

# 虚拟化技术调研报告

BigBear

2009/11/26

## 当前各虚拟化技术的分析

### ■ VMWare

#### EXSi

免费, Bare-Metal 的虚拟化的产品。

ESX Server 是一个单内核的产品,其主要特点是硬件的驱动程序集中在 Hypervisor 一层,被 Hypervisor 上的所有的虚机所共同使用。当一个虚机的 OS 需要访问硬件时,它通过 Hypervisor 中的 driver model 来访问,这种单内核的 Hypervisor 能够提供很好的性能,但是它在安全性和兼容性上存在缺陷。由于驱动程序和一些第三方代码跑在一个很敏感的区域,这种模式继承式了拥有了一个很大的被攻击面。

VMware ESXi 的详尽特性见附录 2。

#### vSphere

简单的说,把 ESXi 当做资源, vSphere 实现管理资源。

VMware vSphere 号称是业界第一款云操作系统,它利用虚拟化的强大能力将数据中心转变为极为简化的云计算基础架构,使 IT 组织能够利用内部和外部资源,安全和低风险地提供新一代灵活可靠的 IT 服务。

VMware vSphere 的关键组件:

基础架构服务组件,它们全面虚拟化服务器、存储设备和网络资源,聚合这些资源,并基于业务优先级将它们准确地按需分配给应用程序。

应用程序服务组件,无论应用程序类型或操作系统是什么,都将为运行于 VMware vSphere 上的所有应用程序提供内置的服务级别控制。

VMware vCenter, 提供基础架构和应用程序服务管理,实现日常操作任务自动化,并且可以深度监测大型和小型 VMware vSphere 环境的每个方面。

#### vCloud

基于 vSphere (并可跨越多个 vSphere 进行管理), 提供用户、计费、客户自服务等管理功能。

提供二次开发 API。

提供接口、自动化与管理的功能集,从而让企业和服务提供商以 Web 服务的形式来提供基于 VMware 平台的基础设施资源。

最终用户可以自助式的创建,使用和管理虚拟机和虚拟应用,而 IT 人员只需要管理用户的帐单并适当控制用户的使用。

### ■ Xen

## Xen Hypervisor

免费、开源的服务器虚拟化产品。

目前版本 3.4.2, 对 Intel nehalem 架构全面的支持。

下一个版本是 4.0

## XCP

免费的 Xen Cloud Platform, 基于 Xen 3.4.1, 采用 Linux 2.6.27 Kernel。

目前版本是 Version 0.1, 1.0 还没有发布, 预计在 2010 年的 Q1。

提供对虚拟化的基础设置的管理, 提供开放的 API, 提供租赁、SLA 保障和按需的细粒度计费模式。

将提供经微软认证的 Windows PV Drivers。

## XCI ( Xen Client Initiative )

*Xen Hypervisor for Client Devices*

嵌入式的 Xen Hypervisor。

主要是为了支持使用 Intel 或 AMD 体系机构的 Laptop 和 mobile devices。

Xen.org 桌面虚拟化方面的路线。

## ■ Citrix

### Xen Desktop

桌面虚拟化的解决方案, 提供像 windows 体验般的桌面系统。

*Any device, anytime, anywhere*

*HDX user experience - high definition experience on any device*

*On-demand apps*

*FlexCast delivery*

*Open architecture*

### Xen Server

免费, 服务器虚拟化产品, 基于 Xen。

Xen Server 和 VMware ESXi 的特性对比见附录 1。

## ■ Microsoft

## Hyper - V

微软提出的一种系统管理程序虚拟化技术。它主要作用就是管理、调度虚拟机的创建和运行，并提供硬件资源的虚拟化。

*Hyper-V 的体系结构见附录 3。*

### 虚拟化技术方面的发展重点

#### ■ CPU 和 MEM 的热插拔技术：

云计算的基础设施服务无非是提供：计算能力、网络带宽能力、存储能力。计算能力的动态扩展，也就是虚拟 CPU 和 MEM 的热插拔技术是其中最难的部分，这不仅涉及到硬件对热插拔的支持，还涉及到虚拟化技术对热插拔的支持，及 Guest OS 对热插拔的支持。

*目前，VMWare vSphere 支持 (1) Hot add CPU and memory (2) Hot plug virtual storage and network devices (3) Hot extend of virtual disks。*

*Xen 开源团队正在努力实现在 PVM 和 HVM 中的 CPU 和 MEM 热插拔技术，可能是 Hot add。*

*一旦具备了该功能，相应可以在我们的服务中真正实现虚拟机的动态扩容，实现真正的弹性，能在不停客户服务的情况下变更计算能力。*

#### ■ 软交换技术：

*VMWare 有自己的软交换技术，用来提供虚拟机之间的虚拟网络数据通信。目前 Xen 也在开发自己的软交换程序，但他目前是二层交换。*

*这些运行于 Domain0 中的软交换设备通过上联到一个真实交换机上实现一个级联的交换网络，可以提供 Vlan 划分，限速，流量控制等功能。*

#### ■ 动态迁移与异地动态迁移技术：

*动态迁移是提供 HA 的基础，而目前实现动态迁移的基本条件是底层的共享存储；正像 VMWare 那样共享存储是标配的，在次基础上才能支持 VMotion 等其它高级特性。*

*Xen Summit 2009 上，Xen 的主席 Ian Pratt (Ian Pratt is the chief architect of the Xen project, and chairman of Xen.org) 演示了一次跨数据中心的 Live Migration。*

*该特性对异地容灾很有意义。*

#### ■ 备份与容灾技术：

*数据安全性是云计算需要解决的最关键的任务。*

#### ■ IaaS API：

目前，虚拟化技术的两个主要发展方向是服务器虚拟化和桌面虚拟化，而桌面虚拟化是将整个 IT 带入云计算时代的关键，这意味这可以通过各种便携式设备访问云端的服务。

而这些服务必须具备统一的访问协议、API。

IaaS 需要为上层服务、第三方开发者提供统一的访问协议。

■ 与分布式计算的结合：

虚拟化技术是将计算能力进行“分”，使得每一个虚拟计算单元的能力小于支撑它的物理硬件。而目前稍微具备规模的应用其对计算能力的需求往往大于一个单独的物理硬件，这就需要分布式计算的支持，也就是“合”。“分”能带来很多好处，例如资源复用、方便的管理、HA 等，但如果我们一味的只是“分”而不对“分”以后的资源有效的“合”，那么我们就只能满足小客户的应用，不能满足大客户的应用。

现阶段，分布式计算比较主流的就是 Apache Hadoop，国内阿里巴巴正在做这方面的努力，我想我们的产品也应该关注这方面的趋势。

附录：

1、XenServer 和 VMware ESXi 的一个对比

Features included at no cost	Citrix XenServer	VMware ESXi
Windows® and Linux® guests	✓	✓
Unlimited servers, virtual machines, memory	✓	✓
P2V & V2V conversion	✓	✓
Shared SAN and NAS storage	✓	✓
Centralized multi-server management	✓	
Resilient distributed management architecture	✓	
Live motion	✓	
Shared virtual machine template library	✓	
Centralized configuration management	✓	
Virtual infrastructure patch management	✓	
Intelligent server maintenance mode	✓	

2、VMware ESXi 的详尽特性

关键特性

- 在性能方面创下纪录：每秒可执行多达 8,900 个数据库事务，每秒可完成 200,000 项 I/O 操作，每台物理主机上可承载多达 16,000 个 Exchange 邮箱
- 支持多达八路虚拟 SMP（对称多处理），因而可实现多处理器工作负载的虚拟化
- 可实现内存过量使用和重复数据消除，因而可提高整合率
- 可对各种管理程序提供最广泛的操作系统支持，使 IT 部门可以虚拟化众多版本的 Windows®、Linux®、Solaris®、NetWare® 及其他操作系统。
- 通过网卡绑定和 HBA 多路径功能来针对硬件组件故障提供保护，因此具备固有的高可用性。
- 每个主机具备多达 64 个逻辑处理核，256 个虚拟 CPU 和 1 TB RAM，从而可以获得更高的整合率

## 体系结构

- **64 位裸机管理程序体系结构。**使用经过生产验证、直接在服务器硬件上运行的管理程序技术，无需主机操作系统，即可实现接近于本机水平的虚拟机性能、可靠性和可扩展性。
- **虚拟磁盘文件。**使用虚拟机磁盘 (VMDK) 文件为虚拟机提供对自己专用数据存储区的访问，同时可以使 IT 管理员以功能完备、可驻留在共享存储设备上的独立文件形式灵活地创建、管理和迁移虚拟机存储。
- **VMware vStorage VMFS。**通过使用 [VMware vStorage Virtual Machine File System](#) (“VMFS”，一种允许多个 VMware ESX 主机同时访问单个 VMDK 文件的集群文件系统) 为虚拟机实施共享存储，可消除单点故障并平衡存储资源。vStorage VMFS 在光纤通道 SAN、iSCSI SAN 和 NAS 等多种不同的存储阵列上均受支持，这种支持对应用程序所有者和最终用户而言都是透明的。下载 [vStorage VMFS 产品介绍](#) 可了解有关 vStorage VMFS 的更多内容，其中介绍了一些新的增强功能，例如动态增加 vStorage VMFS 卷大小。
- **从 SAN 启动。**由于是在刀片服务器和机架安装式服务器的无盘配置上运行 VMware ESX 主机，因此无需单独备份本地连接的服务器磁盘。
- **虚拟网络连接。**借助 VMware ESX 和 ESXi 中的虚拟网络连接功能，客户可以在驻留在单个主机或跨多个 VMware ESX 及 ESXi 主机安装的虚拟机之间构建复杂网络，以进行生产部署或进行开发与测试。为每个虚拟机配置一个或多个虚拟网卡，且每个网卡都有其自己的 IP 地址和 MAC 地址，从而使虚拟机与物理机别无二致。可使用连接虚拟机的虚拟交换机在 VMware ESX 主机内创建一个模拟网络。使用虚拟局域网可在物理局域网的上层覆盖一个逻辑局域网来隔离网络流量，以实现安全性和负载分离。无需更改实际布线和交换机设置即可修改网络配置。

## 高级资源管理

VMware ESX 和 VMware ESXi 提供了一些高级资源管理功能，用以提高性能和整合率。

- **针对虚拟机的资源管理。**为虚拟机定义高级资源分配策略以提高软件应用程序的服务级别。为 CPU、内存、磁盘和网络带宽确定最小、最大和按比例的资源共享。在虚拟机运行的同时修改分配。
- **智能 CPU 虚拟化。**在物理机上所有可用的 CPU 之间采用智能进程调度和负载平衡，以此方式管理虚拟机进程的执行。
- **RAM 过量使用。**配置虚拟机内存以使其安全地超过物理服务器的内存量，以此方式提高内存利用率，使 VMware ESX 或 ESXi 主机上能够运行更多的虚拟机。
- **透明页共享（内存重复数据消除）。**通过将多台虚拟机中完全相同的内存页仅存储一次，更加高效地利用物理 RAM。
- **内存释放。**将 RAM 从空闲虚拟机动态转移到活动工作负载。内存释放在空闲虚拟机中制造内存紧张的假象，迫使它们使用其自己的分页区域，从而释放内存供活动虚拟机使用。

- **网络通信量调整。**确保关键虚拟机能够优先获得网络带宽。来自虚拟机的网络通信量可以根据“公平共享”机制划分优先级。Network Traffic Shaper 负责管理虚拟机网络通信以满足峰值带宽、平均带宽和突发事件所需带宽限制。
- **存储 I/O 通信量优先级划分。**通过根据“公平共享”机制划分 I/O 通信量的优先级来确保关键虚拟机能够优先访问存储设备。
- **改进的电源管理。**由于可动态地调整电压和频率并且支持 Intel SpeedStep® 和 AMD PowerNow!，因此可提高能效。

## 性能和可扩展性

VMware ESX 和 VMware ESXi 提供了无与伦比的性能和可扩展性，使资源占用量最大的生产应用程序也能实现虚拟化。

- **针对虚拟化工作负载的性能优化。**VMware ESX 和 ESXi 4.0 针对 Oracle 数据库、Microsoft SQL Server、Microsoft Exchange 等特定的业务关键型应用程序进行了性能优化。在性能方面达到如下水平：每秒可执行多达 8,900 个数据库事务，每秒可完成 200,000 项 I/O 操作，每台主机可承载多达 16,000 个 Exchange 邮箱。
- **针对 iSCSI 存储的性能改进。**通过结合运用客户操作系统内针对虚拟化进行了优化的新型 SCSI 驱动程序和 VMkernel 级存储堆栈优化，可显著提高需占用大量 I/O 资源的应用程序（如数据库和邮件应用程序）的性能。
- **支持强大的服务器硬件。**可利用在单台主机上具有多达 64 个物理 CPU 核心、256 个虚拟 CPU、1 TB RAM 以及多达数百个虚拟机的硬件系统来推动大规模整合和灾难恢复项目。
- **支持更大的虚拟机。**可配置具有多达 255 GB RAM 的虚拟机。
- **支持 8 路虚拟 SMP。**VMware 虚拟对称多处理 (SMP) 功能通过让单个虚拟机最多同时使用 8 个物理处理器，增强了虚拟机的性能。VMware Virtual SMP 可使 CPU 资源占用量最大的企业级应用程序（如数据库、ERP 和 CRM）实现虚拟化。
- **原始设备映射。**为了支持应用程序集群和基于阵列的快照技术，同时利用 VMware vStorage VMFS 的可管理性优势，也可以选择将 SAN LUN 直接映射到虚拟机。
- **支持硬件虚拟化。**VMware ESX 和 ESXi 为下一代虚拟化硬件辅助技术（如 AMD 的快速虚拟化索引 (Rapid Virtualization Indexing®) 或 Intel 的扩展页表）提供了业界领先的支持。
- **支持大容量内存页。**只有 VMware ESX 和 ESXi 这两个管理程序支持大容量内存页，以提高客户操作系统的内存访问效益。
- **网络连接性能优化。**VMware ESX 和 ESXi 支持多种性能卸载技术，包括 TCP 分段卸载 (TSO)、VLAN 和校验和卸载，以及用于减少与网络 I/O 处理关联的 CPU 开销的巨型帧。此外，还支持针对虚拟化进行优化的 I/O 性能提升功能，NetQueue 就属于这种功能，它在 10 千兆以太网虚拟化环境中可大大提升性能。
- **支持新型高性能设备和协议。**VMware ESX 和 ESXi 支持 10 千兆以太网网卡和存储阵列以及 Infiniband 技术，以提高虚拟机性能。
- **支持半虚拟化。**VMware ESX 和 ESXi 支持半虚拟化的 Linux 客户操作系统（Linux 内核 2.6.21 以上），以提升虚拟机性能。
- **针对虚拟机的 VMDirectPath I/O 技术。**通过允许所选虚拟机直接访问底层硬件设备，提高了需要频繁访问 I/O 设备的应用程序的 CPU 效率。虚拟机使用此功能时，将无法使用 VMware VMotion™ 和物理 I/O 设备共享等其他虚拟化功能，亦不具备硬件独立性。

## 高可用性

VMware ESX 和 ESXi 为虚拟机提供了数据中心级的高可用性。

- **内置的存储访问多路径功能。**通过将 SAN 多路径功能用于光纤通道或 iSCSI SAN，确保共享存储的可用性。

- **网卡绑定。**为网络中的每台虚拟机提供内置的网卡故障切换和负载均衡功能，从而实现更高的硬件可用性和容错能力。新的网卡绑定策略允许用户配置多个活动和备用适配器。
- **支持 Microsoft 集群服务。**可将跨物理主机、运行 Microsoft Windows 操作系统的多个虚拟机组成集群。

## 互操作性

VMware ESX 和 VMware ESXi 跨 IT 的所有层次进行了优化并经过了严格测试和认证，可用于实现企业范围的标准化。

- **服务器硬件。**VMware ESX 和 ESXi 经认证可用于 Dell、Fujitsu Siemens、HP、IBM、NEC、Sun Microsystems 和 Unisys 生产的业界领先的机架安装式、塔式和刀片式服务器。
- **存储硬件。**VMware ESX 和 ESXi 经认证可用于 Dell、EMC、Fujitsu、Fujitsu Siemens、HP、Hitachi Data Systems、IBM、NEC、Network Appliance、StorageTek、Sun Microsystems 和 3PAR 出品的众多存储系统。它们支持内部 SATA 驱动器、直连存储 (DAS)、网络连接存储 (NAS) 以及光纤通道 SAN 和 iSCSI SAN。
- **操作系统。**VMware ESX 和 ESXi 支持最广泛的未修改操作系统，包括 Windows、Linux、Solaris、Novell NetWare 等。VMware 最近又增加了对 20 种新型客户操作系统的支持。
- **软件应用程序。**任何软件应用程序无需修改都可以在 VMware 虚拟机中运行。
- **虚拟机格式。**VMware ESX 和 ESXi 可以运行以非 VMware 格式创建的虚拟机。使用免费的 VMware vCenter Converter，用户可以转换 Microsoft Virtual Server、Microsoft Virtual PC 以及 Symantec LiveState Recovery 虚拟机，使它们能够在 VMware ESX 和 ESXi 主机上运行。

## 安全性

VMware ESX 和 ESXi 中的高级安全功能可保护虚拟环境中存储的数据。

- **VMware VMsafe™。**VMware VMsafe 是一项新型安全技术，可帮助以原先使用物理机时不可能实现的方式来保护虚拟化的工作负载。VMsafe 提供了一组安全性 API，借助这些 API，第三方安全产品能够像 VMware ESX 或 ESXi 那样清晰地洞察虚拟机的运行情况，以便发现并消除恶意软件，如病毒、特洛伊木马和击键记录程序。这种高级保护功能是通过精确洞察虚拟机的硬件资源（如内存、CPU、磁盘以及其 I/O 系统）状况实现的。
- **VMkernel 保护。**通过确保 VMkernel（此管理程序的一个核心组件）的完整性来保护 VMware ESX 和 ESXi 免遭常见的攻击和漏洞利用。ESX 和 ESXi 中的磁盘完整性技术利用可信平台模块（TPM，服务器中嵌入的一种硬件设备）来保护此管理程序的启动。加载到磁盘和内存中的 VMkernel 模块在加载期间会进行数字签名和验证，以确保动态加载的代码的真实性和完整性，并防范恶意软件在 VMkernel 持久保存到磁盘上时企图对它进行修改。VMkernel 在加载时也会结合使用内存完整性技术和微处理器功能，以保护其自身免遭企图利用正在运行的代码的常见缓冲区溢出攻击。
- **加密：**使用 SSL 加密确保安全连接到 VMware ESX 和 ESXi 主机。
- **针对 iSCSI 设备启用身份验证：**VMware ESX 和 ESXi 要求，只要主机企图访问目标 LUN 上的数据，主机或 iSCSI 启动器就必须经过 iSCSI 设备或目标的身份验证，以此方式来保护 iSCSI 设备免遭不利侵袭。
- **网络安全策略。**在以太网层强制执行虚拟机安全策略。不允许使用杂乱模式进行网络通信嗅探，不允许更改 MAC 地址和假冒源发送 MAC。

## 可管理性

提供了多个管理界面，以便更加高效地管理 VMware ESX 和 ESXi 环境。VMware ESX 和 ESXi 管理员使用的核心管理界面包括：

- **VMware vSphere Client。** 可使用 VMware vSphere Client 的通用用户界面管理 VMware ESX 或 ESXi 主机、虚拟机以及（可选）VMware vCenter Server。vSphere Client 以免费下载的形式提供，可以针对 VMware ESX 或 ESXi 主机使用以便管理单台主机，也可以针对 VMware vCenter Server 使用以便管理多台主机。
- **VMware vCenter Server。** 实现了对 VMware ESX 和 ESXi 主机及其虚拟机的集中管理。为使用 VMware vCenter Server 管理 ESX 或 ESXi 主机，相应主机需要具备 VMware vCenter Agent 许可证（所有版本的 VMware vSphere 中均包含此许可证）。VMware vSphere 还包含可以提高业务连续性并使运营效益最大化的许多其他管理功能，例如实时迁移、自动负载均衡、硬件故障防护，以及虚拟机备份和恢复功能。

VMware ESX 和 ESXi 的其他管理工具包括：

- **VMware vSphere Command-Line Interface 4.0 (vCLI)。** 通过远程执行环境管理 VMware ESX 和 ESXi。最新版本的 vCLI 新增了许多命令，在 VMware ESX 4.0 和 VMware ESXi 4.0 上均受支持。查看《[vSphere Command-Line Interface 安装和参考指南](#)》。
- **VMware vSphere Power Command-Line Interface 4.0 (PowerCLI)。** 使用这款基于 Microsoft PowerShell 技术、功能强大但易于使用的界面可管理和配置数以千计的虚拟机。PowerCLI 允许 IT 管理员通过一个脚本编写界面来管理 VMware ESX 或 ESXi，可管理使用 VMware vSphere Client 时执行的相同任务。
- **VMware vSphere Management Assistant。** VMware vSphere Management Assistant 是一种虚拟机，它包含 VMware vSphere 命令行界面及其他的预打包软件，开发人员和管理员可以使用这些软件来运行代理和脚本，以管理 VMware ESX 和 ESXi 主机。
- **使用 CIM 进行无代理硬件管理。** 通用信息模型 (CIM) 提供了一种用于通过 VMware vCenter Server 或与 CIM 兼容的第三方工具监控硬件运行情况和状态的协议。

### 3、微软 Hyper - V 架构

