云计算与虚拟化技术

第06章：Storage Devices

https://internet.hactcm.edu.cn

河南中医药大学信息技术学院（智能医疗行业学院）智能医疗教研室
河南中医药大学医疗健康信息工程技术研究所

2025年2月

讨论提纲

存储基础知识

- 企业级存储、本地存储和共享存储、存储阵列
- 存储性能：RAID、去重、复制、物理存储设备类型、SSDs 和 AFAs

vSphere Storage 基本应用

- vSphere 支持的存储
- ESXi 支持的物理存储和逻辑存储
- VM 支持的物理存储和逻辑存储
- vSphere 存储的配置：FC、FCoE、iSCSI、NFS

vSphere Storage 高级功能

- SIOC 和 storage DRS、数据存储集群、VMFS 6
- 高级存储功能：自动空间回收、即时和链接克隆、裸设备映射、多路径等

vSAN

- 规划与设计 vSAN 方案
- vSAN 应用：配置、健康检测、策略

1. 存储基础知识

- 不同类型的存储具有不同协议、体系结构、扩展功能和用途。
- 虚拟化环境下能够满足性能、可弹性扩展、可靠的存储解决方案，需要使用企业级存储产品实现。
- 企业级存储（Enterprise-class storage）的分类：
  - Direct-Attached Storage (DAS)
  - Network Attached Storage (NAS)
  - Storage Area Network (SAN)

  } 可以运行VMs

  - Object-based storage/cloud storage

  不可以运行VMs，可用于备份等

1. 存储基础知识

- 从架构角度：企业存储有三种扩展解决方案
  - 规模内扩展或规模上扩展：Scale-in or scale-up
    - 通过添加新的磁盘架来增加存储容量，但性能没有改变。
    - 适用于中小规模存储需求，在不引入新的存储系统的情况下增加容量和性能。
  - 规模外扩展：Scale-out
    - 将更多的存储节点添加到存储集群中，形成一个更大的存储池。
    - 通过添加新的存储阵列或节点来扩展存储容量和性能，使存储系统具有更好的横向扩展性。
    - 通常与分布式存储系统或软件定义存储（SDS）解决方案结合使用。
  - 混合扩展：
    - 将规模内扩展和规模外扩展结合起来，根据实际需求采取不同的扩展策略。
    - 在已有的存储系统上进行硬件升级，在需要时添加新存储节点以实现更大规模的扩展。

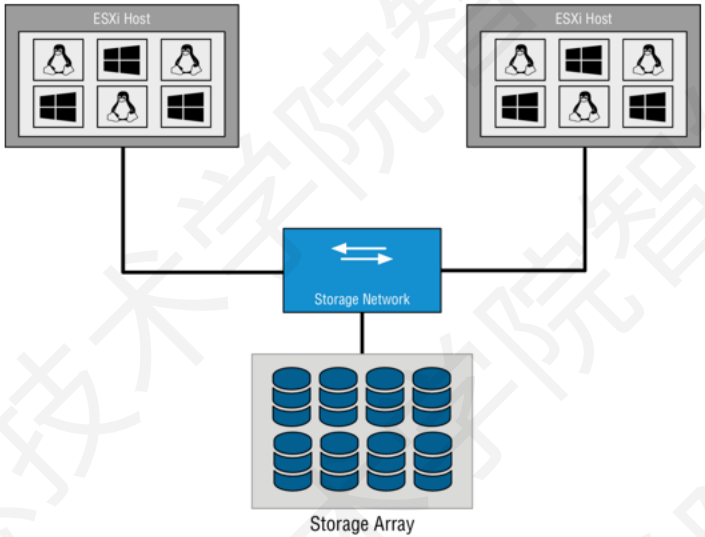
1. 存储基础知识

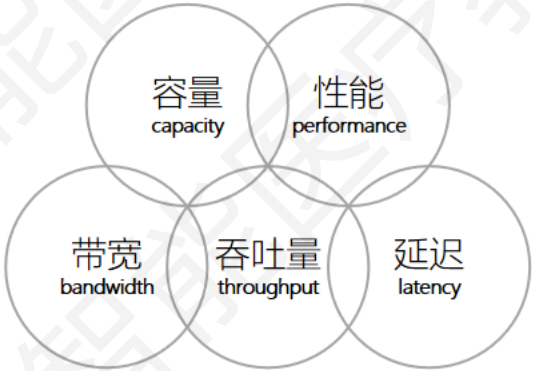
- 从性能角度：企业级存储通常分为4个等级
  - 第 0 层：
    - 非常高性能的存储，例如全闪存阵列 （AFA），提供无与伦比的 I/O 性能。
    - 随着企业级 SSD 价格的下降，第 0 层存储越来越常见。
  - 第 1 层或主存储：
    - 通常是主存储，应用于 vShpere 中。
  - 第 2 层或辅助存储：
    - 通常不应用于生产环境中，主要存储文档、备份、冷数据等。
  - 第 3 层：
    - 长期和离线归档的存储，例如磁带或公有云存储上的备份副本。

1. 存储基础知识

- 存储是虚拟基础设施中最关键的部分。
  - 存储设计方案对于数据中心非常重要。
  - 存储为整个集群及其所有 VM 提供支撑，并直接决定性能。
  - 数据中心存储由本地存储和共享存储两大类。
  - 共享存储的重要性体现在三个方面：
    - 先进功能的支持：
      - vSphere 高可用性（HA）、分布式资源调度器（DRS）、容错（FT）以及 VMware Site Recovery Manager 的部分功能都依赖共享存储。
    - 性能：
      - 虚拟机（VM）和整个 vSphere 集群的整体性能都取决于共享存储。
    - 可用性：
      - 虚拟化整体可用性，以及运行的 VM，取决于共享存储基础结构。

1. 存储基础知识


ESXi Host
ESXi Host
Storage Network
Storage Array


容量
capacity
性能
performance
带宽
bandwidth
吞吐量
throughput
延迟
latency

1. 存储基础知识

存储阵列 Storage arrays

华为存储基础课程
https://www.bilibili.com/video/BV1Zt4y1a7Rj

1. 存储基础知识

AFAs
flash memories

Hybrid
Array
flash and HDD

1. 存储基础知识

Concerning the protocols used for frontend interfaces

前端接口面板

将磁盘阵列接入到虚拟化服务器或者存储网络中

1. 存储基础知识

VMware vSphere 支持的存储接口及应用案例

Protocol type	Type of service	Interface speed	Typical usage
SAS	Block	6 or 12 Gbps	Shared storage with limited host scaling
FC	Block	8, 16, 32 Gbps	Shared storage, typically for enterprises
FCoE	Block	10, 25, 40, 50, 100 Gbps	Shared storage, typically for enterprises and mid-sized businesses
iSCSI	Block	1, 10, 25, 40, 50, 100 Gbps	Shared storage
NFS	File	1, 10, 25, 40, 50, 100 Gbps	Shared storage

1. 存储基础知识

- 存储性能最主要指标是：IOPS
  - Read and Write I/O per second
  - IOPS（Input/Output Operations Per Second）是衡量存储系统性能的一种指标，表示系统每秒钟能够执行的输入/输出操作数量。
    - 通常用于评估存储设备（如硬盘驱动器、固态硬盘或存储阵列）的性能。
    - 较高的 IOPS 值通常表示存储系统具有更快的数据访问速度和更好的性能表现。
    - IOPS 是评估存储系统性能和选择适当存储解决方案时的重要指标之一。
  - 影响存储性能的因素有很多，例如：
    - 存储使用的物理磁盘或 SSD
    - 存储控制器的成熟度
    - 物理介质
    - 使用的协议和 IO 大小

1. 存储基础知识

影响存储性能的因素

Physical disk or SSDs

physical media

the maturity of the storage controller

the IO size

the protocols used

最重要的影响因素:

- The ***Redundant Array of Independent Disks*** (RAID) level of the logical volume presented to the ESXi servers
- The device type that is used to form a logical volume (逻辑层设备类型)

1. 存储基础知识

- RAID （Redundant Array of Independent Disks）
- 独立磁盘冗余阵列，通常简称为磁盘阵列。
  - RAID 是由多个独立的高性能磁盘驱动器组成的磁盘子系统，提供比单个磁盘更高的存储性能和数据冗余的技术。
  - RAID 是一种多磁盘管理技术，其向主机环境提供了成本适中、数据可靠性高的高性能存储。
  - RAID 的初衷是为大型服务器提供高端的存储功能和冗余的数据安全。
  - RAID 的两个关键目标是提高数据可靠性和 I/O 性能。
  - RAID 的三个关键概念和技术：
    - 镜像（Mirroring）
    - 数据条带（Data Stripping）
    - 数据校验（Data parity）

1. 存储基础知识

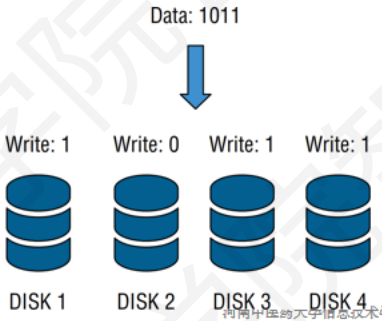
- RAID level
  - RAID使用多个物理存储设备形成逻辑卷，使用逻辑卷有两个原因：
    - 冗余：在一个逻辑卷中有多个设备（HDD 或 SSD），基于 RAID 可以在一个或多个物理存储设备故障时不影响逻辑卷的可用性。
    - 性能：RAID 形成的逻辑卷比物理存储设备具有更好的性能和容量。

- 最常用的RAID等级：

RAID 技术全解（技术文章）
https://cloud.tencent.com/developer/article/2304179

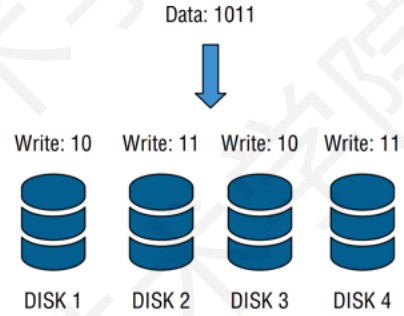
1. 存储基础知识

- RAID level
  - RAID 0:
    - 又称Stripe或Striping，条带化。
    - 将两块以上的硬盘合并成一块，数据连续地分割在每块盘上。
    - 因为带宽加倍，所以读/写速度加倍， 但RAID 0在提高性能的同时，并没有提供数据保护功能，只要任何一块硬盘损坏就会丢失所有数据。



1. 存储基础知识

- RAID level
  - RAID 1 , 1+0 , 0+1 :
    - RAID 1是将一个两块硬盘所构成RAID磁盘阵列，其容量仅等于一块硬盘的容量，因为另一块只是当作数据“镜像”。
    - RAID 1是最可靠的一种阵列，总是保持一份完整的数据备份。性能不如RAID 0但高于单一硬盘。RAID 1写入速度较慢，因为数据得分别写入两块硬盘中并做比较。

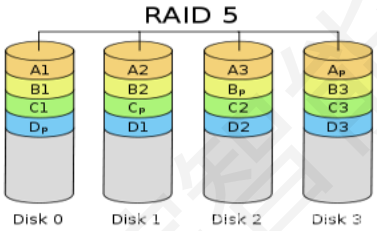


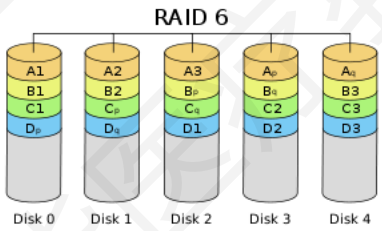
RAID 01 为 RAID 0+1：
先进行条带存放（RAID 0），再进行镜像（RAID 1）

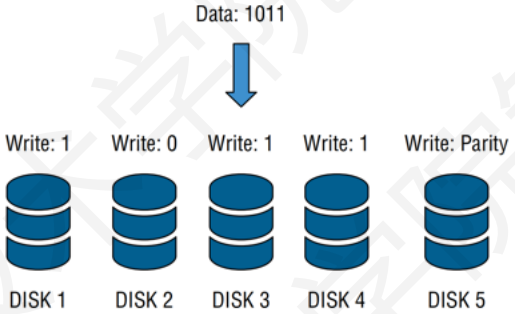
RAID 10 为 RAID 1+0：
先进行镜像（RAID 1），再进行条带存放（RAID 0）

1. 存储基础知识

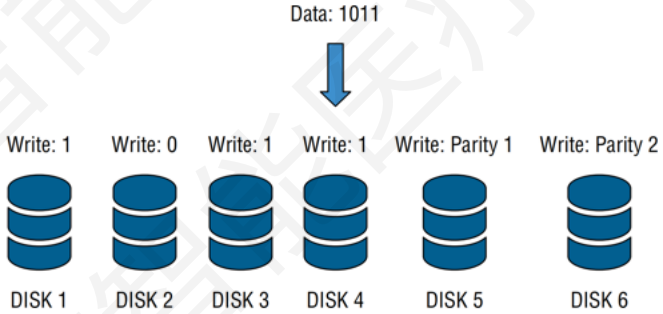
- RAID level
  - RAID 5 , 6:
    - RAID 5 是最常见的RAID 等级。
    - RAID 5 磁盘上同时存储数据和校验数据，数据块和对应的校验信息保存在不同的磁盘上，当一个数据盘损坏时，系统可以根据同一条带的其他数据块和对应校验数据来重建损坏的数据。
    - RAID 5 具备良好扩展性，当磁盘数量增加时，并行操作量的能力也随之增长。
    - RAID 5 兼顾存储性能、数据安全和存储成本等各方面因素，可简单理解为 RAID 0 和 RAID 1 的折中方案，是目前综合性能最佳的数据保护解决方案。推荐数据中心使用RAID 5。
    - RAID 6 引入双重校验的概念，同时出现两个磁盘失效时，存储仍能够正常工作，不会发生数据丢失。
    - RAID 6 是在 RAID 5 的基础上为了进一步增强数据保护而设计的一种 RAID 方式，可以看作是一种扩展的 RAID 5 等级。







RAID 5
4+1



RAID 6
4+2

1. 存储基础知识

- 数据去重：Deduplication
  - 是一种数据存储优化技术，旨在消除数据中重复内容，减少存储空间需求。
  - 通过去除重复的数据块或文件副本，存储系统可以在物理上只保留一份数据副本，而在逻辑上仍然使得所有需要访问该数据的应用程序都能够正常工作。
  - 这种技术对于大型数据集、备份和存档等场景尤其有用，可以显著节省存储空间并提高存储效率。
  - 数据去重有两种方式：
    - In-line：数据在写入其他存储时进行重复数据删除。将数据写入持久缓存，计算哈希值并与存储的块进行比较，如果尚未存储，再将数据写入存储本身。
    - Off-line：数据按原样写入存储，并按定义的时间周期调用重复数据删除任务以删除重复数据。

1. 存储基础知识

- 复制：Replication
  - 复制通常是中端或高端存储阵列软件堆栈的一部分，用于灾难恢复。
  - 复制允许透明地在两个（有时甚至更多）物理存储阵列之间复制数据。
  - 复制复制有两种类型：
    - 同步，Synchronous：
      - 当 ESXi 服务器向底层存储发出存储命令时，数据首先写入主存储，然后再写入辅助存储。
      - 只有当辅助存储确认数据已写入时，才会将确认发送到 VM。
      - 对于同步复制，需要高性能存储区域网络，因为两个存储之间的互连导致的任何延迟都会影响 VM 的整体延迟。
      - 当主存储发生故障时，不会丢失任何数据。
    - 异步，Asynchronous：
      - 收到存储命令后，会将其写入主存储，并立即向 VM 发送确认。然后，在预定义的复制间隔之后，数据将同步到辅助存储。
      - 如果主存储发生故障，将丢失尚未复制到辅助阵列的数据。

1. 存储基础知识

- 物理存储设备接口也是决定存储性能的重要因素
  - 两种常见的物理存储接口标准：SATA 和 SAS
    - SATA（Serial ATA，串行ATA）：
      - SATA 是一种通用存储接口，主要用于个人计算机、消费级存储设备和一些中小型企业级存储系统。
      - SATA 相对便宜，易于使用和部署，适合于一般性的存储需求，在处理大并发时性能不及 SAS。
    - SAS（Serial Attached SCSI，串行连接 SCSI）：
      - SAS 是一种专业的存储接口，通常用于企业级存储系统和高性能计算环境。
      - SAS 接口提供了更高的带宽和更低的访问延迟，适合处理高吞吐量和低延迟的应用程序。
      - SAS 设备通常具有更高的可靠性和更长的寿命，支持更高的并发连接数和更多的同时访问。
      - SAS 支持诸多高级功能，如多路径访问、热插拔、端口多路复用等，有助于提高可用性和可管理性。
    - 对于对存储性能要求较高的企业级应用，SAS 接口往往是更好的选择，而对于一般性的存储需求，SATA 接口则是一个经济实惠的选择。

1. 存储基础知识

- 使用闪存技术提升存储性能
  - 两种使用闪存技术的存储方案：SSDs 和 AFAs
    - SSDs（固态硬盘）：
      - SSDs 是单个存储设备，使用闪存芯片来存储数据。
      - 通常用于替代传统的硬盘驱动器（HDDs），提供更快的数据访问速度、更低的延迟和更高的吞吐量。
      - 可以在个人电脑、服务器和数据中心中使用，提供更高性能和更低功耗的存储解决方案。
    - AFAs（全闪存阵列）：
      - AFA 是一种专门设计的企业级存储解决方案，完全由固态硬盘组成，旨在提供极高的性能和可靠性。
      - 采用全闪存的架构，以实现更快的数据访问速度、更低的延迟和更高的吞吐量。
      - 适用于需要大量随机 I/O 访问和快速数据处理的应用，如数据库、虚拟化、人工智能和大数据分析。
    - SSDs 提供了单个存储设备的闪存解决方案，而 AFAs 是专门设计的全闪存存储阵列，用于提供高性能、高可靠性的企业级存储解决方案。

2. vSphere Storage 基本应用

VMware vSphere 支持的存储体系结构

技术	协议	传输	接口
光纤通道	FC/SCSI、FC/NVMe	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	■ iSCSI HBA 或启用 iSCSI 的网卡（硬件 iSCSI） ■ 网络适配器（软件 iSCSI）
NAS	IP/NFS	文件（无直接 LUN 访问）	网络适配器

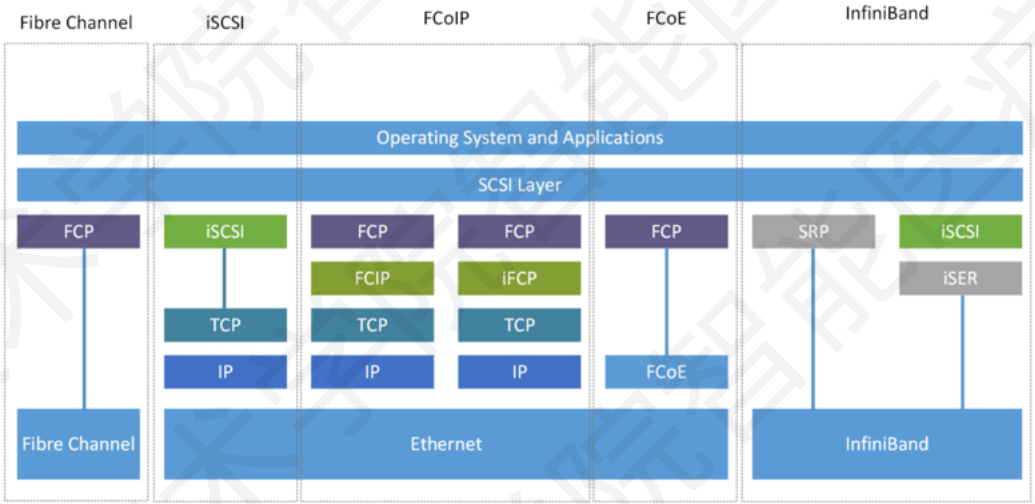
2. vSphere Storage 基本应用

不同存储体系结构支持的 VMware vSphere 特性

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

2. vSphere Storage 基本应用

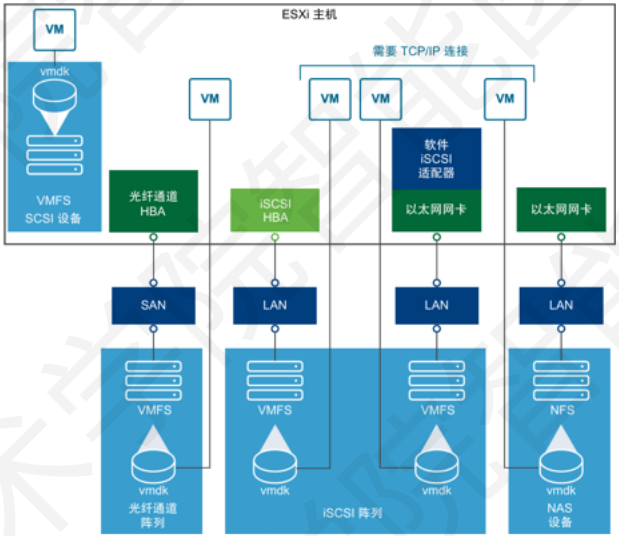
SCSI 命令从 VM Guest OS 到物理网络的流程



FCP：Fibre Channel Protocol，光纤通道协议。

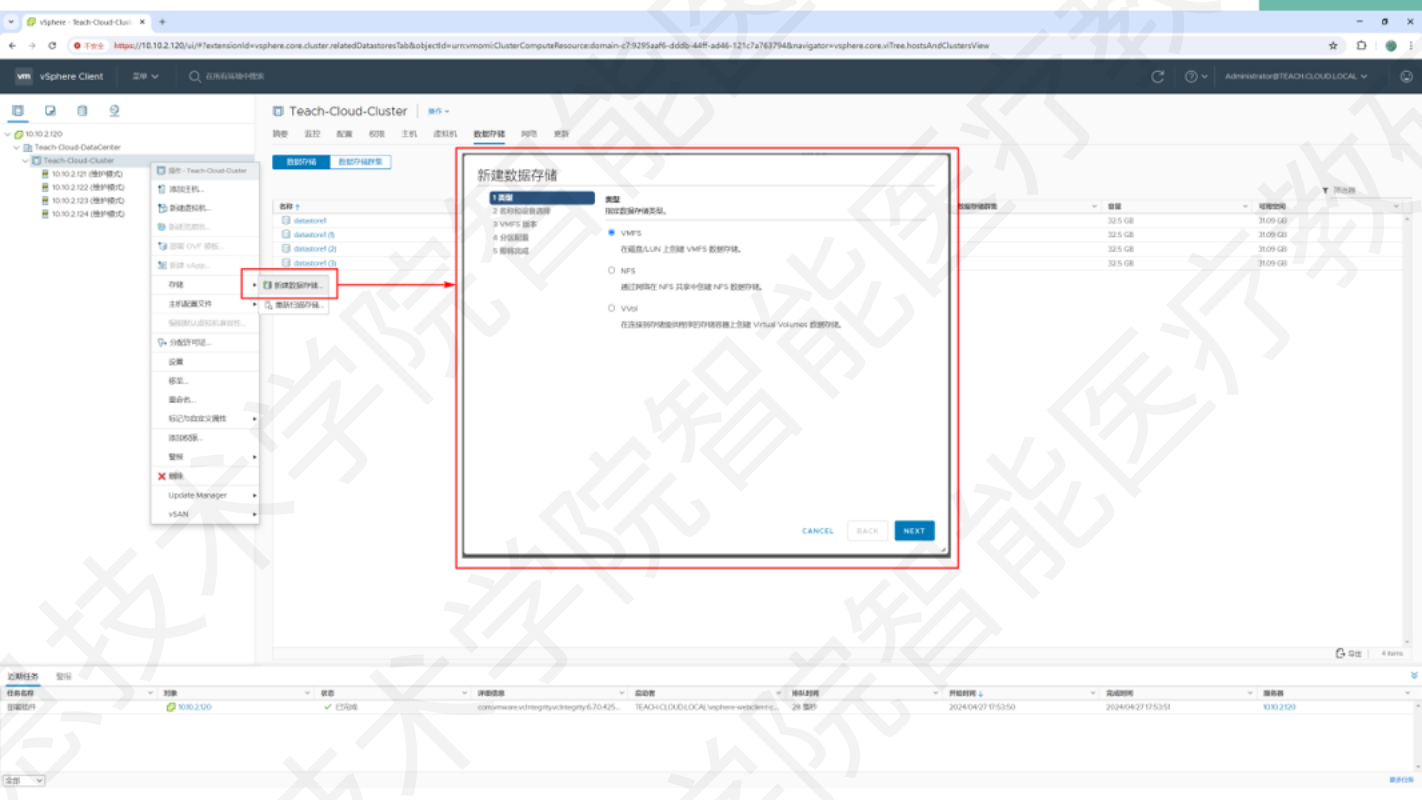
2. vSphere Storage 基本应用

SCSI 命令从 VM Guest OS 到不同体系结构存储的流程



2. vSphere Storage 基本应用

- ESXi 支持的四种数据存储类型
  - VMFS（VMware文件系统）：
    - 专门为虚拟化环境设计的文件系统，用于存储虚拟机的虚拟磁盘文件和其他相关文件。支持多路径访问、高可用性（通过存储 I/O 控制和多路径访问）、支持快照和克隆、以及支持存储 I/O 控制。
  - NFS（网络文件系统）：
    - 基于网络的文件共享协议，允许 ESXi 主机通过网络连接到远程存储设备，并将其挂载为数据存储，通常用于共享存储资源。
  - vVOL（虚拟 Volumes）：
    - 将物理硬件资源抽象到逻辑容量池来虚拟化存 储设备。
  - vSAN（虚拟 SAN）：
    - 软件定义存储解决方案，利用 ESXi 主机上的本地存储资源创建虚拟 SAN，用于存储虚拟机和其他数据。

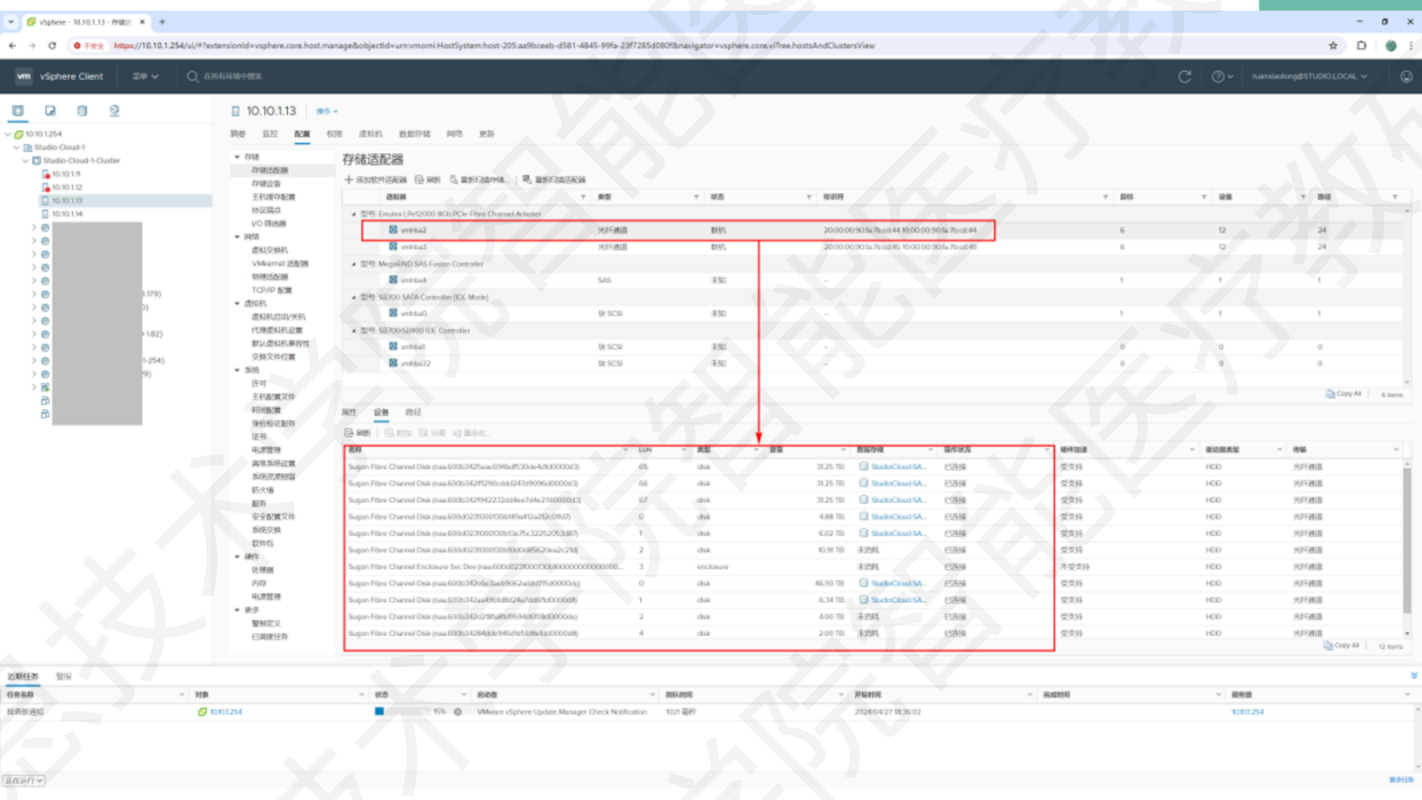

vSphere - Teach-Cloud-Clust
https://10.10.2.120/ui/#?extensionId=vsphere.core.cluster.relatedDatastoresTab&objectId=urn:vmomi:ClusterComputeResource:domain-c7:9295aaf6-dddb-44ff-ad46-121c7a763794&navigator=vsphere.core.viTree.hostsAndClustersView
vSphere Client
菜单
在所有环境中搜索
Administrator@TEACH.CLOUD.LOCAL
Teach-Cloud-Cluster
操作
摘要
监控
配置
权限
主机
虚拟机
数据存储
网络
更新
数据存储
数据存储群集
10.10.2.120
Teach-Cloud-DataCenter
Teach-Cloud-Cluster
10.10.2.121 (维护模式)
10.10.2.122 (维护模式)
10.10.2.123 (维护模式)
10.10.2.124 (维护模式)
操作 - Teach-Cloud-Cluster
添加主机...
新建虚拟机...
部署 OVF 模板...
新建 vApp...
存储
主机配置文件
分配许可证...
设置
移至...
重命名...
标记为自定义属性
添加权限...
警报
删除
Update Manager
vSAN
新建数据存储...
选择已装存储...
新建数据存储
1 类型
2 名称和设备选择
3 VMFS 版本
4 分区配置
5 即将完成
类型
指定数据存储类型。
VMFS
在磁盘/LUN 上创建 VMFS 数据存储。
NFS
通过网络在 NFS 共享中创建 NFS 数据存储。
VVol
在连接到存储提供程序的存储容器上创建 Virtual Volumes 数据存储。
CANCEL
BACK
NEXT
名称
datastore1
datastore1 (1)
datastore1 (2)
datastore1 (3)
容量
31.09 GB
32.5 GB
4 items
近期任务
警报
任务名称
对象
状态
详细信息
启动者
排队时间
开始时间
完成时间
服务器
部署组件
10.10.2.120
已完成
com.vmware.vcIntegrity.vcIntegrity.6.70.425...
TEACH.CLOUD.LOCAL\vsphere-webclient-c...
28 毫秒
2024/04/27 17:53:50
2024/04/27 17:53:51
10.10.2.120

2. vSphere Storage 基本应用

- ESXi 在物理层上通过三种方式访问存储
  - Block-based storage accessed by a hardware adapter:
    - 使用硬件适配器访问块存储
    - 例如：DAS、FC-SAN，需要服务器配置相应的存储卡
  - Block-based storage accessed by a software adapter:
    - 使用软件适配器访问块存储
    - 例如：iSCSI-SAN，需要正确配置网络连接。
  - NFS storage:
    - NFS 存储
    - 需要正确配置 IP 网络，能够访问到 NFS 数据存储。

2. vSphere Storage 基本应用

- 使用物理存储适配器时，ESXi 支持五种存储通讯协议和技术
  - Fibre Channel Host Bus Adapter (FC HBA)：https://www.unicaca.com/info/detail/244.html
  - iSCSI HBA：https://sniansfblog.org/tag/iscsi
    - 专用的PCIe卡，可完全在硬件中实现整个iSCSI堆栈，从而减少主机CPU的负载。
  - CNA adapters for FCoE or iSCSI：https://blog.csdn.net/wuzhimang/article/details/52399064
    - 主要是 10 Gbps（或更高）以太网适配器。
    - 在融合（或专用）网络上提供硬件（或硬件辅助）FCoE 或 iSCSI 功能。
  - RDMA over Converged Ethernet (RoCE)：https://zhuanlan.zhihu.com/p/361740115
    - 通过以太网进行远程直接内存访问 （RDMA）。
  - InfiniBand HCA：https://www.cnblogs.com/D-Tec/p/3157582.html
    - Mellanox Technologies InfiniBand HCA。

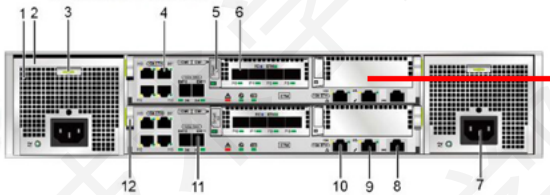

vSphere Client
10.10.1.13
存储适配器
Emulex LPe12000 8Gb PCIe Fibre Channel Adapter
vmhba2
vmhba3
MegaRAID SAS Fusion Controller
vmhba4
SAS
vmhba0
vmhba1
vmhba32
Copy All
Sugon Fibre Channel Disk
Sugon Fibre Channel Enclosure Svc Dev
StudioCloud-SA...
HDD
VMware vSphere Update Manager Check Notification

2U 12 盘位控制框前视图

1	硬盘模块	2	保险箱盘标识
3	信息牌（带ESN号）	4	控制框ID显示器
5	电源指示灯/电源开关按钮	6	硬盘模块卡扣

存储阵列

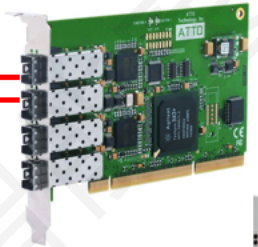
2200 V3 控制框后视图（双控制器）



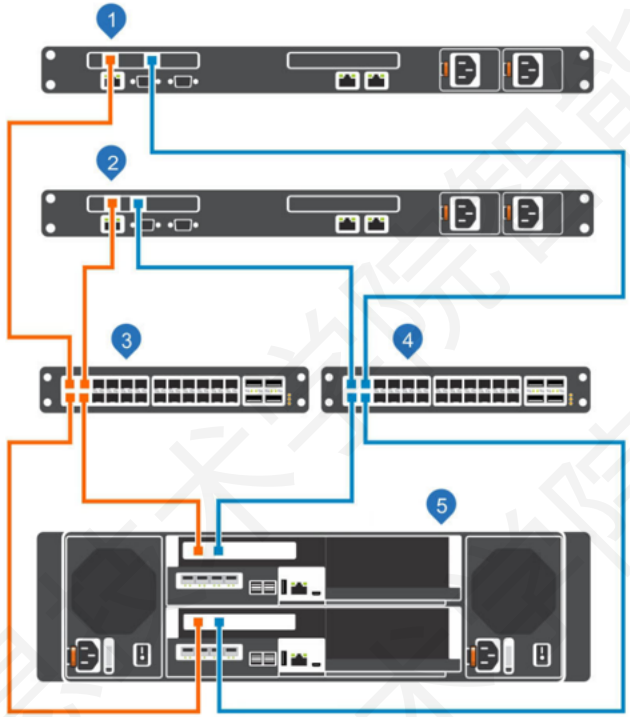
1	电源模块	2	电源模块拉手
3	电源模块卡扣	4	GE电端口
5	接口模块拉手	6	SmartIO端口
7	电源模块插座	8	串口
9	维护网口	10	管理网口
11	mini SAS HD级联端口	12	控制器拉手

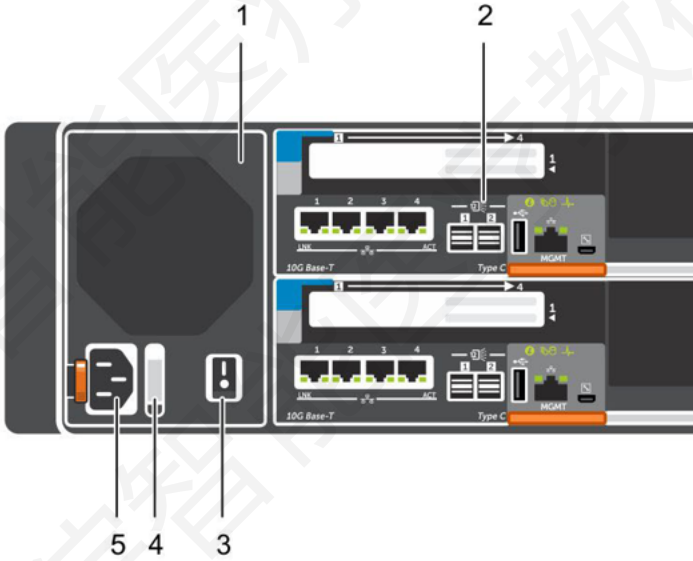
FC 存储交换机

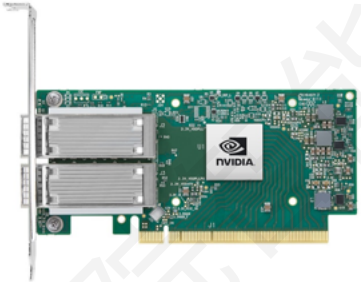
FC HBA（Fibre Channel Host Bus Adapter）



服务器（ESXi 主机）


1
2
3
4
5


1
2
5
4
3
10G Base-T
Type C
MGMT



NVIDIA Quantum-2 Switches

Orderable Part Number (OPN)	Description
MQM9790-NS2F	NVIDIA Quantum-2-based 400Gb/s InfiniBand switch, 64 400Gb/s ports, 32 OSFP ports, non-blocking switching capacity of 51.2Tb/s, two power supplies (AC), standard depth, unmanaged, power-to-connector (P2C) airflow, rail kit
MQM9790-NS2R	NVIDIA Quantum-2-based 400Gb/s InfiniBand switch, 64 400Gb/s ports, 32 OSFP ports, non-blocking switching capacity of 51.2Tb/s, two power supplies (AC), standard depth, unmanaged, connector-to-power (C2P) airflow, rail kit
MQM9700-NS2F	NVIDIA Quantum-2-based 400Gb/s InfiniBand switch, 64 400Gb/s ports, 32 OSFP ports, non-blocking switching capacity of 51.2Tb/s, two power supplies (AC), standard depth, managed, P2C airflow, rail kit
MQM9700-NS2R	NVIDIA Quantum-2-based 400Gb/s InfiniBand switch, 64 400Gb/s ports, 32 OSFP ports, non-blocking switching capacity of 51.2Tb/s, two power supplies (AC), standard depth, managed, C2P airflow, rail kit

ConnectX-7 Adapters

PCIe Standup Adapters

Orderable Part Number (OPN)	Description
MCX755106AC-HEAT	NVIDIA ConnectX-7 HHHL adapter card, 200GbE (default mode) / NDR200 IB, dual-port QSFP112, PCIe 5.0 x16 with x16 PCIe extension option, crypto enabled, Secure Boot enabled, tall bracket
MCX755106AS-HEAT	NVIDIA ConnectX-7 HHHL adapter card, 200GbE (default mode) / NDR200 IB, dual-port QSFP112, PCIe 5.0 x16 with x16 PCIe extension option, crypto disabled, Secure Boot enabled, tall bracket
MCX75310AAC-NEAT	NVIDIA ConnectX-7 HHHL adapter card, 400GbE / NDR IB (default mode), single-port OSFP, PCIe 5.0 x16, crypto enabled, Secure Boot enabled, tall bracket

2. vSphere Storage 基本应用

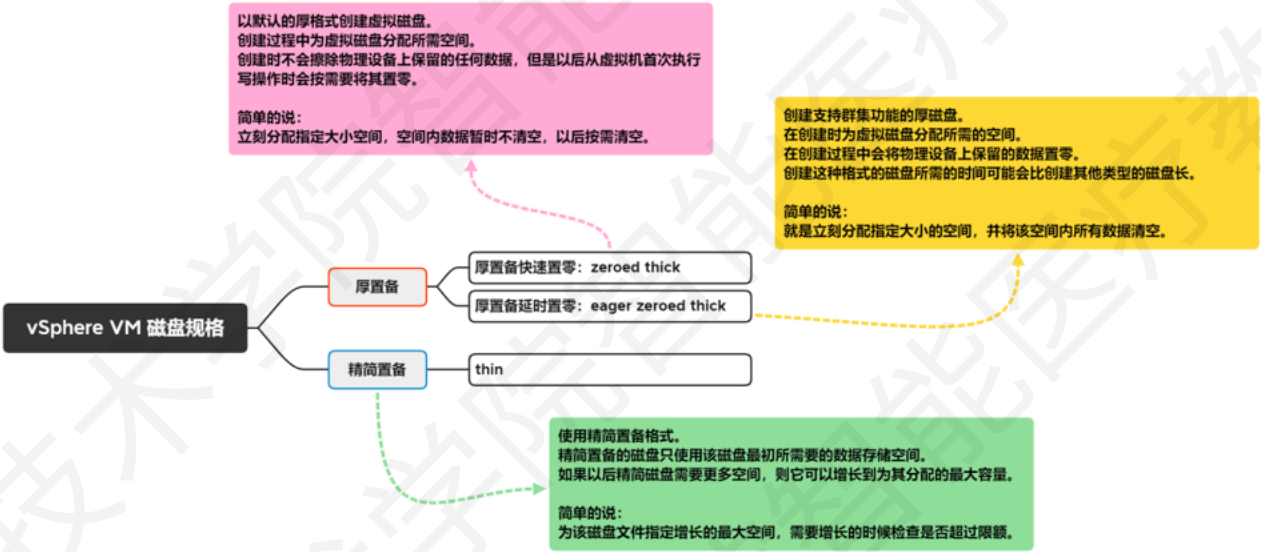
- Thick Provision Lazy Zeroed：厚置备延迟置零
- Thick Provision Eager Zeroed：厚置备快速置零
- Thin Provision：精简置备

2. vSphere Storage 基本应用

- VM 逻辑存储级别（即虚拟磁盘）有四种类型
  - Thick Provision Lazy Zeroed（厚置备延迟置零）：
    - 在创建虚拟磁盘时，会预先分配所需的存储空间，并将磁盘初始化为零值，但只有在虚拟机首次写入数据时才会执行初始化。
    - 提供了最佳的性能，并且避免了创建磁盘时的延迟，但可能会浪费一些存储空间。
  - Thick Provision Eager Zeroed（厚置备快速置零）：
    - 在创建虚拟磁盘时，会预先分配所需的存储空间，并将磁盘立即初始化为零值。
    - 提供了最佳的性能，并且确保磁盘的内容是安全的，但可能会消耗一些额外的时间来初始化磁盘。

2. vSphere Storage 基本应用

- VM 逻辑存储级别（即虚拟磁盘）有四种类型
  - Thin Provision（精简置备）：
    - 在创建虚拟磁盘时，只会预留所需的存储空间，并不会立即分配实际的存储空间。存储空间只有在虚拟机写入数据时才会被动态分配。
    - 可以节省存储空间，并且可以根据需要动态增长，但可能会对性能产生一定的影响，尤其是在存储空间不足时。
  - Raw Device Mapping (RDM)（原始设备映射）：
    - 允许虚拟机直接访问存储阵列上的物理磁盘或逻辑单元（逻辑卷），而不是使用虚拟磁盘文件（VMDK）。
    - RDM 提供了更接近物理硬件的访问方式，并且可以通过兼容性模式或原始模式进行配置，以满足不同的需求。


以默认的厚格式创建虚拟磁盘。
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，但是以后从虚拟机首次执行写操作时会按需要将其置零。
简单的说：
立刻分配指定大小空间，空间内数据暂时不清空，以后按需清空。
创建支持群集功能的厚磁盘。
在创建时为虚拟磁盘分配所需的空间。
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长。
简单的说：
就是立刻分配指定大小的空间，并将该空间内所有数据清空。
vSphere VM 磁盘规格
厚置备
厚置备快速置零：zeroed thick
厚置备延时置零：eager zeroed thick
精简置备
thin
使用精简置备格式。
精简置备的磁盘只使用该磁盘最初所需要的数据存储空间。
如果以后精简磁盘需要更多空间，则它可以增长到为其分配的最大容量。
简单的说：
为该磁盘文件指定增长的最大空间，需要增长的时候检查是否超过限额。

2. vSphere Storage 基本应用

2. vSphere Storage 基本应用

- VM 支持的存储控制器类型
  - BusLogic 并行：
    - 最早在VMware ESX中可用的模拟 SCSI 控制器之一。
    - 主要用于旧版操作系统，不支持大于 2TB 的 VMDK。
  - LSI Logic 并行：
    - 是一个可用的 SCSI 虚拟控制器。
    - 用于操作系统，如 Windows Server 2003。
  - LSI Logic SAS：
    - vSphere 4.0 引入，并且是并行驱动程序的演进。
    - 作为SAS虚拟控制器工作，用于 Windows Server 2008 或更新版本。

2. vSphere Storage 基本应用

- VM 支持的存储控制器类型
  - VMware paravirtual（或PVSCSI）：
    - VMware 准虚拟化，vSphere 4.0 引入的。
    - 是一个 SCSI 虚拟控制器，旨在以最小的处理成本支持非常高的吞吐量，工作不是在模拟模式下，而是在 paravirtual 模式下（需要VMware工具才能识别）。
  - NVMe：
    - 虚拟 NVMe 设备降低了客户端 I/O 处理开销。
    - 允许每个主机承载更多的虚拟机或每分钟执行更多的事务。
    - 每个VM支持四个 NVMe 控制器，每个控制器最多支持 15 个设备。
  - SATA：
    - 通常用于桌面虚拟化和测试环境，不支持高性能和高可用性需求。

Others virtual controllers are also possible in a VM, such as AHCI SATA (introduced in vSphere 5.5), IDE, and also USB controllers, but usually for specific cases (for example, SATA or IDE are usually used for virtual DVD drives).

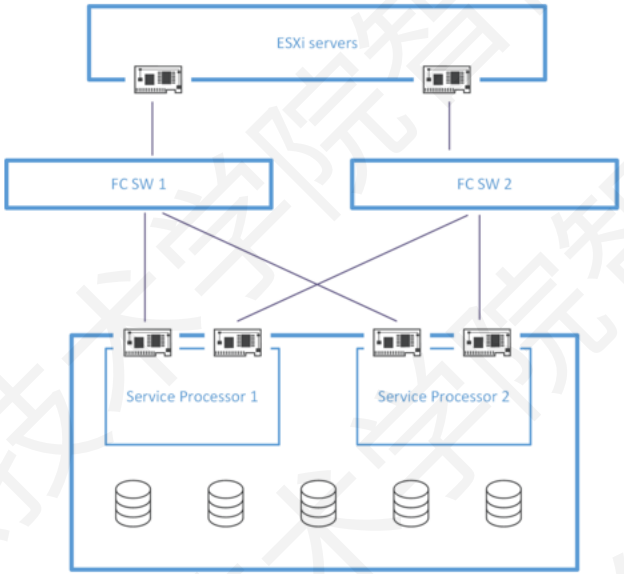
VMware 存储控制器兼容性

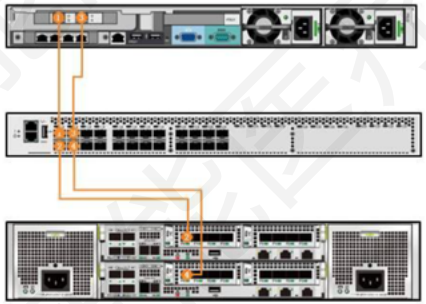
现有控制器	添加的控制器						
	BusLogic 并行	LSI Logic	LSI Logic SAS	VMware 准虚拟 SCSI	AHCI SATA	IDE	NVMe
BusLogic 并行	是	是	是	是	是	是	是
LSI Logic	是	是	是	是	是	是	是
LSI Logic SAS	需要 BIOS 设置	需要 BIOS 设置	通常生效	通常生效	需要 BIOS 设置	是	通常生效
VMware 准虚拟 SCSI	需要 BIOS 设置	需要 BIOS 设置	通常生效	通常生效	需要 BIOS 设置	是	通常生效
AHCI SATA	需要 BIOS 设置	需要 BIOS 设置	是	是	是	是	是
IDE	是	是	是	是	是	不适用	是
NVMe	需要 BIOS 设置	需要 BIOS 设置	通常生效	通常生效	需要 BIOS 设置	是	通常生效

SCSI、SATA 和 NVMe 存储控制器条件、限制和兼容性：

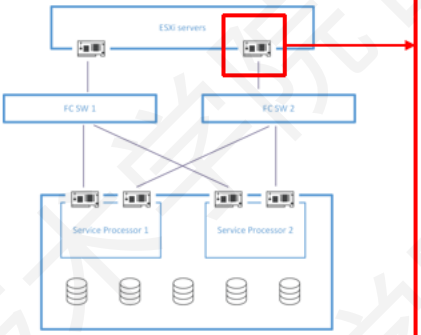
https://docs.vmware.com/cn/VMware-vSphere/7.0/com.vmware.vsphere.vm_admin.doc/GUID-5872D173-A076-42FE-8D0B-9DB0EB0E7362.html

2. vSphere Storage 基本应用

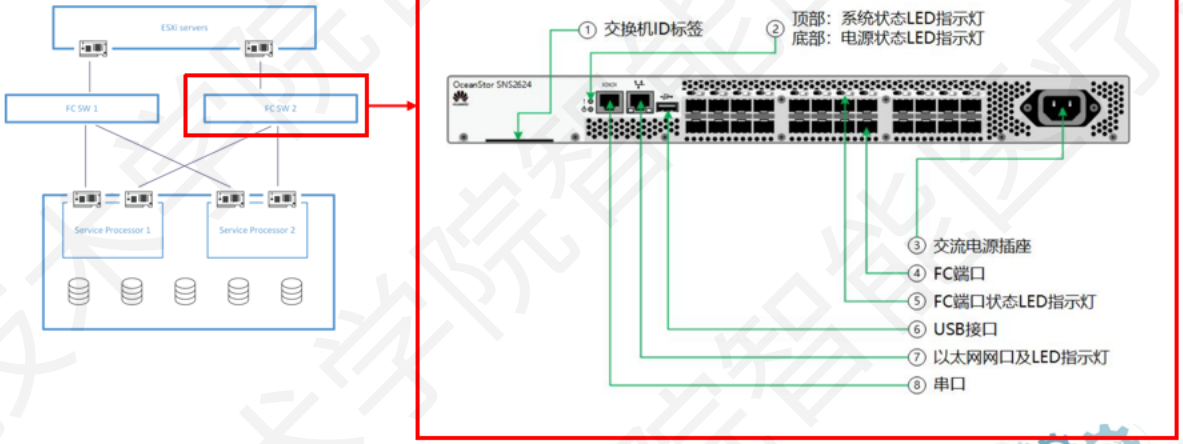




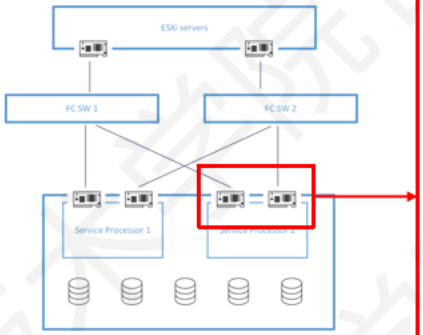
2. vSphere Storage 基本应用



2. vSphere Storage 基本应用



2. vSphere Storage 基本应用



2. vSphere Storage 基本应用

- FCoE：
  - 以太网上封装FC帧的技术。
  - FCoE 网络在第1层和第2层使用 10Gbps（或更高）的以太网网络，但其余仍然是FC协议栈。
    - FC 是一个完整的网络栈，不使用 IP、UDP 或 TCP 协议。
    - FCoE 仍支持 FC 的特性。
  - 如何使用FCoE呢?
    - ESXi Host 需要安装 CAN 卡，使用 NPAR 协议 (Network partition)。
    - 需要使用 FCoE 专用交换机。
    - 存储阵列和通信网络可以不做改变直接接入 FCoE。
    - 在 vSphere 中进行 FCoE 的配置。

2. vSphere Storage 基本应用

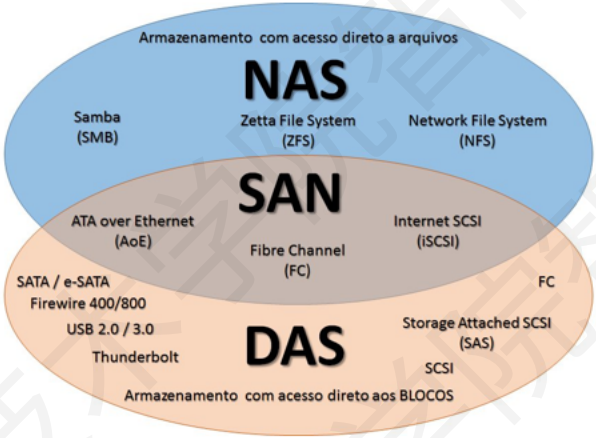
华为CloudEngine系列交换机 FC 和 FCoE 特性介绍
https://e.huawei.com/cn/videolist/video/25b1591c674c4881a4f8ac371a2ac199

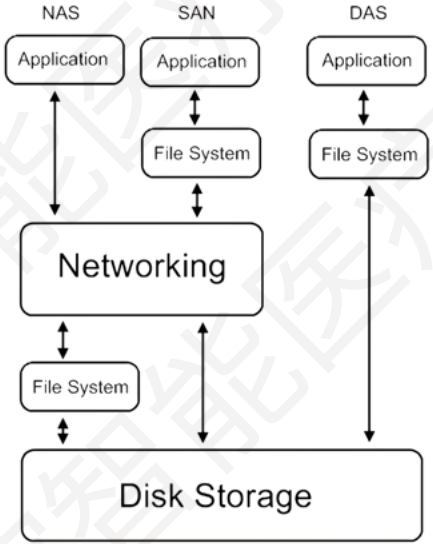
2. vSphere Storage 基本应用

- iSCSI
  - ESXi 支持以下三种 iSCSI
    - Software iSCSI adapter：
      - 利用一个或多个 VMkernel 网络接口和虚拟交换机来管理整个iSCSI流量。
      - 使用软件 iSCSI 适配器，无需购买专用硬件。
    - Dependent hardware iSCSI adapter：
      - 使用硬件的 iSCSI 适配器（iSCSI 卡），但由 VMware vSphere 进行管理和配置。
    - Independent hardware iSCSI adapter or iSCSI HBA：
      - 如同 FC HBA。
      - 所有网络堆栈都是在适配器内部的硬件中实现的。
      - 在 ESXi 端，将看到一个或多个 vmhba，就像其他块存储适配器一样。
      - 必须使用 BIOS 管理或特定工具在卡级别执行网络配置。

2. vSphere Storage 基本应用

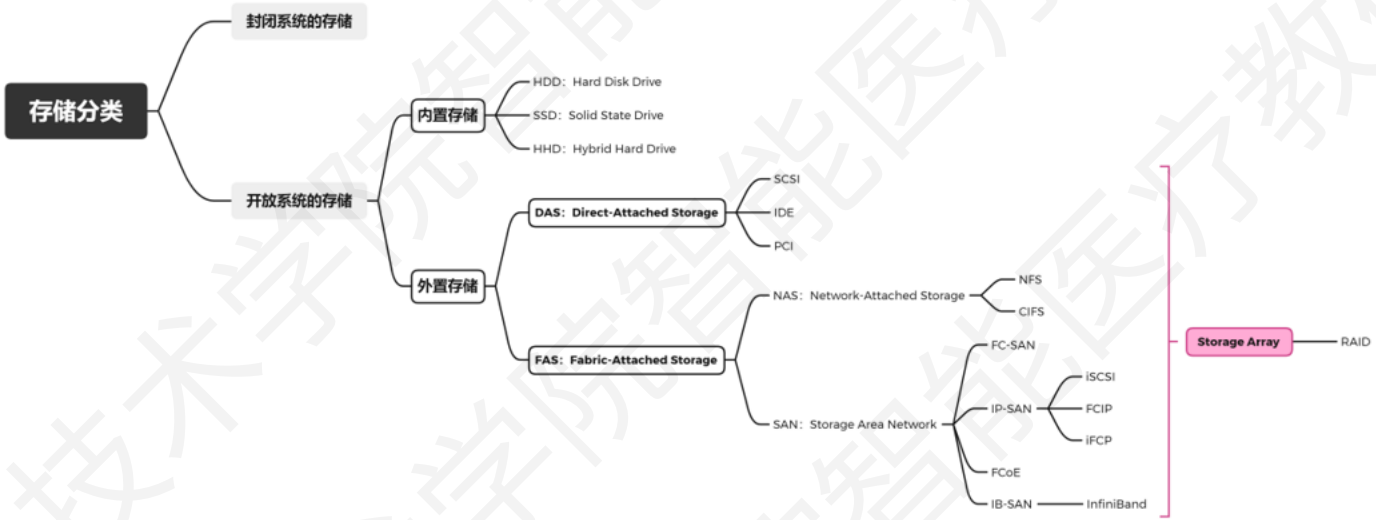
- NFS：
  - ESXi 支持的 NAS 协议是 NFS。
    - 不支持 SMB。
  - 支持NFS协议版本是：
    - NFS 3
    - NFS 4.1（vSphere6.0 及以后版本）


Armazenamento com acesso direto a arquivos
NAS
Samba
(SMB)
Zetta File System
(ZFS)
Network File System
(NFS)
SAN
ATA over Ethernet
(AoE)
Fibre Channel
(FC)
Internet SCSI
(iSCSI)
SATA / e-SATA
Firewire 400/800
USB 2.0 / 3.0
Thunderbolt
FC
Storage Attached SCSI
(SAS)
DAS
SCSI
Armazenamento com acesso direto aos BLOCOS


NAS
SAN
DAS
Application
Application
Application
File System
File System
Networking
File System
Disk Storage

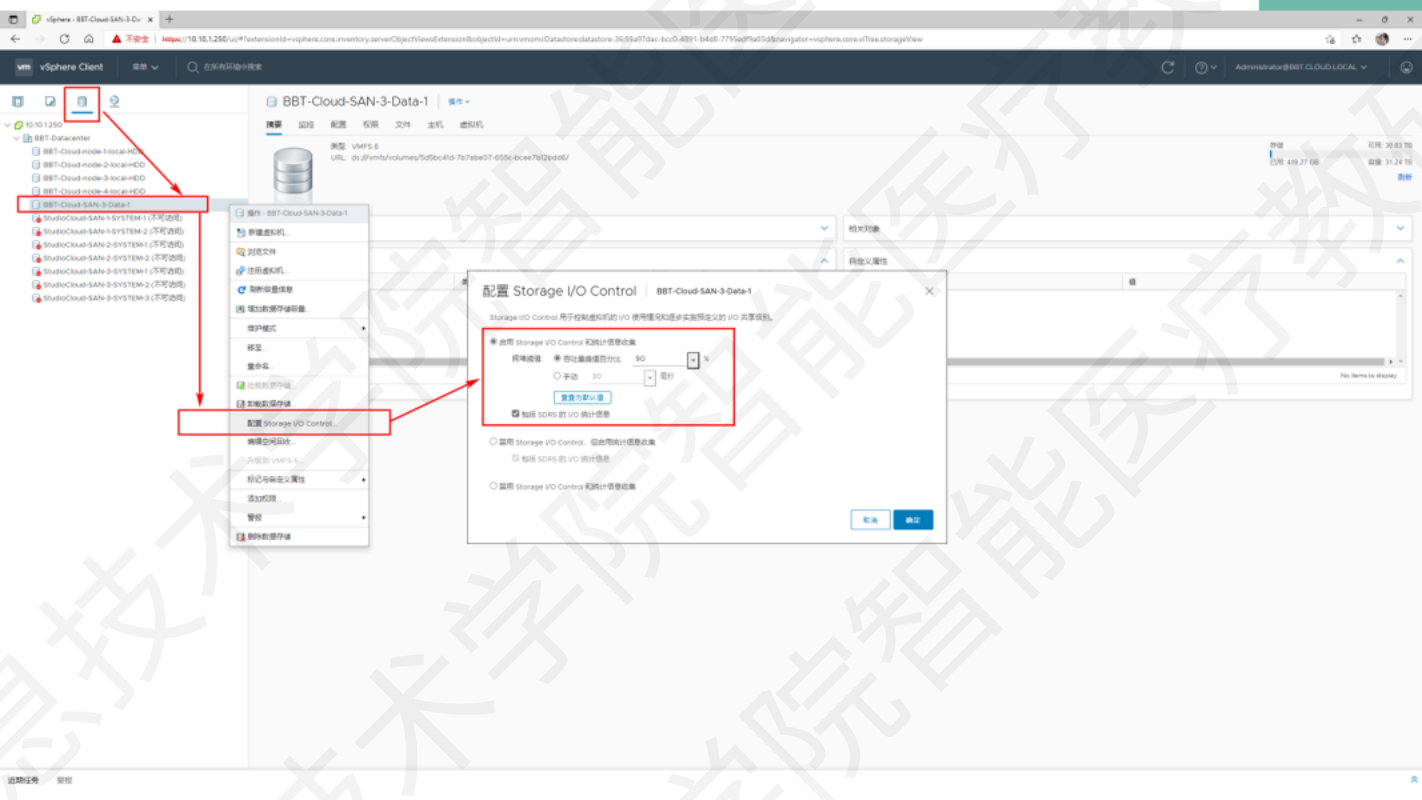
- 存储区域网络 (Storage Area Network, SAN) 是企业最常用的存储网络架构，要求高吞吐量和低延迟的业务关键型应用往往采用这类架构运行。
- SAN 将数据存储在集中式共享存储中，使企业能够运用一致的方法和工具来实施安全防护、数据保护和灾难恢复。
- SAN 是一种基于块的存储，利用高速架构将服务器与其逻辑磁盘单元 (Logical Disk Unit, LUN) 相连。LUN 是一系列通过共享存储池配置的块，以逻辑磁盘的形式呈现给服务器。服务器会对这些块进行分区和格式化，通常使用文件系统，以便可以像在本地磁盘存储上一样在 LUN 上存储数据。
- SAN 约占网络存储市场总额三分之二。SAN 的设计可消除单点故障，具有极高的可用性和故障恢复能力，设计完善的 SAN 可以轻松承受多个组件或设备的故障。

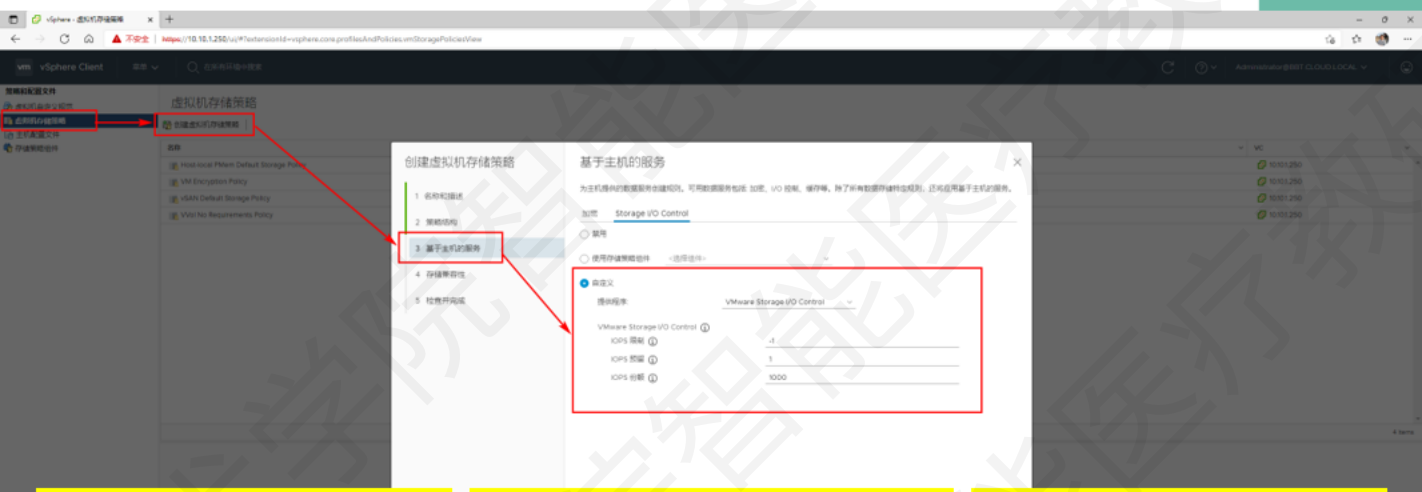
- NAS（Network Attached Storage）网络附加存储。
- 在NAS存储结构中，存储系统不再通过I/O总线附属于某个服务器或客户机，而直接通过网络接口与网络直接相连，由用户通过网络访问。
- NAS实际上是一个带有瘦服务器的存储设备，其作用类似于一个专用的文件服务器。这种专用存储服务器去掉了通用服务器原有的不适用的大多数计算功能，而仅仅提供文件系统功能。
- 与传统以服务器为中心的存储系统（如DAS）相比，数据不再通过服务器内存转发，直接在客户机和存储设备间传送，服务器仅起控制管理的作用。


存储分类
封闭系统的存储
开放系统的存储
内置存储
HDD：Hard Disk Drive
SSD：Solid State Drive
HHD：Hybrid Hard Drive
外置存储
DAS：Direct-Attached Storage
SCSI
IDE
PCI
FAS：Fabric-Attached Storage
NAS：Network-Attached Storage
NFS
CIFS
SAN：Storage Area Network
FC-SAN
IP-SAN
iSCSI
FCIP
iFCP
FCoE
IB-SAN
InfiniBand
Storage Array
RAID

3. vSphere Storage 高级功能

- SIOC 和 Storage DRS是解决存储资源在高负载情况下的争用冲突。
  - 启用SIOC，可以在VM对存储资源使用发生争用时，对虚拟机访问存储资源的优先级和访问压力进行限制管理。
  - 由于存储在所有虚拟机之间是共享的，当特定虚拟机大量且无限制的进行I/O操作视，可能会大量占用存储容量和资源，导致所有虚拟机性能变差。


vSphere - BBT-Cloud-SAN-3-Da
不安全
https://10.10.1.250/ui/#?extensionId=vsphere.core.inventory.serverObjectViewsExtension&objectId=urn:vmomi:Datastore:datastore-3655a97dac-bcc0-4891-b4d8-7795ed79a33d&navigator=vsphere.core.vmTree.storageView
vSphere Client
菜单
在所有环境中搜索
Administrator@BBT.CLOUD.LOCAL
10.10.1.250
BBT-Datacenter
BBT-Cloud-node-1-local-HDD
BBT-Cloud-node-2-local-HDD
BBT-Cloud-node-3-local-HDD
BBT-Cloud-node-4-local-HDD
BBT-Cloud-SAN-3-Data-1
StudioCloud-SAN-1-SYSTEM-1 (不可访问)
StudioCloud-SAN-1-SYSTEM-2 (不可访问)
StudioCloud-SAN-2-SYSTEM-1 (不可访问)
StudioCloud-SAN-2-SYSTEM-2 (不可访问)
StudioCloud-SAN-3-SYSTEM-1 (不可访问)
StudioCloud-SAN-3-SYSTEM-2 (不可访问)
StudioCloud-SAN-3-SYSTEM-3 (不可访问)
BBT-Cloud-SAN-3-Data-1
操作
摘要
监控
配置
权限
文件
主机
虚拟机
类型: VMFS 6
URL: ds:///vmfs/volumes/5d5bc41d-7b7ebe07-655c-bcee7bf2bdd6/
容量
已用 419.27 GB
可用 30.83 TB
容量 31.24 TB
刷新
相关对象
自定义属性
No items to display
操作 - BBT-Cloud-SAN-3-Data-1
新建虚拟机...
浏览文件
注册虚拟机...
刷新容量信息
增加数据存储容量...
维护模式
移至...
重命名...
取消数据存储
配置 Storage I/O Control...
编辑空间回收...
升级到 VMFS-6...
标记与自定义属性
添加权限...
警报
删除数据存储
配置 Storage I/O Control
BBT-Cloud-SAN-3-Data-1
Storage I/O Control 用于控制虚拟机的 I/O 使用情况和提供实现预定义的 I/O 共享级别。
启用 Storage I/O Control 和统计信息收集
拥堵阈值
吞吐量峰值百分比
90
%
手动
30
毫秒
重置为默认值
包括 SDRS 的 I/O 统计信息
禁用 Storage I/O Control，但启用统计信息收集
包括 SDRS 的 I/O 统计信息
禁用 Storage I/O Control 和统计信息收集
取消
确定
近期任务
警报



Limits 限制

是对资源的硬性上限。

VM（或虚拟磁盘）将不会消耗超过限制所配置的资源量。

Reservations 预留

是 VM （或在这种情况下是虚拟磁盘）始终可用的资源数量。

如果未使用预留的资源，剩余的资源将被用于满足其他 VM 的需求。一旦 VM 要访问预留资源，即使其他 VM 正在使用，也会保证其可用性。

Shares 份额

只在发生拥塞时才起作用。

当VM想要访问的资源超过存储所能提供的资源时，就会出现这种情况。在这种情况下，每个 VM 将获得一定数量的资源份额。

3. vSphere Storage 高级功能

- SIOC 和 Storage DRS是解决存储资源在高负载情况下的争用冲突。
  - Storage DRS：Storage Distributed Resource Scheduler (SDRS)
  - 启用 SDRS 后，VM 优化和资源分配依据两个指标：space（容量）和 I/O
  - Storage DRS 管理数据存储群集的聚合资源。
    - 如果启用 Storage DRS，会对虚拟机磁盘放置和迁移提出建议，以平衡数据存储群集中各个数据存储之间的空间和 I/O 资源。
    - 启用 Storage DRS 时，可以使用如下存储集群的功能：
      - 数据存储群集中数据存储之间的空间负载平衡
      - 数据存储群集中数据存储之间的 I/O 负载平衡
      - 基于空间和 I/O 工作负载的虚拟磁盘的初始放置
    - Storage DRS 仅应用于数据存储集群，不能在单个数据存储中使用。

3. vSphere Storage 高级功能

- 数据存储集群是一组数据存储
  - 是一种逻辑上的存储资源池，将多个数据存储（Datastore）组合成一个集群，并提供统一的管理和分配。
  - 为虚拟机提供了更高的存储可用性和性能，同时简化了存储管理的复杂性。
  - 优点与特性：
    - 高可用性：可以跨多个物理存储设备分布数据，从而提高了存储资源的可用性。
    - 负载均衡：可以在多个存储设备之间动态平衡虚拟机的负载，确保每个存储设备的负载均衡，并提供了更高的存储性能。
    - 简化管理：可以在一个集群级别上管理和配置存储资源，简化了存储管理。
    - 存储策略应用：可以与存储策略（Storage Policy）结合使用，根据业务需求定义不同的存储策略，从而实现对存储资源的灵活管理和配置。

3. vSphere Storage 高级功能

- VMFS版本
  - VMFS 1：
    - 用于 ESX Server 1.x，不支持集群。不支持对多个服务器的并发访问。集群不可用，一次只能在一台服务器上使用。
  - VMFS 2：
    - 用于ESX Server 2.x 或 3.x，不支持目录结构。
  - VMFS 3：
    - 适用于vSphere 3.x 或 4.x，支持目录结构。至多50TB容量。VMFS 3 不适用 ESXi 7.0。
  - VMFS 5：
    - 用于vSphere 5.x 和 6.x。单个盘区大小增加到 64TB，VMDK文件大小增加到 62TB。
  - VMFS 6：
    - 在 vSphere 6.5 发布，vSphere 6.7 可以使用。支持4Kn和512n存储设备和本地设备。

3. vSphere Storage 高级功能

VMFS版本

VMFS 6 的特点和优势

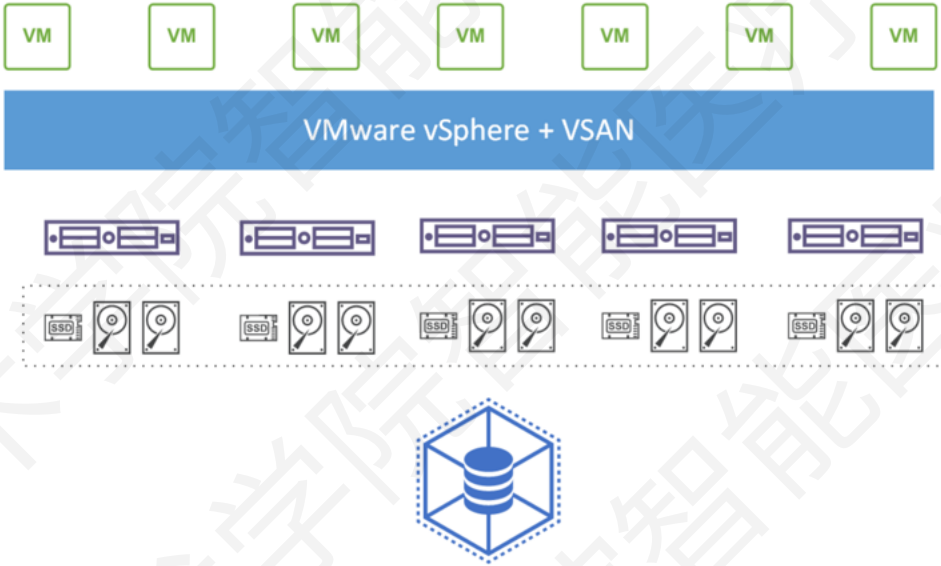
- 64位文件系统：
  - 引入 64 位文件系统，支持更大的文件和存储容量，可以有效管理大规模的虚拟化环境。
- 增强的存储性能：
  - 通过优化和改进了存储访问算法，提供了更高的存储性能和更低的延迟，提升了虚拟机的性能和响应速度。
- 更强的可靠性和一致性：
  - 引入了新的存储校验和功能，可以检测和纠正数据损坏，提高了存储的可靠性和一致性。
- 增强的管理功能：
  - 提供更丰富的管理，包括在线扩容、在线存储迁移、存储I/O控制和存储策略应用等，简化了存储管理的操作和流程。
- 与其他VMFS版本的兼容性：
  - 与早期版本的 VMFS 文件系统兼容，并且可以无缝地升级文件系统到 VMFS 6，而不会影响虚拟机的正常运行。

3. vSphere Storage 高级功能

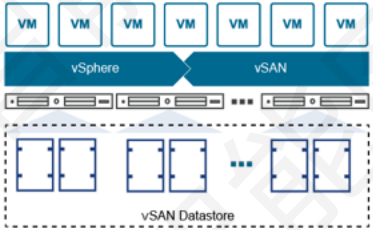
- vSphere 存储管理上新增了诸多高级特性
  - Instant clones versus linked clones：快速克隆与链接克隆的存储支持
  - Permanent Device Loss (PDL) and All-Paths Down (APD)：
    - VM Component Protection (VMCP) 能够管理的两种 VM 存储风险
  - Flash Read Cache：vFlash，闪存读缓存
  - Storage integration：存储集成
  - Multipathing：多路径访问LUN
  - VMware vStorage API for Array Integration (VAAI)：用于和存储设备进行深度整合
  - VMware vSphere APIs for I/O Filtering (VAIO)：用于I/O 过滤 vSphere API框架

4. VMware vSAN

- VMware vSAN 是软件定义的企业存储解决方案，支持超融合基础架构 (Hyper-Converged Infrastructure, HCI) 系统。
  - vSAN 与 VMware vSphere 完全集成在一起，是 ESXi 内置的功能。
  - vSAN 使用本地存储设备或直接连接的存储设备，创建 vSAN 集群中所有主机之间共享的单个存储池。
    - 混合 vSAN 集群将闪存设备用于缓存层，将磁盘驱动器用于容量层。
    - 全闪存 vSAN 集群将闪存设备用于缓存层和容量层。
    - 可用于软件定义数据中心（SDDC），且是经过闪存优化的弹性共享数据存储。
  - vSAN 无需外部共享存储，并通过基于存储策略的管理 (Storage Policy-Based Management, SPBM) 简化了存储配置。


VM
VM
VM
VM
VM
VM
VM
VMware vSphere + VSAN
SSD
SSD
SSD
SSD
SSD
VSAN Datastore

4. VMware vSAN



VMware vSAN:
https://docs.vmware.com/cn/VMware-vSAN/index.html

4. VMware vSAN

VMware vSAN 介绍
https://www.bilibili.com/video/BV1j4411K7ZV（3分40秒）
https://www.bilibili.com/video/BV1Ls411K7AJ（31分51秒）

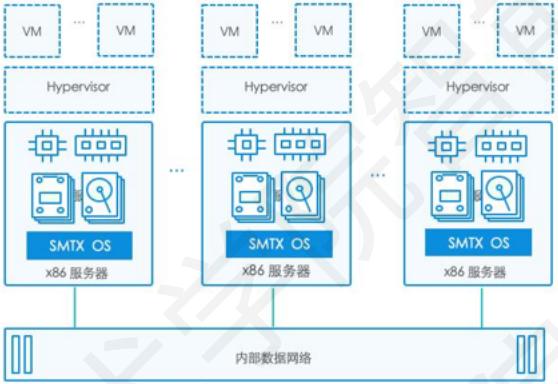
4. VMware vSAN

- vSAN实现的两种设备选型方案：
  - Hybrid deployment：混合部署
    - 使用固态硬盘作为缓存层，磁盘作为容量层。
    - 不需要使用 RAID 5/6 进行去重和压缩，以及存储容灾。
  - All-flash：全闪存（推荐方案）
    - 使用 High-performance SSD （例如 NVMe）作为缓存层，使用 SSDs 作为容量层。
    - All-flash 能够提供非常高的 I/O 性能。
  - 网络要求：
    - 建议使用成对的 10 GbE Network Cards。
    - 将 vSAN 流量进行隔离，并为 vSAN 提供充足带宽。

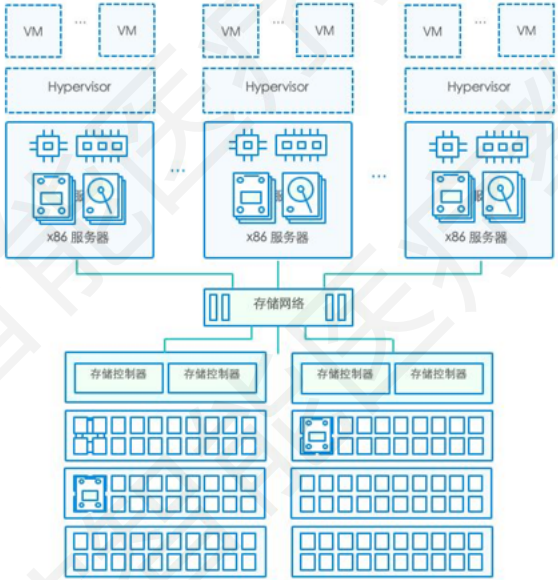
4. VMware vSAN

现场演示：
vSAN 的实现与应用

SDDC的虚拟化数据中心设计方案


VM
VM
Hypervisor
SMTX OS
x86 服务器
内部数据网络

超融合的架构


VM
Hypervisor
x86 服务器
存储网络
存储控制器

传统架构

智能运维课程体系


棒棒堂智能运维
UID: 510401430
抖音
@棒棒堂智能运维
抖音号：ID2003079563
只要思想不滑坡，办法总比困难多。
保存图片到相册
打开抖音搜索页扫一扫