

## 服务器技术基础



# 前言

- 服务器是一切服务平台的立命之本，云计算也不例外，那么什么是服务器？它有哪些关键技术需要我们了解？本章将一一为您揭开，云计算基础技术的学习也将从这里启程。

# 目标

- 学完本课程后，您将能够：
  - 了解什么是服务器
  - 了解服务器的类型
  - 认识服务器的硬件组成与基本原理
  - 熟悉服务器的关键技术

# 目录

---

## 1. 服务器介绍

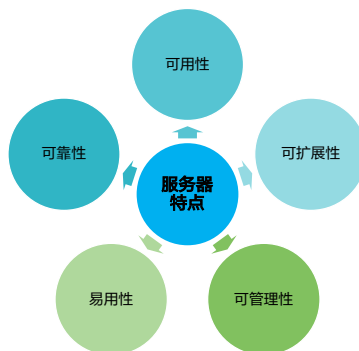
- 什么是服务器
  - 服务器发展历程
  - 服务器的类型
  - 服务器硬件介绍

## 2. 服务器关键技术

# 服务器定义和特点

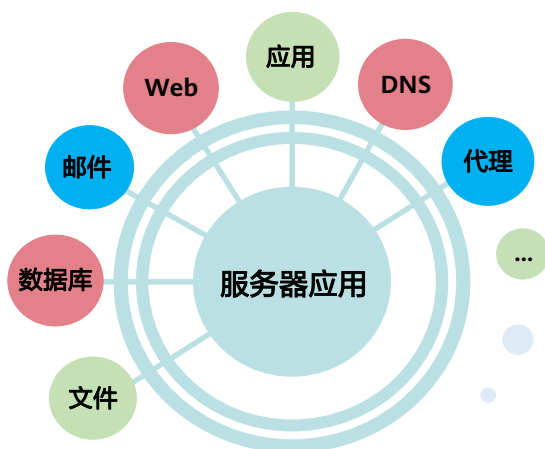
- 服务器定义

- 服务器是计算机的一种。它比普通计算机运行速度更快、负载更高且价格更高
- 服务器是为用户提供服务的计算机，通常分为文件服务器、数据库服务器和应用程序服务器



- 服务器是20世纪90年代迅速发展的主流计算产品，能为网络用户提供集中计算、信息发布及数据管理等服务，也可以将与其相连的如硬盘、打印机、Modem等各种专用通讯设备给网络上的用户进行共享。
- 服务器的主要特点包括：
  - R: Reliability—可靠性，能连续正常运行多长时间；
  - A: Availability—可用性，系统正常运行时间和使用时间的百分比；
  - S: Scalability—可扩展性，包括两方面，一方面是硬件的可扩展性，另一方面是软件对操作系统的支持能力；
  - U: Usability—易用性，服务器的硬件和软件易于维护和修复；
  - M: Manageability—可管理性，对服务器运行情况能进行监控，报警，对一些故障的自动智能化处理。

## 服务器使用场景



- 服务器已经广泛应用在电信运营商、政府、金融、教育、企业、电子商务等各个行业领域，为用户提供文件、数据库、邮件、Web等服务。
- 服务器应用部署架构：
  - C/S：Client/Server的缩写。通常也称为客户端/服务器架构。服务器端运行服务端程序，客户端安装客户端软件。在此架构里服务端和客户端分别完成不同的任务，客户端处理用户的前端界面和交互操作，服务端处理后台业务逻辑和请求数据，这使得两端的通讯速度和通讯的效率大大提高。例如，我们在文件服务器上（服务器端）安装vsftpd程序，并启动服务；在用户的计算机中安装FileZilla或WinSCP等客户端工具后，用户就可以通过客户端工具进行文件的上传和下载。
  - B/S：Browser/Server的缩写。通常也称为浏览器/服务器架构。B/S架构中，用户只需安装浏览器即可，而将应用逻辑集中在服务器和中间件上，可以提高数据处理性能。例如，我们在访问一个网站时，只需要在自己的浏览器中输入网站的域名，如www.huawei.com，就可以看到该网站的后台服务器给我们提供的Web服务。而网站的后台服务器端可能有很多服务器提供服务，如数据库服务，代理服务，缓存服务等等，这些都不需要用户关心，用户只需要通过浏览器输入网址就可以看到相应的界面。

# 目录

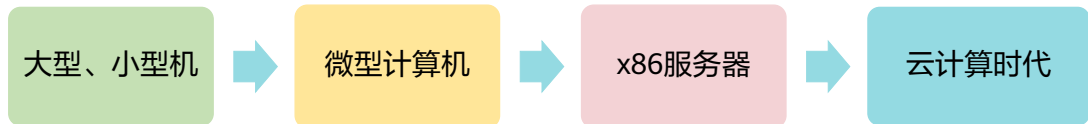
---

## 1. 服务器介绍

- 什么是服务器
- 服务器发展历程
- 服务器的类型
- 服务器硬件介绍

## 2. 服务器关键技术

## 服务器发展历程



- 大型主机阶段

- 20世纪40-50年代，是第一代电子管计算机。经历了电子管数字计算机、晶体管数字计算机、集成电路数字计算机和大规模集成电路数字计算机的发展历程，计算机技术逐渐走向成熟；

- 小型计算机阶段

- 20世纪60-70年代，是对大型主机进行的第一次“缩小化”，可以满足中小企业事业单位的信息处理要求，成本较低，价格可被接受；

- 微型计算机阶段

- 20世纪70-80年代，是对大型主机进行的第二次“缩小化”，1976年美国苹果公司成立，1977年就推出了Apple II计算机，大获成功。1981年IBM推出IBM-PC，此后它经历了若干代的演进，占领了个人计算机市场，使得个人计算机得到了很大的普及；

- x86服务器时代

- 1978年，英特尔推出第一代x86架构处理器—8086中央处理器。
- 1993年，英特尔正式推出Pentium(奔腾)系列，该系列的推出，将x86架构处理器带上了一个新的性能高度。
- 1995年，英特尔推出Pentium Pro—为服务器而生的x86处理器，从此开启了x86的至强时代，其标准化开放性也促成了市场发展，为云计算时代打下了坚实基础。



- 云计算时代

- 从2008年起，云计算（Cloud Computing）概念逐渐流行起来，它正在成为一个通俗和大众化（Popular）的词语。云计算被视为“革命性的计算模型”，因为它使得超级计算能力通过互联网自由流通成为了可能。企业与个人用户无需再投入昂贵的硬件购置成本，只需要通过互联网来购买租赁计算力，用户只为自己需要的功能付费。云计算让用户脱离技术与部署上的复杂性而获得应用。云计算囊括了开发、架构、负载平衡和商业模式等，是软件业的未来模式。

## 计算产业在进化：从计算1.0到计算3.0



- 计算产业，历经近半个世纪的发展，持续改变着社会，改变着产业；但计算产业本身，也在不断的进化。
- 从最早期的大型机、小型机时代，是专用计算，称之为计算1.0；到了x86时代，在Intel的带领下，在摩尔定律的驱动下，计算由专用走向了通用，大量数据中心开始出现，也是当前计算产业所处的阶段，称之为计算2.0；随着数字化程度的加速发展，世界逐步走向智能化，计算已经不仅仅局限于数据中心，也开始走向全栈全场景，称之为计算3.0时代，而这个时代的主要特征就是“智能”，所以也称之为“智能计算”。

# 目录




---

## 1. 服务器介绍

- 什么是服务器
- 服务器发展历程
- 服务器的类型
- 服务器硬件介绍

## 2. 服务器关键技术

## 服务器分类方式 - 硬件形态

硬件形态		
塔式服务器	机架服务器	刀片服务器
		

- **塔式服务器：**

- 有的塔式服务器采用大小与普通立式计算机大致相当的机箱，有的采用大容量的机箱，像个硕大的柜子，可以安装的扩展板卡以及硬盘的数量都比较多，因此，当用户应用不断递增时，通过添加组件方式提升服务器的处理能力，从而可以有效的满足用户需求的增长，并保护了用户的投资。

- **机架服务器：**

- 机架式服务器的外形看来不像计算机，而像交换机，有1U（1U=1.75英寸）、2U、4U等规格。机架式服务器安装在标准的19英寸机柜里面。这种结构的多为功能型服务器。特点是机箱尺寸比较小巧，在机柜中可以同时放置多台服务器，从而获得更高的处理能力。

- **刀片服务器：**

- 特点是每个服务器都是一个插板，在插板上配备有处理器、内存、硬盘以及相关组件。由于刀片服务器的特殊架构，所以刀片服务器的使用还需要与刀片服务器的专用机箱结合，通常在一个机箱中可以容纳几个到几十个刀片服务器，所以对于高性能计算、多种应用的前台服务器，应用服务器，以及后台的中心数据库应用都可以满足。

- 大型机和小型机请参考前文描述。

# 目录

---

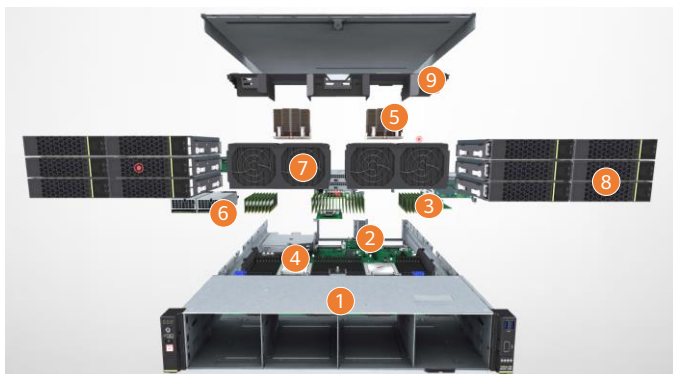
## 1. 服务器介绍

- 什么是服务器
- 服务器发展历程
- 服务器的类型
  - 服务器硬件介绍

## 2. 服务器关键技术

## 服务器硬件结构

- 以华为TaiShan 200服务器的硬件结构为样例



- 1 机箱
- 2 主板
- 3 内存
- 4 CPU
- 5 散热器
- 6 电源
- 7 风扇
- 8 硬盘
- 9 导风罩

- 华为TaiShan 200服务器3D模型展示：  
<https://info.support.huawei.com/computing/server3D/res/server/taishan2280/index.html?lang=cn>。

## CPU定义和组成

- CPU的定义

- CPU: Central Processing Unit, 中央处理器, 是一台计算机的运算核心和控制核心
- CPU、内部存储器和输入/输出设备是电子计算机三大核心部件
- CPU主要功能是解释计算机指令以及处理计算机软件中的数据

- CPU的组成

- CPU由逻辑运算单元、控制单元和存储单元构成



- CPU: 是服务器上的核心处理单元, 而服务器是网络中的重要设备, 要处理大量的访问需求。因此对服务器具有大数据量的快速吞吐、超强的稳定性、长时间运行等严格要求。所以说CPU是计算机的“大脑”, 是衡量服务器性能的首要指标。
- 控制器: 计算机是根据事先存储的程序对全机实行控制, 而程序是指能实现某一功能的指令序列, 控制器就是根据指令来对各种逻辑电路发布命令的机构, 它是计算机的指挥中心, 控制整个CPU的工作, 决定计算机运行过程的自动化。
- 运算器: 计算机中执行各种算术和逻辑运算操作的部件。运算器的基本操作包括加、减、乘、除四则运算, 与、或、非、异或等逻辑操作, 以及移位、比较和传送等操作, 亦称算术逻辑部件。
- 寄存器: 寄存器的主要作用是用来暂时存放参与运算的数据和运算结果, 具有接收数据、存放数据和输出数据的功能。

## CPU的频率

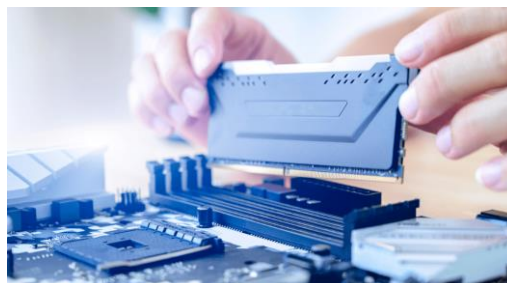
- 主频
  - 主频也叫时钟频率，单位是兆赫（MHz）或吉赫（GHz），用来表示CPU的运算、处理数据的速度
- 外频
  - 外频是CPU的基准频率，单位是MHz。CPU的外频决定着整块主板的运行速度
- 总线频率
  - 总线频率直接影响CPU与内存数据交换的速度
- 倍频系数
  - 倍频系数是指CPU主频与外频之间的相对比例关系



# 内存

- 内存定义

- 存储器按用途可分为主存储器和辅助存储器。主存储器又称内存（简称内存），是CPU能直接寻址的存储空间
- 内存的作用是用于暂时存放CPU中的运算数据，以及与硬盘等外部存储器交换的数据
- 内存是计算机中重要的部件之一，它是与CPU进行沟通的桥梁
- 内存由内存芯片、电路板、金手指等部分组成



- 在计算机的组成结构中，有一个很重要的部分就是存储器。存储器是用来存储程序和数据的部件，对于计算机来说，有了存储器，才有记忆功能，才能保证正常工作。
- 内存是电脑中的主要部件，它是相对于外存而言的。我们平常使用的程序，如Windows操作系统、打字软件、游戏软件等，一般都是安装在硬盘等外存上的，必须把它们调入内存中运行，才能真正使用其功能。我们平时输入一段文字，或玩一个游戏，其实都是在内存中进行的。就好比在一个书房里，存放书籍的书架和书柜相当于电脑的外存，而我们工作的办公桌就是内存。通常我们把要永久保存的、大量的数据存储在外部存储设备上，而把一些临时的或少量的数据和程序放在内存上。
- 内存是计算机中重要的部件之一，它是与CPU进行沟通的桥梁。计算机中所有程序的运行都是在内存中进行的，因此内存的性能对计算机的影响非常大。内存由内存芯片、电路板、金手指等部分组成。
- 服务器内存插槽及配置原则：
  - 同一台服务器必须使用相同型号的DIMM。
  - CPU对应的内存槽位上必须至少配置一根内存条。
  - 当服务器配置完全平衡的内存条时，可实现最佳的内存性能。不平衡配置会降低内存性能，因此不推荐使用。

## 硬盘简介

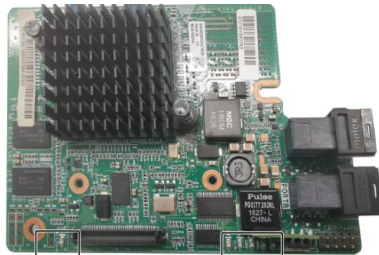
- 硬盘是计算机最主要的存储设备
- 硬盘接口是硬盘与主机系统间的连接部件，作用是在硬盘缓存和主机内存之间传输数据。不同的硬盘接口决定着硬盘与计算机之间的连接速度，直接影响着程序运行快慢和系统性能好坏

	SATA	SAS	NL-SAS	SSD
主流转速 (RPM)	7,200	15,000/10,000	7,200	NA
串行/并行	串行	串行	串行	串行
主流容量 (TB)	1T/2T/3T	0.6T/0.9T	2T/3T/4T	0.6T/0.8T/1.2T/1.6T
MTBF (h)	1,200,000	1,600,000	1,200,000	2,000,000
备注	由ATA硬盘发展而来，采用串行的方式传输，SATA 3.0实现600 MB/s最高数据传输率  SATA硬盘年故障率大约2%	SAS专为满足高性能企业需求而设计，并且兼容SATA硬盘。能够提供3.0 Gb/s到6.0 Gb/s的传输率，未来规划到12.0 Gb/s  SAS硬盘年故障率小于2%	带有SAS接口的“企业级SATA驱动器”，适用于在一个磁盘阵列系统中实现分级存储，简化了磁盘阵列系统的设计  NL-SAS硬盘年故障率大约2%	固态硬盘是用固态电子存储芯片阵列而制成的硬盘，由控制单元和存储单元（FLASH芯片、DRAM芯片）组成  固态硬盘的接口规范和定义、功能及使用方法与普通硬盘完全相同，在产品外形和尺寸上也完全与普通硬盘一致

- MTBF: Mean Time Between Failure，平均无故障工作时间。
- 在价格方面，通常来说SATA和NL-SAS硬盘较为便宜，SAS硬盘较贵一些，SSD最为昂贵。

## RAID卡

- RAID卡又名磁盘阵列卡，简称阵列卡
- RAID卡的作用：
  - 可以将若干硬盘驱动器按照一定要求组成一个整体、由阵列控制器管理的系统
  - 可以提高磁盘子系统的性能及可靠性



LSI SAS3108

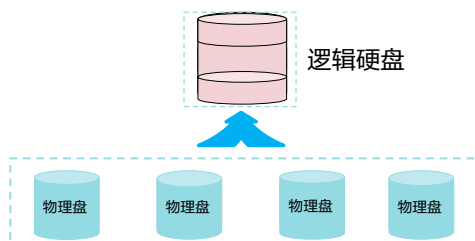
- LSI SAS3108 RAID控制卡用户指南：  
<https://support.huawei.com/enterprise/zh/doc/EDOC1100048779/6e4f4d85>。

# RAID定义

- RAID定义

- RAID: Redundant Array of Independent Disks, 独立硬盘冗余阵列, 旧称廉价磁盘冗余阵列 (Redundant Array of Inexpensive Disks), 简称磁盘阵列。利用虚拟化存储技术把多个硬盘组合起来, 成为一个或多个硬盘阵列组, 目的为提升性能或数据冗余, 或是两者同时提升

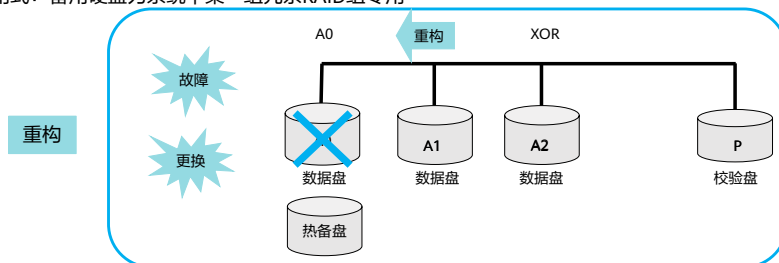
——维基百科



- 关于RAID技术原理, 我们会在存储基础技术中详细讲解, 这里了解概念、定义即可。

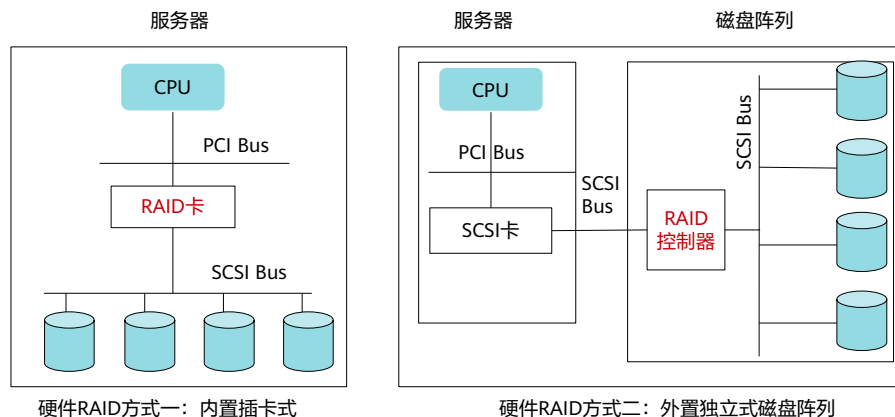
## RAID热备和重构概念

- 热备（Hot Spare）的定义
  - 当冗余的RAID组中某个硬盘失效时，在不干扰当前RAID系统正常使用的情况下，用RAID系统中另外一个正常的备用硬盘自动顶替失效硬盘，及时保证RAID系统的冗余性
- 热备一般分为两种
  - 全局式：备用硬盘为系统中所有的冗余RAID组共享
  - 专用式：备用硬盘为系统中某一组冗余RAID组专用



- 数据校验，利用冗余数据进行数据错误检测和修复，冗余数据通常采用海明码、异或操作等算法来计算获得。利用校验功能，可以很大程度上提高磁盘阵列的可靠性、高性能和容错能力。不过，数据校验需要从多处读取数据并进行计算和对比，会影响系统性能。
- 一般来说，RAID不可作为数据备份的替代方案，它对非磁盘故障等造成的数据丢失无能为力，比如病毒、人为破坏、意外删除等情形。此时的数据丢失是相对操作系统、文件系统、卷管理器或者应用系统来说的，对于RAID而言，数据都是完好的，没有发生丢失。所以，数据备份、灾备等数据保护措施是非常必要的，与RAID相辅相成，保护数据在不同层次的安全性，防止发生数据丢失。

## RAID的实现 - 硬件方式



- 基于硬件的RAID是利用硬件RAID适配卡来实现的。
- 硬件RAID又可分为内置插卡式和外置独立式磁盘阵列。
- RAID卡上集成了处理器，能够独立于主机对RAID存储子系统进行控制。因为拥有自己独立的处理器和存储器，RAID卡可以自己计算奇偶校验信息并完成文件定位，减少对主机CPU运算时间的占用，提高数据并行传输速度。

## RAID的实现 - 软件方式

- 定义
  - 软件RAID是指在操作系统中，通过安装软件的方式来实现RAID相应的功能
- 特点
  - 软件RAID不需要昂贵的RAID控制卡，提供了廉价的解决办法
  - RAID功能完全依靠CPU执行，主机的CPU占用严重，如RAID 5的大量异或（XOR）操作

- 软件RAID中不能提供如下功能：
  - 硬盘热拔插；
  - 硬盘热备份；
  - 远程阵列管理；
  - 可引导阵列支持；
  - 在硬盘上实现阵列配置；
  - SMART硬盘支持。

## RAID的实现 - 方式比较

方式	软件RAID	内置插卡式RAID	外置独立磁盘阵列RAID
特点	全部RAID功能靠CPU执行，主机CPU占用严重，系统性能下降	减少了密集RAID操作对主机CPU的占用，从而提高了性能	通过标准控制器与服务器相连，RAID功能全部由外部RAID存储子系统上的微处理器来实现，独立于操作系统
优点	<ul style="list-style-type: none"><li>实现成本低</li><li>配置灵活</li></ul>	<ul style="list-style-type: none"><li>具有数据保护和高速特点</li><li>与软件RAID相比有更强的容错性和更好的性能</li><li>与外置相比价格相对低廉</li><li>支持可引导阵列</li></ul>	<ul style="list-style-type: none"><li>为高端服务器提供超大容量存储系统</li><li>可配置双控制器，提高数据吞吐率，或供双机共享存储</li><li>支持热插拔</li><li>更好的可扩展性</li></ul>



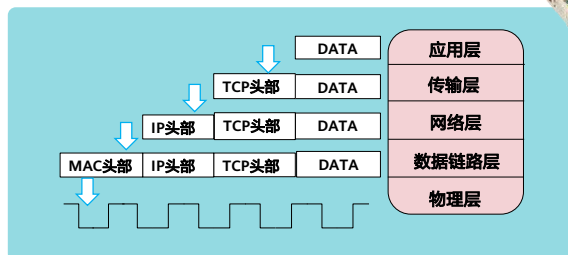
# 网卡定义和功能

- 网卡定义

- 网卡，又称为网络适配器或网络接口卡NIC（Network Interface Card），是构成计算机网络系统中最基本的、最重要的、必不可少的连接设备，计算机主要通过网卡接入网络

- 网卡主要功能

- 代表固定的网络地址
- 数据的发送与接收
- 数据的封装与解封
- 链路管理
- 编码与译码



## 华为服务器网卡

- 板载网卡
  - 集成在服务器主板的PCH芯片中，不可更换
  - 对外提供2个GE电口+2个10G光口/电口，集成网卡不占用PCIe插槽
- PCIe标卡
  - 华为有自研及外购的PCIe标卡网卡，可以配置在标准PCIe插槽中
- 灵活IO插卡
  - 华为自研网卡，非标准PCIe卡形态，只能用于华为机架服务器
- Mezz卡
  - 刀片服务器计算节点专用Mezz卡，只能用于华为E9000刀片服务器



板载网卡



PCIe标卡



灵活IO插卡



Mezz卡

- PCIe（PCI-Express）是继ISA和PCI总线之后的第三代I/O总线，即3GIO。由Intel在2001年的IDF上提出，由PCI-SIG（PCI特殊兴趣组织）认证发布后才改名为“PCI-Express”。它的主要优势就是数据传输速率高，另外还有抗干扰能力强，传输距离远，功耗低等优点。
- PCIe标卡：对于华为服务器，指的是位于PCIe插槽的网卡。
- 下方链接为PCIe卡装卸视频，可以对PCIe卡有更直观的认识。

<https://support.huawei.com/enterprise/zh/doc/EDOC1100002168?section=o00d>

## 电源和风扇

- 支持服务器的电力负载
- 支持冗余，防止电源故障
  - 故障预警和防止
  - 故障之前的预防性维护
  - 保证服务器持续运行
- 电源子系统包括：
  - 智能电源
  - 风扇
- 冗余电源和风扇



电源模块



风扇模块

- 电源冗余特性：
  - 1+1，此时每个模块承担50%的输出功率，当一个模块拔出时，另一个模块承担100%输出功率。
  - 2+1，有三个模块，每个模块承担输出功率的1/3，拔出一个模块，其余两个模块各承担50%的输出功率。

# 目录

---

## 1. 服务器介绍

## 2. 服务器关键技术

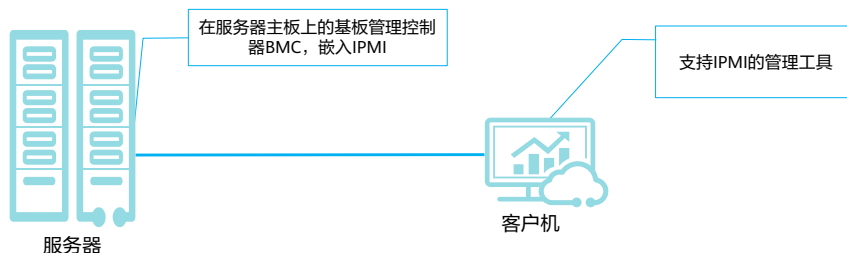
- BMC介绍

- BIOS

# 什么是IPMI

- IPMI定义

- IPMI: Intelligent Platform Management Interface, 智能平台管理接口, 是一种开放标准的硬件管理接口规格, 定义了嵌入式管理子系统进行通信的特定方法
- IPMI信息通过基板管理控制器 ( BMC ) 进行交流, 使用低级硬件智能管理而不使用操作系统进行管理



- IPMI是智能型平台管理接口 ( Intelligent Platform Management Interface ) 的缩写, 是管理基于Intel结构的企业系统中所使用的外围设备采用的一种工业标准, 该标准由英特尔、惠普、NEC、美国戴尔电脑和SuperMicro等公司制定。用户可以利用IPMI监视服务器的物理健康特征, 如温度、电压、风扇工作状态、电源状态等。而且更为重要的是IPMI是一个开放的免费标准, 用户无需为使用该标准而支付额外的费用。
- IPMI的发展:
  - 1998年Intel、DELL、HP及NEC共同提出IPMI规格, 可以透过网路远端控制温度、电压。
  - 2001年IPMI从1.0版改版至1.5版, 新增PCI Management Bus等功能。
  - 2004年Intel发表了IPMI 2.0的规格, 能够向下相容IPMI 1.0及1.5的规格。新增了Console Redirection, 并可以通过Port、Modem以及LAN远端管理服务器, 也加强了安全、VLAN 和刀片服务器的支持性。

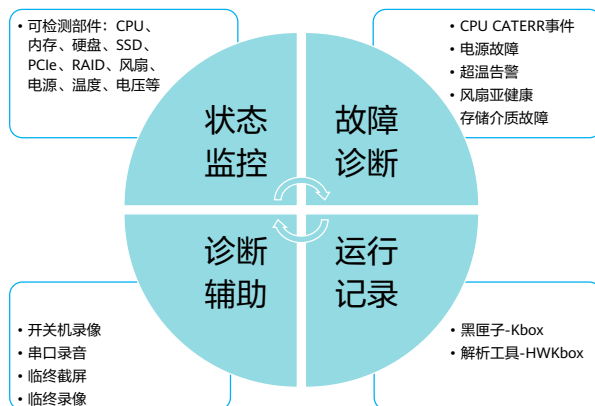
## BMC介绍

- BMC定义
  - BMC: Baseboard Management Controller, 主板管理控制单元, 是IPMI规范的核心, 负责各路传感器的信号采集、处理、储存, 以及各种器件运行状态的监控。BMC向机框管理板提供被管理对象的硬件状态及告警等信息, 从而实现机框管理模块对被管理对象的设备管理功能

- BMC主要实现以下功能:
  - 远程控制;
  - 告警管理;
  - 状态检测;
  - 设备信息管理;
  - 散热控制;
  - 支持ipmitool工具;
  - 支持Web界面管理;
  - 支持集中账号管理。

## iBMC介绍

- 华为服务器智能管理系统（Huawei Intelligent Baseboard Management Controller，iBMC）是具有完全自主知识产权，面向服务器全生命周期的服务器嵌入式管理系统



- iBMC提供硬件状态监控、部署、节能、安全等系列管理工具，标准化接口构建服务器管理更加完善的生态系统。iBMC基于华为自研的管理芯片Hi1710，采用多项创新技术，全面实现服务器的精细化管理。
- iBMC提供了丰富的用户接口，如命令行、基于Web界面的用户接口、IPMI集成接口、SNMP集成接口、Redfish集成接口，并且所有用户接口都采用了认证机制和高度安全的加密算法，保证接入和传输的安全性。

# 目录

---

## 1. 服务器介绍

## 2. 服务器关键技术

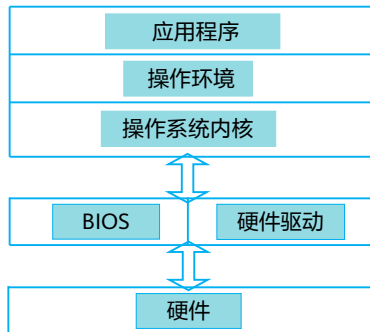
- BMC介绍

- BIOS



## BIOS简介

- BIOS：Basic Input/Output System，基本输入输出系统
- BIOS实际上是固化到计算机中的一组程序，为计算机系统提供最底层的、最直接的硬件控制服务，解决硬件的实时需求



BIOS功能：

- ▣ 检测和初始化硬件
- ▣ 操作系统引导
- ▣ 高级电源管理

- BIOS是系统内核和硬件层之间桥梁。
- BIOS特点：
  - ▣ 软件升级、加载和装载功能；
  - ▣ 基本OAM功能；
  - ▣ 串口管理功能；
  - ▣ 故障恢复功能；
  - ▣ ECC管理功能；
  - ▣ 硬件诊断功能。

## 思考题

1. （多选）以下关于华为服务器网卡的描述，正确的是哪些项？
  - A. 板载网卡集成在服务器主板的PCH芯片中，不可更换
  - B. 华为自研的PCIe标卡网卡，可以配置在标准PCIe插槽中
  - C. 灵活IO插卡集成在服务器面板上，可用于前端业务连接
  - D. Mezz卡可用于华为机架服务器
2. （判断）主板管理控制单元（BMC）是IPMI规范的核心，负责各路传感器的信号采集、处理、储存，以及各种器件运行状态的监控。
  - A. 正确
  - B. 错误

- 参考答案：

- AB
- A

## 本章总结

- 本章我们一起学习了服务器的基本概念和发展历程，掌握了服务器的硬件组成和基本原理，了解了服务器中涉及到的关键技术。在接下来的课程中我们将继续学习基础技术中的存储技术部分，敬请期待。

## 学习推荐

- 华为Learning网站
  - <http://support.huawei.com/learning/Index!toTrainIndex>
- 华为Support案例库
  - <http://support.huawei.com/enterprise/servicecenter?lang=zh>

## 缩略语

- BIOS: Basic Input/Output System, 基本输入输出系统
- BMC: Baseboard Management Controller, 主板管理控制单元
- B/S: Browser/Server, 通常也称为浏览器/服务器架构
- C/S: Client/Server, 通常也称为客户端/服务器架构
- CPU: Central Processing Unit, 中央处理器
- iBMC: Huawei Intelligent Baseboard Management Controller, 华为服务器智能管理系统
- IPMI: Intelligent Platform Management Interface, 智能平台管理接口
- MTBF: Mean Time Between Failure, 平均无故障工作时间
- NIC: Network Interface Card, 网络接口卡
- RAID: Redundant Array of Independent Disks, 独立硬盘冗余阵列

# Thank you.

把数字世界带入每个人、每个家庭、  
每个组织，构建万物互联的智能世界。

Bring digital to every person, home, and  
organization for a fully connected,  
intelligent world.

Copyright©2023 Huawei Technologies Co., Ltd.  
All Rights Reserved.

The information in this document may contain predictive statements including, without limitation, statements regarding the future financial and operating results, future product portfolio, new technology, etc. There are a number of factors that could cause actual results and developments to differ materially from those expressed or implied in the predictive statements. Therefore, such information is provided for reference purpose only and constitutes neither an offer nor an acceptance. Huawei may change the information at any time without notice.



## 存储技术基础



# 前言

- 数据是用户最重要的资产，在云计算中，数据如何被保存？保存数据的介质有哪些？有哪些数据存储的关键技术？本章将一一为您揭开。



# 目标

- 学完本课程后，您将能够：
  - 了解主流的数据存储方式及组网
  - 掌握RAID和华为RAID 2.0+块级虚拟化技术
  - 区分集中式存储和分布式存储
  - 了解存储的相关协议及适用场景

# 目录

---

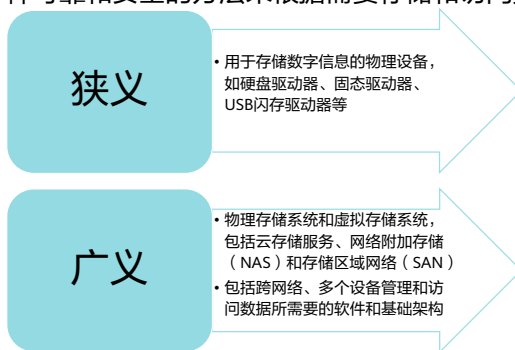
## 1. 存储基础介绍

- 什么是存储
- 存储发展历程
- 主流硬盘类型
- 存储组网类型
- 存储形态简介

## 2. 存储关键技术

## 什么是存储

- 存储是指在计算机或其他电子设备上存储和管理数字信息的过程。它既可以包括物理存储设备，如硬盘和固态硬盘，也可以包括虚拟存储系统，如云存储服务
- 存储的目标是提供一种可靠和安全的方法来根据需要存储和访问数据



- 狭义上的存储：CD、DVD、ZIP，磁带，硬盘等存储介质
- 广义上的存储：
  - 存储硬件系统（磁盘阵列，控制器，磁盘柜，磁带库等）
  - 存储软件（备份软件，管理软件，快照，复制等增值软件）
  - 存储网络（HBA卡，光纤交换机，FC/SAS线缆等）
  - 存储解决方案（集中存储，归档，备份，容灾等）

# 目录

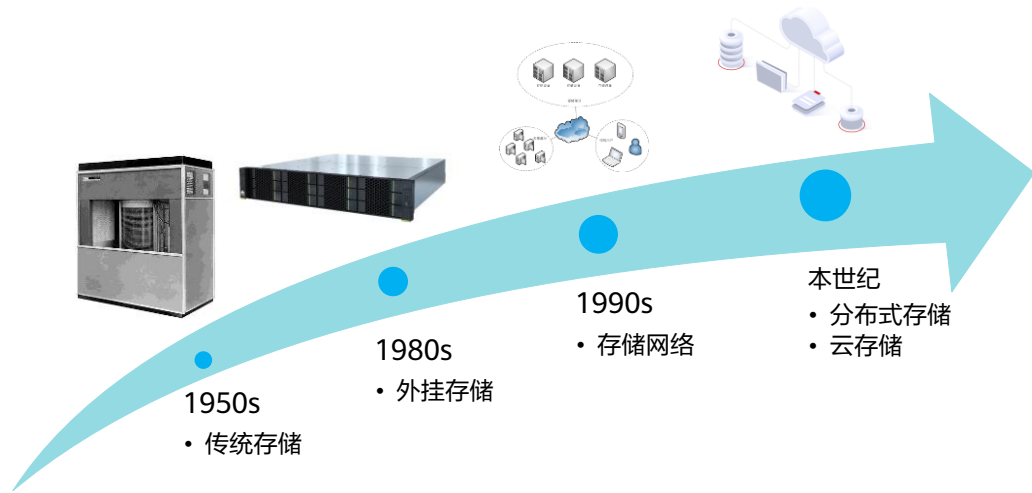
---

## 1. 存储基础介绍

- 什么是存储
- 存储架构历程
- 主流硬盘类型
- 存储组网类型
- 存储形态简介

## 2. 存储关键技术

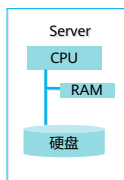
## 存储架构历程



- 存储架构的发展历程经历了传统存储、外挂存储、存储网络、分布式存储和云存储几个阶段。
- 传统存储是由单硬盘组成，1956年IBM发明了世界上第一块机械硬盘，这个硬盘有50个24英寸盘片，容量只有5 MB。体积与两台冰箱差不多大，重量超过一吨。被用于当时的工业领域，且独立于主机以外。
- 外挂存储即直连存储，最早的形态是JBOD，仅仅是将一些磁盘串联在一起，被称为JBOD（Just a Bunch Of Disks，磁盘簇），主机看到的就是一堆独立的硬盘，只增加了容量，无法提供安全保障。
- SAN是典型的存储网络，主要使用FC网络传输数据，随后出现了IP存储区域网络。
- 分布式存储采用通用服务器硬件构建存储资源池，更适合云计算的场景。在后续的課程中会进行介绍。

# 从硬盘到磁盘阵列

硬盘在服务器内部

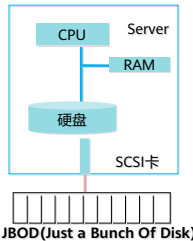


## 局限性:

- 硬盘成为系统性能瓶颈
- 有限硬盘槽位，容量小
- 单硬盘存放数据，可靠性差
- 存储空间利用率低
- 本地存储，数据分散



早期外挂存储



JBOD (Just a Bunch Of Disk)

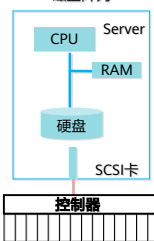
在逻辑上把几块物理磁盘串联在一起，其目的纯粹是为了增加磁盘的容量

## 解决问题:

- 有限硬盘槽位，容量小



磁盘阵列



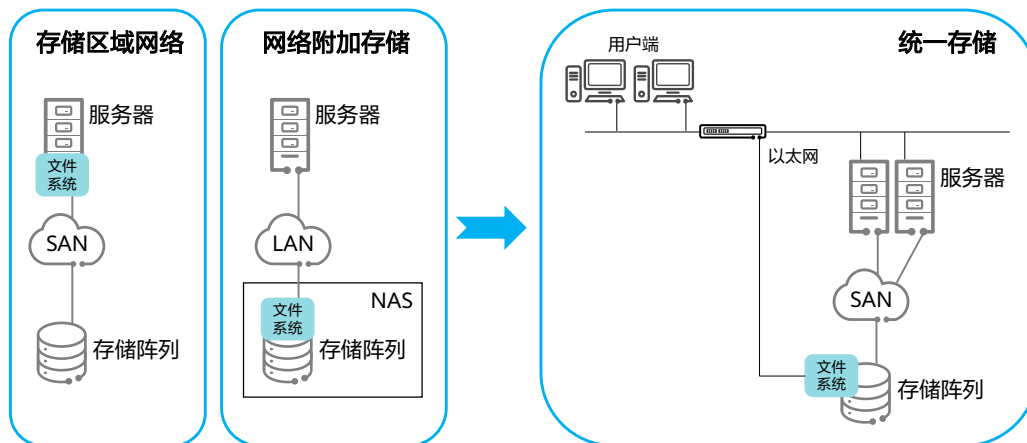
控制器中包含RAID功能、大容量Cache，同时使得磁盘阵列具有多种实用的功能，如提高数据的读写性能、数据的安全性等

## 解决问题:

- 硬盘成为系统性能瓶颈
- 有限硬盘槽位，容量小
- 单硬盘存放数据，可靠性差，读写性能低



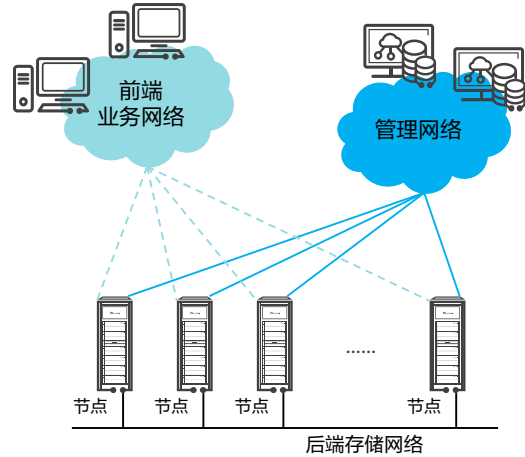
## 从分离到融合



- DAS ( Direct-Attached Storage，直连式存储 ) 的特点：
  - 数据分散
  - 存储空间利用率低
- 存储发展需求：
  - 数据共享
  - 提升资源利用率
  - 距离延长
- 网络的出现给存储带来了新的生机
  - SAN ( Storage Area Network，存储区域网络 )：在存储设备与服务器之间组成网络，主要提供块存储服务。
  - NAS ( Network Attached Storage，网络附加存储 )：使用网络连接服务器与带有文件系统的存储，主要提供文件存储服务。
- 2011年左右，支持SAN和NAS多协议的统一存储迅速成为热点，融合的趋势再次出现：通过整合降低成本，将NAS和SAN结合起来，同时提供数据库与文件共享服务，简化存储管理，提高存储利用率。

## 分布式存储

- 通过软件将物理资源组织起来构成高性能逻辑存储池，在保证可靠性的同时提供多种存储服务
- 一般而言，分布式存储是将数据分散存储在多台独立的设备上，采用可扩展的系统结构、利用多台存储服务器分担存储负荷、利用位置服务器定位存储信息

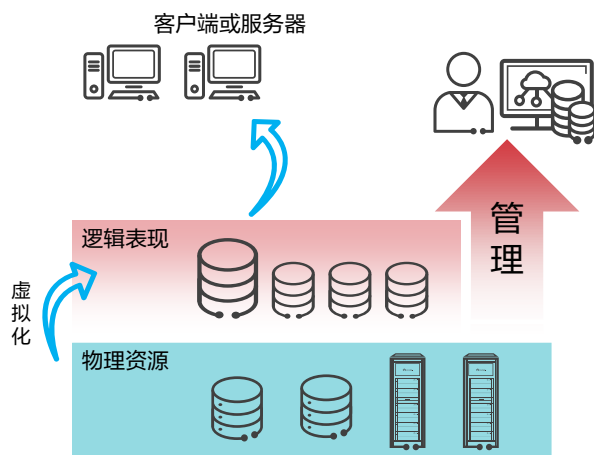


- 传统架构扩展能力和灵活性有限。
- 分布式存储架构：
  - 通用硬件、架构统一、存储和网络不被绑定；
  - 性能、容量线性扩展，目前可达数千个节点；
  - 资源弹性伸缩，利用率高。



## 存储虚拟化

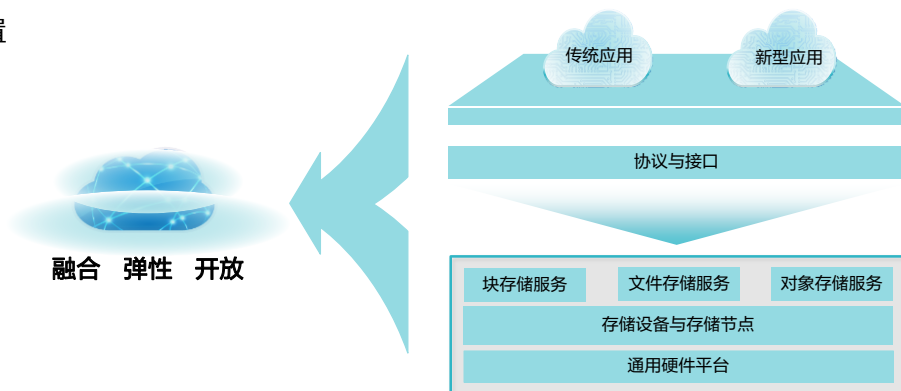
- 存储虚拟化可以将存储设备进行抽象，以逻辑资源的方式呈现，统一提供全面的存储服务
- 可以在不同的存储形态或设备类型之间提供统一的功能



- 在物理存储系统和服务器之间增加一个虚拟层，它管理和控制所有存储并对服务器提供存储服务。
- 存储硬件的增减、调换、分拆、合并对服务器完全透明。
- 作用：
  - 隐藏物理设备复杂度；
  - 集成使用现有的功能；
  - 摆脱物理容量的局限。

## 云存储

- 云存储系统是一个多存储设备、多应用、多服务协同工作的集合体，它使用高度虚拟化的多租户基础设施为企业提供可扩展的存储资源，可以根据组织的要求动态配置



- 通过集群应用、网格技术或分布式文件系统等功能，将网络中不同类型的存储设备通过应用软件集合起来协同工作，共同对外提供数据存储和业务访问功能。

# 目录

---

## 1. 存储基础介绍

- 什么是存储
- 存储发展历程
- 主流硬盘类型
- 存储组网类型
- 存储形态简介

## 2. 存储关键技术

## 硬盘简介

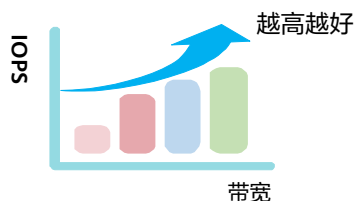
- 硬盘是计算机最主要的存储设备
- 硬盘接口是硬盘与主机系统间的连接部件，作用是在硬盘缓存和主机内存之间传输数据。不同的硬盘接口决定着硬盘与计算机之间的连接速度，直接影响着程序运行快慢和系统性能好坏

	SATA	SAS	NL-SAS	SSD
主流转速 (RPM)	7,200	15,000/10,000	7200	NA
串行/并行	串行	串行	串行	串行
主流容量 (TB)	1T/2T/3T	0.6T/0.9T	2T/3T/4T	0.6T/0.8T/1.2T/1.6T
MTBF (h)	1,200,000	1,600,000	1,200,000	2,000,000
备注	由ATA硬盘发展而来，采用串行方式传输，SATA 3.0实现600MB/s最高数据传输率 SATA硬盘年故障率大约2%	SAS专为满足高性能企业需求而设计，并且兼容SATA硬盘。能够提供3.0 Gb/s到6.0Gb/s的传输率，未来规划到12.0Gb/s SAS硬盘年故障率小于2%	带SAS接口的“企业级SATA驱动器”，适用于在一个磁盘阵列系统中实现分级存储，简化了磁盘阵列系统的设计 NL-SAS硬盘年故障率大约2%	固态硬盘 (Solid State Disk) 用固态电子存储芯片阵列而制成的硬盘，由控制单元和存储单元 (FLASH芯片、DRAM芯片) 组成。固态硬盘的接口规范和定义、功能及使用方法与普通硬盘完全相同，在产品外形和尺寸上也完全与普通硬盘一致

- MTBF: Mean Time Between Failure，平均无故障工作时间。
- 在价格方面，通常来说SATA和NL-SAS硬盘较为便宜，SAS硬盘较贵一些，SSD最为昂贵。

## 硬盘关键指标

- 硬盘容量（Disk Capacity）
- 转速（Rotational Speed）（机械硬盘指标）
- 平均访问时间（Average Access Time）
- 数据传输率（Date Transfer Rate）
- IOPS（Input/Output Operations Per Second）



硬盘类型	IOPS (4K随机写)	带宽 (128K顺序读)
SATA	330	200MB/s
SAS 10K	350	195MB/s
SAS 15K	450	290MB/s
SATA SSD	30~60K	540MB/s
SAS SSD	155K	1000MB/s
NVMe SSD	300K	3500MB/s

- **硬盘容量（Disk Capacity）**：容量的单位为兆字节（MB）或千兆字节（GB）。影响硬盘容量的因素有单碟容量和碟片数量。
- **转速（Rotational speed）**：硬盘的转速指硬盘盘片每分钟转过的圈数，单位为RPM（Rotation Per Minute）。一般硬盘的转速都达到5400RPM/7200RPM。SCSI接口硬盘转速可达10000 – 15000RPM。
- **平均访问时间（Average Access Time）** = 平均寻道时间 + 平均等待时间。
- **数据传输率（Date Transfer Rate）**：硬盘的数据传输率是指硬盘读写数据的速度，单位为兆字节每秒（MB/s）。硬盘数据传输率包括内部传输率和外部传输率两个指标。
- **IOPS（Input/Output Per Second）**：即每秒的输入输出量（或读写次数），是衡量磁盘性能的主要指标之一。随机读写频繁的应用，如OLTP（Online Transaction Processing），IOPS是关键衡量指标。另一个重要指标是数据吞吐量（Throughput），指单位时间内可以成功传输的数据数量。对于大量顺序读写的应用，如电视台的视频编辑，视频点播等则更关注吞吐量指标。

# 目录

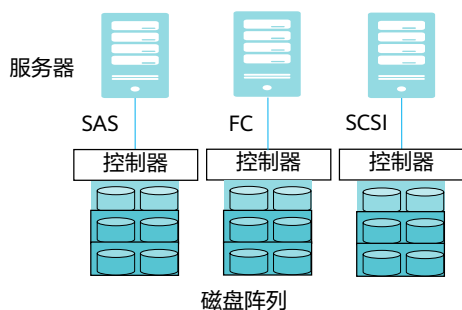
---

## 1. 存储基础介绍

- 什么是存储
- 存储发展历程
- 主流硬盘类型
- 存储组网类型
- 存储形态简介

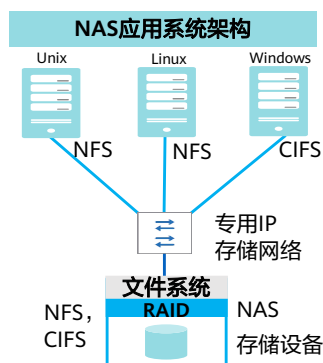
## 2. 存储关键技术

## DAS存储简介



- **DAS ( Direct Attached Storage )**
- **时间：**70年代
- **背景：**用户最早因为数据量的增多而产生存储的需求，从而产生最早最简单的存储架构：直连附加存储DAS
- **连接方式：**FC，SCSI，SAS
- **访问方式：**直连式存储与服务器主机之间的连接通道通常采用SAS连接
- **链路速率：**3Gbit/s、6Gbit/s、12Gbit/s
- **提供快照、备份等功能**

## NAS存储简介（1）



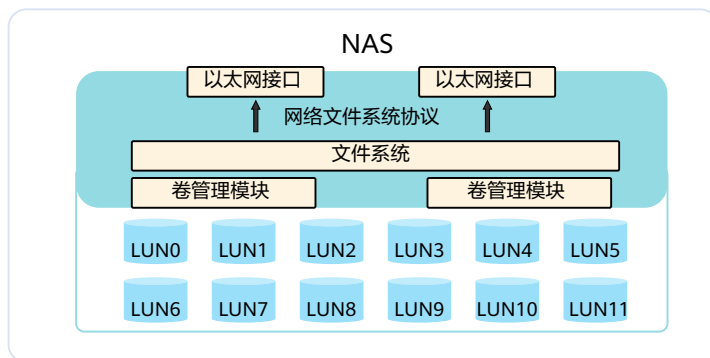
- **NAS（Network Attached Storage）**
- **时间：**90年代初
- **背景：**网络飞速发展，大量数据需要共享和交换，出现专用的NAS存储设备，成为数据共享与交换的核心
- **访问方式：**多台前端服务器共享后端存储设备，后端NAS设备上的存储空间通过CIFS（Windows系统）、NFS（类Unix系统）协议共享给前端主机，可同时对同一目录或文件进行并发读写
- **文件系统位于后端存储设备**

- NFS（Network File System）是Sun Microsystems在1984年创建的Internet标准协议。开发NFS是为了允许在局域网上的系统之间共享文件。
- Linux NFS客户端支持三个版本的NFS协议：NFSv2 [RFC1094]、NFSv3 [RFC1813]和NFSv4 [RFC3530]。其中NFSv2使用UDP协议，数据访问和传输能力有限，已经过时；
  - NFSv3版本在1995年发布，添加了TCP协议作为传输选项，被广泛使用；
  - NFSv4版本在2003年发布，已获得更好的性能和安全性。
- NFS的工作机制：主要是采用远程过程调用RPC机制。
  - RPC提供了一组与机器、操作系统以及低层传送协议无关的存取远程文件的操作，允许远程客户端以与本地文件系统类似的方式，来通过网络进行访问。
  - NFS客户端向NFS服务器端发起RPC请求，服务器将请求传递给本地文件访问进程，进而读取服务器主机上的本地磁盘文件，返回给客户端。
- CIFS（Common Internet File System）是一种网络文件系统协议，用于在网络上的机器之间提供对文件和打印机的共享访问。现在主要实现在Windows主机之间进行网络文件共享功能。



## NAS存储简介（2）

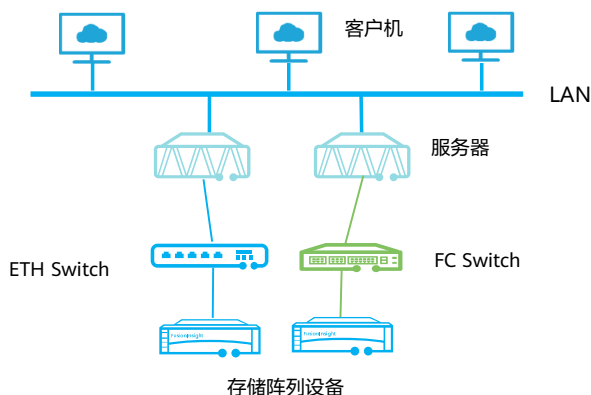
- NAS: Network Attached Storage, 网络附加存储, 是一种将分布、独立的数据进行整合, 集中化管理, 以便于对不同主机和应用服务器进行访问的技术



- NAS可作为网络节点, 直接接入网络中, 理论上NAS可支持各种网络技术, 支持多种网络拓扑, 但是以太网是目前最普遍的一种网络连接方式, 我们主要讨论是以以太网为网络基础的NAS环境。
- NAS本身能够支持多种协议 (如NFS、CIFS等), 而且能够支持各种操作系统。通过任何一台工作站, 采用IE或Netscape浏览器就可以对NAS设备进行直观方便的管理。

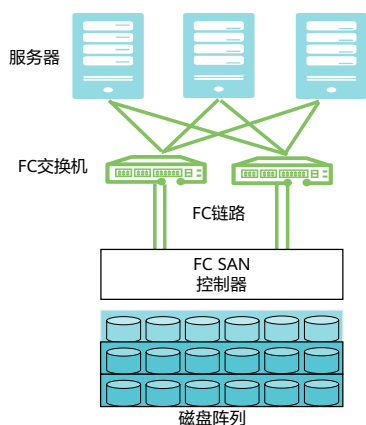
## SAN存储简介

- SAN：Storage Area Networks，存储区域网络，是通过专用高速网将一个或多个网络存储设备和服务器连接起来的专用存储系统



- 存储区域网络（Storage Area Networks，SAN）：一个存储网络是一个用在服务器和存储资源之间的、专用的、高性能的网络体系。SAN是独立于LAN的服务器后端存储专用网络。SAN采用可扩展的网络拓扑结构连接服务器和存储设备，每个存储设备不隶属于任何一台服务器，所有的存储设备都可以在全部的网络服务器之间作为对等资源共享。
- SAN主要利用Fibre Channel Protocol（光纤通道协议），通过FC交换机建立起与服务器和存储设备之间的直接连接，因此我们通常也称这种利用FC连接建立起来的SAN为FC-SAN。FC特别适合这项应用，原因在于一方面它可以传输大块数据，另一方面它能够实现较远距离传输。SAN主要应用在对于性能、冗余度和数据的可获得性都有很高要求的高端、企业级存储应用上。
- 随着存储技术的发展，目前基于TCP/IP协议的IP-SAN也得到很广泛的应用。IP-SAN具备很好的扩展性、灵活的互通性，并能够突破传输距离的限制，具有明显的成本优势和管理维护容易等特点。
- NAS和SAN最大的区别就在于NAS有文件操作和管理系统，而SAN却没有这样的系统功能，其功能仅仅停留在文件管理的下一层，即数据管理。SAN和NAS并不是相互冲突的，是可以共存于一个系统网络中的，但NAS通过一个公共的接口实现空间的管理和资源共享，SAN仅仅是为服务器存储数据提供一个专门的快速后方存储通道。

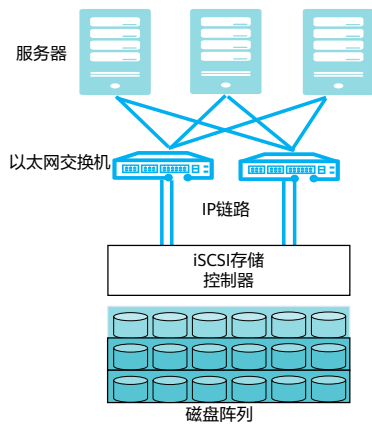
## FC-SAN简介



- **FC-SAN ( Fiber Channel Storage Area Network )**
- **时间:** 90年代中后期
- **背景:** 为解决DAS扩展性差的问题, 将存储设备网络化, 可以同时连接上百台服务器
- **连接方式:** FC光纤, 使用专用的FC交换机
- **访问方式:** 后端一台存储设备的存储空间可以划分为多个LUN, 供前端服务器使用
- **链路速率:** 2Gbit/s、4Gbit/s 、8Gbit/s
- **提供快照、容灾等高级数据保护功能**

- **FC:** Fiber Channel, 光纤通道, 是指一种用于在光纤或者铜缆上传输100Mbit/s到4.25Gbit/s信号的标准数据存储网络。用于建立存储区域网的高速传输技术。光纤通道能够用于支持ATM, IP等协议的一般网络, 但它主要用途是从服务器上传输小型计算机系统接口 ( SCSI ) 流量到磁盘阵列。

## IP-SAN简介



- **IP-SAN ( IP Storage Area Network )**
- **时间:** 2001年
- **背景:** 为解决FC-SAN在价格及管理上的诸多门坎而产生
- **连接方式:** 采用以太网作为连接链路, 以太网交换机
- **访问方式:** 后端一台存储设备的存储空间可以划分为多个LUN, 供前端服务器使用
- **链路速率:** 1Gbit/s 、10Gbit/s
- 提供快照、容灾等高级数据保护功能
- iSCSI 被看好的原因
  - 可以采用非常成熟的IP网络管理工具和基础建设
  - IP网络使用普遍, 可为企业节省大笔建设、管理及人事成本

- iSCSI: Internet Small Computer System Interface, Internet小型计算机系统接口, 是一种基于因特网及SCSI-3协议下的存储技术, 它将原来只用于本机的SCSI协议透过TCP/IP网络发送, 使连接距离可作无限延伸。在后续课程中我们会了解到它的协议封装、工作原理及使用场景。

## 三种存储组网总结对比

	DAS	NAS	SAN	
			FC-SAN	IP-SAN
传输类型	SCSI、FC、SAS	IP	FC	IP
数据类型	块级	文件级	块级	块级
典型应用	任何	文件服务器	数据库应用	视频监控
优点	易于理解 兼容性好	易于安装 成本低	高扩展性、高性能 高可用性	高扩展性 成本低
缺点	难管理，扩展性有限 存储空间利用率不高	性能较低 对某些应用不适合	较昂贵，配置复杂 组网兼容性问题	性能较低

# 目录

---

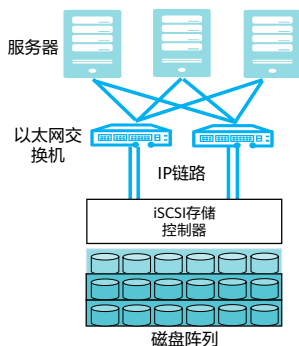
## 1. 存储基础介绍

- 什么是存储
- 存储发展历程
- 主流硬盘类型
- 存储组网类型
- 存储形态简介

## 2. 存储关键技术

## 集中式存储

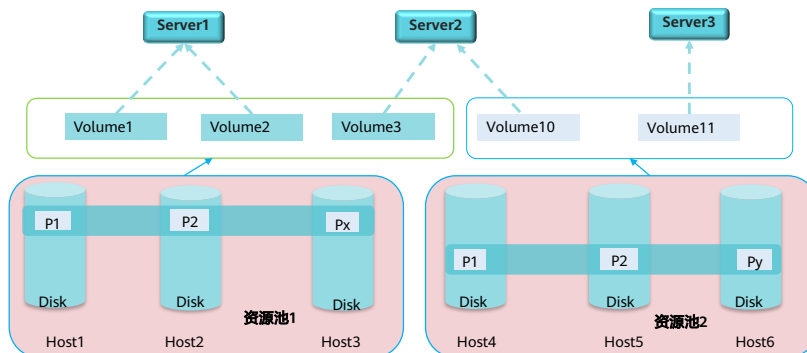
- 集中式存储系统，即整个存储是集中在一个系统中的。企业级的存储设备一般都是集中式存储。但集中式存储并不是一个单独的设备，是集中在一套系统当中的多个设备。以华为公司的存储为例，整个存储系统可能需要几个机柜来存放。集中式存储按照技术架构可以划分为SAN和NAS，其中SAN又可以细分为FC-SAN、IP-SAN
- 集中式存储最大的特点就是部署结构简单，无需考虑如何对服务进行多个节点的部署，也就不用考虑多个节点之间的分布式协作问题



- 集中式存储的缺点：
  - 孤立的存储资源，存储通过专用网络连接到有限数量的服务器；
  - 集中式纵向扩容通过增加硬盘框实现，硬件控制器性能（单控制器带盘能力）成为瓶颈；
  - 集中式存储横向扩容需要通过控制器全连接实现，硬件控制器性能成为扩容瓶颈；
  - 集中式存储资源缺乏共享，存储设备和资源往往由不同厂家提供，设备之间无法进行资源共享，数据中心看到的是一个个孤立的存储池；
  - 集中式存储采用集中式元数据管理方式，系统所能提供的并发操作能力将受限于元数据服务的性能，元数据服务也将会成为系统的性能瓶颈；
- 如何解决传统集中式存储扩容、性能瓶颈的问题？

## 分布式存储

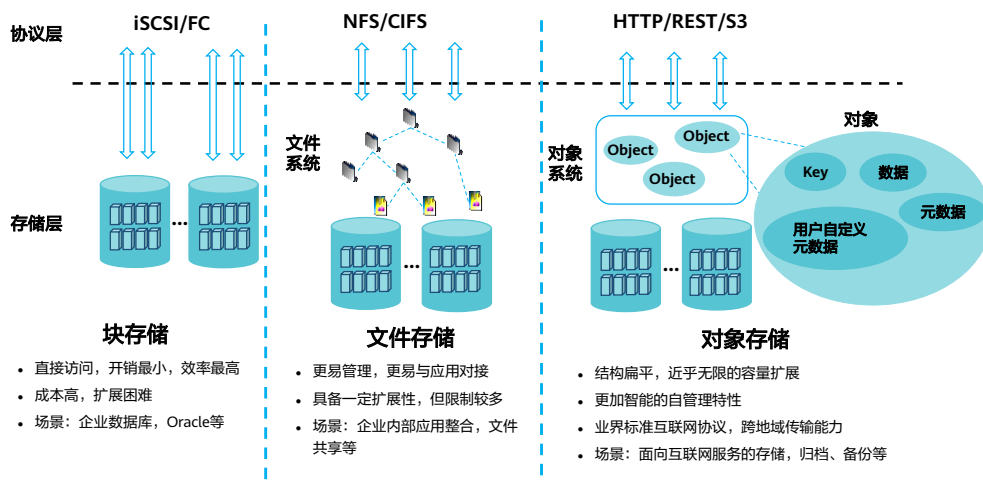
- 分布式存储系统，是将数据分散存储在多台独立的服务器上。分布式网络存储系统采用可扩展的系统结构，利用多台存储服务器分担存储负荷，它不但提高了系统的可靠性、可用性和存取效率，还易于扩展。随着分布式存储越来越普及，现在一些对性能要求比较高的应用也开始使用分布式存储，比如金融系统的数据库



- 分布式存储利用软件重构存储服务形式，通过软件模拟原先硬件控制器实现功能的同时，规避硬件控制器的种种弊端。
- 资源池：类似于SAN的RAID组概念。



## 存储业务类型



- 对象存储是一种新兴存储技术, 对象存储系统综合了NAS和SAN的优点, 同时具有SAN的高速直接访问和NAS的数据共享等优势, 提供了高可靠性、跨平台性以及安全的数据共享的存储体系结构。对象存储与块存储、文件存储的对比如下:
  - 块存储对存储层直接访问, 开销最小, 效率最高, 速度最快。但成本最高, 扩展困难。块存储采用iSCSI/FC协议, 很难跨网络传输。适合的应用场景是企业数据库, 如运行Oracle等;
  - 文件存储是在块存储之上构建了文件系统, 采用目录-目录-文件的方式组织数据, 更容易管理。因为大多数应用程序都是对文件进行操作, 因此文件存储更容易和应用系统对接。文件系统受目录树的限制, 扩展性受限, 一般最多扩展到几十PB。文件系统适用于企业内部应用整合, 文件共享场景;
  - 对象存储是在块存储之上构建了对象管理层, 与文件系统相比, 对象系统层是扁平的, 扩展限制少, 因此拥有近乎无限的扩展性。对象由唯一的Key, 文件, 数据(文件), 元数据, 自定义元数据构成, 由于包含了自管理信息, 更加智能。对象存储采用兼容标准的互联网协议接口, 可以跨地域传输。对象存储适用于面向互联网服务的存储场景, 以及企业内部的归档、备份场景。

# 目录

---

1. 存储基础介绍
2. 存储关键技术
  - RAID技术
  - 存储协议

## RAID基本概念

- RAID: Redundant Array of Independent Disks, 独立硬盘冗余阵列, RAID技术将多个单独的物理硬盘以不同的方式组合成一个逻辑硬盘, 从而提高了硬盘的读写性能和数据安全性
  - 根据不同的组合方式可以分为不同的RAID级别

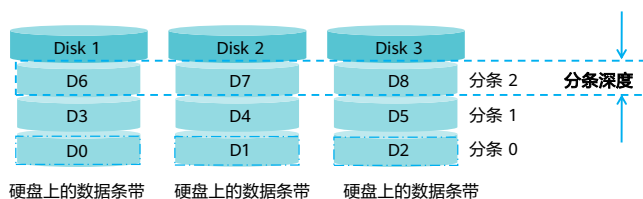
RAID 0	数据条带化, 无校验
RAID 1	数据镜像, 无校验
RAID 3	数据条带化读写, 校验信息存放于专用硬盘
RAID 5	数据条带化, 校验信息分布式存放
RAID 6	数据条带化, 分布式校验并提供两级冗余

- 同时采用两种不同的RAID方式还能组合成新的RAID级别

RAID 0+1	先做RAID 0, 后做RAID 1, 同时提供数据条带化和镜像
RAID 10	类似于RAID 0+1, 区别在于先做RAID 1, 后做RAID 0
RAID 50	先做RAID 5, 后做RAID 0, 能有效提高RAID 5的性能

## RAID的数据组织形式

- 硬盘的条带化：将硬盘空间按照设定的大小分为多个条带，数据写入时也按照条带的大小来划分数据模块
- 条带（strip）：硬盘中单个或者多个连续的扇区构成一个条带，是组成分条的元素
- 分条（stripe）：同一硬盘阵列中的多个硬盘驱动器上的相同“位置”（或者说是相同编号）的条带



- 分条宽度
  - 指一个磁盘组在进行分条时所使用的磁盘的个数，例如一个由3个成员盘组成的磁盘组，其分条宽度为“3”。
- 分条深度
  - 指一个分条的容量大小。

## RAID的数据保护方式

- 方法一：镜像（Mirroring），在另一块冗余的硬盘上保存数据的副本
- 方法二：奇偶校验算法（XOR）
  - XOR运算广泛地使用在数字电子和计算机科学中
  - XOR校验的算法—相同为假，相异为真：
    - $0\oplus 0=0$ ；  $0\oplus 1=1$ ；  $1\oplus 0=1$ ；  $1\oplus 1=0$

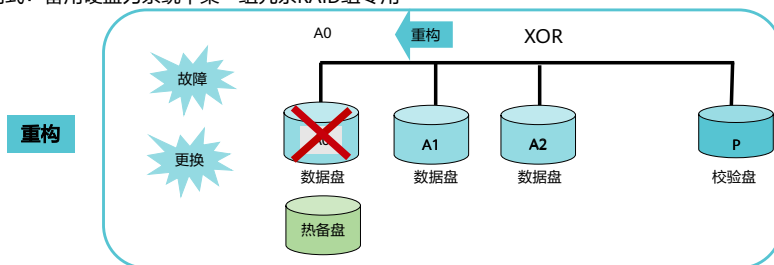


异或校验冗余备份

- RAID技术通常有2种不同的方式进行数据保护：
  - 在另一块冗余的硬盘上保存数据的副本，提高可靠性，并可以提高读性能。
  - 使用奇偶校验算法。奇偶校验码是使用用户数据计算出的额外信息。对于使用奇偶校验的RAID类型，它意味着需要额外的校验硬盘。奇偶校验采用的是异或（XOR的计算符号 $\oplus$ ）算法。

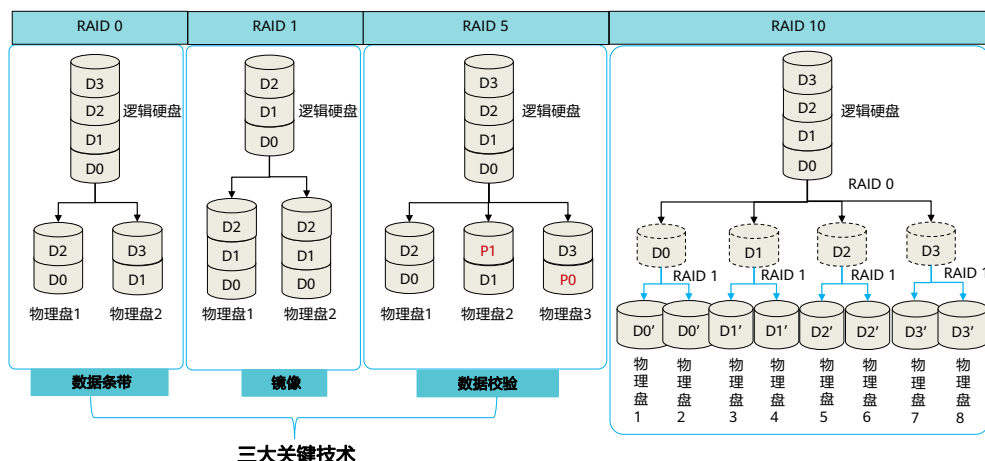
## RAID热备和重构概念

- 热备（Hot Spare）的定义
  - 当冗余的RAID组中某个硬盘失效时，在不干扰当前RAID系统正常使用的情况下，用RAID系统中另外一个正常的备用硬盘自动顶替失效硬盘，及时保证RAID系统的冗余性
- 热备一般分为两种
  - 全局式：备用硬盘为系统中所有的冗余RAID组共享
  - 专用式：备用硬盘为系统中某一组冗余RAID组专用



- 数据校验，利用冗余数据进行数据错误检测和修复，冗余数据通常采用海明码、异或操作等算法来计算获得。利用校验功能，可以很大程度上提高磁盘阵列的可靠性、高性能和容错能力。不过，数据校验需要从多处读取数据并进行计算和对比，会影响系统性能。
- 一般来说，RAID不可作为数据备份的替代方案，它对非磁盘故障等造成的数据丢失无能为力，比如病毒、人为破坏、意外删除等情形。此时的数据丢失是相对操作系统、文件系统、卷管理器或者应用系统来说的，对于RAID而言，数据都是完好的，没有发生丢失。所以，数据备份、灾备等数据保护措施是非常必要的，与RAID相辅相成，保护数据在不同层次的安全性，防止发生数据丢失。

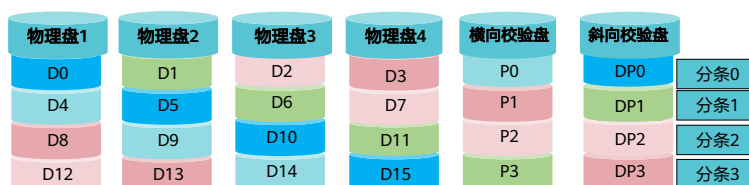
## 常见RAID级别介绍



- RAID 0：没有容错设计的条带硬盘阵列，以条带形式将RAID组的数据均匀分布在各个硬盘中。
- RAID 1：又称镜像（Mirror），数据同时一致写到主硬盘和镜像硬盘。
- RAID 3：带有校验的并行数据传输阵列，数据条带化分布在数据盘中，同时使用专用校验硬盘存放校验数据。
- RAID 5：与RAID 3机制类似，但校验数据均匀分布在各数据硬盘上，RAID成员硬盘上同时保存数据和校验信息，数据块和对应的校验信息保存在不同硬盘上。RAID 5是最常用的RAID方式之一。
- RAID 10：将镜像和条带进行两级组合的RAID级别，第一级是RAID 1镜像对，第二级为RAID 0。RAID 10也是一种应用比较广泛的RAID级别。

## RAID 6 DP工作原理

- DP: Double Parity, 就是在RAID 4所使用的一个行XOR校验硬盘的基础上又增加了一个硬盘用于存放斜向的XOR校验信息
- 横向校验盘中P0-P3为各个数据盘中横向数据的校验信息
  - 例:  $P0 = D0 \text{ XOR } D1 \text{ XOR } D2 \text{ XOR } D3$
- 斜向校验盘中DP0-DP3为各个数据盘及横向校验盘的斜向数据校验信息
  - 例:  $DP0 = D0 \text{ XOR } D5 \text{ XOR } D10 \text{ XOR } D15$



- RAID 6 DP也有两个独立的校验数据块。第一个校验信息与RAID 6 P+Q的第一个校验值是相同的，第二个不同于RAID 6 P+Q，采用的是斜向异或运算得到行对角奇偶校验数据块。行奇偶校验值是同一分条的用户数据异或运算获得到，如图所示：P0是由分条0上的D0，D1，D2和D3异或运算得到，P1由分条1上的D4，D5，D6，D7异或运算得到。所以， $P0 = D0 \oplus D1 \oplus D2 \oplus D3$ ， $P1 = D4 \oplus D5 \oplus D6 \oplus D7$ ，如此类推。
- 第二个校验数据块是由阵列的对角线数据块进行异或运算。数据块的选择过程比较复杂。DP0是由硬盘1的分条0上的D0，硬盘2的分条1上的D5，硬盘3上的分条2的D10，和硬盘4分条3上的D15异或操作得到。DP1是对硬盘2的分条0上的D1，硬盘3的分条1上的D6，硬盘4上分条2的D11，和第一块校验硬盘分条3上的P3进行异或运算得到。DP2是硬盘3分条0上的D2，硬盘4上的分条1的D7，奇偶硬盘分条2的P2，和硬盘1分条3上的D12进行异或运算得到。所以， $DP0 = D0 \oplus D5 \oplus D10 \oplus D15$ ， $DP1 = D1 \oplus D6 \oplus D11 \oplus P3$ ，如此类推。
- 一个RAID 6阵列能够容忍双硬盘失效。
- 一个RAID 6组的性能，无论算法是DP还是P+Q，相对都比较慢。因此，RAID 6适用两种场景：
  - 数据非常重要，需要尽可能长的时间处于在线和可使用的状态。
  - 使用的硬盘容量非常大（通常超过2 T）。大容量硬盘的重建时间较长，两个硬盘都失效是会造成数据较长时间不能访问。在RAID 6中，可以实现一个硬盘重构时另一个硬盘失效。一些企业希望在使用大容量硬盘后，存储阵列的供应商使用一个双重保护的RAID组。



## RAID 2.0技术介绍

- **RAID 2.0**

- 增强型RAID技术，有效解决了机械硬盘容量越来越大，重构一块机械硬盘所需时间越来越长，传统RAID组重构窗口越来越大而导致重构期间又故障一块硬盘而彻底丢失数据风险的问题

- **RAID 2.0+**

- 在RAID 2.0的基础上提供了更细粒度（可达几十KB粒度）的资源颗粒，形成存储资源的标准分配及回收单位，类似计算虚拟化中的虚拟机，我们称之为虚拟块技术

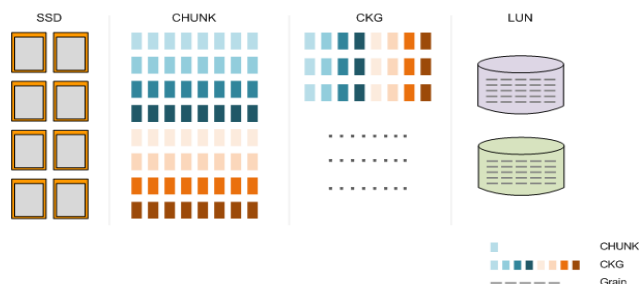
- **华为RAID 2.0+**

- 是华为针对传统RAID的缺点，设计的一种满足存储技术虚拟化架构发展趋势的全新的RAID技术，其传统固定管理模式为两层虚拟化管理模式，在底层块级虚拟化（Virtual for Disk）硬盘管理的基础之上，通过一系列Smart效率提升软件，实现了上层虚拟化（Virtual for Pool）的高效资源管理

- 块级虚拟化的含义是：将系统中的硬盘划分成若干个连续的固定大小的存储空间，称为存储块，即Chunk，或简称CK。

## RAID 2.0+块虚拟化技术

- 存储的数据最终都会存储到SSD盘上，如果有些盘片上存放的数据不均匀，就可能导致某些压力大的SSD盘成为系统的瓶颈
- 为应对该问题，华为存储系统采用新型的RAID 2.0+块虚拟化技术，通过更细粒度的划分，实现所有LUN的数据均衡地分布在每块SSD盘上，实现盘的负载均衡



1. 多块硬盘组成一个存储池。
  2. 存储系统将存储池中的每个SSD硬盘划分为固定大小的CK（通常大小为4MB）进行逻辑空间管理。
  3. 来自不同SSD硬盘的CHUNK按照用户在DeviceManager上设置的“RAID策略”组成CKG。
  4. CKG再被划分为更细粒度的Grain（通常为8KB）。用户创建LUN时，存储系统以Grain为单位映射到LUN，实现对存储资源的精细化管理。
- 相比传统RAID机制，RAID2.0+具备如下优势：
    - 业务负载均衡，避免热点。数据打散到资源池内所有硬盘上，没有热点，硬盘负荷平均，避免个别盘因为承担更多的写操作而提前达到寿命的上限。
    - 快速重构，缩小风险窗口。当硬盘故障时，故障盘上的有效数据会被重构到资源池内除故障盘外的所有盘上，实现了多对多的重构，速度快，大幅缩短数据处于非冗余状态的时间。
    - 全盘参与重构。资源池内所有硬盘都会参与重构，每个盘的重构负载很低，重构过程对上层应用无影响。

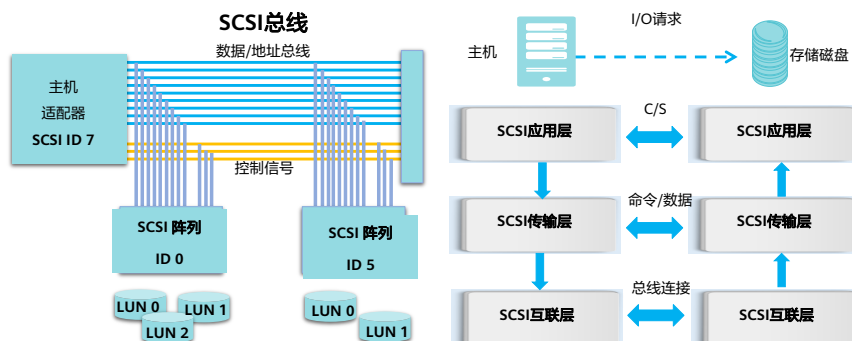
# 目录

---

- 1. 存储基础介绍
- 2. 存储关键技术
  - RAID技术
  - 存储协议

## SCSI协议

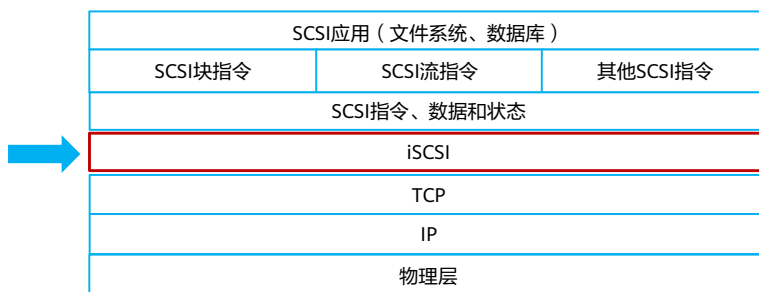
- SCSI: Small Computer System Interface, 小型计算机系统接口, 最初是一种为小型机研制的接口技术, 用于主机与外部设备之间的连接



- 计算机与存储系统之间的通信是通过总线来完成的。总线就是从源设备传输数据到目标设备的路径。在最简单的情况下, 控制器的高速缓存作为源, 将数据传输给目标磁盘。控制器首先向总线处理器发出请求使用总线的信号。该请求被接受之后, 控制器高速缓存就开始执行发送操作。在这个过程中, 控制器占用了总线, 总线上所连接的其它设备都不能使用总线。由于总线具备中断功能, 所以总线处理器可以随时中断这一传输过程并将总线控制权交给其它设备, 以便执行更高优先级的操作。
- 计算机中布满了总线—从一个位置向另一个位置传输信息和电力的高速通道。例如, 将MP3或数码相机连接到计算机时, 您可能会使用通用串行总线 (USB) 端口。对于存储图片、音乐等的小型电子设备, USB端口完全可以胜任传输数据和充电的工作。但是, 这种总线还不足以同时支持整台计算机和服务器的其他许多设备。
- 在这种情况下, 就需要使用SCSI这样的总线。SCSI直译为小型计算机系统专用接口 (Small Computer System Interface) 是一种连结主机和外围设备的接口, 支持包括磁盘驱动器、磁带机、光驱、扫描仪在内的多种设备。它由SCSI 控制器进行数据操作, SCSI 控制器相当于一块小型CPU, 有自己的命令集和缓存。SCSI是一种特殊的总线结构, 可以对计算机中的多个设备进行动态分工操作, 对于系统同时要求的多个任务可以灵活机动地适当分配, 动态完成。

## iSCSI协议

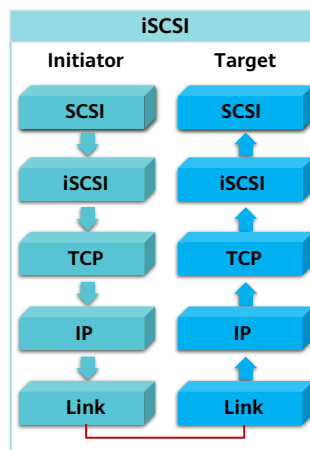
- iSCSI: Internet SCSI, 把SCSI命令和块状数据封装在TCP中在IP网络中传输, 基本出发点是利用成熟的IP网络技术来实现和延伸SAN



- 通过SCSI控制卡的使用可以连接多个设备, 形成自己的“网络”, 但是这个“网络”仅局限于与所附加的主机进行通信, 并不能在以太网上共享。那么, 如果能够通过SCSI协议组成网络, 并且能够直接挂载到以太网上, 作为网络节点和其它设备进行互联共享, 那么SCSI就可以得到更为广泛的应用。所以, 经过对SCSI的改进, 就推出了iSCSI这个协议。基于iSCSI协议的IP-SAN是把用户的请求转换成SCSI代码, 并将数据封装进IP包内在以太网中进行传输。
- iSCSI方案最早是由Cisco和IBM两家发起, 并且由Adaptec、Cisco、HP、IBM、Quantum等公司共同倡导。它提供基于TCP传输, 将数据驻留于SCSI设备的方法。iSCSI标准草案在2001年推出, 并经过多次论证和修改, 于2002年提交IETF, 在2003年2月, iSCSI标准正式发布。iSCSI技术的重要贡献在于其对传统技术的继承和发展: 其一, SCSI (Small Computer System Interface, 小型计算机系统接口) 技术是被磁盘、磁带等设备广泛采用的存储标准, 从1986年诞生起到现在仍然保持着良好的发展势头; 其二, 沿用TCP/IP协议, TCP/IP在网络方面是最通用、最成熟的协议, 且IP网络的基础建设非常完善。这两点为iSCSI的无限扩展提供了坚实的基础。
- IP网络的普及性将使得数据可以通过LAN、WAN或者是通过Internet利用新型IP存储协议传输, iSCSI既是在这个思想的指导下进行研究和开发的。iSCSI是基于IP协议的技术标准, 实现了SCSI和TCP/IP协议的融合, 对众多的以太网用户而言, 只需要极少的投资, 就可以方便、快捷地对信息和数据进行交互式传输和管理。

## iSCSI的发起端与目标端

- 发起端（Initiator）
  - SCSI层负责生成CDB（命令描述符块），将CDB传给iSCSI
  - iSCSI层负责生成iSCSI PDU（协议数据单元），并通过IP网络将PDU发给target
- 目标器（Target）
  - iSCSI层收到PDU，将CDB传给SCSI层
  - SCSI层负责解释CDB的意义，必要时发送响应



- iSCSI的通信体系仍然继承了SCSI的部分特性，在iSCSI通信中，具有一个发起I/O请求的启动器设备（Initiator）和响应请求并执行实际I/O操作的目标器设备（Target）。在Initiator和Target建立连接后，Target在操作中作为主设备控制整个工作过程。
- iSCSI Initiator：iSCSI启动器，可分为三种，即软件Initiator驱动程序、硬件的TOE（TCP Offload Engine，TCP卸载引擎）卡以及iSCSI HBA卡。就性能而言，软件Initiator驱动程序最差、TOE卡居中、iSCSI HBA卡最佳。
- iSCSI Target：iSCSI目标器，通常为iSCSI磁盘阵列、iSCSI 磁带库等。
- iSCSI协议为Initiator和Target定义了一套命名和寻址方法。所有的iSCSI节点都是通过其iSCSI名称被标识的。这种命名方式使得iSCSI名称不会与主机名混淆。
- iSCSI使用iSCSI Name来唯一鉴别启动设备和目标设备。地址会随着启动设备和目标设备的移动而改变，但是名字始终是不变的。建立连接时，启动设备发出一个请求，目标设备接收到请求后，确认启动设备发起的请求中所携带的iSCSI Name是否与目标设备绑定的iSCSI Name一致，如果一致，便建立通信连接。每个iSCSI节点只允许有一个iSCSI Name，一个iSCSI Name可以被用来建立一个启动设备到多个目标设备的连接，多个iSCSI Name可以被用来建立一个目标设备到多个启动设备的连接。

## 问题研讨：

- 之前讲过FC-SAN和IP-SAN，如果当前使用的是FC-SAN，但是外部网络是TCP/IP网络，此时需要两个站点的存储互相通信，如何能做到呢
  - FC协议与TCP协议融合



## FC协议与TCP协议融合

- 从现有的情况来看，以太网技术和FC技术都在飞速发展，IP-SAN和FC-SAN在很长的一段时间内都将是并存且互为补充的
- TCP/IP网络承载FC信道方式：
  - iFCP
  - FCoE

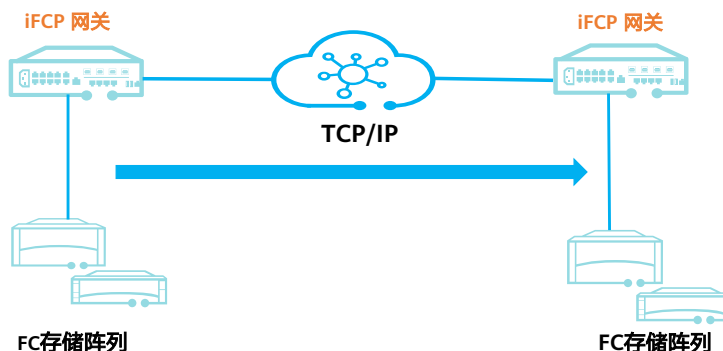


- FCIP: Fibre Channel over IP，基于IP的光纤通道，是连接TCP/IP网络链路上的光纤通道架构的一项IETF建议标准。利用目前的IP协议和设施来连接两个异地FC SAN的隧道，用以解决两个FC SAN的互连问题。
- iFCP: Internet Fibre Channel Protocol，Internet光纤信道协议，是一种网关到网关的协议，为TCP/IP网络上的光纤设备提供光纤信道通信服务。iFCP 使用 TCP提供拥塞控制、差错监测与恢复功能。iFCP主要目标是使现有的光纤信道设备能够在IP网络上以线速互联与组网。此协议及其定义的帧地址转换方法允许通过透明网关（transparent gateway）将光纤信道存储设备附加到基于IP的网络结构。
- FCoE: Fibre Channel over Ethernet，利用以太网路，传送光纤通道（Fibre Channel）的信号，让光纤通信的资料可以在10 Gigabit以太网路网络骨干中传输，但仍然是使用光纤通道的协定。
- IPFC: IP over Fiber Channel，在光纤通道上的IP。IPFC使用在两个服务器之间的光纤通道连接作为IP数据交换的媒介。为此，IPFC定义了如何通过光纤通道网络传送IP分组。跟所有的应用协议一样，IPFC实现为在操作系统中的一个设备驱动器。对本地IP配置的连接使用“ifconfig”“或“ipconfig”。然后IPFC驱动器寻址光纤通道主机适配卡，就可以在光纤通道上发送IP分组。



## iFCP协议

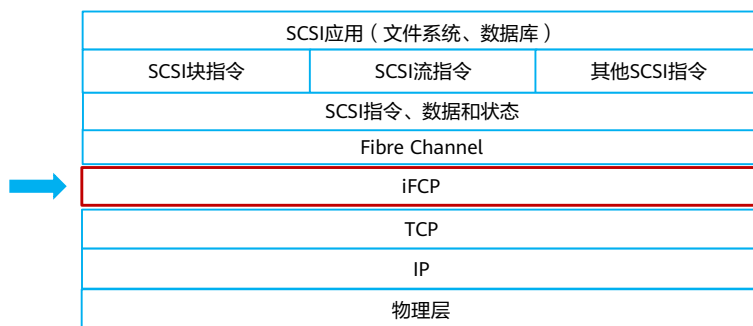
- iFCP: Internet Fibre Channel Protocol, 是一种网关到网关协议, 为TCP/IP网络中的光纤通道设备提供光纤通道通信服务, 实现端到端IP连接



- iFCP: Internet Fibre Channel Protocol, 是一种网关到网关协议, 为TCP/IP网络中的光纤通道设备提供光纤通道通信服务, 实现端到端IP连接。FC存储设备、HBA卡和交换机可以与iFCP网关直连。iFCP协议通过TCP协议提供流量控制、错误检测和错误恢复功能。iFCP的目的是使当前的光纤通道设备在IP网络上以线路速率互连和网络。
- iFCP协议定义的帧地址转换方法, 允许FC存储设备通过透明网关加入到TCP/IP网络中。iFCP可以替代光纤通道, 通过iFCP设备连接和分组光纤通道设备。但iFCP不支持独立SAN的合并, 因此无法形成逻辑SAN。iFCP的突出之处在于支持SAN互联和网关分区, 允许隔离故障区域, 打破点对点隧道的限制。同时, iFCP实现了光纤通道设备之间的端到端连接。因此, TCP连接中断只会对通信产生影响。采用iFCP的SAN能够实现故障隔离和安全管理, 比采用FCIP的SAN具有更高的可靠性。

## iFCP协议栈

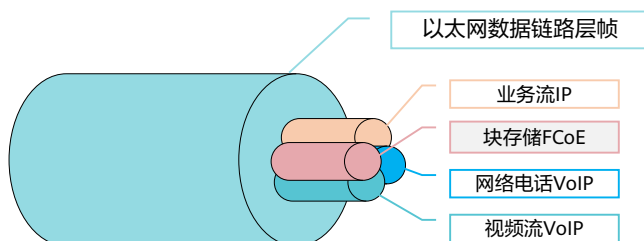
- iFCP协议位于光纤通道和TCP/IP之间，这意味着iFCP可以与光纤通道或TCP/IP进行交互



- iFCP协议层的主要功能是在本地和远程连接的N\_Ports之间传输光纤通道帧图像。当传输帧到远程N\_Port时，iFCP层封装并路由由每个光纤通道信息单元组成的光纤通道帧，通过预先确定的TCP连接在IP网络上传输。
- 在采用iFCP协议的IP SAN中，iFCP设备代替了光纤交换机，即iFCP交换机也可以作为iSNS (Internet Storage Name Server)，为终端节点提供名称发现服务。iFCP交换机为每个FC终端节点分配一个4字节的IP地址。当光纤通道设备发送SNS名称查询请求时，iFCP交换机将截获请求，并由iSNS服务器进行解析。

## FCoE协议

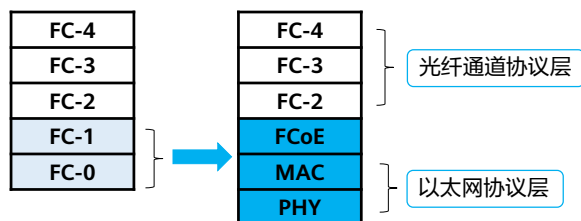
- FCoE: Fibre Channel over Ethernet, 以太网光纤通道。FCoE允许在一根通信线缆上传输LAN和FC SAN通信, 融合网络可以支持LAN和SAN数据类型, 减少数据中心设备和线缆数量, 同时降低供电和制冷负载, 收敛成一个统一的网络后, 需要支持的点也跟着减少了, 有助于降低管理负担。
- FCoE把FC帧封装在以太网帧中, 允许LAN和SAN的业务流量在同一个以太网中传送



- FCoE ( Fibre Channel over Ethernet ) 可以提供标准的光纤通道原有服务, 如发现、全局名称命名、分区等, 而且这些服务都可以照标准原有的运作, 保有FC原有的低延迟性、高性能。
- 从FC协议的角度来看, FCoE就是把FC承载在一种新型的链路上, 即以太网二层链路; 从以太网的角度来看, FCoE仅是其承载的另外一种上层协议 ( 类似于IP/IPX ) 。

## FCoE协议封装

- FCoE将FC-2及以上层的内容封装成以太网报文进行传输



- 光纤通道协议栈有五层。FC-0定义了媒体类型，FC-1定义了帧编解码方式，FC-2定义了帧划分协议和流量控制机制，FC-3定义了通用业务，FC-4定义了上层协议到光纤通道的映射。

## 问题研讨：

- 什么场景适合使用FCoE
- 什么场景适合使用iFCP



## 思考题

1. （多选）以下关于FC SAN的描述，正确的是哪些项？
  - A. 需要使用光纤交换机
  - B. 需要使用以太网交换机
  - C. 存储设备之间连接不能使用FC链路
  - D. 数据报文遵循FC协议栈
2. （判断）SATA硬盘的性能优于SAS硬盘。
  - A. 正确
  - B. 错误

- 参考答案：

- AD
- B

## 本章总结

- 本章我们一起学习了主流的数据存储方式及组网，掌握了RAID和华为RAID 2.0+块级虚拟化技术，了解了集中存储和分布式存储的区别与联系、存储的相关协议及适用场景。在接下来的课程中我们将继续学习基础技术中的网络技术部分，敬请期待。

## 学习推荐

- 华为Learning网站
  - <http://support.huawei.com/learning/Index!toTrainIndex>
- 华为Support案例库
  - <http://support.huawei.com/enterprise/servicecenter?lang=zh>



## 缩略语

- FC: Fiber Channel, 光纤通道
- FCIP: Fibre Channel over IP, 基于IP的光纤通道
- FCoE: Fibre Channel over Ethernet, 以太网光纤通道
- iFCP: Internet Fibre Channel Protocol, Internet光纤信道协议
- iSCSI: Internet Small Computer System Interface, Internet小型计算机系统接口
- IPFC: IP over Fiber Channel, 在光纤通道上的IP
- IOPS: Input/Output Per Second, 每秒的输入输出量或读写次数
- MTBF: Mean Time Between Failure, 平均无故障工作时间
- NAS: Network Attached Storage, 网络附加存储
- RAID: Redundant Array of Independent Disks, 独立磁盘冗余阵列
- SAN: Storage Area Networks, 存储区域网络
- SCSI: Small Computer System Interface, 小型计算机系统接口

# Thank you.

把数字世界带入每个人、每个家庭、  
每个组织，构建万物互联的智能世界。

Bring digital to every person, home, and  
organization for a fully connected,  
intelligent world.

Copyright©2023 Huawei Technologies Co., Ltd.  
All Rights Reserved.

The information in this document may contain predictive statements including, without limitation, statements regarding the future financial and operating results, future product portfolio, new technology, etc. There are a number of factors that could cause actual results and developments to differ materially from those expressed or implied in the predictive statements. Therefore, such information is provided for reference purpose only and constitutes neither an offer nor an acceptance. Huawei may change the information at any time without notice.



## 网络技术基础



# 前言

- 网络技术是一切平台、服务互联的基础，那么什么是网络？网络通信的基础原理是怎样的？我们常用的网络技术有哪些？本章将带您一起了解。

# 目标

- 学完本课程后，您将能够：
  - 了解IP地址的分类及子网划分
  - 了解网络通信基础原理
  - 熟悉交换机、路由器的运作原理
  - 掌握VLAN技术原理及基础配置方式

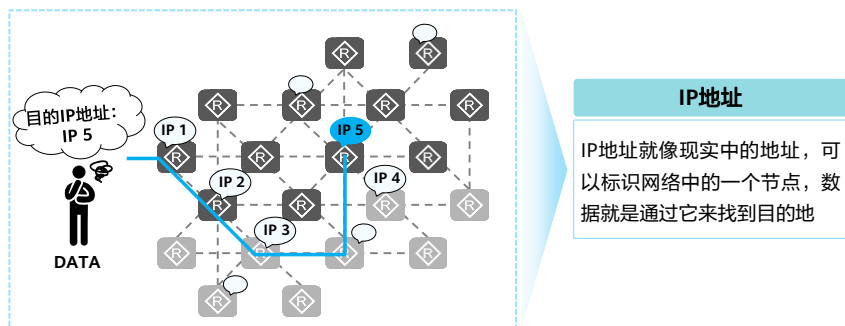
# 目录

---

1. IP地址基础
2. 网络技术介绍
3. 交换技术基础
4. 路由技术基础

# 什么是IP地址

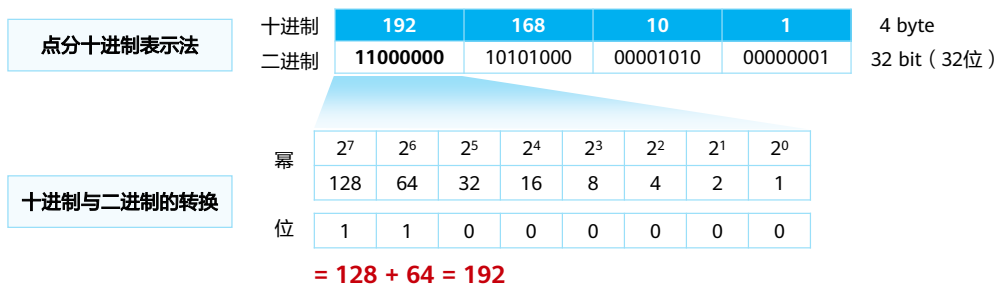
- IP地址是指在网络中用于标识发送或接收数据报文设备的唯一的逻辑地址
- IP地址的主要作用：
  - 标识主机或网络设备（标识其网络接口，提供其在网络中的位置）
  - 网络寻址



- 在IP网络上，如果用户要将一台计算机连接到Internet上，就需要申请一个IP地址。IP地址就像现实中的地址，可以标识网络中的一个节点，数据就是通过它来找到目的地的。即我们通过IP地址实现全球范围内的网络通信。
- IP地址是网络设备接口的属性，不是网络设备本身的属性。当我们说给某台设备分配一个IP地址时，实质上是指给这台设备的某个接口分配一个IP地址。如果设备有多个接口，通常每个接口都至少需要一个IP地址。
- 注：需要使用IP地址的接口，通常是路由器和计算机的接口。

## IP地址表示

- 一个IPv4地址有32 bit
- IPv4地址通常采用“点分十进制”表示



- IPv4地址范围：0.0.0.0~255.255.255.255

- IP地址表示：
  - IP地址的长度是32 bit，由4个字节组成。为了阅读和书写方便，IP地址通常采用点分十进制数来表示。
- 点分十进制表示法：
  - 这种IP地址表现形式因为方便理解所以广泛使用，但通信设备在对IP地址进行计算时使用的是二进制的操作方式，因此掌握十进制、二进制的转换运算非常有必要。
- IPv4地址范围：
  - 00000000.00000000.00000000.00000000~11111111.11111111.11111111.11111111，即0.0.0.0~255.255.255.255。



## IP地址构成

- **网络部分**：用来标识一个网络
- **主机部分**：用来区分一个网络内的不同主机



- IPv4地址由如下两部分组成：
  - 网络部分 (网络号)：用来标识一个网络。
    - IP地址不能反映任何有关主机位置的地理信息，只能通过网络号码字段判断出主机属于哪个网络。
    - 对于网络号相同的设备，无论实际所处的物理位置如何，它们都是处在同一个网络中。
  - 主机部分 (主机号)：用来区分一个网络内的不同主机。
- 子网掩码 (Subnet Mask)，又称网络掩码 (Netmask)：
  - 网络掩码为32 bit，与IP地址的位数一样，通常也以点分十进制数来表示。
  - 网络掩码不是一个IP地址，在二进制的表示上是一堆连续的1、后面接一堆连续的0。
  - 通常将网络掩码中1的个数称为这个网络掩码的长度。如：掩码0.0.0.0的长度是0，掩码252.0.0.0的长度是6。
  - 网络掩码一般与IP地址结合使用，其中值为1的比特对应IP地址中的网络位；值为0的比特对应IP地址中的主机位，以此来辅助我们识别一个IP地址中的网络位与主机位。即网络掩码中1的个数就是IP地址的网络号的位数，0的个数就是IP地址的主机号的位数。

## IP地址分类（有类编址）

- 为了方便IP地址的管理及组网，IP地址分成五类

A类	0NNNNNNN	NNNNNNNN	NNNNNNNN	NNNNNNNN	0.0.0.0~127.255.255.255	} 分配主机使用
B类	10NNNNNN	NNNNNNNN	NNNNNNNN	NNNNNNNN	128.0.0.0~191.255.255.255	
C类	110NNNNN	NNNNNNNN	NNNNNNNN	NNNNNNNN	192.0.0.0~223.255.255.255	
D类	1110NNNN	NNNNNNNN	NNNNNNNN	NNNNNNNN	224.0.0.0~239.255.255.255	用于组播
E类	1111NNNN	NNNNNNNN	NNNNNNNN	NNNNNNNN	240.0.0.0~255.255.255.255	用于研究

- A/B/C类默认子网掩码

- A类：8 bit，0.0.0.0~127.255.255.255/8
- B类：16 bit，128.0.0.0~191.255.255.255/16
- C类：24 bit，192.0.0.0~223.255.255.255/24

网络部分

主机部分

- 为了方便IP地址的管理及组网，IP地址分成五类：

- A、B、C、D、E类的类别字段分别是二进制数0、10、110、1110、1111，通过网络号码字段的前几个比特就可以判断IP地址属于哪一类，这是区分各类地址最简单的方法。
- A、B、C三类地址是单播IP地址（除一些特殊地址外），只有这三类地址才能分配给主机接口使用。
- D类地址属于组播IP地址。
- E类地址专门用于特殊的实验目的。
- 本节内容，只关注A、B、C三类地址。

- A、B、C类地址比较：

- 使用A类地址的网络称为A类网络；使用B类地址的网络称为B类网络；使用C类地址的网络称为C类网络。
- A类网络的网络号为8 bit，个数很少，但所允许的主机接口的个数很多；首位恒定为0，地址空间为：0.0.0.0~127.255.255.255。
- B类网络的网络号为16 bit，介于A类和C类网络之间；首两位恒定为10，地址空间为：128.0.0.0~191.255.255.255。
- C类网络的网络号为24 bit，个数很多，但所允许的主机接口的个数就很少；首三位恒定为110，地址空间为：192.0.0.0~223.255.255.255。

- 注：

- 主机（Host），通常指路由器和计算机的统称。并且常把主机的某个接口的IP地址简称为主机IP地址。
- 组播地址：组播能实现一对多传递消息。

## 公网和私网IP地址

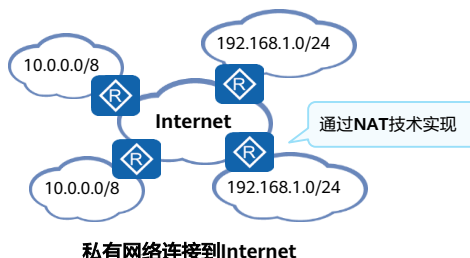
### • 公网IP地址

- IP地址是由ICANN统一分配的，以保证任何一个IP地址在Internet上的唯一性。这里的IP地址是指公网IP地址

### • 私网IP地址

- 实际上一些网络不需要连接到Internet，比如一个大学的封闭实验室内的网络，只要同一网络中的网络设备的IP地址不冲突即可。在IP地址空间里，A、B、C三类地址中各预留了一些地址专门用于上述情况，称为私网IP地址

- A类：10.0.0.0~10.255.255.255
- B类：172.16.0.0~172.31.255.255
- C类：192.168.0.0~192.168.255.255



- 为了解决IP地址短缺的问题，提出了私有地址的概念。私有地址是指内部网络或主机地址，这些地址只能用于某个内部网络，不能用于公共网络。
  - 公网IP地址：连接到Internet的网络设备必须具有由ICANN分配的公网IP地址。
  - 私网IP地址：私网IP地址的使用使得网络可以得到更为自由的扩展，因为同一个私网IP地址是可以在不同的私有网络中重复使用的。
- 私有网络连接到Internet：私有网络由于使用了私网IP地址，是不允许连接到Internet的。后来在实际需求的驱动下，许多私有网络也希望能够连接到Internet上，从而实现私网与Internet之间的通信，以及通过Internet实现私网与私网之间的通信。私网与Internet的互联，必须使用网络地址转换（NAT）技术实现。
- 注：
  - NAT（Network Address Translation），网络地址转换，其基本作用是实现私网IP地址与公网IP地址之间的转换。
  - ICANN（Internet Corporation for Assigned Names and Numbers），互联网名称与数字地址分配机构

## 特殊IP地址

- IP地址空间中，有一些特殊的IP地址，这些IP地址有特殊的含义和作用，举例如下：

特殊IP地址	地址范围	作用
有限广播地址	255.255.255.255	可作为目的地址，发往该网段所有主机（受限网关）
任意地址	0.0.0.0	“任何网络”的网络地址； “这个网络上这个主机接口”的IP地址
环回地址	127.0.0.0/8	测试设备自身的软件系统
本地链路地址	169.254.0.0/16	当主机自动获取地址失败后，可使用该网段中的某个地址进行临时通信

- 255.255.255.255
  - 这个地址称为有限广播地址，它可以作为一个IP报文的目的IP地址使用。
  - 路由器接收到目的IP地址为有限广播地址的IP报文后，会停止对该IP报文的转发。
- 0.0.0.0
  - 如果把这个地址作为网络地址，它的意思就是“任何网络”的网络地址；如果把这个地址作为主机接口地址，它的意思就是“这个网络上主机接口”的IP地址。
  - 例如：当一个主机接口在启动过程中尚未获得自己的IP地址时，就可以向网络发送目的IP地址为有限广播地址、源IP地址为0.0.0.0的DHCP请求报文，希望DHCP服务器在收到自己的请求后，能够给自己分配一个可用的IP地址。
- 127.0.0.0/8
  - 这个地址为环回地址，它可以作为一个IP报文的目的IP地址使用。其作用是测试设备自身的软件系统。
  - 一个设备产生的、目的IP地址为环回地址的IP报文是不可能离开这个设备本身的。
- 169.254.0.0/16
  - 如果一个网络设备获取IP地址的方式被设置成了自动获取方式，但是该设备在网络上又没有找到可用的DHCP服务器，那么该设备就会使用169.254.0.0/16网段的某个地址来进行临时通信。
- 注：DHCP（Dynamic Host Configuration Protocol），动态主机配置协议，用于动态分配网络配置参数，如IP地址。

## 子网掩码及主机可用地址

- 我们通常把一个网络号所定义的网络范围称为一个网段
- **子网掩码**：用于计算IP地址中的网络号（网络地址）和主机号（主机地址），和IP地址一起使用

例如：192.168.10.0/24

192	168	10	00000000
-----	-----	----	----------

- **广播地址**：用于向该网络中的所有主机发送数据的特殊地址

例如：192.168.10.255/24

192	168	10	11111111
-----	-----	----	----------

- **可用地址**：可分配给网络中的节点或网络设备接口的地址

例如：192.168.10.1/24

192	168	10	00000001
-----	-----	----	----------

### 注意

- 网络地址和广播地址不能直接被节点或网络设备所使用。
- 一个网段可用地址数量为： $2^n - 2$ （n：主机部分的比特位数）

- 可用地址
  - 又称主机地址，可以分配给具体的主机接口使用。
- 一个网段可用地址数量计算
  - 一个网段的主机位为n位，则IP地址数为： $2^n$ ，可用IP地址数为： $2^n - 2$  (减去网络地址和广播地址)。

## IP地址计算

- 计算172.16.10.1/16这个B类地址的网络地址、广播地址以及可用地址数

	172	16	00001010	00000001	
IP地址	1 0 1 0 1 1 0 0	0 0 0 1 0 0 0 0	0 0 0 0 1 0 1 0	0 0 0 0 0 0 0 1	
子网掩码	1 1 1 1 1 1 1 1	1 1 1 1 1 1 1 1	0 0 0 0 0 0 0 0	0 0 0 0 0 0 0 0	
网络地址	1 0 1 0 1 1 0 0	0 0 0 1 0 0 0 0	0 0 0 0 0 0 0 0	0 0 0 0 0 0 0 0	主机位全为0，得出网络地址 172.16.0.0
广播地址	1 0 1 0 1 1 0 0	0 0 0 1 0 0 0 0	1 1 1 1 1 1 1 1	1 1 1 1 1 1 1 1	主机位全为1，得出广播地址 172.16.255.255
IP地址数	$2^{16}=65536$				
可用IP地址数	$2^{16}-2=65534$				
可用IP地址范围	172.16.0.1~172.16.255.254				

**练习**  
请计算10.128.20.10/8这个A类地址的网络地址、广播地址以及可用地址数

- 网络地址：将IP地址的主机位全设为0，所得结果是该IP地址所在网络的网络地址。
- 广播地址：将IP地址的主机位全设为1，所得结果是该IP地址所在网络的广播地址。
- IP地址数： $2^n$ ，n为主机位位数。
- 可用IP地址数： $2^n-2$ ，n为主机位位数。
- 练习题答案：
  - 网络地址：10.0.0.0
  - 广播地址：10.255.255.255
  - IP地址数：224
  - 可用IP地址数：224-2
  - 可用IP地址范围：10.0.0.1~10.255.255.254

## 子网划分

- 为什么需要子网划分
- 子网划分使用的技术：变长子网掩码VLSM（Variable-Length Subnet Mask）
  - VLSM允许一个组织根据网络规模的大小来划分网络。将一个网络再划分为更小的多个子网，以供不同组织部门使用
- 举例：某公司分配到C类地址201.222.5.0。假设需要20个子网，每个子网有5台主机，我们该如何划分

子网地址	可用的主机地址
201.222.5.8/29	201.222.5.9~201.222.5.14
201.222.5.16/29	201.222.5.17~201.222.5.22
.....	.....
201.222.5.232/29	201.222.5.233~201.222.5.238
201.222.5.240/29	201.222.5.241~201.222.5.246

- 为什么需要子网划分？
- 现实中，如果将一个A类网络分配给一个组织使用，但是这个组织内部的计算机又没有16777214这么多台，这就会造成大量的IP地址闲置与浪费。因此，需要有一种更加灵活的方式根据网络规模的大小来划分网络。思路就是将一个网络再划分为更小的多个子网，以供不同组织使用。
- 这种划分网络的方式所使用的技术就是可变长度子网掩码（VLSM）。可变长度子网掩码在公网与企业内部中都可以使用。
- 在这个例子中，网段地址是一个C类地址：201.222.5.0/24。假设需要20个子网，其中每个子网5个主机，就要把主机地址的最后一个八位组分成子网部分和主机部分。
- 子网部分的位数决定了子网的数目。在这个例子中，因为是C类地址，所以子网部分和主机部分总共是8位，因为 $2^4 < 20 < 2^5$ ，所以子网部分占有5位，最大可提供32（ $2^5$ ）个子网。剩余3位为主机部分 $2^3=8$ 。主机部分全是0的IP地址，是子网网络地址；主机部分全是1的IP地址是本子网的广播地址。这样就剩余6个主机地址，可以满足需要。
- 每个网段分别为：
  - 201.222.5.0~201.222.5.7
  - 201.222.5.8~201.222.5.15
  - 201.222.5.16~201.222.5.23
  - .....
  - 201.222.5.232~201.222.5.239
  - 201.222.5.240~201.222.5.247
  - 201.222.5.248~201.222.5.255



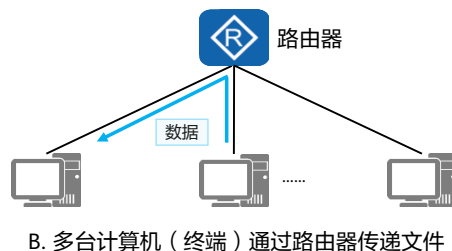
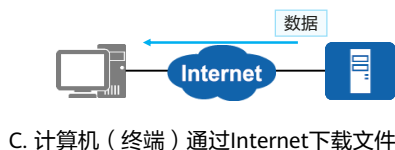
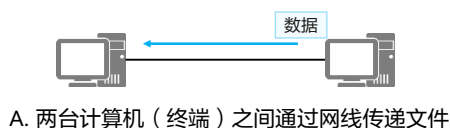
# 目录

---

1. IP地址基础
- 2. 网络技术介绍**
  - 网络基础
    - 常用的网络设备
    - 常用协议介绍
3. 交换技术基础
4. 路由技术基础

## 网络通信基本概念

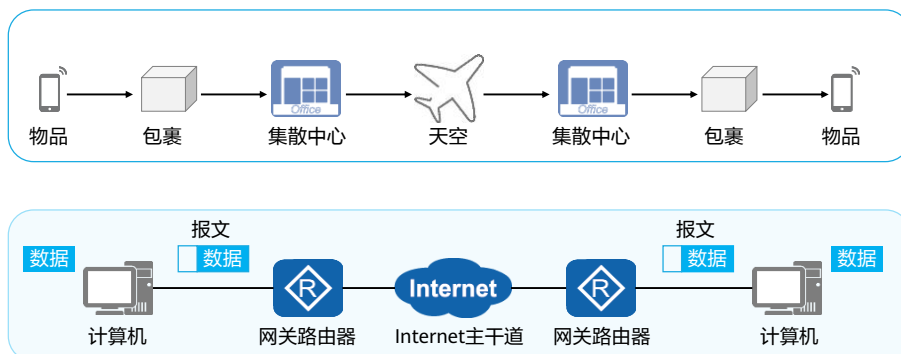
- 通信：是指人与人、人与物、物与物之间通过某种媒介和行为进行的信息传递与交流
- 网络通信：是指终端设备之间通过计算机网络进行的通信
- 网络通信的举例：



- 网络通信的例子：
  - A：两台计算机通过一根网线相连，就组成了一个最简单的网络。
  - B：由一台路由器（或交换机）和多台计算机设备组成的小型网络。在这样的网络中，通过路由器的中转，每两台计算机之间都可以自由地传递文件。
  - C：当计算机想从某个网址获取文件时，必须先接入Internet，然后才能下载。
- Internet（译名：英特网、互联网、网际网等），是目前世界上规模最大的计算机网络，其前身诞生于1969年的ARPAnet（Advanced Research Projects Agency Network）。Internet的广泛普及和应用是当今信息时代的标志性内容之一。

## 信息传递过程

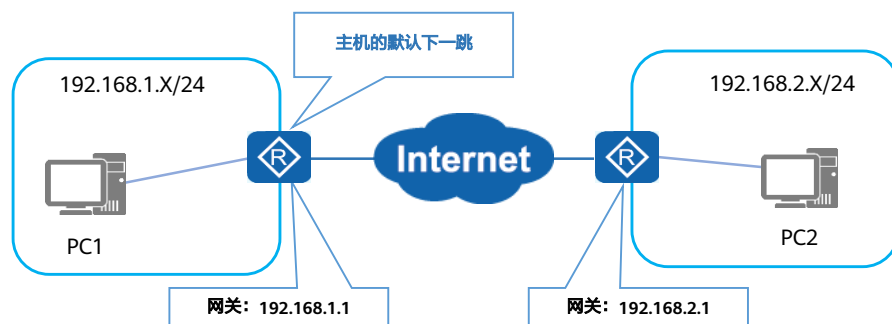
- 虚拟的信息传递与真实的物品传递过程有许多相似之处



- 快递过程与网络通信过程的对比：
- 需要快递的物品：
  - 应用程序生成需要传递的信息（或数据）。
- 物品被包装起来形成包裹，并粘贴含有收货人姓名、地址的快递单：
  - 应用程序将数据打包成原始的“数据载荷”，并添加“头部”和“尾部”形成报文，报文中的重要信息是接收者的地址信息，即“目的地址”。
  - 在一个信息单元的基础上，增加一些新的信息段，使其形成一个新的信息单元，这个过程称为“封装”。
- 包裹被送到集散中心，集散中心对包裹上的目的地址进行分检，去往同一个城市的物品被放入同一架飞机，并飞向天空：
  - 报文通过网线到达“网关”，网关收到报文后，对其“解封装”，读取目的地址，再重新封装，并根据目的地址不同，送往不同的“路由器”，通过网关及路由器的传递，报文最终离开本地网络，进入Internet的干道进行传输。
  - 其中，网线所起的作用跟公路一样，它是信息传输的介质。

- 飞机抵达目的机场后，包裹被取出进行分检，去往同一地区的包裹，被送到了同一集散中心：
  - 报文经过Internet干道的传输，到达目的地址所在的本地网络，本地网络的网关或路由器对报文进行解封装和封装，并根据目的地址决定发往相应的下一台路由器，最终到达目的计算机所在网络的网关。
- 集散中心根据包裹上的目的地址进行分检，快递员送包裹上门，收件人拆开包裹，确认物品完好无损后收下。整个快递过程完成。
  - 报文到达目的计算机所在网络的网关，解封装和封装，然后根据目的地址发往相应的计算机。计算机收到报文后，对报文进行校验处理，校验无误后，接收下报文，并将其中的数据载荷交由相应的应用程序进行处理。一次完整的网络通信过程就结束了。

## 什么是网关



- 网关（Gateway）又称网间连接器、协议转换器。默认网关在网络层以上实现网络互连。
- 大家都知道，从一个房间走到另一个房间，必然要经过一扇门。同样，从一个网络（网段）向另一个网络（网段）发送信息，也必须经过一道“关口”，这道关口就是网关。
- 网关主要意义：
  - 网关设备主要意义：在主机（无论是PC、服务器、路由器还是防火墙，这里统称主机）想要访问另一网段地址时，负责反馈ARP报文，并接收、转发随后而来的数据报文。
  - 网关配置主要意义：在主机上生成默认路由，下一跳指向网关设备。

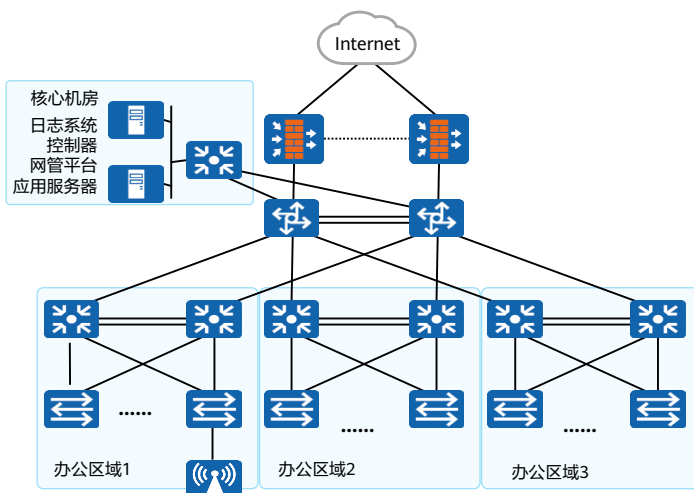
## 通信网络基本组成架构

- 通信网络

由路由器、交换机、防火墙，以及  
个人电脑、网络打印机、服务器等  
设备构成的通信网络

- 功能

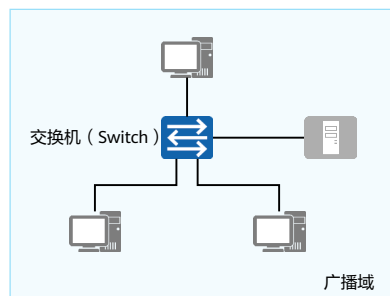
数据通信网络最基本的功能是实现  
数据互通



- 以企业数据中心网络为例，企业对于数据中心的主体需求主要是业务运行与计算、数据存储以及服务访问。
- 那么围绕这几点需求，数据中心网络的主体作用就是提供设备与设备、用户与设备之间的互联并提供业务的外部访问能力。其中，就需要接入交换机设备连接各办公区的用户主机，汇聚交换机汇聚接入交换机侧的流量，路由器负责不同办公区、内部与外部不同网络间的通信转发，防火墙负责不同安全等级的区域、内部与外部网络的安全接入与访问控制。每种网络设备在整体架构中各司其职，共同完成整个数据中心网络的通信需求。

## 网络设备 - 交换机

- 交换机：距离终端用户最近的设备，用于终端用户接入网络、对数据帧进行交换等
  - 实现终端设备（PC、服务器等）的网络接入
  - 隔离冲突域
  - 广播转发未知报文
  - MAC地址学习、维护MAC地址表
  - 基于MAC地址表转发报文

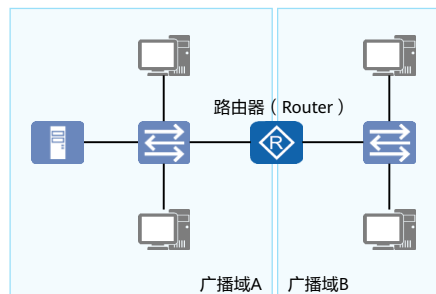


- 交换机：
  - 在园区网络中，交换机一般来说是距离终端用户最近的设备，接入层的交换机一般为二层交换机，又称为以太网交换机，二层是指TCP/IP参考模型的数据链路层。
  - 以太网交换机可以实现：数据帧的交换、终端用户设备的接入、基本的接入安全功能、二层链路的冗余等。
  - 广播域：一个节点发送一个广播报文其余节点都能够收到的节点的集合。
  - 冲突域：在同一个网络上两个设备同时进行传输会产生冲突，在网络报文转发中产生与发生冲突的区域称为冲突域。
  - MAC（Media Access Control）地址：在网络中唯一标识一个网卡，每个网卡都需要且有唯一的一个MAC地址。
  - MAC地址表：每台交换机中都有一个MAC地址表，存放了MAC地址与交换机端口编号之间的映射关系。

## 网络设备 - 路由器

- 路由器：网络层设备，可以在因特网中进行数据报文转发。路由器根据所收到的报文的目地地址选择一条合适的路径，将报文传送到下一个路由器或目的地，路径中最后的路由器负责将报文送交目的主机

- 实现同类型网络或异种网络之间的通信
- 隔离广播域
- 维护路由表（Routing Table）、运行路由协议
- 路径（路由信息）选择、IP报文转发
- 广域网接入、网络地址转换
- 连接通过交换机组建的二层网络



- 路由器：
  - 路由器工作在TCP/IP参考模型的网络层；
  - 路由器可以实现：维护路由表和路由信息、路由发现及路径选择、数据转发、隔离广播域、广域网接入和网络地址转换及特定的安全功能。



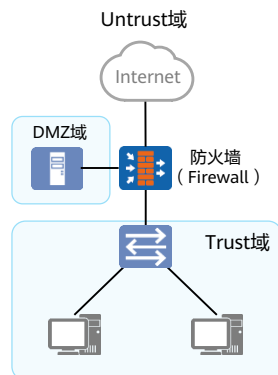
## 问题研讨

- 作为网络连接设备，二层交换机和路由器有着什么样的区别？
- 使用场景分别是什么？



## 网络设备 - 防火墙

- 防火墙：网络安全设备，用于控制两个网络之间的安全通信。它通过监测、限制、更改跨越防火墙的数据流，尽可能地对外部屏蔽网络内部的信息、结构和运行状况，以此来实现对网络的安全保护
  - 隔离不同安全级别的网络
  - 实现不同安全级别的网络之间的访问控制（安全策略）
  - 用户身份认证
  - 实现远程接入功能
  - 实现数据加密及虚拟专用网业务
  - 执行网络地址转换
  - 其他安全功能



- 防火墙：
  - 是位于两个信任程度不同的网络之间（如企业内部网络和Internet之间）的设备，它对两个网络之间的通信进行控制，通过强制实施统一的安全策略，防止对重要信息资源的非法存取和访问，以达到保护系统安全的目的。

# 目录

---

1. IP地址基础
2. 网络技术介绍
  - 网络基础
    - 网络参考模型和数据封装
  - 常用协议介绍
3. 交换技术基础
4. 路由技术基础

## OSI参考模型

- 为了解决网络之间兼容性的问题，帮助各个厂商生产出可兼容的网络设备，国际标准化组织ISO（International Standards Organization）于1984年提出了开放系统互连参考模型OSI/RM（Open System Interconnection Reference Model），它很快成为计算机网络通信的基础模型

7. 应用层	对应用程序提供接口
6. 表示层	进行数据格式的转换，以确保一个系统生成的应用层数据能够被另外一个系统的应用层所识别和理解
5. 会话层	在通信双方之间建立、管理和终止会话
4. 传输层	建立、维护和取消一次端到端的数据传输过程。控制传输节奏的快慢，调整数据的排序等等
3. 网络层	定义逻辑地址；实现数据从源到目的地的转发
2. 数据链路层	将分组数据封装成帧；在数据链路上实现数据的点到点、或点到多点方式的直接通信；差错检测
1. 物理层	在媒介上传输比特流；提供机械的和电气的规约

- OSI模型（Open Systems Interconnection Model），由国际化标准组织ISO（The International Organization for Standardization）收录在ISO 7489标准中并于1984年发布。
- OSI参考模型又被称为七层模型，由下至上依次为：
  - 物理层：在设备之间传输比特流，规定了电平、速度和电缆针脚等物理特性。
  - 数据链路层：将比特组合成字节，再将字节组合成帧，使用链路层地址（以太网使用MAC地址）来访问介质，并进行差错检测。
  - 网络层：定义逻辑地址，供路由器确定路径，负责将数据从源网络传输到目的网络。
  - 传输层：提供面向连接或非面向连接的数据传递以及进行重传前的差错检测。
  - 会话层：负责建立、管理和终止表示层实体之间的通信会话。该层的通信由不同设备中的应用程序之间的服务请求和响应组成。
  - 表示层：提供各种用于应用层数据的编码和转换功能，确保一个系统的应用层发送的数据能被另一个系统的应用层识别。
  - 应用层：OSI参考模型中最靠近用户的一层，为应用程序提供网络服务。

## TCP/IP参考模型

- 因为OSI协议栈比较复杂，且TCP和IP两大协议在业界被广泛使用，所以TCP/IP参考模型成为了互联网的主流参考模型



- TCP/IP模型在结构上与OSI模型类似，采用分层架构，同时层与层之间联系紧密。
- TCP/IP标准参考模型将OSI中的数据链路层和物理层合并为网络接入层，这种划分方式其实是有悖于现实协议制定情况的，故融合了TCP/IP标准模型和OSI模型的TCP/IP对等模型被提出，后面的讲解也都将基于这种模型。
- TCP/IP起源于60年代末美国政府资助的一个分组交换网络研究项目，到90年代已发展成为计算机之间最常用的组网形式。它是一个真正的开放系统，因为协议族的定义及其多种实现，可以不用花钱或花很少的钱就可以公开得到。它成为被称作“全球互联网”或“因特网（Internet）”的基础。
- 与OSI参考模型一样，TCP（Transfer Control Protocol）/IP（Internet Protocol）协议（传输控制协议/网际协议）也分为不同的层次开发，每一层负责不同的通信功能。但是，TCP/IP协议简化了层次设计，只有五层：应用层、传输层、网络层、数据链路层和物理层。从上图可以看出，TCP/IP协议栈与OSI参考模型有清晰的对应关系，覆盖了OSI参考模型的所有层次。应用层包含了OSI参考模型所有高层协议。
- 因为TCP/IP协议栈支持所有标准的物理层和数据链路层协议，而且物理层和数据链路层在前面已经做过简述，所以本章不对TCP/IP协议的物理层和数据链路层做进一步的描述。关于这两层协议和标准的深入细节，在后续章节会有讲解。

- OSI参考模型与TCP/IP协议栈的异同点：

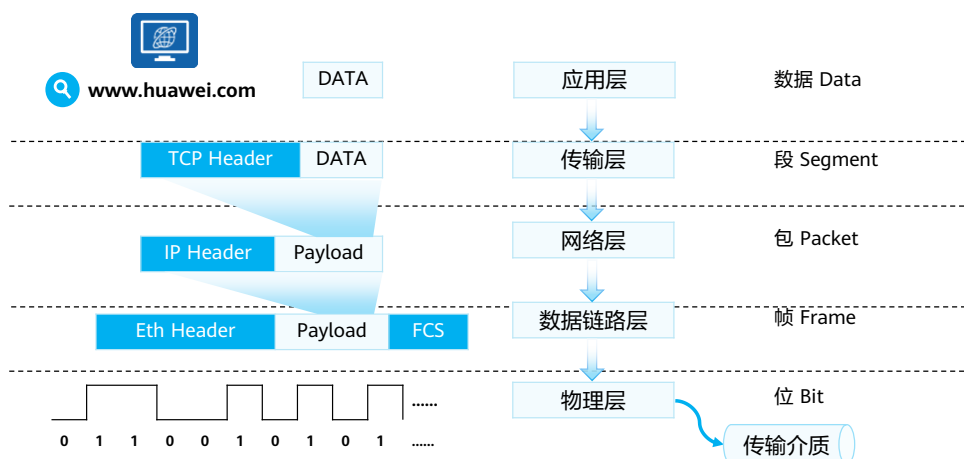
- 相同点

- 都是分层结构，并且工作模式一样，都要求层和层之间具备很密切的协作关系；
    - 有相同的应用层，传输层，网络层，数据链路层，物理层；（注意：这里为了方便比较，才将TCP/IP分为5层，在其他很多文献资料里都把数据链路层和物理层合并为数据链路层或网络接口层：（ Network Access Layer ）；
    - 都使用包交换技术（ Packet-Switched ）；
    - 网络工程师必须都要了解这两个模型。

- 不同点

- TCP/IP把表示层和会话层都归入了应用层；
    - TCP/IP的结构比较简单，因为分层少；
    - TCP/IP标准是在Internet网络不断的发展中建立的，基于实践，有很高的信任度。相比较而言，OSI参考模型是基于理论上的，是作为一种向导！

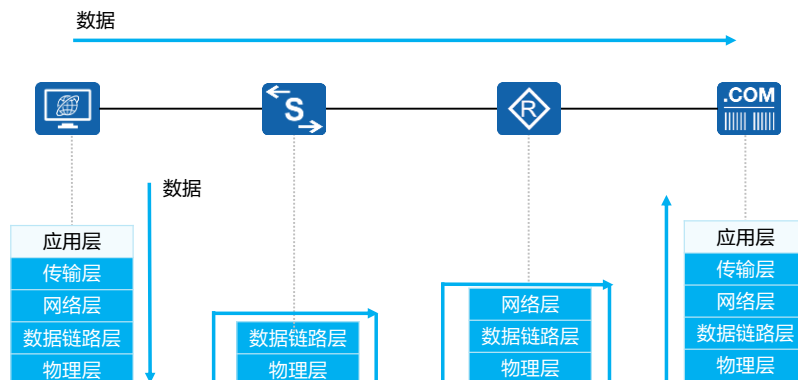
## 发送方数据封装



- 假设你正在通过网页浏览器访问华为官网，当你输入完网址，敲下回车后，计算机内部会发生下列事情：
  - IE浏览器（应用程序）调用HTTP（应用层协议），完成应用层数据的封装（图中DATA还应包括HTTP头部，此处省略）。
  - HTTP依靠传输层的TCP进行数据的可靠性传输，将封装好的数据传递到TCP模块。
  - TCP模块给应用层传递下来的Data添加上相应的TCP头部信息（源端口、目的端口等）。此时的PDU被称作Segment（段）。
  - 在IPv4网络中，TCP模块会将封装好的Segment传递给网络层的IPv4模块（若在IPv6环境，会交给IPv6模块进行处理）。
  - IPv4模块在收到TCP模块传递来的Segment之后，完成IPv4头部的封装，此时的PDU被称为Packet（包）。
  - 由于使用了Ethernet作为数据链路层协议，故在IPv4模块完成封装之后，会将Packet交由数据链路层的Ethernet模块（例如以太网卡）处理。
  - Ethernet模块在收到IPv4模块传递来的Packet之后，添加上相应的Ethernet头部信息和FCS帧尾，此时的PDU被称为Frame（帧）。
  - 在Ethernet模块封装完毕之后，会将数据传递到物理层。
  - 根据物理介质的不同，物理层负责将数字信号转换成电信号，光信号，电磁波（无线）信号等。
  - 转换完成的信号在网络中开始传递。

## 中间网络数据传输

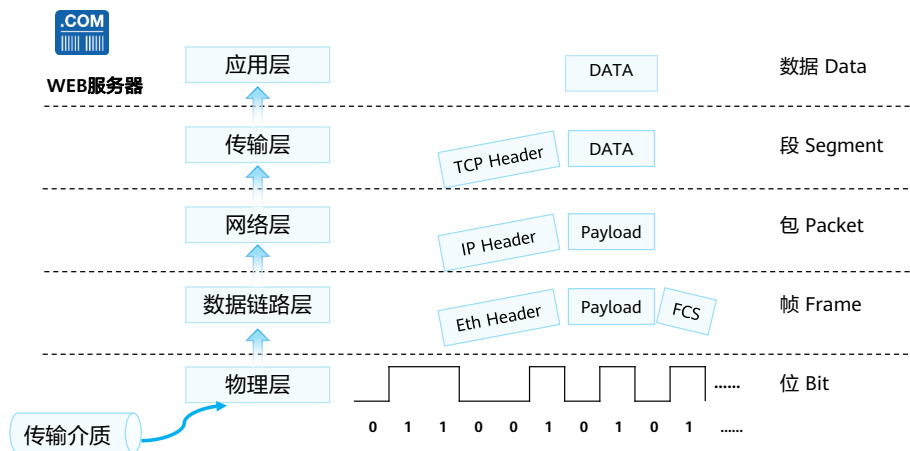
- 封装好的完整数据，将会在网络中被传递



- 一般情况下：
  - 网络中的二层设备（如以太网交换机）只会解封装数据的二层头部，根据二层头部的信息进行相应的“交换”操作。
  - 网络中的三层设备（如路由器）只会解封装到三层头部，并且根据三层头部的信息进行相应的“路由”操作。
  - 注：“交换”和“路由”的详细细节和原则，将会在后面的课程中详细介绍。



## 接收方数据解封装



- 经过中间网络传递之后，数据最终到达目的服务器。根据不同的协议头部的信息，数据将被一层层地解封装并做相应的处理和传递，最终交由WEB服务器上的应用程序进行处理。

# 目录

---

1. IP地址基础
- 2. 网络技术介绍**
  - 网络基础
  - 网络参考模型和数据封装
    - 常用协议介绍
3. 交换技术基础
4. 路由技术基础

## TCP/IP常见协议

- TCP/IP协议栈定义了一系列的标准协议

应用层	Telnet	FTP	TFTP	SNMP
	HTTP	SMTP	DNS	DHCP
传输层	TCP		UDP	
网络层	ICMP		IGMP	
	IP			
数据链路层	PPPoE			
	Ethernet		PPP	
物理层	.....			

- 协议介绍：
  - HTTP: Hypertext Transfer Protocol, 超文本传输协议, 用来访问在网页服务器上的各种页面。
  - FTP: File Transfer Protocol, 文件传输协议, 为文件传输提供了途径, 它允许数据从一台主机传送到另一台主机上。
  - DNS: Domain Name Service, 域名称解析服务, 用于实现从主机域名到IP地址之间的转换。
  - TCP: Transmission Control Protocol, 传输控制协议, 为应用程序提供可靠的面向连接的通信服务。目前, 许多流行的应用程序都使用TCP。
  - UDP: User Datagram Protocol, 用户数据报协议, 提供了无连接通信, 且不对传送数据包进行可靠性的保证。
  - IP: Internet Protocol, 互联网协议, 将传输层的数据封装成数据包并完成源站点到目的站点的转发, 提供无连接的、不可靠的服务。
  - IGMP: Internet Group Management Protocol, 因特网组管理协议, 负责IP组播成员管理的协议。它用来在IP主机和与其直接相邻的组播路由器之间建立、维护组播组成员关系。
  - ICMP: Internet Control Message Protocol, 网际报文控制协议, 基于IP协议在网络中发送控制消息, 提供可能发生在通信环境中的各种问题反馈。通过这些信息, 使管理者可以对所发生的问题作出诊断, 然后采取适当的措施解决。
  - ARP: Address Resolution Protocol, 地址解析协议, 是根据IP地址获取数据链路层地址的一个TCP/IP协议, 其主要功能: 将IP地址解析为MAC地址、维护IP地址与MAC地址的映射关系的缓存 (ARP表)、实现网段内重复IP地址的检测。

# TCP协议

- TCP为应用程序提供一种面向连接的、可靠的服务

- TCP的可靠性

- 面向连接的传输
- 最大报文段长度
- 传输确认机制
- 首部和数据的检验和
- 流量控制

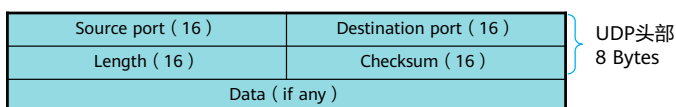
Source port ( 16 )		Destination port ( 16 )	
Sequence number ( 32 )			
Acknowledgement number ( 32 )			
Header length ( 4 )	Reserved ( 6 )	Control bits ( 6 )	Window ( 16 )
Checksum ( 16 )		Urgent ( 16 )	
Options			
Data ( varies )			

TCP头部  
20 Bytes

- TCP传输控制协议为应用程序提供可靠的、面向连接的服务。
- TCP通过如下几方面提供其可靠性：
  - 面向连接的传输：TCP的任何一方进行数据传输之前必须在双方之间建立连接。
  - MSS（最大报文段长度）：表示TCP发往另一端的最大报文段长度。当一个连接建立时，连接的双方都要通告各自的MSS，以充分利用带宽资源。
  - 传输确认机制：当TCP传输一个数据段后，它启动一个定时器，等待目的端确认收到这个报文段。如果不能及时收到目的端的确认报文，TCP将重传该数据段。
  - 首部和数据的检验和：TCP将保持首部和数据的检验和，这是一个端到端的检验。目的是检测数据在传输过程中的变化。如果收到段的检验和有所差错，TCP将丢弃该报文段并不确认收到此报文段。因此TCP将启动重传机制。
  - 流量控制：TCP还能提供流量控制。TCP连接的每一方都有一个固定大小的缓冲空间。TCP接收端只允许另一端发送接收缓冲区所能够容纳的数据。这能够防止较快的主机致使较慢的主机缓冲区溢出。

## UDP协议

- UDP为应用程序提供面向无连接的服务。传输数据之前，源端和目的端不需要建立连接
- 不需要维护连接状态，收发状态等，因此服务器可同时向多个客户端传输相同的消息
- UDP适用于对传输效率要求高的运用



- UDP为应用程序提供面向无连接的服务，所以在传数据之前源端和目的端之间不必要像TCP一样需要事先建立连接。正由于UDP是基于无连接的传输协议，所以UDP不需要去维护连接状态和收发状态，因此服务器可同时向多个客户端传输相同的消息。
- UDP适用于对传输效率要求高的应用，应用层能提供可靠性的保障。如常用于认证计费的Radius协议、RIP路由协议都是基于UDP的。

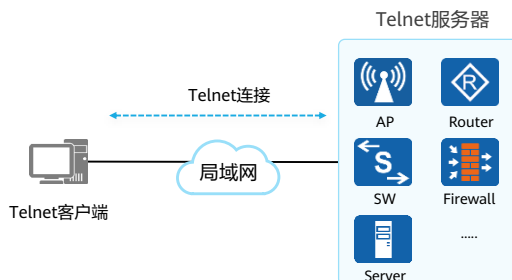
## TCP VS UDP

TCP	UDP
<ul style="list-style-type: none"><li>• 面向连接</li><li>• 可靠传输，采用流量控制和拥塞控制</li><li>• 头部封装20-60 B</li><li>• 适用于要求可靠传输的应用，例如文件传输等</li></ul>	<ul style="list-style-type: none"><li>• 面向无连接</li><li>• 不可靠传输，报文可靠性依靠上层应用保证</li><li>• 头部封装小，8 B</li><li>• 适用于实时应用，例如视频会议等</li></ul>

- TCP协议可靠，但也是因为它的可靠机制，导致它传输报文效率不高，封装开销也较高。
- UDP面向无连接，不可靠，但传输效率相对较高。

## Telnet协议

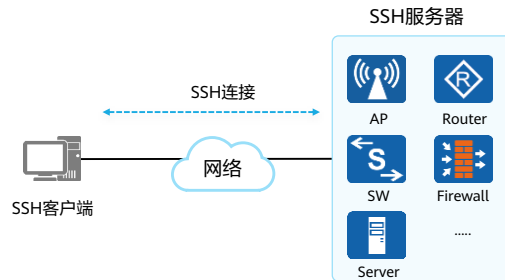
- Telnet是数据网络中提供远程登录服务的标准协议，为用户提供了在本地计算机上完成远程设备工作的能力，其数据采用明文方式传输
- 用户通过Telnet客户端程序连接到Telnet服务器。用户在Telnet客户端中输入命令，这些命令会在服务器端运行，就像直接在服务端的控制台上输入一样



- Telnet协议的优势在于通过它可以远程地登录到网络设备上，对网络设备进行配置，为网络管理员异地管理网络设备提供了极大的方便。
- 但是Telnet协议存在三个弱点：
  - 数据传输采用明文方式，传输的数据没有任何机密性可言。
  - 认证机制脆弱。用户的认证信息在网络上以明文方式传输，很容易被窃听；Telnet 只支持传统的密码认证方式，很容易被攻击。
  - 客户端无法真正识别服务器的身份，攻击者很容易进行“伪服务器欺骗”。SSH协议正是为了克服Telnet协议存在的问题而诞生的。

# SSH协议

- SSH协议是一种在不安全的网络环境中，通过加密和认证机制，实现安全的远程访问以及文件传输等业务的网络安全协议
- 为解决Telnet协议安全性上的不足，SSH协议应运而生，为用户安全接入远端设备管理提供了保障
- SSH协议采用客户端/服务器架构，分为传输层、认证层和连接层



- SSH协议层：
  - 传输层：主要用来在客户端和服务端之间建立一条安全的加密通道，为用户认证、数据交互等对数据传输安全性要求较高的阶段提供足够的机密性保护。
  - 认证层：运行在传输层协议之上，完成服务器对登录用户的认证。
  - 连接层：负责在加密通道上划分出若干逻辑通道，以运行不同的应用。它运行在认证层协议之上，提供交互会话、远程命令执行等服务。
- SSH的报文交互主要有以下几个阶段：
  - 连接建立
  - 版本协商
  - 算法协商
  - 密钥交换
  - 用户认证
  - 服务请求
  - 数据传输和连接关闭



## Telnet VS SSH

Telnet	SSH
<ul style="list-style-type: none"><li>• 数据传输采用明文方式</li><li>• 认证机制脆弱，用户的认证信息在网络上以明文方式传输</li><li>• 只支持传统的密码认证方式</li><li>• 客户端无法真正识别服务器的身份</li></ul>	<ul style="list-style-type: none"><li>• 数据传输采用密文的方式</li><li>• 用户的认证信息以密文的方式传输</li><li>• 除了传统的密码认证，SSH服务器还可以采用多种方式对用户进行认证（如安全性级别更高的公钥认证）</li><li>• 客户端和服务端之间通信使用动态生成的加解密密钥</li><li>• 为客户端提供了认证服务器的功能</li></ul>

- SSH协议在发送数据时会对数据进行加密操作，数据传输更安全，常用在需要加密认证的场景。
- 但在一些测试、无需加密的场合（如局域网），Telnet协议仍常被使用。

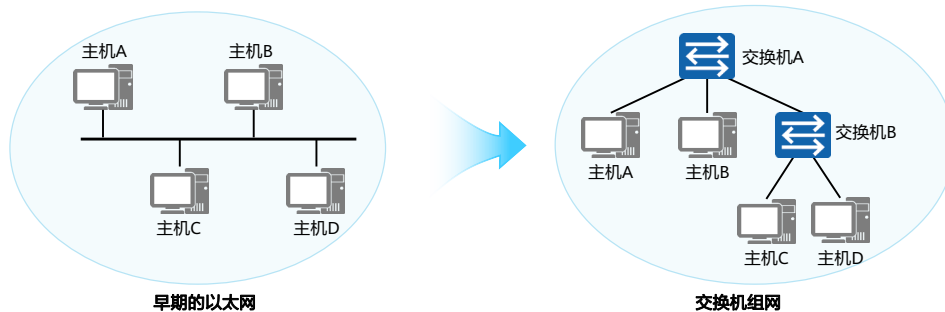
# 目录

---

1. IP地址基础
2. 网络技术介绍
- 3. 交换技术基础**
  - 以太网交换基础
    - VLAN技术基础
    - VLAN基础配置
4. 路由技术基础

# 以太网协议

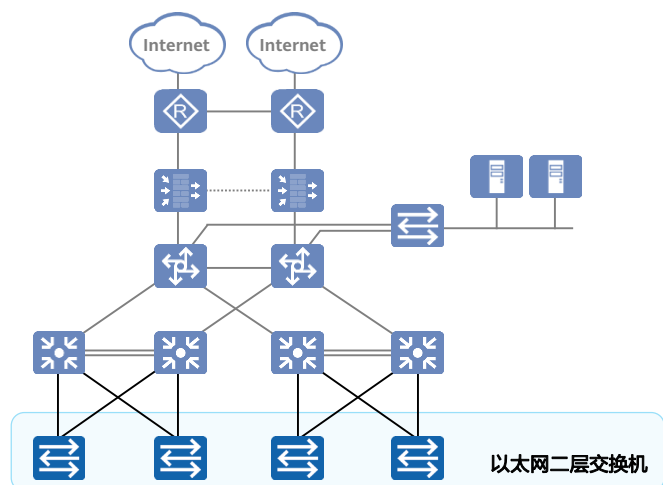
- 以太网是当今现有局域网（Local Area Network，LAN）采用的最通用的通信协议标准，该标准定义了局域网中采用的电缆类型和信号处理方法



- 早期的以太网：
  - 以太网是建立在CSMA/CD机制上的广播型网络。冲突的产生是限制以太网性能的重要因素，早期的以太网设备如集线器HUB是物理层设备，不能隔绝冲突扩散，限制了网络性能的提高。
- 交换机组网：
  - 交换机作为一种能隔绝冲突的二层网络设备，极大地提高了以太网的性能，并替代HUB成为主流的以太网设备。但是交换机对网络中的广播数据流量不做任何限制，这也影响了网络的性能。

## 以太网二层交换机

- 以太网二层交换机转发数据的端口都是以太网口，并且只能针对数据的二层头部（以太网数据帧头）中的MAC地址进行寻址并转发数据。



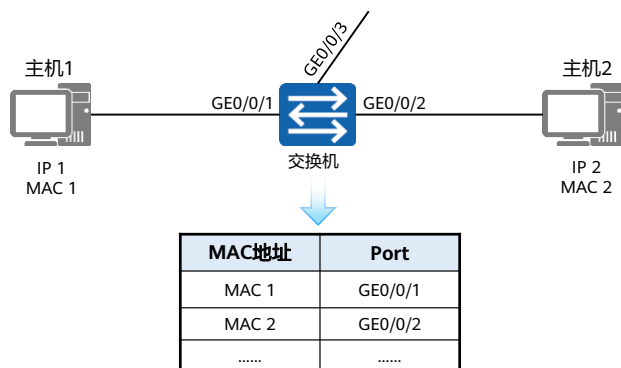
42 Huawei Confidential

HUAWEI

- 通信网络架构及组成在之前的课件中已介绍，以太网二层交换机主要用来接入用户及终端设备，处在通信网络架构的边缘，属于接入设备。
- 以太网二层交换机：
  - 在园区网络中，交换机一般来说是距离终端用户最近的设备，用于终端接入园区网，接入层的交换机一般为二层交换机。
  - 二层交换设备工作在TCP/IP对等模型的第二层，即数据链路层，它对数据包的转发是建立在MAC（Media Access Control）地址基础之上的。
- 以太网三层交换机：
  - 不同局域网之间的网络互通需要由路由器来完成。随着数据通信网络范围的不断扩大，网络业务的不断丰富，网络间互访的需求越来越大，而路由器由于自身成本高、转发性能低、接口数量少等特点无法很好地满足网络发展的需求。因此出现了三层交换机这样一种能实现高速三层转发的设备。
- 注意：本课程中所涉及的交换机，均指以太网二层交换机。

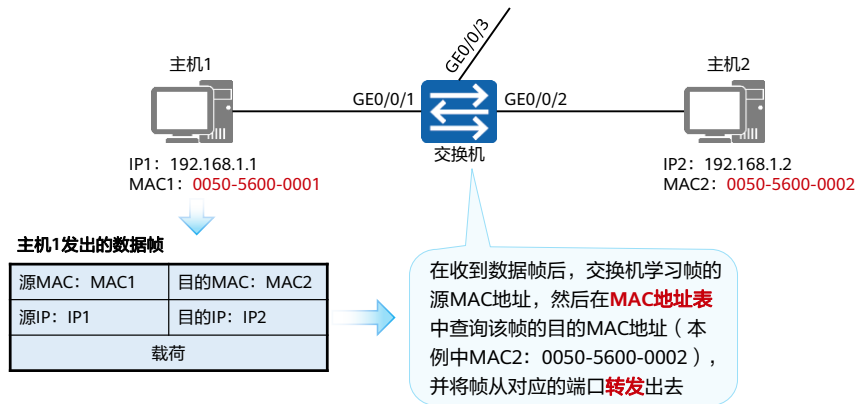
## MAC地址表

- 每台交换机中都有一个MAC地址表，存放了MAC地址与交换机端口编号之间的映射关系。



- MAC地址表记录了交换机学习到的其他设备的MAC地址与接口的对应关系。交换机在转发数据帧时，根据数据帧的目的MAC地址查询MAC地址表。如果MAC地址表中包含与该帧目的MAC地址对应的表项，则直接通过该表项中的出接口转发该报文；如果MAC地址表中没有包含该帧目的MAC地址对应的表项时，交换机将采取泛洪方式在除接收接口外的所有接口发送该报文。

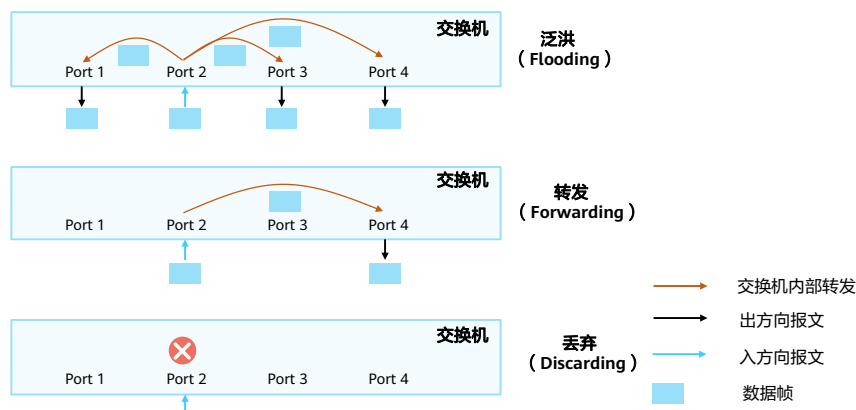
## 交换机的工作原理



- 二层交换机工作在数据链路层，它对数据帧的转发是建立在MAC地址基础之上的。交换机不同的接口发送和接收数据是独立的，各接口属于不同的冲突域，因此有效地隔离了网络中的冲突域。
- 二层交换设备通过学习以太网数据帧的源MAC地址来维护MAC地址与接口的对应关系（保存MAC与接口对应关系的表称为MAC地址表），通过其目的MAC地址来查找MAC地址表决定向哪个接口转发。

## 交换机的3种数据帧处理行为

- 交换机对于从传输介质进入某一端口的帧的处理行为一共有3种：



- 交换机会通过传输介质将进入其端口的每一个帧都进行转发操作，交换机的基本作用就是用来转发数据帧。
- 交换机对帧的处理行为一共有三种：泛洪（Flooding），转发（Forwarding），丢弃（Discarding）。
  - 泛洪：交换机把从某一端口进来的帧通过所有其它的端口转发出去（注意，“所有其它的端口”是指除了这个帧进入交换机的那个端口以外的所有端口）。
  - 转发：交换机把从某一端口进来的帧通过另一个端口转发出去（注意，“另一个端口”不能是这个帧进入交换机的那个端口）。
  - 丢弃：交换机把从某一端口进来的帧直接丢弃。

# 目录

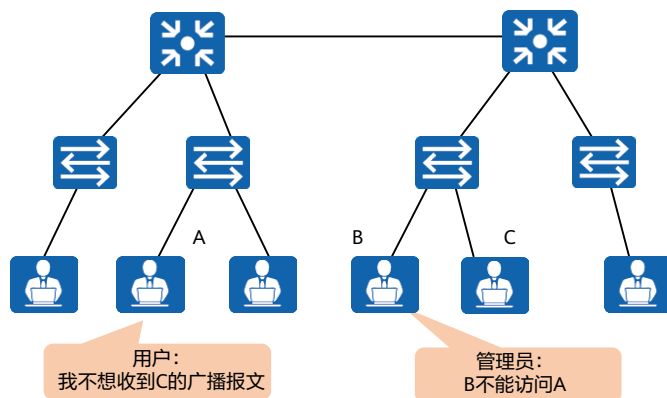
---

1. IP地址基础
2. 网络技术介绍
- 3. 交换技术基础**
  - 以太网交换基础
  - VLAN技术基础
  - VLAN基础配置
4. 路由技术基础



## VLAN的产生原因

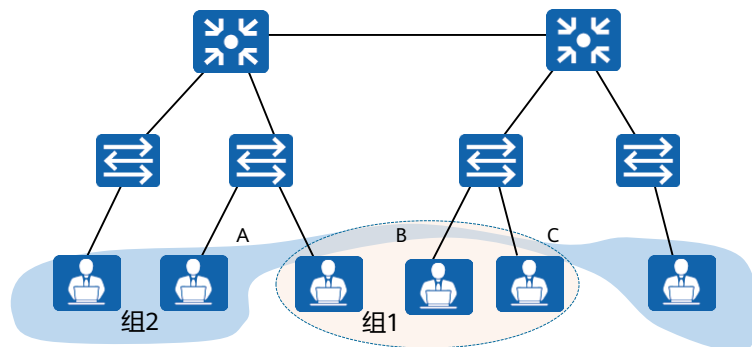
- 以太网缺少转发控制手段，广播影响范围广



- 传统的以太网交换机在转发数据时，采用源地址学习的方式，自动学习各个端口连接的主机的MAC地址，形成转发表，然后依据此表进行以太网帧的转发，整个转发的过程自动完成，所有端口都可以互访，维护人员无法控制端口之间的转发，例如实现B主机不能访问A主机。该网络存在如下缺陷：
  - 网络的安全性差。由于各个端口之间可以直接互访，增加了用户进行网络攻击的可能性。
  - 网络效率低。用户可能收到大量不需要的报文，这些报文同时消耗网络带宽资源和客户主机CPU资源，例如不必要的广播报文。
  - 业务扩展能力差。网络设备平等的对待每台主机的报文，无法实现有差别的服务，例如无法优先转发用于网络管理的以太网帧。

## VLAN技术的目标

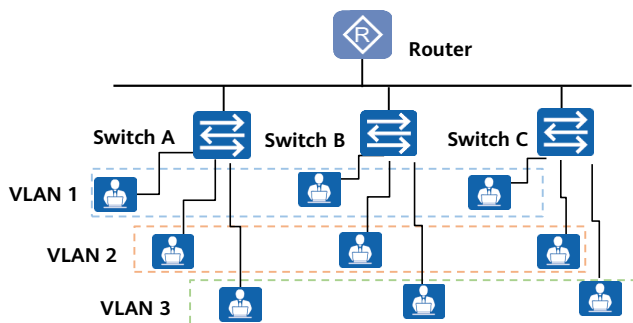
- VLAN技术把用户划分成多个逻辑的网络（组），组内可以通信，组间不允许通信，二层转发的单播、组播、广播报文只能在组内转发。同时，VLAN技术可以很容易地实现组成员的添加或删除



- VLAN技术提供了一种管理手段，控制终端之间的互通。如上图，组1和组2的PC无法相互通信。

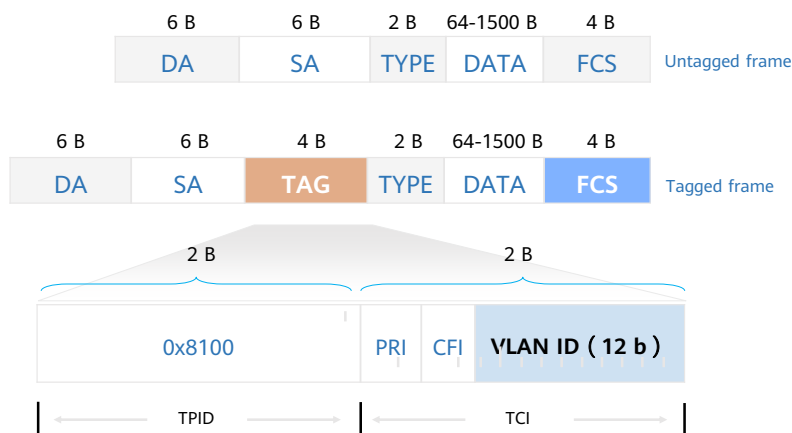
## 什么是VLAN

- VLAN (Virtual Local Area Network) 即虚拟局域网, 是将一个物理的LAN在逻辑上划分成多个广播域 (多个VLAN) 的通信技术



- VLAN内的主机间可以直接通信, 而VLAN间不能直接互通, 从而将广播报文限制在一个VLAN内。由于VLAN间不能直接互访, 从而提高了网络安全性。例如, 同一个写字楼的不同企业客户, 若建立各自独立的LAN, 企业的网络投资成本将很高; 若共用写字楼已有的LAN, 又会导致企业信息安全无法保证。采用VLAN, 可以实现各企业客户共享LAN设施, 同时保证各自的网络信息安全。
- 上图是一个典型的VLAN应用组网图。3台交换机放置在不同的地点, 比如写字楼的不同楼层。每台交换机分别连接3台计算机, 他们分别属于3个不同的VLAN, 比如不同的企业客户。

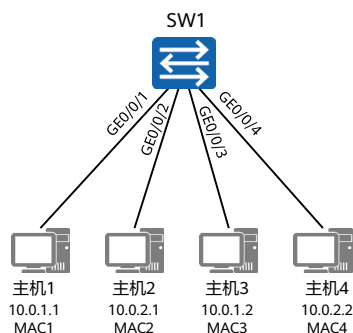
## VLAN帧格式



- VLAN标签长4个字节，直接添加在以太网帧头中，IEEE802.1Q文档对VLAN标签作出了说明。
- TPID: Tag Protocol Identifier, 2字节，固定取值，0x8100，是IEEE定义的新类型，表明这是一个携带802.1Q标签的帧。
- TCI: Tag Control Information, 2字节。帧的控制信息，详细说明如下：
  - Priority: 3比特，指示以太网帧的优先级。一共有8种优先级，0-7，用于提供有差别的转发服务。
  - CFI: Canonical Format Indicator, 1比特。用于令牌环/源路由FDDI介质访问中指示地址信息的比特次序信息，即先传送的是低特位还是高比特位。
  - VLAN Identifier: VLAN ID, 12比特，取值从0到4095。结合交换机端口的VLAN配置，能够控制以太网帧的转发。
- 使用VLAN标签后，在交换网络环境中，以太网的帧有两种格式：没有加上这四个字节标志的，称为标准以太网帧（untagged frame）；有四个字节的以太网帧，称为带有VLAN标记的帧（tagged frame）。
- 本课程仅仅讨论VLAN标签中的VLAN ID。

## VLAN的划分方式

- 整个网络是如何划分VLAN的

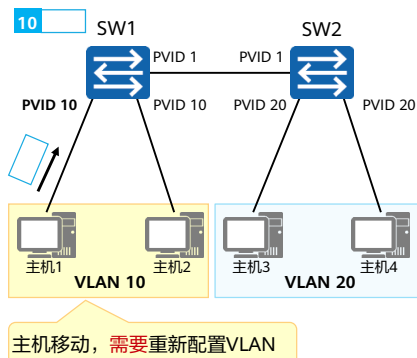


VLAN划分方式	VLAN 10	VLAN 20
基于接口	GE0/0/1, GE0/0/3	GE0/0/2, GE0/0/4
基于MAC地址	MAC 1, MAC 3	MAC 2, MAC 4
基于IP子网划分	10.0.1.*	10.0.2.*
基于协议划分	IP	IPv6
基于策略	10.0.1.* + GE0/0/1 + MAC 1	10.0.2.* + GE0/0/2 + MAC 2

- 计算机发出的数据帧不带任何标签。对已支持VLAN特性的交换机来说，当计算机发出的Untagged帧一旦进入交换机后，交换机必须通过某种划分原则把这个帧划分到某个特定的VLAN中去。
- VLAN的划分包括如下5种方法：
  - 基于接口划分：根据交换机的接口来划分VLAN。
    - 网络管理员预先给交换机的每个接口配置不同的PVID，当一个数据帧进入交换机时，如果没有带VLAN标签，该数据帧就会被打上接口指定PVID的标签，然后数据帧将在指定VLAN中传输。
  - 基于MAC地址划分：根据数据帧的源MAC地址来划分VLAN。
    - 网络管理员预先配置MAC地址和VLAN ID映射关系表，当交换机收到的是Untagged帧时，就依据该表给数据帧添加指定VLAN的标签，然后数据帧将在指定VLAN中传输。
  - 基于IP子网划分：根据数据帧中的源IP地址和子网掩码来划分VLAN。
    - 网络管理员预先配置IP地址和VLAN ID映射关系表，当交换机收到的是Untagged帧，就依据该表给数据帧添加指定VLAN的标签，然后数据帧将在指定VLAN中传输。

- 基于协议划分：根据数据帧所属的协议（族）类型及封装格式来划分VLAN。
  - 网络管理员预先配置以太网帧中的协议域和VLAN ID的映射关系表，如果收到的是Untagged帧，就依据该表给数据帧添加指定VLAN的标签，然后数据帧将在指定VLAN中传输。
- 基于策略划分：根据配置的策略划分VLAN，能实现多种组合的划分方式，包括接口、MAC地址、IP地址等。
  - 网络管理员预先配置策略，如果收到的是Untagged帧，且匹配配置的策略时，给数据帧添加指定VLAN的标签，然后数据帧将在指定VLAN中传输。

## 基于接口的VLAN划分



### 基于接口的VLAN划分

#### • 原理

- 根据交换机的接口来划分VLAN
- 网络管理员预先给交换机的每个接口配置不同的PVID，将该接口划入PVID对应的VLAN
- 当一个数据帧进入交换机时，如果没有带VLAN标签，该数据帧就会被打上接口指定PVID的Tag，然后数据帧将在指定PVID中传输

#### • 缺省VLAN，PVID

- Port VLAN ID，是接口上的缺省VLAN
- 取值：1~4094

### • 划分原则：

- 将VLAN ID配置到交换机的物理接口上，从某一个物理接口进入交换机的、由终端计算机发送的Untagged数据帧都被划分到该接口的VLAN ID所表明的那个VLAN。

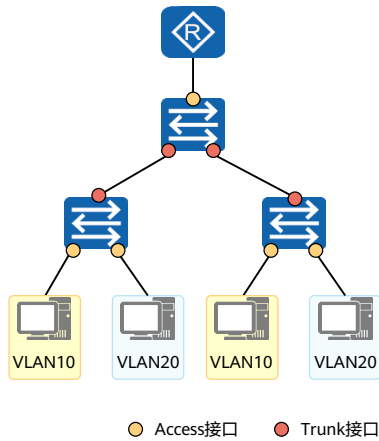
### • 特点：

- 这种划分原则简单而直观，实现容易，是目前实际网络应用中最为广泛的划分VLAN的方式。
- 当计算机接入交换机的端口发生变化时，该计算机发送的帧的VLAN归属可能会发生变化。

### • 缺省VLAN，PVID（Port VLAN ID）：

- 每个交换机的接口都应该配置一个PVID，到达这个端口的Untagged帧将一律被交换机划分到PVID所指代的VLAN。
- 默认情况下，PVID的值为1。

## VLAN的接口类型



### 接口类型

#### • Access接口

交换机上常用来连接用户PC、服务器等终端设备的接口。Access接口所连接的这些设备的网卡往往只收发无标记帧。Access接口只能加入一个VLAN

#### • Trunk接口

Trunk接口允许多个VLAN的数据帧通过，这些数据帧通过802.1Q Tag实现区分。Trunk接口常用于交换机之间的互联，也用于连接路由器、防火墙等设备的子接口

#### • Hybrid接口

Hybrid接口与Trunk接口类似，也允许多个VLAN的数据帧通过，这些数据帧通过802.1Q Tag实现区分。用户可以灵活指定Hybrid接口在发送某个（或某些）VLAN的数据帧时是否携带Tag

- 基于接口的VLAN划分依赖于交换机的接口类型。
- Access接口
  - Access接口一般用于和不能识别Tag的用户终端（如用户主机、服务器等）相连，或者不需要区分不同VLAN成员时使用。
- Trunk接口
  - Trunk接口一般用于连接交换机、路由器、AP以及可同时收发Tagged帧和Untagged帧的语音终端。
- Hybrid接口
  - Hybrid接口既可以用于连接不能识别Tag的用户终端（如用户主机、服务器等），也可以用于连接交换机、路由器以及可同时收发Tagged帧和Untagged帧的语音终端、AP。
  - 华为设备默认的接口类型是Hybrid。



# 目录

---

1. IP地址基础
2. 网络技术介绍
- 3. 交换技术基础**
  - 以太网交换基础
  - VLAN技术基础
    - VLAN基础配置
4. 路由技术基础

## VLAN的基础配置命令

- 创建VLAN

[Huawei] **vlan** *vlan-id*

- 通过此命令创建VLAN并进入VLAN视图，如果VLAN已存在，直接进入该VLAN的视图
- *vlan-id*是整数形式，取值范围是1 ~ 4094

[Huawei] **vlan batch** { *vlan-id1* [ **to** *vlan-id2* ] }

通过此命令批量创建VLAN。其中：

- *batch*：指定批量创建的VLAN ID
- *vlan-id1*：表示第一个VLAN的编号
- *vlan-id2*：表示最后一个VLAN的编号

- **vlan**命令用来创建VLAN并进入VLAN视图，如果VLAN已存在，直接进入该VLAN的视图。
- **undo vlan**用来删除指定VLAN。
- 缺省情况下，将所有接口都加入到一个缺省的VLAN中，该VLAN标识为1。
  - 命令：
  - **vlan** *vlan-id*
    - *vlan-id*：指定VLAN ID。整数形式，取值范围是1 ~ 4094。
  - **vlan batch** { *vlan-id1* [ **to** *vlan-id2* ] }
    - *batch*：指定批量创建VLAN。
    - *vlan-id1 to vlan-id2*：指定批量创建的VLAN ID，其中：
      - *vlan-id1*表示第一个VLAN的编号。
      - *vlan-id2*表示最后一个VLAN的编号。*vlan-id2*的取值必须大于等于*vlan-id1*，它与*vlan-id1*共同确定一个VLAN范围。
    - 如果不指定**to** *vlan-id2*参数，则只创建*vlan-id1*所指定的VLAN。
    - *vlan-id1*和*vlan-id2*是整数形式，取值范围是1 ~ 4094。

## Access接口的基础配置命令

- 配置接口类型

```
[Huawei-GigabitEthernet0/0/1] port link-type access
```

- 在接口视图下，配置接口的链路类型为Access

- 配置Access接口的缺省VLAN

```
[Huawei-GigabitEthernet0/0/1] port default vlan vlan-id
```

- 在接口视图下，配置接口的缺省VLAN并同时加入这个VLAN
- *vlan-id*：配置缺省VLAN的编号。整数形式，取值范围是1～4094

## Trunk接口的基础配置命令

- 配置接口类型

```
[Huawei-GigabitEthernet0/0/1] port link-type trunk
```

- 在接口视图下，配置接口的链路类型为Trunk

- 配置Trunk接口加入指定VLAN

```
[Huawei-GigabitEthernet0/0/1] port trunk allow-pass vlan { { vlan-id1 [ to vlan-id2 ] } | all }
```

- 在接口视图下，配置Trunk类型接口加入的VLAN

- (可选) 配置Trunk接口的缺省VLAN

```
[Huawei-GigabitEthernet0/0/1] port trunk pvid vlan vlan-id
```

- 在接口视图下，配置Trunk类型接口的缺省VLAN

- 命令：port trunk allow-pass vlan { { vlan-id1 [ to vlan-id2 ] | all }
  - vlan-id1 [ to vlan-id2 ]：指定Trunk类型接口加入的VLAN，其中：
    - vlan-id1表示第一个VLAN的编号。
    - to vlan-id2表示最后一个VLAN的编号。vlan-id2的取值必须大于等于vlan-id1的取值。
    - vlan-id1和vlan-id 2为整数形式，取值范围是1 ~ 4094。
  - all：指定Trunk接口加入所有VLAN。
- 命令：port trunk pvid vlan vlan-id，设置Trunk类型接口的缺省VLAN。
  - vlan-id：指定Trunk类型接口的缺省VLAN编号。整数形式，取值范围是1 ~ 4094。

## Hybrid接口的基础配置命令

- 配置接口类型

```
[Huawei-GigabitEthernet0/0/1] port link-type hybrid
```

- 在接口视图下，配置接口的链路类型为Hybrid

- 配置Hybrid接口加入指定VLAN

```
[Huawei-GigabitEthernet0/0/1] port hybrid untagged vlan { { vlan-id1 [ to vlan-id2 ] } | all }
```

- 在接口视图下，配置Hybrid类型接口加入的VLAN，这些VLAN的帧以Untagged方式通过接口

```
[Huawei-GigabitEthernet0/0/1] port hybrid tagged vlan { { vlan-id1 [ to vlan-id2 ] } | all }
```

- 在接口视图下，配置Hybrid类型接口加入的VLAN，这些VLAN的帧以Tagged方式通过接口

- (可选) 配置Hybrid接口的缺省VLAN

```
[Huawei-GigabitEthernet0/0/1] port hybrid pvid vlan vlan-id
```

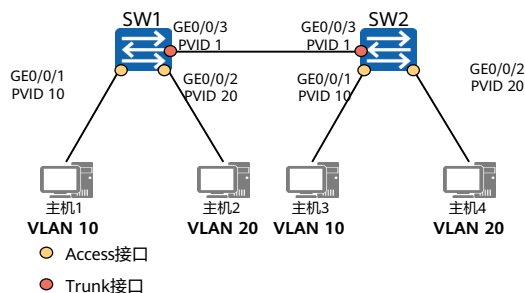
- 在接口视图下，配置Hybrid类型接口的缺省VLAN

- 命令：port hybrid untagged vlan { { vlan-id1 [ to vlan-id2 ] } | all }
  - vlan-id1 [ to vlan-id2 ]：指定Hybrid类型接口加入的VLAN，其中：
    - vlan-id1表示第一个VLAN的编号。
    - to vlan-id2表示最后一个VLAN的编号。vlan-id2的取值必须大于等于vlan-id1的取值。
    - vlan-id1和vlan-id 2为整数形式，取值范围是1 ~ 4094。
  - all：指定Hybrid接口加入所有VLAN。
- 命令：port hybrid tagged vlan { { vlan-id1 [ to vlan-id2 ] } | all }
  - vlan-id1 [ to vlan-id2 ]：指定Hybrid类型接口加入的VLAN，其中：
    - vlan-id1表示第一个VLAN的编号。
    - to vlan-id2表示最后一个VLAN的编号。vlan-id2的取值必须大于等于vlan-id1的取值。
    - vlan-id1和vlan-id 2为整数形式，取值范围是1 ~ 4094。
  - all：指定Hybrid接口加入所有VLAN。
- 命令：port hybrid pvid vlan vlan-id，设置Hybrid类型接口的缺省VLAN。
  - vlan-id：指定Hybrid类型接口的缺省VLAN编号。整数形式，取值范围是1 ~ 4094。

## 配置案例 - 基于接口划分VLAN

- 组网需求

- 某企业的交换机连接有很多用户，且相同业务用户通过不同的设备接入企业网络。为了通信的安全性，企业希望业务相同的用户之间可以互相访问，业务不同的用户不能直接访问
- 可以在交换机上配置基于接口划分VLAN，把业务相同的用户连接的接口划分到同一VLAN。这样属于不同VLAN的用户不能直接进行二层通信，同一VLAN内的用户可以直接互相通信



- 配置思路：

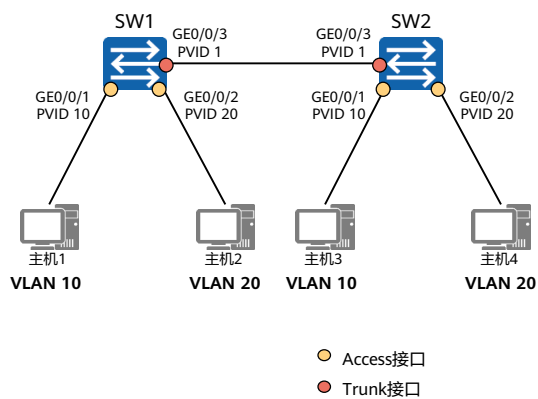
- 创建VLAN并将连接用户的接口加入VLAN，实现不同业务用户之间的二层流量隔离。
- 配置SW1和SW2的各接口类型以及通过的VLAN，实现相同业务用户通过SW1和SW2通信。

# 创建VLAN

创建VLAN：

```
[SW1] vlan 10  
[SW1-vlan10] quit  
[SW1] vlan 20  
[SW1-vlan20] quit
```

```
[SW2] vlan batch 10 20
```



## 配置Access接口和Trunk接口

配置Access接口，并加入对应的VLAN：

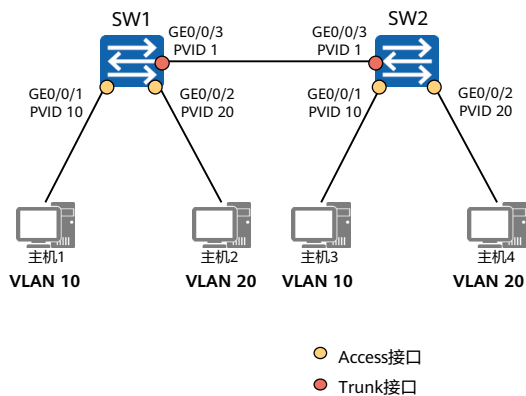
```
[SW1] interface GigabitEthernet 0/0/1
[SW1-GigabitEthernet0/0/1] port link-type access
[SW1-GigabitEthernet0/0/1] port default vlan 10
```

```
[SW1] interface GigabitEthernet 0/0/2
[SW1-GigabitEthernet0/0/2] port link-type access
[SW1] vlan 20
[SW1-vlan20] port GigabitEthernet0/0/2
[SW1-vlan20] quit
```

配置Trunk接口，并创建对应的允许通过列表：

```
[SW1] interface GigabitEthernet 0/0/3
[SW1-GigabitEthernet0/0/3] port link-type trunk
[SW1-GigabitEthernet0/0/3] port trunk pvid vlan 1
[SW1-GigabitEthernet0/0/3] port trunk allow-pass vlan 10 20
```

注：SW2配置与SW1类似





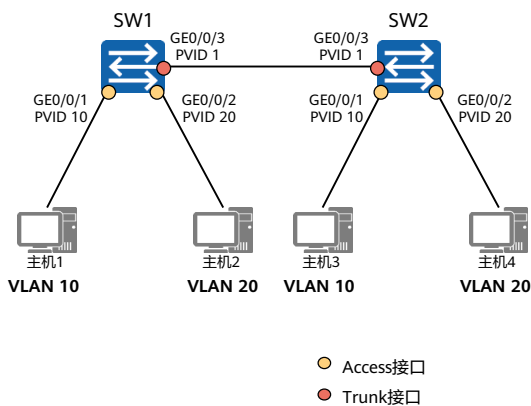
## 验证配置

[SW1]display vlan

The total number of vlans is : 3

U: Up; D: Down; TG: Tagged; UT: Untagged;  
MP: Vlan-mapping; ST: Vlan-stacking;  
#: ProtocolTransparent-vlan; \*: Management-vlan;

VID	Type	Ports
1	common	UT:GE0/0/3(U) .....
10	common	UT:GE0/0/1(U) TG:GE0/0/3(U)
20	common	UT:GE0/0/2(U) TG:GE0/0/3(U)



- 命令：display vlan命令用来查看VLAN的相关信息。
- 输出信息：
  - Tagged/Untagged Port：手动加入本VLAN的接口，分为Tagged和Untagged方式。
  - VID或VLAN ID：VLAN编号。
  - Type或VLAN Type：VLAN类型，common指普通VLAN。
  - Ports：加入该VLAN的接口。

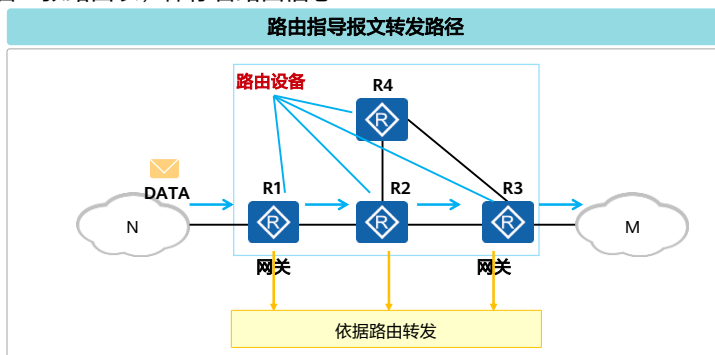
# 目录

---

1. IP地址基础
2. 网络技术介绍
3. 交换技术基础
- 4. 路由技术基础**
  - 路由基础原理
  - 静态和缺省路由

## 路由

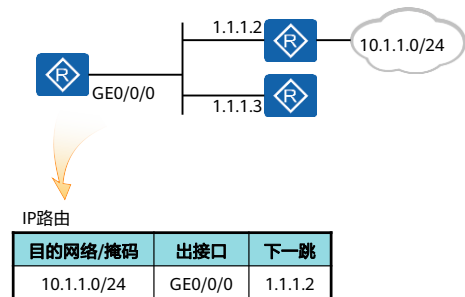
- 路由是指导报文转发的路径信息，通过路由可以确认转发IP报文的路径
- 路由设备是依据路由转发报文到目的网段的网络设备，最常见的路由设备：路由器
- 路由设备维护着一张路由表，保存着路由信息



- 网关以及中间节点（路由器）根据收到的IP报文其目的地址选择一条合适的路径，并将报文转发到下一个路由器。在路径中的最后一跳路由器二层寻址将报文转发给目的主机。这个过程被称为路由转发。
- 中间节点选择路径所依赖的表项称为路由表。
- 路由条目包含明确的出接口以及下一跳，这两项信息指导IP报文转发到相应的下一跳设备上。

## 路由信息介绍

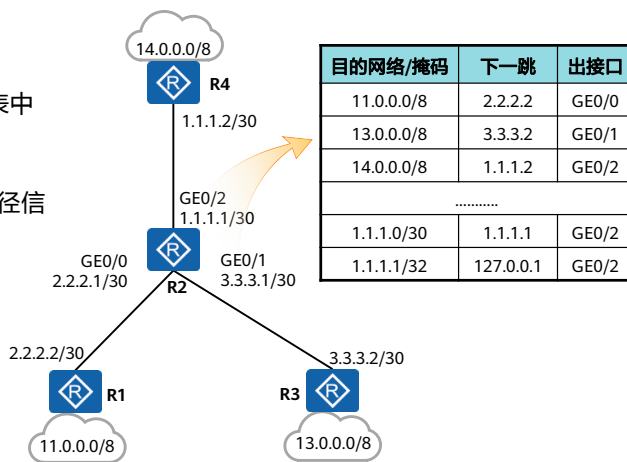
- 路由中包含以下信息：
  - 目的网络：标识目的网段
  - 掩码：与目的地址共同标识一个网段
  - 出接口：数据包被路由后离开本路由器的接口
  - 下一跳：路由器转发到达目的网段的数据包所使用的下一跳地址
- 这些信息标识了目的网段、明确了转发IP报文的路径



- 通过路由中包含的信息，路由设备可以转发IP报文到相应的路径。
- 目的地址、掩码用于识别IP报文目的地址，路由设备将IP报文匹配到相应的路由之后，根据路由的出接口、下一跳确认转发的路径。
- 只有出接口并不能够确认转发IP报文的下一跳设备，还需要明确的下一跳设备地址。

## 路由表

- 路由器通过各种方式发现路由
- 路由器选择最优的路由条目放入路由表中
- 路由表指导设备对IP报文的转发
- 路由器通过对路由表的管理实现对路径信息的管理



- 路由器依据路由表转发报文。
- 路由表由一条条详细的路由条目组成。
- 路由表由路由条目组成，但不代表路由表中保存了所有路由，路由表中只会保存“最优的”路由。
- 对路由表中的路由条目的管理实际上就是路由器维护、管理路由信息的具体实现。

## 查看IP路由表

```
<Huawei> display ip routing-table
Route Flags: R - relay, D - download to fib
```

---

Routing Tables: Public  
Destinations : 6      Routes : 6

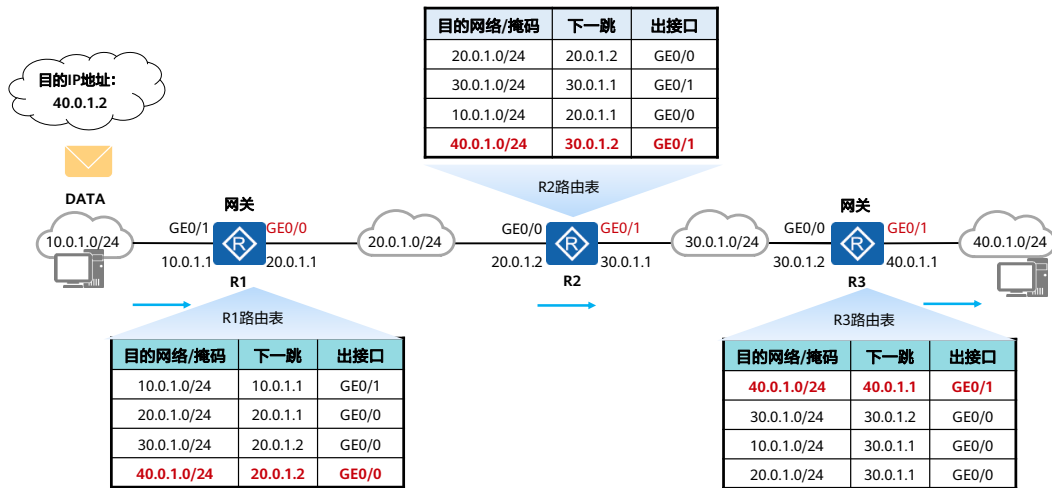
Destination/Mask	Proto	Pre	Cost	Flags	NextHop	Interface
1.1.1.1/32	Static	60	0	D	0.0.0.0	NULL0
2.2.2.2/32	Static	60	0	D	100.0.0.2	Vlanif100
100.0.0.0/24	Direct	0	0	D	100.0.0.1	Vlanif100
100.0.0.1/32	Direct	0	0	D	127.0.0.1	Vlanif100
127.0.0.0/8	Direct	0	0	D	127.0.0.1	InLoopBack0
127.0.0.1/32	Direct	0	0	D	127.0.0.1	InLoopBack0

↓                      ↓                      ↓                      ↓                      ↓                      ↓

目的网络地址/网络掩码长度   协议类型                      路由优先级   开销（度量值）   标志                      下一跳地址                      出接口

- Destination/Mask：表示此路由的目的网络地址与网络掩码。将目的地址和子网掩码“逻辑与”后可得到目的主机或路由器所在网段的地址。例如：目的地址为1.1.1.1，掩码为255.255.255.0的主机或路由器所在网段的地址为1.1.1.0。
- Proto（Protocol）：该路由的协议类型，即路由器是通过什么协议获知该路由的。
- Pre（Preference）：表示此路由的路由协议优先级。针对同一目的地，可能存在不同下一跳、出接口等多条路由，这些不同的路由可能是由不同的路由协议发现的，也可以是手工配置的静态路由。优先级最高（数值最小）者将成为当前的最优路由。
- Cost：路由开销。当到达同一目的地的多条路由具有相同的路由优先级时，路由开销最小的将成为当前的最优路由。
- NextHop：表示对于本路由器而言，到达该路由指向的目的网络的下一跳地址。该字段指明了数据转发的下一个设备。
- Interface：表示此路由的出接口。指明数据将从本路由器的哪个接口转发出去。

## 路由转发流程



- 来自10.0.1.0/24网段的IP报文想要去往40.0.1.0/24网段，首先到达网关，网关查找路由表项，确定转发的下一跳、出接口，之后报文转发给R2。报文到达R2之后，R2通过查找路由表项转发给R3，R3收到后查找路由表项，发现IP报文目的IP属于本地接口所在网段，直接本地转发。

# 目录

---

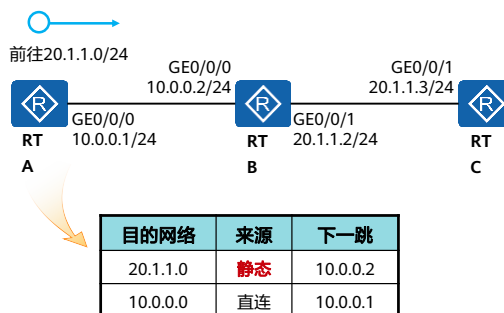
1. IP地址基础
2. 网络技术介绍
3. 交换技术基础
- 4. 路由技术基础**
  - 路由基础原理
  - 静态和缺省路由



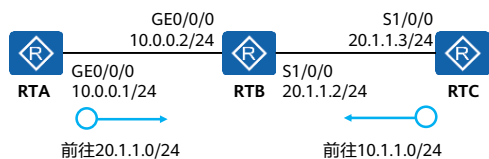
## 静态路由介绍

- 静态路由是由网络管理员手动配置路由条目，配置方便，对系统要求低，适用于拓扑结构简单并且稳定的小型网络
- 缺点是不能自动适应网络拓扑的变化，需要人工干预
- RTA上转发目的地址属于20.1.1.0/24的报文，在只有直连路由的情况下没有路由匹配。此时可以通过手动配置静态路由，使RTA发送前往20.1.1.0/24网段的报文交给下一跳10.0.0.2转发

### 静态路由



## 配置举例



RTA的配置如下:

```
[RTA] ip route-static 20.1.1.0 255.255.255.0 10.0.0.2
```

RTC的配置如下:

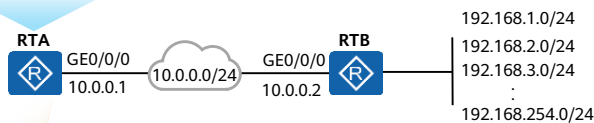
```
[RTC] ip route-static 10.0.0.0 255.255.255.0 S1/0/0
```

- RTA与RTC上配置静态路由，实现10.0.0.0/24与20.1.1.0/24的互通
- 因为报文是逐跳转发的，所以每一跳路由设备上都需要配置到达相应目的地址的路由
- 另外需要注意通信是双向的，针对通信过程中的往返流量，都需关注途径设备上的路由配置

## 缺省路由

- 缺省路由是一种特殊的路由，当报文没有在路由表中找到匹配的具体路由表项时才使用的路由；如果报文的目地址不能与路由表的任何目的地址相匹配，那么该报文将选取缺省路由进行转发
- 缺省路由在路由表中的形式为0.0.0.0/0，缺省路由也被叫做默认路由

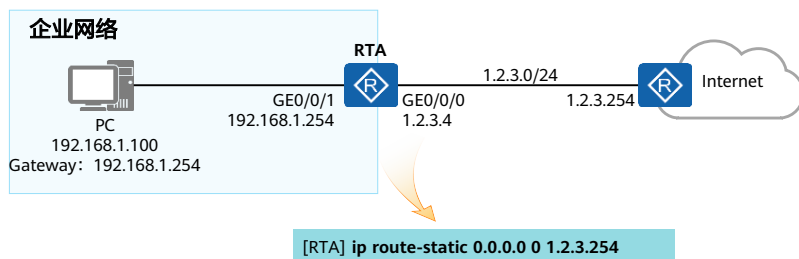
RTA前往非本地直连网段，  
将报文转发给10.0.0.2。



[RTA] ip route-static 0.0.0.0 0 10.0.0.2

## 缺省路由应用场景

- 缺省路由一般用于企业网络出口，配置一条缺省路由让出口设备能够转发前往Internet上任意地址的IP报文



## 本章总结

- 本章我们一起学习了IP地址的基本构成和子网划分，掌握了网络通信的基础原理，了解了网络常用协议的基本运作和使用场景。在接下来的课程中我们将继续学习基础技术中的操作系统基础部分，敬请期待。

## 思考题

1. （多选）以下关于防火墙设备功能的描述，正确的是哪些项？
  - A. 隔离不同安全级别的网络
  - B. 用户身份认证
  - C. 执行网络地址转换
  - D. 路由计算
2. （判断）缺省路由一般用于企业网络出口，配置一条缺省路由让出口设备能够转发前往Internet上任意地址的IP报文。
  - A. 正确
  - B. 错误

- 参考答案：

- ABCD
- A

## 学习推荐

- 华为Learning网站
  - <http://support.huawei.com/learning/Index!toTrainIndex>
- 华为Support案例库
  - <http://support.huawei.com/enterprise/servicecenter?lang=zh>

## 缩略语

- ARP: Address Resolution Protocol, 地址解析协议
- DNS: Domain Name Service, 域名解析服务
- FTP: File Transfer Protocol, 文件传输协议
- HTTP: Hypertext Transfer Protocol, 超文本传输协议
- ICMP: Internet Control Message Protocol, 网际报文控制协议
- IGMP: Internet Group Management Protocol, 因特网组管理协议
- IP: Internet Protocol, 互联网协议
- LAN: Local Area Network, 局域网
- TCP: Transmission Control Protocol, 传输控制协议
- UDP: User Datagram Protocol, 用户数据报协议
- VLAN: Virtual Local Area Network, 虚拟局域网
- VLSM: Variable-Length Subnet Mask, 可变长度子网掩码



# Thank you.

把数字世界带入每个人、每个家庭、  
每个组织，构建万物互联的智能世界。  
Bring digital to every person, home, and  
organization for a fully connected,  
intelligent world.

Copyright©2023 Huawei Technologies Co., Ltd.  
All Rights Reserved.

The information in this document may contain predictive statements including, without limitation, statements regarding the future financial and operating results, future product portfolio, new technology, etc. There are a number of factors that could cause actual results and developments to differ materially from those expressed or implied in the predictive statements. Therefore, such information is provided for reference purpose only and constitutes neither an offer nor an acceptance. Huawei may change the information at any time without notice.



## 操作系统基础



# 前言

- 操作系统在用户与计算机之间的连接中扮演着重要的角色，那么操作系统到底是什么？有哪些类型的操作系统？Linux操作系统的基本操作命令有哪些？本章将一一为您揭开。

# 目标

- 学完本课程后，您将能够：
  - 了解操作系统的定义及组成
  - 熟悉操作系统的分类
  - 掌握Linux操作系统的基本操作

# 目录

---

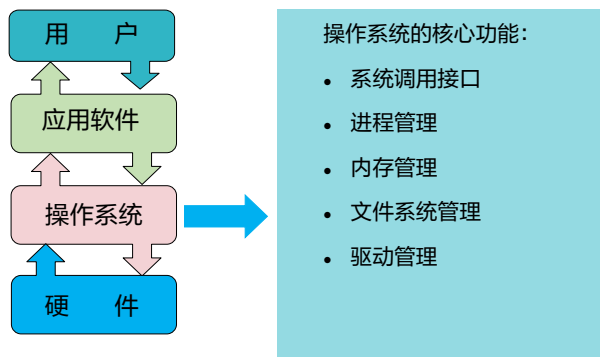
## 1. 操作系统基础

- 操作系统的定义
- 操作系统的组成
- 操作系统的分类

## 2. Linux基础

## 操作系统的定义及作用

- 操作系统（Operating System，简称OS），是管理和控制计算机硬件与软件资源的计算机程序（系统软件）



- 操作系统是一种控制计算机的特殊计算机程序，在计算机和用户之间起到连接的作用，并对硬件资源进行协调管理。这些资源包括CPU、磁盘、内存和打印机等，用户运行程序时需要访问这些资源。
- 主流操作系统
  - 根据应用领域来划分，可分为：
    - 桌面操作系统、服务器操作系统、主机操作系统、嵌入式操作系统。
  - 根据源码开放程度，可分为：
    - 开源操作系统（Linux、Unix）、闭源操作系统（Windows、Mac OS）。

# 目录

---

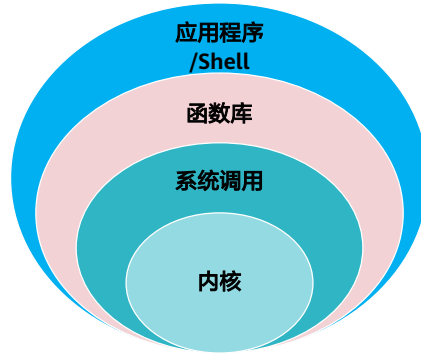
## 1. 操作系统基础

- 操作系统的定义
- 操作系统的组成
- 操作系统的分类

## 2. Linux基础

## 操作系统的组成

- 通常站在用户角度，操作系统由内核和各种应用程序组成。即系统分为用户空间和内核空间
- 用户空间即上层应用程序活动的空间
- 内核本质上是一种软件程序，用来管理计算机硬件资源并提供上层应用程序运行所需要的系统调用接口



- 系统调用：应用程序的执行必须依托于内核提供的资源，包括CPU、存储、I/O资源等，为了使上层应用能够访问到这些资源，内核必须为上层应用提供访问接口—系统调用。（如果写一个汉字代表一个上层应用，那么一个系统调用更像是汉字的一个笔画）。
- 库函数：通过库函数实现对系统调用的封装，将简单的业务逻辑接口呈现给用户，方便用户调用。对于简单的操作我们可以直接调用系统调用来访问资源，对于复杂操作，可以借助库函数来实现。（库函数更像是汉字的偏旁）。
- Shell：是一个特殊的应用程序（俗称命令行），本质上是一个命令解释器，它可以执行符合Shell语法的文本（脚本），实际上这些Shell语句通常就是对系统调用做了一层封装，方便客户使用。



- 内核：向下控制硬件资源，向内管理操作系统资源，向上则向应用程序提供系统调用的接口。
  - 进程调度和管理：内核负责创建和销毁进程，处理他们与外部世界的联系。
  - 内存管理：内核为所有的进程在有限的可用资源上建立一个虚拟地址空间。
  - 文件系统管理：Linux在很大程度上基于文件系统的概念，几乎Linux中的任何东西都可看作一个文件。内核在非结构化的硬件之上建立了一个结构化的文件系统。
  - 设备驱动管理：内核中需要嵌入系统中出现的每个外设的驱动, 从硬盘驱动到键盘和磁带驱动器。
  - 网络资源管理：所有的路由和地址解析动作都在内核中实现。
- 总结：用户态应用程序可以通过三种方式访问内核态资源：
  - 系统调用
  - Shell脚本
  - 库函数

# 目录

---

## 1. 操作系统基础

- 操作系统的定义
- 操作系统的组成
- 操作系统的分类

## 2. Linux基础

## 常用服务器操作系统

### UNIX

一种多用户、多进程的计算机操作系统。支持大型文件系统服务、数据服务应用，功能强大、稳定性和安全性能好

#### 常见UNIX OS:

HP-UX、IBM-AIX、Solaris、Apple UNIX

### GNU/Linux

Linux是类Unix计算机操作系统的统称。Linux系统比较成熟，拥有一套完整的权限机制，安全性与稳定性都很高

#### 常见Linux OS:

SUSE、Kylin、RedFlag、CentOS、Redhat、openEuler

### Windows

Windows Server是微软发布的服务器操作系统，主要应用于服务器场景，拥有友好的用户视窗操作界面

#### 常见Windows Server版本:

2000、2003、2008、2012、2016、2019

- Linux与Unix的关系:

- Linux是一个类似UNIX的操作系统，并在功能和用户体验上进行优化，所以Linux模仿了UNIX，使得Linux在外观和交互上与UNIX非常类似。
- Linux内核最初是由李纳斯·托瓦兹（Linus Torvalds）在赫尔辛基大学读书时出于个人爱好而编写的，当时他觉得教学用的迷你版UNIX操作系统Minix太难用了，于是决定自己开发一个操作系统。第1个版本于1991年9月发布，当时仅有10000行代码。
- UNIX系统大多是与硬件配套的，也就是说，大多数UNIX系统如AIX、HP-UX等是无法安装在x86服务器和个人计算机上的，而Linux则可以运行在多种硬件平台上。
- UNIX是商业软件，而Linux是开源软件，是免费、公开源代码的。

- GNU计划:

- GNU计划，又称革奴计划，是由Richard Stallman在1983年9月27日公开发起的。它的目标是创建一套完全自由的操作系统。
- GNU是“GNU's Not UNIX”的递归缩写。Linux提供了一个操作系统内核，而GNU提供了大量的自由软件来丰富在其之上的各种应用程序。

# 目录

---

## 1. 操作系统基础

## 2. Linux基础

- Linux系统介绍
  - openEuler简介
  - openEuler文件系统简介
  - openEuler基础操作

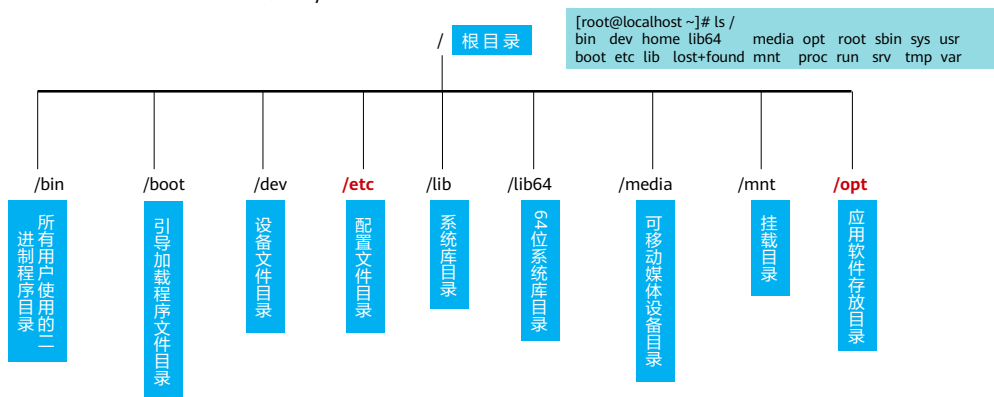
## Linux操作系统特性

- 支持多种平台
  - Linux可运行在多种硬件平台上。此外，Linux还是一种嵌入式操作系统，可运行在掌上电脑、机顶盒等设备
- 多用户多任务
  - 系统资源可以被不同用户各自拥有使用，同时执行多个程序，而且各个程序的运行相互独立
- 自由软件
  - 用户可以随时免费获取它的源代码，并且可以根据自己的需求对源代码进行编辑、修改
- 完全兼容POSIX1.0标准
- 继承了Unix的设计理念
  - 一切皆文件

- POSIX，可移植操作系统接口（Portable Operating System Interface，缩写为POSIX），是IEEE为要在各种UNIX操作系统上运行的软件，而定义API的一系列互相关联的标准的总称，其正式称呼为IEEE 1003，而国际标准名称为ISO/IEC 9945。它基本上是Portable Operating System Interface（可移植操作系统接口）的缩写，而X则表明其对Unix API的传承。
- Linux操作系统是一种非常流行的多任务、多用户操作系统，应用非常广泛。Linux的主要特点为：
  - 多任务（Multi-tasking）：Linux是一个多任务操作系统，在它内部允许有多个任务同时运行。而DOS操作系统是单任务的操作系统，不能同时运行多个任务。系统在执行多任务时，CPU在某一时刻只能执行一个任务，但Linux操作系统将CPU时间分片，并把这些时间片分别安排给多个进程。因为CPU运行很快，在操作者看来，所有程序（进程）都在同时运行。
  - 多用户（Multi-users）：Linux又是一个多用户操作系统，它允许多个用户同时使用。在Linux中，每位用户运行自己的或公用的程序，好象拥有一台单独的机器。DOS操作系统是单用户的操作系统，只允许一个用户同时使用。
  - 管道：Linux允许一个程序的输出作为另外一个程序输入，多个程序串起来好像一条管道一样。通过各个简单任务的组合，就可以完成比较复杂的任务，并极大提高了操作的方便性。晚期版本的DOS操作系统也借鉴并提供了这种机制。
  - 功能强大的Shell：Linux的命令解释器由Shell实现。Linux提供了多种功能强大的Shell，每种Shell本身就是一种解释型高级语言，通过用户编程就可创造无数命令，使用方便。
  - 安全保护机制：Linux提供了非常强大的安全保护机制，防止系统及其数据未经许可而被非法访问。

# Linux文件目录结构

- 在Linux操作系统中，一切皆是文件
- Linux的文件目录结构是树形，“/”称为根目录



- Linux下的哲学核心思想是“一切皆文件”。“一切皆文件”，指的是对所有文件（目录、字符设备、块设备、套接字、打印机、进程、线程、管道等）操作，读写都可用`fopen()/fclose()/fwrite()/fread()`等函数进行处理。
- 登录系统后，在当前命令窗口下输入命令：`ls /` 你会看到上图所示的回显，那么这些目录是什么意思呢？以下是这些目录的解释：
  - `/bin`：bin是Binary的缩写, 这个目录存放着最经常使用的命令。
  - `/boot`：这里存放的是启动Linux时使用的一些核心文件，包括一些连接文件以及镜像文件。
  - `/dev`：dev是Device（设备）的缩写, 该目录下存放的是Linux的外部设备，在Linux中访问设备的方式和访问文件的方式是相同的。
  - `/etc`：这个目录用来存放所有的系统管理所需要的配置文件和子目录。
  - `/lib`：这个目录里存放着系统最基本的动态连接共享库，其作用类似于Windows里的DLL文件。几乎所有的应用程序都需要用到这些共享库。
  - `/media`：Linux系统会自动识别一些设备，例如U盘、光驱等等，当识别后，Linux会把识别的设备挂载到这个目录下。
  - `/mnt`：系统提供该目录是为了让用户临时挂载别的文件系统的，我们可以将光驱挂载在`/mnt/`上，然后进入该目录就可以查看光驱里的内容了。
  - `/opt`：这是给主机额外安装软件所摆放的目录。比如你安装一个ORACLE数据库则就可以放到这个目录下。默认是空的。

# 目录

---

## 1. 操作系统基础

## 2. Linux基础

- Linux系统介绍
- openEuler简介
- openEuler文件系统简介
- openEuler基础操作

## openEuler出现的背景

- EulerOS是一款基于Linux内核的服务器操作系统，支持x86和ARM等多种处理器架构，适用于数据库、大数据、云计算、人工智能等应用场景
- 在近10年的发展中，EulerOS成功支持了华为各种产品解决方案，以安全、稳定、高效被业界认可
- 随着云计算的兴起和鲲鹏芯片的发展，EulerOS成为与鲲鹏芯片配套最合适的软件基础设施
- 为推动鲲鹏生态的发展，繁荣国内和全球的计算产业，2019年底EulerOS被正式推送开源社区，命名为openEuler





## openEuler操作系统介绍

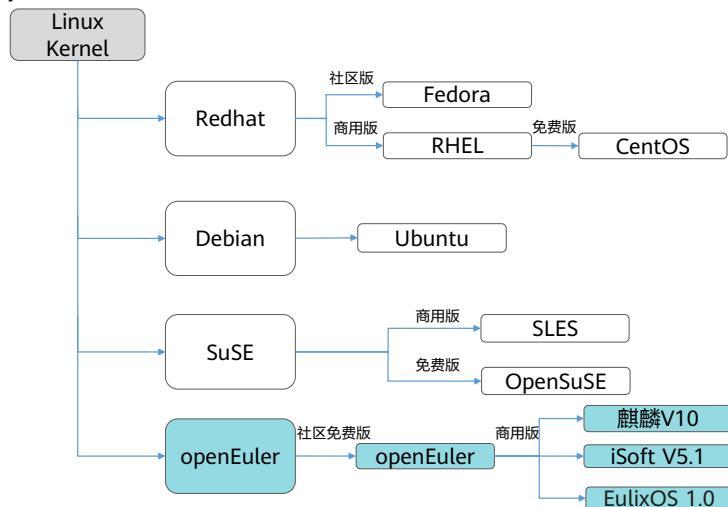
- openEuler是一个开源、免费的Linux发行平台
- 支持x86、ARM、RISC-V等多种处理器架构
- 所有开发者、企业、商业组织都可以使用openEuler社区版本，也可以基于社区版本发布自己二次开发的操作系统版本

<https://openeuler.org/>

<https://gitee.com/openeuler/>



## openEuler和主流OS系的关系



- openEuler与SuSE、Debian、RedHat的上游社区都是kernel社区  
[www.kernel.org](http://www.kernel.org)
- openEuler社区发行LTS免费版本，使能OSV发展商业发行版，如麒麟软件、普华、中科软、万里开源等

# 目录

---

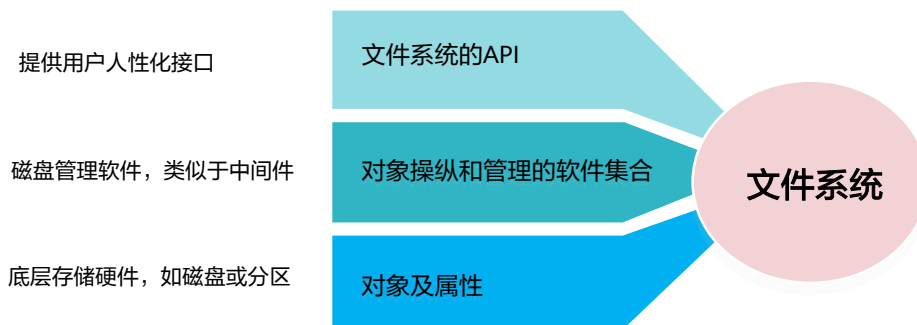
## 1. 操作系统基础

## 2. Linux基础

- Linux系统介绍
- openEuler简介
  - openEuler文件系统简介
- openEuler基础操作

## 文件系统概述

- 文件系统是操作系统用于明确存储设备或分区上的文件的方法和数据结构，即在存储设备上组织文件的方法
- 操作系统中负责管理和存储文件信息的软件机构称为文件管理系统，简称文件系统



- 作用：文件系统是对文件存储设备的空间进行组织和分配，负责文件存储并对存入的文件进行保护和检索的系统。具体地说，它负责为用户建立文件，存入、读出、修改、转储文件，控制文件的存取，当用户不再使用时撤销文件等。文件系统的功能包括：管理和调度文件的存储空间，提供文件的逻辑结构、物理结构和存储方法；实现文件从标识到实际地址的映射，实现文件的控制操作和存取操作，实现文件信息的共享并提供可靠的文件保密和保护措施，提供文件的安全措施。

## openEuler文件系统

- openEuler内核源于Linux，Linux内核支持十多种不同类型的文件系统，比如Btrfs、JFS、ReiserFS、Ext、Ext2、Ext3、Ext4、ISO9660、XFS、Minix、MSDOS、UMSDOS、VFAT、NTFS、HPFS、SMB、SysV、PROC等，表格中列举常用几种并做说明
- openEuler默认文件系统为Ext4

常用文件系统	说明
Ext	专门为Linux设计的文件系统类型，目前最新版本Ext4
XFS	一种高性能的日志文件系统，最早于1993年，由Silicon Graphics为他们的IRIX操作系统而开发，之后被移植到Linux内核上，特别擅长处理大文件，同时提供平滑的数据传输
VFAT	VFAT是Linux对DOS，Windows系统下的FAT（包括FAT16和FAT32）文件系统的统称
ISO 9600	该文件系统中光盘所使用的标准文件系统，Linux对该文件系统也有很好的支持，不仅能读取光盘和光盘ISO映像文件，而且还支持在Linux环境中刻录光盘

# 目录

---

## 1. 操作系统基础

## 2. Linux基础

- Linux系统介绍
- openEuler简介
- openEuler文件系统简介
- openEuler基础操作

# 目录

---

## openEuler基础操作

- Linux命令行基础知识
- openEuler基础命令
- openEuler的文本处理
- openEuler的网络管理

## Linux的GUI与CLI

- GUI: Graphical User Interface, 图形用户界面。用户界面的所有元素图形化, 主要使用鼠标作为输入工具, 使用按钮、菜单、对话框等进行交互, 追求易用
- CLI: Command Line Interface, 命令行界面。用户界面字符化, 使用键盘作为输入工具, 输入命令、选项、参数执行程序, 追求高效

- 示例:
  - 在Windows系统中通过GUI启动“计算器”: 开始 → 所有程序 → 附件 → 计算器, 在这个计算器中, 用鼠标点击上面的按键输入表达式, 或者某些程序要求你输入密码的时候, 弹出一个小键盘, 要求你点击上面的数字.....不得不说, 这种设置十分的人性化, 计算器看起来跟真的很像, 到银行取钱, 也有用于输入密码的小键盘.....不过问题的关键在于, 这种图形界面模拟的键盘是用手来按的, 而不是用鼠标点的。
  - 在CLI下使用计算器, 只要输入bc就可以启动计算器, 输入一个复杂的公式1+1, 回车, 得到结果2。



## 为什么要使用Linux命令行

- 命令行更高效
  - Linux系统中使用键盘操作速度要比鼠标更快
  - 图形化界面不可重复，而命令行可以编写脚本完成所有过程（例如：删除过期日志文件）
- 图形化界面开销大
  - 运行图形化界面会占用很多的系统资源，运行命令行可以将系统资源释放给更应该做的事情
- 命令行有时候是唯一的选择
  - 大部分服务器操作系统不会安装图形界面（GUI）
  - 联网设备的维护管理工具本来就没有图形化界面供我们使用

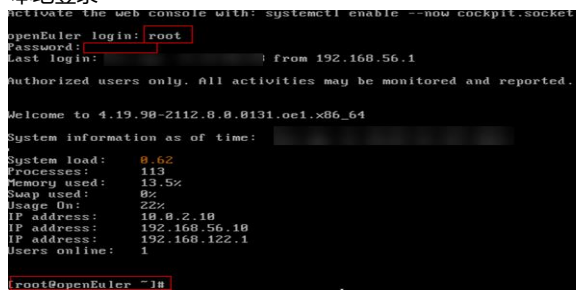
# Linux命令行操作技巧

- Tab键补全
  - 可以用来自动补全命令或文件名，省时准确
  - 未输入命令状态下，连按两次“Tab”键列出所有可用命令
  - 已输入部分命令名或文件名，按“Tab”键自动补全
- 光标
  - “up方向键上”：可以调出输入历史执行记录，快速执行命令
  - “down方向键下”：配合“up”选择历史执行记录
  - “Home”：移动光标到本行开头
  - “Ctrl + A”：移动光标到行首
  - “Ctrl + E”：移动光标到行尾
  - “Ctrl + L”：清理屏幕显示

## 如何登录Linux

- Linux登录方式主要有如下两种：

- 本地登录



```
activate the web console with: systemctl enable --now cockpit.socket
openEuler login: root
Password:
Last login:  from 192.168.56.1
Authorized users only. All activities may be monitored and reported.

Welcome to 4.19.98-2112.8.0.0131.oe1.x86_64
System information as of time:
System load: 0.62
Processes: 113
Memory used: 13.5%
Swap used: 0%
Usage On: 22%
IP address: 192.168.56.10
IP address: 192.168.122.1
Users online: 1
root@openEuler ~#
```

- 远程登录

- 可以通过PuTTY、Xshell等工具远程登录到openEuler

- 以root管理员登录系统后，提示符为“#”。

## 修改密码

- 密码关系到系统和数据的安全
- 为保障系统安全，您需要做：
  - 初次登录系统时修改密码
  - 定期修改密码
  - 设置高复杂度的密码，如八位以上字符，且包含大写字母、小写字母、数字和特殊字符中三种及以上的字符

- |   |   |
|---|---|
| <pre>[root@openEuler ~]# passwd Changing password for user root. New password: Retype new password: passwd: all authentication tokens updated successfully</pre>  | <pre>#修改当前用户自己的密码 #此处输入新密码 #重复输入新密码</pre> |
| <pre>[root@openEuler ~]# passwd test1 Changing password for user test1. New password: BAD PASSWORD: The password is a palindrome Retype new password: passwd: all authentication tokens updated successfully.</pre> | <pre>#管理员root修改普通用户的密码</pre>              |

- 为了安全起见，openEuler系统默认在输入密码时不会显示密码，而且也不会用任何字符进行位数反馈。

# Linux用户分类

- Linux系统中，使用UID来表示不同的用户身份，且每个用户的UID是唯一的
- Linux系统根据用户UID的不同，将用户分为以下不同的类型（以openEuler为例）：
  - 超级用户
    - 也叫超级管理员，其UID=0，拥有系统所有权限，类似Windows的Administrator
  - 系统用户
    - 也叫程序用户，其  $1 \leq \text{UID} \leq 999$ ，一般都是由程序创建，用于程序或者服务运行时候的身份
  - 普通用户
    - 一般是由管理员root创建，用于对系统进行有限的管理维护操作， $1000 \leq \text{UID} \leq 60000$

## Linux用户创建及删除

- 用户创建（默认普通用户）： useradd username
- 用户查看： id username
- 用户切换： su - username
- 用户删除： userdel username

```
[root@openEuler ~]# useradd user01                #创建用户user01
[root@openEuler ~]# id user01                      #管理员查看user01信息
uid=1001(user01) gid=1001(user01) groups=1001(user01)
[root@openEuler ~]# su - user01                    #切换到user01用户，提示符为$符号
[user01@openEuler ~]$ id                           #id命令默认查看当前用户信息
uid=1001(user01) gid=1001(user01) groups=1001(user01)
[user01@openEuler ~]$ exit                          #退出当前用户
logout
[root@openEuler ~]# userdel user01                  #删除user01用户
```

# 目录

---

## openEuler基础操作

- Linux命令行基础知识
- openEuler基础命令
- openEuler的文本处理
- openEuler的网络管理

## 电源命令 - 关机与重启

- shutdown命令的作用是关闭计算机，使用权限是超级用户
  - 主要参数：
    - -h：关机后关闭电源
    - -r：关机后打开电源（相当于重启）
    - -p：明确表示将关闭系统并切断主电源
- reboot命令的作用是重新启动计算机，使用权限是系统管理者
  - 主要参数：
    - -w：并不会真的重开机，只是把记录写到/var/log/wtmp文件里
    - -d：不把记录写到/var/log/wtmp文件里
    - -i：关闭网络设置之后再重新启动系统

- shutdown命令可以安全地将系统关机，使用直接断掉电源的方式来关闭Linux系统十分危险。
- Linux与Windows不同，其后台运行着许多进程，所以强制关机可能会导致进程的数据丢失，使系统处于不稳定的状态，甚至在有的系统中会损坏硬件设备。
- 在系统关机前使用shutdown命令，系统管理员会通知所有登录的用户系统将要关闭，并且login指令会被冻结，即新的用户不能再登录。



## 文件路径表示

- 绝对路径：指从根 “/” 路径开始来表示文件，例如：/root/Desktop
- 相对路径：指从当前路径开始来表示文件，例如：./Desktop
  - 其中 “./或.” 表示当前路径； “../或..” 表示当前路径的上一级目录
- 查看当前路径命令：pwd
- 路径切换命令：cd

```
[root@localhost ~]# pwd                #查看当前路径
/root
[root@localhost ~]# cd /root/Desktop    #进入Desktop目录（绝对路径表示）
[root@localhost Desktop]# cd           #cd回车默认进入当前用户的家目录
[root@localhost ~]#
[root@localhost ~]# cd ./Desktop        #进入Desktop目录（相对路径表示）
[root@localhost Desktop]# pwd
/root/Desktop
[root@localhost Desktop]# cd ~          #切换到用户家目录。~表示用户的家目录
[root@localhost ~]# pwd
/root
```

- cd命令用于改变当前工作目录。
- 语法：cd [dir]
  - cd /usr：进入目录/usr中。
  - cd ..：进入（退到）上一层目录，两个点代表父目录。
  - cd .：进入当前目录
  - cd：不带参数，则默认回到家目录。
  - cd -：进入前一个目录，适用于在两个目录之间快速切换。
  - cd ~：进入家目录

## 文件的查看

- 查看目录里的内容：ls

```
[root@localhost ~]# ls -a                #显示所有文件及目录
ifcfg-lo  ifdown-eth  ifdown-isdn  ifdown-routes
[root@localhost ~]# ls -l                #显示文件型态、权限、拥有者、文件大小等详细信息
total 228
-rw-r----- 1 root root  86 Jun 15 19:03 ifcfg-eth0
-rw-r----- 1 root root 254 Jun 15 19:03 ifcfg-lo
```

- 普通文件的查看：cat、tail、head

```
[root@localhost ~]# cat ifcfg-eth0        #读取该文件全部内容
DEVICE="eth0"
BOOTPROTO="dhcp"
ONBOOT="yes"
TYPE="Ethernet"
PERSISTENT_DHCLIENT="yes"
[root@localhost ~]# tail -2 ifcfg-eth0     #查看该文件最后2行内容
TYPE="Ethernet"
PERSISTENT_DHCLIENT="yes"
[root@localhost ~]# head -2 ifcfg-eth0     #查看文件前2行内容
DEVICE="eth0"
BOOTPROTO="dhcp"
```

- ls命令：用于查看目录里的文件：
  - a：查看所有文件，包含隐藏文件；
  - l：以长格式显示文件的详细信息；
  - R：递归列出目录文件中的文件；
  - t：按文件的修改时间排序列出。
- cat命令：一般用于查看小文件，将文件从第一行到最后一行显示到屏幕。
- tail命令：默认查看文件的后10行内容：
  - n：后面接数字（如5），表示查看文件的后5行，可以省略-n参数直接接数字；
  - f：动态显示文件的变化，一般用户查看日志文件。
- head命令：默认查看文件的前10行内容。
- less和more：一般用于查看大文件。“q”表示退出，可以在文件中进行搜索查找，按斜杠“/”+关键字。

## 文件的创建

- 目录（文件夹）创建：mkdir

- 选项-p：递归级联创建目录

```
[root@localhost ~]# mkdir my_dir_01          #创建文件夹my_dir_01
[root@localhost ~]# ls
anaconda-ks.cfg  my_dir_01
[root@localhost ~]# mkdir -p my_dir_02/sub_dir  #创建文件夹my_dir_02及子目录sub_dir
```

- 普通文件的创建：touch

```
[root@localhost ~]# touch test01.log test02.log  #创建文件test01.log test02.log
[root@localhost ~]# ls -lt
total 0
-rw-----, 1 root root 0 Jul 29 15:06 test01.log
-rw-----, 1 root root 0 Jul 29 15:06 test02.log
```

## 文件的复制

- cp命令：拷贝复制文件或目录

- 选项-a：此选项通常在复制目录时使用，它保留链接、文件属性并复制目录下的所有文件
- 选项-r：若给出的源文件是目录文件，此时将复制该目录下所有子目录和文件，并保留属性信息

```
[root@localhost ~]# ls
test01.log test02.log
[root@localhost ~]# cp /etc/passwd passwd.back      #拷贝/etc/passwd文件内容至当前目录并重命名为passwd.back
[root@localhost ~]# cp -r /var/log/audit ./          #拷贝audit目录及目录下所有文件至当前目录
[root@localhost ~]# ls
audit passwd.back test01.log test02.log
[root@localhost ~]# cp -s /etc/passwd passwd_link    #将passwd_link作为passwd文件的链接文件
[root@localhost ~]# ls
audit passwd.back passwd_link test01.log test02.log
[root@localhost ~]# ls -l
total 8
drwx----- 2 root root 4096 Jul 29 15:24 audit
-rw----- 1 root root 2546 Jul 29 15:24 passwd.back
lrwxrwxrwx 1 root root 11 Jul 29 15:25 passwd_link -> /etc/passwd
-rw----- 1 root root 0 Jan 2 19:20 test01.log
-rw----- 1 root root 0 Jul 29 19:20 test02.log
[root@localhost ~]#
```

- cp命令用于复制文件或者目录，可以一次复制单个文件，也可以一次复制多个文件（cp命令属于高危命令，使用不慎就会有丢失数据的危险）。
- 语法：cp [OPTION]... SOURCE... DIRECTORY:
  - -a：此选项通常在复制目录时使用，它保留链接、文件属性，并复制目录下的所有内容；
  - -p：除复制文件的内容外，还把修改时间和访问权限也复制到新文件中；
  - -r：若给出的源文件是一个目录文件，此时将复制该目录下所有的子目录和文件；
  - -l：不复制文件，只是生成链接文件；
  - -s：把目标文件建立为源文件的软链接文件，而不是复制源文件。
- cp f1 f2：把文件f1复制一份，新文件名为f2。
- cp f1 d1/：复制f1到目录d1下，新文件名字不变。
- cp f1 f2 f3 d1/：复制多个文件到同一个目录中。
- cp -i f1 f2：如果f2已经存在，则覆盖之前等用户确认。
- cp -r d1 d2：复制目录时需要-r参数。
- cp -rv d1 d2：-v参数显示复制的过程。
- cp -s d1 d2：把目标文件建立为源文件的软链接文件，而不是复制源文件。
- cp -a f1 f2：-a参数保留原文件的属性，可用于复制块设备，字符设备，管道文件等。
- cp命令默认情况下在覆盖文件之前不会询问用户，所以通常情况下shell里面已经对它做了别名，加入了-i参数。cp命令的-f参数不是强制覆盖的意思。

## 文件的移动和重命名

- mv命令：移动或重命名文件
  - mv命令用于移动文件或者目录（高危命令，使用不慎就会有丢失数据的危险）
  - 若原文件和目标文件在同一个目录里面，则mv命令的效果就相当于给文件重命名

```
[root@localhost ~]# ls
passwd_link test01.log test02.log
[root@localhost ~]# mv test02.log test03.log    #将文件test02.log改名为test03.log
[root@localhost ~]# ls
passwd_link test01.log test03.log
[root@localhost ~]# mv test01.log /root/test    #将文件test01.log移动到/root/test目录
[root@localhost ~]# mv -f test01.log test03.log #以test01.log内容强制覆盖test03.log
```

- mv命令用于移动文件或者目录（高危命令，使用不慎就会有丢失数据的危险）。
- 如果原文件和目标文件在同一个父目录里面，则mv命令的效果就相当于给文件改名。
- 语法：mv [选项] 源文件或目录 目标文件或目录。
  - -b：若需覆盖文件，则覆盖前先行备份。
  - -f：force强制的意思，如果目标文件已经存在，不会询问而直接覆盖。
  - -i：若目标文件（destination）已经存在时，就会询问是否覆盖。
  - -u：若目标文件已经存在，且source比较新，才会更新（update）。

## 文件的删除

- rm命令：删除文件或目录

- rm命令属于高危命令，没有一个工具能够100%恢复rm命令删除的文件，rm命令删除文件时并不是把文件放到类似图形界面的“回收站”里，所以没有“撤销删除”操作可用

```
[root@localhost ~]# ls
audit_back  passwd.back test01.log  test03.log
[root@localhost ~]# rm test01.log                #删除test01.log文件，删除前询问
rm: remove regular empty file 'test01.log'? yes
[root@localhost ~]# rm -rf test03.log            #强制删除test03.log文件
[root@localhost ~]# rm -rf audit_back/          #删除mail.bak目录及其目录下所有文件和目录
[root@localhost ~]# ls
passwd.back
[root@localhost ~]#
```

- 语法：rm [OPTION] file\_or\_dir

- -f, --force：忽略不存在的文件，从不给出提示。
- -i, --interactive：进行交互式删除。
- -r, -R, --recursive：指示rm将参数中列出的全部目录和子目录均递归地删除。
- -v, --verbose：详细显示进行的步骤。

## 命令如何获取帮助

- help命令：获取命令的简单帮助
  - 在Linux系统中，命令太多，记全几乎不太可能。但是，我们可以通过help命令获取帮助
  - 语法：command --help或help command

```
[root@localhost ~]# help pwd
pwd: pwd [-LP]
Print the name of the current working directory.
Options:
  -L      print the value of $PWD if it names the current working directory
  -P      print the physical directory, without any symbolic links

By default, 'pwd' behaves as if '-L' were specified.

Exit Status:
Returns 0 unless an invalid option is given or the current directory cannot be read.
[root@localhost ~]# systemctl --help
systemctl [OPTIONS...] {COMMAND} ...

Query or send control commands to the systemd manager.

-h --help          Show this help
--version          Show package version
--system           Connect to system manager
-H --host=[USER@]HOST.....
```

- 在Linux系统中，命令太多，记全几乎不太可能，但是，我们可以通过help命令获取帮助。
- 命令语法：
  - help [选项] [命令]
- 其中，选项内容包括如下：
  - -d：显示命令简短主题描述
  - -s：显示命令简短语法描述

# 目录

---

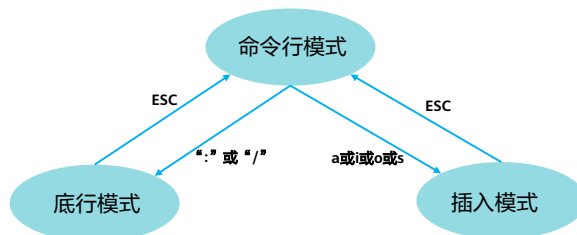
## openEuler基础操作

- Linux命令行基础知识
- openEuler基础命令
  - openEuler的文本处理
- openEuler的网络管理



## Linux文本编辑器 - vim

- vim是从Vi（VisualEditor）发展出来的一个高度可定制的文本编辑器，在Vi原有的基础上改进和增加了很多特性
- vim常用的操作：
  - 命令行模式（普通模式）：文本复制粘贴、删除、撤销、光标来回跳转等操作
  - 插入模式（编辑模式）：文本的编辑、修改等操作
  - 底行模式（普通模式的一种）：在普通模式下按“:”冒号，可以进行文本的保存、退出、搜索替换等操作



- openEuler 20.03 LTS系统安装后默认没有安装vim，需要手动安装vim。

## vim编辑器命令模式

- 命令格式：默认情况下使用vim打开文件就会进入到命令模式（普通模式）下

<code>vim [options] [file]...</code>	编辑指定文件
<code>vim [options] -</code>	从标准输入（stdin）读取文本
<code>vim [options] -t tag</code>	编辑tag定义出的文件
<code>vim [options] -q [errorfile]</code>	编辑第一个出错处的文件

- 常见参数：
  - `-c`：打开文件前先执行指定的命令
  - `-R`：以只读方式打开，但是可以强制保存
  - `-M`：以只读方式打开，不可以强制保存
  - `-r`：恢复崩溃的会话
  - `+num`：从第`num`行开始

## vim命令模式常见操作

- 快速移动光标
  - 上下左右键或k、j、h、l键上下左右移动光标
  - 0: 移动到行首
  - g0: 移到光标所在屏幕行行首
  - :n: 移动到第n行
  - gg: 到文件头部
  - G: 到文件尾部
- 数据操作
  - yy or Y: 复制整行文本
  - y[n]w: 复制一 (n) 个词
  - d[n]w: 删除 (剪切) 1 (n) 个单词
  - [n] dd: 删除 (剪切) 1 (n) 行

## vim编辑器插入模式

- 命令行使用vim filename打开文件，默认进入命令模式，输入i/I/a/A/o/O等命令就会进入到编辑模式
- 如果filename文件存在，则会打开文件并显示文件内容
- 如果filename文件不存在，vim会在下面提示 “[New File]”，并且会在第一次保存时创建该文件
- 按Ecs可以退出插入模式，回到命令模式（普通模式）

```
[root@openEuler ~]# vim test.txt           #默认进入命令模式
~
~
"test.txt" [New File]
[root@openEuler ~]# vim test.txt           #按i/I/a/A/o/O等命令会进入到编辑模式
~
~
-- INSERT --
```

## vim编辑器底行模式

- 查找
  - `:/word` 或者 `/word`：查找一个字符串word，按“n”向后继续搜索，“shift+n”向上搜索
- 替换
  - `:1,5s/word1/word2/g`：将文档中1-5行的word1替换为word2，不加“g”则只替换每行的第一个word1
  - `%s/word1/word2/gi`：将文档所有的word1替换为word2，不区分大小写
- 保存退出
  - `:w`：保存文档不退出
  - `:wq`：保存文档并退出
  - `:q`：退出不保存文档
  - `:q!`：文档修改后退出不保存强制退出
  - `:wq!`：强制保存退出

# 目录

---

## openEuler基础操作

- Linux命令行基础知识
- openEuler基础命令
- openEuler的文本处理
- openEuler的网络管理

## openEuler系统网络重要概念

- **主机网络设备**：主机内的网卡
- **接口**
  - 为使用设备，驱动程序在设备上创建了接口
- **广播地址**
  - 到达本网段上所有主机的地址
- **子网掩码**
  - 将IP地址划分成网络地址和主机地址两部分的掩码
- **路由**
  - IP包跨网段访问时候的下一跳地址
- **链路**：设备到网络的连接

## 查看IP地址相关命令

- 使用ip、ifconfig命令可以查看当前主机的IP地址信息
- 查看主机所有网卡信息

```
[root@openEuler ~]# ifconfig -a  
[root@openEuler ~]# ip addr show
```

- 查看主机某个接口信息

```
[root@openEuler ~]# ifconfig enp0s3  
[root@openEuler ~]# ip addr show enp0s3
```

- 查看所有接口的当前IP地址和子网掩码：ip addr。



## 配置静态IP地址 - 修改网卡配置文件

- 网卡配置文件存放路径

```
[root@openEuler ~]# ls /etc/sysconfig/network-scripts/ifcfg-*  
/etc/sysconfig/network-scripts/ifcfg-enp0s3 /etc/sysconfig/network-scripts/ifcfg-enp0s8
```

- 配置文件参数说明

参数	说明
TYPE	配置文件接口类型
BOOTPROTO	系统启动地址协议
ONBOOT	系统启动时是否激活
IPADDR	IP地址
NETMASK	子网掩码
GATEWAY	网关地址
BROADCAST	广播地址
HWADDR/MACADDR	MAC地址，只需设置其中一个，同时设置时不能相互冲突
PEERDNS	是否指定DNS，如果使用DHCP协议，默认为yes
DNS{1, 2}	DNS地址
USERCTL	用户权限控制
NAME	网络连接名称
DEVICE	物理接口名称

## 配置IP地址 - 配置文件示例

- 配置enp0s3接口的静态IP地址：192.168.56.100/24

```
TYPE=Ethernet
BOOTPROTO=static
NAME=enp0s3
DEVICE=enp0s3
ONBOOT=yes
IPADDR=192.168.56.100
NETMASK=255.255.255.0
```

- 重启网络

```
[root@openEuler ~]# nmcli connection reload enp0s3
[root@openEuler ~]# nmcli connection up enp0s3
Connection successfully activated (D-Bus active path: /org/freedesktop/NetworkManager/ActiveConnection/12)
```

- 有时过多配置项反而会增加网络排障难度。

## 配置静态IP地址 - nmcli命令配置

- 查看当前主机网络连接

```
[root@openEuler ~]# nmcli connection show
NAME      UUID                                  TYPE      DEVICE
enp0s3    3c36b8c2-334b-57c7-91b6-4401f3489c69  ethernet  enp0s3
enp0s8    00cb8299-feb9-55b6-a378-3fdc720e0bc6  ethernet  enp0s8
```

- 配置静态IP地址

```
[root@openEuler ~]# nmcli connection modify enp0s3 ipv4.method manual ipv4.addresses "10.0.2.10/24" ipv4.gateway "10.0.2.2"
```

- 重启网络

```
[root@openEuler network-scripts]# nmcli connection reload enp0s3
[root@openEuler network-scripts]# nmcli connection up enp0s3
Connection successfully activated (D-Bus active path: /org/freedesktop/NetworkManager/ActiveConnection/18)
```

- 查看IP

```
[root@openEuler ~]# ip addr show enp0s3
2: enp0s3: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc fq_codel state UP group default qlen 1000
    link/ether 08:00:27:7d:e1:a5 brd ff:ff:ff:ff:ff:ff
    inet 10.0.2.10/24 brd 10.0.2.255 scope global noprefixroute enp0s3
        valid_lft forever preferred_lft forever
```

- conn: connection, 表示要对连接进行操作
- add: 添加（连接）
- type: 类型
- con-name: 连接名称
- ifname: 网卡名称
- mod: modify, 表示要修改（连接）

## 路由介绍

- 要让不同子网的两台主机能够相互通讯，就需要有一种能够描述如何从一台主机到另一台主机的机制，这一机制称为：路由选择（Routing），路由选择通过路由项进行描述
- 路由项是一对预先定义的地址，包括：目的地（Destination）和网关（Gateway）
- 路由项的含义是通过网关能够完成与目的地的通讯；路由表是多个路由项的集合

- 关于系统中路由的详细描述：

[https://www.freebsd.org/doc/zh\\_CN/books/handbook/network-routing.html](https://www.freebsd.org/doc/zh_CN/books/handbook/network-routing.html)

## 路由管理与配置

- openEuler中，使用route命令查看、配置、管理本机路由
- 除了route命令，ip命令也可以用来管理系统路由
- 这些命令作用于系统中的路由表，系统运行时，路由表加载到内存中，由系统内核进行维护

- route、ip、nmcli都可以用来管理路由，这里以route为例介绍。

## route查看路由表

- 使用route命令查看路由表，如下：

```
[root@openEuler ~]# route -n
Kernel IP routing table
Destination Gateway Genmask Flags Metric Ref Use Iface
0.0.0.0 192.168.110.254 0.0.0.0 UG 100 0 0 enp4s0
192.168.110.0 0.0.0.0 255.255.255.0 U 100 0 0 enp4s0
192.168.122.0 0.0.0.0 255.255.255.0 U 0 0 0 virbr0
```

- 使用route -n参数显示路由时，Destination字段显示的是IP。
- 采用route命令查看路由时，共有八个字段，字段意义说明如下：
  - U Up，表示此路由当前为启动状态；
  - H Host，表示此网关为一主机；
  - G Gateway，表示此网关为一路由器；
  - R Reinstall Route，使用动态路由重新初始化的路由；
  - D Dynamically，此路由是动态性地写入；
  - M Modified，此路由是由路由守护程序或导向器动态修改；
  - !，表示此路由当前为关闭状态。

## route新增路由

- 新增一条到网段或主机的路由，如下（临时生效）：

```
route [-f] [-p] [Command [Destination] [mask Netmask] [Gateway] [metric Metric]] [if Interface]]
```

- 举例：

```
[root@openEuler ~]# route add -net 192.168.101.0 netmask 255.255.255.0 dev enp4s0
[root@openEuler ~]# route add -host 192.168.100.10 dev enp4s0
[root@openEuler ~]# route
```

Kernel IP routing table

Destination	Gateway	Genmask	Flags	Metric	Ref	Use	Iface
default	_gateway	0.0.0.0	UG	100	0	0	enp4s0
192.168.100.10	0.0.0.0	255.255.255.255	UH	0	0	0	enp4s0
192.168.101.0	0.0.0.0	255.255.255.0	U	0	0	0	enp4s0
192.168.110.0	0.0.0.0	255.255.255.0	U	100	0	0	enp4s0
192.168.122.0	0.0.0.0	255.255.255.0	U	0	0	0	virbr0

- 使用route命令可以新增路由数据，新增路由数据保存在内存中，重启后失效。
- route add -net 192.168.101.0 netmask 255.255.255.0 dev enp3s0，代表新增到192.168.101.0/24网段路由，经由enp3s0发送出去。
- route add -host 192.168.101.100 dev enp3s0，代表新增到192.168.101.100主机路由，经由enp3s0发送出去。
- 通过route查看路由表，可知，到主机的路由优先于到网段的路由。

## route删除路由

- 使用route del命令删除到网段或者主机的路由
- 语法：

```
route del [-net|-host] [netmask Nm] [gw Gw] [[dev] If]
```

- 举例：

```
[root@openEuler ~]# route del -host 192.168.100.10 dev enp4s0
[root@openEuler ~]# route
Kernel IP routing table
Destination Gateway Genmask Flags Metric Ref Use Iface
default _gateway 0.0.0.0 UG 100 0 0 enp4s0
192.168.101.0 0.0.0.0 255.255.255.0 U 0 0 0 enp4s0
192.168.110.0 0.0.0.0 255.255.255.0 U 100 0 0 enp4s0
192.168.122.0 0.0.0.0 255.255.255.0 U 0 0 0 virbr0
```

- route del -net 192.168.101.0 netmask 255.255.255.0 dev enp3s0，代表删除到192.168.101.0/24网段路由。删除网段路由时，网段和掩码这两个参数一定需要携带，dev参数可以不携带。
- route del -host 192.168.101.100 dev enp3s0，代表删除到192.168.101.100主机路由。后面的dev参数可以不带。
- 删除route文件中的路由，采用vi编辑，然后重启网络即可。



## 主机名

- 主机名是用来在一个局域网中标识一台机器的名称
- 这台机器不一定是一台物理机器，也可以是一台虚拟机
- 主机名信息存放在文件/etc/hostname中

```
[root@openEuler ~]# cat /etc/hostname  
openEuler
```

## 设置主机名

- 临时设置主机名：hostname new-name
- 永久设置主机名：hostnamectl set-hostname new-name
- 通过修改文件设置主机名：在/etc/hostname文件中写入new-name

```
[root@openEuler ~]# hostname  
openEuler  
[root@openEuler ~]# hostname huawei  
[root@openEuler ~]# hostname  
huawei  
[root@openEuler ~]# hostnamectl set-hostname openEuler01  
[root@openEuler ~]# hostname  
openEuler01  
[root@openEuler ~]# echo "HCIA-openEuler" > /etc/hostname
```

- 设置完成后并不会立即生效，重新登录后才能在命令提示符中看到，或者source .bashrc。
- 使用hostname命令查看当前系统的主机名。

## hosts文件介绍

- 局域网中存在着一些主机，我们可以通过IP地址去访问它们
- 随着局域网中主机数量增加，IP地址较难记忆，我们希望通过主机名直接访问其他主机
- 这时我们可以通过一张记录主机名和IP地址对应的表找到这些主机，这张表就是hosts文件

```
[root@openEuler ~]# cat /etc/hosts
127.0.0.1    localhost    localhost.localdomain    localhost4    localhost4.localdomain4
::1         localhost    localhost.localdomain    localhost6    localhost4.localdomain6
```

- hosts是一个没有扩展名的系统文件，其基本作用就是将一些常用的网址域名与其对应的IP地址建立一个关联“数据库”。
- 当用户在浏览器中输入一个需要登录的网址时，系统会首先自动从hosts文件中寻找对应的IP地址，一旦找到，系统就会立即打开对应网页。
- 如果没有找到，则系统会将网址提交DNS域名解析服务器进行IP地址的解析。
- 查看hosts文件：cat /etc/hosts。

## 修改hosts文件

- 可以用编辑器修改hosts文件内容，格式：

```
# Ip domain.com  
192.168.10.20 www.example.com
```

- 若想去掉某条记录，加#注释即可，如：

```
#ip domain.com  
#192.168.10.20 www.example.com
```

## 思考题

1. （单选）以下关于文件系统的描述，错误的是哪一项？
  - A. 文件系统是操作系统用于明确存储设备或分区上的文件的方法和数据结构
  - B. 操作系统中负责管理和存储文件信息的软件结构称为文件管理系统
  - C. 负责管理和控制计算机硬件与软件资源
  - D. 对文件存储设备的空间进行组织和分配，负责文件存储并对存入的文件进行保护和检索
2. （判断）Linux是一个多用户操作系统，可以同时接受多个用户登录，还允许一个用户进行多次登录。
  - A. 正确
  - B. 错误

- 参考答案：

- C
- A

## 本章总结

- 本章节我们一起学习了操作系统的基本组成、操作系统的类型以及Linux操作系统的基本操作，至此，计算、存储、网络和操作系统基础技术部分的学习就告一段落了。那么在云计算中我们应该如何利用、管理这些资源来为应用提供服务呢？在接下来虚拟化技术的课程中我们将一起学习。

## 学习推荐

- 华为Learning网站
  - <http://support.huawei.com/learning/Index!toTrainIndex>
- 华为Support案例库
  - <http://support.huawei.com/enterprise/servicecenter?lang=zh>

## 缩略语

- CLI: Command Line Interface, 命令行界面
- GUI: Graphical User Interface, 图形用户界面
- POSIX: Portable Operating System Interface, 可移植操作系统接口



# Thank you.

把数字世界带入每个人、每个家庭、  
每个组织，构建万物互联的智能世界。  
Bring digital to every person, home, and  
organization for a fully connected,  
intelligent world.

Copyright©2023 Huawei Technologies Co., Ltd.  
All Rights Reserved.

The information in this document may contain predictive statements including, without limitation, statements regarding the future financial and operating results, future product portfolio, new technology, etc. There are a number of factors that could cause actual results and developments to differ materially from those expressed or implied in the predictive statements. Therefore, such information is provided for reference purpose only and constitutes neither an offer nor an acceptance. Huawei may change the information at any time without notice.

