

# 云计算与虚拟化技术



## 第7讲：Storage Devices and vSAN

阮晓龙

13938213680 / rxl@hactcm.edu.cn

<http://cloud.xg.hactcm.edu.cn>  
<http://www.51xueweb.cn>

河南中医药大学信息管理与信息系统教研室  
信息技术学院网络与信息系统科研工作室

2021.10

# 讨论提纲

- Storage basics
  - 要重视数据中心的存储设计
  - 存储阵列 Storage arrays: performance、RAID、SSDs、AFAs、ALUA
- vSphere storage types
  - Storage types at the ESXi logical level and physical level
  - Storage types at VM logical levels and physical level
- vSphere storage configuration
  - FC storage、iSCSI storage、NFS storage
  - SIOC and storage DRS
  - Advanced storage features
- VMware vSAN
  - vSAN的设计方案
  - vSAN的配置



# 1. Storage basics

## 1.1 要重视数据中心的存储设计

- 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。
  - vSphere 许多高级功能依赖于共享存储，例如HA、DRS、FT、VMware Site Recovery Manager等。
- 数据中心存储解决方案的基本要求：高可用、高性能、弹性扩展。
- Enterprise-class storage can be classified in different ways:
  - Direct-Attached Storage (DAS)
  - Network Attached Storage (NAS)
  - Storage Area Network (SAN)
  - Object-based storage/cloud storage

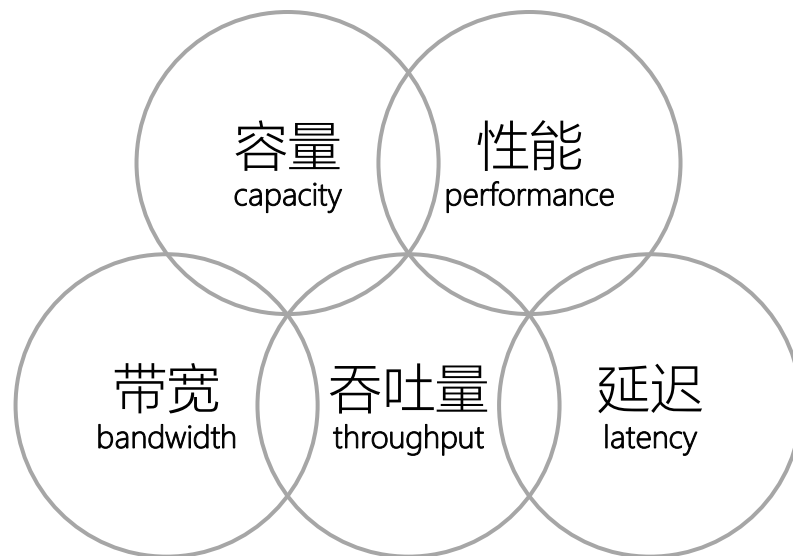
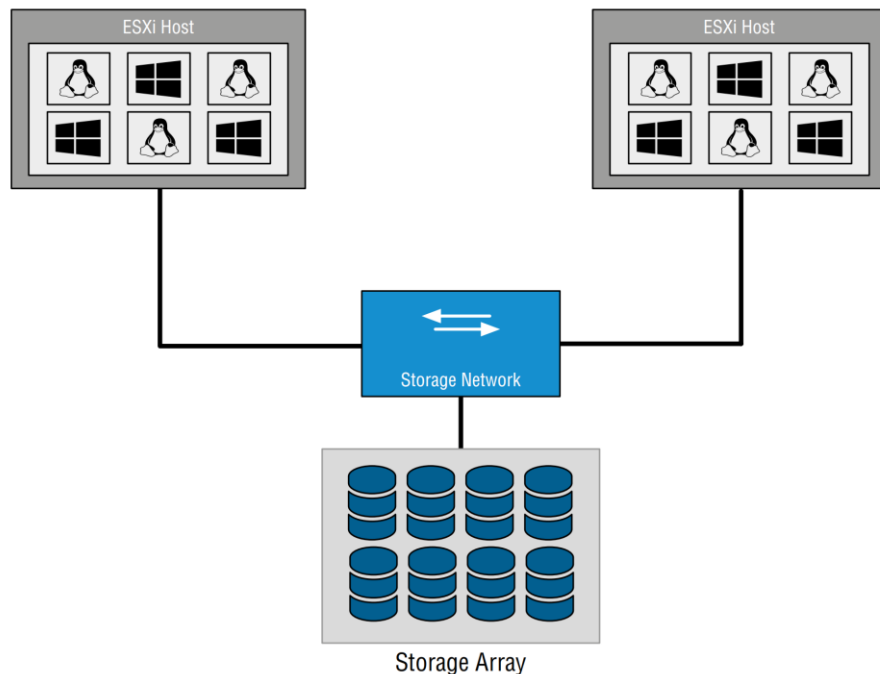
the only solutions that can be used for running VMs

could be used by other solutions (such as backup products)



# 1. Storage basics

## 1.1 要重视数据中心的存储设计



The ESXi hosts and the VMs will be contending for the shared storage asset.



# 1. Storage basics

## 1.1 要重视数据中心的存储设计

- 仅考虑性能，存储通常被分为以下几个等级。
  - Tier 0:
    - Very high-performance storage, such as the All-Flash Array (AFA). Tier 0 storage is getting more and more popular as the prices of Enterprise class SSDs are dropping. It provides unmatched I/O performance.
  - Tier 1 or primary storage:
    - This is usually the main storage and corresponds to the VMware side in different datastores.
  - Tier 2 or secondary storage:
    - This is storage not (usually) used from VMware production environments. It stores online archives, backups, cold data, and so on.
  - Tier 3:
    - This can be long-term and offline archival storage repositories, such as tapes, or copies of backups on public cloud storage.



# 1. Storage basics

## 1.1 要重视数据中心的存储设计

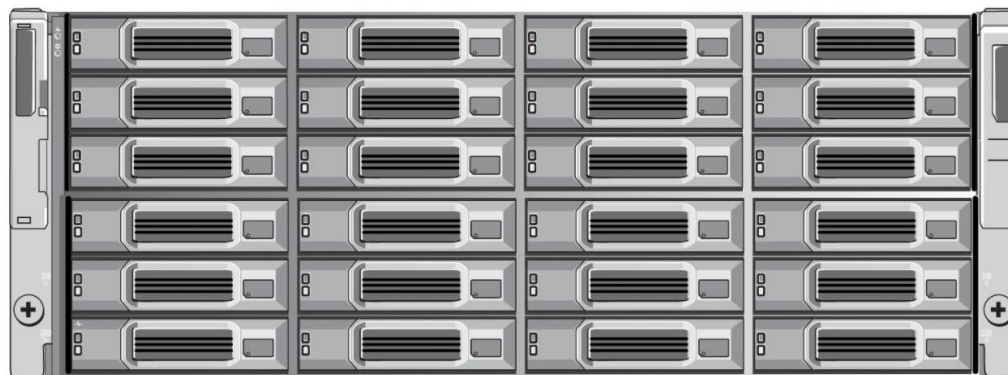
- ESXi Host能够支持的存储类型有：
    - Local SAS/SATA/SCSI storage
    - Fibre Channel
    - Fibre Channel over Ethernet (FCoE)
    - iSCSI using software and hardware initiators
    - NAS (specifically, NFS)
    - InfiniBand
- Local Storage
- Shared Storage
- 



# 1. Storage basics

## 1.2 Storage arrays

### 什么是存储阵列 Storage arrays?

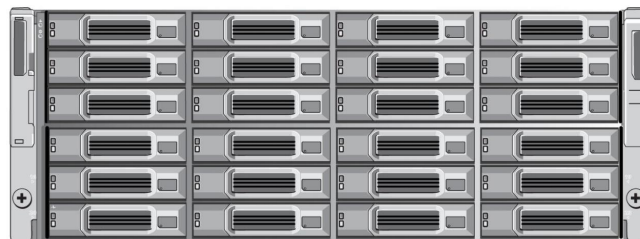


- 通过华为交互式3D在线体验华为存储产品  
<https://info.support.huawei.com/storage/3d-center>
- 华为存储基础课程  
<https://www.bilibili.com/video/BV1Zt4y1a7Rj>



# 1. Storage basics

## 1.2 Storage arrays



**AFAs**  
flash memories

**Hybrid  
Array**  
flash and HDD





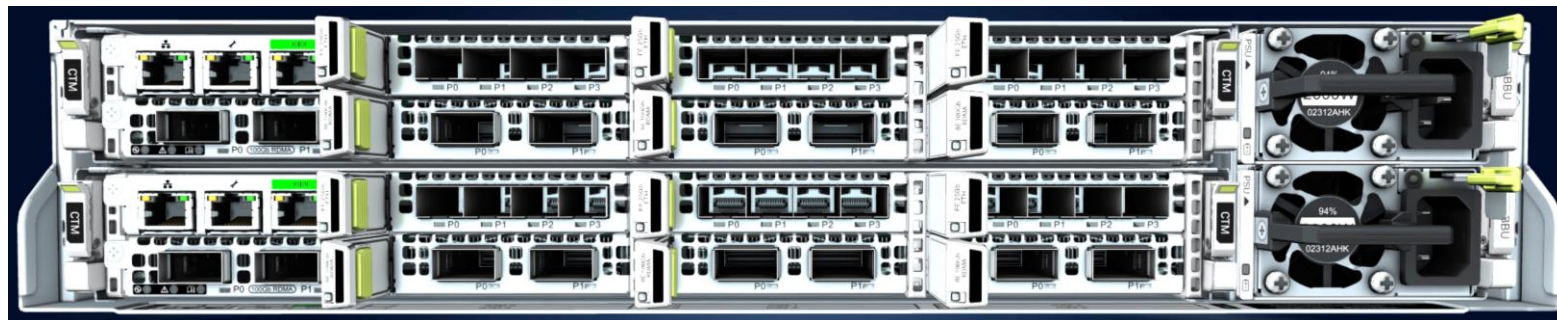
# 1. Storage basics

## 1.2 Storage arrays

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



## 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. Storage basics

## 1.2 Storage performance

### 影响存储性能的因素

Physical disk or SSDs  
physical media      the maturity of the storage controller  
the protocols used      the IO size

#### 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. Storage basics

## 1.3 RAID

- RAID ( Redundant Array of Independent Disks ) 即独立磁盘冗余阵列，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，从而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 Striping )
    - 数据校验 ( Data parity )

**RAID choices as a significant factor in your design**



# 1. Storage basics

## 1.3 RAID

### □ 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技术全解图解 <https://blog.csdn.net/ensp1/article/details/81318135>

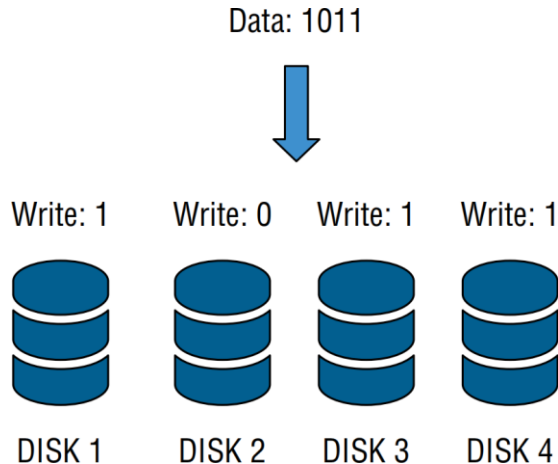


# 1. Storage basics

## 1.3 RAID

### □ 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。

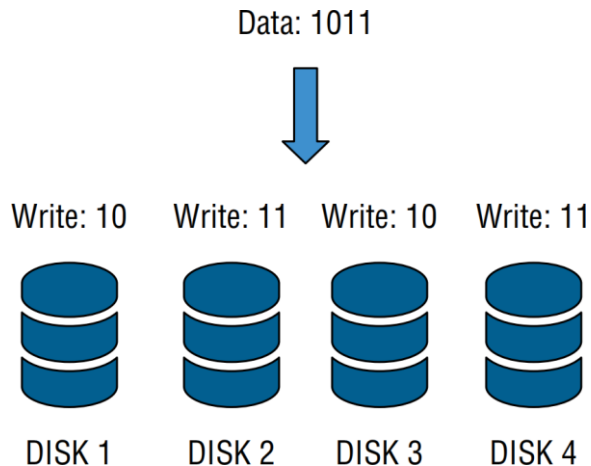


# 1. Storage basics

## 1.3 RAID

### □ 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. Storage basics

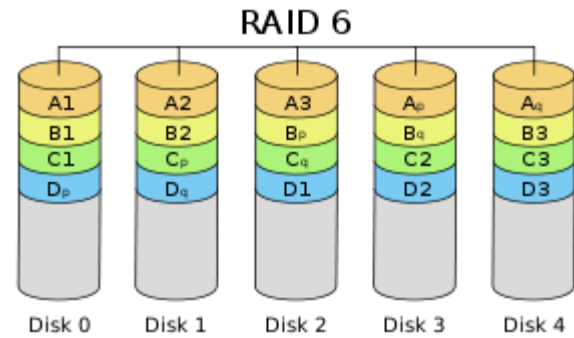
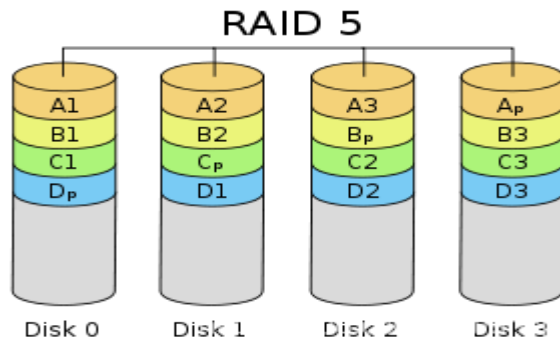
## 1.3 RAID

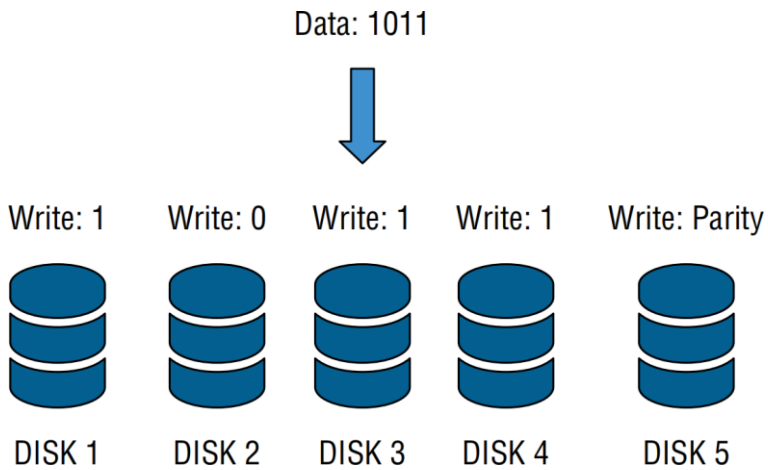
### □ RAID 5,6:

- 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。

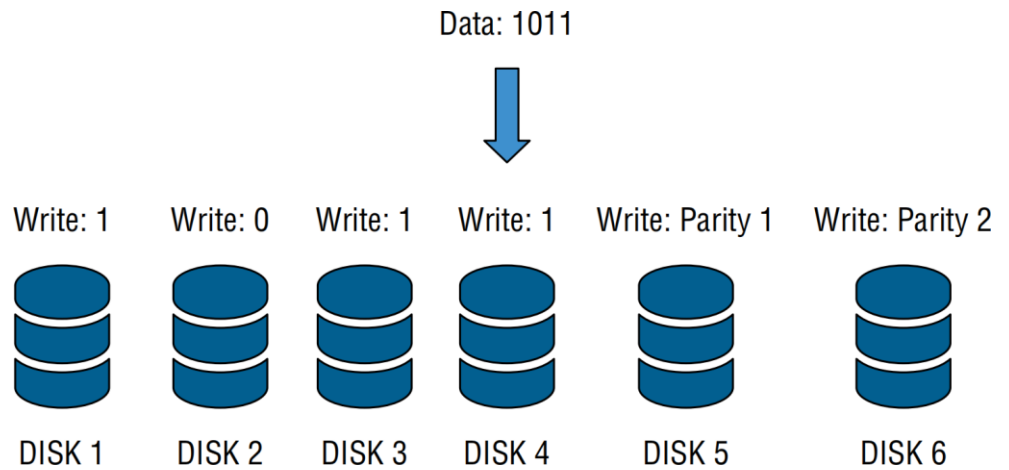








RAID 5 4+1



RAID 6 4+2



## 2. vSphere storage types

### 2.1 存储体系结构的支持情况

VMware vSphere supports different types of storage architecture

Technology	Protocols	Transfer	Interface
FC	FC/SCSI	Block access	FC HBA
FCoE	FCoE/SCSI	Block access	Converged network adapter (hardware FCoE) NIC with FCoE support (software FCoE)
iSCSI	IP/SCSI	Block access	iSCSI HBA-or iSCSI-enabled NIC (hardware iSCSI) Network adapter (software iSCSI)
NAS	IP/NFS	File access	Network adapter



## 2. vSphere storage types

### 2.1 存储体系结构的支持情况

Different storage types have different vSphere characteristics

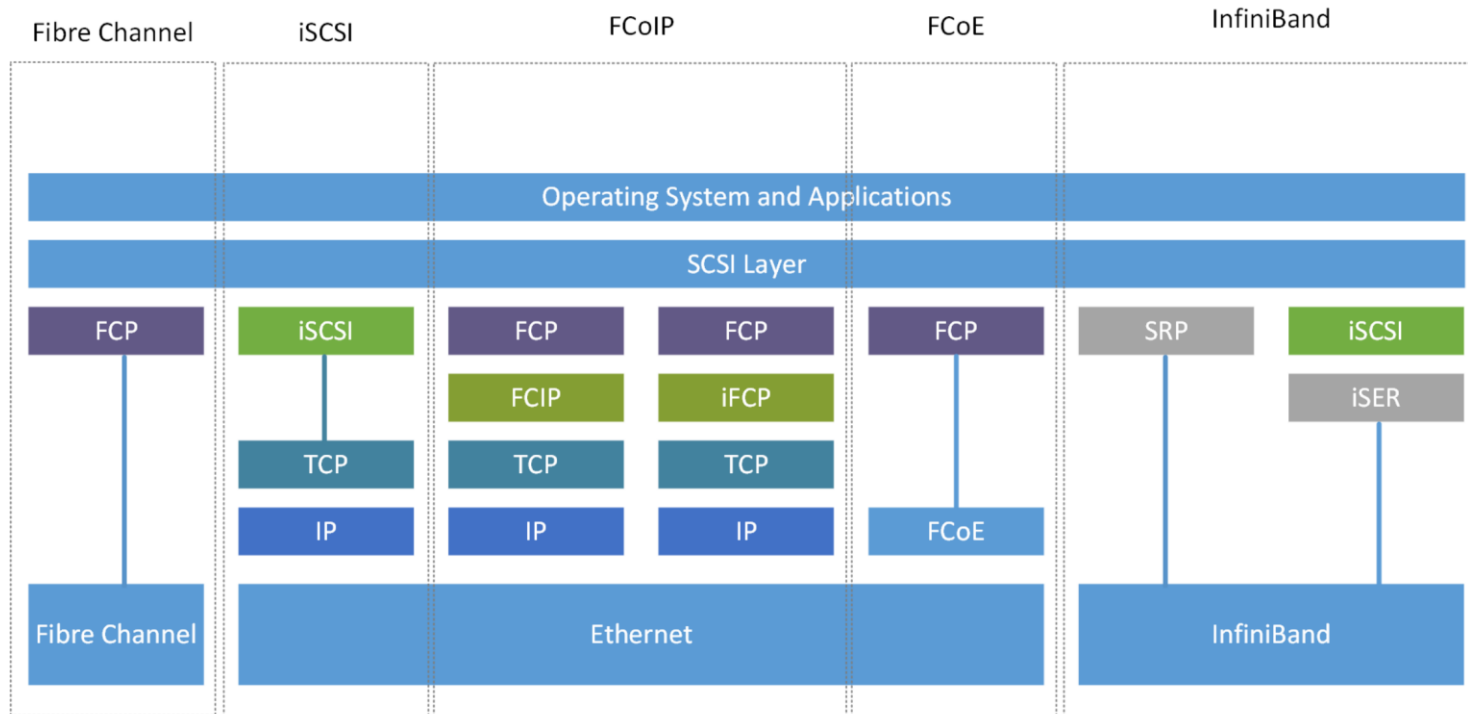
Storage type	Boot ESXi	Run VMs	vMotion	Datastore	RDM	VM clustering (guest OS)	VMware HA and DRS	Storage APIs – data protection
Local Storage	Yes	Yes	No	VMFS	No	Yes	No	Yes
Fibre Channel and FCoE	Yes*	Yes	Yes	VMFS	Yes	Yes	Yes	Yes
iSCSI	Yes*	Yes	Yes	VMFS	Yes	Yes	Yes	Yes
NAS over NFS	No	Yes	Yes	NFS3 and NFS 4.1	No	No	Yes	Yes



## 2. vSphere storage types

### 2.1 存储体系结构的支持情况

the flow of the SCSI command from the guest OS down to the physical network



## 2. vSphere storage types

### 2.2 ESXi支持的存储类型

- vSphere 6.x支持四种数据存储类型
  - VMware File System (VMFS) datastores:
    - All block-based storage must be first formatted with VMFS to transform a block service to a file and folder oriented services.
  - Network File System (NFS) datastores:
    - This is for NAS storage.
  - VVol:
    - This is introduced in vSphere 6.0 and is a new paradigm for accessing SAN and NAS storage in a uniform way and by better integrating and consuming storage array capabilities.
  - vSAN datastore:
    - If you are using a vSAN solution, all your local storage devices could be polled together in a single shared vSAN datastore.

**Storage types at the ESXi logical level**



## 2. vSphere storage types

### 2.2 ESXi支持的存储类型

- vSphere 6.x支持三种物理存储阵列 (不包含vSAN)
  - Block-based storage accessed by a **hardware adapter**:
    - This includes DAS storage or SAN FC storage.
  - Block-based storage accessed by a **software adapter**:
    - This is like SAN iSCSI storage when the software initiator is used. In this case, you first need to configure the network connectivity properly. After that, it becomes very similar to the first case.
  - **NFS storage**:
    - This is where you first have to configure the IP network connectivity to your storage and then connect the NFS datastore.

**Storage types at the ESXi physical level**



## 2. vSphere storage types

### 2.2 ESXi支持的存储类型

- vSphere 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>
    - This is the standard way to implement FC-based storage using a dedicated full fabric.
  - iSCSI HBA: <https://sniansfblog.org/tag/iscsi/>
    - These are specialized PCIe cards that implement the entire iSCSI stack completely in the hardware, reducing the load on the host CPU.
  - CNA adapters for FCoE or iSCSI: <https://blog.csdn.net/wuzhimang/article/details/52399064>
    - These are mostly 10 Gbps (or higher) Ethernet adapters providing hardware (or hardware-assisted) FCoE or iSCSI functionality on converged (or dedicated) networks.
  - RDMA over Converged Ethernet (RoCE): <https://baijiahao.baidu.com/s?id=1689933693620352512>
    - This is a network protocol that allows remote direct memory access (RDMA) over an Ethernet network. Starting with vSphere 6.5, RoCE certified adapters can be used for converged networks.
  - InfiniBand HCA: <https://www.cnblogs.com/D-Tec/p/3157582.html>
    - Mellanox Technologies InfiniBand HCA device drivers are available directly from Mellanox Technologies. Mostly used for the network instead of the storage, they can be interesting in converged networks, and also in vSAN implementation.

**Storage types at the ESXi physical level**





- 10.10.1.250
- BBT-Datacenter
  - 10.10.1.15
  - 10.10.1.16
  - 10.10.1.17
  - 10.10.1.18
  - VM-VMware vCenter Server Appliance-10.10.1...

10.10.1.15 操作

摘要 监控 配置 权限 虚拟机 资源池 数据存储 网络 更新

- 存储
  - 存储适配器
  - 存储设备
  - 主机库存储配置
  - 协议端点
  - I/O 筛选器
- 网络
  - 虚拟交换机
  - VMkernel 适配器
  - 物理适配器
  - TCP/IP 配置
- 虚拟机
  - 虚拟机启动/关机
  - 代理虚拟机设置
  - 默认虚拟机兼容性
  - 交换文件位置
- 系统
  - 许可
  - 主机配置文件
  - 时间配置
  - 身份验证服务
  - 证书
  - 电源管理
  - 高级系统设置
  - 系统资源预留
  - 防火墙
  - 服务
  - 安全配置文件
  - 系统交换
  - 软件包
- 硬件
  - 处理器
  - 内存
  - 电源管理
- 更多
  - 警报定义
  - 已调度任务

## 存储适配器

+ 添加软件适配器 刷新 重新扫描存储... 重新扫描适配器

适配器	类型	状态	标识符	目标	设备	路径
型号: Emulex LPe1250 8Gb PCIe Fibre Channel Adapter						
vmhba2	光纤通道	联机	20:00:00:90:fa:5e:f3:86 10:00:00:90:fa:5e:f3:86	6	12	24
型号: ISCSI Software Adapter						
vmhba64	iSCSI	联机	iqn.1998-01.com.vmware.bbt-cloud-node-1-15005bfc	1	0	0
型号: LSI2008						
vmhba3	SAS	未知	--	1	1	1
型号: SB700 SATA Controller [IDE Mode]						
vmhba0	块 SCSI	未知	--	0	0	0
型号: SB700/SB800 IDE Controller						
vmhba1	块 SCSI	未知	--	1	1	1
vmhba32	块 SCSI	未知	--	0	0	0

Copy All 6 items

未选择任何项目



## 2. vSphere storage types

### 2.3 VM支持的存储类型

- 自vSphere 4.0之后，VM的存储类型基本确定。
  - 根据VM存储的分配方法不同，一般有三种格式。
    - Eager zeroed thick Virtual Machine Disk (VMDK): **厚置备快速置零**
      - An eager zeroed thick disk has all the space allocated and wiped clean of any previous content on the physical media at creation time. Such disks may take a long time to create compared to other disk formats. The entire disk space is reserved and unavailable for use by other VMs.
    - Thick or lazy zeroed thick VMDK: **厚置备延迟置零**
      - A thick disk has all space allocated at creation time. This space may contain stale data on the physical media. Before writing to a new block, a zero has to be written, increasing the input/output operation per second (IOPS) on new blocks compared to eager disks. The entire disk space is reserved and unavailable for use by other VMs.
    - Thin VMDK: **精简置备**
      - The space required for the thin-provisioned virtual disk is allocated and zeroed on demand as space is used. Unused space is available for use by other VMs.

**Storage types at VM logical levels**



## 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

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間。

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。

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。



vm vSphere Client 菜单 在搜索框中输入搜索

Administrator@BBT-CLOUD.LOCAL

10.10.1.15 操作

摘要 监控 配置 权限 虚拟机 资源池 数据存储 网络 更新

管理程序: VMware ESXi 6.7.0, 13006603  
型号: 24000635/24000741  
处理单元: AMD Opteron(tm) Processor 6320  
逻辑处理器: 16  
网卡: 2  
虚拟机: 0  
状况: 已连接  
正常运行时间: 7天

CPU 可用: 44.72 GHz  
已用: 88 MHz 容量: 44.8 GHz  
内存 可用: 129.45 GB  
已用: 2.53 GB 容量: 127.96 GB  
存储 可用: 31.09 TB  
已用: 158.34 TB 容量: 169.43 TB

主机硬件风扇状态  
主机硬件温度状态

硬件

制造商  
型号  
CPU  
内存  
虚拟闪存资源  
网络  
存储

标记

已分配的标记

分配 移除

自定义属性

属性 值  
AutoDeploy.MachineIdentity

### 新建虚拟机

- 1 选择创建类型
- 2 选择名称和文件夹
- 3 选择计算资源
- 4 选择存储
- 5 选择兼容性
- 6 选择客户机操作系统
- 7 自定义硬件**
- 8 即将完成

自定义硬件  
配置虚拟机硬件

虚拟硬件 虚拟机选项

添加新设备

新硬盘	16	GB
大小上限	30.83 TB	
虚拟机存储策略	数据存储默认值	
位置	与虚拟机存储在同一目录中	
磁盘置备	厚置备延迟置零	
共享	厚置备快速置零	
份额	精简置备	
	LEDS	
限制 - IOPs	不受限制	
虚拟闪存读取缓存	0	MB
磁盘模式	从属	

兼容性: ESXi 6.7 及更高版本 (虚拟机版本 14)

CANCEL BACK NEXT

合规 (从未检查过)  
修复状态未知 (从未检查过)

近期任务 警告

vSphere - 10.10.1...

11:49  
21/10/22

## 2. vSphere storage types

### 2.3 VM支持的存储类型

- ❑ 虚拟机访问存储设备需要添加存储控制器，VM支持五种存储控制器类型：
  - **BusLogic** : This was one of the first emulated SCSI virtual controllers available in VMware ESX. It's now a legacy controller used mainly for legacy operating systems. It does not support VMDKs larger than 2 TB.
  - **LSI** : logic parallel: This was formally known as LSI Logic and was the other SCSI virtual controller available formerly in VMware ESX. It is used for operating systems such as Windows Server 2003.
  - **LSI logic SAS** : This was introduced in vSphere 4.0 and is the evolution of the parallel driver. It works as an SAS virtual controller and is used in Windows Server 2008 or newer.
  - **VMware paravirtual (or PVSCSI)** : This was introduced in vSphere 4.0. It's a SCSI virtual controller designed to support very high throughput with minimal processing cost, working not in emulation mode, but in paravirtual mode (it requires VMware Tools to be recognized).
  - **NVMe** : Virtual NVMe devices have reduced guest I/O processing overheads (over 50% compared to AHCI SATA SCSI device), which allows more VMs per host or more transactions per minute. Each VM supports four NVMe controllers and up to 15 devices per controller.

**Storage types at the VM physical level**



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



vm vSphere Client 菜单 在搜索框中输入搜索

Administrator@EBBT-CLOUD.LOCAL

10.10.115 操作

摘要 监控 配置 权限 虚拟机 资源池 数据存储 网络 更新

管理程序: VMware ESXi 6.7.0, 13006603  
型号: 24000635/24000741  
处理器类型: AMD Opteron(tm) Processor 6320  
逻辑处理器: 16  
网卡: 2  
虚拟机: 0  
状况: 已连接  
正常运行时间: 7 天

CPU 可用: 44.72 GHz  
已用: 88 MHz 容量: 44.8 GHz  
内存 可用: 129.45 GB  
已用: 2.83 GB 容量: 127.66 GB  
存储 可用: 31.09 TB  
已用: 168.34 TB 容量: 169.43 TB

主机硬件风扇状态  
主机硬件温度状态

硬件

制造商  
型号  
CPU  
内存  
虚拟闪存资源  
网络  
存储

标记

已分配的标记

分配... 移除...

自定义属性

属性 值  
AutoDeploy.MachineIdentity

### 新建虚拟机

- 1 选择创建类型
- 2 选择名称和文件夹
- 3 选择计算资源
- 4 选择存储
- 5 选择兼容性
- 6 选择客户机操作系统
- 7 自定义硬件
- 8 即将完成

自定义硬件  
配置虚拟机硬件

虚拟硬件 虚拟机选项

添加新设备

份额	正常	1000
限制 - IOPs	不受限制	
虚拟闪存读取缓存	0	MB
磁盘模式	从属	
虚拟设备节点	新 SCSI 控制器	SCSI (0:0) 新硬盘
新 SCSI 控制器	VMware 准虚拟	
更改类型	VMware 准虚拟	
SCSI 总线共享	BusLogic 并行 LSI Logic 并行 LSI Logic SAS VMware 准虚拟	<input type="checkbox"/> 连接...
新网络		
新的 CD/DVD 驱动器	客户机设备	

兼容性: ESXi 6.7 及更高版本 (虚拟机版本 14)

CANCEL BACK NEXT

合规 (从未检查过)  
修复状态未知 (从未检查过)

近期任务 警告

vSphere - 10.10.1...

12:02  
2/11/22

10.10.115 操作

摘要 监控 配置 权限 虚拟机 资源池 数据存储 网络 更新

管理程序: VMware ESXi 6.7.0, 13006603  
型号: 24000635/24000741  
处理单元: AMD Opteron(tm) Processor 6320  
逻辑处理器: 16  
网卡: 2  
虚拟机: 0  
状况: 已连接  
正常运行时间: 7 天

CPU 可用: 44.72 GHz  
已用: 85 MHz 容量: 44.8 GHz  
内存 可用: 129.45 GB  
已用: 2.53 GB 容量: 127.96 GB  
存储 可用: 31.09 TB  
已用: 158.34 TB 容量: 169.43 TB

主机硬件风扇状态  
主机硬件温度状态

硬件

- 制造商
- 型号
- CPU
- 内存
- 虚拟闪存资源
- 网络
- 存储

标记

已分配的标记

分配 移除

自定义属性

属性	值
AutoDeploy.MachineIdentity	

### 新建虚拟机

- 1 选择创建类型
- 2 选择名称和文件夹
- 3 选择计算资源
- 4 选择存储
- 5 选择兼容性
- 6 选择客户机操作系统
- 7 自定义硬件**
- 8 即将完成

自定义硬件 配置虚拟机硬件

虚拟硬件 虚拟机选项

> CPU	1	G
> 内存	2	G
> 硬盘 *	16	G
大小上限	30.83 TB	
虚拟机存储策略	数据存储默认值	
位置	与虚拟机存储在同一...	
磁盘策略	厚置备延迟置零	
共享	未指定	
份额	正常	1000
限制 - IOPs	不受限制	

兼容性: ESXi 6.7 及更高版本 (虚拟机版本 14)

添加新设备

- CD/DVD 驱动器
- 主机 USB 设备
- 硬盘
- RDM 磁盘
- 现有硬盘
- 网络适配器
- SCSI 控制器
- USB 控制器
- SATA 控制器
- NVDIMM
- NVMe 控制器
- 共享 PCI 设备
- PCI 设备
- 串行端口

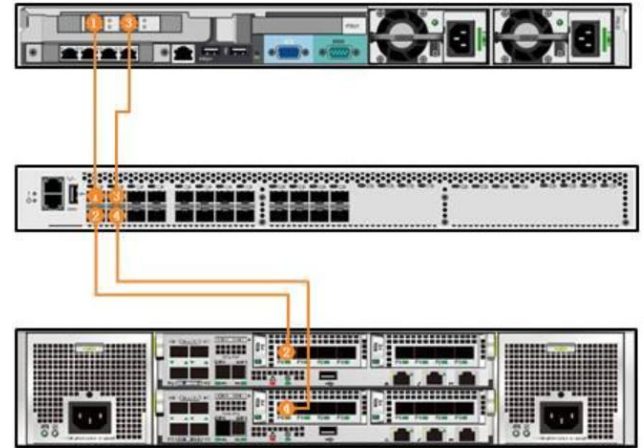
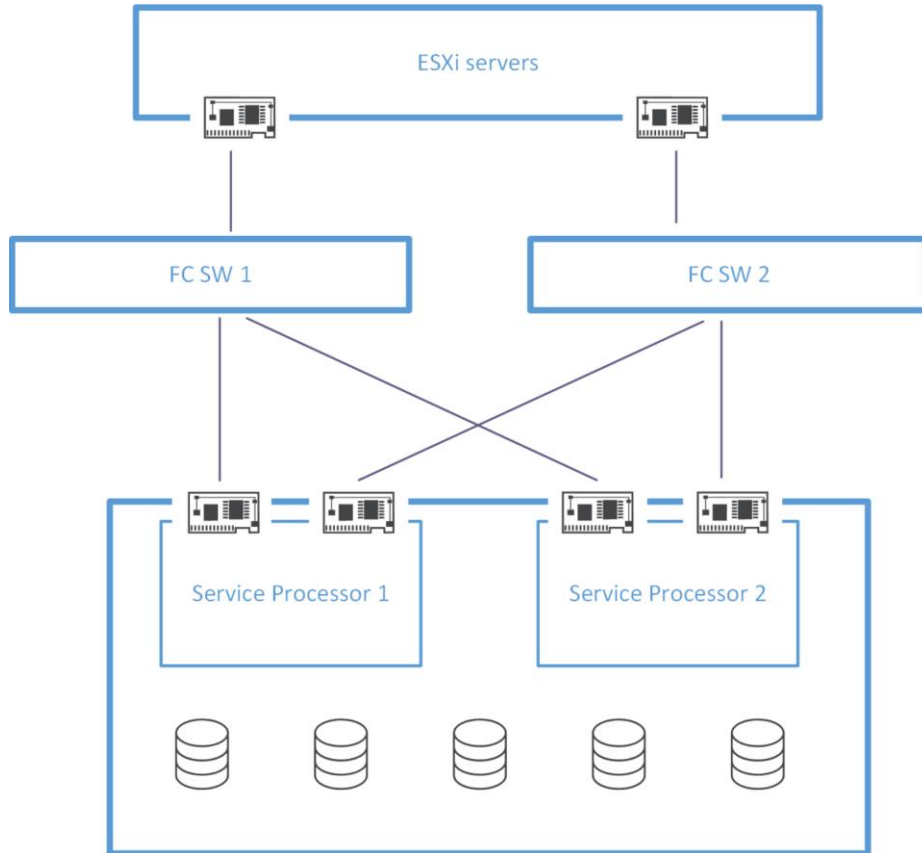
CANCEL BACK NEXT





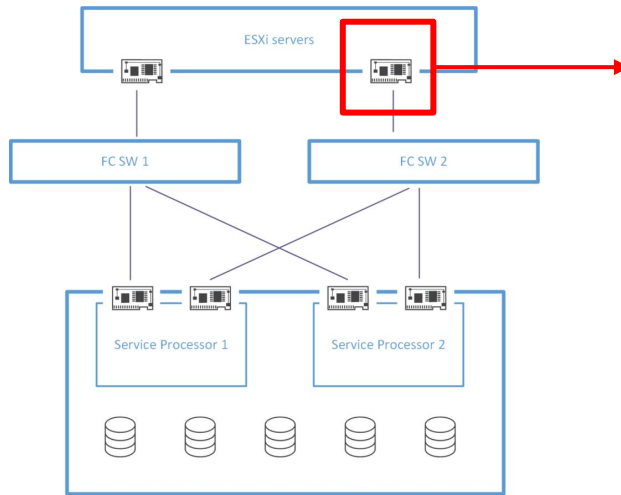
# 3.vSphere storage configuration

## 3.1 FC storage



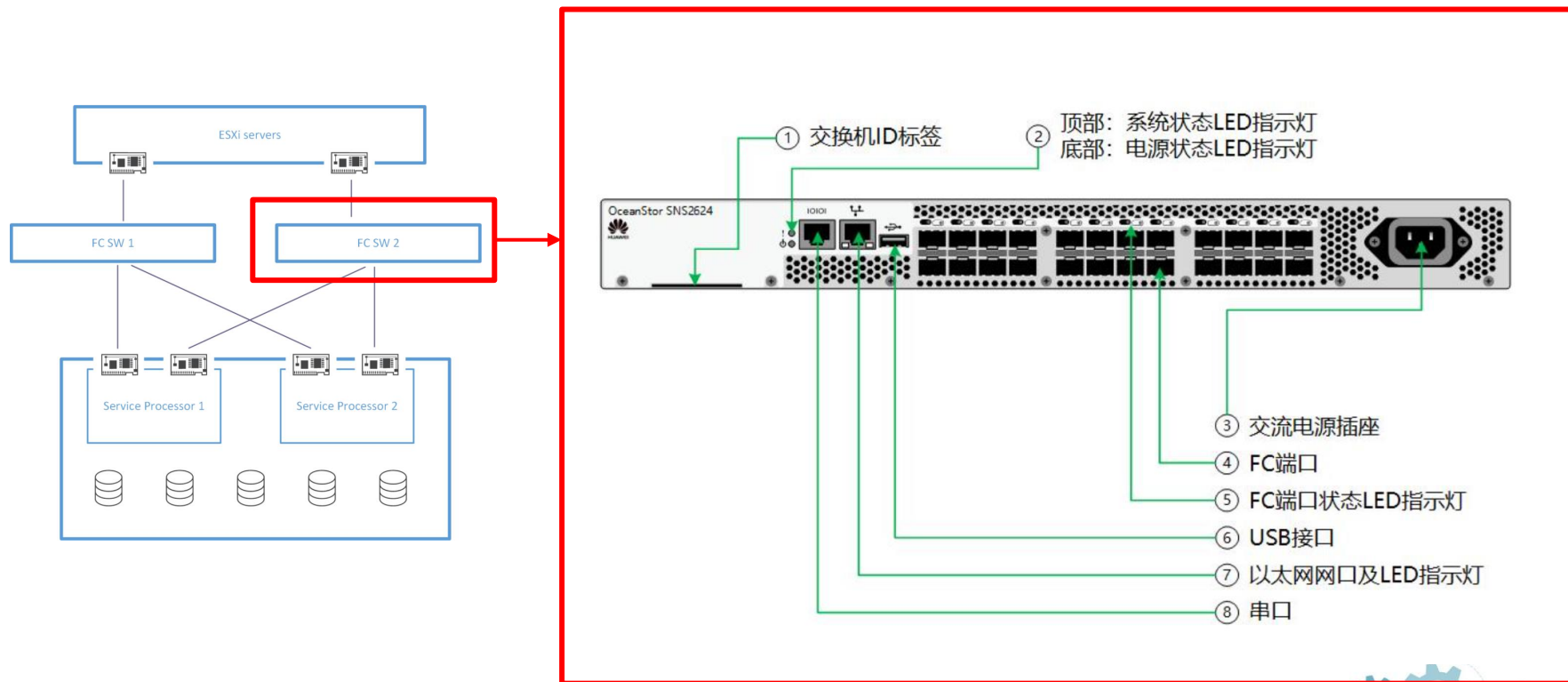
# 3.vSphere storage configuration

## 3.1 FC storage



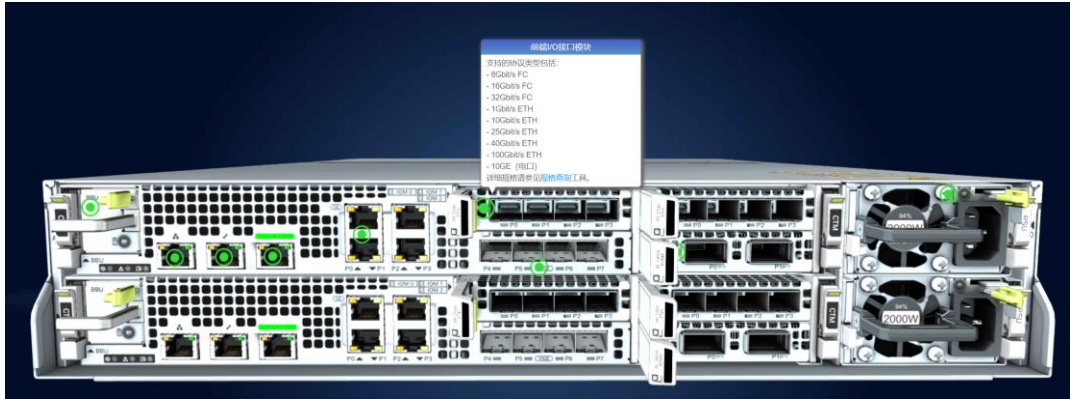
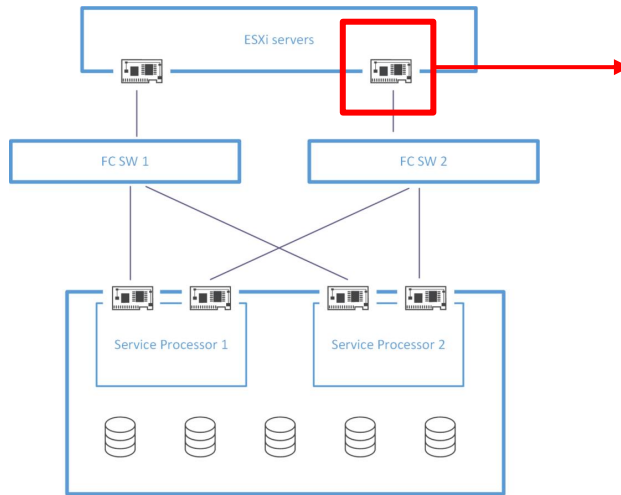
# 3.vSphere storage configuration

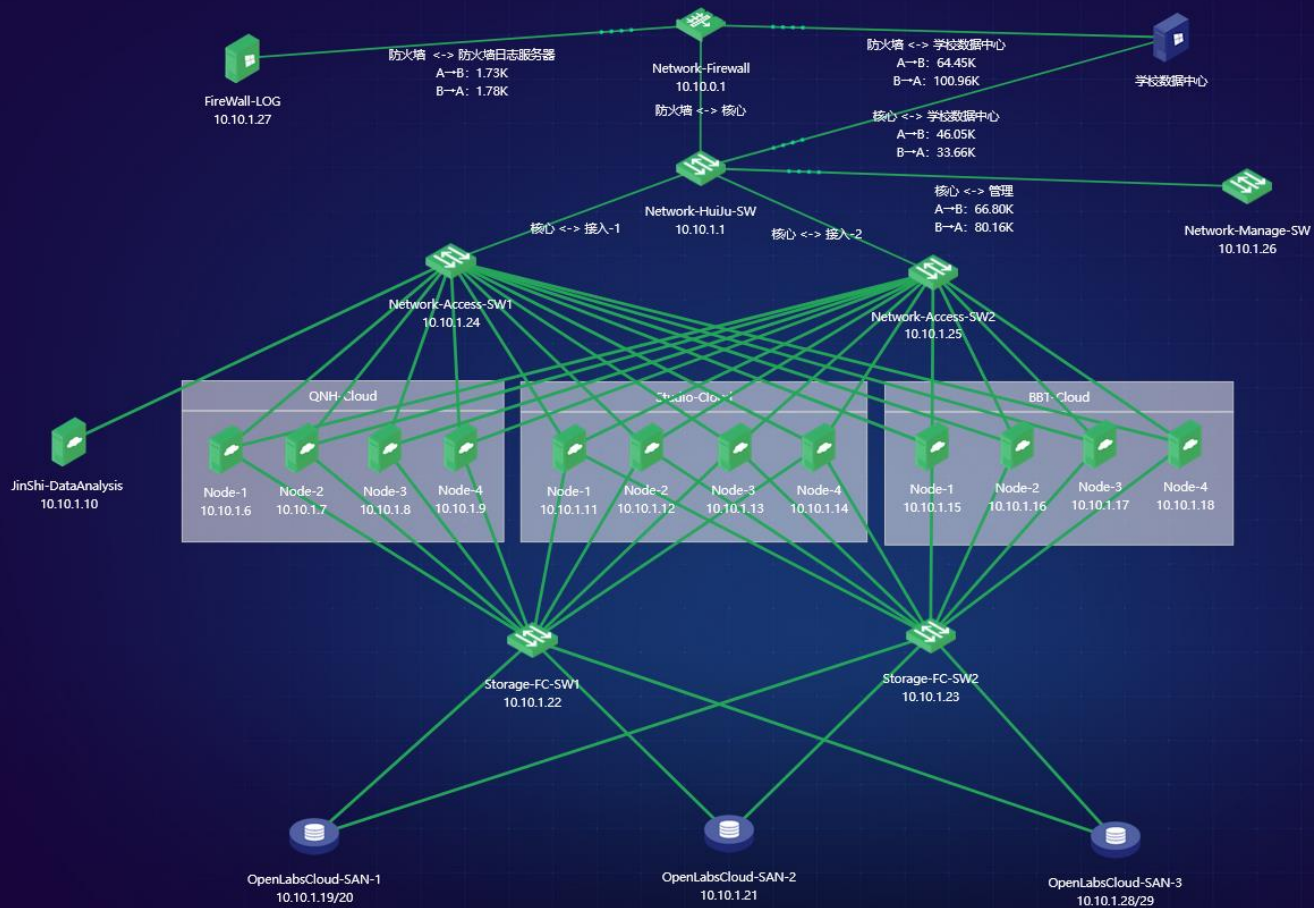
## 3.1 FC storage



# 3.vSphere storage configuration

## 3.1 FC storage





## 3.vSphere storage configuration

### 3.2 FCoE storage

- 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。



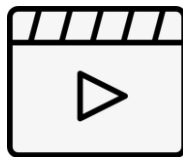
## 3.vSphere storage configuration

### 3.2 FCoE storage

关于FCoE推荐阅读

以太网光纤通道(FCoE)技术综述

[http://www.h3c.com/cn/d\\_201012/705332\\_30008\\_0.htm](http://www.h3c.com/cn/d_201012/705332_30008_0.htm)



华为CloudEngine系列交换机FC和FCoE特性介绍

<https://e.huawei.com/cn/videolist/video/25b1591c674c4881a4f8ac371a2ac199>



## 3.vSphere storage configuration

### 3.3 iSCSI storage

#### □ iSCSI:

- Internet Small Computer System Interface, Internet 小型计算机系统接口
  - 是一种在TCP/IP上进行数据块传输的标准。
  - 是由CISCO和IBM两家公司发起, 并得到存储厂商的广泛支持。
  - iSCSI将SCSI命令封装于TCP/IP包之中, 使用了一个SCSI帧头。
  - iSCSI可以实现在IP网络上运行SCSI协议, 使其能够在诸如高速千兆以太网上进行快速的数据存取备份操作, 也就是基于网络的存储。
  - iSCSI又称为IP-SAN (基于IP网络的存储区域网), 是一种基于因特网及SCSI-3协议下的存储技术, 是实现SAN的另一种方式。

iSCSI详解及 iSCSI配置: <https://blog.csdn.net/tjiyu/article/details/52811458>





## 3.vSphere storage configuration

### 3.3 iSCSI storage

- iSCSI:
  - ESXi can be one of the following iSCSI initiator types:
    - **Software iSCSI adapter** : Use one or more VMkernel network interfaces and the virtual switches to manage the entire iSCSI traffic. With the software iSCSI adapter, you can use iSCSI technology without purchasing specialized hardware.
    - **Dependent hardware iSCSI adapter** : VMware manages iSCSI management and configuration, and it may also be the part of the network that must be implemented at the virtual switch level. Ethernet NIC with iSCSI offload capabilities falls into this category. At the ESXi level, those NICs are presented with two different components—a hardware iSCSI adapter and a corresponding standard networking NIC.
    - **Independent hardware iSCSI adapter or iSCSI HBA** : This is like the FC HBA. All of the network stack is implemented in hardware inside the adapter. On the ESXi side, you will see one or more vmhba, like with all other block storage adapters. Network configuration must be performed at the card level, using BIOS management, or specific tools (there are also plugins for vCenter to manage the configuration inside vSphere).





10.10.1.250

BBT-Datacenter

10.10.1.15

10.10.1.16

10.10.1.17

10.10.1.18

10.10.1.15 操作 ▾

摘要 监控 配置 权限 虚拟机 资源池 数据存储 网络 更新

## ▼ 存储

存储适配器

存储设备

主机缓存配置

协议端点

I/O 筛选器

## ▼ 网络

虚拟交换机

VMkernel 适配器

物理适配器

TCP/IP 配置

## ▼ 虚拟机

虚拟机启动/关机

代理虚拟机设置

默认虚拟机兼容性

交换文件位置

## ▼ 系统

许可

主机配置文件

时间配置

身份验证服务

证书

电源管理

高级系统设置

系统资源预留

防火墙

服务

安全配置文件

系统交换

软件包

## ▼ 硬件

## 存储适配器

+ 添加软件适配器 刷新 重新扫描存储... 重新扫描适配器

适配器	类型	状态	标识符	目标	设备	路径
型号: Emulex LPe1250 8Gb PCIe Fibre Channel Adapter						
vmhba2	光纤通道	联机	20:00:00:90:fa:5a:f3:86 10:00:00:90:fa:5a:f3:86	6	12	24
型号: iSCSI Software Adapter						
vmhba64	iSCSI	联机	iqn.1998-01.com.vmware:bbt-cloud-node-1-15005bfc	1	0	0
型号: LSI2008						
vmhba3	SAS	未知	--	1	1	1
型号: SB700 SATA Controller [IDE Mode]						
vmhba0	块 SCSI	未知	--	0	0	0

Copy All | 6 items

属性 设备 路径 动态发现 静态发现 网络端口绑定 高级选项

## 适配器状态

状态 已启用

禁用

## 常规

名称 vmhba64

型号 iSCSI Software Adapter

iSCSI 名称 iqn.1998-01.com.vmware:bbt-cloud-node-1-15005bfc

iSCSI 别名

目标发现 发送目标, 静态目标

编辑...

## 身份验证

方法 无

编辑...



## 案例：在vSphere中通过iSCSI使用IP-SAN (基于CentOS 8)



## 3.vSphere storage configuration

### 3.4 NFS storage

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



新建数据存储

1 类型  
2 选择 NFS 版本  
3 名称和配置  
4 主机的可访问性  
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。

容量	可用空间
271 GB	259.59 GB
271 GB	269.59 GB
271 GB	269.59 GB
271 GB	269.59 GB
31.24 TB	30.83 TB
46.93 TB	0 B
6.34 TB	0 B
6.02 TB	0 B
4.88 TB	0 B
31.25 TB	0 B
31.25 TB	0 B
31.25 TB	0 B

CANCEL BACK NEXT

## 案例：在vSphere中使用NFS网络存储





**DAS**



**SAN**



**NAS**





# SAN

- 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 将数据存储集中在共享存储中，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，以便可以像在本地的磁盘存储上一样在 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。



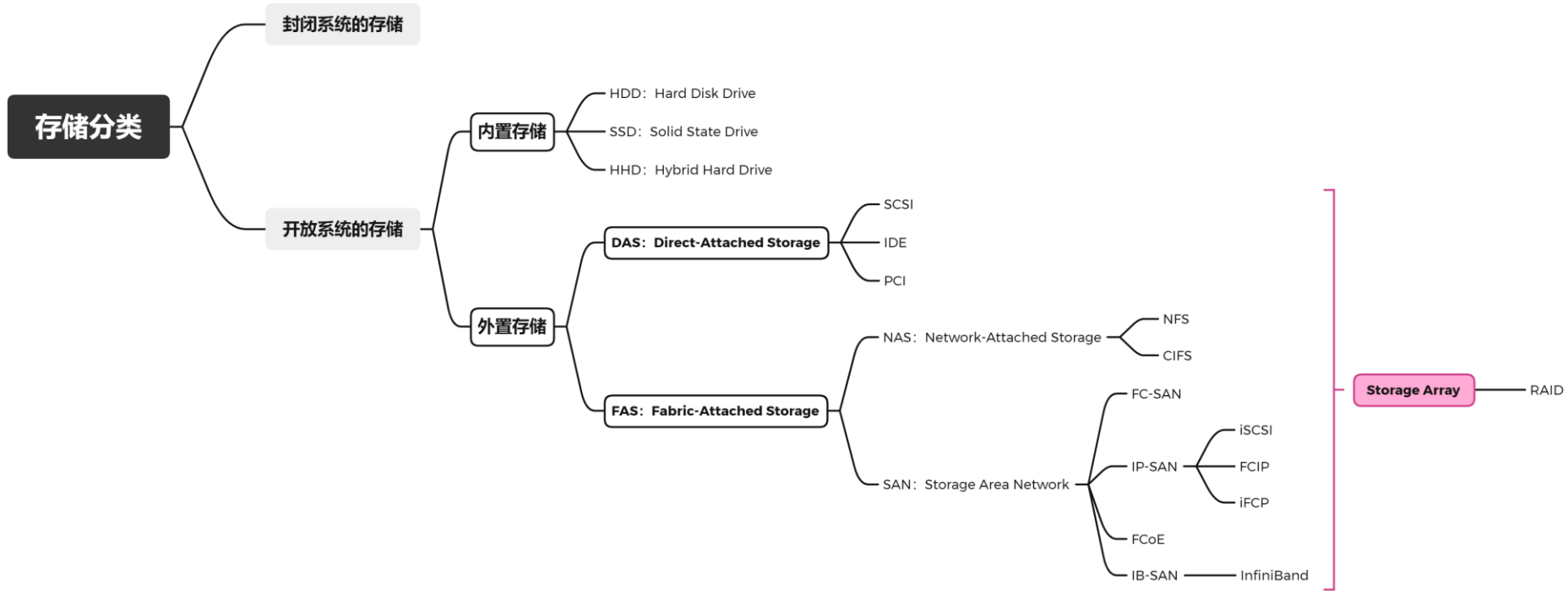




# NAS

- 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。





## 3.vSphere storage configuration

### 3.5 SIOC and storage DRS

- 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操作时，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。



vm vSphere Client 菜单 在所有环境中搜索

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-bcee7b12bdd6/

存储 可用 30.83 TB  
已用 419.27 GB 磁盘 31.24 TB 刷新

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...
- 挂载数据存储...
- 卸载数据存储...
- 配置 Storage I/O Control...
- 编辑空间回收...
- 升级到 VMFS-5...
- 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 使用情况和逐步实施预定义的 I/O 共享级别。

启用 Storage I/O Control 和统计信息收集

抖动阈值  吞吐盘峰值百分比 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](#)



策略和配置文件

虚拟机自定义视图

虚拟机存储策略

主机配置文件

存储策略组件

### 虚拟机存储策略

名称

- Host-local PMem Default Storage Policy
- VM Encryption Policy
- vSAN Default Storage Policy
- VMware No Requirements Policy

#### 创建虚拟机存储策略

- 名称和描述
- 策略结构
- 基于主机的服务
- 存储兼容性
- 检查并完成

#### 基于主机的服务

为主机提供的数据服务创建规则。可用数据服务包括 加密、I/O 控制、缓存等。除了所有数据存储服务特定规则，还将应用基于主机的服务。

加密 **Storage I/O Control**

禁用

使用存储策略组件 <选择组件>

自定义

提供程序: VMware Storage I/O Control

VMware Storage I/O Control ①

IOPS 限制 ①	-1
IOPS 预置 ①	1
IOPS 份额 ①	1000

取消 上一页 下一页

### Limits 限制

This is a hard cap on resources. The VM (or virtual disk) will not consume more than the resources configured by a limit.



**创建虚拟机存储策略**

1 名称和描述  
2 策略结构  
3 基于主机的服务  
4 存储兼容性  
5 检查并完成

**基于主机的服务**

为虚拟机提供的数据库服务创建规则。可用数据库服务包括 加密、I/O 控制、缓存等。除了所有数据库存储特定规则，还将应用基于主机的服务。

加密 Storage I/O Control

禁用

使用存储策略组件 <选择组件>

自定义

提供程序: VMware Storage I/O Control

VMware Storage I/O Control

IOPS 限制	-1
IOPS 预留	1
IOPS 份额	1000

取消 上一页 下一页

## Reservations 预留

Reservation is the number of resources that will always be available to the VM (or virtual disk respectively in this case). If the reserved resources are not used, the remaining number of resources will be used to satisfy the needs of other VMs, but once a machine with a reservation wants to access those resources, it will even if others are using them as well.



创建虚拟机存储策略

1 名称和描述

2 策略结构

3 基于主机的服务

4 存储兼容性

5 检查并完成

基于主机的服务

为主机提供的数据服务创建规则。可用数据服务包括 加密、I/O 控制、缓存等。除了所有数据存储特定规则，还将应用基于主机的服务。

加密 Storage I/O Control

禁用

使用存储策略组件 <选择组件>

自定义

提供程序: VMware Storage I/O Control

VMware Storage I/O Control

IOPS 限制	-1
IOPS 预留	1
IOPS 份额	1000

取消 上一页 下一页

### Shares 份额

Shares come into play only when congestion occurs. VMs want to access more resources than the storage can provide. In this situation, each VM will get a fair number of resources calculated using shares.



## 3.vSphere storage configuration

### 3.5 SIOC and storage DRS

- 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, 它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仅应用于数据存储群集, 不能在单个数据存储中使用。





# 3.vSphere storage configuration

## 3.6 Advanced Storage Features

- vSphere在存储管理上新增了许多新功能，提供了诸多高级特性：
  - Virtual Machine File System (VMFS) 6
    - <http://cormachogan.com/2017/08/16/vmfs-6-large-small-file-blocks>
  - Automatic space reclaim: 自动空间回收
  - Instant clones versus linked clones: 快速克隆与链接克隆的存储支持
  - RDM: 直接访问存储设备LUN
  - 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
  - VMware vSphere SPBM
  - Pluggable Storage Architecture (PSA)
  - 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
  - VASA: 存储管理集成编程接口
  - VVols: 虚拟卷



## 4. VMware vSAN

### 4.1 vSAN的方案

- 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，作为 ESXi Hypervisor 内的分布式软件层。
  - vSAN 可聚合本地数据存储设备或直接连接的数据存储设备，以创建在 vSAN 集群中的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并且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。
    - 此架构可创建用于软件定义数据中心 (Software-Define Data Center, SDDC) 且经过闪存优化的弹性共享数据存储。
  - vSAN 无需外部共享存储，并且通过基于存储策略的管理 (Storage Policy-Based Management, SPBM) 简化了存储配置。



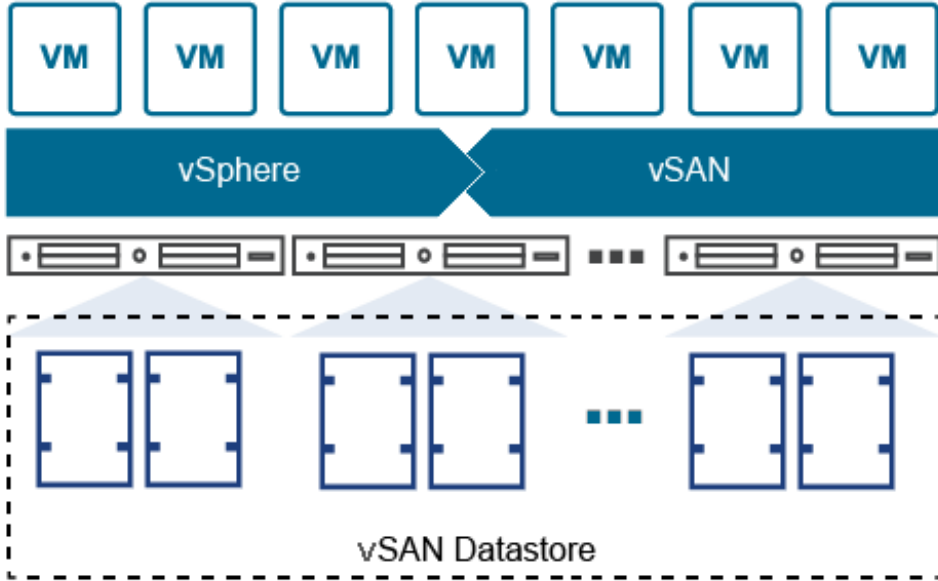


VMware vSphere + VSAN



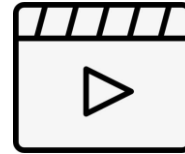
VSAN Datastore





## VMware vSAN:

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



VMware vSAN 介绍

<https://www.bilibili.com/video/BV1j4411K7ZV>

<https://www.bilibili.com/video/BV1Ls411K7AJ>



## 4. VMware vSAN

### 4.1 vSAN的方案

- vSAN实现的两种设备选型方案：
  - Hybrid deployment: 混合部署
    - A mixture of fast caching devices (SSDs) and magnetic devices for capacity.
    - In hybrid deployments, you can't use features such as erasure coding (RAID 5/6) or deduplication and compression.
  - All-flash: 全闪存
    - High-performance SSD is used as a caching device (usually NVMe)
    - standard SSDs are used for a capacity tier.
    - All-flash configuration provides unmatched I/O performance.
    - 推荐方案
  - 网络要求：
    - consider using dedicated pairs of 10 GbE network cards
    - isolating vSAN traffic and provide adequate bandwidth for vSAN traffic

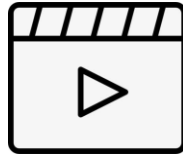


## 4. VMware vSAN

### 4.1 vSAN的配置

- 使用VMware 云中实验室进行vSAN的体验：
  - <https://www.vmware.com/cn/try-vmware/try-vmware-lightning-hands-on-labs.html>
  - 基于VMware WorkStation Pro进行vSAN实验





基于VMware WorkStation Pro

通过vSphere + vSAN实现超融合架构

<http://dms.it.hactcm.edu.cn/api/h/f?m=9929704efc050243-1-0>



## SDDC的虚拟化数据中心设计方案

**中央  
存储**

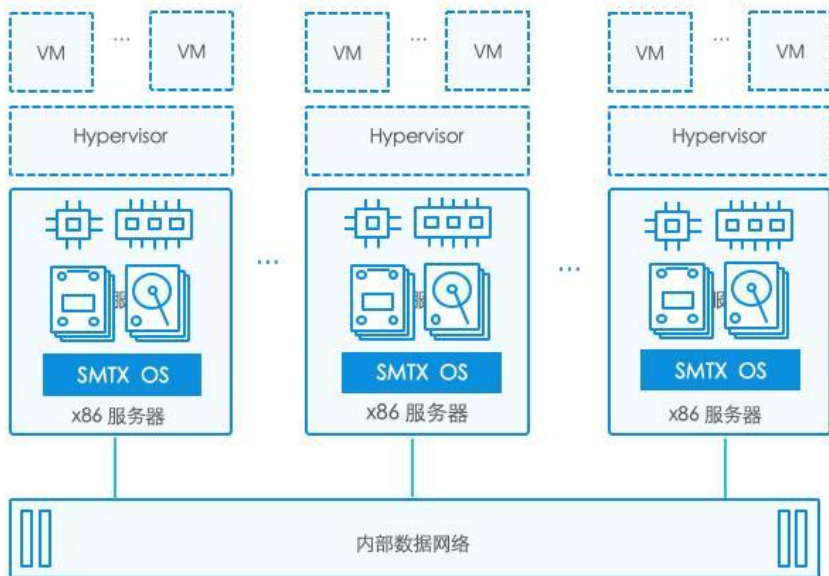
虚拟化服务器  
存储阵列  
虚拟化软件

**超  
融合**

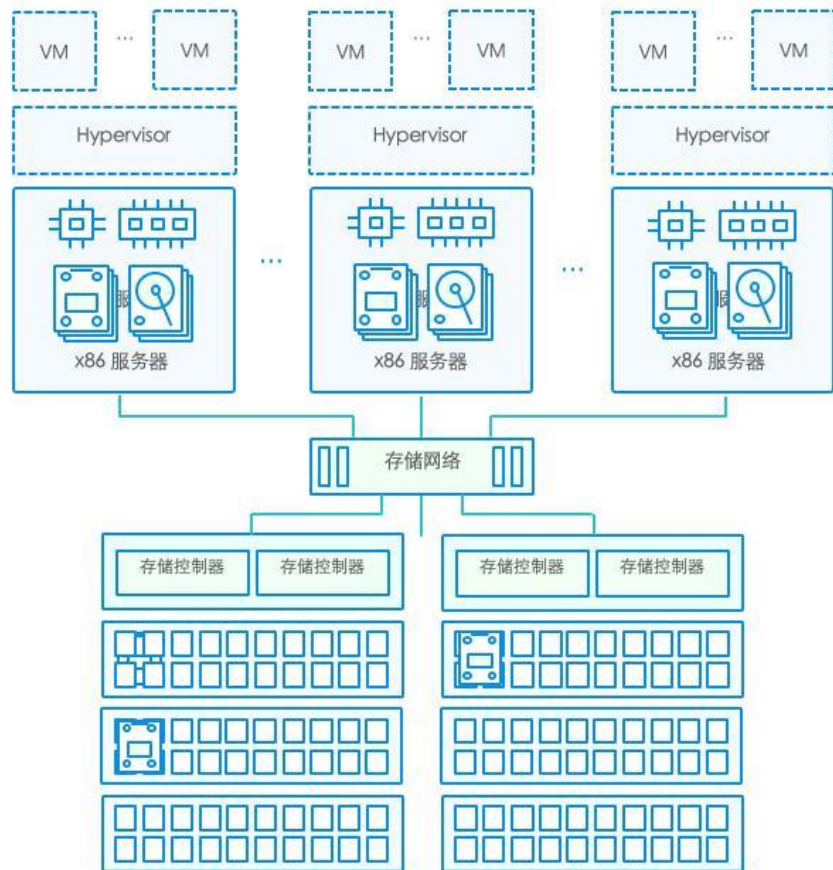
虚拟化服务器  
虚拟化软件(vSAN)







超融合的架构



传统架构



