不可用时，状态可能变为“不合规”。您可以更改主机集群的物理配置，以使数据存储合规。例如，通过将主机或磁盘添加到集群。如果其他资源符合存储策略，状态将变为“合规”。</p> <p>当无法将不合规的数据存储变为合规时，请将文件或虚拟磁盘迁移到兼容的数据存储。请参见<a href="#">为不合规虚拟机查找兼容存储资源</a>。</p>
不适用	存储策略引用不受数据存储支持的数据存储功能。

## 默认存储策略

在数据存储上置备虚拟机时，必须为虚拟机分配兼容的虚拟机存储策略。如果未明确配置存储策略并将其分配给虚拟机，系统将使用默认存储策略。

### VMware 提供的默认存储策略

ESXi 提供的通用的默认存储策略适用于所有数据存储，并且不包含特定于任何存储类型的规则。

此外，ESXi 还提供适用于基于对象的数据存储、vSAN 或 Virtual Volumes 的默认存储策略。这些策略可保证在基于对象的存储中以最佳方式放置虚拟机对象。

有关 Virtual Volumes 的默认存储策略的信息，请参见 [Virtual Volumes 和虚拟机存储策略](#)。

VMFS 和 NFS 数据存储没有特定的默认策略，可以使用通用的默认策略或您为其定义的自定义策略。

### 用户定义的默认存储策略

您可以创建与 vSAN 或 Virtual Volumes 兼容的虚拟机存储策略。然后，可以将此策略指定为 vSAN 和 Virtual Volumes 数据存储的默认策略。用户定义的默认策略将替换 VMware 提供的默认存储策略。

每个 vSAN 和 Virtual Volumes 数据存储一次只能有一个默认策略。但是，您可以创建包含多个放置规则集的单个存储策略，以便该存储策略匹配多个 vSAN 和 Virtual Volumes 数据存储。您可以将此策略指定为所有数据存储的默认策略。

虚拟机存储策略成为数据存储的默认策略时，除非解除其与数据存储的关联，否则无法删除该策略。

## 更改数据存储的默认存储策略

对于 Virtual Volumes 和 vSAN 数据存储，VMware 提供的存储策略将在虚拟机置备期间用作默认策略。可以更改选定 Virtual Volumes 或 vSAN 数据存储的默认存储策略。

---

**注** 请勿将包含复制规则的存储策略指定为默认存储策略。否则，该策略将阻止您选择复制组。

---

### 前提条件

创建与 Virtual Volumes 或 vSAN 兼容的存储策略。可以创建与两种类型的存储匹配的策略。

### 步骤

- 1 在 vSphere Client 中，导航到数据存储。
- 2 依次单击**配置**选项卡和**常规**。
- 3 在“默认存储策略”窗格中，单击**编辑**。
- 4 从可用存储策略的列表中，选择要指定为默认策略的策略，然后单击**确定**。

### 结果

所选存储策略将成为数据存储的默认策略。未选择其他策略时，系统会将此策略分配给在数据存储上置备的任何虚拟机对象。

# 使用存储提供程序

# 13

存储提供程序是由 VMware 提供或由第三方通过 vSphere APIs for Storage Awareness (VASA) 开发的软件组件。存储提供程序也可以称为 VASA 提供程序。存储提供程序可与包含外部物理存储和存储抽象的各种存储实体（例如 vSAN 和 Virtual Volumes）集成。存储提供程序也可以支持软件解决方案，如 I/O 筛选器。

本章讨论了以下主题：

- 关于存储提供程序
- 存储提供程序和数据呈现形式
- 存储提供程序要求和注意事项
- 注册存储提供程序
- 查看存储提供程序信息
- 管理存储提供程序

## 关于存储提供程序

通常，vCenter Server 和 ESXi 使用存储提供程序获取有关存储配置、状态和环境中提供的存储数据服务的信息。此信息显示在 vSphere Client 中。此信息可帮助您制定恰当的虚拟机放置决策、设置存储要求以及监控存储环境。

### 持久性存储提供程序

管理阵列和存储抽象的存储提供程序称为持久性存储提供程序。支持 Virtual Volumes 或 vSAN 的提供程序属于此类别。除了存储，持久性提供程序还可以提供其他数据服务，如复制。

### 数据服务提供程序

另一种提供程序类别是 I/O 筛选器存储提供程序，也可称为数据服务提供程序。这些提供程序提供数据服务，包括基于主机的缓存、压缩和加密。

这两个持久性存储和数据服务提供程序可以属于以下类别之一。

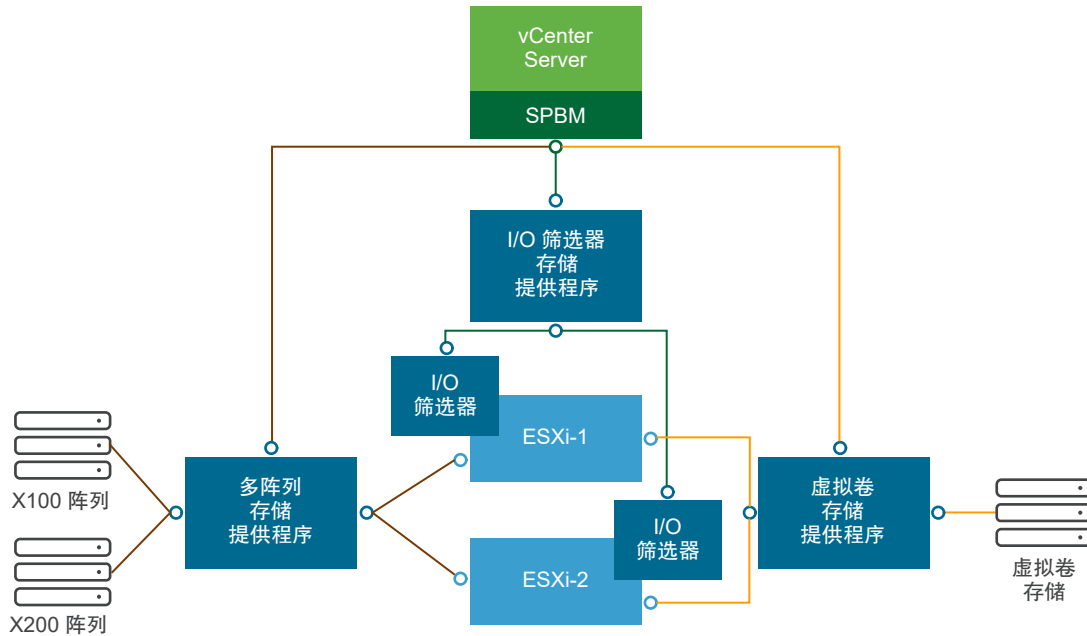
### 内置存储提供程序

内置存储提供程序由 VMware 提供。通常情况下，无需注册。例如，支持 vSAN 或 I/O 筛选器的存储提供程序为内置提供程序，并会自动变为已注册的提供程序。

### 第三方存储提供程序

如果存储提供程序由第三方提供，则通常必须注册此提供程序。Virtual Volumes 提供程序即是此类提供程序。使用 vSphere Client 注册和管理每个存储提供程序组件。

下图说明了不同类型的存储提供程序如何帮助实现 vCenter Server 与 ESXi 以及存储环境中的其他组件之间的通信。例如，这些组件可能包括存储阵列、Virtual Volumes 存储和 I/O 筛选器。



## 存储提供程序和数据呈现形式

vCenter Server 和 ESXi 与存储提供程序通信，以获取存储提供程序从基础物理存储和软件定义的存储收集的信息或从可用 I/O 筛选器收集的信息。然后 vCenter Server 可在 vSphere Client 中显示存储数据。

存储提供程序提供的信息可以分为以下几类：

- 存储数据服务和功能。此类信息对 vSAN、Virtual Volumes 和 I/O 筛选器等功能至关重要。呈现这些功能的存储提供程序与基于存储策略的管理 (SPBM) 机制集成。存储提供程序将收集有关基础存储实体或可用 I/O 筛选器提供的数据服务的信息。

在存储策略中定义虚拟机和虚拟磁盘的存储要求时，请引用这些数据服务。根据您的环境，SPBM 机制可确保虚拟机的适当存储放置或启用虚拟磁盘的特定数据服务。有关详细信息，请参见[创建和管理虚拟机存储策略](#)。

- 存储状态。此类别包含有关各种存储实体的状态报告。它还包含用于通知有关配置变化的警报和事件。这种类型的信息可以帮助您解决存储连接和性能问题，还可以帮助您将阵列生成的事件和警报与阵列上相应的性能和负载变化相关联。

- 块设备或文件系统上 Distributed Resource Scheduling 的 Storage DRS 信息。此信息有助于确保 Storage DRS 所做的决定与存储系统内部的资源管理决定兼容。

## 存储提供程序要求和注意事项

使用第三方存储提供程序时，适用特定的要求和注意事项。

通常，供应商负责提供存储提供程序。VMware VASA 计划定义将第三方存储提供程序集成到 vSphere 环境中的架构，以便 vCenter Server 和 ESXi 主机可以与存储提供程序通信。

要使用存储提供程序，请遵循以下要求：

- 确保所使用的每个存储提供程序都经过 VMware 的认证并进行正确部署。有关部署存储提供程序的信息，请联系存储供应商。
- 确保存储提供程序与 vCenter Server 和 ESXi 版本兼容。请参见《VMware 兼容性指南》。
- 请勿将 VASA 提供程序安装到 vCenter Server 所在的系统上。
- 如果您的环境包含旧版的存储提供程序，现有功能可以继续运行。但是，要使用新功能，应将存储提供程序升级到新版本。
- 将存储提供程序升级到更高的 VASA 版本时，您必须取消注册并重新注册此提供程序。注册后，vCenter Server 可以检测并使用新的 VASA 版本的功能。

## 注册存储提供程序

要在 vCenter Server 和存储提供程序之间建立连接，必须注册存储提供程序。可使用 vSphere Client 为集群中的每个主机注册单独的存储提供程序。

将存储提供程序升级到更高的 VASA 版本时，您必须取消注册并重新注册此提供程序。注册后，vCenter Server 可以检测并使用更高的 VASA 版本的功能。

---

**注** 如果您使用 vSAN，则注册 vSAN 的存储提供程序，并且该提供程序会自动显示在存储提供程序列表中。vSAN 不支持手动注册存储提供程序。请参见《管理 VMware vSAN》文档。

---

### 前提条件

确认是否在存储端安装了存储提供程序组件，并从存储管理员那里获取其凭据。

### 步骤

- 1 导航到 vCenter Server。
- 2 依次单击**配置**选项卡和**存储提供程序**。
- 3 单击**添加**图标。
- 4 输入存储提供程序的连接信息，其中包括名称、URL 和凭证。

## 5 指定安全方法。

操作	描述
指示 vCenter Server 使用存储提供程序证书	选择 <b>使用存储提供程序证书</b> 选项并指定证书的位置。
使用存储提供程序证书的指纹	如果不指示 vCenter Server 使用提供程序证书，则将显示证书指纹。可以检查指纹并批准它。vCenter Server 将证书添加到信任库，然后继续进行连接。

当 vCenter Server 首次连接到提供程序时，存储提供程序会将 vCenter Server 证书添加到其信任库。

## 6 单击**确定**。

### 结果

vCenter Server 可注册存储提供程序，并建立与该存储提供程序的安全 SSL 连接。

### 后续步骤

要对存储提供程序注册进行故障排除，请参见 VMware 知识库文章 <https://kb.vmware.com/s/article/49798>。

## 查看存储提供程序信息

将存储提供程序组件注册到 vCenter Server 之后，存储提供程序将显示在存储提供程序列表中。某些存储提供程序会自行注册，并在您设置好它们呈现的实体后自动显示在此列表上，例如 vSAN 或 I/O 筛选器。

可以使用 vSphere Client 查看存储提供程序的一般信息以及每个存储组件的详细信息。

### 步骤

- 1 导航到 vCenter Server。
- 2 依次单击**配置**选项卡和**存储提供程序**。
- 3 在“存储提供程序”列表中，查看向 vCenter Server 注册的存储提供程序。  
此列表显示一般信息，如存储提供程序的名称、其 URL 和状态、VASA API 的版本、提供程序呈现的存储实体等。
- 4 要显示其他信息，请从列表中选择特定的存储提供程序或其组件。

**注** 一个存储提供程序可支持多个不同供应商的存储系统。

## 管理存储提供程序

可使用 vSphere Client 对注册的存储提供程序执行多项管理操作。

### 步骤

- 1 导航到 vCenter Server。

- 2 依次单击**配置**选项卡和**存储提供程序**。
- 3 从存储提供程序列表中选择存储提供程序，然后单击以下图标之一。

选项	描述
同步存储提供程序	将所有存储提供程序与环境的当前状态同步。
重新扫描	更新提供程序的存储数据。 vCenter Server 定期更新其数据库中的存储数据。这些更新不完整，仅反映当时存储提供程序传递到 vCenter Server 的更改。可以根据需要为所选存储提供程序执行数据库完全同步。
移除	取消注册不使用的存储提供程序。执行此操作后，vCenter Server 会关闭连接并从其配置中移除该存储提供程序。  <b>注</b> 无法手动取消注册 VMware 提供的一些存储提供程序，例如 vSAN 存储提供程序。  将存储提供程序升级到更高的 VASA 版本时，此选项也很有用。在这种情况下，您必须先取消注册该提供程序，然后再重新注册。注册后，vCenter Server 可以检测并使用更高的 VASA 版本的功能。
刷新证书	当分配给存储提供程序的证书即将到期时，vCenter Server 将向您发出警告。您可以刷新证书以继续使用该提供程序。 如果在到期前无法刷新证书，vCenter Server 将中止使用该提供程序。

## 结果

vCenter Server 会终止连接并从其配置中移除该存储提供程序。

# 使用 VMware vSphere Virtual Volumes

# 14

VMware vSphere Virtual Volumes（也称为 vVols）通过将物理硬件资源抽象到逻辑容量池来虚拟化 SAN 和 NAS 设备。Virtual Volumes 功能可将存储管理范式从管理数据存储内的空间更改为管理存储阵列处理的抽象存储对象。

本章讨论了以下主题：

- 关于 Virtual Volumes
- Virtual Volumes 概念
- Virtual Volumes 架构
- Virtual Volumes 和 VMware Certificate Authority
- 虚拟卷快照
- 在启用 Virtual Volumes 之前
- 配置 Virtual Volumes
- 配置 NVMe-oF-vVols
- 在 Virtual Volumes 数据存储中置备虚拟机
- Virtual Volumes 和复制
- 使用 Virtual Volumes 的最佳做法
- 对 Virtual Volumes 进行故障排除

## 关于 Virtual Volumes

使用 Virtual Volumes，单个虚拟机而不是数据存储将成为存储管理单元，而存储硬件可完全控制虚拟磁盘内容、布局和管理。

vSphere 存储管理以前采用以数据存储为中心的方法。使用这种方法，存储管理员和 vSphere 管理员可预先讨论虚拟机的基础存储要求。然后，存储管理员设置 NVMe 命名空间、LUN 或 NFS 共享并将其提供给 ESXi 主机。vSphere 管理员基于 NVMe 命名空间、LUN 或 NFS 创建数据存储，并将这些数据存储用作虚拟机存储。从存储角度而言，数据存储通常是发生数据管理的最低粒度级别。然而，单个数据存储包含可能具有不同要求的多个虚拟机。传统方法很难满足单个单个虚拟机的要求。



**Virtual Volumes** 功能有助于改善数据管理的粒度。它提供了一种新的存储管理方法，可以帮助您在每个应用程序级别对虚拟机服务进行不同的处理。**Virtual Volumes** 根据单个虚拟机的需求安排存储，而不是根据存储系统的功能安排存储，从而使存储以虚拟机为中心。

**Virtual Volumes** 可将虚拟磁盘及其衍生内容、克隆、快照和副本直接映射到存储系统上的对象（即虚拟卷）。此映射使 **vSphere** 可以将快照、克隆和复制等密集型存储操作转移到存储系统，从而减轻负担。

通过为每个虚拟磁盘创建一个卷，您可以在最佳级别设置策略。您可以预先决定应用程序的存储要求，并将这些要求传达给存储系统。存储系统会基于这些要求创建相应的虚拟磁盘。例如，如果虚拟机需要主动-主动存储阵列，则不必再选择支持主动-主动模式的数据存储，只需创建可自动放置在主动-主动阵列中的单个虚拟卷。

## Virtual Volumes 概念

使用 **Virtual Volumes** 时，抽象的存储容器将替换基于 **NVMe** 命名空间、**LUN** 或 **NFS** 共享的传统存储卷。在 **vCenter Server** 中，存储容器以 **Virtual Volumes** 数据存储表示。**Virtual Volumes** 数据存储存储虚拟卷，即封装虚拟机文件的对象。

要了解有关 **Virtual Volumes** 功能的不同组件的详细信息，请观看视频。



( **Virtual Volumes** 第 1 部分：概念 )

## 虚拟卷对象

虚拟卷用于封装虚拟机文件、虚拟磁盘及其衍生内容。

虚拟卷存储在通过以太网或 **SAN** 连接到 **ESXi** 主机的存储系统本机内部。虚拟卷由合规存储系统导出为对象，并完全通过存储端的硬件进行管理。通常，唯一的 **GUID** 可标识虚拟卷。系统不会预置备虚拟卷，而是在执行虚拟机管理操作时自动创建虚拟卷。这些操作包括虚拟机创建、克隆和快照。**ESXi** 和 **vCenter Server** 会将一个或多个虚拟卷关联到虚拟机。

## 虚拟卷的类型

系统将为构成虚拟机的核心元素创建以下类型的虚拟卷：

### 数据 vVol

直接与每个虚拟磁盘 **.vmdk** 文件对应的数据虚拟卷。作为传统数据存储上的虚拟磁盘文件，虚拟卷将作为 **SCSI** 磁盘提供给虚拟机。数据 **vVol** 可以厚置备，也可以精简置备。

### 配置 vVol

配置虚拟卷或主目录，它代表一个小目录，其中包含虚拟机的元数据文件。此类文件包括 **.vmx** 文件、虚拟磁盘的描述符文件、日志文件等等。配置虚拟卷将使用文件系统进行格式化。**ESXi** 使用 **SCSI** 或 **NVMe** 协议连接到存储时，配置虚拟卷将使用 **VMFS** 进行格式化。使用 **NFS** 协议，配置虚拟卷将显示为 **NFS** 目录。它通常精简置备。

### 交换 vVol

首次打开虚拟机电源时创建。该虚拟卷用于保存内存中无法保留的虚拟机内存页副本。它的大小由虚拟机的内存大小决定。它默认为厚置备。

## 快照 vVol

用于保存快照的虚拟机内存内容的虚拟内存卷。厚置备。

## 其他

用于特定功能的虚拟卷。例如，系统会为基于内容的读缓存 (CBRC) 创建一个摘要虚拟卷。

通常，虚拟机至少会创建三个虚拟卷：数据 vVol、配置 vVol 和交换 vVol。最大虚拟卷数量取决于虚拟机上驻留的虚拟磁盘数和快照数。

例如，以下 SQL Server 有六个虚拟卷：

- 配置 vVol
- 用于操作系统的数据 vVol
- 用于数据库的数据 vVol
- 用于日志的数据 vVol
- 打开电源时创建的交换 vVol
- 快照 vVol

通过为不同虚拟机组件使用不同虚拟卷，可在最佳粒度级别应用和操作存储策略。例如，与虚拟机引导磁盘的虚拟卷相比，包含虚拟磁盘的虚拟卷具有一组更丰富的服务。同样，与当前虚拟卷相比，快照虚拟卷可以使用不同的存储层。

## 磁盘置备

Virtual Volumes 功能支持精简置备虚拟磁盘和厚置备虚拟磁盘的概念。但是，从 I/O 角度而言，通过阵列实现和管理精简置备或厚置备对 ESXi 主机来说是透明的。ESXi 会将与精简置备相关的任何功能卸载到存储阵列。在数据路径中，ESXi 以相同方式处理精简置备虚拟卷和厚置备虚拟卷。

您可以在创建虚拟机时为虚拟磁盘选择精简置备或厚置备类型。如果磁盘为精简置备类型并驻留在 Virtual Volumes 数据存储上，则您以后无法通过扩充磁盘来更改其类型。

## 共享磁盘

您可以将共享磁盘放置到对 Virtual Volumes 支持 SCSI 永久预留的 Virtual Volumes 存储上。您可以使用此磁盘作为仲裁磁盘，并消除 MSCS 集群中的 RDM。有关详细信息，请参见《vSphere 资源管理》文档。

## Virtual Volumes 存储提供程序

Virtual Volumes 存储提供程序也称为 VASA 提供程序，是一款在 vSphere 环境中充当存储感知服务的软件组件。该提供程序可调节一端的 vCenter Server 和 ESXi 主机与另一端的存储系统之间的带外通信。

存储提供程序通过 VMware APIs for Storage Awareness (VASA) 实现，并用于管理 Virtual Volumes 存储的各个方面。存储提供程序与 vSphere 附带的存储监控服务 (SMS) 集成，从而与 vCenter Server 和 ESXi 主机进行通信。

存储提供程序可提供来自基础存储容器的信息。在 vCenter Server 和 vSphere Client 中都可以查看存储容器功能。反过来，存储提供程序会将虚拟机存储要求传达给存储层，您可以采用存储策略的形式定义这些要求。此集成过程可确保在存储层中创建的虚拟卷符合策略中概述的要求。

通常，供应商负责提供可与 vSphere 集成的存储提供程序，并为 Virtual Volumes 提供支持。每个存储提供程序必须经过 VMware 的认证并进行正确部署。有关部署 Virtual Volumes 存储提供程序以及将其升级到与当前 ESXi 版本兼容的版本的版本的信息，请联系存储供应商。

部署存储提供程序之后，必须在 vCenter Server 中注册该存储提供程序，才能使其通过 SMS 与 vSphere 进行通信。

## Virtual Volumes 存储容器

与基于传统块（SCSI 或 NVMe）或文件（NFS）的存储不同，Virtual Volumes 功能不需要存储端的预配置卷。相反，Virtual Volumes 需要使用存储容器。存储容器是原始存储容量池，或者说是存储系统可提供给虚拟卷的存储容量的聚合。

存储容器是逻辑存储结构的一部分，并且是基础硬件的逻辑单元。存储容器将根据管理需求对虚拟卷进行逻辑分组。例如，存储容器可包含为多租户部署中的租户或企业部署中的部门创建的所有虚拟卷。每个存储容器作为一个虚拟卷存储，且虚拟卷将在存储容器容量外进行分配。

通常，存储端的存储管理员将定义存储容器。存储容器的数量、存储容量和大小取决于供应商特定的实施。每个存储系统至少需要一个容器。

---

**注** 单个存储容器无法跨越不同的物理阵列。

---

注册与存储系统关联的存储提供程序后，vCenter Server 会发现所有已配置的存储容器及其存储功能配置文件、协议端点和其他属性。单个存储容器可导出多个功能配置文件。因此，具有不同需求和不同存储策略设置的虚拟机可以作为同一存储容器的一部分。

最初，所有发现的存储容器不会连接到任何特定主机，且您无法在 vSphere Client 中查看这些存储容器。要挂载存储容器，必须将其映射到 Virtual Volumes 数据存储。

## 协议端点

尽管存储系统管理虚拟卷的各个方面，但 ESXi 主机无法直接访问存储端上的虚拟卷。ESXi 主机使用称为协议端点的逻辑 I/O 代理与虚拟卷及虚拟卷封装的虚拟磁盘文件通信。ESXi 使用协议端点按需建立从虚拟机到其各自虚拟卷的数据路径。

每个虚拟卷都会绑定到特定的协议端点。当主机上的虚拟机执行 I/O 操作时，协议端点会将 I/O 定向到相应的虚拟卷。一般来说，存储系统只需要几个协议端点。一个协议端点可以连接到数百或数千个虚拟卷。

在存储端，存储管理员可为每个存储容器配置一个或多个协议端点。协议端点是物理存储架构的一部分。存储系统通过存储提供程序导出协议端点以及关联的存储容器。将存储容器映射到 Virtual Volumes 数据存储后，ESXi 主机可以发现协议端点，并且这些协议端点会显示在 vSphere Client 中。在存储重新扫描期间也可发现协议端点。多个主机可以发现和挂载协议端点。

在 vSphere Client 中，可用的协议端点列表类似于主机存储设备列表。不同的存储传输可用于向 ESXi 公开协议端点。如果使用基于 SCSI 的传输，则协议端点表示由基于 T10 的 LUN WWN 定义的代理 LUN。对于 NFS 协议，协议端点是挂载点，例如 IP 地址和共享名称。您可以在基于 SCSI 的协议端点上配置多路径，但不能在基于 NFS 的协议端点上执行此操作。无论使用何种协议，存储阵列都可以提供多个协议端点以实现可用性。

协议端点按阵列管理。ESXi 和 vCenter Server 假设针对阵列报告的所有协议端点都与该阵列上的所有容器关联。例如，如果某个阵列有两个容器和三个协议端点，ESXi 会假设这两个容器上的虚拟卷可以绑定到全部三个协议点。

## 虚拟协议端点

Virtual Volumes over NVMe 使用虚拟协议端点。这相当于将协议端点用于 SCSI。ANA 组提供称为虚拟协议端点 (vPE) 的代理表示形式。它们用于表示 PSA 中的协议端点。它们整合表示 vVols 的命名空间的数据路径和访问状态 (PE LUN 对 SCSI vVols 的作用)。NVMe 阵列将 vVol 命名空间放置在与多个控制器 (已优化、未优化、不可访问) 关联的 ANA 组中。主机会检测哪些控制器最适合访问成员命名空间。

虚拟协议端点本质上是动态的。您可以根据需要创建和销毁虚拟协议端点。此操作完全自动化。

例如，当 ESXi 主机尝试访问 ANA 组中的 vVol/命名空间时，您无法创建 vPE 设备。但是，当它尝试访问 ANA 组中的第一个 vVol/命名空间时，您可以在 NVMe 子系统中为该 ANA 组创建一个 vPE。此 vPE 将用于访问该 ANA 组中的所有命名空间。

## 绑定和解除绑定 Virtual Volumes

虚拟卷在创建时是被动实体，不会立即就可供进行 I/O 操作。要访问虚拟卷，ESXi 或 vCenter Server 会发送绑定请求。

存储系统使用变为虚拟卷访问点的协议端点 ID 进行响应。协议端点接受发送到虚拟卷的所有 I/O 请求。ESXi 发送虚拟卷解除绑定请求之前，此绑定一直存在。

对于相同虚拟卷上的后续绑定请求，存储系统会返回不同的协议端点 ID。

使用 NVMe 协议时，绑定虚拟卷响应会提供 NVMe 子系统 NQN 和命名空间/vVol 对象的 namespace-id (nsid)。ESXi 主机使用这些信息，并将其解析为子系统 ANA 组。如果与此 ANA 组相对应的 vPE 不存在，则会创建 vPE。vPE 用于将所有 I/O 请求定向到 Virtual Volumes。

从多个 ESXi 主机接收发送到虚拟卷的并行绑定请求时，存储系统会向每个请求 ESXi 主机返回相同或不同的端点绑定。也就是说，存储系统会通过不同端点将不同并行主机绑定到相同虚拟卷。

解除绑定操作会移除虚拟卷的 I/O 访问点。存储系统可能会立即解除绑定协议端点中的虚拟卷，也可能在延迟一段时候或采取其他一些操作之后解除绑定。解除绑定后，才能删除绑定的虚拟卷。

## Virtual Volumes 数据存储

Virtual Volumes 数据存储表示 vCenter Server 和 vSphere Client 中的存储容器。

vCenter Server 发现存储系统导出的存储容器后，您必须将它们挂载为 Virtual Volumes 数据存储。Virtual Volumes 数据存储不按传统方法进行格式化，例如，VMFS 数据存储。您仍需创建此类数据存储，因为所有 vSphere 功能（包括 FT、HA、DRS 等）均需要数据存储构造才能正常运行。

可以在 vSphere Client 中使用数据存储创建向导将存储容器映射到 Virtual Volumes 数据存储。您创建的 Virtual Volumes 数据存储直接对应于特定的存储容器。

从 vSphere 管理员角度而言，Virtual Volumes 数据存储类似于任何其他数据存储且用于保存虚拟机。与其他数据存储一样，Virtual Volumes 数据存储可供浏览并按虚拟机名称列出虚拟卷。与传统数据存储一样，Virtual Volumes 数据存储支持卸载和挂载操作。但是，升级和调整大小等操作不适用于 Virtual Volumes 数据存储。存储管理员在 vSphere 之外配置 Virtual Volumes 数据存储容量。

您可以将 Virtual Volumes 数据存储与传统 VMFS 和 NFS 数据存储以及 vSAN 配合使用。

---

**注** 虚拟卷的大小必须是 1 MB 的倍数，最小大小为 1 MB。因此，在 Virtual Volumes 数据存储中置备的所有虚拟磁盘必须为 1 MB 的偶数倍。如果迁移到 Virtual Volumes 数据存储的虚拟磁盘大小不是 1 MB 的偶数倍，请将磁盘大小扩展到最接近的 1 MB 的偶数倍。

---

## Virtual Volumes 和虚拟机存储策略

在 Virtual Volumes 数据存储中运行的虚拟机需要具有虚拟机存储策略。

虚拟机存储策略是一组包含虚拟机放置和服务质量要求的规则。该策略将在 Virtual Volumes 存储内强制实施虚拟机的适当位置放置，并保证存储可满足虚拟机的需求。

使用“虚拟机存储策略”界面创建 Virtual Volumes 存储策略。将新策略分配给虚拟机时，该策略将强制 Virtual Volumes 存储满足相关要求。

### Virtual Volumes 默认存储策略

对于 Virtual Volumes，VMware 提供不包含规则或存储要求的默认存储策略，称为 Virtual Volumes 无要求策略。如果没有为 Virtual Volumes 数据存储上的虚拟机指定其他策略，则向虚拟机对象应用此策略。使用“无要求”策略时，存储阵列可决定虚拟机对象的最佳放置。

VMware 提供的默认“无要求”策略具有以下特性：

- 无法删除、编辑或克隆此策略。
- 此策略仅与 Virtual Volumes 数据存储兼容。
- 您可以为 Virtual Volumes 创建虚拟机存储策略，并将其指定为默认策略。

## Virtual Volumes 和存储协议

Virtual Volumes 存储系统提供可在物理存储架构上发现的协议端点。ESXi 主机使用协议端点连接到存储上的虚拟卷。协议端点的操作取决于向 ESXi 主机公开端点的存储协议。

Virtual Volumes 支持 NFS 版本 3 和 4.1、iSCSI、光纤通道、FCoE 和 NVMe over FC。

不管使用何种存储协议，协议端点均会提供同时访问 SAN 和 NAS 存储的统一权限。虚拟卷类似于其他传统数据存储上的文件，可以作为 SCSI 磁盘提供给虚拟机。

---

**注** 存储容器专用于 SAN (SCSI/NVMe) 或 NAS，不能在这些协议类型之间共享。阵列可以提供一个具有 SCSI 协议端点的存储容器和另一个具有 NFS 协议端点的容器。容器不能同时使用 SCSI、NVMe 和 NFS 存储访问协议。

---

## Virtual Volumes 和基于 SCSI 的传输

在磁盘阵列上，Virtual Volumes 支持光纤通道、FCoE 和 iSCSI 协议。

如果使用基于 SCSI 的协议，则协议端点表示由基于 T10 的 LUN WWN 定义的代理 LUN。

和任何基于块的 LUN 一样，协议端点是使用标准 LUN 发现命令发现的。ESXi 主机定期重新扫描新设备，并异步发现基于块的协议端点。可以从多个路径访问协议端点。这些路径上的流量遵从已知的路径选择策略，这对 LUN 来说是典型的行为。

在基于 SCSI 的磁盘阵列上，在创建虚拟机时，ESXi 会创建虚拟卷，并将其格式化为 VMFS。该小容量虚拟卷存储所有虚拟机元数据文件，称为配置 vVol。配置 vVol 用作 vSphere 的虚拟机存储定位符。

磁盘阵列上的 Virtual Volumes 支持的 SCSI 命令集与 VMFS 相同，并使用 ATS 锁定机制。

## 对 iSCSI 端点支持 CHAP

Virtual Volumes 对 iSCSI 目标支持质询握手访问协议 (CHAP)。借助此支持，ESXi 主机可与 Virtual Volumes 存储提供程序（也称为 VASA 提供程序）共享 CHAP 启动器凭据，且 Virtual Volumes 存储提供程序能够发出系统事件，通知 vCenter Server 有关对存储阵列上的 CHAP 目标凭据所做的更改。

每个 ESXi 主机可以具有多个 HBA，并且每个 HBA 都可以配置属性。其中一个属性是 HBA 必须使用的身份验证方法。身份验证是可选项，但如果实施，则必须同时受启动器和目标支持。CHAP 是一种可在启动器和目标之间双向使用的身份验证方法。

有关不同 CHAP 身份验证方法的详细信息，请参见[选择 CHAP 身份验证方法](#)。要在 ESXi 主机上配置 CHAP，请参见[为 iSCSI 或 iSER 存储适配器配置 CHAP 参数](#)。

## Virtual Volumes 和 NFS 传输

对于 NAS 存储，协议端点是 ESXi 主机使用 IP 地址或 DNS 名称和共享名称挂载的 NFS 共享。Virtual Volumes 支持使用 NFS 版本 3 和 4.1 访问 NAS 存储。同时支持 IPv4 地址和 IPv6 格式。

无论使用哪个版本，存储阵列都可以提供多个协议端点以实现可用性。

此外，NFS 版本 4.1 引入了支持负载均衡和多路径的中继机制。

NAS 设备上的 Virtual Volumes 支持的 NFS 远程过程调用 (RPC) 与 ESXi 主机在连接到 NFS 挂载点时使用的 RPC 相同。

在 NAS 设备上，配置 vVol 是与配置 vVolID 对应的目录子树。配置 vVol 必须支持 NFS 所必需的目录和其他操作。

## Virtual Volumes 和 NVMe

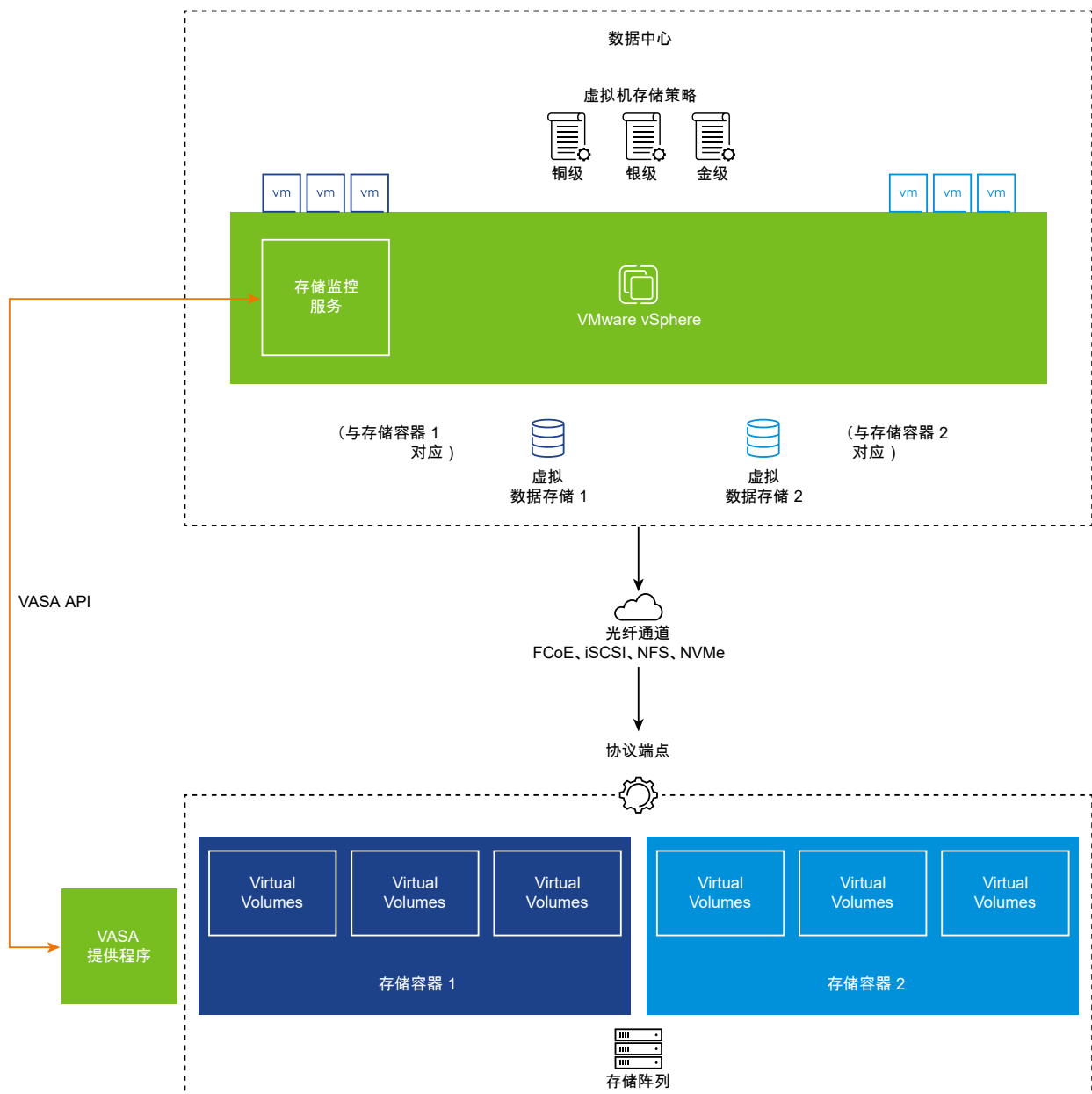
Virtual Volumes 支持 NVMe over Fibre Channel 协议。vVol 对象映射到 NVMe 子系统命名空间。NVMe 子系统命名空间中的 ANA 组被视为 ESXi 主机上的虚拟协议端点。

虚拟协议端点用于在 ANA 组状态更改时进行路径状态管理。ESXi 主机按需动态发现 ANA 组。这意味着，仅当 ESXi 主机需要对 NVMe 子系统命名空间中的 vVol/命名空间进行 I/O 访问时，才会创建虚拟协议端点。NVMe 上的 Config-vVol 类似于使用 VMFS 格式化的 SCSI。它们还用于存储虚拟机元数据文件。

要在 ESXi 主机上配置 NVMe-oF-vVol，请参见 [配置 NVMe-oF-vVols](#)。

## Virtual Volumes 架构

架构图提供 Virtual Volumes 功能的所有组件如何互相交互的概览。





虚拟卷是由合规的存储系统导出的对象，且通常与虚拟机磁盘和其他虚拟机相关文件一一对应。虚拟卷由 VASA 提供程序创建，并在带外而不是在数据路径中操作。

VASA 提供程序或存储提供程序通过 vSphere APIs for Storage Awareness 进行开发。存储提供程序可实现一端上的 ESXi 主机、vCenter Server 和 vSphere Client 与另一端上的存储系统之间的通信。VASA 提供程序在存储端运行并与 vSphere 存储监控服务 (SMS) 集成以管理 Virtual Volumes 存储的各个方面。VASA 提供程序可将虚拟磁盘对象及其衍生内容（如克隆、快照和副本）直接映射到存储系统上的虚拟卷中。

ESXi 主机无权直接访问虚拟卷存储。该主机通过数据路径的中间点（称为协议端点）访问虚拟卷。协议端点按需建立从虚拟机到其各自虚拟卷的数据路径。协议端点作为 ESXi 主机与存储系统之间的直接带内 I/O 的网关。ESXi 可以使用 NVMe over Fibre Channel、SCSI over Fibre Channel、FCoE、iSCSI 和 NFS 协议进行带内通信。

虚拟卷驻留在逻辑上表示存储系统上的物理磁盘池的存储容器内。在 vCenter Server 和 ESXi 端，存储容器以 Virtual Volumes 数据存储表示。单个存储容器可以导出多个存储功能集，并可以向不同虚拟卷提供不同的服务水平。

有关 Virtual Volumes 架构的信息，请观看视频。



( Virtual Volumes 第 2 部分：架构 )

## Virtual Volumes 和 VMware Certificate Authority

vSphere 包含 VMware Certificate Authority (VMCA)。默认情况下，VMCA 将创建 vSphere 环境中使用的所有内部证书。它会为新添加的 ESXi 主机以及管理或代表 Virtual Volumes 存储系统的存储 VASA 提供程序生成证书。

与 VASA 提供程序的通信受 SSL 证书保护。这些证书可以来自 VASA 提供程序或 VMCA。

- 证书可以直接由 VASA 提供程序提供以供长期使用。证书可以自行生成和自行签名，也可以派生自外部证书颁发机构。
- 证书可由 VMCA 生成以供 VASA 提供程序使用。

注册主机或 VASA 提供程序后，VMCA 会自动按照以下步骤操作，而无需 vSphere 管理员参与。

- 1 先将 VASA 提供程序添加到 vCenter Server 存储管理服务 (SMS) 后，它会生成自签名证书。
- 2 对证书进行验证后，SMS 会向 VASA 提供程序请求证书签名请求 (CSR)。
- 3 接收并验证 CSR 后，SMS 会代表 VASA 提供程序将其提供给 VMCA 以请求 CA 签名证书。

VMCA 可配置为充当独立 CA 或充当企业 CA 的辅助机构。如果将 VMCA 设置为辅助 CA，则 VMCA 会通过完整链对 CSR 进行签名。

- 4 包含根证书的已签名证书将传递到 VASA 提供程序。VASA 提供程序可以对将来所有源自 vCenter Server 和 ESXi 主机上的 SMS 的安全连接进行身份验证。

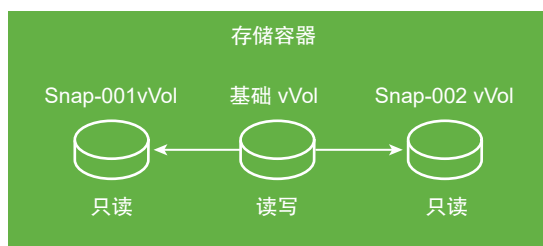


## 虚拟卷快照

当您执行快照时将保留虚拟机的状况和数据。在需要多次恢复至相同虚拟机状况而又不想创建多个虚拟机时，快照会很有用。虚拟卷快照有许多用途。您可以使用它们创建用于备份或存档的静默副本，或者为应用程序创建测试和回滚环境。您还可以使用它们即时置备应用程序映像。

在 **Virtual Volumes** 环境中，快照由 **ESXi** 和 **vCenter Server** 管理，但由存储阵列执行。

每个快照都会创建额外的虚拟卷对象（快照虚拟卷），以用来保存虚拟机内存的内容。原始虚拟机数据复制到此对象后会保持只读状态，这样就可以防止客户机操作系统写入快照。您不能调整快照虚拟卷的大小。一般情况下，当您复制虚拟机时，也会复制它的快照虚拟卷。



基础虚拟卷保持活动或可读写状态。创建另一个快照时，它会保留您生成快照时虚拟机的新状态和数据。

删除快照时，仅保留基础虚拟卷，而丢弃快照虚拟卷对象。基础虚拟卷表示虚拟机的最新状态。与传统数  
据存储上的快照不同，快照虚拟卷不需要将内容提交到基础虚拟卷。



有关创建和管理快照的信息，请参见《《vSphere 虚拟机管理》》文档。

## 在启用 Virtual Volumes 之前

要使用 **Virtual Volumes**，请确保正确设置存储和 **vSphere** 环境。

### 为 Virtual Volumes 准备存储系统

要为 **Virtual Volumes** 准备存储系统环境，请遵循以下准则。有关其他信息，请与存储供应商联系。

- 使用的存储系统或存储阵列必须支持 **Virtual Volumes** 并通过 **vSphere APIs for Storage Awareness (VASA)** 与 **vSphere** 组件集成。存储阵列必须支持精简置备和快照。
- 必须部署 **Virtual Volumes** 存储提供程序。
- 必须在存储端配置以下组件：
  - 协议端点（FC-NVMe 除外）

- 存储容器
- 存储配置文件
- 复制配置（如果计划使用 Virtual Volumes 进行复制）。请参见 [Virtual Volumes 复制的要求](#)。
- 如果使用 NVMeoF，请创建 NVM 子系统，并将 vVols 主机 NQN 与相关的 NVM 子系统相关联。此配置步骤取决于合作伙伴特定的 NVMe 设计。有关详细信息，请联系阵列供应商。

## 准备 vSphere 环境

- 确保遵循您使用的存储类型、FC-NVMe、光纤通道、FCoE、iSCSI 或 NFS 的相应设置准则。如果需要，请在 ESXi 主机上安装和配置存储适配器。
  - 如果使用 iSCSI，则请激活 ESXi 主机上的软件 iSCSI 适配器。配置动态发现，然后输入 Virtual Volumes 存储系统的 IP 地址。请参见[配置软件 iSCSI 适配器](#)。
- 将存储阵列中的所有组件与 vCenter Server 和所有 ESXi 主机同步。使用网络时间协议 (NTP) 来执行此同步。

有关详细信息，请联系供应商并参见《VMware 兼容性指南》。

## 将 vSphere Storage 环境与网络时间服务器同步

如果使用 Virtual Volumes，请配置网络时间协议 (NTP) 以确保 vSphere 网络中的所有 ESXi 主机均同步。

### 步骤

- 1 在 vSphere Client 中，导航到 ESXi 主机。
- 2 单击**配置**选项卡。
- 3 在**系统**下，选择**时间配置**。
- 4 单击**编辑**并设置 NTP 服务器。
  - a 选择**使用网络时间协议 (启用 NTP 客户端)**。
  - b 设置 NTP 服务启动策略。
  - c 输入要与其同步的 NTP 服务器的 IP 地址。
  - d 在“NTP 服务状态”部分中单击**启动**或**重新启动**。
- 5 单击**确定**。

此时，主机将与 NTP 服务器同步。

## 配置 Virtual Volumes

要配置 Virtual Volumes 环境，请遵循几个步骤。

### 前提条件

请遵循在启用 [Virtual Volumes](#) 之前中的准则。

### 步骤

#### 1 为 Virtual Volumes 注册存储提供程序

Virtual Volumes 环境必须包括存储提供程序（也称为 VASA 提供程序）。通常，第三方供应商通过 VMware APIs for Storage Awareness (VASA) 开发存储提供程序。存储提供程序可促进 vSphere 与存储端之间的通信。可使用 vSphere Client 注册 Virtual Volumes 存储提供程序。

#### 2 创建 Virtual Volumes 数据存储

您可以使用[新建数据存储](#)向导创建 Virtual Volumes 数据存储。

#### 3 查看和管理协议端点

ESXi 主机使用称为协议端点的逻辑 I/O 代理与虚拟卷及虚拟卷封装的虚拟磁盘文件通信。协议端点和关联的存储容器将由存储系统通过存储提供程序导出。将存储容器映射到 Virtual Volumes 数据存储后，协议端点将在 vSphere Client 中变得可见。您可以查看协议端点的属性并修改特定设置。这不适用于 NVMe over Virtual Volumes。

#### 4 （可选）更改协议端点的路径选择策略

如果 ESXi 主机使用基于 SCSI 的传输与表示存储阵列的协议端点进行通信，则可以修改分配给协议端点的默认多路径策略。使用[编辑多路径策略](#)对话框更改路径选择策略。

### 后续步骤

您现在可以在 Virtual Volumes 数据存储中置备虚拟机。有关创建虚拟机的信息，请参见在 [Virtual Volumes 数据存储中置备虚拟机](#)和《vSphere 虚拟机管理》文档。

## 为 Virtual Volumes 注册存储提供程序

Virtual Volumes 环境必须包括存储提供程序（也称为 VASA 提供程序）。通常，第三方供应商通过 VMware APIs for Storage Awareness (VASA) 开发存储提供程序。存储提供程序可促进 vSphere 与存储端之间的通信。可使用 vSphere Client 注册 Virtual Volumes 存储提供程序。

注册后，Virtual Volumes 提供程序将与 vCenter Server 进行通信。提供程序将报告存储系统提供的基础存储和数据服务（例如，复制）的特性。这些特性将显示在“虚拟机存储策略”界面中，可用于创建与 Virtual Volumes 数据存储兼容的虚拟机存储策略。对虚拟机应用此存储策略后，该策略将被推送到 Virtual Volumes 存储。策略将在 Virtual Volumes 存储内强制实施虚拟机最佳位置放置，并保证存储可满足虚拟机的需求。如果您的存储提供额外服务，例如缓存或复制，则策略将为虚拟机启用这些服务。

### 前提条件

确认在存储端安装了相应版本的 Virtual Volumes 存储提供程序。获取存储提供程序的凭据。

步骤

- 1 导航到 vCenter Server。
- 2 依次单击配置选项卡和存储提供程序。
- 3 单击添加图标。
- 4 输入存储提供程序的连接信息，其中包括名称、URL 和凭证。
- 5 指定安全方法。

操作	描述
指示 vCenter Server 使用存储提供程序证书	选择使用存储提供程序证书选项并指定证书的位置。
使用存储提供程序证书的指纹	如果不指示 vCenter Server 使用提供程序证书，则将显示证书指纹。可以检查指纹并批准它。vCenter Server 将证书添加到信任库，然后继续进行连接。

当 vCenter Server 首次连接到提供程序时，存储提供程序会将 vCenter Server 证书添加到其信任库。

- 6 要完成注册，请单击确定。

结果

vCenter Server 将发现并注册 Virtual Volumes 存储提供程序。

创建 Virtual Volumes 数据存储

您可以使用新建数据存储向导创建 Virtual Volumes 数据存储。

步骤

- 1 在 vSphere Client 对象导航器中，浏览到主机、集群或数据中心。
- 2 从右键单击菜单中，选择存储 > 新建数据存储。
- 3 选择 VVol 作为数据存储类型。
- 4 输入数据存储名称，并从存储容器列表中选择备用存储容器。  
确保使用的名称不会与数据中心环境中的其他数据存储名称重复。  
如果将同一 Virtual Volumes 数据存储挂载到多台主机上，则该数据存储的名称必须在所有主机中保持一致。
- 5 选择需要访问数据存储的主机。
- 6 检查配置选项，然后单击完成。

后续步骤

创建 Virtual Volumes 数据存储后，可以执行重命名数据存储、浏览数据存储文件、卸载数据存储等数据存储操作。

无法将 Virtual Volumes 数据存储添加到数据存储集群中。

## 查看和管理协议端点

ESXi 主机使用称为协议端点的逻辑 I/O 代理与虚拟卷及虚拟卷封装的虚拟磁盘文件通信。协议端点和关联的存储容器将由存储系统通过存储提供程序导出。将存储容器映射到 **Virtual Volumes** 数据存储后，协议端点将在 vSphere Client 中变得可见。您可以查看协议端点的属性并修改特定设置。这不适用于 NVMe over Virtual Volumes。

### 步骤

- 1 在 vSphere Client 中，导航到 ESXi 主机。
- 2 单击**配置**选项卡。
- 3 在**存储**下，单击**协议端点**。
- 4 要查看特定项目的详细信息，请从列表中选择该项目。
- 5 使用“协议端点详细信息”下的选项卡可访问其他信息，并可修改所选协议端点的属性。

选项卡	描述
属性	查看项目属性和特性。对于 SCSI（块）项目，请查看和编辑多路径策略。
路径（仅限 SCSI 协议端点）	显示协议端点的可用路径。禁用或启用所选路径。更改路径选择策略。
数据存储	显示相应的 Virtual Volumes 数据存储。执行数据存储管理操作。

## 更改协议端点的路径选择策略

如果 ESXi 主机使用基于 SCSI 的传输与表示存储阵列的协议端点进行通信，则可以修改分配给协议端点的默认多路径策略。使用**编辑多路径策略**对话框更改路径选择策略。

### 步骤

- 1 在 vSphere Client 中，导航到 ESXi 主机。
- 2 单击**配置**选项卡。
- 3 在**存储**下，单击**协议端点**。
- 4 选择要更改其路径的协议端点，然后单击**属性**选项卡。
- 5 在“多路径策略”下，从**操作**菜单中选择**编辑多路径**。
- 6 选择路径策略并配置其设置。这些选项会发生更改，具体取决于所使用存储设备的类型。

可供您选择的路径策略取决于存储供应商的支持。

- 有关 SCSI 设备的路径策略的信息，请参见**路径选择插件和策略**。
- 有关 NVMe 设备的路径机制的信息，请参见 **VMware 高性能插件和路径选择方案**。

- 7 要保存设置并退出对话框，请单击**确定**。

## 配置 NVMe-oF-vVols

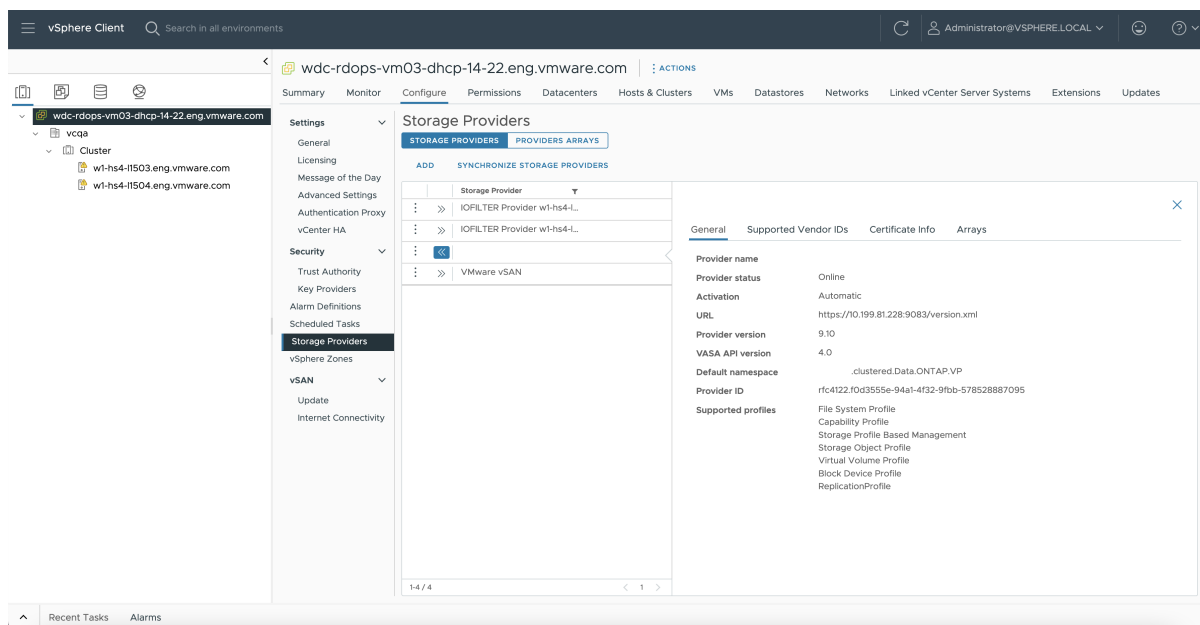
本部分介绍在 ESXi 主机上配置 NVMe over Virtual Volumes 的步骤。

### 前提条件

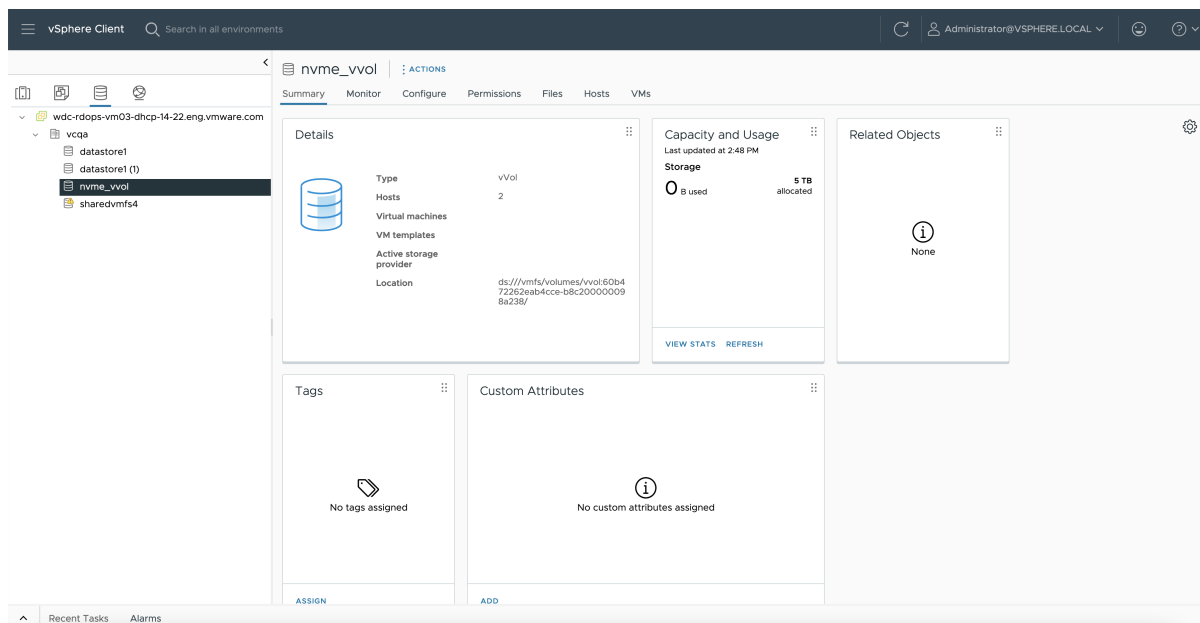
- 确保已准备好 vSphere 环境和存储系统。有关详细信息，请参见在启用 [Virtual Volumes](#) 之前

### 步骤

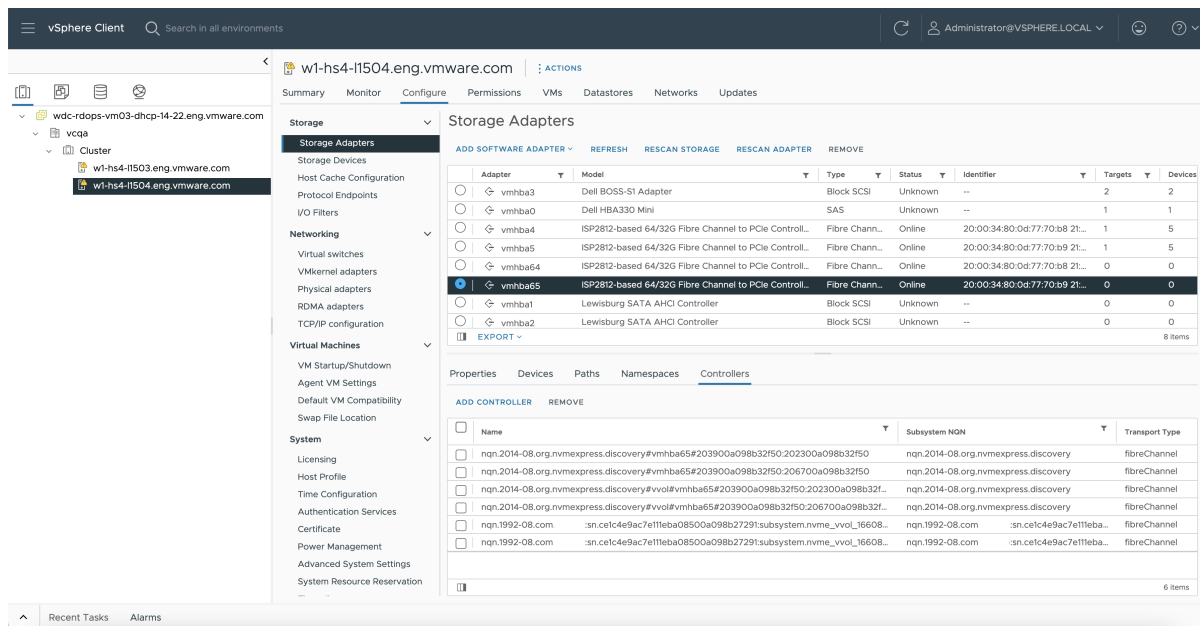
- 1 注册支持 VASA 4.0 的 VASA 提供程序。请确保将其提供给具有 NVMe-oF-vVol 功能的 VASA 存储阵列。



- 2 使用 VP 中支持 NVMe-oF 功能的存储容器将 NVMe vvol 数据存储挂载到所需的主机。



### 3 控制器配置如下所示



**注** 对于 Virtual Volumes，控制器发现和管理完全自动和动态执行。可按需创建和删除控制器。

## 在 Virtual Volumes 数据存储中置备虚拟机

您可以在 Virtual Volumes 数据存储中置备虚拟机。

**注** 在 Virtual Volumes 数据存储中置备的所有虚拟磁盘必须为 1 MB 的偶数倍。

在 Virtual Volumes 数据存储中运行的虚拟机需要相应的虚拟机存储策略。

置备虚拟机后，可以执行典型的虚拟机管理任务。有关信息，请参见《vSphere 虚拟机管理》文档。

#### 步骤

#### 1 为 Virtual Volumes 定义虚拟机存储策略。

VMware 为 Virtual Volumes 提供了默认的“无要求”存储策略。如果需要，您可以创建与 Virtual Volumes 兼容的自定义存储策略。

请参见为 [Virtual Volumes 创建虚拟机存储策略](#)。

#### 2 将 Virtual Volumes 存储策略分配给虚拟机。

要确保在分配虚拟机时 Virtual Volumes 数据存储满足特定存储要求，请将 Virtual Volumes 存储策略与虚拟机相关联。

请参见[将存储策略分配给虚拟机](#)。

#### 3 更改 Virtual Volumes 数据存储的默认存储策略。

对于已在 Virtual Volumes 数据存储上置备的虚拟机，VMware 将提供默认的“无要求”策略。您无法编辑该策略，但可以将新创建的策略指定为默认策略。

请参见[更改数据存储的默认存储策略](#)。

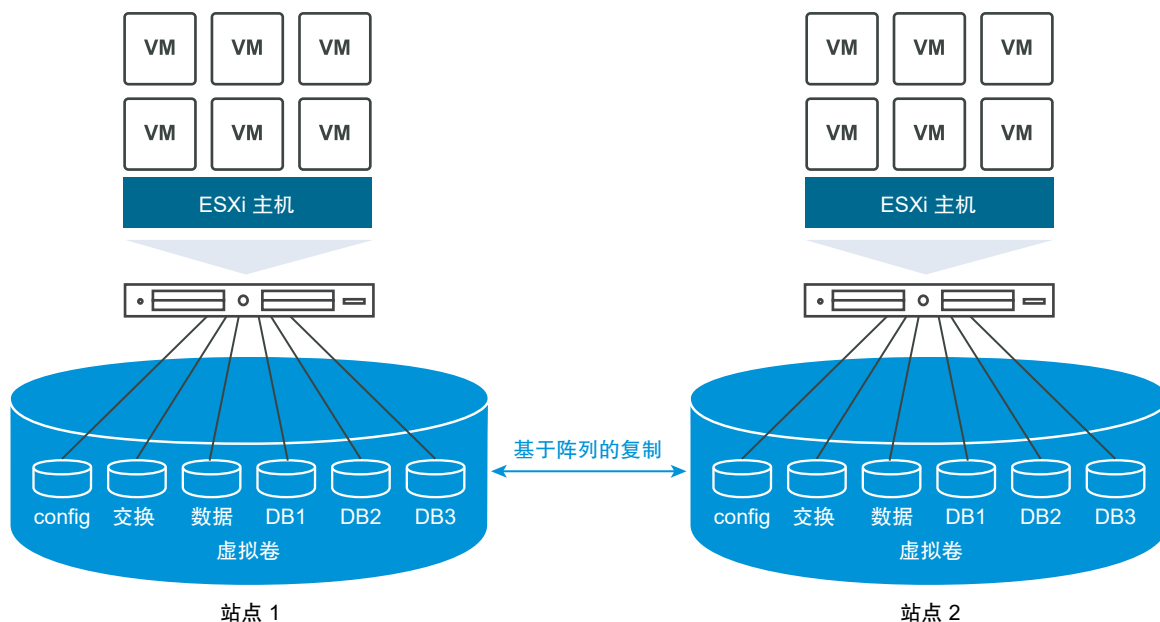
## Virtual Volumes 和复制

Virtual Volumes 支持复制和灾难恢复。借助基于阵列的复制，可将虚拟机复制卸载到存储阵列，并使用阵列的全部复制功能。您可以复制单个虚拟机对象，如虚拟磁盘。还可以分组多个虚拟机对象或虚拟机以将它们作为一个单元进行复制。

基于阵列的复制受策略驱动。配置 Virtual Volumes 存储进行复制后，存储提供程序从阵列提供有关复制功能和复制组的信息。此信息显示在 vCenter Server 的“虚拟机存储策略”界面中。

可以使用虚拟机存储策略描述虚拟机的复制要求。在存储策略中指定的参数取决于阵列如何实施复制。例如，您的虚拟机存储策略可能包括复制调度、复制频率或恢复点目标 (RPO) 等参数。策略还必须指明复制目标（即在其中复制虚拟机的辅助站点），或者指定是否必须删除副本。

通过在虚拟机置备期间分配复制策略，可以为虚拟机请求复制服务。然后，阵列将接管所有复制调度和进程的管理工作。



## Virtual Volumes 复制的要求

启用 Virtual Volumes 复制后，除了一般的 Virtual Volumes 要求，您的环境还必须满足几项特定的必备条件。

有关一般 Virtual Volumes 要求，请参见在[启用 Virtual Volumes 之前](#)。

### 存储要求

Virtual Volumes 复制的实现取决于您的阵列，并且可能因存储供应商而异。一般而言，以下要求适用于所有供应商。

- 用于实现复制的存储阵列必须与 Virtual Volumes 兼容。



- 这些阵列必须与 Virtual Volumes 复制兼容的存储 (VASA) 提供程序版本相集成。
- 存储阵列必须支持复制，且必须配置为使用供应商提供的复制机制。典型的复制通常包括一个或两个复制目标。还必须在存储端执行任何需要的配置，例如配对复制站点和目标站点。
- 如果适用，必须在存储端预配置 Virtual Volumes 的复制组和故障域。

有关详细信息，请联系供应商并参见《VMware 兼容性指南》。

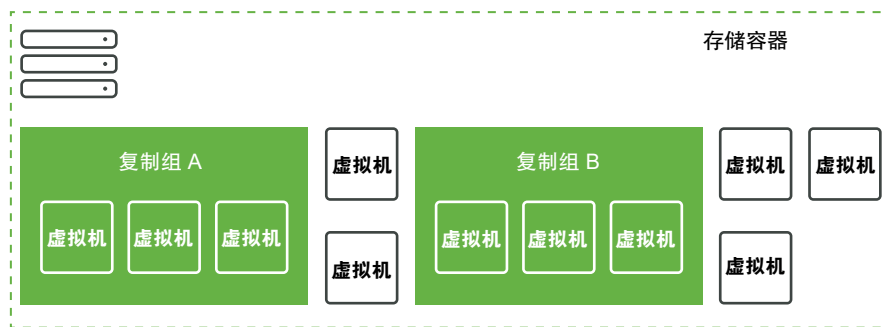
## vSphere 要求

- 使用支持 Virtual Volumes 存储复制的 vCenter Server 和 ESXi 版本。早于 6.5 版本的 vCenter Server 和 ESXi 主机不支持复制的 Virtual Volumes 存储。在不兼容主机上创建复制虚拟机的所有尝试均失败，并显示错误。有关信息，请参见《VMware 兼容性指南》。
- 如果计划迁移虚拟机，请确保目标资源（例如 ESXi 主机和 Virtual Volumes 数据存储）支持存储复制。

## Virtual Volumes 和复制组

当您的存储提供复制服务时，除了存储容器和协议端点，存储管理员还可以在存储端上配置复制组。

vCenter Server 和 ESXi 可以发现复制组，但不管理其生命周期。复制组（也称为一致性组）指明需要一起复制到目标站点的虚拟机和虚拟磁盘。您可以将同一虚拟机的组件（例如虚拟机配置文件和虚拟磁盘）分配给不同的预配置复制组。或者从复制中排除某些虚拟机组件。



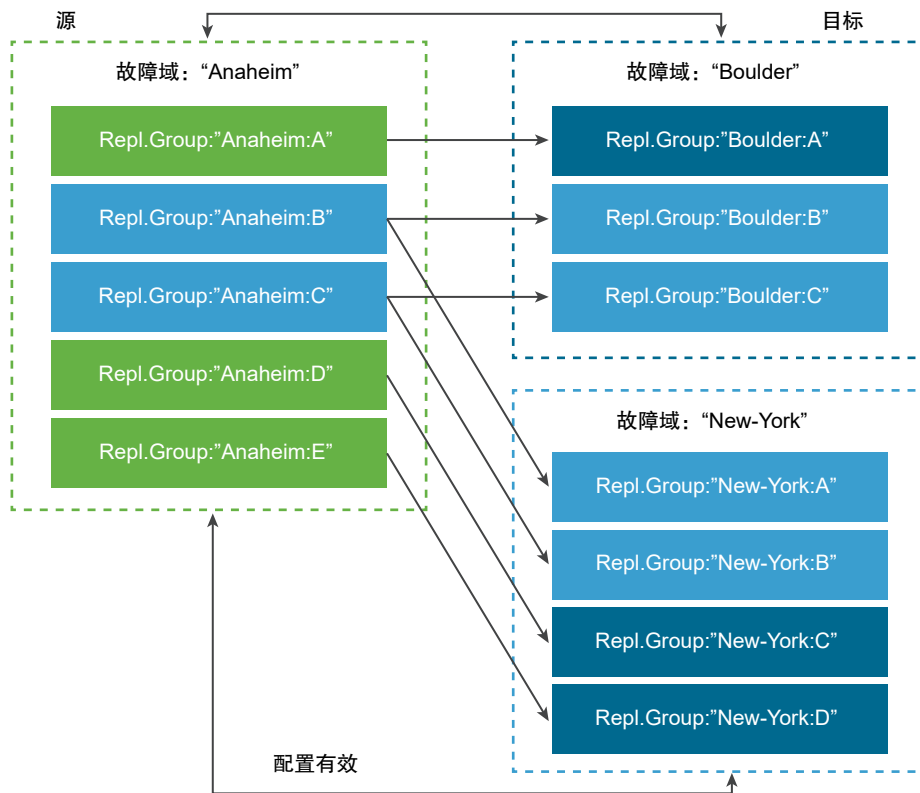
如果没有预配置的组可用，Virtual Volumes 可以使用自动方法。Virtual Volumes 可以使用自动方法根据需要创建复制组，并将此组与正在置备的 Virtual Volumes 对象相关联。如果使用自动复制组，虚拟机的所有组件都将分配给该组。对于同一虚拟机的组件，不能混合预配置复制组和自动复制组。

## Virtual Volumes 和故障域

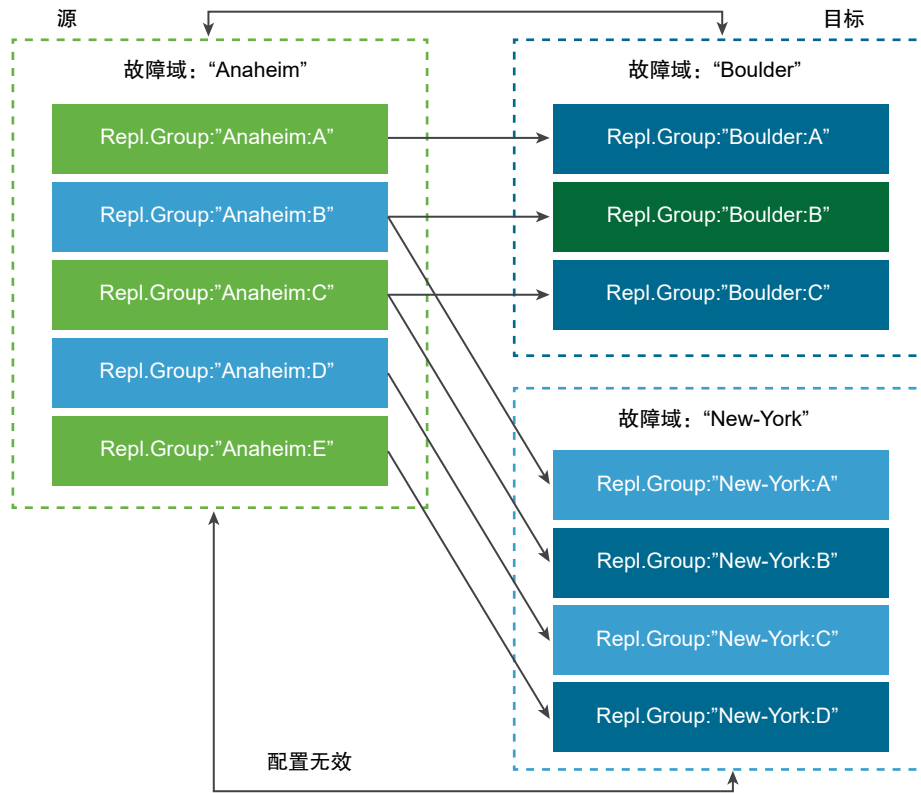
在 Virtual Volumes 环境中，故障域定义如何必须在从源站点复制到目标站点时组合特定复制组。

故障域由存储阵列配置和报告，但不在 vSphere Client 中公开。基于存储策略的管理 (Storage Policy Based Management, SPBM) 机制发现故障域，并在创建虚拟机期间使用它们进行验证。

例如，置备一个包含两个磁盘的虚拟机，其中一个磁盘与复制组 Anaheim:B 关联，另一个磁盘与复制组 Anaheim:C 关联。SPBM 可以验证置备，因为这两个磁盘都复制到相同的目标故障域中。



现在置备一个包含两个磁盘的虚拟机，其中一个磁盘与复制组 Anaheim:B 关联，另一个磁盘与复制组 Anaheim:D 关联。此配置无效。这两个复制组都会复制到 New-York 故障域，但只有一个会复制到 Boulder 故障域。



## Virtual Volumes 复制 workflow

如果 vCenter Server 中显示了有关 Virtual Volumes 存储阵列复制功能的信息，则可以为您的虚拟机激活复制。

为虚拟机激活复制的工作流包括在 Virtual Volumes 存储上置备虚拟机的典型步骤。

- 1 定义与复制存储兼容的虚拟机存储策略。策略中基于数据存储的规则必须包含复制组件。请参见为 [Virtual Volumes 创建虚拟机存储策略](#)。

配置包含复制的存储策略后，vCenter Server 会发现可用的复制组。

- 2 将复制策略分配给虚拟机。配置后，选择兼容的复制组，或者使用自动分配。请参见 [将存储策略分配给虚拟机](#)。

## 复制准则和注意事项

通过 Virtual Volumes 进行复制时，应注意以下特定注意事项。

- 只能将复制存储策略应用于配置虚拟卷和数据虚拟卷。其他虚拟机对象通过以下方式继承复制策略：
  - 内存虚拟卷继承配置虚拟卷的策略。
  - 摘要虚拟卷继承数据虚拟卷的策略。
  - 交换虚拟卷（打开虚拟机电源时存在）从复制中排除。
- 如果未向某个虚拟机磁盘应用复制策略，将不会复制该磁盘。

- 复制存储策略不应作为数据存储的默认存储策略。否则，该策略将阻止您选择复制组。
- 复制会保留快照历史记录。如果创建并复制快照，则可以恢复到应用程序一致的快照。
- 您可以复制链接克隆。如果复制链接克隆但不复制其父项，则它会成为完整克隆。
- 如果描述符文件属于一个虚拟机的虚拟磁盘，但驻留在另一个虚拟机的虚拟机主页中，那么这两个虚拟机必须在同一复制组中。如果这两个虚拟机在不同的复制组中，那么这两个复制组必须同时进行故障切换。否则，描述符可能会在故障切换后不可用，并导致无法打开虚拟机的电源。
- 在具有复制功能的 **Virtual Volumes** 环境中，您可能会定期运行测试故障切换工作流，确保恢复的工作负载在故障切换后能够正常运行。

测试故障切换期间创建的测试虚拟机功能完全正常，并且适用于一般管理操作。尽管如此，仍需注意以下事项：

- 在测试故障切换停止之前，必须删除测试故障切换期间创建的所有虚拟机。删除操作可确保虚拟机的任何快照或快照相关的虚拟卷（如快照虚拟卷）不会影响测试故障切换的停止操作。
- 您可以创建测试虚拟机的完整克隆。
- 仅当应用于新虚拟机的策略包含与被克隆的虚拟机相同的复制组 ID 时，才可以创建快速克隆。尝试将子虚拟机置于父虚拟机复制组外部的操作失败。

## 使用 Virtual Volumes 的最佳做法

将 Virtual Volumes 与 ESXi 和 vCenter Server 结合使用时，请遵循下列建议。

- **使用 Virtual Volumes 时的准则和限制**  
要在使用 Virtual Volumes 功能时获得最佳体验，必须遵循特定准则。
- **存储容器置备最佳做法**  
在 Virtual Volumes 阵列端置备存储容器时，请遵循以下最佳做法。
- **Virtual Volumes 性能最佳做法**  
为了确保实现 Virtual Volumes 最佳性能，请遵循以下建议。

## 使用 Virtual Volumes 时的准则和限制

要在使用 Virtual Volumes 功能时获得最佳体验，必须遵循特定准则。

Virtual Volumes 支持以下功能、特性和 VMware 产品：

- 借助 Virtual Volumes，您可以使用高级存储服务，其中包括对单个虚拟磁盘执行复制、加密、去重和压缩。联系存储供应商了解有关他们支持的 Virtual Volumes 服务的信息。

- Virtual Volumes 功能支持使用 vSphere APIs - Data Protection 的备份软件。虚拟卷根据虚拟磁盘建模。使用 vSphere APIs - Data Protection 的备份产品在虚拟卷上完全受支持，就像其在 LUN 上的 VMDK 文件上完全受支持一样。备份软件使用 vSphere APIs - Data Protection 创建的快照被 vSphere 和备份软件视为非虚拟卷快照。

---

**注** Virtual Volumes 不支持 SAN 传输模式。vSphere APIs - Data Protection 会自动选择其他数据传输方法。

---

有关与 vSphere Storage APIs - Data Protection 集成的更多信息，请咨询备份软件供应商。

- Virtual Volumes 支持 vSphere vMotion、Storage vMotion、快照、链接克隆和 DRS 等 vSphere 功能。
- 您可以将集群产品（如 Oracle Real Application Clusters）与 Virtual Volumes 结合使用。要使用这些产品，可以为存储在 Virtual Volumes 数据存储上的虚拟磁盘激活多写入设置。

有关更多详细信息，请参见位于 <http://kb.vmware.com/kb/2112039> 的知识库文章。有关 Virtual Volumes 功能支持的功能和产品的列表，请参见 VMware 产品互操作性列表。

## Virtual Volumes 限制

了解以下限制有助于获得更好的 Virtual Volumes 使用体验：

- 由于 Virtual Volumes 环境需要 vCenter Server，因此无法将 Virtual Volumes 和独立主机配合使用。
- Virtual Volumes 功能不支持 RDM。
- Virtual Volumes 存储容器不能跨多个物理阵列。某些供应商将多个物理阵列呈现为单个阵列。在这种情况下，从技术方面来讲，仍使用一个逻辑阵列。
- 包含 Virtual Volumes 数据存储的主机配置文件特定于 vCenter Server。提取此类型的主机配置文件后，只能将其附加到由管理引用主机的相同 vCenter Server 管理的主机和集群。有关详细信息，请参见“最高配置”文档。

## NVMe 的 Virtual Volumes 限制

特定于 vSphere 8.0 版本的 NVMe over Virtual Volumes 不支持以下功能：

- 复制和 Site Recovery Manager 操作。
- NVMe 预留。
- NVMe 解除分配和取消映射操作。
- 在虚拟机之间共享磁盘。
- NVMe 带内迁移。

## 存储容器置备最佳做法

在 Virtual Volumes 阵列端置备存储容器时，请遵循以下最佳做法。

### 根据相应限制创建容器

由于存储容器在对虚拟卷进行分组时应用逻辑限制，因此容器必须与要应用的边界匹配。

例如，多租户部署中为某个租户创建的容器，或者企业部署中用于某个部门的容器。

- 组织或部门，例如，人力资源和财务
- 团队或项目，例如，团队 A 和红队。
- 客户

### 将所有存储功能集中到一个容器中

每个存储容器都是一个数据存储。一个存储容器可以导出多个存储功能配置文件。因此，具有不同需求和不同存储策略设置的虚拟机可以作为同一存储容器的一部分。

更改存储配置文件必须在阵列端执行，而不是将存储迁移到另一容器。

### 避免过度置备存储容器

置备存储容器时，在容器配置过程中应用的空间限制仅为逻辑限制。置备容器时，不要超过满足预期用途所需的大小。如果之后增加容器的大小，无需重新设置其格式或重新进行分区。

### 使用存储特定的管理 UI 置备协议端点

每个存储容器都需要 ESXi 主机能够访问的协议端点 (PE)。

使用块存储时，PE 表示一个代理 LUN，该代理 LUN 通过基于 T10 的 LUN WWN 定义。对于 NFS 存储，PE 是一个挂载点，如 IP 地址或 DNS 名称以及共享名称。对于 NVMe 存储，PE 是一个 ANA 组，代理命名空间是虚拟的。

通常，PE 配置特定于阵列。配置 PE 时，可能需要将其与特定存储处理器或特定主机关联。为避免创建 PE 时出错，请勿进行手动配置。在可行的情况下，请使用存储特定的管理工具。

### 不要向协议端点 LUN 分配大于 Disk.MaxLUN 的 ID

默认情况下，ESXi 主机可以访问 0 到 1023 范围内的 LUN ID。如果配置的协议端点 LUN ID 大于或等于 1024，则主机可能会忽略此 PE。

如果您的环境使用大于 1023 的 LUN ID，请通过 `Disk.MaxLUN` 参数更改已扫描 LUN 的数量。请参见[更改扫描的存储设备的数量](#)。

## Virtual Volumes 性能最佳做法

为了确保实现 Virtual Volumes 最佳性能，请遵循以下建议。

### 对每个 Virtual Volumes 组件使用不同的虚拟机存储策略

默认情况下，Virtual Volumes 环境中虚拟机的所有组件使用一个虚拟机存储策略。但是，不同组件可能具有不同的性能特征，例如，数据库虚拟磁盘和对应的日志虚拟磁盘。根据性能要求，可以向每个虚拟磁盘以及虚拟机主页文件或者配置 vVol 分配不同的虚拟机存储策略。

使用 vSphere Client 时，不能更改交换 vVol、内存 vVol 或快照 vVol 的虚拟机存储策略分配。

请参见为 [Virtual Volumes 创建虚拟机存储策略](#)。

### 获取 Virtual Volumes 的主机配置文件

获取 Virtual Volumes 的主机配置文件的最佳方法是配置一个引用主机，然后提取其配置文件。如果在 vSphere Client 中手动编辑现有主机配置文件并将编辑后的配置文件附加到新主机，可能会触发合规性错误。还可能发生其他不可预知的问题。有关更多详细信息，请参见 [VMware 知识库文章 2146394](#)。

### 监控单个协议端点的 I/O 负载

- 所有虚拟卷 I/O 都会通过协议端点 (PE)。阵列从 ESXi 主机可以访问的多个 PE 中选择协议端点。阵列可以实现负载均衡并更改连接虚拟卷和 PE 的绑定路径。请参见[绑定和解除绑定 Virtual Volumes](#)。
- 在块存储上，由于可能存在大量虚拟卷，ESXi 会向 I/O 提供较大的队列深度。  
Scsi.ScsiVVolPESNRO 参数控制可以针对 PE 进行排队的 I/O 数。您可以在 vSphere Client 的“高级系统设置”页面配置该参数。

### 监控阵列限制

一个虚拟机可能会占用多个虚拟卷。请参见[虚拟卷对象](#)。

假设您的虚拟机拥有两个虚拟磁盘，并且您生成了两个内存快照。您的虚拟机最多可能占用 10 个 Virtual Volumes 对象：1 个配置 vVol、1 个交换 vVol、2 个数据 vVol、4 个快照 vVol 和 2 个内存快照 vVol。

### 确保存储提供程序可用

要访问 Virtual Volumes 存储，ESXi 主机需要一个存储提供程序（VASA 提供程序）。它内置在存储阵列固件中，或者以虚拟机的形式部署为设备。如果您的 VASA 提供程序是基于虚拟机的设备，则存储提供程序始终可用。要确保存储提供程序始终可用，请遵循以下准则：

- 请勿将存储提供程序虚拟机迁移到 Virtual Volumes 存储。
- 备份存储提供程序虚拟机。
- 必要时，使用 vSphere HA 或 Site Recovery Manager 保护存储提供程序虚拟机。

## 对 Virtual Volumes 进行故障排除

故障排除主题为您在使用 Virtual Volumes 时可能遇到的问题提供解决方案。

- **Virtual Volumes 和 esxcli 命令**

可以使用 `esxcli storage vvol` 命令对 Virtual Volumes 环境进行故障排除。

- **收集 Virtual Volumes 的统计信息**

可以在 ESXi 主机中使用 `vvol stats` 命令来跟踪性能统计信息。

- **Virtual Volumes 数据存储无法访问**

创建 Virtual Volumes 数据存储后，一直无法访问。

- **将虚拟机迁移到 Virtual Volumes 数据存储或将 VM OVF 部署到该数据存储时失败**

尝试将虚拟机迁移到 Virtual Volumes 数据存储或将 VM OVF 部署到该数据存储失败。

## Virtual Volumes 和 esxcli 命令

可以使用 `esxcli storage vvol` 命令对 Virtual Volumes 环境进行故障排除。

可以使用以下命令选项。

**表 14-1. esxcli storage vvol 命令**

命名空间	命令选项	描述
<code>esxcli storage core device</code>	<code>list</code>	确定协议端点。输出条目 <code>Is VVOL PE: true</code> 指示存储设备是协议端点。
<code>esxcli storage vvol daemon</code>	<code>unbindall</code>	解除绑定所有 VASA 提供程序中 ESXi 主机已知的所有虚拟卷。
<code>esxcli storage vvol protocolendpoint</code>	<code>list</code>	列出主机可以访问的所有协议端点。
<code>esxcli storage vvol storagecontainer</code>	<code>list</code> <code>abandonedvvol scan</code>	列出所有可用的存储容器。 扫描用于已放弃 Virtual Volumes 的指定存储容器。
<code>esxcli storage vvol vasacontext</code>	<code>get</code>	显示与主机关联的 VASA 环境 (VC UUID)。
<code>esxcli storage vvol vasaprovider</code>	<code>list</code>	列出与主机关联的所有存储 (VASA) 提供商。

## 收集 Virtual Volumes 的统计信息

可以在 ESXi 主机中使用 `vvol stats` 命令来跟踪性能统计信息。

可以使用以下命令选项。



命令	描述	选项
<code>esxcli storage vvol stats get</code>	获取所有 VASA 提供程序（默认），或指定命名空间或给定命名空间中的实体的统计信息。	<code>-e --entity=str</code> 输入实体 ID。 <code>-n --namespace=str</code> 输入节点命名空间表达式。 <code>-r --raw</code> 启用原始格式输出。
<code>esxcli storage vvol stats list</code>	列出所有统计信息节点（默认），或指定命名空间下的节点。	<code>-n --namespace=str</code> 输入节点命名空间表达式。
<code>esxcli storage vvol stats enable</code>	为整个命名空间启用统计信息跟踪。	
<code>esxcli storage vvol stats disable</code>	为整个命名空间禁用统计信息跟踪。	
<code>esxcli storage vvol stats add</code>	为特定命名空间下的特定实体启用统计信息跟踪。	<code>-e --entity=str</code> 输入实体 ID。 <code>-n --namespace=str</code> 输入节点命名空间表达式。
<code>esxcli storage vvol stats remove</code>	移除指定命名空间下用于统计信息跟踪的特定实体。	<code>-e --entity=str</code> 输入实体 ID。 <code>-n --namespace=str</code> 输入节点命名空间表达式。
<code>esxcli storage vvol stats reset</code>	重置指定统计信息命名空间或实体的统计信息计数器。	<code>-e --entity=str</code> 输入实体 ID。 <code>-n --namespace=str</code> 输入节点命名空间表达式。

## Virtual Volumes 数据存储无法访问

创建 Virtual Volumes 数据存储后，一直无法访问。

### 问题

vSphere Client 显示数据存储无法访问。无法将数据存储用于虚拟机置备。

### 原因

无法为映射到虚拟数据存储的基于 SCSI 的存储容器配置协议端点时，或者无法为映射到虚拟数据存储的基于 NVMe 的 vVol 配置 ANA 组时，可能会出现此问题。与传统 LUN 类似，需要对 SCSI 协议端点进行配置，这样 ESXi 主机才能检测到这些端点。

### 解决方案

为基于 SCSI 的容器创建虚拟数据存储之前，请确保在存储端配置协议端点。

## 将虚拟机迁移到 Virtual Volumes 数据存储或将 VM OVF 部署到该数据存储时失败

尝试将虚拟机迁移到 Virtual Volumes 数据存储或将 VM OVF 部署到该数据存储失败。

### 问题

正在从非虚拟数据存储迁移的 OVF 模板或虚拟机可能包括大型附加文件，如 ISO 磁盘映像、DVD 映像和映像文件。如果这些附加文件导致配置虚拟卷超出其 4-GB 限制，则迁移或部署到虚拟数据存储的操作将会失败。

### 原因

配置虚拟卷（简称为配置 vVol）包含各种虚拟机相关的文件。在传统的非虚拟数据存储上，这些文件存储在虚拟机主目录中。与虚拟机主目录类似，配置 vVol 通常包括虚拟机配置文件、虚拟磁盘和快照描述符文件、日志文件和锁定文件等等。

在虚拟数据存储上，所有其他大型文件（如虚拟磁盘、内存快照、交换和摘要）均存储为单独的虚拟卷。

配置 vVol 以 4-GB 虚拟卷的形式创建。配置 vVol 的一般内容通常仅占用这 4-GB 分配量的一小部分，因此配置 vVol 往往经过精简置备，可以节省备份空间。任何大型附加文件（如 ISO 磁盘映像、DVD 映像和映像文件）都有可能会导致配置 vVol 超出其 4-GB 限制。如果此类文件包含在 OVF 模板中，则将 VM OVF 部署到 Virtual Volumes 存储的操作将会失败。如果这些文件是现有虚拟机的一部分，则将该虚拟机从传统数据存储迁移到 Virtual Volumes 存储的操作也会失败。

### 解决方案

- 对于虚拟机迁移。在将虚拟机从传统数据存储迁移到虚拟数据存储之前，将多余的内容从虚拟机主目录中移除，以使配置 vVol 不超出 4-GB 限制。
- 对于 OVF 部署。由于无法将包含多余文件的 OVF 模板直接部署到虚拟数据存储，因此请先将虚拟机部署到非虚拟数据存储。将任何多余的内容从虚拟机主目录中移除，然后将生成的虚拟机迁移到 Virtual Volumes 存储。

I/O 筛选器是可以安装到 ESXi 主机上的软件组件，并可以向虚拟机提供其他数据服务。筛选器处理在虚拟机的客户机操作系统与虚拟磁盘之间移动的 I/O 请求。

I/O 筛选器可以由 VMware 提供，也可以由第三方通过 vSphere APIs for I/O Filtering (VAIO) 创建。

本章讨论了以下主题：

- 关于 I/O 筛选器
- 使用闪存设备与缓存 I/O 筛选器
- I/O 筛选器系统要求
- 在 vSphere 环境中配置 I/O 筛选器
- 在虚拟磁盘上启用 I/O 筛选器数据服务
- 管理 I/O 筛选器
- I/O 筛选器准则和最佳实践
- 处理 I/O 筛选器安装故障

## 关于 I/O 筛选器

通过 I/O 筛选器，可以直接访问虚拟机 I/O 路径。您可以在单个虚拟磁盘级别启用 I/O 筛选器。I/O 筛选器独立于存储拓扑。

VMware 提供某些类别的 I/O 筛选器。此外，第三方供应商可以创建 I/O 筛选器。通常，它们以软件包形式分发，可提供安装程序以在 vCenter Server 和 ESXi 主机集群上部署筛选器组件。

部署 I/O 筛选器之后，vCenter Server 会为集群中的每个主机配置并注册 I/O 筛选器存储提供程序（也称为“VASA 提供程序”）。这些存储提供程序与 vCenter Server 进行通信，并使 I/O 筛选器提供的数据服务在“虚拟机存储策略”界面中可见。在定义虚拟机策略的常用规则时，可以引用这些数据服务。将虚拟磁盘与此策略关联后，会在虚拟磁盘上启用 I/O 筛选器。

## 数据存储支持

I/O 筛选器可以支持所有数据存储类型，包括：

- VMFS
- NFS 3

- NFS 4.1
- VVol
- vSAN

## I/O 筛选器的类型

VMware 提供了某些类别的 I/O 筛选器，它们安装在您的 ESXi 主机上。此外，VMware 合作伙伴可以通过 vSphere APIs for I/O Filtering (VAIO) 开发人员计划创建 I/O 筛选器。I/O 筛选器有多个用途。

支持的筛选器类型包括：

- 复制。将所有写入 I/O 操作复制到外部目标位置，如其他主机或集群。
- 加密。由 VMware 提供。为虚拟机提供加密机制。有关详细信息，请参见《《vSphere 安全性》》文档。
- 缓存。为虚拟磁盘数据实现缓存。该筛选器可以使用本地闪存设备缓存数据并提高虚拟磁盘的 IOPS 和硬件利用率。如果使用缓存筛选器，则您可能需要配置虚拟闪存资源。
- 存储 I/O 控制。由 VMware 提供。限制数据存储的 I/O 负载并控制在 I/O 拥堵期间分配给虚拟机的存储 I/O 量。有关详细信息，请参见《《vSphere 资源管理》》文档。

---

**注** 您可以在 ESXi 主机上安装多个相同类别的筛选器，如缓存。但是，只能为每个虚拟磁盘安装一个相同类别的筛选器。

---

## I/O 筛选组件

I/O 筛选进程涉及多个组件。

I/O 筛选的基本组件包括：

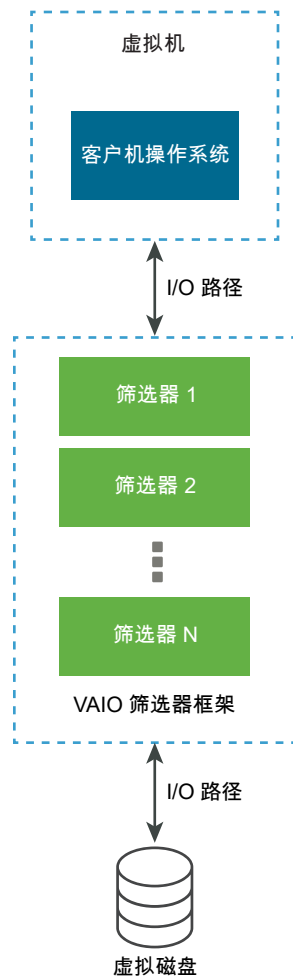
### VAIO 筛选器框架

用户环境和 ESXi 提供的 VMkernel 基础架构的组合。借助框架，您可以将筛选器插件添加到往返虚拟磁盘的 I/O 路径。基础架构包含 I/O 筛选器存储提供程序（VASA 提供程序）。该提供程序与基于存储策略的管理 (SPBM) 系统集成，并将其筛选器功能导出到 vCenter Server。

### I/O 筛选器插件

由 VMware 提供或由 VMware 合作伙伴开发的软件组件，用于拦截并筛选在虚拟磁盘与客户机操作系统之间传输的 I/O 数据。如果 I/O 筛选器由 VMware 合作伙伴开发，该筛选器可能包含附加的可选组件，这些组件有助于配置和管理该筛选器。

下图说明了 I/O 筛选组件以及客户机操作系统与虚拟磁盘之间的 I/O 流。



虚拟机的所有虚拟机可执行 (VMX) 组件均包含一个筛选器框架，可管理附加到虚拟磁盘的 I/O 筛选器插件。当 I/O 请求在客户机操作系统与虚拟磁盘之间移动时，该筛选器框架会调用筛选器。另外，筛选器还会拦截在运行的虚拟机之外发生的任何虚拟磁盘 I/O 访问。

筛选器按特定顺序依次运行。例如，复制筛选器在缓存筛选器之前执行。多个筛选器可以对虚拟磁盘进行操作，但是每种类别仅限一个筛选器。

特定磁盘的所有筛选器都验证 I/O 请求后，请求将到达其目标位置，即虚拟机或虚拟磁盘。

由于筛选器在用户空间中运行，因此任何筛选器故障都只会影响虚拟机，而不会对 ESXi 主机产生影响。

## I/O 筛选器的存储提供程序

在 ESXi 主机上安装 I/O 筛选器后，I/O 筛选器框架会为集群中的每个主机配置并注册存储提供程序（也称为 VASA 提供程序）。

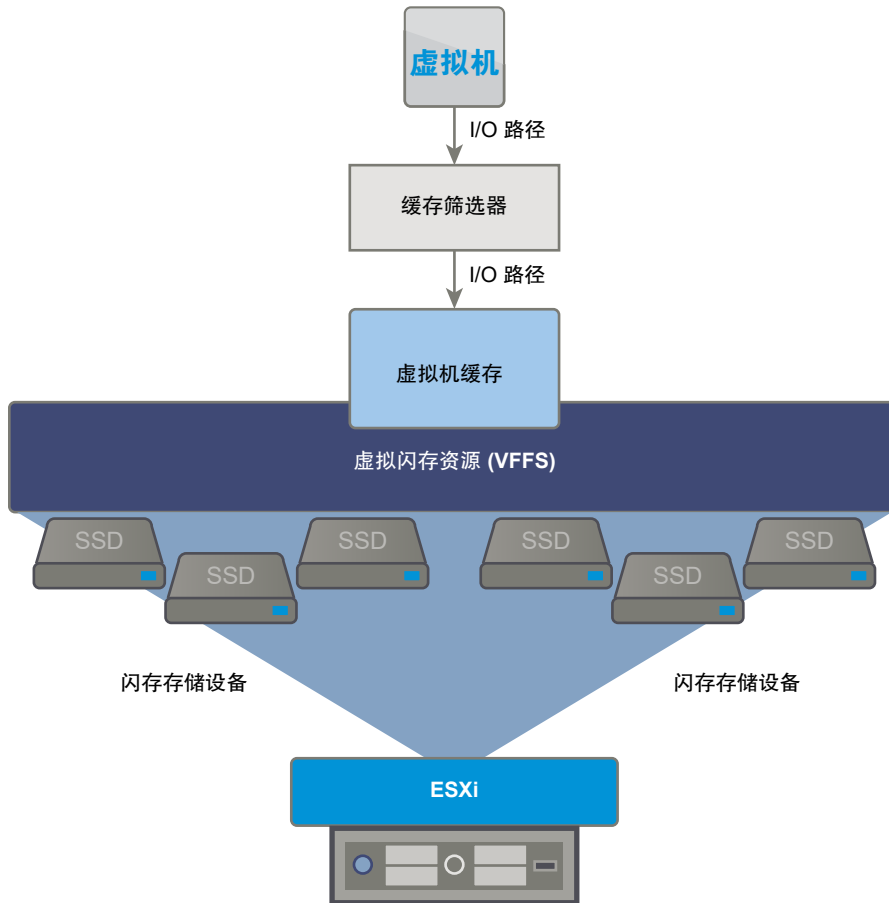
用于 I/O 筛选的存储提供程序是由 vSphere 提供的软件组件。它们与 I/O 筛选器集成并向 vCenter Server 报告 I/O 筛选器支持的数据服务功能。

此功能填充“虚拟机存储策略”界面并且可以在虚拟机存储策略中引用。然后，您可以将此策略应用到虚拟磁盘，以便 I/O 筛选器可以为磁盘处理 I/O。

## 使用闪存设备与缓存 I/O 筛选器

缓存 I/O 筛选器可以使用本地闪存设备缓存虚拟机数据。

如果缓存 I/O 筛选器使用本地闪存设备，您需要配置虚拟闪存资源（也称为 VFFS 卷）。请在激活筛选器之前在 ESXi 主机上配置资源。处理虚拟机读取 I/O 时，该筛选器会创建虚拟机缓存并将其放置在 VFFS 卷上。



要设置虚拟闪存资源，请使用连接到主机的闪存设备。要增加虚拟闪存资源的容量，您可以添加更多闪存驱动器。单个闪存驱动器必须以独占方式分配给虚拟闪存资源，不能与任何其他 vSphere 服务（例如 vSAN 或 VMFS 共享。请参见[设置虚拟闪存资源](#)。

## I/O 筛选器系统要求

要在环境中使用 I/O 筛选器，必须遵循特定的要求。

下列要求适用。

- 使用与 I/O 筛选器兼容的最新 ESXi 和 vCenter Server 版本。较早的版本可能不支持 I/O 筛选器或只提供部分支持。
- 检查各个合作伙伴解决方案可能具有的任何其他要求。在特定情况下，环境可能需要闪存设备、额外的物理内存或网络连接和带宽。有关信息，请与供应商或 VMware 代表联系。

- 托管合作伙伴软件包以用于筛选器安装的 Web 服务器。服务器必须在初始安装后保持可用。在新的主机加入集群时，服务器会向该主机推送相应的 I/O 筛选器组件。

## 在 vSphere 环境中配置 I/O 筛选器

要设置 I/O 筛选器为您的虚拟机提供的数据服务，请遵循几个步骤。

### 前提条件

- 创建至少包含一个 ESXi 主机的集群。
- 有关第三方提供的 I/O 筛选器的信息，请与供应商或 VMware 代表联系。

### 步骤

#### 1 在集群中安装 I/O 筛选器

如果使用第三方提供的 I/O 筛选器，则将该 I/O 筛选器安装在 ESXi 主机集群上。

#### 2 查看 I/O 筛选器和存储提供程序

可以使用 vSphere Client 查看环境中可用的 I/O 筛选器并验证 I/O 筛选器提供程序是否按预期显示并且处于活动状态。

## 在集群中安装 I/O 筛选器

如果使用第三方提供的 I/O 筛选器，则将该 I/O 筛选器安装在 ESXi 主机集群上。

VMware 合作伙伴通过 vSphere APIs for I/O Filtering (VAIO) 开发人员计划创建 I/O 筛选器。

筛选器软件包以解决方案包 ZIP 软件包的形式进行分发，可以包括 I/O 筛选器守护进程、I/O 筛选器库、CIM 提供程序和其他关联的组件。

通常，您可以运行供应商提供的安装程序部署筛选器。在 ESXi 集群级别执行安装。无法直接在所选主机上安装筛选器。

### 前提条件

- 所需特权：**主机.配置.查询修补程序**。
- 确认 I/O 筛选器解决方案已经过 VMware 认证。

### 步骤

- ◆ 运行供应商提供的安装程序。

安装程序在 vCenter Server 上部署相应的 I/O 筛选器扩展，并在集群内的所有主机上部署筛选器组件。

系统会为集群中的每个 ESXi 主机自动注册存储提供程序（也称为 VASA 提供程序）。成功自动注册 I/O 筛选器存储提供程序后，会触发主机级别的事件。如果存储提供程序自动注册失败，系统会在主机上引发警报。

## 查看 I/O 筛选器和存储提供程序

可以使用 vSphere Client 查看环境中可用的 I/O 筛选器并验证 I/O 筛选器提供程序是否按预期显示并且处于活动状态。

安装第三方 I/O 筛选器后，系统会为集群中的每个 ESXi 主机自动注册一个存储提供程序（也称为 VASA 提供程序）。成功自动注册 I/O 筛选器存储提供程序后，会触发主机级别的事件。如果存储提供程序自动注册失败，系统会在主机上引发警报。

### 步骤

- 1 验证 I/O 筛选器存储提供程序是否按预期显示且处于活动状态。
  - a 导航到 vCenter Server。
  - b 依次单击配置选项卡和存储提供程序。
  - c 检查 I/O 筛选器的存储提供程序。

I/O 筛选器提供程序经过正确注册后，筛选器提供的功能和数据服务将填充到“虚拟机存储策略”界面中。

- 2 验证 I/O 筛选器组件是否在集群和 ESXi 主机上列出。

选项	操作
查看集群上的 I/O 筛选器	<ul style="list-style-type: none"><li>a 导航到集群。</li><li>b 单击配置选项卡。</li><li>c 在配置下面，单击 I/O 筛选器，以查看集群中安装的筛选器。</li></ul>
查看主机上的 I/O 筛选器	<ul style="list-style-type: none"><li>a 导航到主机。</li><li>b 单击配置选项卡。</li><li>c 在存储下面，单击 I/O 筛选器，以查看主机上安装的筛选器。</li></ul>

## 在虚拟磁盘上启用 I/O 筛选器数据服务

启用 I/O 筛选器提供的数据服务过程分为两步。先根据 I/O 筛选器提供的数据服务创建虚拟机策略，然后将此策略附加到虚拟机。

### 前提条件

对于缓存 I/O 筛选器，请在激活筛选器之前先在 ESXi 主机上配置虚拟闪存资源。请参见[设置虚拟闪存资源](#)。

### 步骤

- 1 根据 I/O 筛选器服务定义虚拟机策略。
  - 确保虚拟机策略列出 I/O 筛选器提供的数据服务。
  - 请参见[为基于主机的数据服务创建虚拟机存储策略](#)。



## 2 将 I/O 筛选器策略分配给虚拟机。

要激活 I/O 筛选器提供的数据服务，请将 I/O 筛选器策略与虚拟磁盘关联。可以在置备虚拟机时分配该策略。

请参见[将 I/O 筛选器策略分配给虚拟机](#)。

### 后续步骤

如果稍后要禁用虚拟机的 I/O 筛选器，您可以从虚拟机存储策略中移除筛选器规则，然后重新应用该策略。请参见[编辑或克隆虚拟机存储策略](#)。或者，可以编辑虚拟机的设置，选择不包含该筛选器的其他存储策略。

## 将 I/O 筛选器策略分配给虚拟机

要激活 I/O 筛选器提供的数据服务，请将 I/O 筛选器策略与虚拟磁盘关联。您可以在创建或编辑虚拟机时分配该策略。

您可以在虚拟机初始部署期间分配 I/O 筛选器策略。本主题介绍在创建新的虚拟机时如何分配该策略。有关其他部署方法的信息，请参见《《vSphere 虚拟机管理》》文档。

---

**注** 迁移或克隆虚拟机时，不能更改或分配 I/O 筛选器策略。

---

### 前提条件

验证虚拟机在其上运行的 ESXi 主机上是否已安装 I/O 筛选器。

### 步骤

- 1 启动虚拟机置备流程并按照相应的步骤操作。
- 2 将同一存储策略分配给所有虚拟机文件和磁盘。
  - a 在[选择存储](#)页面上，从[虚拟机存储策略](#)下拉菜单中选择存储策略。
  - b 从兼容数据存储列表中选择数据存储，然后单击[下一步](#)。

该数据存储将成为虚拟机配置文件及所有虚拟磁盘的目标存储资源。该策略还会为虚拟磁盘激活 I/O 筛选器服务。
- 3 更改虚拟磁盘的虚拟机存储策略。

使用此选项可仅为您的虚拟磁盘启用 I/O 筛选器。

  - a 在[自定义硬件](#)页面上，展开[新硬盘](#)窗格。
  - b 从[虚拟机存储策略](#)下拉菜单中，选择要分配给虚拟磁盘的存储策略。
  - c （可选）更改虚拟磁盘的存储位置。

使用此选项可将虚拟磁盘存储在除虚拟机配置文件所在的数据存储以外的数据存储上。
- 4 完成虚拟机置备流程。

### 结果

创建虚拟机后，[摘要](#)选项卡将显示已分配的存储策略及其合规性状态。

## 后续步骤

稍后可以更改虚拟策略分配。请参见[更改虚拟机文件和磁盘的存储策略分配](#)。

## 管理 I/O 筛选器

可以运行供应商提供的安装程序来安装、卸载或升级 I/O 筛选器。

使用 I/O 筛选器时，需要考虑以下注意事项：

- vCenter Server 使用 ESX Agent Manager (EAM) 安装和卸载 I/O 筛选器。作为管理员，切勿为由 vCenter Server 创建或使用的 EAM 代理直接调用 EAM API。所有与 I/O 筛选器相关的操作均必须通过 VIM API。如果意外修改了由 vCenter Server 创建的 EAM 代理，则必须恢复所做的更改。如果意外破坏了由 I/O 筛选器使用的 EAM 代理，则必须调用 `Vim.IoFilterManager#uninstallIoFilter` 卸载受影响的 I/O 筛选器。卸载后，执行全新安装。
- 当新主机加入具有 I/O 筛选器的集群时，集群上安装的筛选器将会部署在该主机上。vCenter Server 为该主机注册 I/O 筛选器存储提供程序。任何集群更改都会显示在 vSphere Client 的“虚拟机存储策略”界面中。
- 将主机从集群中移出或将其从 vCenter Server 中移除时，I/O 筛选器会从主机上卸载。vCenter Server 取消注册 I/O 筛选器存储提供程序。
- 如果使用无状态 ESXi 主机，可能会在重新引导过程中丢失其 I/O 筛选器 VIB。vCenter Server 将在重新引导后检查安装在主机上的包，并根据需要将 I/O 筛选器 VIB 推送到主机。

## 从集群卸载 I/O 筛选器

您可以卸载 ESXi 主机集群中部署的 I/O 筛选器。

### 前提条件

- 所需特权：**Host.Config.Patch**。

### 步骤

- 1 通过运行供应商提供的安装程序，卸载 I/O 筛选器。

在卸载过程中，第三方 I/O 筛选器安装程序会自动将主机置于维护模式。

如果卸载成功，则筛选器以及任何相关的组件将从主机中移除。

- 2 验证 I/O 筛选器组件是否已从 ESXi 主机正确卸载。使用下列方法之一：

- 运行 `esxcli software vib list` 命令。
- 在 vSphere Client 中查看 I/O 筛选器。请参见[查看 I/O 筛选器和存储提供程序](#)。

卸载后的筛选器不再出现在该列表中。

## 升级集群中的 I/O 筛选器

升级 ESXi 主机后，请使用 I/O 筛选器供应商提供的安装程序升级 ESXi 主机集群中部署的 I/O 筛选器。

升级包括卸载旧筛选器组件并将其替换为新筛选器组件。为了确定安装是否属于升级，vCenter Server 将检查现有筛选器的名称和版本。如果现有筛选器的名称与新筛选器的名称匹配、但是版本有所不同，则安装被视为升级。

### 前提条件

- 所需特权：**Host.Config.Patch**。
- 将主机升级到 ESXi 8.0。如果使用 vSphere Lifecycle Manager 进行升级，请参见《管理主机和集群生命周期》文档。

### 步骤

- 1 要升级筛选器，请运行供应商提供的安装程序。  
在升级过程中，第三方 I/O 筛选器安装程序会自动将主机置于维护模式。  
安装程序将标识任何现有筛选器组件并在安装新筛选器组件之前将其移除。
- 2 验证 ESXi 主机上是否已正确升级 I/O 筛选器组件。使用下列方法之一：
  - 运行 `esxcli software vib list` 命令。
  - 在 vSphere Client 中查看 I/O 筛选器。请参见[查看 I/O 筛选器和存储提供程序](#)。

### 结果

升级后，系统会将主机返回操作模式。

## I/O 筛选器准则和最佳实践

在您的环境中使用 I/O 筛选器时，请遵循特定的准则和最佳做法。

- 由于 I/O 筛选器未指定数据存储，因此所有类型的数据存储（包括 VMFS、NFS、Virtual Volumes 和 vSAN）都与 I/O 筛选器兼容。
- I/O 筛选器在虚拟兼容模式下支持 RDM。在物理兼容模式下不支持 RDM。
- 迁移或克隆虚拟机时，不能更改或分配 I/O 筛选器策略。完成迁移或克隆后，可以更改该策略。
- 将具有 I/O 筛选器策略的虚拟机从一台主机克隆或迁移到另一台主机上时，请确保目标主机上已安装兼容筛选器。此要求适用于管理员或者 HA 或 DRS 等功能启动的迁移。
- 将模板转换为虚拟机且使用 I/O 筛选器策略配置模板时，目标主机必须已安装兼容的 I/O 筛选器。
- 如果使用 vCenter Site Recovery Manager 复制虚拟磁盘，则恢复站点上生成的磁盘不具有 I/O 筛选器策略。您必须在恢复站点中创建 I/O 筛选器策略，并将其重新附加到复制磁盘。
- 创建虚拟机时，可以将加密 I/O 筛选器附加到新的虚拟磁盘上。无法将加密筛选器附加到现有的虚拟磁盘上。
- 如果虚拟机具有与之关联的快照树，则您无法添加、更改或移除虚拟机的 I/O 筛选器策略。

## 通过 I/O 筛选器迁移虚拟机

迁移具有 I/O 筛选器的虚拟机时，需要注意特定事项。

如果使用 Storage vMotion 迁移具有 I/O 筛选器的虚拟机，则必须将目标数据存储连接到已安装兼容 I/O 筛选器的主机。

您可能需要跨不同类型的数据存储迁移具有 I/O 筛选器的虚拟机，例如，在 VMFS 与 Virtual Volumes 之间。如果的确如此，请确保虚拟机存储策略包含您计划使用的每种类型的数据存储的规则集。例如，如果在 VMFS 与 Virtual Volumes 数据存储之间迁移虚拟机，请创建包含以下规则的混合虚拟机存储策略：

- 适用于 I/O 筛选器的常用规则
- 适用于 VMFS 数据存储的规则集 1。由于基于存储策略的管理不提供明确的 VMFS 策略，因此该规则集必须包括适用于 VMFS 数据存储的基于标记的规则。
- 适用于 Virtual Volumes 数据存储的规则集 2

当 Storage vMotion 迁移虚拟机时，将会选择对应于目标数据存储的正确规则集。I/O 筛选器规则持不变。

如果您没有为数据存储指定规则，仅为 I/O 筛选器定义了常用规则，则系统会为数据存储应用默认存储策略。

## 处理 I/O 筛选器安装故障

通常，集群中的所有 ESXi 主机都安装一组相同的 I/O 筛选器。有时，安装期间可能会发生故障。

如果在主机上安装 I/O 筛选器失败，则系统会生成事件来报告故障。此外，主机上的警报还会显示故障原因。故障示例包括：

- 无法从主机访问 VIB URL。
- VIB 的格式无效。
- VIB 要求主机进入维护模式以便进行升级或卸载。
- VIB 要求主机在安装或卸载之后重新引导。
- 尝试将主机置于维护模式失败，因为无法从主机上撤出虚拟机。
- VIB 需要手动安装或卸载。

vCenter Server 可以解决一些故障。您可能需要干预其他故障。例如，您可能需要编辑 VIB URL，手动撤出虚拟机或关闭虚拟机电源，或者手动安装或卸载 VIB。

## 在单个 ESXi 主机中安装 I/O 筛选器

要进行故障排除，可以下载 I/O 筛选器的 ESXi 组件（打包为 VIB 文件），并将其安装在 ESXi 主机上。使用 `esxcli` 命令安装 VIB 文件。

### 前提条件

安装 ESXCLI。请参见《ESXCLI 入门》。要进行故障排除，请在 ESXi Shell 中运行 `esxcli` 命令。

## 步骤

- 1 运行以下命令安装 VIB:

```
esxcli software vib install --depot path_to_VMware_vib_ZIP_file
```

通过 `install` 命令选项，您可以执行试运行、指定特定的 VIB 及跳过接受程度验证等。请勿跳过对生产系统的验证。请参见《ESXCLI 参考指南》文档。

- 2 验证 VIB 是否已安装在 ESXi 主机上。

```
esxcli software vib list
```

使用硬件加速功能可以将 ESXi 主机与合规存储系统集成。主机可以将某些虚拟机和存储管理操作卸载到存储系统。利用存储硬件的辅助，主机可以更快地执行这些操作并且占用更少的 CPU、内存和存储结构带宽。

块存储设备、光纤通道和 iSCSI 以及 NAS 设备均支持硬件加速。

有关其他详细信息，请参见 VMware 知识库文章，网址为 <http://kb.vmware.com/kb/1021976>。

本章讨论了以下主题：

- 硬件加速的优点
- 硬件加速要求
- 硬件加速支持状态
- 块存储设备的硬件加速
- NAS 设备上的硬件加速
- 硬件加速注意事项

## 硬件加速的优点

当支持硬件加速功能时，主机可以获得硬件支持，从而可以更快和更有效地执行多个任务。

主机可以通过以下活动获取帮助：

- 通过 Storage vMotion 迁移虚拟机
- 从模板部署虚拟机
- 克隆虚拟机或模板
- 虚拟机文件的 VMFS 集群锁定和元数据操作
- 置备厚虚拟磁盘
- 创建容错虚拟机
- 在 NFS 数据存储上创建和克隆厚磁盘

## 硬件加速要求

仅当结合使用相应的主机和存储阵列时，硬件加速功能才起作用。

表 16-1. 硬件加速存储要求

ESXi	块存储设备	NAS 设备
ESXi	支持用于阵列集成的 T10 SCSI 标准或块存储插件 (VAAI)	支持用于阵列集成的 NAS 插件

**注** 如果 SAN 或 NAS 存储结构在支持硬件加速的存储系统前使用了中间设备，则该中间设备必须也支持硬件加速并且已得到相应认证。该中间设备可能是存储虚拟化设备、I/O 加速设备、加密设备等。

## 硬件加速支持状态

对于每个存储设备和数据存储，vSphere Client 都将显示硬件加速支持状态。

状态值为“未知”、“支持”和“不支持”。初始值为“未知”。

对于块设备，在主机成功执行卸载操作后，状态将更改为“支持”。如果卸载操作失败，状态将更改为“不支持”。如果设备提供部分硬件加速支持，则状态仍为“未知”。

使用 NAS 后，当存储可以执行至少一个硬件卸载操作时，状态将变为“支持”。

当存储设备不支持主机操作或为主机操作提供部分支持时，主机将恢复使用其本机方法以执行不支持的操作。

## 块存储设备的硬件加速

利用硬件加速，您的主机可以与块存储设备、光纤通道或 iSCSI 相集成，并且可以使用某些存储阵列操作。

ESXi 硬件加速支持以下阵列操作：

- 完全复制，也称为克隆块或复制卸载。使存储阵列可以对阵列内的数据进行完全复制，而无需使主机读取和写入数据。此操作可减少在克隆虚拟机、从模板进行置备或使用 vMotion 进行迁移时所耗费的时间和网络负载。
- 块置零，也称为同写。使存储阵列可以将大量块置零以提供新分配的存储（不包含先前写入的数据）。此操作可减少在创建虚拟机和格式化虚拟磁盘时所耗费的时间和网络负载。
- 硬件辅助锁定，也称为原子测试和设置 (ATS)。支持离散虚拟机锁定，而不使用 SCSI 预留。此操作允许按扇区进行磁盘锁定，取代了使用 SCSI 预留对整个 LUN 进行磁盘锁定。

请向您的供应商确认是否提供硬件加速支持。某些存储阵列要求在存储器端激活该支持。

在主机上，硬件加速功能在默认情况下已启用。如果存储器不支持硬件加速，可以将其禁用。

除了硬件加速支持外，ESXi 还支持阵列精简置备。有关信息，请参见 [ESXi 和阵列精简置备](#)。

## 禁用块存储设备的硬件加速

在主机上，块存储设备的硬件加速功能在默认情况下已启用。可以使用 vSphere Client 高级设置来禁用硬件加速操作。

与任何高级设置一样，在禁用硬件加速之前，请咨询 VMware 支持团队。

### 步骤

- 1 在 vSphere Client 中，导航到 ESXi 主机。
- 2 单击**配置**选项卡。
- 3 在**系统**下，单击**高级系统设置**。
- 4 将任意选项的值更改为 0（已禁用）：
  - VMFS3.HardwareAcceleratedLocking
  - DataMover.HardwareAcceleratedMove
  - DataMover.HardwareAcceleratedInit

## 管理块存储设备上的硬件加速

为了与块存储阵列集成，vSphere 使用称为 Storage APIs - Array Integration (VAAI) 的 ESXi 扩展。通过此集成，vSphere 可以使用阵列硬件操作。

在 vSphere 5.x 及更高版本中，这些扩展是以 T10 SCSI 命令形式实现的。因此，通过支持 T10 SCSI 标准的设备，ESXi 主机可以直接通信而不需要 VAAI 插件。

如果设备不支持 T10 SCSI 或提供部分支持，则 ESXi 将恢复使用安装在主机上的 VAAI 插件。主机也可以组合使用 T10 SCSI 命令和插件。VAAI 插件是特定于供应商的，可能是 VMware 或合作伙伴开发的。要管理支持 VAAI 的设备，主机可将 VAAI 筛选器和特定于供应商的 VAAI 插件附加到该设备。

有关存储是否需要 VAAI 插件或通过 T10 SCSI 命令支持硬件加速的信息，请参见《《VMware 兼容性指南》》或联系存储供应商。

可以使用多个 `esxcli` 命令查询存储设备，以获取硬件加速支持信息。对于需要 VAAI 插件的设备，也可以使用声明规则命令。有关 `esxcli` 命令的信息，请参见《ESXCLI 入门》。

### 显示硬件加速插件和筛选器

要与不支持 T10 SCSI 标准的设备进行通信，主机需要使用单个 VAAI 筛选器和特定于供应商的 VAAI 插件。使用 `esxcli` 命令可查看当前加载到系统中的硬件加速筛选器和插件。

### 前提条件

安装 ESXCLI。请参见《ESXCLI 入门》。要进行故障排除，请在 ESXi Shell 中运行 `esxcli` 命令。



**步骤**

- ◆ 运行 **esxcli storage core plugin list --plugin-class=value** 命令。

对于 *value*，输入以下参数之一：

- 键入 VAAI 可显示插件。

此命令的输出与以下示例类似：

```
#esxcli storage core plugin list --plugin-class=VAAI
Plugin name      Plugin class
VMW_VAAIP_EQL    VAAI
VMW_VAAIP_NETAPP VAAI
VMW_VAAIP_CX     VAAI
```

- 键入 Filter 可显示筛选器。

此命令的输出与以下示例类似：

```
esxcli storage core plugin list --plugin-class=Filter
Plugin name  Plugin class
VAAI_FILTER Filter
```

**验证硬件加速支持状态**

使用 **esxcli** 命令可验证特定存储设备的硬件加速支持状态。

**前提条件**

安装 ESXCLI。请参见《ESXCLI 入门》。要进行故障排除，请在 ESXi Shell 中运行 **esxcli** 命令。

**步骤**

- ◆ 运行 **esxcli storage core device list -d=device\_ID** 命令。

输出显示硬件加速（或 VAAI）状态，该状态可以为未知、受支持或不受支持。

```
# esxcli storage core device list -d naa.XXXXXXXXXXXXX4c
naa.XXXXXXXXXXXXX4c
Display Name: XXXX Fibre Channel Disk(naa.XXXXXXXXXXXXX4c)
Size: 20480
Device Type: Direct-Access
Multipath Plugin: NMP
XXXXXXXXXXXXXXXXXX
Attached Filters: VAAI_FILTER
VAAI Status: supported
XXXXXXXXXXXXXXXXXX
```

## 验证硬件加速支持详细信息

使用 `esxcli` 命令查询块存储设备是否提供硬件加速支持。

### 前提条件

安装 ESXCLI。请参见《ESXCLI 入门》。要进行故障排除，请在 ESXi Shell 中运行 `esxcli` 命令。

### 步骤

- ◆ 运行 `esxcli storage core device vaai status get -d=device_ID` 命令。

如果由 VAAI 插件管理该设备，则输出会显示连接到该设备的插件的名称。此外，输出还会显示每个基于 T10 SCSI 的原语的支持状态（如果可用）。输出如以下示例中所示：

```
# esxcli storage core device vaai status get -d naa.XXXXXXXXXXXXX4c
naa.XXXXXXXXXXXXX4c
VAAI Plugin Name: VMW_VAAIP_SYMM
ATS Status: supported
Clone Status: supported
Zero Status: supported
Delete Status: unsupported
```

## 列出硬件加速声明规则

VAAI 插件管理的每个块存储设备都需要两个声明规则。一个声明规则指定硬件加速筛选器，另一个指定设备的硬件加速插件。可以使用 `esxcli` 命令列出硬件加速筛选器和插件声明规则。

### 步骤

- 1 要列出筛选器声明规则，请运行 `esxcli storage core claimrule list --claimrule-class=Filter` 命令。

在此示例中，筛选器声明规则指定 VAAI\_FILTER 筛选器声明的设备。

```
# esxcli storage core claimrule list --claimrule-class=Filter
Rule Class Rule Class Type Plugin Matches XCOPY Use Array
Reported Values XCOPY Use Multiple Segments XCOPY Max Transfer Size KiB
Filter 65430 runtime vendor VAAI_FILTER vendor=EMC
model=SYMMETRIX False
False 0
Filter 65430 file vendor VAAI_FILTER vendor=EMC
model=SYMMETRIX False
False 0
Filter 65431 runtime vendor VAAI_FILTER vendor=DGC
model=* False
False 0
Filter 65431 file vendor VAAI_FILTER vendor=DGC
model=* False
False 0
```

## 2 要列出 VAAI 插件声明规则，请运行

**esxcli storage core claimrule list --claimrule-class=VAAI** 命令。

在此示例中，VAAI 声明规则指定 VAAI 插件声明的设备。

```
esxcli storage core claimrule list --claimrule-class=VAAI
Rule Class Rule Class Type Plugin Matches XCOPY Use
Array Reported Values XCOPY Use Multiple Segments XCOPY Max Transfer Size KiB
VAAI 65430 runtime vendor VMW_VAAIP_SYMM vendor=EMC
model=SYMMETRIX False
False 0
VAAI 65430 file vendor VMW_VAAIP_SYMM vendor=EMC
model=SYMMETRIX False
False 0
VAAI 65431 runtime vendor VMW_VAAIP_CX vendor=DGC
model=* False
False 0
VAAI 65431 file vendor VMW_VAAIP_CX vendor=DGC
model=* False
False 0
```

## 添加硬件加速声明规则

要为新的阵列配置硬件加速，请添加两个声明规则：一个用于 VAAI 筛选器，另一个用于 VAAI 插件。为了激活新声明规则，请先定义规则，然后将其加载到系统中。

### 前提条件

安装 ESXCLI。请参见《ESXCLI 入门》。要进行故障排除，请在 ESXi Shell 中运行 **esxcli** 命令。

### 步骤

- 运行以下命令为 VAAI 筛选器定义新的声明规则：  
**esxcli storage core claimrule add --claimrule-class=Filter --plugin=VAAI\_FILTER。**
- 运行以下命令为 VAAI 插件定义新的声明规则：  
**esxcli storage core claimrule add --claimrule-class=VAAI。**
- 运行以下命令加载这两个声明规则：  
**esxcli storage core claimrule load --claimrule-class=Filter**  
**esxcli storage core claimrule load --claimrule-class=VAAI**
- 使用以下命令运行 VAAI 筛选器声明规则：  
**esxcli storage core claimrule run --claimrule-class=Filter。**

**注** 只需运行筛选器类规则。当 VAAI 筛选器声明设备时，会自动查找要附加的适当 VAAI 插件。

### 示例：定义硬件加速声明规则

此示例显示如何为使用 VMW\_VAAIP\_T10 插件的 IBM 阵列配置硬件加速。按以下顺序运行命令。有关命令使用的选项的信息，请参见[添加多路径声明规则](#)。

```
# esxcli storage core claimrule add --claimrule-class=Filter --
plugin=VAAI_FILTER --type=vendor --vendor=IBM --autoassign
# esxcli storage core claimrule add --claimrule-class=VAAI --
plugin=VMW_VAAIP_T10 --type=vendor --vendor=IBM --autoassign
# esxcli storage core claimrule load --claimrule-class=Filter
# esxcli storage core claimrule load --claimrule-class=VAAI
# esxcli storage core claimrule run --claimrule-class=Filter
```

### 配置 XCOPY 参数

XCOPY 是 VAAI 基本类型之一，用于将任务卸载到存储阵列。例如，可以使用 XCOPY 将诸如虚拟机迁移或克隆等操作卸载到阵列，从而不必使用 vSphere 资源来执行这些任务。

可以将 XCOPY 机制与所有支持 VMW\_VAAIP\_T10 插件（由 VMware 开发，基于 SCSI T10）的存储阵列结合使用。要启用 XCOPY 机制，请创建 VAAI 类的声明规则。

#### 前提条件

安装 ESXCLI。请参见《ESXCLI 入门》。要进行故障排除，请在 ESXi Shell 中运行 esxcli 命令。

#### 步骤

- ◆ 使用以下命令并输入 XCOPY 选项：

```
esxcli storage core claimrule add --claimrule-class=VAAI
```

有关命令使用的选项的信息，请参见[添加多路径声明规则](#)。

选项	描述
-a --xcopy-use-array-values	将阵列报告的值用于 XCOPY 命令。
-s --xcopy-use-multi-segs	将多个段用于 XCOPY 命令。仅当指定 --xcopy-use-array-values 时才有效。
-m --xcopy-max-transfer-size	使用不同于阵列所报告的传输大小时 XCOPY 命令的最大传输大小 (MB)。仅当指定 --xcopy-use-array-values 时才有效。
-k --xcopy-max-transfer-size-kib	使用不同于阵列所报告的传输大小时 XCOPY 命令的最大传输大小 (KiB)。仅当指定 --xcopy-use-array-values 时才有效。

### 示例：配置 XCOPY

```
■ # esxcli storage core claimrule add -r 914 -t vendor -V XtremIO -M XtremApp -P
VMW_VAAIP_T10 -c VAAI -a -s -k 64
```

```
# esxcli storage core claimrule add -r 65430 -t vendor -V EMC -M SYMMETRIX -P  
VMW_VAAIP_SYMM -c VAAI -a -s -m 200
```

## 删除硬件加速声明规则

使用 `esxcli` 命令可删除现有硬件加速声明规则。

### 前提条件

安装 ESXCLI。请参见《ESXCLI 入门》。要进行故障排除，请在 ESXi Shell 中运行 `esxcli` 命令。

### 步骤

◆ 运行以下命令：

```
esxcli storage core claimrule remove -r claimrule_ID --claimrule-  
class=Filter
```

```
esxcli storage core claimrule remove -r claimrule_ID --claimrule-  
class=VAAI
```

## NAS 设备上的硬件加速

通过硬件加速，ESXi 主机可以与 NAS 设备相集成并使用 NSA 存储提供的多个硬件操作。硬件加速使用 vSphere APIs for Array Integration (VAAI) 来方便主机与存储设备之间的通信。

VAAI NAS 框架支持这两种版本的 NFS 存储：NFS 3 和 NFS 4.1。

VAAI NAS 使用一组存储原语将存储操作从主机卸载到阵列。以下列表显示支持的 NAS 操作：

### 完整文件克隆

支持 NAS 设备克隆虚拟磁盘文件的功能。此操作与 VMFS 块克隆类似，不同之处在于 NAS 设备克隆的是整个文件而不是文件段。受益于完整文件克隆操作的任务包括虚拟机克隆、Storage vMotion 和从模板部署虚拟机。

当 ESXi 主机使用 VAAI NAS 复制数据时，不需要从 NAS 中读取数据并将数据写回 NAS。主机仅发送复制命令，将数据卸载到 NAS。复制过程在 NAS 中完成，从而减少了主机上的负载。

### 快速文件克隆

此操作（也称为基于阵列的快照或本机快照）会将虚拟机快照和链接克隆的创建操作卸载到阵列。

### 预留空间

支持存储阵列作为厚格式的虚拟磁盘文件分配空间的功能。

通常，在 NFS 数据存储上创建虚拟磁盘时，NAS 服务器会确定分配策略。大多数 NAS 服务器上的默认分配策略是精简格式，不保证将存储备份到文件。但是，预留空间操作可以指示 NAS 设备使用供应商特定的机制来为虚拟磁盘预留空间。因此，如果备用 NAS 服务器支持预留空间操作，则可以在 NFS 数据存储上创建厚虚拟磁盘。

### 扩展后的统计信息

支持对 NAS 设备上的空间使用情况的可见性。利用此操作，可以查询 NFS 数据存储上虚拟磁盘的空间利用率详细信息。详细信息包括虚拟磁盘的大小和虚拟磁盘的空间消耗。此功能对于精简置备非常有用。

使用 NAS 存储设备，可以通过供应商特定的 NAS 插件实施硬件加速集成。这些插件通常由供应商创建并以供应商软件包的形式进行分发。这些 NAS 插件无需声明规则便可以正常运行。

有多种工具可用于安装和更新 NAS 插件。这些工具包括 `esxcli` 命令和 vSphere Lifecycle Manager。有关详细信息，请参见《VMware ESXi 升级》和《管理主机和集群生命周期》文档。有关安装和更新建议，请参见[知识库文章](#)。

---

**注** NAS 存储供应商可能会提供其他设置，这些设置可能会影响 VAAI 的性能和操作。请遵循供应商的建议，并在 NAS 存储阵列和 ESXi 上配置相应的设置。有关详细信息，请参见存储供应商文档。

---

## 在虚拟机上启用 NAS 本机快照

如果您的部署包括支持 vSphere APIs for Array Integration (VAAI) 的 NAS 阵列，您可以使用快速文件克隆技术（也称为本地 NFS 快照）创建虚拟机快照。利用此技术，NFS 设备可复制虚拟机，而无需 ESXi 主机读取和写入数据。此操作可能会缩短创建虚拟机快照时的时间并减少网络负载。

默认情况下，所有新创建的虚拟机都支持传统 ESXi 快照技术。要使用 NFS 本机快照技术，请为虚拟机启用该技术。

### 前提条件

- 验证 NAS 阵列是否支持对 VAAI NAS 程序执行快速文件克隆操作。
- 在 ESXi 主机上，安装供应商特定的 NAS 插件，该插件支持使用 VAAI 进行文件快速克隆。
- 请遵循 NAS 存储供应商的建议，在 NAS 阵列和 ESXi 上配置任何所需设置。有关详细信息，请参见存储供应商文档。

### 步骤

- 1 在 vSphere Client 中，右键单击虚拟机，然后选择**编辑设置**。
- 2 单击**虚拟机选项**选项卡，然后展开**高级**菜单。
- 3 单击“配置参数”旁边的**编辑配置**。
- 4 配置 `snapshot.alwaysAllowNative` 参数。

如果该参数存在，请确保其值设置为 True。如果该参数不存在，请添加并将其值设置为 True。

名称	值
<code>snapshot.alwaysAllowNative</code>	True

## 硬件加速注意事项

在 ESXi 中使用硬件加速功能时，需注意某些事项。

多个原因可能导致硬件加速操作失败。

对于阵列未实施的任何原语，阵列均会返回错误。错误会触发 ESXi 主机尝试使用其本机方法执行操作。

发生以下任一情况时，VMFS 数据移动程序不会利用硬件下载，而是使用软件数据移动：

- 源 VMFS 数据存储和目标 VMFS 数据存储具有不同的块大小。
- 源文件类型为 RDM，目标文件类型为非 RDM（常规文件）。
- 源 VMDK 类型为 eagerzeroedthick，目标 VMDK 类型为精简格式。
- 源 VMDK 或目标 VMDK 为稀疏或托管格式。
- 源虚拟机拥有快照。
- 请求的操作中的逻辑地址和传输长度与存储设备所需的最低对齐要求不一致。使用 vSphere Client 创建的所有数据存储将自动对齐。
- VMFS 具有多个 LUN 或数据区，并且它们位于不同的阵列。

阵列之间的硬件克隆（即使在同一 VMFS 数据存储内）无法工作。

vSphere 支持两种模型的存储置备：厚置备和精简置备。

## 厚置备

厚置备是传统存储置备模型。对于厚置备，预先提供大量存储空间以满足未来的存储需要。但是，空间可能一直未被使用，这样会导致无法充分利用存储容量。

## 精简置备

此方法与厚置备相反，通过以灵活的按需方式分配存储空间，可帮助您消除无法充分利用存储的问题。通过 ESXi，可以使用两种模型的精简置备（阵列级别和虚拟磁盘级别）。

精简置备允许您报告比真实物理容量更多的虚拟存储空间。这种差异可能会导致存储超额订购，也称为过度置备。当使用精简置备时，要监控实际存储使用情况，避免在用尽物理存储空间时出现各种状况。

本章讨论了以下主题：

- [虚拟磁盘精简置备](#)
- [ESXi 和阵列精简置备](#)
- [存储空间回收](#)

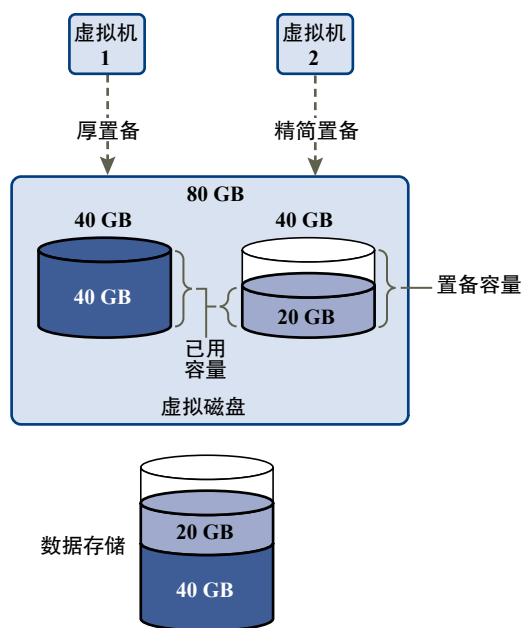
## 虚拟磁盘精简置备

创建虚拟机时，会在数据存储上为虚拟磁盘文件置备一定量的存储空间。

默认情况下，ESXi 为虚拟机提供传统的存储置备方法。使用这种方法，您可以先估算虚拟机完成其整个生命周期可能需要的存储空间。然后，预先为虚拟机虚拟磁盘置备固定量（如 40 GB）的存储空间。并将整个置备空间提交给虚拟磁盘。立即占据整个置备空间的虚拟磁盘为厚磁盘。

ESXi 支持精简置备虚拟磁盘。使用磁盘级别精简置备功能，可以采用精简格式创建虚拟磁盘。对于精简虚拟磁盘，ESXi 为磁盘的当前和未来活动置备所需的整个空间，例如 40 GB。但是，精简磁盘仅使用磁盘初始操作所需的存储空间大小。在此示例中，精简置备的磁盘只占用 20 GB 的存储空间。如果磁盘需要更多空间，可以扩展到其全部 40 GB 的置备空间。





## 关于虚拟磁盘置备策略

执行某些虚拟机管理操作时，您可以为虚拟磁盘文件指定置备策略。这类操作包括创建虚拟磁盘、将虚拟机克隆为模板或者迁移虚拟机。

带有硬件加速功能的 NFS 数据存储和 VMFS 数据存储支持以下磁盘置备策略。在不支持硬件加速功能的 NFS 数据存储上，只有精简格式可用。

可以使用 Storage vMotion 或跨主机 Storage vMotion 将虚拟磁盘从一种格式转换为另一种格式。

### 厚置备延迟置零

以默认的厚格式创建虚拟磁盘。在创建虚拟磁盘时分配该磁盘所需的空間。创建过程中不会清除物理设备上保留的数据，但以后首次从虚拟机写入时则会按需置零。虚拟机不会从物理设备读取失效数据。

### 厚置备快速置零

一种厚虚拟磁盘类型，可支持集群功能，如 Fault Tolerance。在创建时为虚拟磁盘分配所需的空間。与厚置备延迟置零格式相反，创建虚拟磁盘时，会将物理设备上保留的数据置零。创建这种格式的虚拟磁盘所需的时间可能会比创建其他类型的磁盘所用时间长。增加厚置备快速置零虚拟磁盘的大小会导致虚拟机关闭时间延长。

### 精简置备

使用此格式可节省存储空间。对于精简磁盘，可以根据输入的虚拟磁盘大小值置备磁盘所需的数据存储空间。但是，精简磁盘开始时很小，只使用与初始操作所需的大小完全相同的存储空间。如果精简磁盘以后需要更多空间，它可以增长到其最大容量，并占据为其置备的整个数据存储空间。

精简置备是创建虚拟磁盘的最快方法，因为它创建的磁盘仅具有头文件信息。它不会分配存储块或将其置零。初次访问存储块时，才分配存储块并将其置零。

**注** 如果虚拟磁盘支持集群解决方案（如 Fault Tolerance），请勿将磁盘设置为精简格式。

## 创建精简置备虚拟磁盘

为了节省存储空间，您可创建精简置备格式的虚拟磁盘。精简置备虚拟磁盘开始时很小，它随着磁盘空间的使用需求而扩展。您只能在支持磁盘级别精简置备的数据存储上创建精简磁盘。

该过程假设正在创建新虚拟机。有关信息，请参见《vSphere 虚拟机管理》文档。

### 步骤

#### 1 创建虚拟机。

- a 右键单击属于虚拟机的有效父对象的任何清单对象，例如数据中心、文件夹、集群、资源池或主机，然后选择**新建虚拟机**。
- b 选择**创建新的虚拟机**，然后单击**下一步**。
- c 遵循创建虚拟机所需的步骤。

#### 2 配置精简虚拟磁盘。

- a 在“自定义硬件”页面上，单击**虚拟硬件**选项卡。
- b 单击**新硬盘**三角形符号展开硬盘选项。
- c （可选）调整默认磁盘大小。

对于精简虚拟磁盘，磁盘大小值表示为磁盘置备和保证提供的空间。开始时，虚拟磁盘可能不会使用整个置备空间。实际存储使用值可以小于虚拟磁盘大小。

- d 针对磁盘置备选择**精简置备**。

#### 3 完成虚拟机创建。

### 结果

此时即已创建带有精简格式磁盘的虚拟机。

### 后续步骤

如果创建了精简格式的虚拟磁盘，则以后可以将其增加到最大大小。

## 查看虚拟机存储资源

可以查看虚拟机的数据存储存储空间的分配方式。

### 步骤

- 1 浏览到虚拟机。
- 2 双击虚拟机，然后单击**摘要**选项卡。
- 3 检查**摘要**选项卡右上方区域的存储使用信息。

### 结果

**存储使用情况**显示虚拟机文件（包括配置和日志文件、快照、虚拟磁盘等等）占用的数据存储空间。当虚拟机正在运行时，使用的存储空间还包括交换文件。

对于带有精简磁盘的虚拟机，实际存储使用值可能小于虚拟磁盘大小。

## 确定虚拟机的磁盘格式

可以确定虚拟磁盘是厚格式还是精简格式。

### 步骤

- 1 右键单击虚拟机，然后选择**编辑设置**。
- 2 单击**虚拟硬件**选项卡。
- 3 单击**硬盘**三角形符号展开硬盘选项。  
**类型**文本框会显示虚拟磁盘格式。

### 后续步骤

如果虚拟磁盘为精简格式，则可以将其扩充到其最大容量。

## 扩充精简虚拟磁盘

如果创建的是精简格式的虚拟磁盘，可以将其格式更改为厚磁盘。

使用数据存储浏览器扩充精简虚拟磁盘。

### 前提条件

- 请确保虚拟机驻留的数据存储具有足够的空间。
- 请确保虚拟磁盘为精简磁盘。
- 移除快照。
- 关闭虚拟机电源。

### 步骤

- 1 在 vSphere Client 中，导航到要扩充的虚拟磁盘的文件夹。
  - a 导航到虚拟机。
  - b 单击**数据存储**选项卡。  
此时将列出存储虚拟机文件的数据存储。
  - c 右键单击数据存储，然后选择**浏览文件**。  
数据存储浏览器将显示该数据存储的内容。
- 2 扩展虚拟机文件夹，浏览到要转换的虚拟磁盘文件。  
该文件使用 .vmdk 扩展名，并带有虚拟磁盘图标标记 (🗄️)。
- 3 选择虚拟磁盘文件，然后单击**扩充**。

---

**注** 如果虚拟磁盘为厚磁盘或当虚拟机正在运行时，该选项可能不可用。

---

结果

扩充的虚拟磁盘将占据最初为其置备的整个数据存储空间。

处理数据存储超额订购

由于为精简磁盘置备的空间可能大于提交的空间，因此可能发生数据存储超额订购，从而导致数据存储上的虚拟机磁盘总置备空间超过实际容量。

通常，所有附带精简磁盘的虚拟机不会同时需要整个置备数据存储空间，因此可能发生超额订购。但是，如果要避免数据存储超额订购，则可以设置警报，它会在置备空间达到特定阈值时通知您。

有关设置警报的信息，请参见《vCenter Server 和主机管理》文档。

如果虚拟机需要更多空间，则根据先来先服务的原则分配数据存储空间。当数据存储空间不足时，可以增加更多的物理存储器，并增加数据存储空间。

请参见[增加 VMFS 数据存储容量](#)。

ESXi 和阵列精简置备

您可以将精简置备的存储阵列用于 ESXi。

ESXi 主机与基于块的存储集成并执行以下任务：

- 主机可以识别底层精简置备 LUN 并监控空间使用情况以避免物理空间耗尽。LUN 空间可能会更改，例如，如果 VMFS 数据存储扩展，或者如果您使用 Storage vMotion 将虚拟机迁移到精简置备的 LUN 时。主机会提醒您物理 LUN 空间中的违规和空间不足状况。
- 主机可以自动从 VMFS6 和虚拟机客户机操作系统运行 T10 unmap 命令以从阵列回收未使用的空间。VMFS5 支持手动空间回收方法。

**注** ESXi 不支持在存储设备上启用和禁用精简置备。

要求

要使用精简置备报告和空间回收功能，请遵循以下要求：

- 使用适当的 ESXi 版本。

表 17-1. ESXi 版本和精简置备支持

支持的精简置备组件	ESXi 6.0 和更早版本	ESXi 6.5 及更高版本
精简置备	是	是
源自 VMFS 的取消映射命令	对于 VMFS5，需要手动发出。请使用 <code>esxcli storage vmfs unmap</code> 。	对于 VMFS6，可以自动发出。
源自客户机操作系统的取消映射命令	是。仅提供有限的支持。	是 (VMFS6)

- 使用支持基于 T10 的 vSphere Storage APIs - Array Integration (VAAI)（包括精简置备和空间回收）的存储系统。有关信息，请联系您的存储提供商并查看《VMware 兼容性指南》。

## 监控空间使用情况

精简置备集成功能可帮助您监控精简置备 LUN 上的空间使用情况并避免耗尽空间。

以下示例流程演示了 ESXi 主机和存储阵列如何进行交互来为精简置备 LUN 生成空间违规和空间不足的警告。使用 Storage vMotion 将虚拟机迁移到精简置备 LUN 时也可使用相同的机制。

- 1 通过使用特定于存储的工具，存储管理员置备精简 LUN 并设置软阈值限制，在达到该限制时会触发警告。此步骤是特定于供应商的。
- 2 通过使用 vSphere Client，在精简置备 LUN 上创建 VMFS 数据存储。数据存储可增大到 LUN 报告的最大逻辑大小。
- 3 由于数据存储所使用的空间增加且达到了设置的软阈值，因此会执行以下操作：

- a 存储阵列向主机报告空间违规。
- b 主机触发数据存储的警告警报。

您可以与存储管理员联系，请求更多的物理空间。或者，在 LUN 耗尽容量之前，使用 Storage vMotion 清空虚拟机。

- 4 如果没有剩余空间以分配到精简置备的 LUN，将执行以下操作：

- a 存储阵列向主机报告空间不足状况。

---

**小心** 在某些情况下，如果 LUN 已满，则可能会处于脱机状态或从主机取消映射。

---

- b 主机暂停虚拟机并生成空间不足警报。

通过向存储管理员请求更多的物理空间，可以解决永久的空间不足状况。

## 识别精简置备的存储设备

使用 `esxcli` 命令验证特定存储设备是否已精简置备。

### 前提条件

安装 ESXCLI。请参见《ESXCLI 入门》。要进行故障排除，请在 ESXi Shell 中运行 `esxcli` 命令。

### 步骤

- ◆ 运行 `esxcli storage core device list -d=device_ID` 命令。

### 结果

以下精简置备状态表示该存储设备已精简置备。

```
# esxcli storage core device list -d naa.XXXXXXXXXXXXX4c
naa.XXXXXXXXXXXXX4c
Display Name: XXXX Fibre Channel Disk(naa.XXXXXXXXXXXXX4c)
Size: 20480
Device Type: Direct-Access
Multipath Plugin: NMP
-----
```

```
Thin Provisioning Status: yes
```

```
-----
```

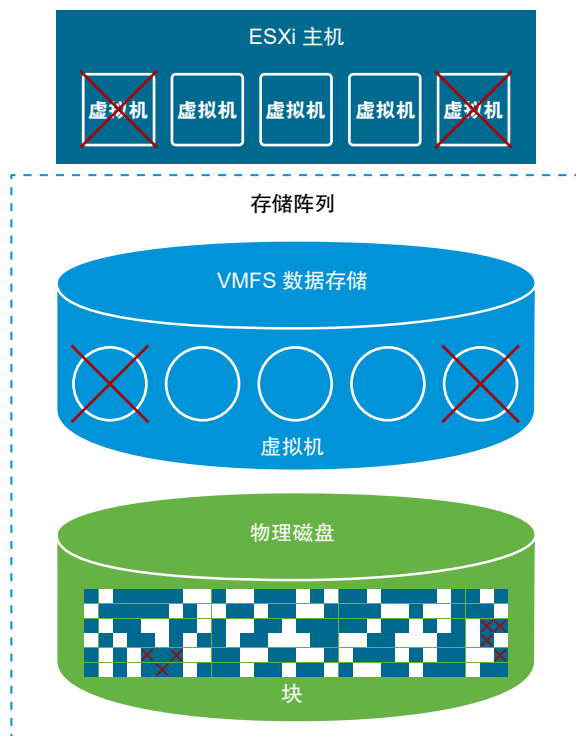
未知状态表示某个存储设备为厚格式。

**注** 有些存储系统将所有设备都显示为精简置备，无论这些设备是精简格式还是厚格式。其精简置备状态始终为 yes。有关详细信息，请咨询存储供应商。

## 存储空间回收

ESXi 支持从 VMFS 数据存储或虚拟机客户机操作系统发出的空间回收命令（也称为 SCSI 取消映射命令）。该命令可帮助精简置备存储阵列从 VMFS 数据存储和数据存储中的精简虚拟磁盘回收未使用的空间。VMFS6 数据存储可自动发送空间回收命令。通过 VMFS5 数据存储，可以手动回收存储空间。

在执行诸如删除或迁移虚拟机、整合快照等操作时，可释放 VMFS 数据存储内的存储空间。在虚拟机中，当删除精简虚拟磁盘上的文件时会释放存储空间。这些操作会将未使用的空间块保留在存储阵列上。但是，在阵列未感知到数据已从块中删除时，阵列仍将分配这些块，直到数据存储将其释放为止。VMFS 使用 SCSI 取消映射命令来指示阵列存储块包含已删除数据，以便阵列可取消分配这些块。



该命令也可从客户机操作系统直接发出。VMFS5 和 VMFS6 数据存储都可对从客户机操作系统进行的取消映射命令提供支持。但是，VMFS5 上的支持级别是有限的。

根据 VMFS 数据存储的类型不同，可以使用不同的方法来配置数据存储和虚拟机的空间回收。

观看以下视频，了解空间回收的工作原理。



## (VMFS 空间回收)

### ■ VMFS 数据存储上的空间回收

从 VMFS 数据存储删除或移除文件将释放文件系统内的空间。此可用空间会映射到某个存储设备，直到文件系统释放或取消映射该空间为止。ESXi 支持可用空间回收，也称为取消映射操作。

### ■ 来自客户机操作系统的空间回收请求

ESXi 支持使用直接从客户机操作系统发出的取消映射命令回收存储空间。支持级别和要求取决于虚拟机所在的数据存储的类型。

## VMFS 数据存储上的空间回收

从 VMFS 数据存储删除或移除文件将释放文件系统内的空间。此可用空间会映射到某个存储设备，直到文件系统释放或取消映射该空间为止。ESXi 支持可用空间回收，也称为取消映射操作。

此操作有助于存储阵列回收未使用的可用空间。然后，取消映射的空间可以用于其他存储分配请求和需求。

### 自动回收 VMFS6 数据存储上的可用空间

在 VMFS6 数据存储上，ESXi 支持自动式异步回收可用空间。VMFS6 可以在后台运行取消映射命令，以释放支持取消映射操作的精简置备存储阵列上的可用存储空间。

自动取消映射操作有多项优势：

- 取消映射请求以恒定速率发送，有助于避免备份阵列上出现任何即时加载。
- 释放的区域按批量一起取消映射。
- 其他工作负载的 I/O 性能不受取消映射命令的影响。

对于 VMFS6 数据存储，可以配置以下空间回收参数。

### 空间回收粒度

粒度定义底层存储可以回收的已释放空间扇区的最小大小。存储无法回收大小低于指定粒度的扇区。

对于 VMFS6，回收粒度等于块大小。如果指定块大小为 1 MB，则粒度也是 1 MB。不会回收大小小于 1 MB 的存储扇区。

---

**注** 某些存储阵列会建议最佳取消映射粒度。ESXi 支持自动取消映射过程，在阵列推荐的取消映射粒度为 1 MB 或更高版本，例如 16 MB。在最佳粒度为 1 MB 及更低的阵列上，如果粒度为 1 MB 的因数，则支持取消映射操作。例如，1 MB 可以被 512 字节、4 KB、64 KB 等整除。

---

### 空间回收方法

可以采用优先级回收方法，也采用固定回收方法。使用优先级回收方法时，需要配置优先级别。对于固定方法，必须以 MB/秒为单位指定回收速率。

创建 VMFS6 数据存储时，可以使用 vSphere Client 设置优先级方法。要启用固定方法，稍后可以修改现有数据存储的空间回收设置。

或者，也可以使用 `esxcli storage vmfs reclaim config set` 命令设置空间回收参数。

## 空间回收优先级速率

此参数定义使用优先级回收方法时执行空间回收操作的速率。通常，VMFS6 可以突然发送大量取消映射命令，也可以零星发送，具体取决于工作负载和配置。对于 VMFS6，可以指定以下选项之一。

空间回收优先级	描述	配置
无	为数据存储禁用取消映射操作。	vSphere Client esxcli 命令
低（默认设置）	以低速率（25-50 MB/秒）发送取消映射命令。	vSphere Client esxcli 命令
中等	以比低速率快两倍的速率（50-100 MB/秒）发送命令。	esxcli 命令
高	以比低速率快三倍的速率（100 MB/秒以上）发送命令。	esxcli 命令

启用空间回收后，VMFS6 数据存储只能在它至少具有一个打开的文件时才能开始释放未使用的空间块。例如，打开数据存储上其中一个虚拟机的电源时，可以满足这种条件。

## 创建 VMFS6 数据存储时配置优先级回收

空间回收设置控制如何回收支持数据存储的设备中已删除或未映射的块。在 vSphere Client 中创建 VMFS6 数据存储时，只能指定优先级这一空间回收方法。此外，还可以在创建数据存储时禁用空间回收。

### 步骤

- 1 在 vSphere Client 对象导航器中，浏览到主机、集群或数据中心。
- 2 从右键单击菜单中，选择**存储 > 新建数据存储**。
- 3 遵循创建 VMFS6 数据存储所需的步骤。
- 4 在**分区配置**页面上，指定空间回收参数。

这些参数定义执行空间回收操作时采用的粒度和优先级比率。您还可以使用此页面为数据存储禁用空间回收。

选项	描述
块大小	VMFS 数据存储中的块大小定义最大文件大小和文件占用的空间量。VMFS6 支持的块大小为 1 MB。
空间回收粒度	为取消映射操作指定粒度。取消映射粒度等于块大小，即 1 MB。 不会回收大小小于 1 MB 的存储扇区。
空间回收优先级	选择下列任一选项。 <ul style="list-style-type: none"> <li>■ 低（默认值）。使用固定优先级方法进行空间回收。以较低优先级比率启用取消映射操作。</li> <li>■ 无。如果您希望为数据存储禁用空间回收操作，请选择此选项。</li> </ul>

**注** 在 vSphere Client 中，空间回收优先级的唯一可用设置是“低”和“无”。要将设置更改为“中”或“高”，请使用 `esxcli` 命令。



## 5 完成数据存储创建过程。

### 结果

启用空间回收后，VMFS6 数据存储只能在它至少具有一个打开的文件时才能开始释放未使用的空间块。例如，打开数据存储上其中一个虚拟机的电源时，可以满足这种条件。

## 为现有 VMFS6 数据存储配置固定回收

要启用固定方法，请修改数据存储的空间回收设置。

### 步骤

- 1 在 vSphere Client 中，导航到数据存储。
- 2 从右键单击菜单中，选择**编辑空间回收**。
- 3 指定空间回收设置。

选项	描述
以固定速率启用自动空间回收	使用固定方法进行空间回收。以 MB/秒为单位指定回收速率。最小可用值为 10 MB/秒。
禁用自动空间回收	不回收已删除或未映射的块。 如果您希望为数据存储禁用空间回收操作，请选择此选项。

- 4 单击**确定**以保存新设置。
- 5 卸载并重新挂载数据存储以使更改生效。  
请参见 [挂载数据存储](#)。
- 6 对访问数据存储的所有 ESXi 主机重复此过程。

### 结果

修改的空间回收优先级值显示在数据存储的**常规**页面上。

## 使用 ESXCLI 命令更改 VMFS6 上的空间回收

可以更改默认空间回收优先级、粒度及其他参数。

### 步骤

- 1 使用以下命令设置空间回收参数。

```
esxcli storage vmfs reclaim config set
```

该命令采用以下选项：

选项	描述
-b --reclaim-bandwidth	空间回收固定带宽，以 MB/秒为单位。
-g --reclaim-granularity	自动空间回收的最小粒度，以字节为单位。

选项	描述
<b>-m --reclaim-method</b>	自动空间回收的方法。支持的选项包括： <ul style="list-style-type: none"> <li>■ 优先</li> <li>■ 固定</li> </ul>
<b>-pl --reclaim-priority</b>	自动空间回收的优先级。支持的选项包括： <ul style="list-style-type: none"> <li>■ 无</li> <li>■ 低</li> <li>■ 中</li> <li>■ 高</li> </ul>
<b>-l --volume-label</b>	目标 VMFS 卷的标签。
<b>-u --volume-uuid</b>	目标 VMFS 卷的 uuid。

2 如果已选择 `fixed` 作为回收方法，请从所有 ESXi 主机上卸载 VMFS6 数据存储，然后重新挂载。

a 卸载数据存储。

b 挂载数据存储。

此步骤可确保挂载 VMFS6 数据存储的所有 ESXi 主机都切换到数据存储的固定回收方法。

#### 示例：将回收方法设置为固定

使用以下示例将回收方法设置为固定，将速率设置为每秒 100 MB。

```
esxcli storage vmfs reclaim config set --volume-label datastore_name --reclaim-method fixed
-b 100
```

## 验证 VMFS6 上的自动空间回收设置

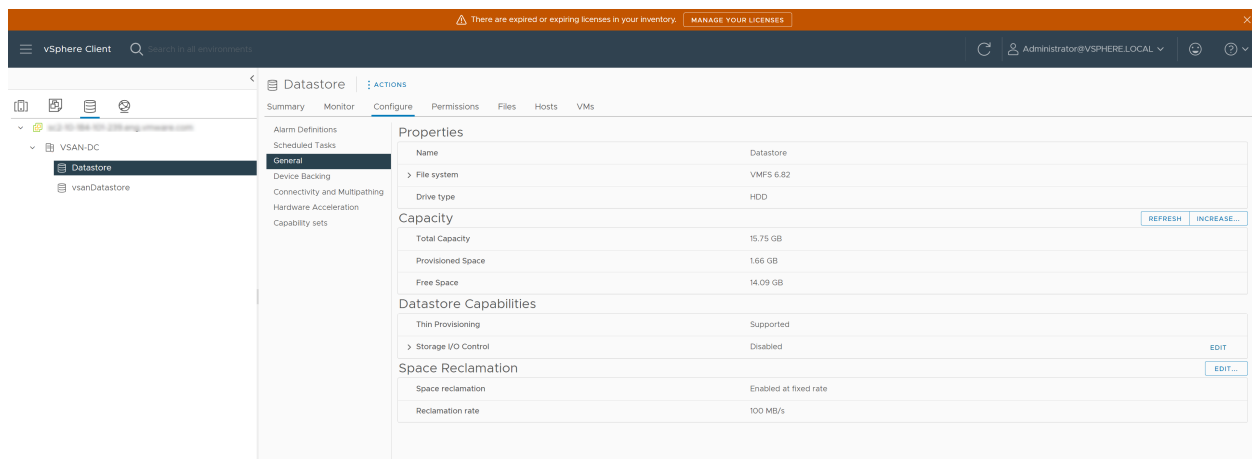
为 VMFS6 数据存储配置或编辑空间回收参数后，可以查看您的设置。

### 步骤

- 1 在 vSphere Client 中，导航到数据存储。
- 2 单击**配置**选项卡。
- 3 单击**常规**。
- 4 在**空间回收**下面，查看空间回收优先级设置。

如果通过 `esxcli` 命令配置了任何值，例如为空间回收优先级配置了 `Medium` 或 `High`，这些值还会显示在 vSphere Client 中。

## 结果



### 示例：获取 VMFS6 空间回收的参数

可以使用 `esxcli storage vmfs reclaim config get -l=VMFS_label|-u=VMFS_uuid` 命令获取空间回收配置的信息。

```
# esxcli storage vmfs reclaim config get -l my_datastore
Reclaim Granularity: 1048576 Bytes
Reclaim Priority: low
```

## 手动回收 VMFS5 上累积的存储空间

在不支持自动空间回收的 VMFS 数据存储（例如 VMFS5）上，可以使用 `esxcli` 命令手动回收未使用的存储空间。

使用此命令时请务必注意，它可能会一次性发送大量取消映射请求。在执行此操作期间可能会锁定某些资源。使用此命令时请务必注意，它可能会一次性发送大量取消映射请求。在执行此操作期间可能会锁定某些资源。

### 前提条件

安装 ESXCLI。请参见《ESXCLI 入门》。要进行故障排除，请在 ESXi Shell 中运行 `esxcli` 命令。

### 步骤

- 1 要回收精简置备设备上未使用的存储块，请运行以下命令：

```
esxcli storage vmfs unmap
```

该命令采用以下选项：

选项	描述
<code>-l --volume-label=volume_label</code>	要取消映射的 VMFS 卷的标签。必需的参数。如果指定了此参数，则请勿使用 <code>-u --volume-uuid=volume_uuid</code> 。
<code>-u --volume-uuid=volume_uuid</code>	要取消映射的 VMFS 卷的 UUID。必需的参数。如果指定了此参数，则请勿使用 <code>-l --volume-label=volume_label</code> 。
<code>-n --reclaim-unit=number</code>	每次迭代时要取消映射的 VMFS 块的数量。可选参数。如果未指定此参数，此命令将使用默认值 200。

2 要验证是否已完成取消映射过程，请在 `vmkernel.log` 文件中搜索取消映射。

## 来自客户机操作系统的空间回收请求

ESXi 支持使用直接从客户机操作系统发出的取消映射命令回收存储空间。支持级别和要求取决于虚拟机所在的数据存储的类型。

例如，在虚拟机中，当删除精简虚拟磁盘上的文件时，会释放存储空间。客户机操作系统通过发送取消映射命令向 VMFS 通知释放的空间。从客户机操作系统发送的取消映射命令会释放 VMFS 数据存储中的空间。然后，该命令进到阵列，以便阵列可以回收释放的空间块。

### VMFS6 虚拟机空间回收

通常，VMFS6 支持从客户机操作系统生成的自动空间回收请求，并将这些请求传递到阵列。许多客户机操作系统都可以发送取消映射命令，而无需任何额外的配置。不支持自动取消映射的客户机操作系统可能需要用户干预。有关支持 VMFS6 上自动空间回收的客户机操作系统的信息，请与供应商联系。

通常，客户机操作系统是基于其播发的取消映射粒度发送取消映射命令。有关详细信息，请参见客户机操作系统随附的文档。

使用 VMFS6 空间回收时，需要考虑以下注意事项：

- 只有在要回收的空间等于 1 MB 或是 1 MB 的倍数时，VMFS6 才会处理来自客户机操作系统的取消映射请求。如果该空间小于 1 MB 或不是 1 MB 倍数，则不会处理取消映射请求。
- 对于快照采用默认 SEsparse 格式的虚拟机，VMFS6 仅在 ESXi 主机版本 6.7 或更高版本上支持自动空间回收。

空间回收仅影响顶部快照，并且在虚拟机打开电源时才能发挥作用。

### VMFS5 虚拟机空间回收

通常，从 VMFS5 上的客户机操作系统生成的取消映射命令无法直接传递到阵列。您必须运行 `esxcli storage vmfs unmap` 命令为阵列触发取消映射。

但是，对于一些数量有限的客户机操作系统，VMFS5 支持自动空间回收请求。

要将取消映射请求从客户机操作系统发送至阵列，虚拟机必须满足以下必备条件：

- 虚拟磁盘必须为精简置备虚拟磁盘。
- 虚拟机硬件版本必须为 11 (ESXi 6.0) 或更高版本。

- 高级设置 EnableBlockDelete 必须设置为 1。
- 客户机操作系统必须能够将虚拟磁盘识别为精简虚拟置备磁盘。

Cloud Native Storage 是一种为有状态应用程序提供全面数据管理的解决方案。使用 Cloud Native Storage 时，您可以创建能够允许重新启动和中断的有状态容器化应用程序。有状态容器利用 vSphere 公开的存储，同时使用此类基本类型作为标准卷、持久性卷，并进行动态置备。

借助 Cloud Native Storage，您可以创建独立于虚拟机和容器生命周期的持久性容器卷。vSphere 存储支持卷，您可以直接在卷上设置存储策略。创建卷后，可以在 vSphere Client 中查看这些卷及支持它们的存储对象，并监控其存储策略合规性。

vSphere Cloud Native Storage 在以下 Kubernetes 发行版中支持持久卷：

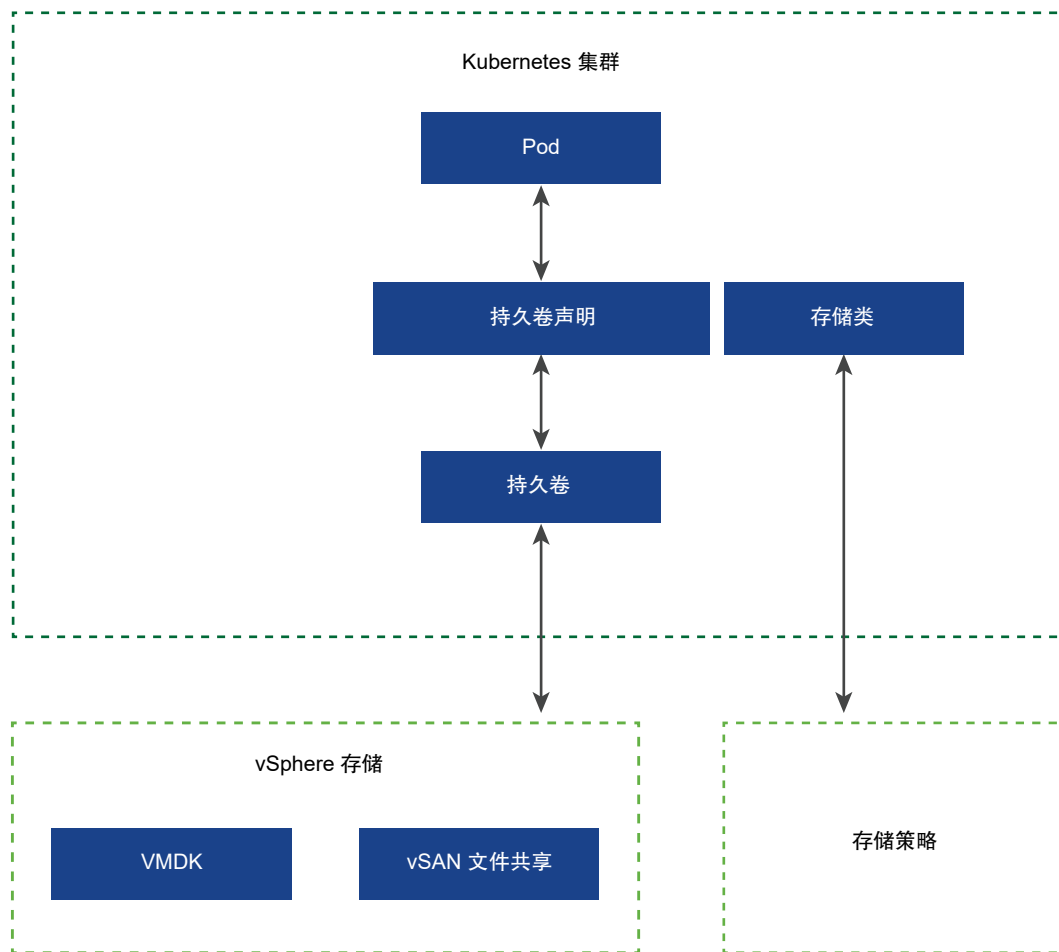
- 从官方存储库安装的通用 Kubernetes（也称为 vanilla）。本《vSphere 存储》文档仅介绍通用 Kubernetes。
- vSphere with Tanzu。有关详细信息，请参见《vSphere with Tanzu 概念和规划》文档。

本章讨论了以下主题：

- [Cloud Native Storage 概念和术语](#)
- [适用于 vSphere 管理员的 Cloud Native Storage](#)

## Cloud Native Storage 概念和术语

熟悉 vSphere Cloud Native Storage 环境的一些基本概念。



### Kubernetes 集群

在 Cloud Native Storage 环境中，可以在虚拟机集群中部署一个通用 Kubernetes 集群。在 Kubernetes 集群上，您可以部署容器化的应用程序。应用程序可以是有状态的和无状态的。

**注** 有关可在 vSphere with Tanzu 环境中运行的 主管 集群和 Tanzu Kubernetes Grid 集群的信息，请参见《vSphere with Tanzu 概念和规划》文档。

### Pod

Pod 是由一个或多个共享存储和网络等资源的容器化应用程序组成的组。Pod 中的容器以组的形式启动、停止和复制。

### 容器 Orchestrator

用于跨主机集群部署、扩展和管理容器化应用程序的开源平台，例如 Kubernetes。这些平台提供了以容器为中心的基础架构。

## 有状态应用程序

由于容器化应用程序从无状态发展为有状态，它们需要持久性存储。无状态应用程序不会在会话之间保存数据，与之不同的是，有状态应用程序会将数据保存到持久性存储中。保留的数据称为应用程序的状态。之后，您可以检索数据，并在下一个会话中使用该数据。大多数应用程序都是有状态的。数据库是有状态应用程序的示例。

### PersistentVolume

有状态应用程序使用 PersistentVolume 存储其数据。PersistentVolume 是一个能够保留其状态和数据的 Kubernetes 卷。PersistentVolume 独立于 Pod，即使删除或重新配置 Pod，它也可以继续存在。在 vSphere 环境中，PersistentVolume 对象使用第一类磁盘 (FCD) 类型的 vSphere 虚拟磁盘或 vSAN 文件共享作为其后备存储。第一类磁盘也称为增强型虚拟磁盘 (IVD) 或受管虚拟磁盘。

- 虚拟磁盘支持以 ReadWriteOnce 形式挂载的卷。这些卷只能由 Kubernetes 中的一个 Pod 使用。

可以使用 vSphere 加密技术保护支持持久卷的 FCD 虚拟磁盘。有关详细信息，请参见[将加密与云原生存储结合使用](#)。

- vSAN 文件共享支持许多节点挂载的 ReadWriteMany 卷。可以在跨 Kubernetes 节点或跨 Kubernetes 集群运行的多个 Pod 或应用程序之间共享这些卷。有关文件共享的可能配置的信息，请参见[使用 vSAN 文件服务置备文件卷](#)。

### StorageClass

Kubernetes 使用 StorageClass 定义不同的存储层，并描述不同类型的支持 PersistentVolume 的存储要求。在 vSphere 环境中，可以将存储类链接到存储策略。作为 vSphere 管理员，您可以创建存储策略来描述不同的存储要求。这些虚拟机存储策略可用作 StorageClass 定义的一部分，以便进行动态卷置备。

以下示例 YAML 文件引用之前使用 vSphere Client 创建的 **Gold** 存储策略。生成的持久卷 VMDK 放置在满足 **Gold** 存储策略要求的兼容数据存储上。

```
kind: StorageClass
apiVersion: storage.k8s.io/v1
metadata:
  name: gold-sc
  annotations:
    storageclass.kubernetes.io/is-default-class: "true"
provisioner: csi.vsphere.vmware.com
parameters:
  storagepolicyname: "Gold"
```

### PersistentVolumeClaim

通常，应用程序或 Pod 可以通过 PersistentVolumeClaim 请求持久存储。PersistentVolumeClaim 可为 PersistentVolume 指定存储的类型和类、访问模式（ReadWriteOnce 或 ReadWriteMany）以及其他参数。然后，该请求可以在 vSphere 环境中动态置备相应的 PersistentVolume 对象和底层虚拟磁盘或 vSAN 文件共享。



创建声明后，PersistentVolume 会自动绑定到声明。Pod 使用声明挂载 PersistentVolume 并访问存储。

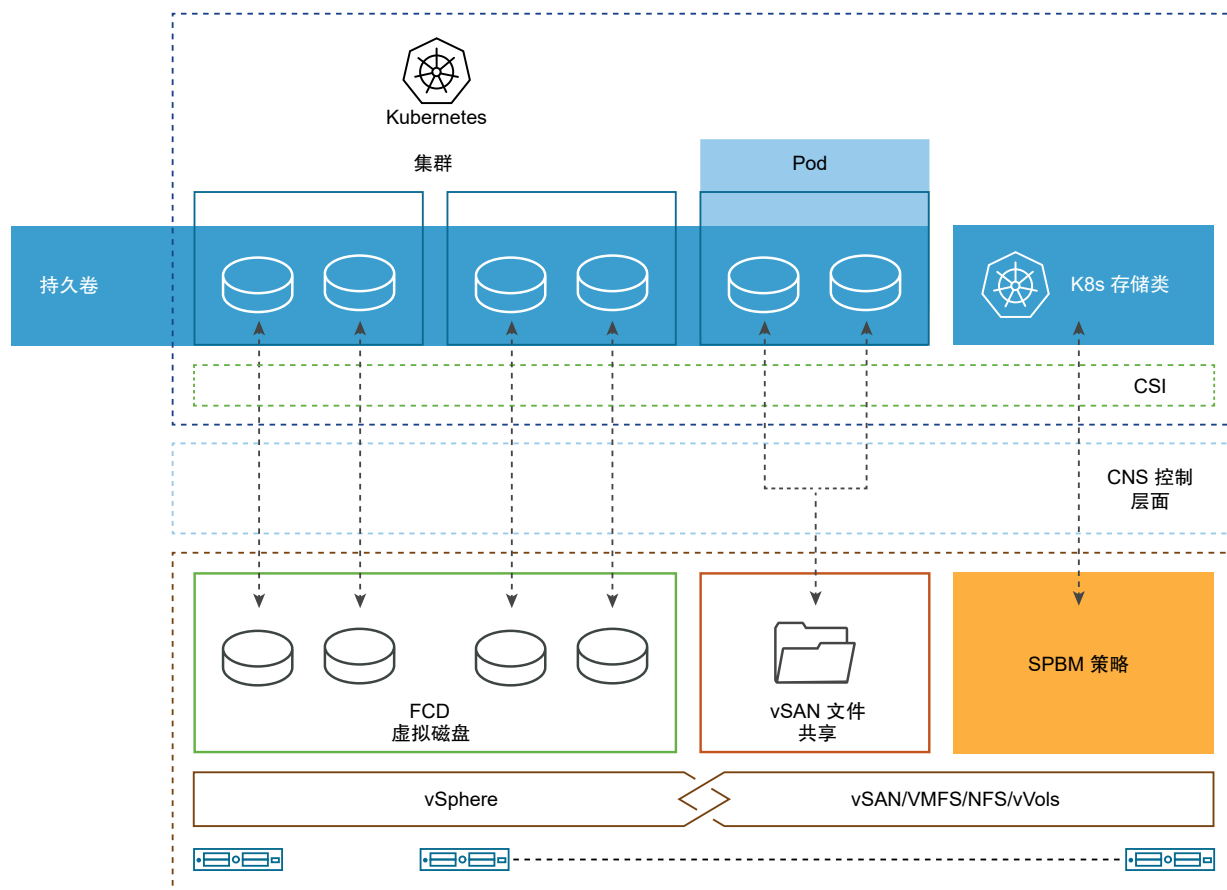
删除此声明时，将删除相应的 PersistentVolume 对象和底层存储。

```
kind: PersistentVolumeClaim
metadata:
  name: persistent-VMDBK
spec:
  accessModes:
    - ReadWriteOnce
  resources:
    requests:
      storage: 5Gi
      storageClassName: gold-sc
```

## Cloud Native Storage 组件

Cloud Native Storage 使用多个组件与 vSphere 存储进行集成。

下图显示了这些组件的交互方式。



## Kubernetes 集群

在 Cloud Native Storage 环境中，会在 vSphere 中运行的虚拟机或节点集群上部署一个通用 Kubernetes 集群。在其上部署有状态应用程序时，Kubernetes 用户直接与集群进行交互。

---

**注** 有关可在 vSphere with Tanzu 环境中运行的 主管 集群和 Tanzu Kubernetes Grid 集群的信息，请参见《vSphere with Tanzu 概念和规划》文档。

---

## 适用于 vSphere 的容器存储接口 (CSI)

要使用底层基础架构资源，集群需要具有 CSI 驱动程序。

vSphere CSI 是一个树外插件，可向容器 Orchestrator（如 Kubernetes）上的容器化工作负载公开 vSphere 存储。该插件支持 vSAN 和其他类型的 vSphere 存储。

vSphere CSI 与 vCenter Server 上的 CNS 控制平面进行通信，以便执行所有存储置备操作。

vSphere CSI 支持以下功能：

- 动态置备容器卷。
- vSphere 第一类磁盘功能。
- Kubernetes 区域。
- 常规挂载和裸挂载。
- 单个 vCenter Server 及多个数据中心和集群。
- 从多个数据存储或数据存储集群置备。
- vSAN 文件服务

在 Kubernetes 上，CSI 驱动程序与树外 vSphere 云提供程序接口 (CPI) 一起使用。CSI 驱动程序以容器映像的形式提供，必须由集群管理员进行部署。有关信息，请参见 [Kubernetes vSphere CSI 驱动程序文档中的驱动程序部署部分](#)。

有关 主管 集群和 Tanzu Kubernetes Grid 集群中所用 CSI 变体的信息，请参见《vSphere with Tanzu 概念和规划》文档。

## Cloud Native Storage 服务器组件

CNS 服务器组件或 CNS 控制平面驻留在 vCenter Server 中。它是 vCenter Server 管理的扩展，用于执行容器卷的置备和生命周期操作。

置备容器卷时，它与 vCenter Server 进行交互，以创建支持这些卷的存储对象。基于存储策略的管理功能可保证提供卷所需的服务级别。

CNS 还执行查询操作，以便允许您通过 vCenter Server 管理和监控容器卷及其备用存储对象。

## 第一类磁盘 (FCD)

也称为增强型虚拟磁盘 (IVD) 或受管虚拟磁盘。它是与虚拟机无关的已命名虚拟磁盘。这些磁盘驻留在 vSAN、VMFS、NFS 或 vVols 数据存储上，并且支持 ReadWriteOnce 容器卷。

FCD 技术允许在虚拟机或 Pod 生命周期之外执行与持久卷相关的生命周期操作。如果虚拟机是 Kubernetes 节点，运行多个基于容器的应用程序，并为许多应用程序使用持久卷和虚拟磁盘，则 CNS 便于以容器和持久卷粒度执行生命周期操作。

## vSAN 文件服务

它是提供文件共享的 vSAN 层。目前支持 NFSv3 和 NFSv4.1 文件共享。Cloud Native Storage 对 ReadWriteMany 类型的持久卷使用 vSAN 文件共享。一个 ReadWriteMany 卷可以挂载到多个节点。卷可在跨 Kubernetes 节点或跨 Kubernetes 集群运行的多个 pod 或应用程序之间共享。

### 基于存储策略的管理

基于存储策略的管理是一种 vCenter Server 服务，它支持根据指定的存储要求置备持久性卷。置备后，服务会监控具有所需策略特性的卷的合规性。

## 使用 vSAN 文件服务置备文件卷

vSAN 文件服务提供 ReadWriteMany (RWM) 类型的持久卷使用的 vSAN 文件共享。一个 RWM 卷可以由多个节点挂载。卷可在跨 Kubernetes 节点或跨 Kubernetes 集群运行的多个 pod 或应用程序之间共享。

Kubernetes pod 请求 RWM 卷时，Cloud Native Storage 会与 vSAN 文件服务进行通信，以创建所请求大小和存储类的基于 NFS 的文件共享。然后，Cloud Native Storage 将 RWM 卷挂载到运行 pod 的 Kubernetes 工作节点。如果多个节点请求访问 RWM 卷，Cloud Native Storage 将确定该特定部署的 RWM 卷已存在，然后将现有卷挂载到节点。

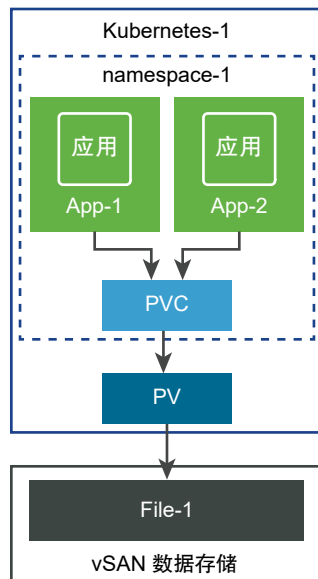
为了能够支持 RWM 卷，您的环境必须包括以下项。

- 使用 vSAN 的 vSphere 7.0 及更高版本
- 已启用 vSAN 文件系统。有关信息，请参见《管理 VMware vSAN》文档中的 [vSAN 文件服务](#)。
- Kubernetes 1.14 及更高版本
- 兼容版本的 CSI。有关信息，请参见 [VMware vSphere Container Storage Plug-in 文档](#)。

您可以为文件卷使用不同的配置。

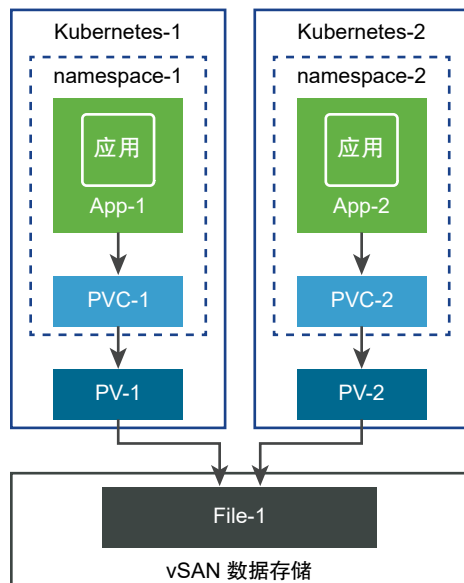
### 跨同一命名空间中的应用程序共享单个文件卷

在此示例中，跨同一命名空间中的不同应用程序使用单个文件卷作为共享存储。使用单个持久卷声明来置备文件卷。



### 跨应用程序和命名空间共享单个文件卷

此示例在不同的应用程序和不同的命名空间之间使用单个文件卷作为共享存储。对于每个命名空间，请创建单独的持久卷声明来置备同一文件卷。



## Cloud Native Storage 用户

在 vSphere Cloud Native Storage 环境中，创建和监控 Kubernetes 卷的过程所涉及的用户类型通常分为两类：Kubernetes 用户和 vSphere 管理员。这两种类型的用户都有权访问不同的工具并执行不同的任务。

### CNS Kubernetes 用户

Kubernetes 用户可以是 Kubernetes 开发人员和应用程序所有者、Kubernetes 管理员，或将两者的功能组合在一起。Kubernetes 用户在 Cloud Native Storage 环境中执行的任务包括：

- 部署和管理 vSphere CSI。有关信息，请参见《[VMware vSphere 容器存储插件入门指南](#)》文档中的 [vSphere 容器存储插件部署](#) 部分。
- 置备持久卷。有关块卷的信息，请参见 [vSphere CSI 驱动程序- 块卷](#)。有关文件卷的信息，请参见 [vSphere CSI 驱动程序 - 文件卷](#)。
- 执行持久性卷的生命周期操作。
- 执行存储类的生命周期操作。

### CNS vSphere 用户

CNS vSphere 用户或 vSphere 管理员有权访问 vSphere Client 以执行以下任务：

- 执行虚拟机存储策略的生命周期操作。例如，创建一个用于 Kubernetes 存储类的虚拟机存储策略，并将其名称传达给 Kubernetes 用户。请参见 [Kubernetes 创建存储策略](#)。
- 使用 vSphere Client 的 Cloud Native Storage 部分监控 Kubernetes 集群中容器卷的运行状况和存储策略合规性。请参见 [监控 Kubernetes 集群中的容器卷](#)。

## 适用于 vSphere 管理员的 Cloud Native Storage

vSphere 管理员向 Kubernetes 团队提供存储资源，并创建用于描述不同存储要求和服务类的虚拟机存储策略。置备具有持久存储的 Kubernetes 工作负载后，vSphere 管理员可以监控支持存储资源的生命周期及其与要求的合规性。

## Cloud Native Storage 的要求和限制

Cloud Native Storage 环境和加入 Kubernetes 集群的虚拟机必须满足多个要求。

### Cloud Native Storage 要求

- 兼容的 vSphere 版本。
- 兼容版本的 Kubernetes。
- 部署在虚拟机上的 Kubernetes 集群。有关部署 vSphere CSI 插件并在 vSphere 上运行 Kubernetes 集群的详细信息，请参见 [VMware vSphere 容器存储插件文档](#)。

## Kubernetes 集群虚拟机的要求

- 硬件版本为 15 或更高版本的虚拟机。在每个节点虚拟机上安装 VMware Tools。
- 虚拟机硬件建议：
  - 根据工作负载要求充分设置 CPU 和内存。
  - 对节点虚拟机上的主磁盘使用 VMware 准虚拟 SCSI 控制器。
- 所有虚拟机都必须有权访问共享数据存储，例如 vSAN。
- 在每个节点虚拟机上设置 `disk.EnableUUID` 参数。请参见[配置 Kubernetes 群集虚拟机](#)。
- 为避免出现错误和不可预知的行为，请勿对 CNS 节点虚拟机生成快照。

## CNS 文件卷的要求

- 将 vSphere 版本 7.0 或更高版本与兼容的 Kubernetes 版本结合使用。
- 使用兼容版本的 CSI。有关信息，请参见 [VMware vSphere 容器存储插件文档](#)。
- 启用并配置 vSAN 文件服务。必须配置必要的文件服务域、IP 池、网络等。有关信息，请参见《管理 VMware vSAN》文档。
- 按照特定准则配置从 Kubernetes 节点中的客户机操作系统到 vSAN 文件共享的网络访问。请参见[配置 vSAN 文件共享的网络访问](#)。

## Cloud Native Storage 限制

Cloud Native Storage 不支持 vSAN 延伸集群和站点容灾。

## 配置 vSAN 文件共享的网络访问

要能够在通用 vSphere Kubernetes 环境中置备 ReadWriteMany 持久卷，请配置 Kubernetes 节点到 vSAN 文件服务网络所需的网络、交换机和路由器。

### 设置网络

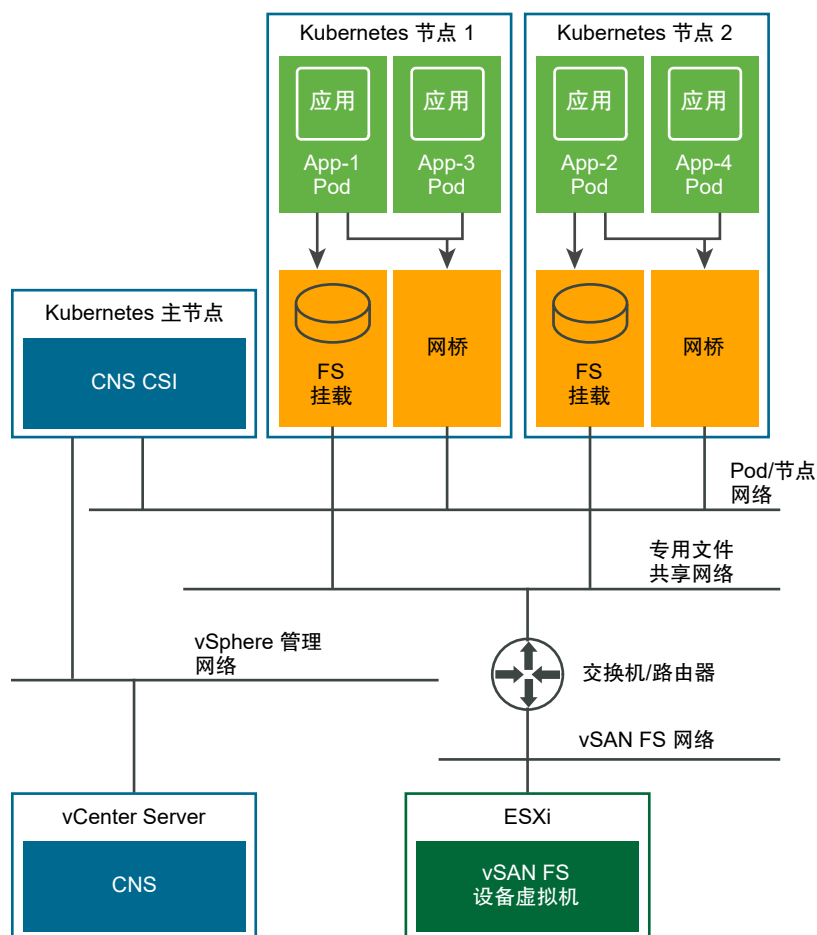
配置网络时，请遵循以下要求：

- 在每个 Kubernetes 节点上，可以使用专用虚拟网卡传输 vSAN 文件共享流量。仅当要对文件卷使用安全数据流量路径时，才需要此选项。
- 如果使用专用虚拟网卡，请确保通过专用虚拟网卡的流量可路由到一个或多个 vSAN 文件服务网络。
- 确保每个 Kubernetes 节点上只有客户机操作系统可以通过文件共享 IP 地址直接访问 vSAN 文件共享。节点中的 pod 不能 ping 或通过 IP 地址访问 vSAN 文件共享。

CNS CSI 驱动程序确保仅那些配置为使用 CNS 文件卷的容器可以通过在客户机操作系统中创建挂载点来访问 vSAN 文件共享。

- 避免在节点虚拟机和 vSAN 文件共享之间产生 IP 地址冲突。

下图是使用 vSAN 文件共享服务的 CNS 网络配置示例。



在该图中，示例网络配置遵循以下准则。

- 该配置对 CNS 环境中的不同项目使用单独的网络。

网络	描述
vSphere 管理网络	通常，在通用 Kubernetes 集群中，每个节点都有权访问此网络。
Pod 或节点网络	Kubernetes 使用该网络进行节点到节点通信或 pod 到 pod 通信。
专用文件共享网络	CNS 文件卷数据流量使用此网络。
vSAN 文件共享网络	启用了 vSAN 文件共享且文件共享可用的网络。

- 每个 Kubernetes 节点都有一个专用虚拟网卡用于文件流量。此虚拟网卡与用于节点到节点通信或 pod 到 pod 通信的虚拟网卡不同。此配置仅用作示例，但并不强制。
- 只有那些配置为使用 CNS 文件共享的应用程序才能通过节点客户机操作系统中的挂载点访问 vSAN 文件共享。例如，在该图中，存在以下行为：
  - App-1 和 App-2 pod 配置为使用文件卷，并且可以通过 CSI 驱动程序创建的挂载点访问文件共享。
  - App-3 和 App 4 未配置文件卷，无法访问文件共享。

- vSAN 文件共享作为容器部署在 ESXi 主机上的 vSAN 文件共享设备虚拟机中。Kubernetes Deployer 是一种可配置、部署和管理 Kubernetes 集群的软件或服务，可配置必要的路由器和交换机，以便 Kubernetes 节点中的客户机操作系统可以访问 vSAN 文件共享。

### 安全限制

尽管专用虚拟网卡可防止未经授权的 pod 直接访问文件共享，但存在某些安全限制：

- CNS 文件功能假设具有 CNS 文件卷 ID 的任何用户都是卷的授权用户。具有 CNS 文件卷 ID 的任何用户都可以访问卷中存储的数据。
- CNS 文件卷仅支持 AUTH\_SYS 身份验证，该身份验证是基于用户 ID 的身份验证。为保护对 CNS 文件卷中数据的访问，必须对访问 CNS 文件卷的容器使用相应的用户 ID。
- 引用 CNS 文件卷的未绑定 ReadWriteMany 持久卷可由任何命名空间下任何 Kubernetes 用户创建的持久卷声明绑定。确保只有授权用户才能访问 Kubernetes 以避免出现安全问题。

### 配置 CSI 驱动程序以访问 vSAN 文件服务集群

根据配置，CSI 驱动程序可以在启用了文件服务的一个或多个 vSAN 集群上置备文件卷。

可以限制仅对启用了文件服务的特定 vSAN 集群进行访问。部署 Kubernetes 集群时，配置 CSI 驱动程序，使其能够访问特定的文件服务 vSAN 集群。因此，CSI 驱动程序只能在这些 vSAN 集群上置备文件卷。

在默认配置中，CSI 驱动程序使用 vCenter Server 中提供的任何文件服务 vSAN 集群进行文件卷置备。CSI 驱动程序不会验证在置备文件卷时可访问的文件服务 vSAN 集群。

## Cloud Native Storage 角色和特权

CNS vSphere 用户必须具有特定的特权才能执行与 Cloud Native Storage 相关的操作。

您可以创建多个角色，以便为加入 Cloud Native Storage 环境的对象分配权限集。

**注** 只需为通用 Kubernetes 集群创建这些角色。如果在 vSphere with Tanzu 环境中工作，请使用工作负载存储主管角色执行存储操作。

有关 vSphere 中的角色和权限以及如何创建角色的详细信息，请参见《vSphere 安全性》文档。

角色名称	特权名称	描述	要求
CNS-Datastore	数据存储 > 低级别文件操作	允许在数据存储浏览器中执行读取、写入、删除和重命名操作。	持久卷所在的共享数据存储。
CNS-HOST-CONFIG-STORAGE	主机 > 配置 > 存储分区配置	允许执行 vSAN 数据存储管理。	在具有 vSAN 文件服务的 vSAN 集群上为必需项。仅对文件卷为必需项。
CNS-VM	虚拟机 > 更改配置 > 添加现有磁盘	允许将现有虚拟磁盘添加到虚拟机。	所有集群节点虚拟机。
	虚拟机 > 更改配置 > 添加或移除设备	允许添加或移除任何非磁盘设备。	



角色名称	特权名称	描述	要求
CNS- SEARCH- AND-SPBM	CNS > 可搜索	允许存储管理员查看云原生存储 UI。	根 vCenter Server。
	配置文件驱动的存储 > 配置文件驱动的存储视图	允许查看定义的存储策略。	
只读	默认角色	具有“只读”角色的对象用户可查看对象的状态和详细信息。例如，具有此角色的用户可以找到所有节点虚拟机均可访问的共享数据存储。  对于区域和拓扑感知环境，节点虚拟机的所有先代（例如，主机、集群和数据中心）都必须针对配置为使用 CSI 驱动程序和 CCM 的 vSphere 用户设置只读角色。要允许读取标记和类别以准备节点拓扑，必须满足此条件。	节点虚拟机所在的所有主机数据中心

## 为 Kubernetes 创建存储策略

支持 Kubernetes 容器化应用程序的 vSphere 存储对象需要满足特定的存储要求。作为 vSphere 用户，您可以根据 Kubernetes 用户提供给您的要求创建虚拟机存储策略。

存储策略将与支持 Kubernetes 容器的虚拟磁盘或 vSAN 文件共享相关联。

如果您的环境中有多于一个 vCenter Server 实例，请在每个实例上创建虚拟机存储策略。在所有实例中使用相同的策略名称。

### 前提条件

- Kubernetes 用户识别将在其中部署有状态容器化应用程序的 Kubernetes 集群。
- Kubernetes 用户收集容器化应用程序的存储要求，并将其传达给 vSphere 用户。
- 所需特权：[虚拟机存储策略.更新](#)和[虚拟机存储策略.查看](#)。

### 步骤

- 1 在 vSphere Client 中，打开[创建虚拟机存储策略](#)向导。
  - a 单击 **菜单 > 策略和配置文件**。
  - b 在 **策略和配置文件** 下，单击 **虚拟机存储策略**。
  - c 单击 **创建**。

- 2 输入策略名称和描述，然后单击下一步。

选项	操作
vCenter Server	选择 vCenter Server 实例。
名称	输入存储策略的名称，例如 <b>高效存储</b> 。
描述	输入存储策略的描述。

- 3 在**策略结构**页面的“数据存储特定规则”下，选择为 **vSAN 存储启用规则**，然后单击下一步。

- 4 在 **vSAN** 页面上，定义策略规则集，然后单击下一步。

- a 在**可用性**选项卡上，定义**站点灾难容错**和**允许的故障数**。

**注** 对于**站点容灾**，请选择**无 - 标准集群**。不要选择与延伸集群相关的选项。Cloud Native Storage 不支持 vSAN 延伸集群和站点容灾。

- b 在**高级策略规则**选项卡上，定义高级策略规则，例如每个对象的磁盘带数和闪存读取缓存预留。

- 5 在**存储兼容性**页面上，检查与此策略匹配的 vSAN 数据存储列表，然后单击下一步。

- 6 在**检查并完成**页面上，检查策略设置，然后单击**完成**。

Edit VM Storage Policy

1 Name and description
2 Policy structure
3 vSAN
4 Storage compatibility
5 Review and finish

Review and finish

General
Name
Description
vCenter Server

Space-Efficient
sc2-rdops-vm08-dhcp-23-199.eng.vmware.com

vSAN
Availability
Site disaster tolerance
Failures to tolerate
Advanced Policy Rules
Number of disk stripes per object
IOPS limit for object
Object space reservation
Flash read cache reservation
Disable object checksum
Force provisioning

None - standard cluster
No data redundancy
1
0
Thin provisioning
0%
No
No

CANCEL
BACK
FINISH

## 后续步骤

现在，您可以将存储策略名称告知 **Kubernetes** 用户。您创建的虚拟机存储策略将用作存储类定义的一部分，以便进行动态卷置备。

## 配置 Kubernetes 群集虚拟机

在每个节点虚拟机上，启用 `disk.EnableUUID` 参数，以便虚拟机可以成功挂载虚拟磁盘。  
对加入群集的每个虚拟机节点执行以下步骤。

### 前提条件

- 为 Kubernetes 群集创建多个虚拟机。有关虚拟机要求，请参见 [Cloud Native Storage 的要求和限制](#)。
- 所需特权：[虚拟机.配置.设置](#)。

**注** 为避免出现错误和不可预知的行为，请勿对 CNS 节点虚拟机生成快照。

### 步骤

- 1 在 vSphere Client 中，右键单击虚拟机，然后选择[编辑设置](#)。
- 2 单击[虚拟机选项](#)选项卡，然后展开[高级](#)菜单。
- 3 单击“配置参数”旁边的[编辑配置](#)。
- 4 配置 `disk.EnableUUID` 参数。

如果该参数存在，请确保其值设置为 True。如果该参数不存在，请添加并将其值设置为 True。

名称	值
<code>disk.EnableUUID</code>	True

## 监控 Kubernetes 集群中的容器卷

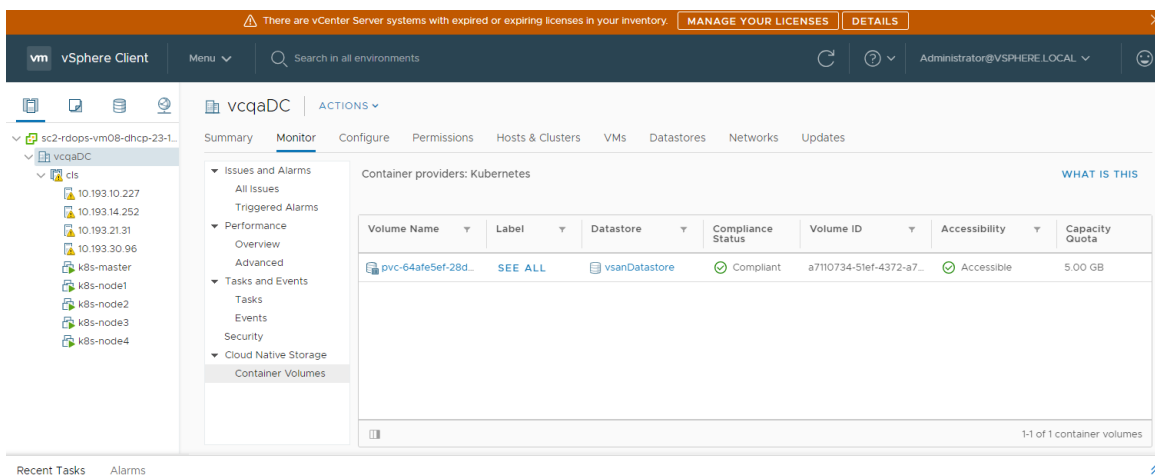
在 Kubernetes 中部署有状态应用程序后，卷及其支持的 vSphere 存储对象将显示在 vSphere Client 中。可以显示和监控卷，并对任何潜在存储问题进行故障排除。

**注** 如果在 Kubernetes CNS 服务器上遇到故障，则在执行完全同步之前，vSphere Client 中的 CNS 对象可能无法正确显示。

### 步骤

- 1 导航到 vCenter Server 实例、数据中心或数据存储。
- 2 单击[监控](#)选项卡，然后单击云原生存储下的[容器卷](#)。

### 3 观察环境中可用的容器卷，并监控其存储策略合规性状态。

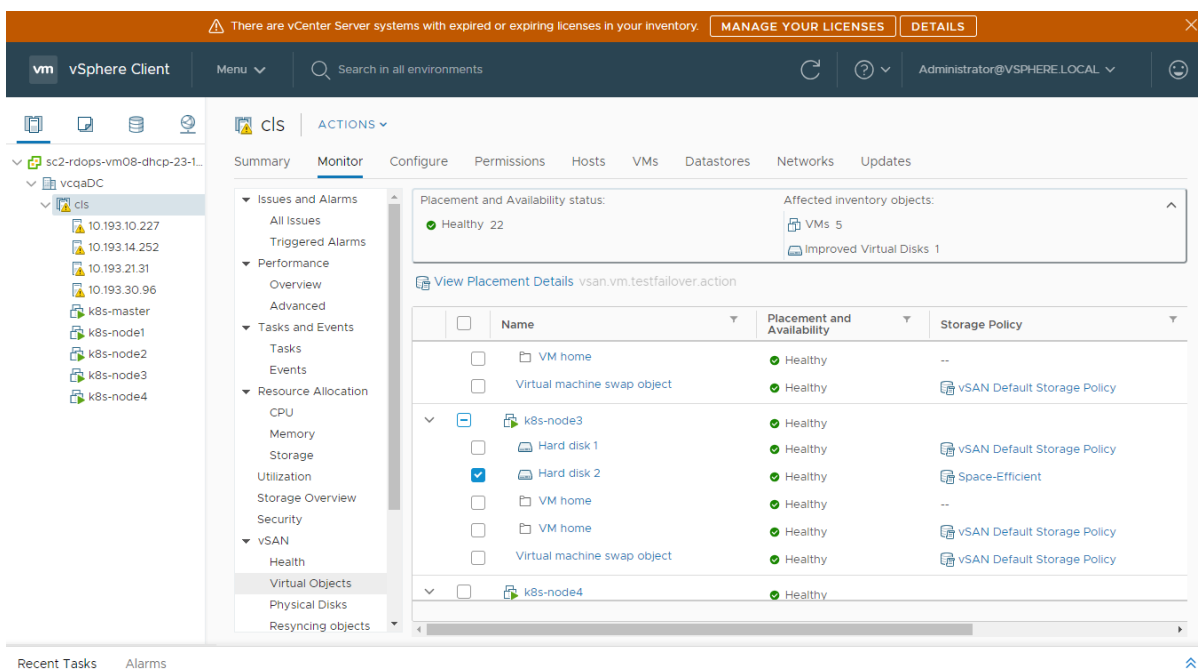


### 4 单击“标签”列中的[查看全部](#)链接以查看其他详细信息。

详细信息包括 PersistentVolumeClaim、StorageClass 等的名称，并帮助您将卷映射到与其关联的 Kubernetes 对象。

### 5 单击卷名称列中的链接，以查看支持卷的各种组件以及放置、合规性和存储策略等详细信息。

**注** 仅当底层数据存储为 vSAN 时，才显示 **Virtual Volumes** 屏幕。



## 将加密与云原生存储结合使用

可以使用 vSphere 加密技术保护支持持久卷的 FCD 虚拟磁盘。

在 vSphere 环境中使用加密需要做好一些准备，包括在 vCenter Server 与密钥提供程序之间设置可信连接。然后，vCenter Server 可以根据需要从密钥提供程序检索密钥。有关参与 vSphere 加密过程的组件的信息，请参见《vSphere 安全性》文档中的 [vSphere 虚拟机加密组件](#)。

### 步骤

- 1 在 vSphere 环境中设置密钥提供程序。

有关信息，请参见[设置密钥管理服务服务器集群](#)。

- 2 对 Kubernetes 集群上的所有节点虚拟机进行加密。

使用 vSphere Client 执行此步骤。

- a 导航到节点虚拟机。
- b 从右键单击菜单中，选择**虚拟机策略 > 编辑虚拟机存储策略**。
- c 从**虚拟机存储策略**下拉菜单中，选择**虚拟机加密策略**，然后单击**确定**。

要加快节点虚拟机的加密过程，可以只加密虚拟机主页。

- 3 使用 vSphere CSI 设置在 Kubernetes 集群中创建加密持久卷。

- a 创建引用虚拟机加密存储策略的 StorageClass。

可参考以下 YAML 文件示例。

```
kind: StorageClass
apiVersion: storage.k8s.io/v1
metadata:
  name: encryption
provisioner: csi.vsphere.vmware.com
parameters:
  storagePolicyName: "VM Encryption Policy"
  datastore: vsanDatastore
```

- b 使用 PersistentVolumeClaim 置备持久卷。

PersistentVolumeClaim 必须在 storageClassName 字段中包含加密存储类的名称。

vmkfstools 是用于管理 VMFS 卷、存储设备和虚拟磁盘的 ESXi Shell 命令之一。可以使用 vmkfstools 命令执行很多存储操作。例如，可以在物理分区上创建和管理 VMFS 数据存储，或操作 VMFS 或 NFS 数据存储中存储的虚拟磁盘文件。

**注** 使用 vmkfstools 进行更改后，vSphere Client 可能不会立即更新。在客户端使用刷新或重新扫描操作。

有关 ESXi Shell 的详细信息，请参见《ESXCLI 入门》。

本章讨论了以下主题：

- [vmkfstools 命令语法](#)
- [vmkfstools 命令选项](#)

## vmkfstools 命令语法

通常，您无需以 root 用户身份登录即可运行 vmkfstools 命令。但是，有些命令可能需要以 root 用户身份登录，如文件系统命令。

vmkfstools 命令支持以下命令语法：

vmkfstools *options target*。

target 是指将命令选项应用到的分区、设备或路径。

**表 19-1. vmkfstools 命令参数**

参数	描述
选项	用于指定 vmkfstools 要执行的活动的的一个或多个命令行选项及关联参数。例如，创建新虚拟磁盘时选择磁盘格式。 输入选项后，指定执行此操作的目标。目标可以是分区、设备或路径。
分区	指定磁盘分配。此参数使用 <i>disk_ID:P</i> 格式，其中， <i>disk_ID</i> 是存储阵列返回的设备 ID， <i>P</i> 是代表分区编号的整数。分区数字必须大于零 (0)，并且必须对应有效 VMFS 分区。

表 19-1. vmkfstools 命令参数（续）

参数	描述
设备	<p>指定设备或逻辑卷。此参数使用 ESXi 设备文件系统中的路径名。路径名以 /vmfs/devices（设备文件系统的挂载点）开头。</p> <p>指定不同设备类型时，请使用以下格式：</p> <ul style="list-style-type: none"><li>■ 本地磁盘或基于 SAN 的磁盘：/vmfs/devices/disks。</li><li>■ ESXi 逻辑卷：/vmfs/devices/lvm。</li><li>■ 一般 SCSI 设备：/vmfs/devices/generic。</li></ul>
路径	<p>指定 VMFS 文件系统或文件。此参数是一个绝对路径或相对路径，用于指定 /vmfs 下面的某个目录符号链接、裸设备映射或文件。</p> <ul style="list-style-type: none"><li>■ 要指定 VMFS 文件系统，请使用以下格式：</li></ul> <div>/vmfs/volumes/file_system_UUID</div> <p>或</p> <div>/vmfs/volumes/file_system_label</div> <ul style="list-style-type: none"><li>■ 要指定 VMFS 数据存储上的某个文件，请使用以下格式：</li></ul> <div>/vmfs/volumes/file_system_label file_system_UUID/[dir]/myDisk.vmdk</div> <p>如果当前工作目录是 myDisk.vmdk 的父目录，请勿输入完整路径。</p>

## vmkfstools 命令选项

vmkfstools 命令具有多个选项。有些选项仅建议高级用户使用。

选项的长格式和单字母格式是等效的。例如，以下命令完全相同。

```
vmkfstools --createfs vmfs6 --blocksize 1m disk_ID:P
vmkfstools -C vmfs6 -b 1m disk_ID:P
```

### -v 子选项

-v 子选项表示命令输出的详细级别。

该子选项的格式如下：

```
-v --verbose number
```

可以指定 *number* 的值，范围是从 1 到 10 的整数。

使用任何 vmkfstools 选项都可以指定 -v 子选项。如果选项的输出不适合于 -v 子选项，则 vmkfstools 将忽略 -v。

**注** 由于可以将 -v 子选项包含在任何 vmkfstools 命令行中，因此 -v 不作为子选项纳入选项描述中。

## 文件系统选项

文件系统选项可用于创建和管理 VMFS 数据存储。这些选项不适用于 NFS。这些任务中有许多是可以通过 vSphere Client 执行的。

### 列出 VMFS 数据存储的属性

使用 `vmkfstools` 命令可以列出 VMFS 数据存储的属性。

```
-P|--queryfs
    -h|--humanreadable
```

当此选项用于任何驻留在 VMFS 数据存储上的文件或目录时，它将列出指定数据存储的属性。列出的属性一般包括文件系统标签、数据存储的数据区数量、UUID 以及每个数据区驻留的设备列表。

**注** 如果任何设备的后备 VMFS 文件系统脱机，则数据区的数量以及可用的空间也将相应更改。

可以在使用 `-P` 选项时指定 `-h|--humanreadable` 子选项。如果这样，则 `vmkfstools` 将以可读性更强的形式列出卷容量。

#### 示例：列出 VMFS 属性的示例

```
~ vmkfstools -P -h /vmfs/volumes/my_vmfs
VMFS-5.81 (Raw Major Version: 14) file system spanning 1 partitions.
File system label (if any): my_vmfs
Mode: public
Capacity 99.8 GB, 97.5 GB available, file block size 1 MB, max supported file size 62.9 TB
UUID: 571fe2fb-ec4b8d6c-d375-XXXXXXXXXXXX
Partitions spanned (on "lvm"):
    eui.3863316131XXXXXX:1
Is Native Snapshot Capable: YES
```

### 创建 VMFS 数据存储或暂存分区

使用 `vmkfstools` 命令创建 VMFS 数据存储或暂存分区。

```
-C|--createfs [vmfs5|vmfs6|vfat]
```

此选项将在指定的 SCSI 分区（例如 `disk_ID:P`）上创建 VMFS 数据存储。该分区将成为数据存储的主分区。对于 VMFS5 和 VMFS6，可用的唯一块大小为 1 MB。

可以与 `-c` 选项一同指定以下子选项。

- `-S|--setfsname` - 为创建的 VMFS 数据存储定义卷标。此子选项只与 `-c` 选项结合使用。指定的卷标最多为 128 个字符，并且在开头和结尾不能包含空格。

**注** 对于所有实体，vCenter Server 都支持 80 个字符的限制。如果某个数据存储名称超过了该限制，则将该数据存储添加到 vCenter Server 中时，名称会缩短。

在定义卷标之后，则可以在为 `vmkfstools` 命令指定 VMFS 数据存储时使用此卷标。卷标将显示在为 `ls -l` 命令生成的列表中，并且作为指向 `/vmfs/volumes` 目录下 VMFS 卷的符号链接。



要更改 VMFS 卷标，请使用 `ln -sf` 命令。可参考以下示例：

```
ln -sf /vmfs/volumes/UUID /vmfs/volumes/datastore
```

*datastore* 是用于 *UUID* VMFS 的新卷标。

---

**注** 如果向 vCenter Server 注册了您的主机，则对 VMFS 卷标所做的任何更改都将被 vCenter Server 覆盖。此操作保证 VMFS 标签在所有 vCenter Server 主机中都是一致的。

---

- `-Y|--unmapGranularity #[bBsSkKmMgGtT]` - 此子选项仅适用于 VMFS6。为取消映射操作定义粒度。默认粒度为 1 MB。与块大小一样，输入单位类型。
- `-O|--unmapPriority <none|low|medium|high>` - 此子选项仅适用于 VMFS6。为取消映射操作定义优先级。

### 示例：创建 VMFS 文件系统的示例

此示例说明如何在 *naa.ID:1* 分区上创建名为 *my\_vmfs* 的 VMFS6 数据存储。

```
~ vmkfstools -C vmfs6 -S my_vmfs /vmfs/devices/disks/naa.ID:1
```

### 将数据区添加到 VMFS 数据存储

使用 `vmkfstools` 命令将数据区添加到 VMFS 数据存储。

添加数据区时，您将主分区的 VMFS 数据存储跨越 *span\_partition* 指定的分区。

```
-Z|--spanfs span_partitionhead_partition
```

您必须指定主分区和跨分区的完整路径名称，例如，`/vmfs/devices/disks/disk_ID:1`。每次使用该选项就会增加一个数据区到 VMFS 数据存储，这样数据存储就会跨越多个分区。

---

**小心** 运行该选项时，您将丢失在 *span\_partition* 中指定的 SCSI 设备上之前存在的所有数据。

---

### 示例：扩展 VMFS 数据存储的示例

在此示例中，您将跨越新的分区扩展 VMFS 数据存储现有的主分区。

```
~ vmkfstools -Z /vmfs/devices/disks/naa.disk_ID_2:1 /vmfs/devices/disks/naa.disk_ID_1:1
```

扩展的数据存储将跨越两个分区，即 *naa.disk\_ID\_1:1* 和 *naa.disk\_ID\_2:1*。在此示例中，*naa.disk\_ID\_1:1* 是主分区的名称。

### 扩展 VMFS 数据存储

无需将数据区添加到 VMFS 数据存储，而是可以增加现有数据存储的大小。请使用 `vmkfstools -G` 命令。

在底层存储增大其自身容量后，您可以增加数据存储大小。

该命令使用以下选项：

```
-G|--growfs devicedevice
```

此选项会扩展 VMFS 数据存储或其特定数据区。例如，

```
vmkfstools --growfs /vmfs/devices/disks/disk_ID:1 /vmfs/devices/disks/disk_ID:1
```

## 虚拟磁盘选项

利用虚拟磁盘选项，您可以设置、迁移和管理存储在数据存储上的虚拟磁盘。其中的大部分任务也可以通过 vSphere Client 执行。

### 支持的磁盘格式

在创建或克隆虚拟磁盘时，可以使用 `-d|--diskformat` 子选项来指定磁盘格式。

您可以选择以下格式：

- **zeroedthick**（默认） - 创建过程中为虚拟磁盘分配所需空间。创建时不会擦除物理设备上保留的任何数据，但是从虚拟机首次执行写入操作时会按需要将其置零。虚拟机不会从磁盘读取失效数据。
- **eagerzeroedthick** - 在创建时为虚拟磁盘分配所需的空間。与 **zeroedthick** 格式相反，在创建过程中会将物理设备上保留的数据置零。创建这种格式的磁盘所需的时间可能会比创建其他类型的磁盘所用时间长。
- **thin** - 精简置备的虚拟磁盘。与 **thick** 格式不同，这种格式在创建时不为虚拟磁盘分配所需空间，而是按需提供空间并置零。
- **rdm:设备** - 虚拟兼容模式裸磁盘映射。
- **rdmp:设备** - 物理兼容模式（直通）裸磁盘映射。
- **2gbsparse** - 最大数据区大小为 2 GB 的稀疏磁盘。您可将该格式磁盘与诸如 VMware Fusion 之类的托管 VMware 产品结合使用。但是，除非首先使用 `vmkfstools` 以兼容格式（如 **thick** 或 **thin**）将磁盘重新导入，否则在 ESXi 主机上稀疏磁盘无法开机。

### NFS 数据存储上的磁盘格式

NFS 仅可使用的磁盘格式为 **thin**、**thick**、**zeroedthick** 和 **2gbsparse**。

通常，**Thick**、**zeroedthick** 和 **thin** 格式的表现相同，因为决定分配策略的是 NFS 服务器而不是 ESXi 主机。大多数 NFS 服务器上的默认分配策略为 **thin**。但是，在支持 Storage APIs - Array Integration 的 NFS 服务器上，可以创建 **zeroedthick** 格式的虚拟磁盘。通过预留空间操作，NFS 服务器可以分配并确保空间。

有关阵列集成 API 的详细信息，请参见 [第 16 章 存储硬件加速](#)。

## 创建虚拟磁盘

使用 `vmkfstools` 命令可以创建虚拟磁盘。

```
-c|--createvirtualdisk size[bB|sS|kK|mM|gG]
-d|--diskformat [thin|zeroedthick|eagerzeroedthick]
-W|--objecttype [file|vsan|vvol]
--policyFile fileName
```

此选项将在数据存储上的指定路径创建虚拟磁盘。指定虚拟磁盘的大小。为 *size* 输入值时，可以加上 k（千字节）、m（兆字节）或 g（千兆字节）等后缀以指明其单位类型。单位类型不区分大小写。

`vmkfstools` 将 k 或 K 解释为千字节。如果不指定单位类型，`vmkfstools` 默认使用字节。

可以与 `-c` 选项一同指定以下子选项。

- `-d|--diskformat` 指定磁盘格式。
- `-W|--objecttype` 指定虚拟磁盘是 VMFS 数据存储上的文件还是 NFS 数据存储上的文件，是 vSAN 数据存储上的对象还是 Virtual Volumes 数据存储上的对象。
- `--policyFile fileName` 指定磁盘的虚拟机存储策略。

### 示例：创建虚拟磁盘的示例

本示例显示如何创建名为 `disk.vmdk` 的 2 千兆字节虚拟磁盘文件。您将在名为 `myVMFS` 的 VMFS 数据存储上创建该磁盘。该磁盘文件表示虚拟机可以访问的空虚拟磁盘。

```
vmkfstools -c 2048m /vmfs/volumes/myVMFS/disk.vmdk
```

## 初始化虚拟磁盘

使用 `vmkfstools` 命令可以初始化虚拟磁盘。

```
-w|--writezeros
```

该选项通过对虚拟磁盘的所有数据清零，从而对其进行清除。完成此命令可能需要花费很长时间，具体取决于您的虚拟磁盘的大小和托管虚拟磁盘的设备的 I/O 带宽。

**小心** 使用此命令时，您将丢失虚拟磁盘上的所有现有数据。

## 扩充精简虚拟磁盘

使用 `vmkfstools` 命令可以扩充精简虚拟磁盘。

```
-j|--inflatedisk
```

该选项将 `thin` 虚拟磁盘转换为 `eagerzeroedthick`，从而保留了所有的现有数据。该选项将分配并置零所有还未分配的块。

## 将 Zeroedthick 虚拟磁盘转换为 Eagerzeroedthick 磁盘

可使用 `vmkfstools` 命令将任何 `zeroedthick` 虚拟磁盘转换为 `eagerzeroedthick` 磁盘。

```
-k|--eagerzero
```

执行转换时，此选项会保留虚拟磁盘上的任何数据。

请按以下示例执行操作：

```
vmkfstools --eagerzero /vmfs/volumes/myVMFS/VMName/disk.vmdk
```

## 移除置零块

使用 `vmkfstools` 命令可以移除置零块。

```
-K|--punchzero
```

该选项会对所有清零块取消分配，只留下那些先前分配且包含有效数据的块。生成的虚拟磁盘为精简格式的虚拟磁盘。

## 删除虚拟磁盘

使用 `vmkfstools` 命令删除 VMFS 卷上指定路径的虚拟磁盘文件。

使用下列选项：

```
-U|--deletevirtualdisk
```

## 重命名虚拟磁盘

使用 `vmkfstools` 命令可对 VMFS 卷上指定路径中的虚拟磁盘文件进行重命名。

您必须指定原始文件名或文件路径 *oldName*，以及新文件名或文件路径 *newName*。

```
-E|--renamevirtualdisk oldName newName
```

## 克隆或转换虚拟磁盘或 RDM

可使用 `vmkfstools` 命令创建您指定的虚拟磁盘或裸磁盘的副本。

非 `root` 用户不能克隆虚拟磁盘或 RDM。您必须指定原始文件名或文件路径 *oldName*，以及新文件名或文件路径 *newName*。

```
-i|--clonevirtualdisk oldName newName  
-d|--diskformat [thin|zeroedthick|eagerzeroedthick|rdm:device|rdmp:device|2gbsparse]  
-W|--objecttype [file|vsan|vvol]  
--policyFile fileName  
-N|--avoidnativeclone
```

可使用以下子选项更改所创建副本的相应参数。

- `-d|--diskformat` 指定磁盘格式。
- `-W|--objecttype` 指定虚拟磁盘是 VMFS 数据存储上的文件还是 NFS 数据存储上的文件，是 vSAN 数据存储上的对象还是 Virtual Volumes 数据存储上的对象。
- `--policyFile fileName` 指定磁盘的虚拟机存储策略。

默认情况下，ESXi 使用其本机方法执行克隆操作。如果您的阵列支持克隆技术，则您可以将操作卸载到阵列。要避免 ESXi 本机克隆，请指定 `-N|--avoidnativeclone` 选项。

### 示例：克隆或转换虚拟磁盘的示例

本示例说明了如何将 templates 存储库中主虚拟磁盘的内容克隆到 myVMFS 文件系统上名为 myOS.vmdk 的虚拟磁盘文件。

```
vmkfstools -i /vmfs/volumes/myVMFS/templates/gold-master.vmdk /vmfs/volumes/myVMFS/myOS.vmdk
```

您可以通过向虚拟机配置文件中添加几行内容，将虚拟机配置为使用此虚拟磁盘，如下示例中所示：

```
scsi0:0.present = TRUE
scsi0:0.fileName = /vmfs/volumes/myVMFS/myOS.vmdk
```

如果您想要转换磁盘的格式，请使用 `-d|--diskformat` 子选项。

当所导入虚拟磁盘的格式与 ESXi 不兼容（例如 2gbsparse 格式）时，此子选项非常有用。转换磁盘后，您可以将此磁盘连接到您在 ESXi 中创建的新虚拟机。

例如：

```
vmkfstools -i /vmfs/volumes/myVMFS/templates/gold-master.vmdk /vmfs/volumes/myVMFS/myOS.vmdk
-d thin
```

## 扩展虚拟磁盘

创建虚拟机之后，可以使用 `vmkfstools` 命令扩展分配给该虚拟机的磁盘大小。

```
-X|--extendvirtualdisk newSize[bBsSkKmMgGtT]
```

指定 `newSize` 参数添加相应的单位后缀。单位类型不区分大小写。vmkfstools 将 k 或 K 解释为千字节。如果不指定单位类型，vmkfstools 将默认为千字节。

上述 `newSize` 参数将重新定义整个磁盘的大小，而不是定义给磁盘增加的大小。

例如，要给 4 G 的虚拟磁盘扩展 1 G，则输入：`vmkfstools -X 5g disk name`。

通过使用 `-d eagerzeroedthick` 选项，可将虚拟磁盘扩展为 `eagerzeroedthick` 格式。

使用 `-x` 选项时，需要考虑以下注意事项：

- 请勿对具有相关快照的虚拟机的基础磁盘进行扩展。否则，您再也不能提交快照或将基础磁盘转换回原始大小。

- 扩展磁盘之后，您可能需要更新磁盘上的文件系统。这样，客户机操作系统才会识别磁盘的新大小并可以使用它。

## 升级虚拟磁盘

此选项可将指定的虚拟磁盘文件从 ESX Server 2 格式转换为 ESXi 格式。

使用此选项可转换 LEGACYSPARSE、LEGACYPLAIN、LEGACYVMFS、LEGACYVMFS\_SPARSE 和 LEGACYVMFS\_RDM 类型的虚拟磁盘。

```
-M|--migratevirtualdisk
```

## 创建虚拟兼容模式裸设备映射

可使用 `vmkfstools` 命令在 VMFS 卷上创建裸设备映射 (RDM) 文件并将裸 LUN 映射到此文件。建立此映射之后，您可以像访问常规 VMFS 虚拟磁盘那样访问 LUN。映射的文件长度与它指向的裸 LUN 的大小相同。

```
-r|--createrdm device
```

指定 *device* 参数时，请使用以下格式：

```
/vmfs/devices/disks/disk_ID:P
```

### 示例：创建虚拟兼容模式 RDM 的示例

在此示例中，您可以创建名为 *my\_rdm.vmdk* 的 RDM 文件，并将 *disk\_ID* 裸磁盘映射到该文件。

```
vmkfstools -r /vmfs/devices/disks/disk_ID my_rdm.vmdk
```

您可以通过在虚拟机配置文件中添加以下行，将虚拟机配置为使用 *my\_rdm.vmdk* 映射文件：

```
scsi0:0.present = TRUE
scsi0:0.fileName = /vmfs/volumes/myVMFS/my_rdm.vmdk
```

## 创建物理兼容模式裸设备映射

可使用 `vmkfstools` 命令将直通裸设备映射到 VMFS 卷上的文件。通过此映射，虚拟机可以在访问虚拟磁盘时绕过 ESXi SCSI 命令筛选。当虚拟机必须发送专有 SCSI 命令时，例如当 SAN 感知软件在虚拟机上运行时，此类型的映射非常有用。

```
-z|--createrdmpassthru deviceexample.vmdk
```

建立此类型的映射后，您可以使用它像访问任何其他 VMFS 虚拟磁盘那样访问裸磁盘。

指定 *device* 路径时，请使用以下格式：

```
/vmfs/devices/disks/device_ID
```

对于.vmdk 名称, 请使用以下格式。使用此命令前, 请务必创建数据存储。

```
/vmfs/volumes/datastore_name/example.vmdk
```

例如,

```
vmkfstools -z /vmfs/devices/disks/naa.600a000000000000... /vmfs/volumes/datastore1/  
mydisk.vmdk
```

## 列出 RDM 的属性

使用 vmkfstools 命令列出裸磁盘映射的属性。这些属性帮助标识 RDM 文件要映射到的存储设备对象。

```
-q|--queryrdm my_rdm.vmdk
```

该选项可打印裸磁盘 RDM 的名称。该选项也可打印裸磁盘的其他标识信息, 例如, 磁盘 ID。

### 示例: 列出 RDM 属性的示例

```
# vmkfstools -q /vmfs/volumes/VMFS/my_vm/my_rdm.vmdk  
  
Disk /vmfs/volumes/VMFS/my_vm/my_rdm.vmdk is a Passthrough Raw Device Mapping  
  
Maps to: vml.02000000006005076801900207700000000000005323134352020
```

## 显示虚拟磁盘几何形状

使用 vmkfstools 命令获取有关虚拟磁盘几何形状的信息。

```
-g|--geometry
```

输出的格式为: Geometry information C/H/S, 其中 C 表示柱面数量, H 表示磁头数量, S 表示扇区数量。

---

**注** 将虚拟磁盘从托管的 VMware 产品导入到 ESXi 主机时, 您可能会看到磁盘几何形状不匹配的错误消息。磁盘几何形状不匹配也会在您加载客户机操作系统或运行新创建的虚拟机时触发问题。

---

## 检查和修复虚拟磁盘

如果虚拟磁盘损坏, 可使用 vmkfstools 命令检查或修复。

```
-x|--fix [check|repair]
```

例如,

```
vmkfstools -x check /vmfs/volumes/my_datastore/my_disk.vmdk
```

## 对磁盘链执行一致性检查

可使用 `vmkfstools` 命令检查整个快照链。您可以确定链中的任何链接是否损坏或者是否存在任何无效的父子关系。

```
-e|--chainConsistent
```

## 存储设备选项

可以使用 `vmkfstools` 命令的设备选项执行物理存储设备的管理任务。

### 管理 LUN 的 SCSI 预留

使用 `vmkfstools` 命令可为 ESXi 主机预留专用的 SCSI LUN。您也可以释放预留，以便其他主机可以访问该 LUN，还可以重置预留，强制释放目标主机上的所有预留。

```
-L|--lock [reserve|release|lunreset|targetreset|busreset|readkeys|readresv] device
```

**小心** 使用 `-L` 选项会中断 SAN 上其他服务器的操作。仅在排除集群设置故障时使用 `-L` 选项。

除非 VMware 建议您在托管 VMFS 卷的 LUN 上使用该选项，否则请勿使用。

可以使用多种方式指定 `-L` 选项：

- `-L reserve` - 预留指定的 LUN。预留之后，只有预留该 LUN 的服务器可以访问它。如果其他服务器尝试访问该 LUN，则会显示预留错误。
- `-L release` - 释放对指定 LUN 的预留。其他服务器可以再次访问该 LUN。
- `-L lunreset` - 通过清除对指定 LUN 的任何预留并将其标记为所有服务器可再次访问，可以重置该 LUN。此重置操作不会影响设备上的任何其他 LUN。如果预留了设备上的另一个 LUN，它会保持预留状态。
- `-L targetreset` - 重置整个目标。此重置操作将对与该目标关联的所有 LUN 清除所有预留，让所有服务器可再次访问这些 LUN。
- `-L busreset` - 重置总线上的所有可访问目标。此重置操作将对可以通过总线访问的所有 LUN 清除任何预留，让所有服务器可以再次访问它。
- `-L readkeys` - 读取向 LUN 注册的预留密钥。适用于 SCSI-III 持久组预留功能。
- `-L readresv` - 读取 LUN 上的预留状态。适用于 SCSI-III 持久组预留功能。

输入 *device* 参数时，请使用以下格式：

```
/vmfs/devices/disks/disk_ID:P
```

## 解除设备锁定

可使用 `vmkfstools` 命令解除特定分区上的设备锁定。

```
-B|--breaklock device
```



输入 *device* 参数时，请使用以下格式：

```
/vmfs/devices/disks/disk_ID:P
```

您可以在主机执行某项数据存储操作（例如扩展数据存储、添加数据区或重新签名）失败时使用此命令。运行此命令时，请确保其他主机未持有锁定。