
AMD-VMWare 虚拟化解决方案

1 企业IT的虚拟化趋势

IT 基础架构规模的日趋庞大，越来越高的复杂性，以及无法避免的异构性，向企业 IT 部门提出了更高的要求。IT 经理或者企业的 CIO 需要在考虑如何应对不断出现的问题的同时，也考虑到未来企业 IT 的运营模式。比如：如何使规模庞大的 IT 系统运行的更有效率；如何从 IT 的角度增加企业的竞争力。

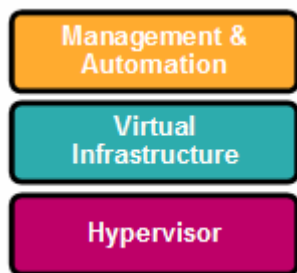
基于企业 IT 运营的需求，一种新的技术应运而生，这就是虚拟化技术。虚拟化技术为企业 IT 运营模式提供了新的解决方案。

- 一、增强业务灵活性
- 二、增强可管理性
- 三、增强业务连续性
- 四、减少整体拥有成本（TCO）
- 五、投资保护

应该说，虚拟化技术是一个新的工具，它将帮助企业从粗放的 IT 管理向精细化管理跃迁，对虚拟化模式下的 IT 管理的精细挖掘，以及有效使用，将极大提升企业的竞争力。

2 VMWare虚拟化技术介绍

虚拟化技术在发展过程中形成了如下几个层次。



Hypervisor 是运行于单机的虚拟化软件，它能够从指令集层面集成虚拟化功能，在一台服务器上运行多个操作系统，为多个操作系统分配资源并且进行管理。

Virtual Infrastructure 是虚拟化平台，是在多个物理服务器的基础上建立的统一的虚拟化平台，虚拟机可以方便的在不同的硬件平台上迁移，而不影响应用的运行。VI 使企业的 IT 设备从孤立的一群设备变成一个统一的平台，或者设备池，为采用更先进的管理方式提供了可能。

Management & Automation 是智能化的虚拟化管理。基于 VI 建立的平台，VMWare 或者 Xen 等虚拟化管理软件，都或多或少的集成了先进的管理方式，能够根据设备的符合，自动分配资源，并进而提升设备利用率、可靠性，并达到节省功耗的目的。

3 AMD-V™虚拟化技术特点

基于 x86 架构的虚拟化技术，出现在上世纪 90 年代，然而由于企业级应用要求具有良好的安全性，以及不错的性能，而传统的纯软件的虚拟化解决方案，难以具备企业级应用的要求。

AMD-V™技术在硬件级别集成了一系列独有的虚拟化技术，使虚拟化具备了高性能和高安全性。下面从性能和安全性两个方面，介绍 AMD 虚拟化技术的一些主要特性。

3.1 高性能的虚拟化技术

AMD-V™技术在性能优化方面主要包括如下一些技术。

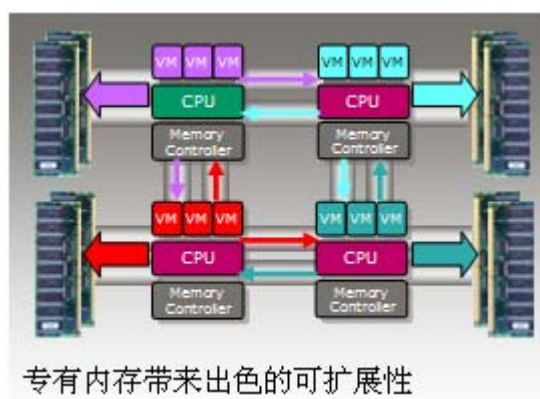
-
- ✓ AMD-V™虚拟化指令集
 - ✓ 直连架构 (DCA)
 - ✓ Tagged TLB
 - ✓ 快速虚拟化索引 (RVI)

3.1.1 AMD-V™虚拟化指令集

AMD 在 CPU 中增加了新的虚拟化指令 ,这些指令使 Hypervisor 能够高效的启动和管理虚拟机 , 以及进行资源的分配。

3.1.2 直连架构

直连架构使 AMD 的 CPU 相对传统的 FSB 架构获得了更高的内存访问带宽和更低的访问时延 ,而由于同时在一台服务器上需要运行多个操作系统 ,因此虚拟化技术本身对内存的性能要求比较高 , 这样直连架构就良好的适应了虚拟化对硬件平台的要求。



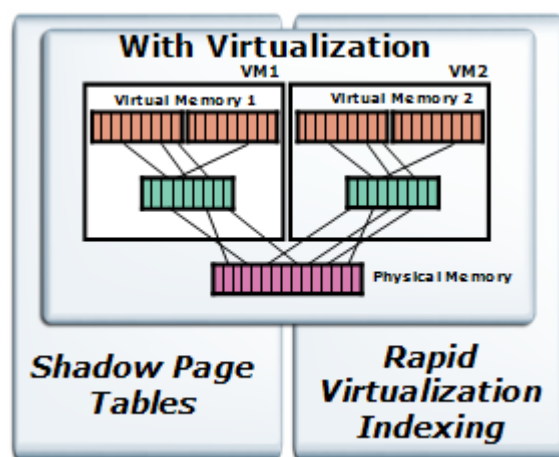
在直连架构的平台上 ,通过合理的分配 CPU 和内存资源 (例如上图) ,虚拟机将获得更高效的运行效率。

3.1.3 快速虚拟化索引RVI

快速虚拟化索引(Rapid Virtualization Index ,RVI)也是 AMD 所独有的一种虚拟化技术。

在采用虚拟化技术后，应用程序的内存访问模式将变的更加复杂（如下图），原先操作系统的物理地址，将需要通过 Hypervisor 再进行一次转换，才能映射到实际的物理地址上。

传统的模式是采用 Shadow Page Tables，既由 Hypervisor 来管理内存映射关系，每一次内存访问都需要 Hypervisor 进行运算才能访问到实际的物理地址。而 RVI 技术，就是将这种转换直接通过嵌套 TLB 的方式实现，既通过硬件直接实现转换，极大的提升了虚拟机的工作效率。

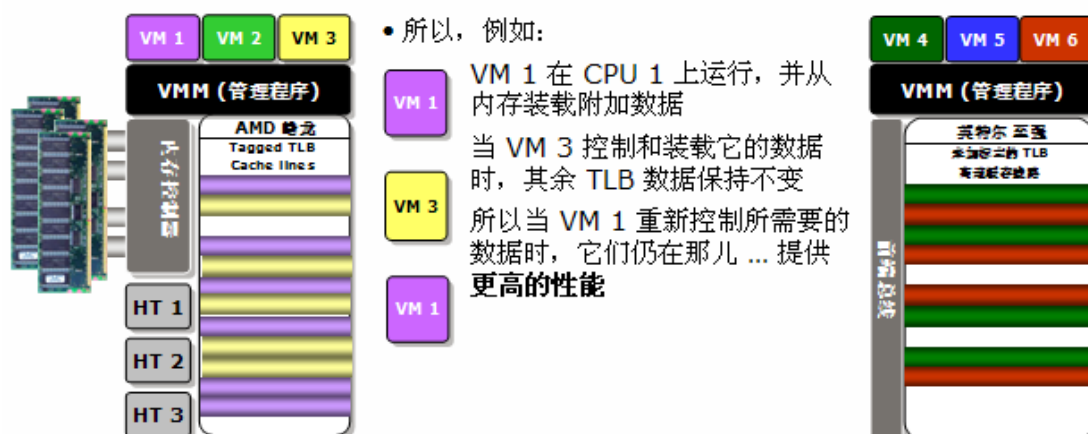


通过测试，使用 RVI 技术后，最多可以减少 75%的 Hypervisor 工作时间，虚拟机切换（world-switch）时间能够减少 25%。

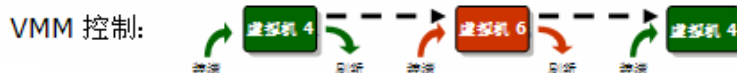
3.1.4 带标志的（Tagged）TLB

Tagged TLB 是又一种 AMD 的创新技术。TLB 是 CPU 中的表格，包含内存页面的虚拟和现实地址间的相互参照。传统上虚拟机之间转换时需要清洗 TLB。与许多 RISC 架构相似，AMD 虚拟化具有带标志的 TLB，地址空间可以在不同虚拟机之间转换，包括 Hypervisor 和虚拟机之间，而无需刷新 TLB。

- “加标记”意味着数据了解它属于哪台虚拟计算机



- 至强 在每次转换时都必须刷新和重新装载TLB，降低了性能



AMD 带标记的 TLB 相比 Intel 不带标记的 TLB 能够提供更出色的内存访问和虚拟机切换性能。

3.2 高安全性的虚拟化技术

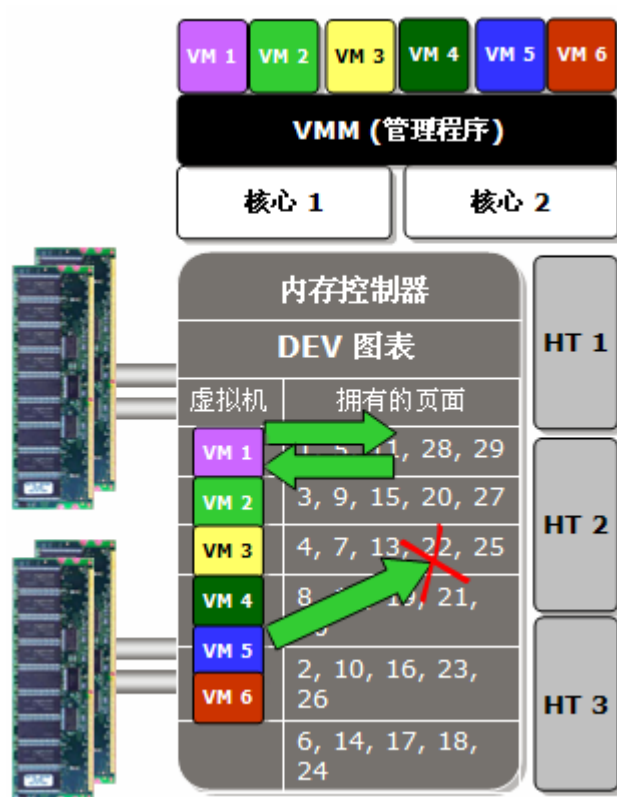
对于企业级应用，高性能固然重要，然而高安全性才是其必要保证，AMD-V™技术在高安全性方面采用了如下一些技术，确保企业级应用的安全。

- ✓ Secure VMM (Secure Init)
- ✓ Secure Memory boundary (嵌套页表)
- ✓ Secure DMA (DEV)
- ✓ Secure SMM

在 Hypervisor 启动 (boot sequence) 时，AMD 采用了 Secure Init 技术，来确保在 Hypervisor 启动过程的安全性。在 Hypervisor 运行过程中，仍然是有可能收到攻击的，AMD 采用了 Secure SMM (系统管理模式) 来保证运行过程中的安全性。

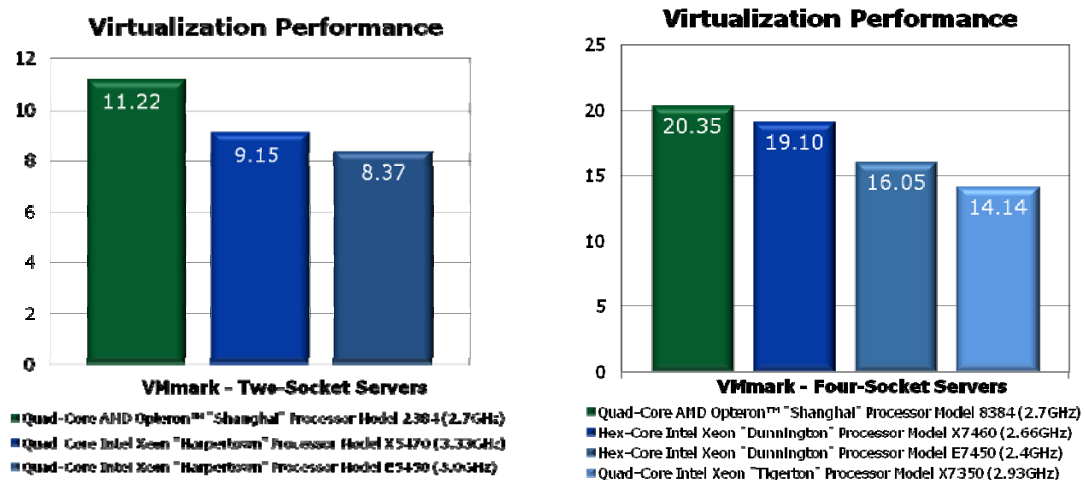
在内存访问的安全性方面，首先 RVI 技术，从硬件角度进行内存地址的映射，提高了内存访问的安全性。此外，AMD 还采用了集成到硬件中的设备排除列表 DEV 技术，使基于 AMD-V™的虚拟化可以以最高的安全性和可靠性运行。

DEV 使 Hypervisor 能够了解虚拟机是否被允许访问到的页面，所以 DEV 通过拒绝未经授权的内存访问而增强了虚拟化的安全性。如下图所示，VM1 被允许访问第 28 个页面，而第 22 页面属于 VM3，则不允许 VM5 进行访问。



DEV 技术是虚拟化必须具备的技术，所不同的是，在 AMD 平台中，DEV 是基于硬件的，因而可以获得更好的安全性。

3.3 “上海” 内核提供了No.1 的虚拟化性能



在 VMWare 的 VMMark 虚拟化基准测试对比中，“上海”内核的皓龙处理器，无论在 2 路还是 4 路服务器领域，都获得了 No.1 的虚拟化性能。如上图。

其中 2 路虚拟化性能比至强最高分值的高功耗产品 X5470 高出 22.6%，比标准功耗的至强 E5450 高 34%。

在 4 路服务器虚拟化性能对比中，“上海”内核的皓龙处理器，比高功耗的至强 6 核心产品 X7460 高 6.5%，每瓦特性能优势达到 84%。比标准功耗的至强 6 核心产品，性能高出 27%。

4 典型的配置方案

考虑到为虚拟化平台预留足够的扩展性和性能的冗余空间，虚拟化的适合平台主要集中在 4 路到 8 路的服务器；在小型企业的运作中，考虑到成本，也可以采用双路的服务器。基于 AMD 在银行、保险、电信、制造业领域的成功实施，我们给出典型的虚拟化配置方案。

4.1 大中型企业虚拟化配置方案

对于大中型企业的虚拟化配置方案，根据平台规模和成本，可以选择采用多台 4 路或者 8 路的服务器，一方面可以作为热备提高可靠性，另一方面增加扩展性和业务部署的灵活性。

下面给出高扩展性、普通应用的参考配置。

4.1.1 高扩展性参考配置

		指标项	技术规格要求
1	外观	服务器外观	机架式（4U 到 7U）
2	处理器	处理器	四核 AMD64 Opteron 8384
3		处理器配置数目	8 个
4	总线	系统总线	HyperTransport 技术
5	内存	内存类型	DDR2
6		内存配置数目	64GB（扩展能力≥128GB）
7	带宽	内存带宽	≥85GB
8		系统总带宽	≥170GB
9	磁盘	内置硬盘类型	≥10000rpm 热插拔 SAS 硬盘
10		内置硬盘容量及数目	≥4 块，单盘容量≥146GB
11		阵列卡	配置 SAS 阵列控制器，带电池缓存≥256MB；支持 RAID 0、1、10、5
12	IO 接口	外部接口	串口≥1，USB 2.0 接口≥2（前置和后置各不少于 1 个）

13	网络	连接端口	≥2 个千兆网口
14	I/O 扩展	PCI I/O 插槽	≥6 个 PCI-E 插槽
15	可靠性	冗余组件	热插拔冗余电源，冗余风扇

4.1.2 一般应用参考配置

		指标项	技术规格要求
1	外观	服务器外观	机架式 (2U 到 4U)
2	处理器	处理器	四核 AMD64 Opteron 8380
3		处理器配置数目	4 个
4	总线	系统总线	HyperTransport 技术
5	内存	内存类型	DDR2
6		内存配置数目	32GB (扩展能力≥64GB)
7	带宽	内存带宽	≥42GB
8		总带宽	≥90GB
9	磁盘	内置硬盘类型	≥10000rpm 热插拔 SAS 硬盘
10		内置硬盘容量及数目	≥4 块，单盘容量≥146GB
11		阵列卡	配置 SAS 阵列控制器 ,带电池缓存≥256MB ;支持 RAID 0、1、10、5
12	IO 接口	外部接口	串口≥1，USB 2.0 接口≥2 (前置和后置各不少于 1 个)
13	网络	连接端口	≥2 个千兆网口
14	I/O 扩展	PCI I/O 插槽	≥4 个 PCI-E 插槽

15	可靠性	冗余组件	热插拔冗余电源，冗余风扇
----	-----	------	--------------

4.2 中小型企业虚拟化配置方案

对于中小型企业的虚拟化配置方案，依然根据应用需求和成本，可以选择采用多台 4 路或者 2 路的服务器。下面给出标准和低端的参考配置。

4.2.1 标准参考配置

		指标项	技术规格要求
1	外观	服务器外观	机架式（2U 到 4U）
2	处理器	处理器	四核 AMD64 Opteron 8378
3		处理器配置数目	4 个
4	总线	系统总线	HyperTransport 技术
5	内存	内存类型	DDR2
6		内存配置数目	32GB（扩展能力≥64GB）
7	带宽	内存带宽	≥42GB
8		总带宽	≥90GB
9	磁盘	内置硬盘类型	≥10000rpm 热插拔 SAS 硬盘
10		内置硬盘容量及数目	≥4 块，单盘容量≥146GB
11		阵列卡	配置 SAS 阵列控制器，带电池缓存≥256MB；支持 RAID 0、1、10、5

12	IO 接口	外部接口	串口≥1，USB 2.0 接口≥2（前置和后置各不少于 1 个）
13	网络	连接端口	≥2 个千兆网口
14	I/O 扩展	PCI I/O 插槽	≥4 个 PCI-E 插槽
15	可靠性	冗余组件	热插拔冗余电源，冗余风扇

4.2.2 低端参考配置

		指标项	技术规格要求
1	外观	服务器外观	机架式（2U），塔式
2	处理器	处理器	四核 AMD64 Opteron 2384
3		处理器配置数目	2 个
4	总线	系统总线	HyperTransport 技术
5	内存	内存类型	DDR2
6		内存配置数目	16GB（扩展能力≥64GB）
7	带宽	内存带宽	≥21GB
8		总带宽	≥45GB
9	磁盘	内置硬盘类型	≥10000rpm 热插拔 SAS 硬盘
10		内置硬盘容量及数目	≥2 块，单盘容量≥146GB
11		阵列卡	配置 SAS 阵列控制器，带电池缓存≥256MB；支持 RAID 0、1、10、5
12	IO 接口	外部接口	串口≥1，USB 2.0 接口≥2（前置和后置各不少于 1 个）
13	网络	连接端口	≥2 个千兆网口

14	I/O 扩展	PCI I/O 插槽	≥4 个 PCI-E 插槽
15	可靠性	冗余组件	热插拔冗余电源，冗余风扇