第三章
云计算的主流技术

www.huawei.com

版权所有© 2018 华为技术有限公司

HUAWEI

前言

- 本章主要讲述了云集的主流技术，介绍了虚拟化优势以及虚拟化的优势。讲述了Hypervisor的作用。讲述了容器的概念，容器和虚拟化的区别。

泰克教育
TECH EDUCATION

版权所有© 2018 华为技术有限公司

HUAWEI

目标

- 学完本课程后，您将能够：
  - 描述虚拟化和容器的概念；
  - 描述Hypervisor的作用；
  - 区分虚拟化和容器；

泰克教育
TECH EDUCATION

版权所有© 2018 华为技术有限公司

HUAWEI

目录

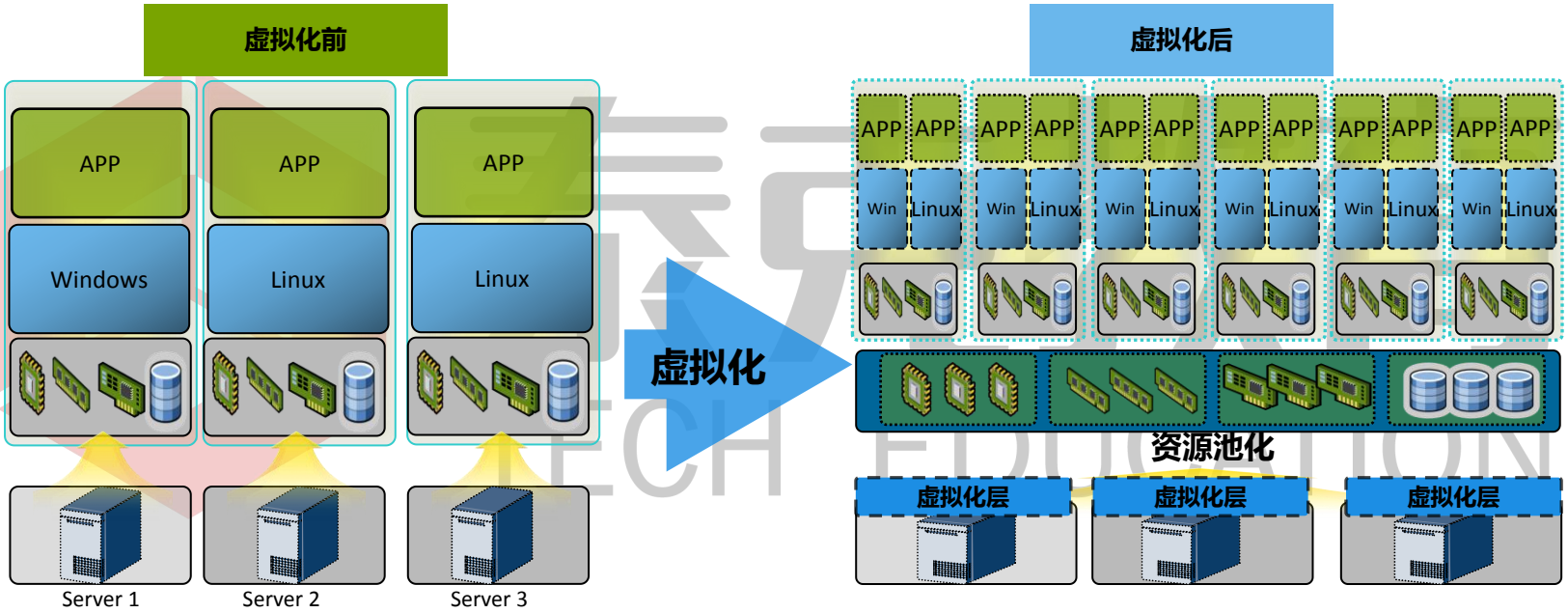
泰克教育 TECH EDUCATION

版权所有© 2018 华为技术有限公司

HUAWEI

什么是虚拟化

- 虚拟化 (Virtualization) 的含义很广泛。将任何一种形式的资源抽象成另一种形式的技术都是虚拟化。虚拟化是资源的逻辑表示，其不受物理限制的约束。

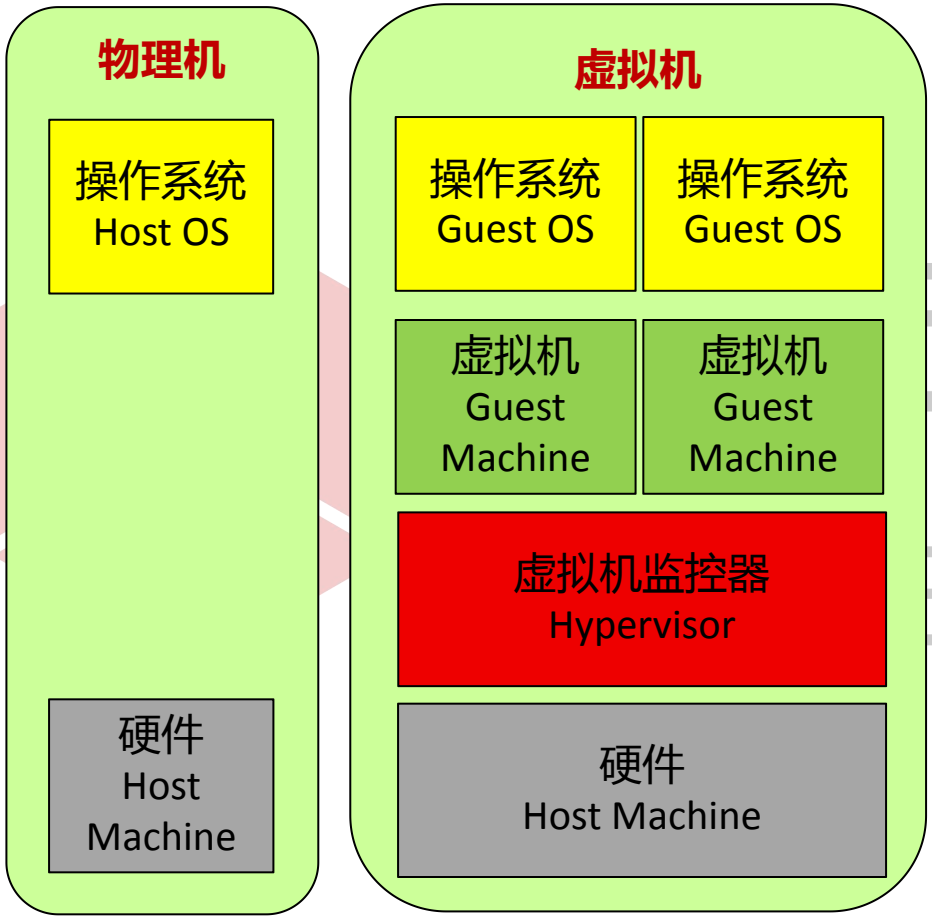


- IT资源独立。
- 操作系统必须与硬件紧耦合。

- 资源抽像成共享资源池。
- 上层操作系统与硬件解耦，操作系统从资源池中分配资源。

版权所有© 2018 华为技术有限公司

虚拟化中的几个重要概念



Guest OS：

运行在虚拟机之上的OS

Guest Machine：

虚拟出来的虚拟机

Hypervisor：

虚拟化软件层/虚拟机监控机

（Virtual Machine Monitor，VMM）

Host OS：

运行在物理机之上的OS

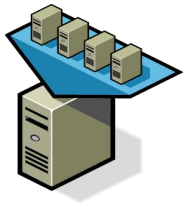
Host Machine：

物理机

版权所有© 2018 华为技术有限公司

虚拟化的特点

分区

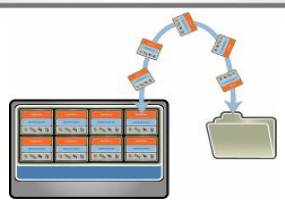


在单一物理服务器上同时运行多个虚拟机。

隔离

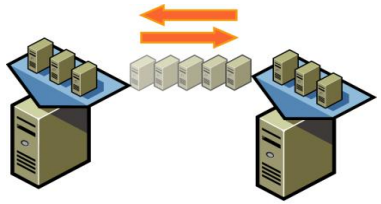
在同一服务器上的多个虚拟机之间相互隔离。

封装



整个虚拟机执行环境封装在独立文件中，可以通过移动文件的方式来迁移该虚拟机。

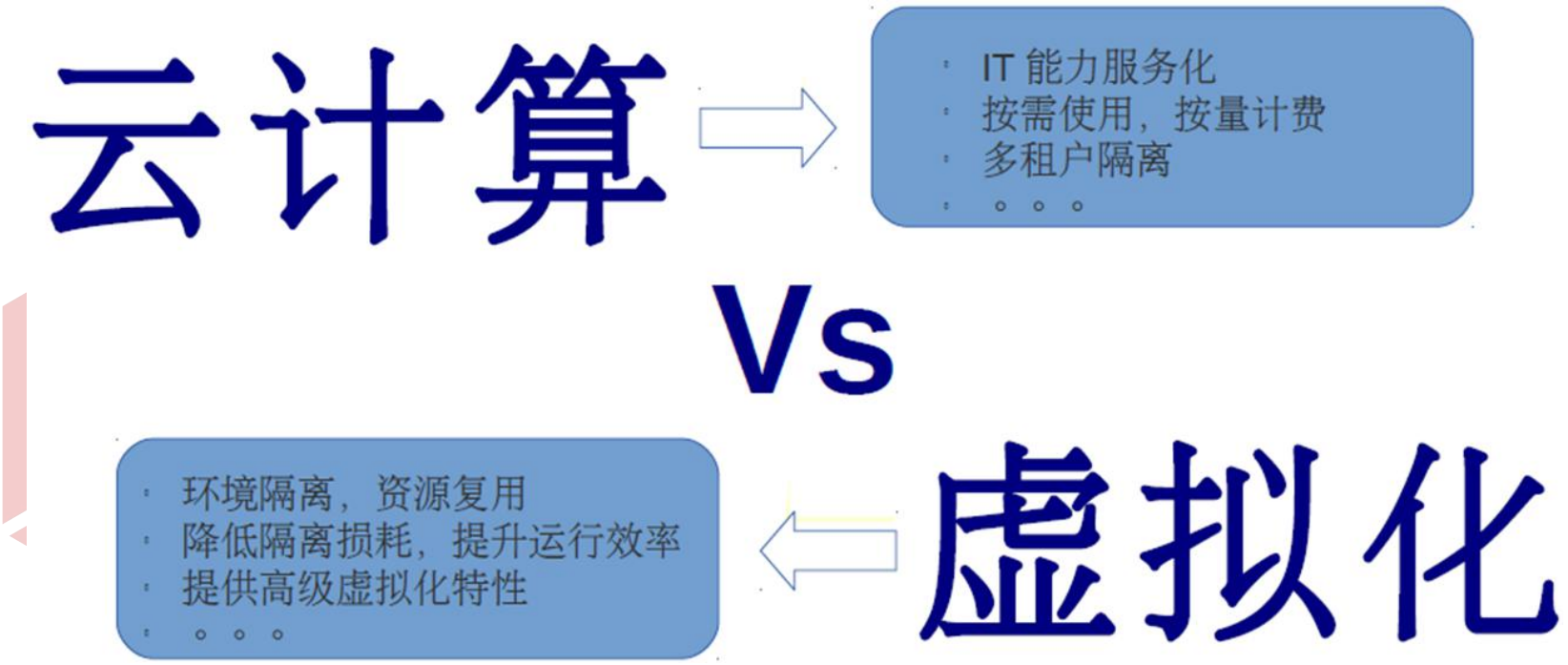
相对于硬件独立



虚拟机无需修改，即可在任何服务器上运行。

版权所有© 2018 华为技术有限公司

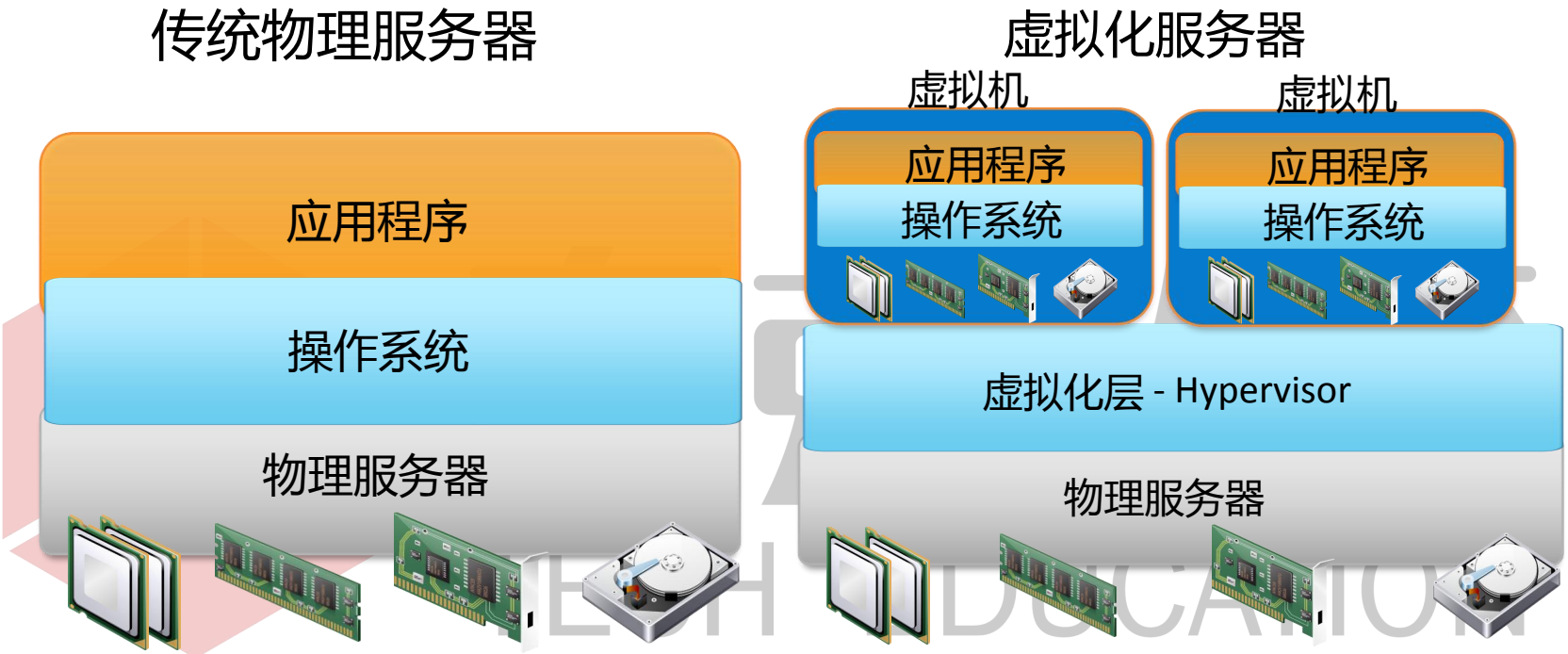
云计算 VS 虚拟化



虚拟化是实现云计算的技术支撑手段之一，但并非云计算的核心关注点。
虚拟化技术是云计算在IaaS层具有商用价值的基础。

版权所有© 2018 华为技术有限公司

虚拟化的优势（1/2）



操作系统与物理服务器绑定

难以迁移　可靠性难以控制

难以扩展　资源利用率低

空间占用高　　难以管理

操作系统与物理服务器分离

易于迁移、扩展，资源整合

标准化的虚拟硬件

由一系列文件组成，易于保护

版权所有© 2018 华为技术有限公司

HUAWEI

虚拟化的优势（2/2）

性能	虚拟化前	虚拟化后
资源利用率	每台主机一个操作系统，系统的资源利用率低。	主机与操作系统不一一对应，按需分配使用，系统的资源利用率高。
独立性	软硬件紧密结合，硬件成本高昂且不够灵活。	操作系统和硬件不相互依赖，虚拟机独立于硬件，能在任何硬件上运行。
程序运行效率	同一台主机上同时运行多个程序容易产生冲突，运行效率较低。	管理操作系统和应用程序被封装成单一个体，不同个体间不冲突。同一台机器上运行同一个程序，效率高。
安全性	安全性较差。	强大的安全和故障隔离。

TECH EDUCATION

版权所有© 2018 华为技术有限公司

HUAWEI

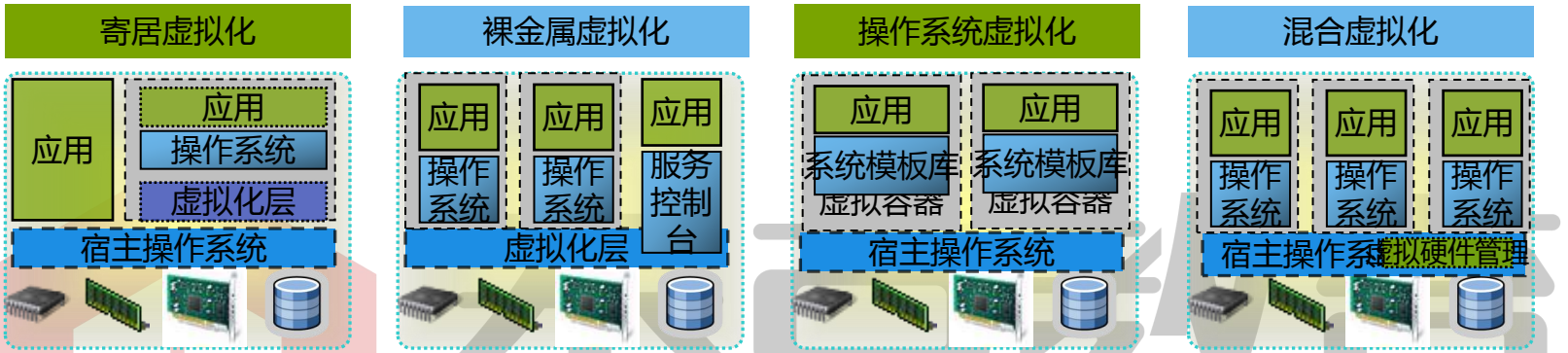
目录

泰克教育 TECH EDUCATION

版权所有© 2018 华为技术有限公司

虚拟化架构（1/2）

根据在整个系统中的位置不同



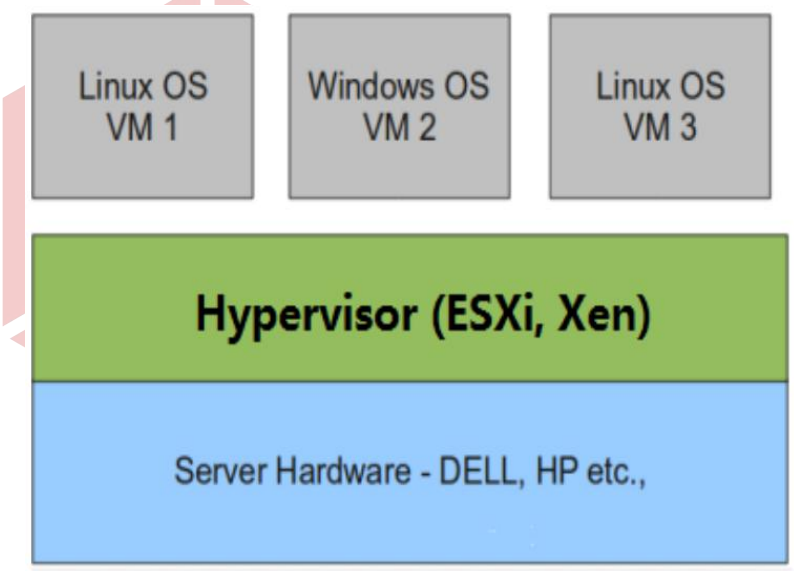
	寄居虚拟化	裸金属虚拟化	操作系统虚拟化	混合虚拟化
优点	• 简单、易于实现。	• 虚拟机不依赖于操作系统。 • 支持多种操作系统，多种应用 。	• 简单、易于实现。 • 管理开销非常低。	• 相对于寄居虚拟化架构，没有冗余，性能高。 • 可支持多种操作系统。
缺点	• 安装和运行应用程序依赖于主机操作系统对设备的支持。 • 管理开销较大，性能损耗大。	• 虚拟层内核开发难度大。	• 隔离性差，多容器共享同一操作系统。	• 需底层硬件支持虚拟化扩展功能。
厂家	• VMware Workstation	• WMware ESXServer • Citrix XenServer • 华为 FusionSphere	• Virtuozzo	• Redhat KVM

版权所有© 2018 华为技术有限公司

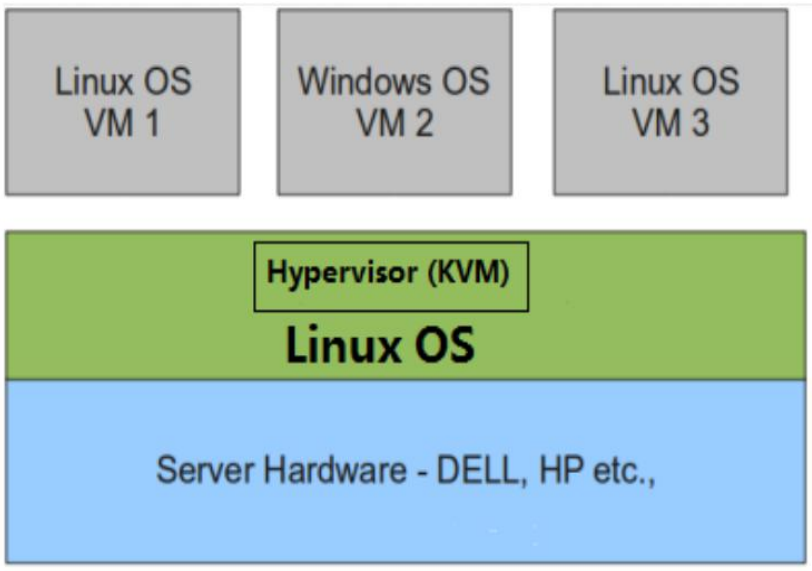
虚拟化架构（2/2）

根据Hypervisor的实现方式和所处的位置

1型虚拟化



2型虚拟化



版权所有© 2018 华为技术有限公司

目录

泰克教育 TECH EDUCATION

版权所有© 2018 华为技术有限公司

HUAWEI

虚拟机与VMM

- 虚拟机 (Virtual Machine) 是由虚拟化层提供的高效、独立的虚拟计算机系统，其皆拥有自己的虚拟硬件（CPU，内存，I/O设备）。
- 通过虚拟化层的模拟，虚拟机在上层软件看来，其就是一个真实的机器。这个虚拟化层一般称为虚拟机监控器 (Virtual Machine Monitor, VMM)，也称Hypervisor。

东方教育 TECH EDUCATION

版权所有© 2018 华为技术有限公司

HUAWEI

VMM的功能

- 虚拟资源

VMM利用底层硬件资源来构建一个包含虚拟CPU、内存和外设等的虚拟环境。在这个环境中，Guest OS认为自己运行在一台真是的计算机上，并唯一拥有这台“虚拟”机器上的所有资源。

- 虚拟环境的调度

VMM可以同时构建多个虚拟机环境，从而允许多个Guest OS并发执行，VMM利用一套策略来有效的调度资源。

- 虚拟化环境的管理接口

VMM提供一组完备的管理接口，来支持虚拟环境的创建、删除、暂停和迁移等功能。上层的管理程序通过调用VMM提供的管理接口，为用户提供管理界面。

泰克教育 TECH EDUCATION

版权所有© 2018 华为技术有限公司

HUAWEI

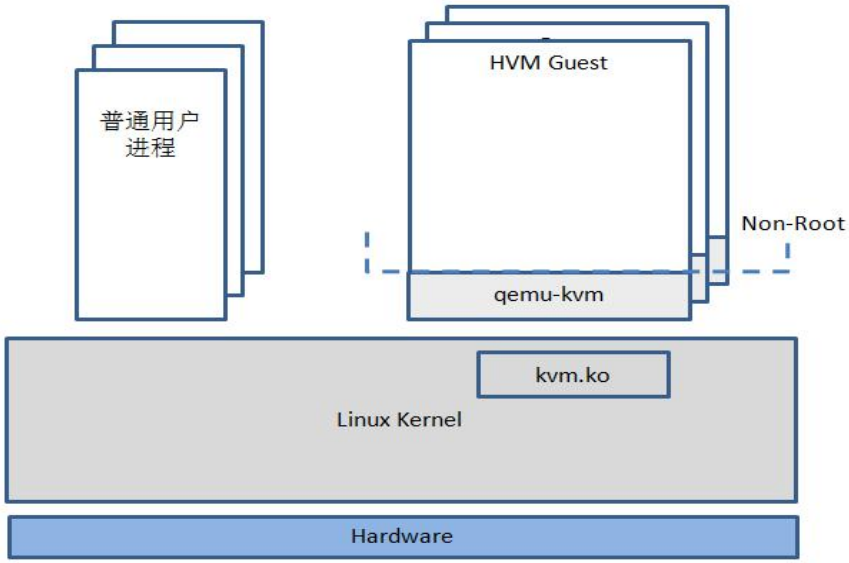
目录

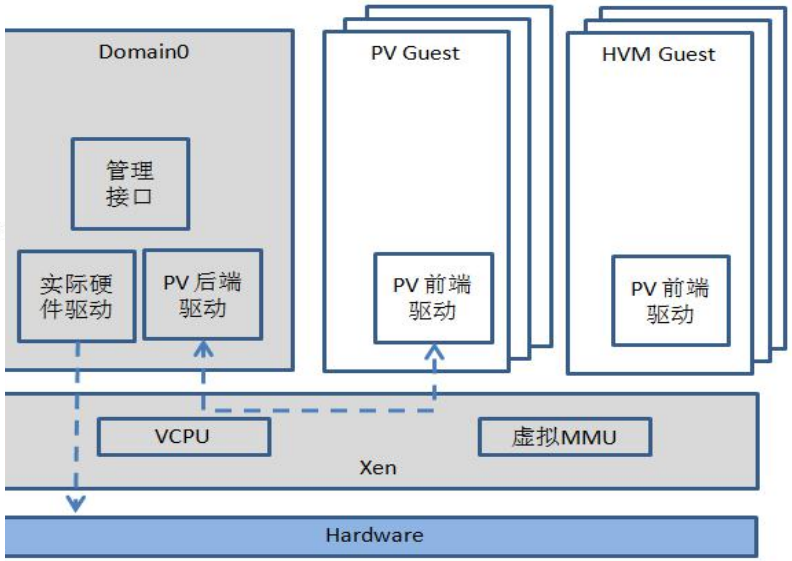
泰克教育 TECH EDUCATION

版权所有© 2018 华为技术有限公司

HUAWEI

KVM架构 VS Xen架构





- **KVM**
  - 内核模块,使得内核成为hypervisor
    - 呈现给用户空间字符设备/dev/kvm
    - 用户空间通过ioctl()访问
  - Guest是一个普通的进程
    - vcpu是一个线程
  - 充分利用linux内核支持
    - 调度，内存共享，QoS，电源管理等
  - 仅支持全虚拟化(Intel VT/AMD-V)

- **Xen**
  - 轻量级hypervisor
    - 直接运行在硬件上
    - CPU虚拟化、内存虚拟化
  - 特权虚拟机Domain0
    - IO虚拟化
    - 虚拟机管理
  - 支持全虚拟化+半虚拟化
    - 非硬件虚拟化平台: PV Guest

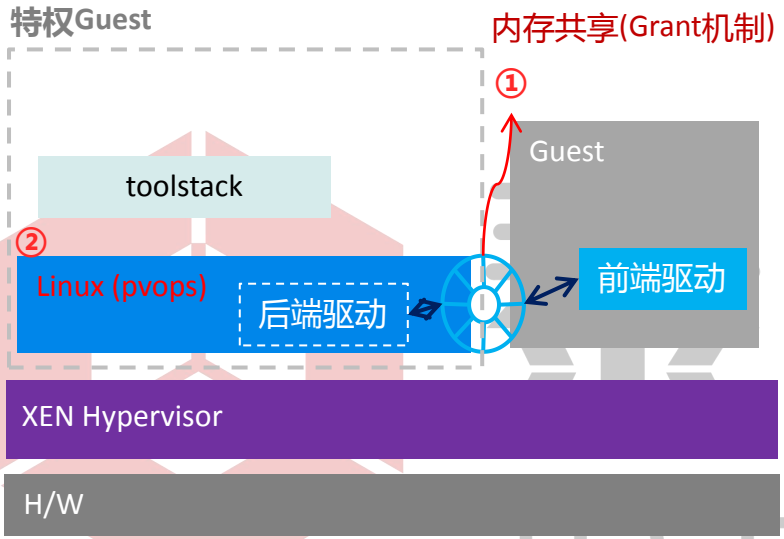
TECH EDUCATION

版权所有© 2018 华为技术有限公司

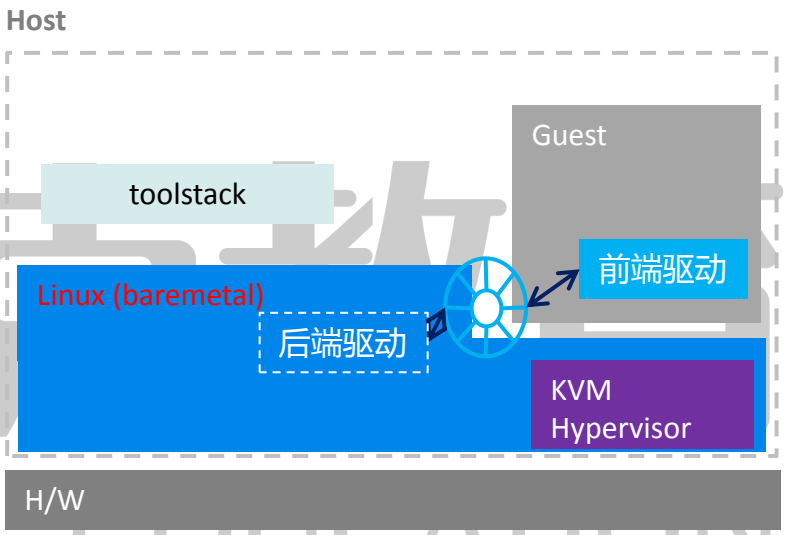
HUAWEI

Xen、KVM架构各有所长

Xen平台



KVM平台



Xen平台架构侧重安全性：

为保证安全性，各Domain之间对共享区域的访问和映射必须通过Hypervisor授权。

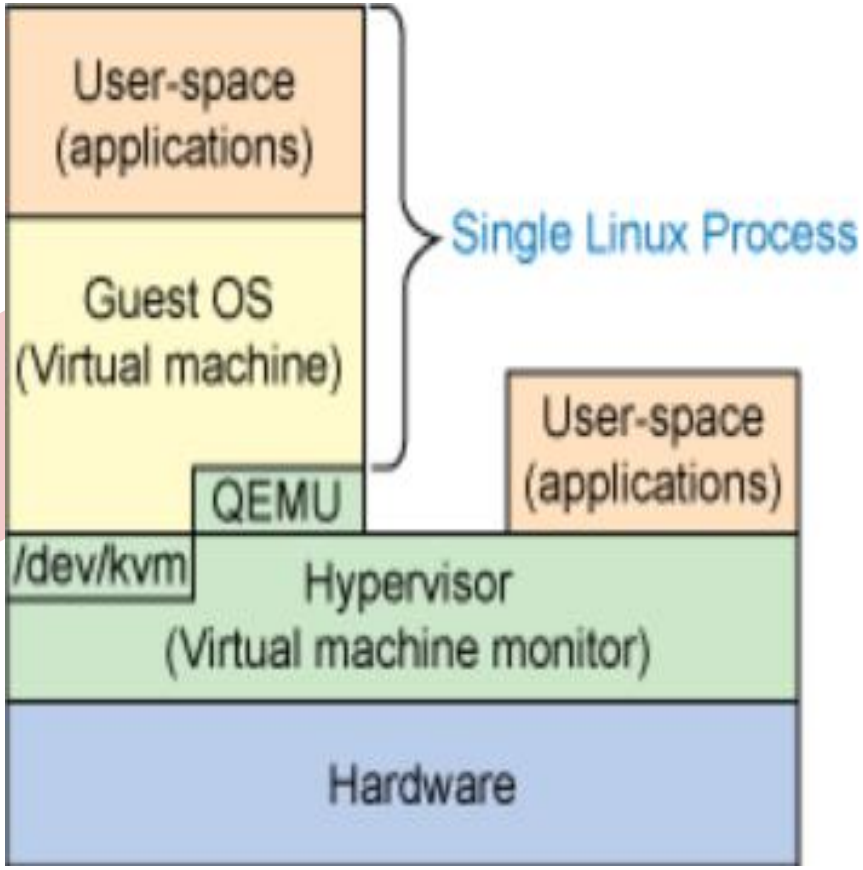
KVM平台架构侧重性能：

VM之间以及与Host Kernel之间对共享区域的访问和映射无需Hypervisor进行授权，故整个访问路径较短

使用Linux baremetal内核，无pvops性能损耗。

版权所有© 2018 华为技术有限公司

KVM



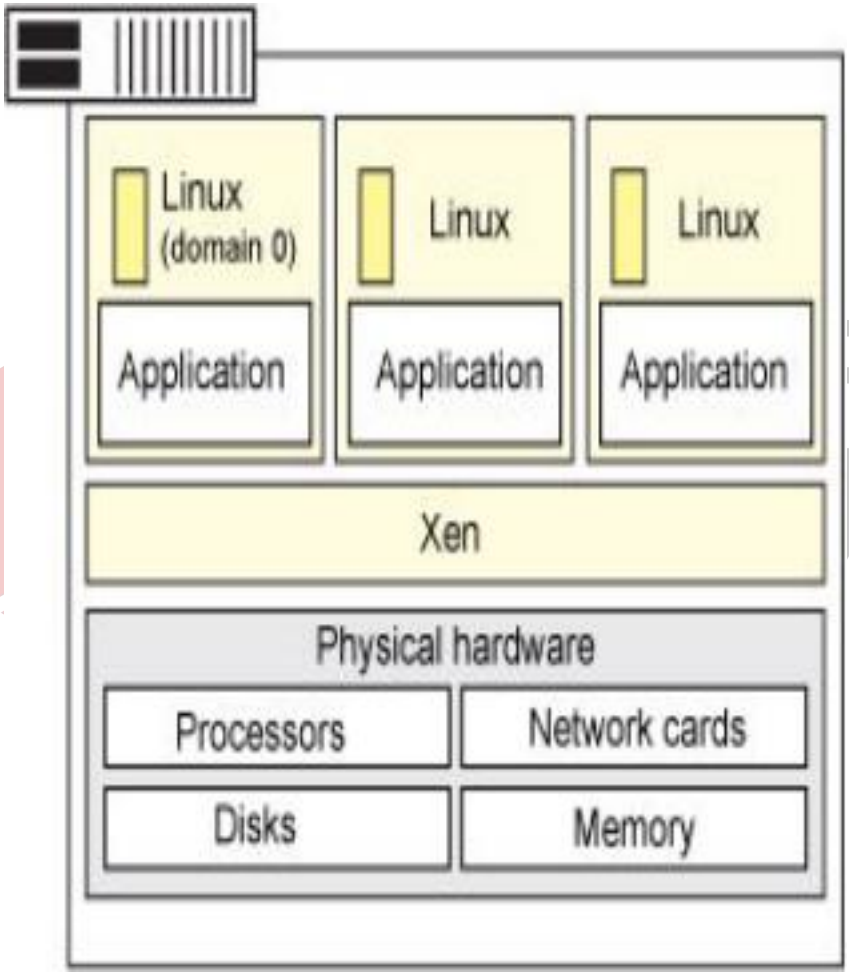
Guest：客户机系统，包括CPU（vCPU）、内存、驱动（Console、网卡、I/O 设备驱动等），被 KVM 置于一种受限制的 CPU 模式下运行。

KVM：运行在内核空间，提供CPU 和内存的虚级化，以及客户机的 I/O 拦截。Guest 的 I/O 被 KVM 拦截后，交给 QEMU 处理。

QEMU：修改过的为 KVM 虚机使用的 QEMU 代码，运行在用户空间，提供硬件 I/O 虚拟化，通过 IOCTL /dev/kvm 设备和 KVM 交互。

版权所有© 2018 华为技术有限公司

XEN



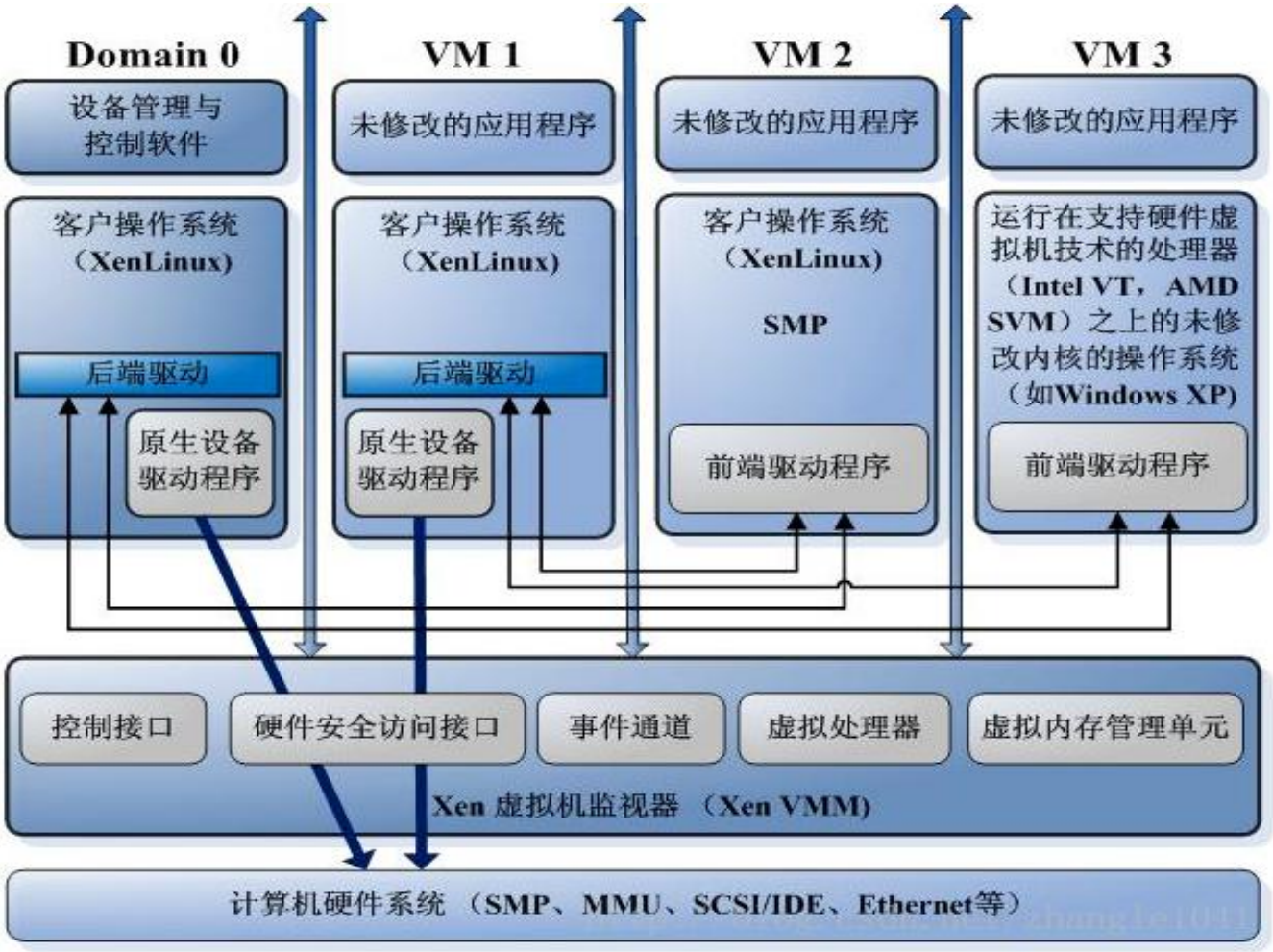
Xen Hypervisor：直接运行于硬件之上，是Xen客户操作系统与硬件资源之间的访问接口。通过将客户操作系统与硬件进行分类，Xen管理系统可以允许客户操作系统安全，独立的运行在相同硬件环境之上。

Domain 0：运行在Xen管理程序之上，具有直接访问硬件和管理其他客户操作系统的特权的客户操作系统。

Domain U：运行在Xen管理程序之上的普通客户操作系统或业务操作系统，不能直接访问硬件资源（如：内存，硬盘等），但可以独立并行的存在多个。

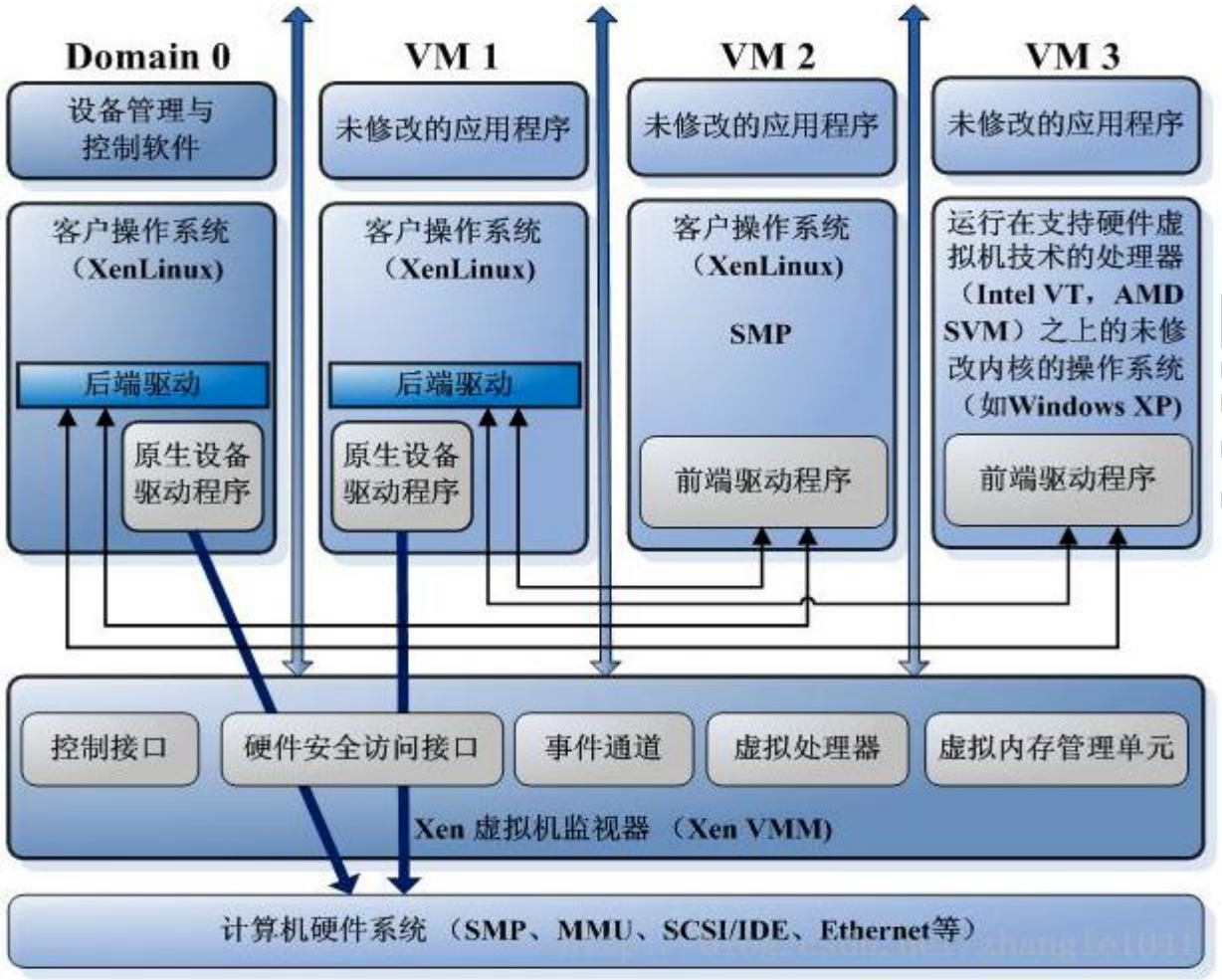
版权所有© 2018 华为技术有限公司

Xen架构简介



版权所有© 2018 华为技术有限公司

Xen架构简介



版权所有© 2018 华为技术有限公司

目录

泰克教育 TECH EDUCATION

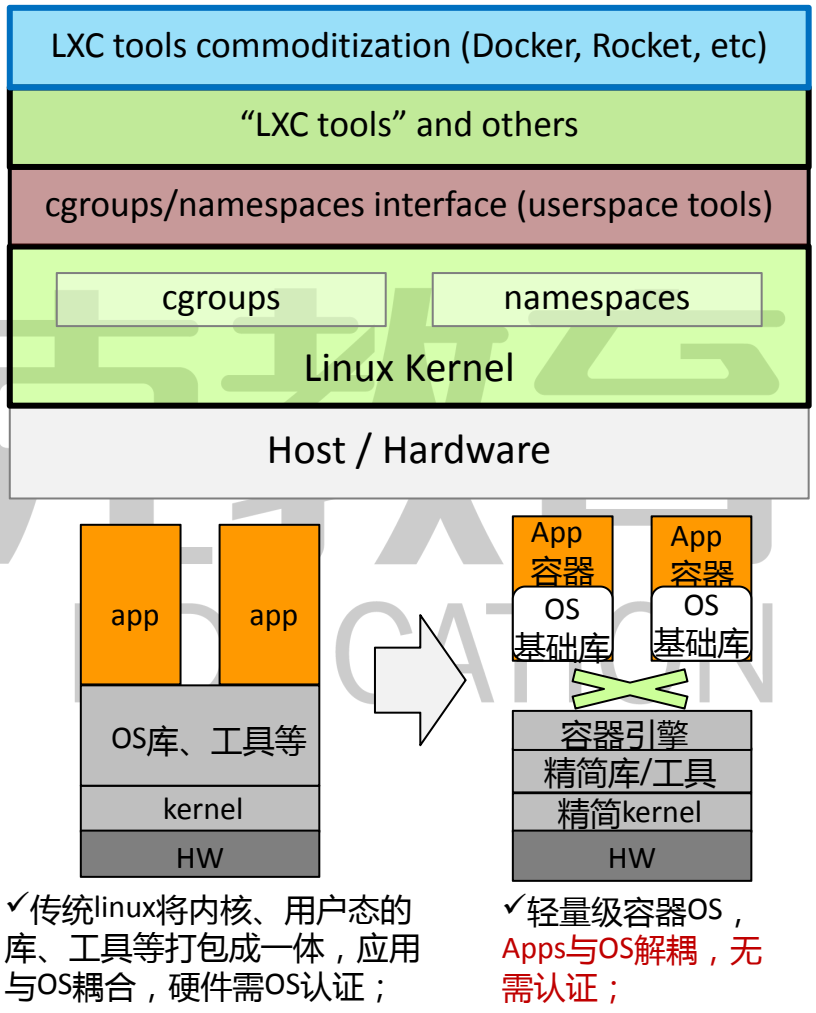
版权所有© 2018 华为技术有限公司

HUAWEI

容器是操作系统内核自带能力

- 容器是基于一些 Linux 内核的特性构建而成，Docker并没有重新发明这些特性。
  - cgroups: 主要做资源控制
  - namespaces: 主要做访问隔离
- LXC (Linux Containers) tools
  - Linux Container容器是一种内核虚拟化技术，可以提供轻量级的虚拟化，以便隔离进程和资源。
- 已有的容器技术
  - Docker
  - Rocket

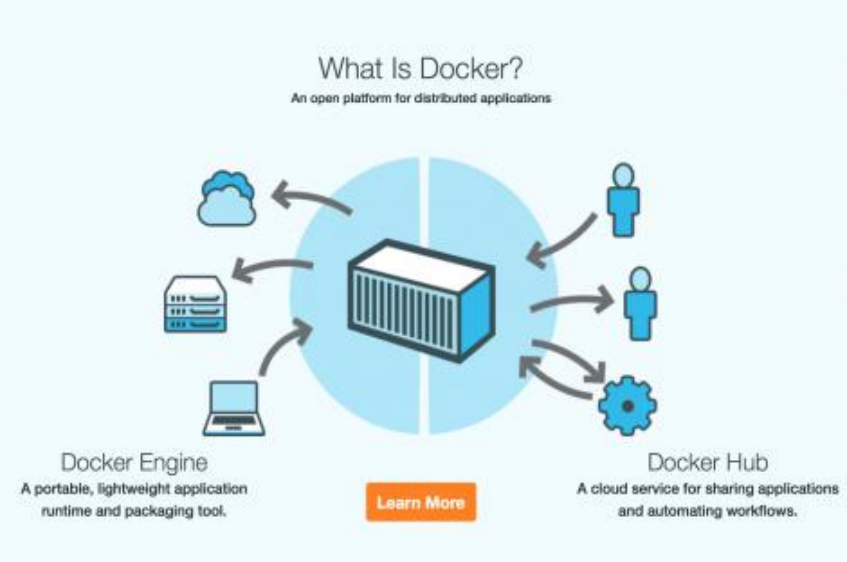
1. Docker 并没有发明容器，更像是容器的前端和外围工具。
2. Docker 核心在于实现应用与运行环境整体打包以及打包格式统一。
3. Docker 并不是容器技术的唯一选择。



版权所有© 2018 华为技术有限公司

Docker VS Container

- 传统Container：
  - 缺少自动化、使用复杂。
  - 用法与平台耦合性高、应用范围窄、用户限制大。
  - 只解决了Run，没有解决Build和Ship。
  - 各个容器的实现方式千差万别，缺省统一的标准。
- Docker：
  - 提供了Portable的标准并且提供了实现。
  - 基于该标准的容器Build和Ship机制。



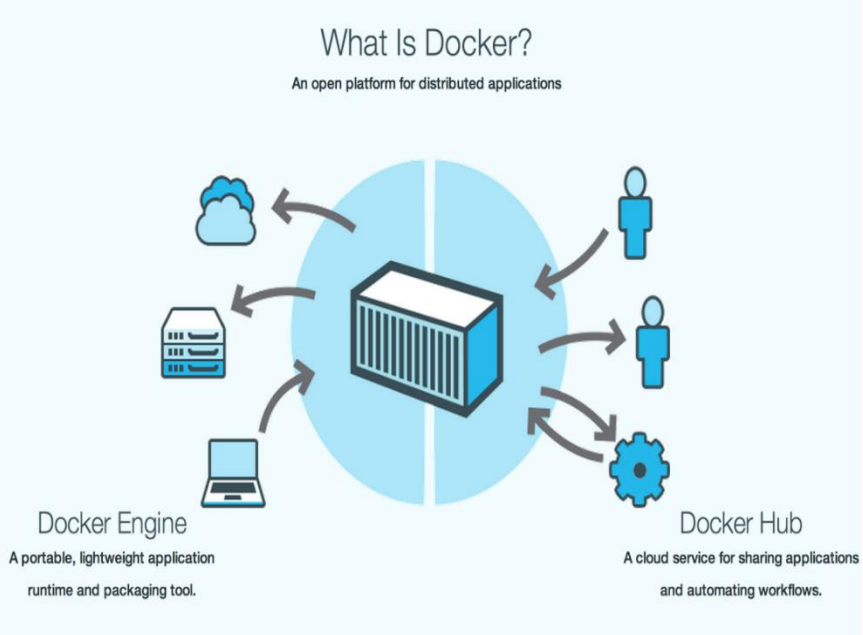
TECH EDUCATION

cgroup + name space + docker image = docker container

版权所有© 2018 华为技术有限公司

Docker容器技术概述

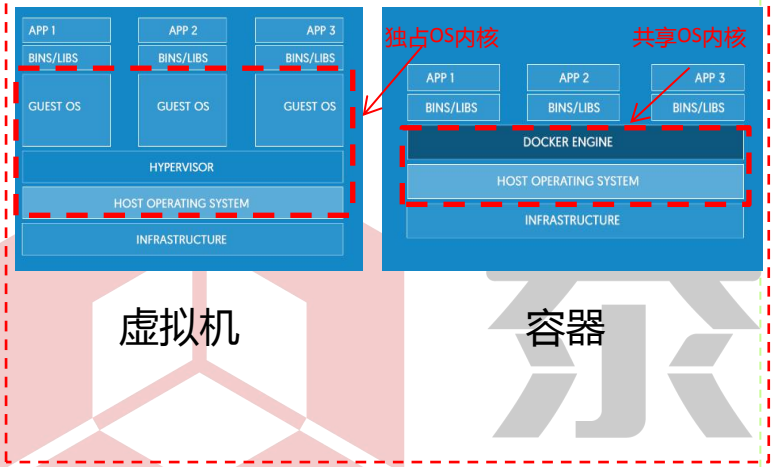
- Docker引擎
  - Docker 是一个开源的应用容器引擎，让开发者可以打包应用以及依赖包到一个可移植的容器中，然后发布到任何流行的 Linux 机器上。
  - 基于Go语言开发，遵从Apache2.0协议开源。



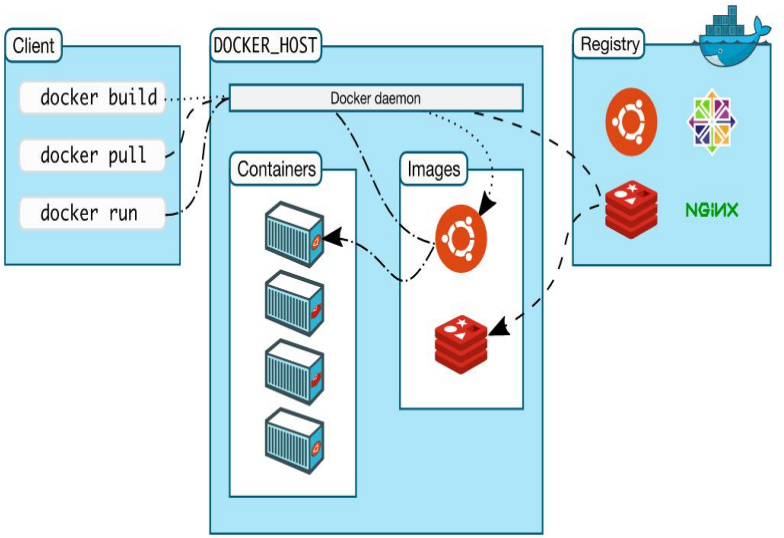
版权所有© 2018 华为技术有限公司

Docker容器技术原理介绍

虚拟机 VS Docker容器



Docker容器技术架构



Docker容器技术主要特点：

快：运行时的性能可以获取极大提升；

灵活：将应用和系统“容器化”，不添加额外的操作系统，支持跨OS部署；

轻便：你会拥有足够的“操作系统”，仅需添加或减小镜像即可，每台服务器可部署100~1000个实例；

廉价：版本开源的，免费的，低成本的；

生态：业界（微软、亚马逊、IBM、Cisco）主流IT厂商逐步使用Docker容器技术、开源社区活跃度非常高，逐步成为未来软件发展趋势；

Docker容器技术使用场景：

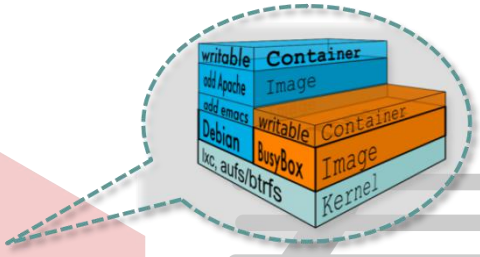
- 简化配置；
- 代码流水线（Code Pipeline）管理；
- 提高开发效率；
- 隔离应用；
- 快速部署；
- 支持多组环境；
- 整合服务器，降低资源成本；

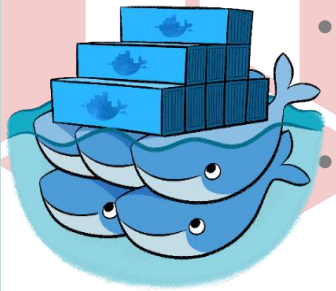
版权所有© 2018 华为技术有限公司

HUAWEI

Docker加速容器技术的发展

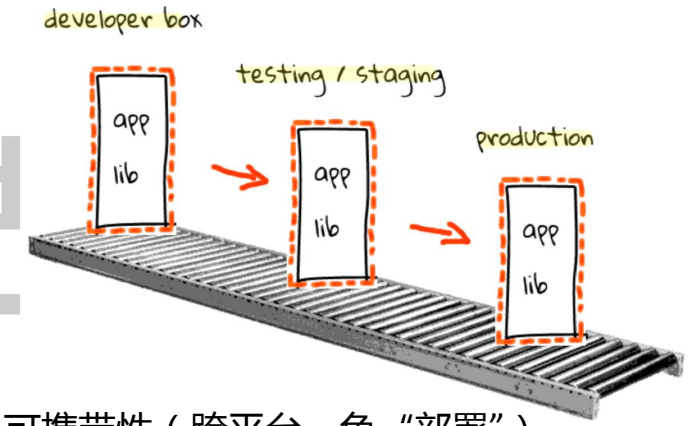
Docker提供软件应用的集装箱，创建即部署





- 集装箱式应用管理环境。
- image分层技术：多副本&重用，节省空间、分发迅速。
- 模块化打包、快速部署、标准化管理，独立&隔离运行。

Docker彻底改变程序的交付方式



- 可携带性（跨平台、免“部署”）
- 一致性（开发&测试&生产）
- 快速分发、复制
- 轻量、隔离、无环境依赖

Docker的核心价值：构建标准化(dockerfile)、交付形态标准化（容器 & Image）、运行环境标准化（Engine）。

版权所有© 2018 华为技术有限公司

容器技术发展

- Docker解决的问题
  - 应用环境管理复杂
    - OS，中间件，各种App
  - 减化环境管理复杂度，减化应用实例部署工作，将应用打成Image部署
    - Web应用，DB应用，Hadoop应用，消息队列
  - 提供分发和标准化管理

东方教育 TECH EDUCATION

版权所有© 2018 华为技术有限公司

HUAWEI

容器技术架构（1）

- Docker三组件
  - Docker Client：用户界面，支持用户与Docker Daemon之间通信。
  - Docker Daemon：运行于主机上，处理服务请求。
  - Docker Registry：支持拥有公有与私有访问权限的Docker容器镜像仓库。
- Docker三要素
  - Docker Containers：负责应用程序的运行，包括操作系统、用户添加的文件以及元数据。
  - Docker Images：构建容器的只读模板，用来运行Docker容器。
  - DockerFile：文件指令集，用来说明如何自动创建Docker镜像。

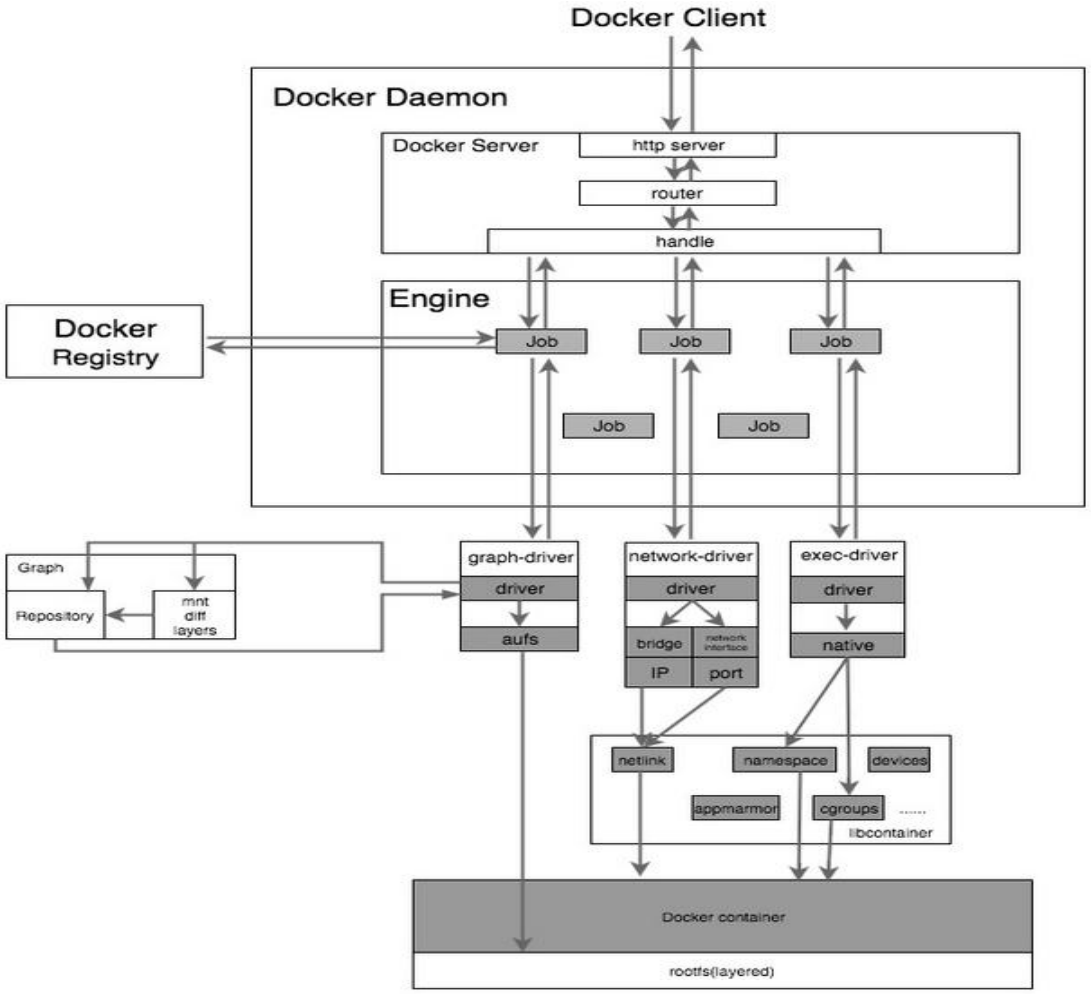
泰克教育 TECH EDUCATION

版权所有© 2018 华为技术有限公司

HUAWEI

容器技术架构（2）

- Docker总体架构
  - Docker daemon
  - Docker registry
  - Libcontainer
  - Docker container
  - Graph driver



版权所有© 2018 华为技术有限公司

容器技术架构（3）

- Docker总体架构
  - Docker 系统使用 C/S架构。
  - Server 端驻守在后台：docker daemon。
  - Docker client通过 REST API 请求 Docker daemon 来管理 Docker 的镜像和容器等。
  - Docker Client是一个 CLI 程序，可以在命令行中通过 Docker 二进制文件进行交互。

东先教育 TECH EDUCATION

版权所有© 2018 华为技术有限公司

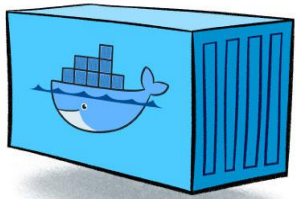
HUAWEI

Docker镜像制作和分发

镜像制作
（标准化）

Dockerfile

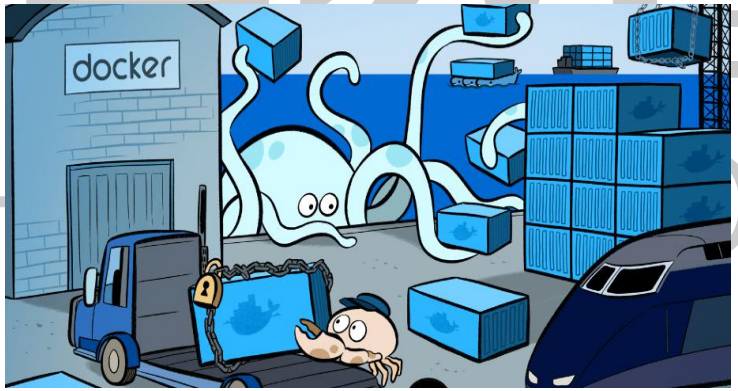
快速生成



镜像分发
（易用性）

本地环境

便捷获取



版权所有© 2018 华为技术有限公司

HUAWEI

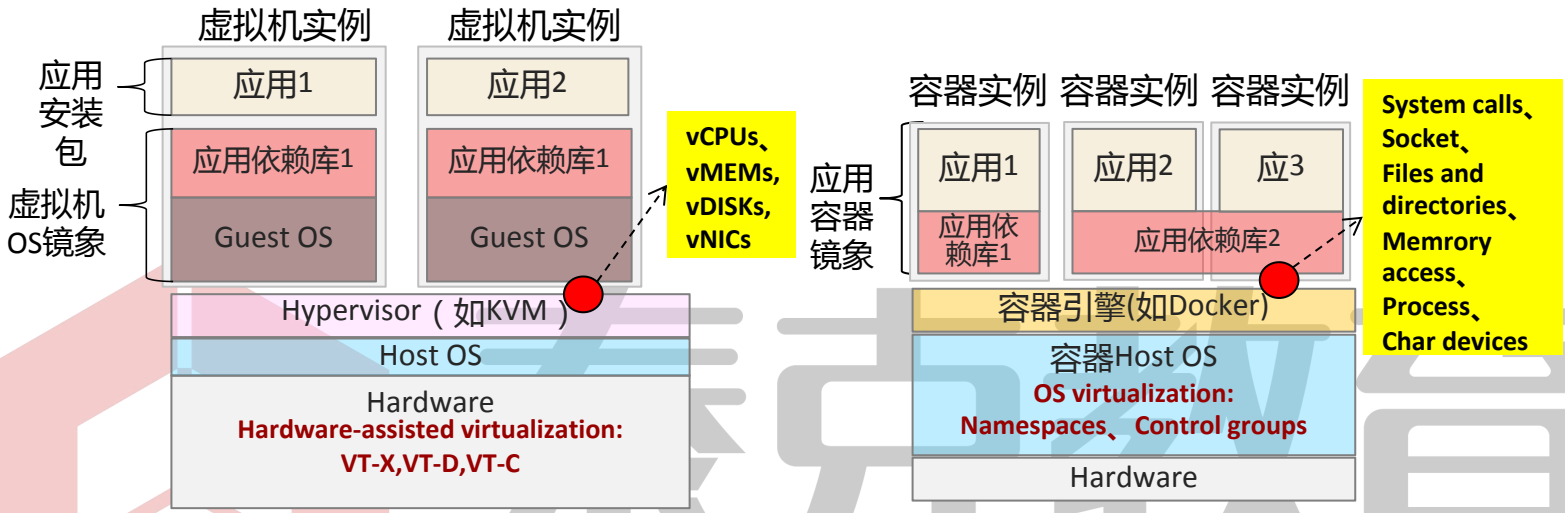
目录

泰克教育 TECH EDUCATION

版权所有© 2018 华为技术有限公司

HUAWEI

容器是一种轻量化的虚拟化技术

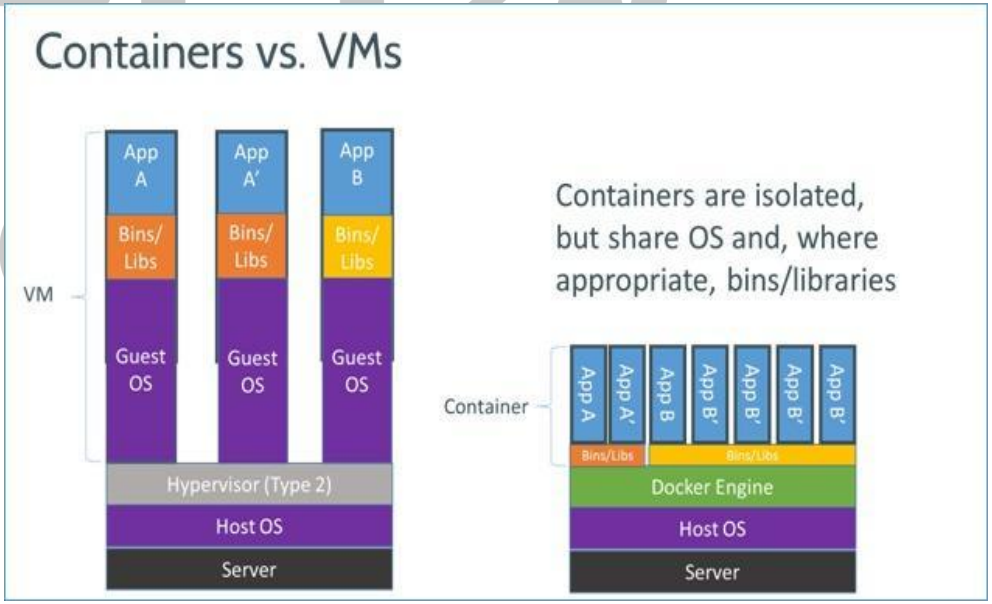


隔离性强，有独立的GUEST OS	共享内核和OS，隔离性弱
虚拟化性能差（>15%）	计算/存储无损耗，无GuestOS内存开销（~200M）
虚拟机镜像庞大（十几G~几十G），且实例化时不能共享	Docker容器镜象 200~300M，且公共基础镜象实例化时可以共享
虚拟机镜象缺乏统一标准	Docker提供了容器应用镜象事实标准，OCI推动进一步标准化
虚拟机创建慢（>2分钟）	秒级创建(<10s)
虚拟机启动慢（>30s）	秒级(<1s，不含应用本身启动)
资源虚拟化粒度低，单机10~100虚拟机	单机支持1000+容器

版权所有© 2018 华为技术有限公司

Docker容器的优势

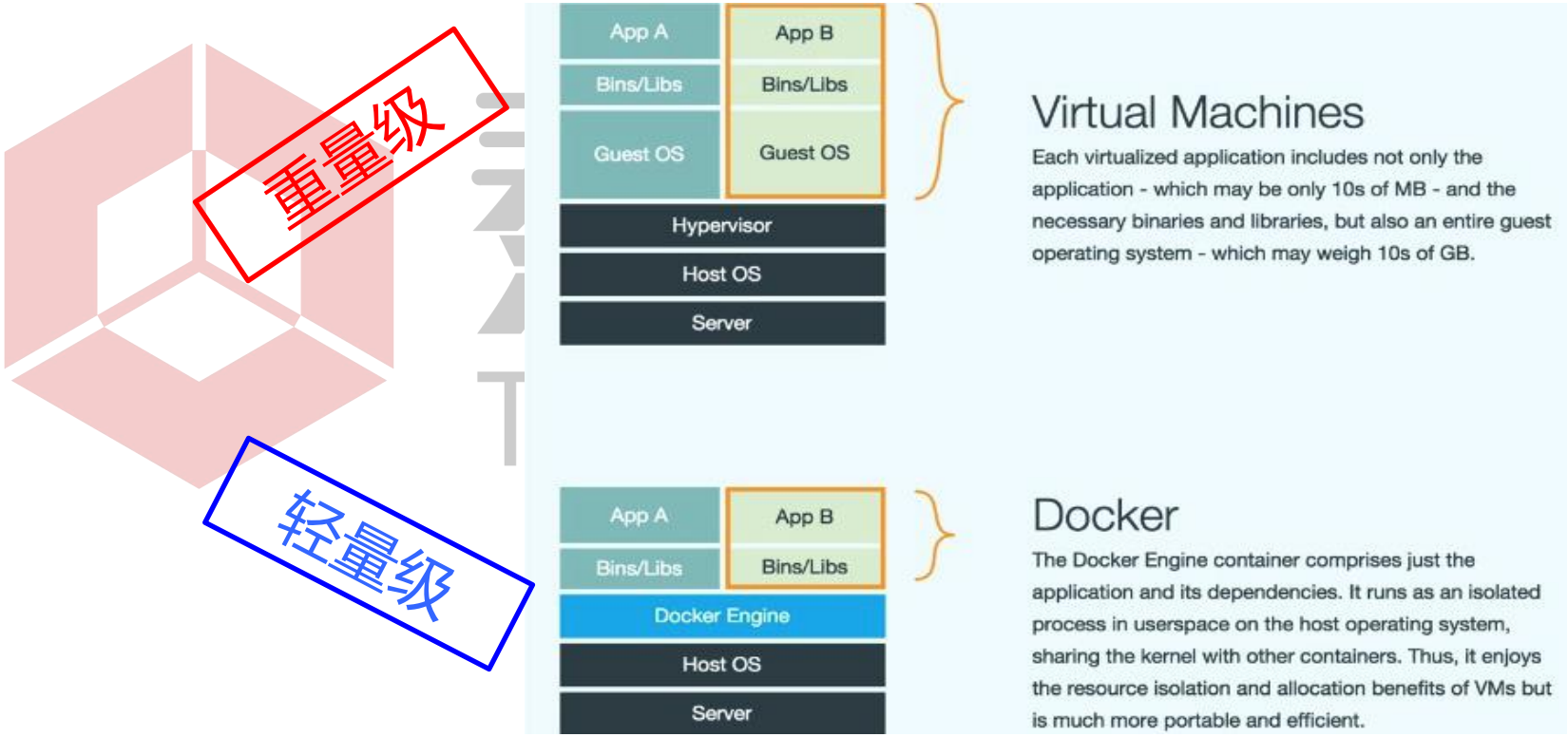
- 轻量级虚拟化
  - Vmware，KVM，XenServer都是重量级虚拟化技术。
  - 用户需要高效运行环境，而非整个机器。
  - 一次构建、到处运行。



版权所有© 2018 华为技术有限公司

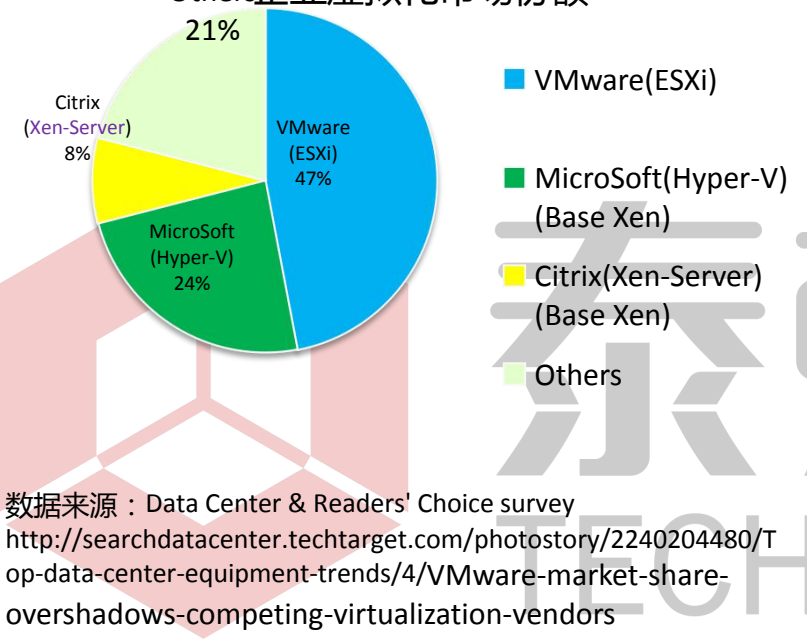
Docker容器虚拟化

- Docker容器是在操作系统层面上实现虚拟化，直接复用本地主机的操作系统，而传统方式则是在硬件层面实现。

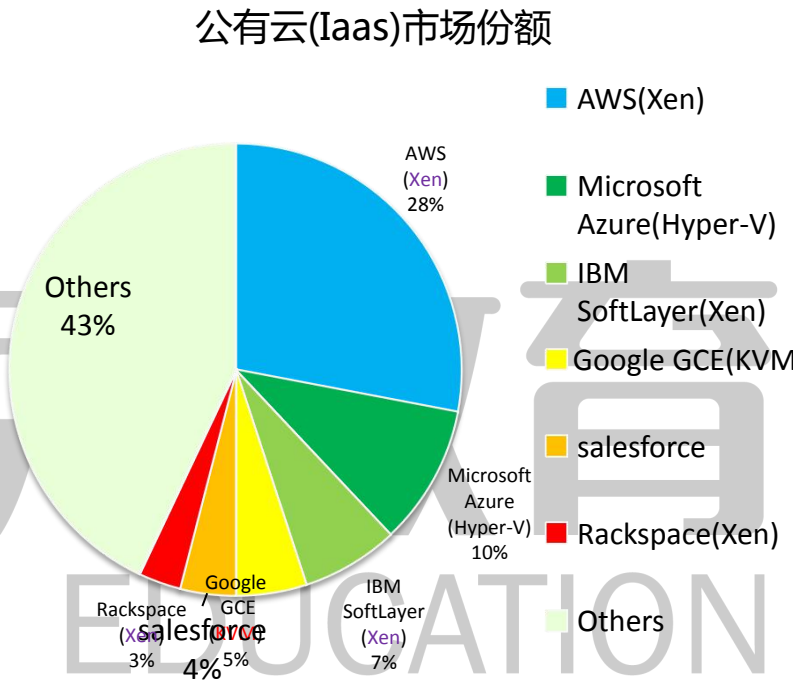


版权所有© 2018 华为技术有限公司

业界虚拟化技术分布



数据来源：Data Center & Readers' Choice survey
http://searchdatacenter.techtarget.com/photostory/2240204480/Top-data-center-equipment-trends/4/VMware-market-share-overshadows-competing-virtualization-vendors



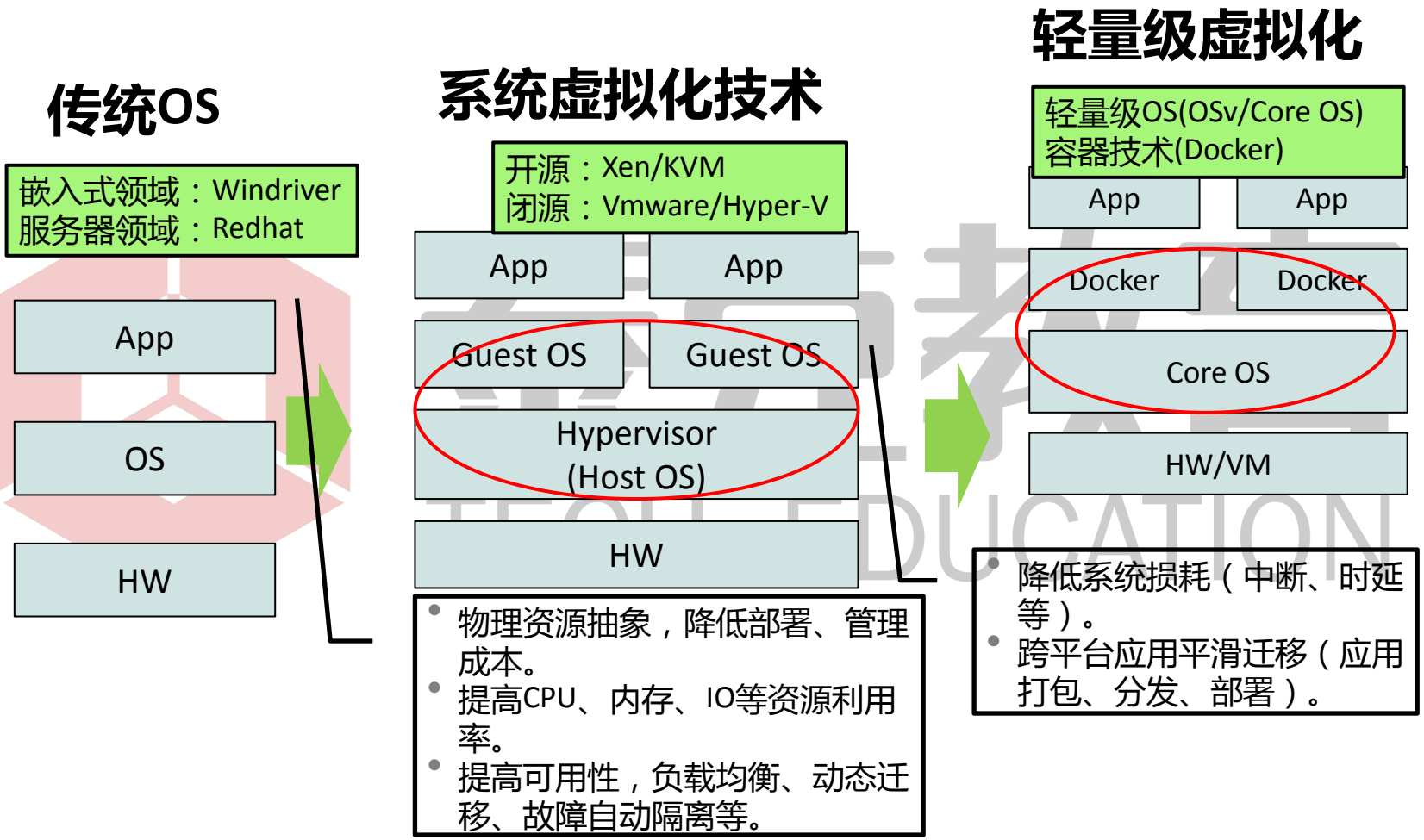
数据来源：Synergy Research
http://www.cloudcomputing-news.net/news/2015/feb/03/aws-hits-five-year-high-cloud-infrastructure-market-share/

从**企业私有云市场**来看，Xen家族（Xen-Server、Hyper-V、Oracle Xen）使用量仅次于VMware，约占40%。KVM商用案例较少。

在**公有云服务市场**公有来看，主流运营商均使用Xen家族平台，包括AWS、Azure、SoftLayer、阿里。

版权所有© 2018 华为技术有限公司

虚拟化技术未来衍生方向：基于轻量级OS + 虚拟化技术



版权所有© 2018 华为技术有限公司

容器与虚拟化

	容器技术	虚拟化技术
占用磁盘空间大小	小，甚至几十KB	非常大，上GB
启动速度	快，几秒钟	慢，几分钟
运行形态	直接运行在宿主机内核，不同容器共享同1个Linux内核	运行在Hypervisor上
并发性	一台宿主机可以运行成千上百个容器	单台机器最多几十台虚拟机
性能	接近于宿主机本地进程	逊于宿主机
资源利用率	高	低

泰克教育 TECH EDUCATION

版权所有© 2018 华为技术有限公司

HUAWEI

容器技术将会与Hypervisor共存，而不是替换

业务满足度	要解决的问题
安全不足	隔离性不如VM，还有**很多子系统**（audit、syslog、fuse、sysfs、time等）**没有namespace支持**。
	很多**权限问题**是通过privileged参数来解决，但是该参数会使得容器权限过大，存在安全风险。
	cgroup**资源隔离还有不完善**的地方，资源隔离不好的话，系统有受到DoS攻击的安全风险。（1）CPU:对CPU配额不够精确;（2）内存：不能限制kernel memory（比如文件系统的inode、dentry等内核数据结构）；（3）网络：只能做发包QoS，不能做收包QoS；（4）IO：只能限制direct-IO，不能限制buffered-IO；（4）存储：不支持存储空间配额；（5）其他：不能为容器单独设置进程上限、打开文件上限，等等。
功能不足	Docker native driver当前还**不支持多网卡配置和多网络平面**。
	容器目前**不支持热迁移**。
	为保证解耦，只能针对Linux应用并且应用不能有kernel改动，包括参数，**不要用自研内核模块**。
	为保证解耦，**不要访问PROC文件/SYS文件**，只应该访问自己镜象中的文件。
性能不足	通过 **Link拼接起来的Docker实例之间通讯效率不如进程间通讯**。通过Linux Bridge和Iptables实现NAT转换和网络隔离，性能堪忧（可能存才50%的损耗）。
扩展性不足	网络连接无法支持单Host的container上百到上千的数量级，数据中心的容器会达到上亿，**大规模容器之间的连接问题**。
成熟度不足	缺省支持的**aufs没有进内核**，且只有ubuntu提供，SLES和Redhat没有提供。
	Docker生态圈的发展还处于起步阶段，**支持企业级应用的工具和项目的还比较缺乏**，而且目前对**容器的资源管理和运维自动化处理尚无统一的或者标准的框架**。

版权所有© 2018 华为技术有限公司

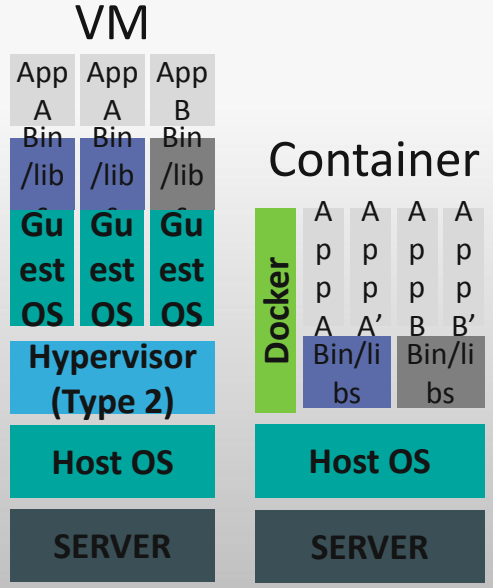
趋势1：容器技术给PaaS带来新的活力

- **容器** 是操作系统内核自带能力，容器是在Linux内核实现在**轻量级高性能资源隔离机制。**
- **Docker** 是容器技术之一，核心在于实现应用与运行环境**整体打包**以及打包**格式统一。**

Docker简化软件打包，加快应用部署及升级，优化DevOps 。

Docker可移植性，跨云部署，无厂商锁定，受企业电信支持。

Docker加速企业环境供应速度，实现资源快速弹性伸缩。



版权所有© 2018 华为技术有限公司

HUAWEI

趋势1(续)：蓬勃发展的容器生态系统加速PaaS的发展与实施

应用分发平台、配套工具	docker Docker Hub, Mar-2013 (initial release)	docker Docker Machine Feb-2015(initial release)		
容器编排	Apache AURORA (Commercial)	Kubernetes Jun-2014 (initial release)	docker Docker Compose Feb-2015 (initial release)	MARATHON
资源调度与集群管理	MESOS 2013 (Apache Project)	docker Docker Swarm Feb-2015(initial release)	CoreOS Aug-2013 (initial release)	etcd Aug-2013 (initial release)
容器执行引擎	docker Mar-2013 (initial release)	rkt Dec-2014 (initial release)	CLOUD FOUNDRY Warden Sep-2011 (initial release)	
容器 Host 操作系统专为运行容器而优化的新一代操作系统	CoreOS Aug-2013 (initial release)	RED HAT ENTERPRISE LINUX ATOMIC HOST Nov-2014 (initial release)	ubuntu Feb-2015 (initial release)	PHOTON vmware Apr-2015 (initial release)

版权所有© 2018 华为技术有限公司

HUAWEI

思考题

1. 以下说法正确的是？（ ）

   A. Xen平台架构侧重安全性

   B. KVM平台架构侧重性能

   C. KVM是在Linux操作系统标准内核中的一个虚拟化模块

   D. Xen直接运行于硬件之上

泰克教育
TECH EDUCATION

版权所有© 2018 华为技术有限公司

HUAWEI

本章总结

- 虚拟化优势
- Hypervisor的作用
- 容器的概念
- 容器和虚拟化的区别

泰克教育
TECH EDUCATION

版权所有© 2018 华为技术有限公司

HUAWEI

谢谢

www.huawei.com

泰克教育
TECH EDUCATION