云计算基础与应用

南通师范高等专科学校
Nantong Normal College

朱亚林

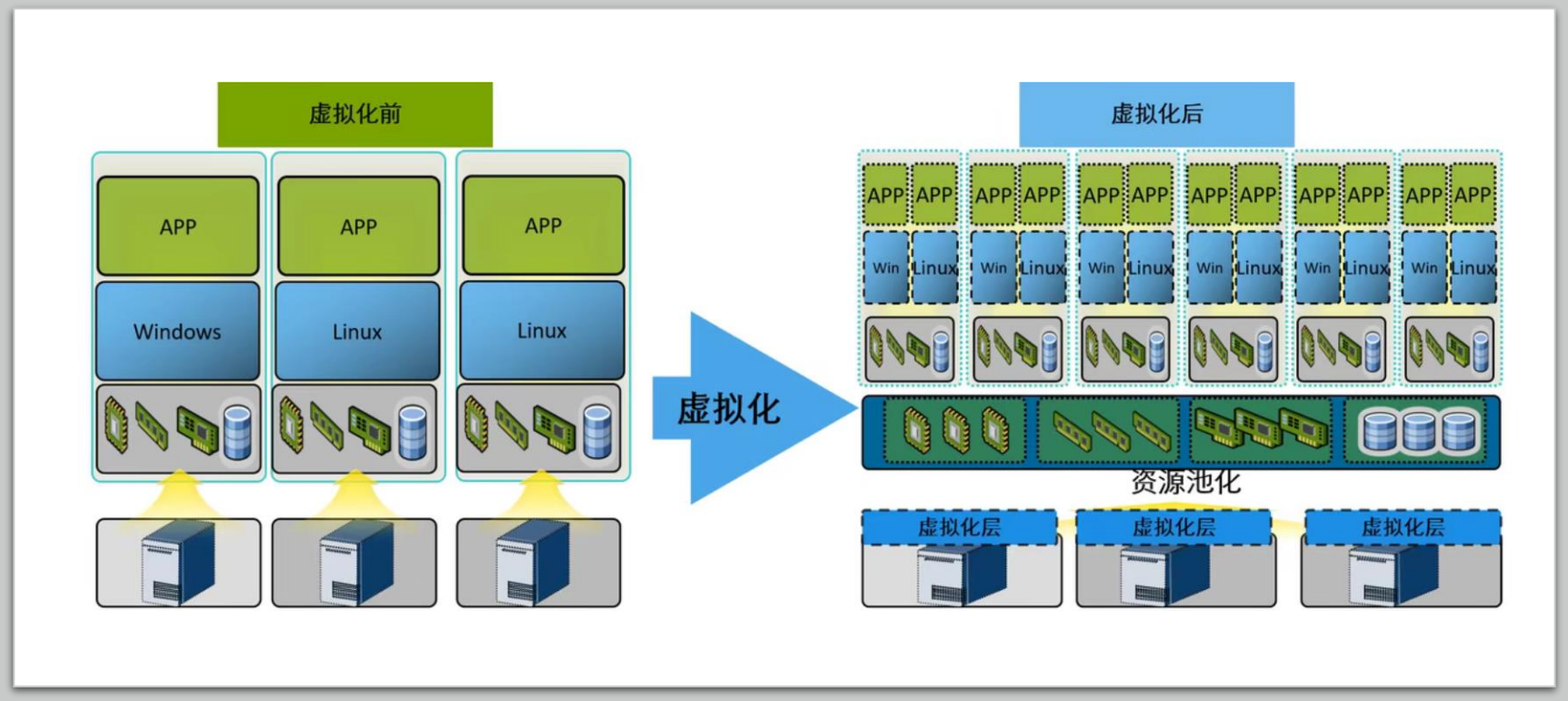
4.1 虚拟化技术

思考两个问题

云计算给你留下的最深刻的印象是什么？

实现云计算最重要的技术是什么？

4.3.1 虚拟化概述

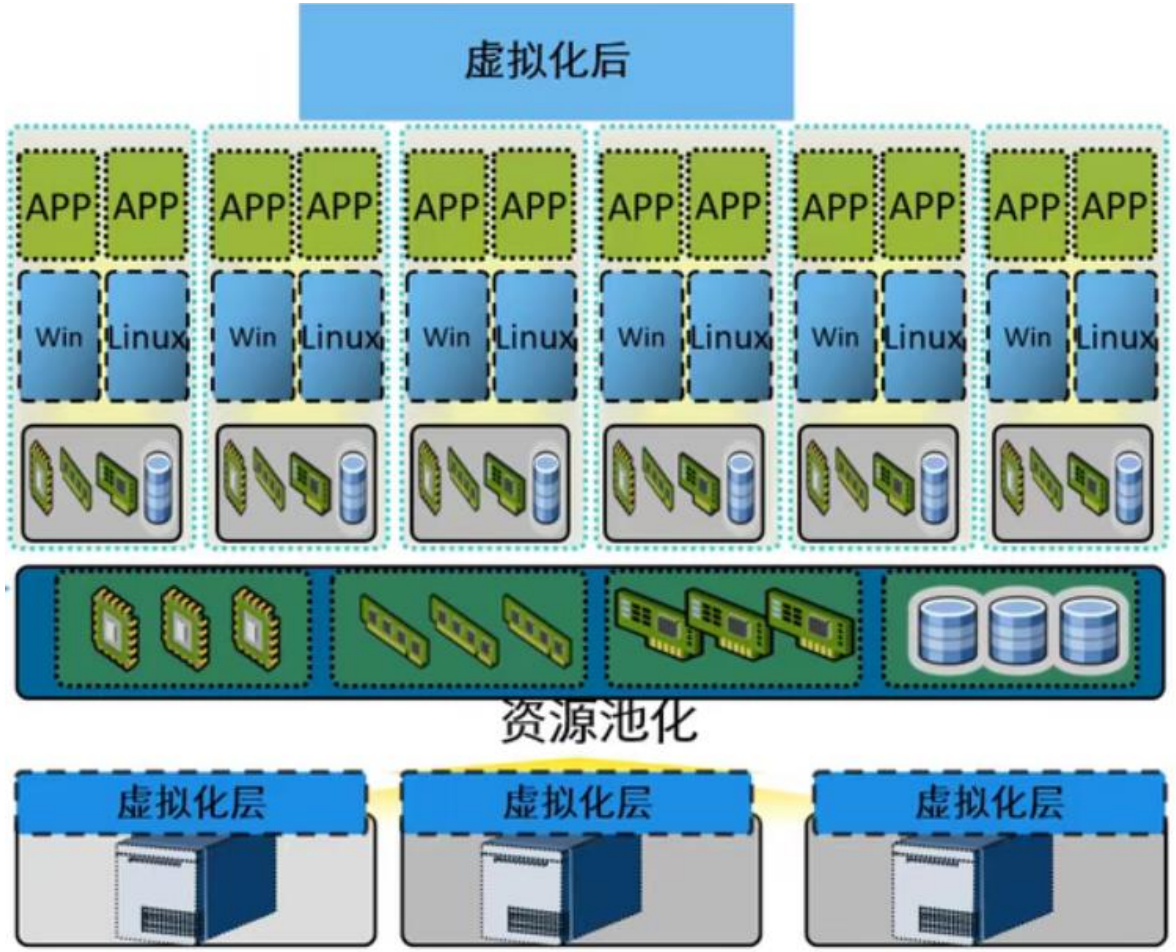


1. 虚拟化定义

- 虚拟化是将信息系统的各种物理资源，如服务器、网络、内存及数据等，进行抽象、转换后呈现出来，打破实体结构间的不可切割的障碍，使用户可以更好地应用这些资源。

1. 虚拟化定义

- 虚拟化资源包括计算能力和存储能力。
- 虚拟化技术实现了软件与硬件的分离，用户不需要考虑后台的具体硬件实现，而只需在虚拟层环境上看待资源和运行自己的系统及软件。

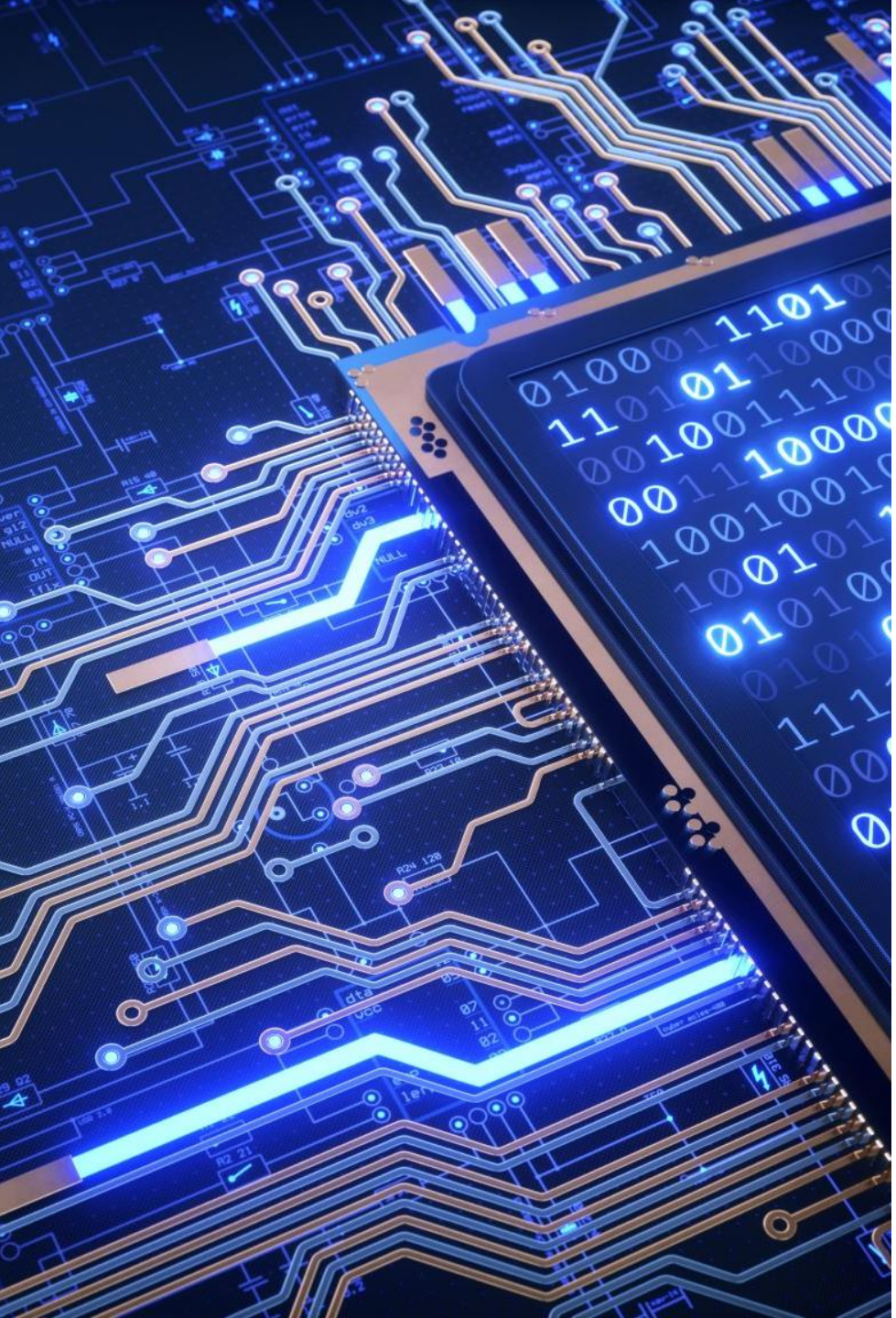


2. 虚拟化的优势

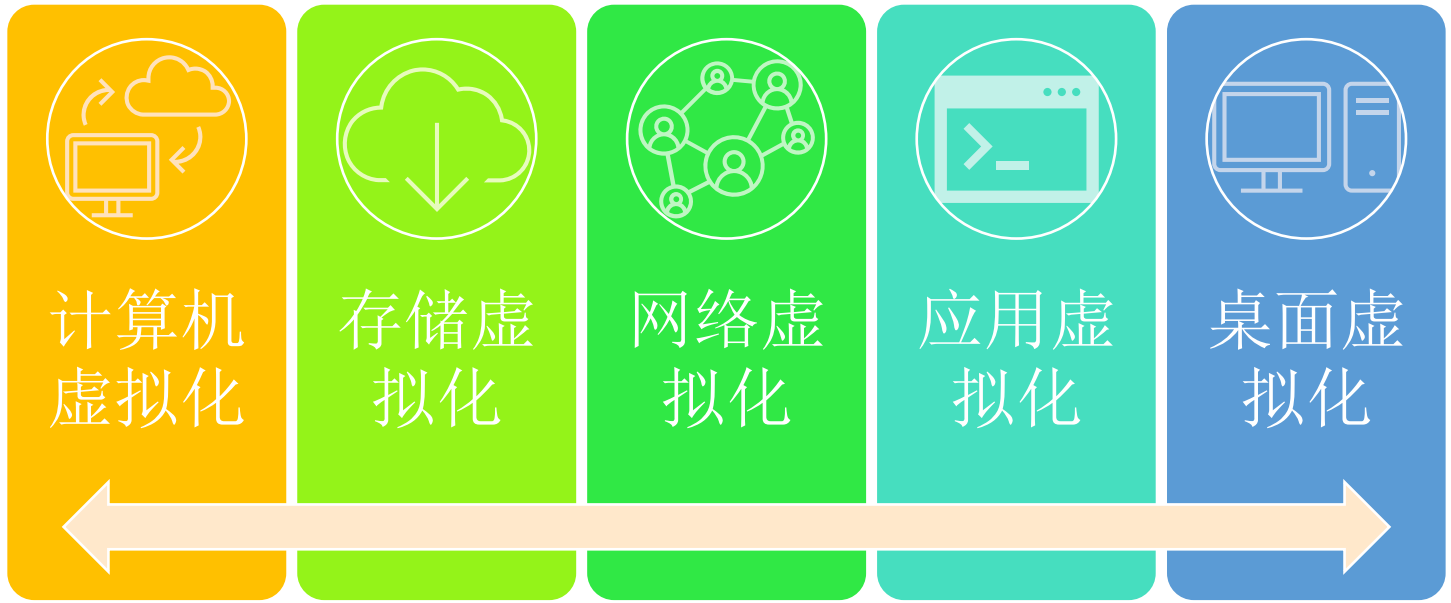
- 虚拟化可以用于资源的重组和隔离。在实际的生产环境中, 虚拟化技术经常用来解决高性能的物理硬件产能过剩和老旧硬件产能过低的重组重用问题，透明化底层物理硬件, 从而最大化利用物理硬件。通过虚拟化可以整合服务资源, 充分利用硬件资源, 大幅提升系统资源利用率。

2. 虚拟化的优势

- 整合服务器,提高资源利用率。通过整合服务器将共用的基础架构资源聚合到池中,打破原有的一台服务器一个应用程序模式。
- 降低成本,节能减排,构建绿色IT。由于服务器及相关IT硬件减少,因而减少了占地空间,也减少了电力和散热需求。管理工具更加出色,可帮助提高服务器/管理员比率,因此所需人员数据也将随之减少。
- 资源池化，提升IT灵活性。
- 统一管理， 提升系统管理效率。
- 完善业务的连续性。



4.3.2 虚拟化分类

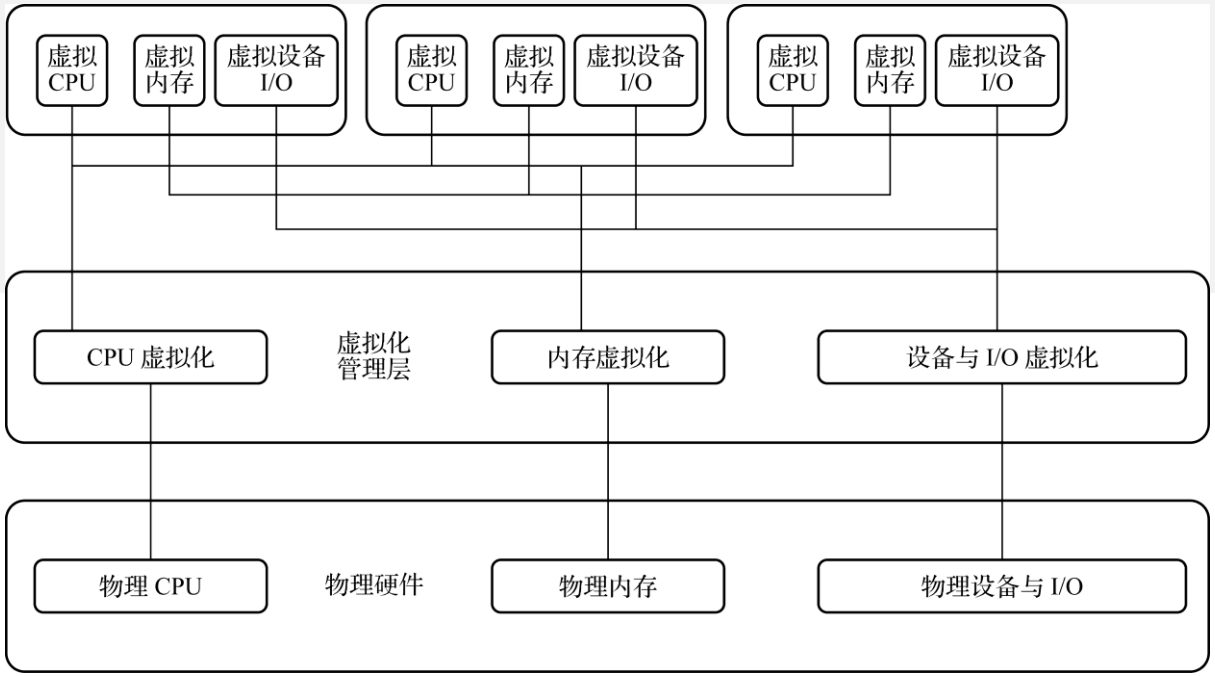


计算机虚拟化

计算机虚拟化也称服务器虚拟化，可以将一个物理计算机虚拟成若干个计算机使用，每个安装在虚拟计算机上的操作系统和运行的应用程序,不会察觉虚拟机与实际硬件的区别。

计算机虚拟化

- 下图说明了计算机虚拟化的原理：物理资源通过虚拟化软件形成虚拟资源池，再由资源池中分出部分资源构成虚拟机。运行多个虚拟计算机可以充分发挥物理计算机的计算潜能，迅速应对数据中心不断变化的需求。计算机虚拟化是基础设施即服务的基础。



计算机虚拟化需要具备的功能和技术

多实例。在一个物理计算机上可以运行多个虚拟计算机。

隔离性。在多实例的计算机虚拟化中,将一个虚拟机与其他虚拟机完全隔离,以保证良好的可靠性及安全性。

CPU虚拟化。把物理CPU抽象成虚拟CPU,任何时间,一个物理CPU只能运行一个虚拟CPU的指令,多个虚拟机同时提供服务将会大大提高物理CPU的利用率。

内存虚拟化。统一管理物理内存,将其包装成多个虚拟的物理内存,分别供给若干个虚拟机使用，使得每个虚拟机拥有各自独立的内存空间,互不干扰。

设备与I/O虚拟化。统一管理物理机的真实设备,将其包装成多个虚拟设备给若干个虚拟机使用响应每个虚拟机的设备访问请求和I/O请求。

无知觉故障恢复。运用虚拟机之间的快速热迁移技术(live migration),可以将故障虚拟机上的用户在没有明显感觉的情况下迅速转移到另一个新开的正常虚拟机上。

负载均衡。利用调度和分配技术,平衡各个虚拟机和物理机之间的利用率。

统一管理。由多个物理计算机支持多个虚拟机的动态,包括实时生成、启动、停止、迁移、调度、负荷、监控等,有一个方便易用的统一管理界面。

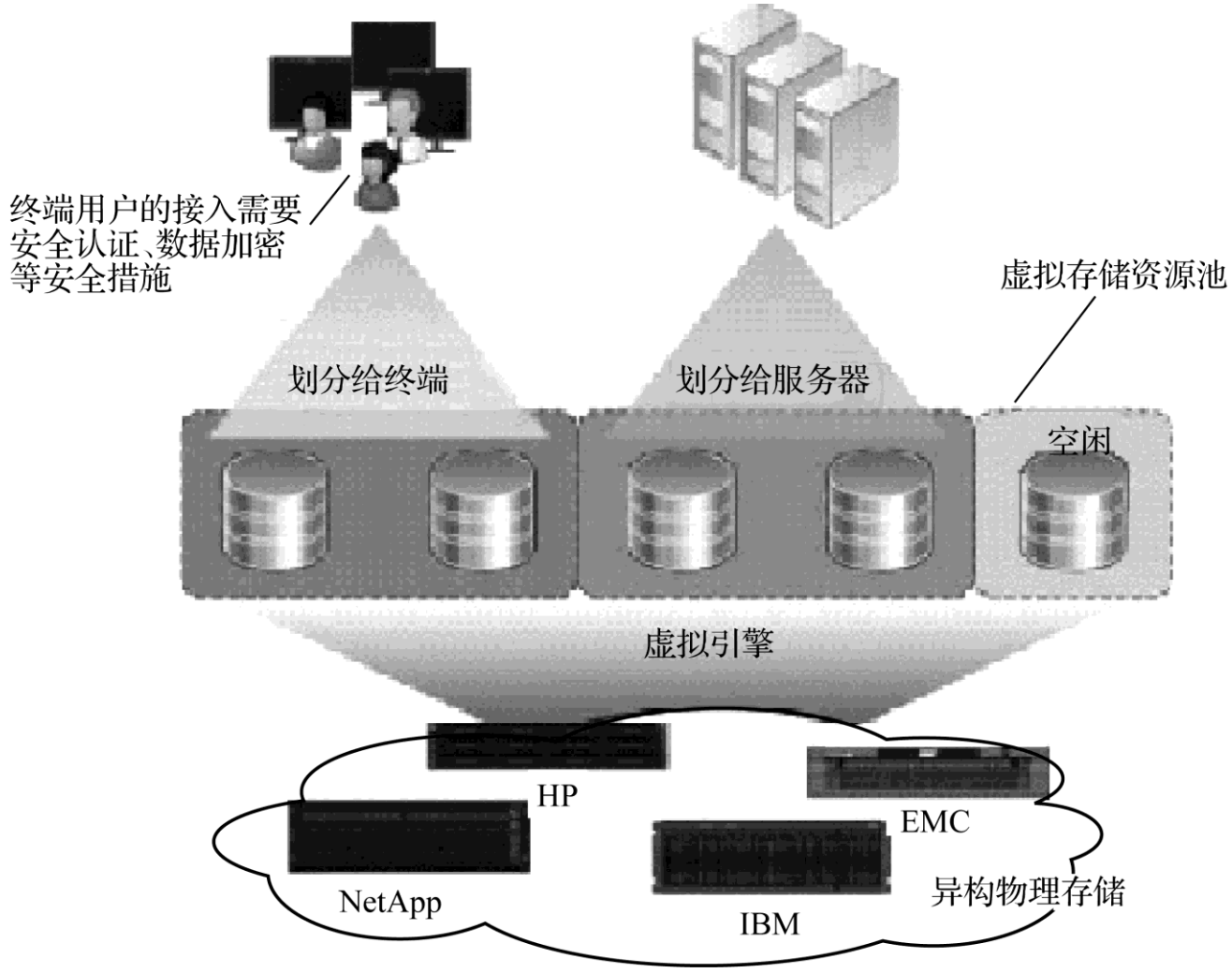
快速部署。整个系统需要一套快速部署机制,对多个虚拟机及其不同的操作系统和应用进行高效部署、更新和升级。

存储虚拟化

- 存储虚拟化就是把多个存储介质模块(如硬盘、RAID)集中管理起来,所有的存储介质模块在一个存储池中统一管理 ,从计算机角度,看到的不是多个硬盘,而是一个分区或者卷,就好像是一个超大容量的硬盘。这种可以将多种、多个存储设备统一管理起来,为使用者提供大容量、高数据传输性能的存储系统,称为虚拟存储。

存储虚拟化

如图4-3所示，虚拟引擎把多种物理存储虚拟成存储资源池，用户可从存储资源池中直接得到存储资源，不需要知道物理存储的细节。



存储虚拟化的功能和特点

集中存储。存储资源统一整合管理，集中存储，形成数据中心模式。

分布式扩展。存储介质易于扩展，由多个异构存储服务器实现分布式存储，以统一模式访问虚拟化后的用户接口。

绿色环保。服务器和磁盘的耗电量巨大，为提供全时段数据访问，存储服务器及磁盘不可停机。但为了节能减排、绿色环保，需要利用更合适的协议和存储模式，尽可能减少开启服务器和磁盘的次数。

安全认证。新建用户加人云存储系统前，必须经过安全认证并获得证书。

数据加密。为保证用户数据的私密性，将数据存储到云存储系统时必须加密。加密后的数据除被授权的特殊用户外，其他人一概 无法解密。

层级管理。支持层级管理模式，即上级可以监控下级的存储数据，而下级无法查看上级或平级的数据。

存储虚拟化

- 虚拟存储设备主要通过大规模的磁盘阵列(RAID)子系统和多个I/O通道连接到服务器上,由智能控制器提供逻辑单元(LUN)访问控制、缓存和其他(如数据复制)的管理功能。存储设备管理员对存储设备拥有完全的控制。存储与服务器系统分开,存储的管理与多种服务器操作系统隔离,可以很容易地调整硬件参数。

网络虚拟化

- 网络虚拟化可分为纵向分割和横向整合

网络虚拟化
——纵向分割

- 早期的网络虚拟化是指虚拟专用网络(VPN)。VPN是对网络连接的概念进行的抽象,允许远程用户访问组织的内部网络,就像物理上连接到该网络一样。网络虚拟化可以帮助保护IT环境,防止来自Internet的威胁,同时使用户能够快速安全地访问应用程序和数据。
- 随后,网络虚拟化技术随着数据中心的业务要求发展为多种应用承载在一张物理网络上,通过网络虚拟化分割(纵向分割)功能使得不同企业机构相互隔离,但可在同一网络上访问其自身应用,从而实现了将物理网络进行逻辑纵向分割,虚拟化为多个网络。
- 如果把一个企业网络分隔成多个不同的网络,让它们使用不同的规则和控制,用户就可以充分利用基础网络的虚拟化功能,而不是部署多套网络来实现各路隔离机制。
- 网络虚拟化并不是什么新概念，因为多年来,虚拟局域网(VLAN)技术作为基本隔离技术已经被广泛应用。当前在交换网络上,通过VLAN来区分不同业务网段、配合防火墙等安全产品划分安全区域，是数据中心基本设计内容之一。

网络虚拟化
——横向整合

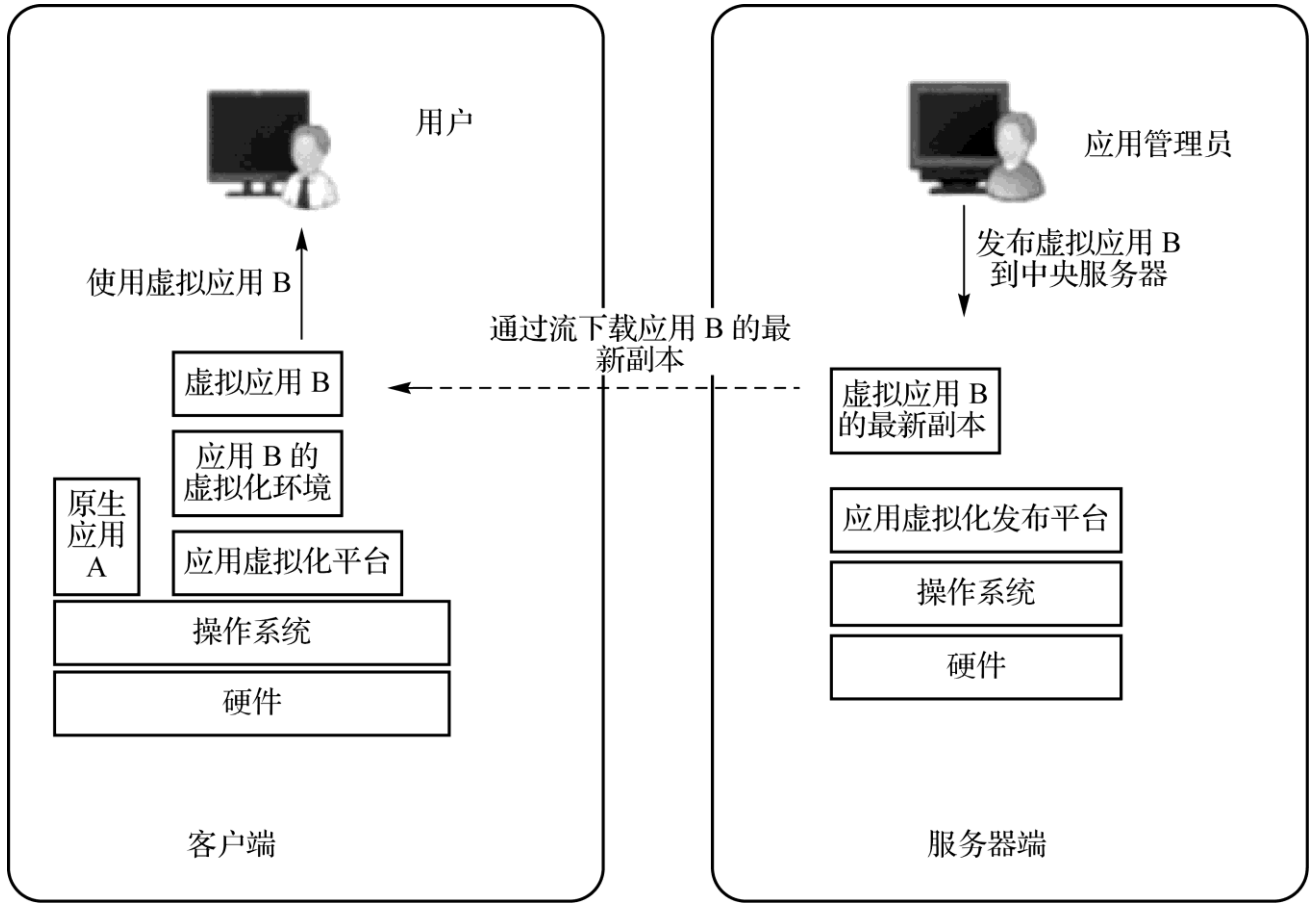
- 多个网络节点（包括网络交换路由设备、服务器等）承载上层应用，基于冗余的网络设计带来复杂性，而将多个网络节点进行整合，虚拟化成一台逻辑设备，在提升数据中心网络可用性、节点性能的同时将极大简化网络架构。
- 使用网络虚拟化技术，用户可以将多台设备连接，横向整合起来组成一个联合设备，并将这些设备视为单一设备进行管理和使用。虚拟化整合后的设备组成一个逻辑单元，在网络中表现为一个网元节点，使管理简单化、配置简单化，可跨设备链路聚合，极大简化了网络架构，同时进一步增强了冗余可靠性。

应用虚拟化

- 应用虚拟化是将应用软件从操作系统中分离出来，将应用对底层系统和硬件的依赖抽象出来，从而解除应用与操作系统和硬件的耦合关系。

应用虚拟化

应用程序运行在本地应用虚拟化环境中时，这个环境为应用程序屏蔽了底层可能与其他应用产生冲突的内容，如图4-4所示。应用虚拟化是SaaS的基础。



应用虚拟化需要具备的功能和特点

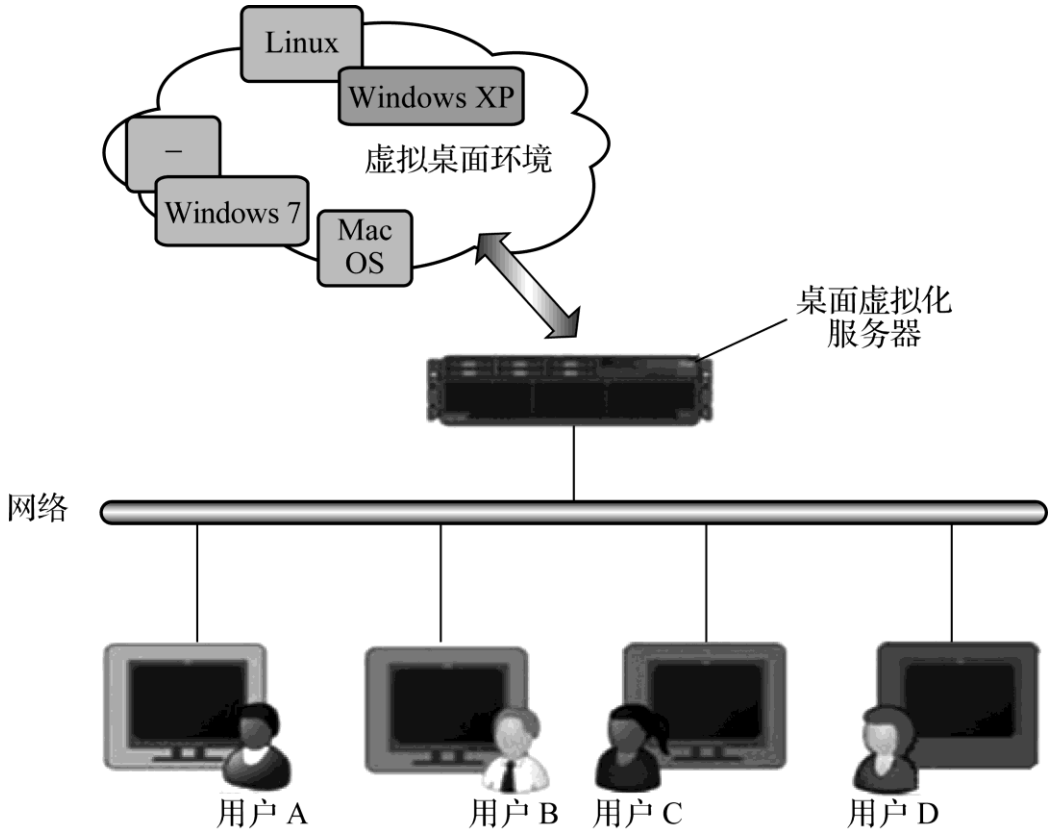
- 解耦合。利用屏蔽底层异构性的技术解除虚拟应用与操作系统和硬件的耦合关系。
- 共享性。应用虚拟化可以使一个 真实应用运行在任何共享的计算资源上。
- 虚拟环境。应用虚拟化为应用程序提供了一个虚拟的运行环境,不仅拥有应用程序的可执行文件,还包括所需的运行环境。
- 兼容性。虚拟应用应屏蔽底层可能与其他应用产生冲突的内容，从而使其具有良好的兼容性。
- 快速升级更新。真实应用可以快速升级更新,通过流的方式将相对应的虚拟应用及环境快速发布到客户端。
- 用户自定义。用户可以选择自己喜欢的虚拟应用的特点及所支持的虚拟环境。

桌面虚拟化

- 桌面虚拟化技术把所有应用客户端系统一次性地部署在数据中心的一台专用服务器上。客户端系统不需要通过网络向每个用户发送实际的数据，只传送虚拟的客户端界面(屏幕图像更新、按键、鼠标移动等)并显示在用户的计算机上。这个过程对最终用户是一目了然的,最终用户的感觉好像是实际的客户端软件正在桌面上运行一样。

桌面虚拟化

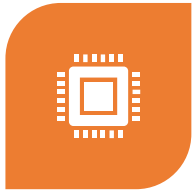
- 桌面虚拟化将用户的桌面环境与其使用的终端设备分开。服务器上存放的是每个用户的完整桌面环境。用户可以使用具有足够处理和显示功能的不同终端设备通过网络访问该桌面环境, 如图4-5所示。



桌面虚拟化具有的功能和接入标准

- 集中管理维护。集中在服务器端管理和配置PC环境及其他客户端需要的软件，可以对企业数据、应用和系统进行集中管理、维护和控制,以减少现场支持工作量。
- 使用连续性。确保终端用户下次在另一个虚拟机上登录时,依然可以继续以前的配置和存储文件内容,让使用具有连续性。
- 故障恢复。桌面虚拟化是将用户的桌面环境保存为一个个虚拟机,通过对虚拟机进行快照和备份,就可以快速恢复用户的故障桌面,并实时迁移到另一个虚拟机上继续进行工作。
- 用户自定义。用户可以选择自己喜欢的桌面操作系统、显示风格、默认环境,以及其他各种自定义功能。

4.3.3
x86 虚拟化技术



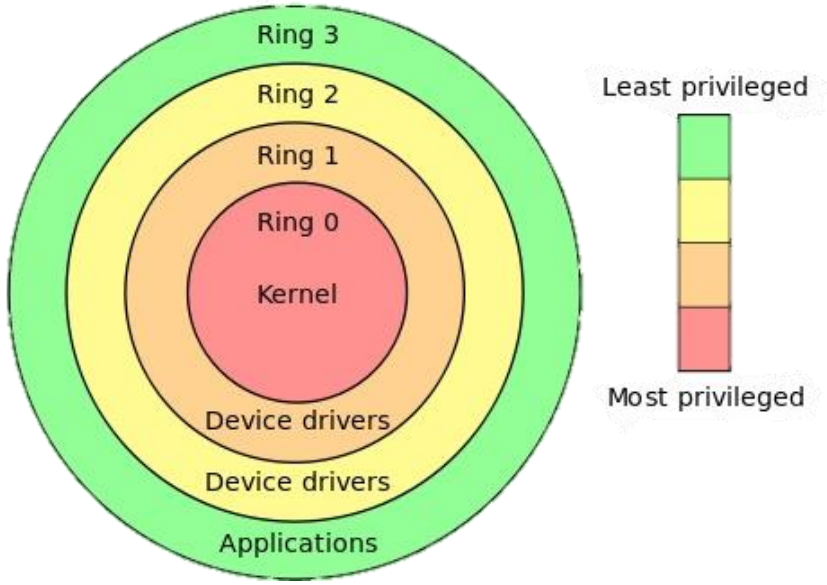
CPU虚拟化

内存虚拟化

I/O虚拟化

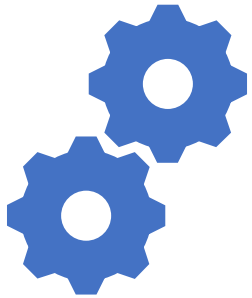
知识链接：CPU的特权级别

- Intel的CPU将特权级别分为4个级别：ring0，ring1，ring2，ring3。
- ring0是指CPU的运行级别，ring0是最高级别，ring1次之，ring2更次之…… 以Linux为例，操作系统（内核）的代码运行在最高运行级别ring0上，可以使用特权指令，控制中断、修改页表、访问设备等等。 应用程序的代码运行在最低运行级别上ring3上，不能做受控操作。如果要做，比如要访问磁盘，写文件，那就要通过执行系统调用（函数），执行系统调用的时候，CPU的运行级别会发生从ring3到ring0的切换，并跳转到系统调用对应的内核代码位置执行，这样内核就为你完成了设备访问，完成之后再从ring0返回ring3。这个过程也称作用户态和内核态的切换。



CPU虚拟化

- CPU虚拟化的目标是使虚拟机上的指令能被正常地执行,而且效率接近物理机。



全虚拟化

- 全虚拟化主要采用优先级压缩(ring compression)和二进制代码翻译技术(binary translation)。优先级压缩能让VMM和Guest运行在不同的特权级下,对x86架构而言,就是VMM运行在最高特权级Ring0下,Guest的内核代码运行在Ring1下,Guest的应用代码运行在Ring3下。通过这种方式能让VMM截获一部分在Guest上执行的特权指令,并对其进行虚拟化。但是有一些对虚拟化不友好的指令则需要二进制代码翻译来处理,它通过扫描并修改Guest的二进制代码将那些难以虚拟化的指令转化为支持虚拟化的指令。

半虚拟化

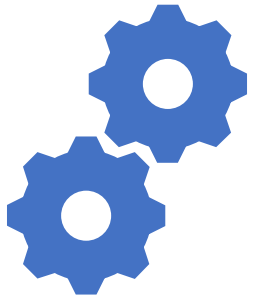
- 通过修改GuestOS的代码,使其将那些和特权指令相关的操作都转换为发给VMM的Hypercall(超级调用),且Hypercall支持批处理和异步这两种优化方式,使得通过Hypercall得到近似于物理机的速度。

硬件辅助虚拟化

- 硬件辅助虚拟化主要包含Intel的VT-x和AMD的AMD-V两种技术,且这两种技术在核心思想上非常相似,都是通过引人新的指令和运行模式,让VMM和GuestOS分别运行在其合适的模式下。在实现方面, VT-x支持两种处理工作方式:第一种称为Root模式, VMM运行于此模式下,用于处理特殊指令;第二种称为Non-Root模式, Guest OS运行于此模式下,当在Non-Root模式Guest执行到特殊指令时,系统会切换到运行Root模式的VMM上,让VMM来处理这个特殊指令。

内存虚拟化

- 内存虚拟化的目标是做好虚拟机内存空间之间的隔离，使每个虚拟机都认为自己拥有了整个内存地址，并且效率也能接近物理机。



全虚拟化

- 影子页表(shadow page table)就是为每个Guest都维护一个影子页表,在这个表中写入虚拟化之后的内存地址映射关系,而Guest OS的页表则无须变动,最后,VMM将影子页表交给MMU进行地址转换。

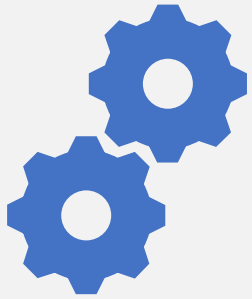
半虚拟化

- 使用页表写入法,当Guest OS创建一个新的页表时,其会向VMM注册该页表,之后在Guest运行时,VMM将不断地管理和维护这个表,使Guest上面的程序能直接访问到合适的地址。

硬件辅助虚拟化

- 扩展页表(extended page table, EPT)通过使用硬件技术,能在原有页表的基础上增加一个EPT页表通过这个页表能够将Guest的物理地址直接翻译为主机的物理地址，降低整个内存虚拟化所需的成本。在EPT推出之前,硬件辅助虚拟化技术在内存虚拟化方面有一个TLB(translation lookaside buffer) Miss 的缺陷。

I/O虚拟化



I/O虚拟化的目标是不仅让虚拟机访问到其所需要的I/O资源，而且要做好它们之间的隔离工作，更重要的是降低由虚拟化所带来的开销。

全虚拟化

- 通过模拟I/O设备(磁盘和网卡等)来实现虚拟化。对GuestOS而言,所能看到就是一组统一的I/O设备,Guest OS每次进行I/O操作时都会陷入VMM，让VMM来执行。这种方式对Guest而言非常透明,无须顾忌底层硬件，如Guest操作的是SCSI的设备，但实际物理机只有SATA的硬盘。

半虚拟化

- 通过前端(front end)/后端(back end)架构,将Guest 的I/O请求通过一个环状队列传递到特权域(privileged domain,也被称为Domain 0)。

硬件辅助虚拟化

- 硬件辅助虚拟化最具代表性的是Intel的VT-d、ADM的IOMMU和PCI-SIG的IOV(I/O virtulization)这三个技术。VT-d 的核心思想是让虚拟机能够直接使用物理设备,但这会涉及I/O地址访问和DMA问题,而VT-d 通过采用DMA重映射(remapping)和I/O页表来解决这两个问题,从而让虚拟机能够直接访问物理设备。

x86虚拟化技术总结

- 如果使用最新的芯片,如45nm的Nehalem和32nm的Westmare,那么硬件辅助虚拟化技术是一个比较好的选择,甚至胜过半虚拟化技术。
- 如果是运行很多TLBMiss的应用(如Java应用),那么应避免使用硬件辅助虚拟化技术。总体而言,就像VMware的白皮书Virtual Machine Monitor ExecutionModes:in VMware vSphere 4.0总结的那样,如果是使用最新45 nm以下的Intel芯片和较新的操作系统,那么推荐使用硬件辅助虚拟化技术,其他使用全虚拟化技术。

x86虚拟化技术总结

- 虽然现在硬件辅助虚拟化有TLBMiss这个缺陷,但随着硬件辅助虚拟化技术不断地发展和优化,其在速度和架构方面的优势更明显。不过,全虚拟化和半虚拟化的一些技术在某些方面还是保持了一定的优势,如半虚拟化的前端和后端架构及全虚拟化的二进制代码翻译技术。所以,今后x86虚拟化技术将会以硬件辅助虚拟化技术为主,同时以全虚拟化技术和半虚拟化技术为辅。

作业

1. 什么是虚拟化技术？虚拟化技术如何进行分类？

2. 采用虚拟化技术后，可以获得哪些优势？