

企业级服务器虚拟化性能评估方法^①

金俊才, 许超超

(浙江省烟草公司台州市公司 信息中心, 台州 318000)

摘要: 随着计算机硬件性能及虚拟化技术的快速发展, 越来越多的企业通过采用服务器虚拟化技术来降低信息中心机房的整合成本, 提高硬件利用率. 然而, 在如何选择最佳的服务器硬件搭建高性能、低成本的虚拟化平台, 尤其是在虚拟化应用层面上的性能评估是目前亟需解决的问题之一. 介绍了四种服务器虚拟化的性能评估方法, 分析比较了不同评估方法的特点和局限及其适用环境, 为企业信息中心机房如何更好地去部署实施服务器虚拟化提供建设思路.

关键词: 虚拟化; 虚拟服务器; 性能评估

Performance Evaluation of the Enterprise-level Virtualization

JIN Jun-Cai, XU Chao-Chao

(Taizhou City Branch, Zhejiang Provincial Tobacco Corporation, Taizhou 318000, China)

Abstract: With the rapid development of computer hardware performance and virtualization technology, server virtualization technology is being used increasingly to reduce the integration cost of information center room or improve hardware utilization. However, how to choose the best server hardware to build high performance and low cost virtualization platform, especially on performance evaluation is one of the urgent problems. This paper briefly introduces four kinds evaluation methods of virtual server. Through the analysis and comparison between different methods, it assess the advantages or disadvantages of evaluation methods and its application environment. Based on the analysis above, it puts forward reference suggestions for better deploying server virtualization system.

Key words: virtualization; virtual server; performance evaluation

虚拟化技术是指在一台高性能的物理服务器上或一套硬件资源上安装虚拟化底层软件, 然后从中虚拟出一个或若干个虚拟机, 不同的服务独立运行在不同的虚拟机上. 作为一种全新的商业服务模式, 虚拟化技术在信息处理和使用领域展现出巨大优势. 它不但为企业节约大量的采购成本和人力成本, 还降低了系统管理成本及新系统开发部署的成本, 同时增强了业务系统的灵活性^[1]. 但是虚拟化技术的发展离不开对虚拟系统性能的评估. 通过评估虚拟系统性能可以查找系统规划、部署问题和资源瓶颈、资源浪费等问题. 另外, 通过评估还可以量化系统服务质量, 指导用户采购, 降低采购风险. 因此, 无论是大型企业数据中

心服务器整合还是中小型企业经济型服务器选型, 虚拟系统性能评估都在其中扮演着越来越重要的角色.

1 研究背景与目的

1.1 虚拟化技术发展

目前, 很多公司提供了比较成熟的虚拟化解决方案, 如 Citrix 公司的 Xen, IBM 公司的 PowerVM、zVM, Microsoft 公司的 Virtual PC、Virtual Server 和 Hyper-V 以及 VMware 公司的 VMware Server、VMware Workstation 等^[2,3].

1.2 虚拟系统性能评估要素

目前, 对于虚拟服务器的性能测量研究工作主要

^① 收稿时间:2014-11-11;收到修改稿时间:2014-12-08

有以下几方面^[3-6]:

虚拟机管理器性能: 虚拟机管理器(VMM)是所有虚拟机的管理和调度中心,具有创建、销毁、启动、重启、关闭、查看、修改、暂停、恢复虚拟机等功能。可以通过在虚拟机中运行某些基准测试程序来比较 Xen、KVM、Vmware 等不同虚拟机管理器之间的性能开销^[7,8]。

可伸缩性能: 可伸缩性(Scalability)主要是指虚拟计算系统能够支持的虚拟机数目及可以运行的应用数目,该指标反映了虚拟系统的规模和服务的可扩展能力^[9]。

整合性能: 服务器整合(Server Consolidation)是指通过虚拟机管理器将若干应用服务器同时运行在同一台宿主机上^[10,11]。

迁移性能: 迁移是指将虚拟机从一台物理机迁移到另一台物理机上。目前,大多数研究是针对迁移策略,比如预拷贝技术、后拷贝技术等^[12,13]。

隔离性能: 隔离性是指虚拟机之间相互影响的程度。虚拟机的设计理念是虚拟机之间完全隔离,相互之间的运行不会影响各自性能。文献[14]设计了简单的基准测试程序揭示虚拟机管理器的隔离性能。文献[15]研究了多核环境中虚拟化的隔离情况,并给出不同虚拟机的搭配组合策略。

1.3 传统虚拟化规划方式

企业在利用虚拟化技术整合信息中心机房服务器过程中,传统方式采用阶段式规划,即 IT 人员凭借主观经验和定性分析,先将负载较轻或者重要性不高的服务器阶段性地迁移至虚拟机中,再逐一迁移剩下的其它服务器,直到整合所有服务器。如表 1 所示,传统的规划方式适合物理服务器数量不多或时间比较宽裕的情况下部署实施。

表 1 传统规划方式

规划方式	特点	局限
阶段规划 阶段实施	易操作、不会承受太大的迁移压力	整个过程主要靠人为主观判断,可能还需要技术人员反复地测试及评估虚拟系统的性能,时间投入成本较高。

1.4 传统虚拟系统性能评测工具

传统的服务器性能评测工具有很多,例如 Lmbench 是一套简易可移植、符合 ANSI/C 标准的针对 UNIX/POSIX 的微型测评工具,主要衡量系统的反应时间和带宽这两个关键特征^[16]。UnixBench 是一个

类 unix 系(Unix, BSD, Linux)统下的性能测试工具,被广泛用于测试 linux 系统主机的性能,主要测试项目有:系统调用、读写、进程、图形化测试、2D、3D、管道、运算、C 库等系统基准性能提供测试数据^[17-19]。Netperf 是一种网络性能及网速带宽的测量工具,主要针对基于 TCP 或 UDP 的传输^[20]。IOZone(www.iozone.org)是一个文件系统的 benchmark 工具,可以测试不同操作系统中的文件系统等^[21]。

但是,传统的基准测试仅侧重于对一台服务器运行单一工作量的检验,而无法获取支持多个虚拟系统状况下的性能或者是在一台物理服务器上同时支持多个工作量时的服务器性能。

1.5 本文研究目的和意义

本文描述了 4 种多负载服务器性能评估方法的基本原理和思路,分别是基于 Capacity Planner 容量规划法、基于 VConsolidate 基准评估法、基于 VMmark 基准评估法、基于 SPEC 的基准评估法。分析比较了这四种评估方法的特点、局限和适用范围,为企业在建设服务器虚拟化过程中提供建设性思路。

2 基于Capacity Planner容量规划法

2.1 VMware Capacity Planner 简介

VMware Capacity Planner 是由 VMware 推出一款虚拟化评估工具,通过对一定数量的当前应用系统资源进行评估,如处理器、内存、磁盘读写、网络等性能,评估这些应用系统移植到指定硬件的 VMware 虚拟平台后该如何分布及后续的运行情况。从而,为客户对当前应用环境存在的潜在虚拟化可能性建立一个初步印象,帮助企业合理规划、部署虚拟服务器,充分挖掘潜在的可利用 IT 资源,提高效率的同时降低成本^[22]。

2.2 容量规划方法概述

容量规划的核心思想是通过迅速收集基础数据、应用数据以及系统性能数据,预测数据中心的生长、评估服务器工作负载和应用性能目标,为最终虚拟服务器架构提供数据依据,该方法主要分为以下几个过程^[23,24]:

(1) 安装 VMware Capacity Planner,对现有在线应用系统进行监控,收集可利用的性能指标,监控时间为系统的一个性能周期(一周/月);

(2) 将收集到的性能参数上传到 VMware 容量规

划站点, 由 VMware 官方数据库进行分析;

(3) 指定 VMware 虚拟化后所用的硬件, 针对特定硬件和配置给出具体的容量规划报告。

容量规划的架构如图 1 所示, 图中给出了容量规划系统各个功能组件的接口连接和数据流动方向。

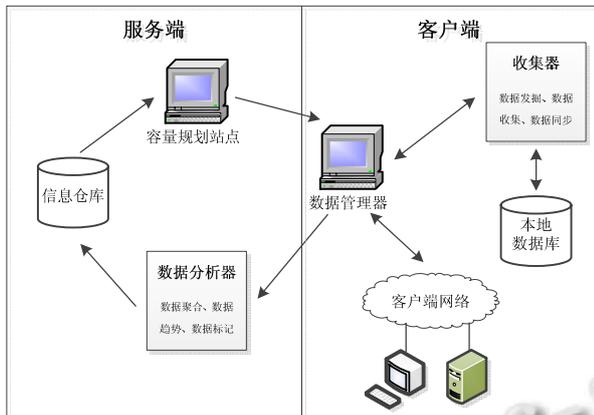


图 1 系统总体框图

数据管理器(Data Manager): 该组件提供了数据收集器的管理接口和用户接口, 为收集的信息数据提供一个直观的可视化界面, 具备以下职能: 设置和维护任务调度、设置目标系统的登陆账号和密码、手工启动某项任务、查看任务进程, 特别是信息数据采集进程、监控收集进程的报错信息、按不同进程管理上传到信息仓库的数据、依据所有被监控系统的信息和执行数据给出总结性报告、启动和停止 Collector 服务等。

数据分析器(Data Analyzer): 该组件是用来对信息数据进行分析 and 标记, 比较数据间的差异性以及分析数据的趋势。

信息仓库(Information Warehouse): 主要用来存储收集器收集上来的具有参考价值的数据, 包括不断增加的可用于企业分析比较的相关数据。利用这些数据指导企业合理地把应用系统部署到虚拟服务器上, 并进行优化。在信息仓库中的信息是匿名的, 尊重个体隐私, 不会对其他个体公开数据。

容量规划站点(Dashboard): 基于 Web 的应用服务是客户端收集上来的数据与服务端存储模块交互的接口, 可以用来查看当前的评估状态。它与数据收集器和信息仓库共同存放于某个具有安全性的中央主机站点。

2.3 评估环境需要

(1) VMware Capacity Planner2.8 机器准备
准备一台虚拟机或物理机能够连接到待评估应用系统所在的局域网和互联网, 机器配置如表 2 所示;

(2) 被监控应用系统环境

VMware Capacity Planner 2.8 支持的被监控系统如表 3 所示。

表 2 CP 服务器需求

类别	需求	说明
操作系统	Windows 2000 (Server or Professional) SP3, Windows XP Pro, Windows Server 2003, Windows Vista, Windows 7, Windows Server 2008	需英文版, 如果数据收集器安装在 XP SP2 上, 必须关掉防火墙。如果被评估系统中含 Windows2000, 那主机系统只能用 Windows 2003 R2 SP2.
系统管理工具或远程连接服务	安装 WMI (Windows Management Instrumentation)或 remote registry	/
处理器	>= 1.5GHz CPU	32 位
内存	>= 1GB	/
本地磁盘	>=2GB	/
网络	带宽>=20KB/s	可访问容量规划站点
浏览器	IE5.5 版本或更高	/

表 3 可被监控的系统环境列表

Windows 系统	Linux/UNIX 系统
Windows NT 4.0(Server, Professional Workstation)	HP-UX 10, 11 (PA-RISC)
Windows 2000 Server/Advanced Server/Datacenter	HP-UX 11.23 (Itanium)
Windows 2000 Server (64-bit Itanium)	Sun Solaris 7 ~ 10 (SPARC) Sun Solaris 9 ~ 10 (x86)
Windows 2000 Professional Workstation	SUSE Linux Enterprise 11
Windows XP Professional (EM64T, AMD 64)	SUSE Linux Enterprise Server9
Windows Server 2003(64-bit Itanium)	SUSE Linux 8 ~ 10
Windows Server 2003 (64-bit x86 / EM64T / AMD64)	Red Hat Linux 8 ~ 9
Windows Vista(Enterprise, Business, Ultimate)	Red Hat Enterprise Linux 6
Windows 7	Red Hat Enterprise Linux (ES/AS/WS) 3 ~ 6
Windows Server 2008, 2008 R2	AIX 5.1 ~ 5.3

3 基于VConsolidate基准评估法

3.1 VConsolidate 简介

VConsolidate 是由 Intel 在 2006 年年底与 IBM 合作共同推出, 该测估工具通过真实地模拟典型虚拟化部署环境中的负载来达到测试服务器虚拟化硬件性能的目的.

3.2 VConsolidate 基准评估指标

VConsolidate 是一项支持多工作量的整合性能指标评估法, 包含了四个可同时运行的性能指标, 分别为 Database、Web、Java 和 Mail. 该评估方法面向虚拟环境, 因而每个组件均在其自己单独的虚拟机和操作系统中运行. 除上述四个性能评估组件之外, 还有第五台虚拟机未运行任何性能评估组件, 由此来模拟闲置的虚拟服务器. 这五台虚拟机构成了一个整合堆栈单元(CSU), 如图 2 所示^[25].

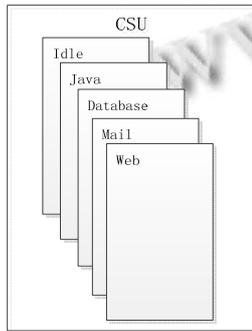


图 2 整合堆栈单元

3.3 VConsolidate 基准评估过程

VConsolidate 基准评估过程为^[25]:

(1)在指定的硬件服务器上生成五台不同的虚拟机(Database、Web、Mail、Java 和 Idle);

(2)VConsolidate 所用的四个组件工作负载如表 4 所示, 工作负载的存储跨整个测试配置进行分布.

表 4 VConsolidate 工作负载

负载类型	负载应用程序
邮件服务器	Windows Exchange Server 2003
Java 服务器	SPECjbb 2005
闲置服务器	无负载程序
网页服务器	WebBench
数据库服务器	Sysbench

(3)将三台客户机与被测服务器连接, 其中两台客户机生成负载(一个用于 Web, 另一个用于邮件), 第三

台客户机来控制运行环境, Java 和 Database 组件不需要运行外部客户机. 随着 CSU 的增加, 所需客户机数将以“三个”为单位递增. 图 3 为 vConsolidate 示意图, 单个 VM 包含各自的工作负载, 这些 VM 文件可以存储到服务器上或存储到外部存储域网络(SAN)上.

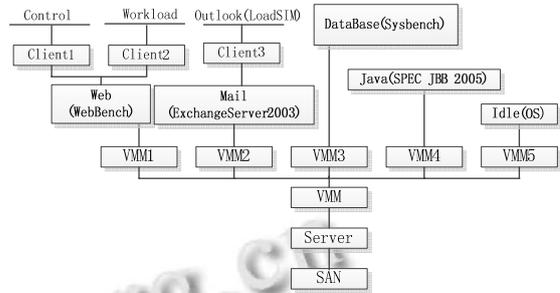


图 3 VConsolidate 示意图

(4)CSU 的数量从 1 开始, 依次递增. 每次计算评估结果得分和处理器的利用率, 直到当 CSU 数目为 N 时评估结果得分小于 CSU 数目为 N-1 时的评估结果得分或者当处理器的使用率为 100%时, 记录该次评估结果, 例如图 4 所示, 图中当 CSU 数目为 4 时, 获得最大得分, 此时处理器的利用率为 78.3%.

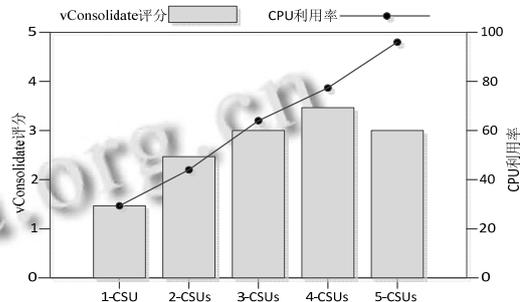


图 4 CSU 评分结果示意图

每次评估得分按以下方法计算:

① 每个负载应用程序的得分通过一个统一的参考系统进行基淮化. 参考系统的得分可以通过在任意系统上运行 1 个 CSU 时的评分获得;

② 计算每个片的总分: 对片内各负载应用程序的基淮化得分计算几何平均值;

③ 计算总得分: 将每片的总分相加.

整个测试过程较长, 一旦到达满负荷的稳定状态时, 将重新执行三次满负荷测试以保证结果的可重复性, 把三次得分的平均值作为最终的结果.

4 基于VMmark基准评估法

4.1 VMmark 简介

2007 年威睿公司针对虚拟化中的服务器整合场景推出了虚拟化测试基准 VMmark. 作为虚拟化技术的测试基准系统, VMmark 可以帮助用户解决: “当前的服务器硬件上, 可以划分几个虚拟机?” 等问题. 服务器供应商通过每台服务器的 VMmark 得分得出可以支持的最大工作负载数, 即该服务器上虚拟机的总体性能. 在服务器整合过程中可以借此帮助企业决策需要多大服务器硬件来运行当前工作负载^[26].

4.2 VMmark 基准评估指标

VMmark 包括六种主要的工作负载: 邮件服务器、Java 服务器、闲置服务器、网页服务器、数据库服务器和文件服务器, 如表 3 所示. 测试时将这六个工作负载运行在单独的虚拟机上, 每个工作量是一个单一的系统运行基准, 被称为“砖瓦(tile)”, 运行时产生的多种 title 得分总和作为整个系统的基准评分. title 的总数是指一个物理系统和虚拟层可调节并衡量的系统整合能力, 通过逐步增加 title 数量来有效评估系统的性能, 最终得分就取决于机器最多能够运行多少块砖瓦^[27].

表 5 VMmark 工作负载

负载类型	负载应用程序
邮件服务器	Windows Exchange Server2003
Java 服务器	SPECjbb2005
闲置服务器	无负载程序
网页服务器	SPECweb2005
数据库服务器	Swingbench
文件服务器	Dbench

4.3 VMmark 基准评估过程

为了测试一个虚拟化平台的最大整合能力, VMmark 采用了类似 TPC-C 服务器基准测试的加压方法, 一步步地增加 title 数量直到平台满负荷运作. 由于每个 title 都是相对重量级的负载单元, 若以 title 数来反映虚拟平台的服务器整合能力则其精度太低, 难以反映不同平台间的微量性能差异. 为解决此问题, title 的数量以及每个虚拟机负载应用程序的性能结果被一起用来计算总分. 此外, 只要整个平台的性能没达到满负荷的稳定状态, title 就会一直增加, 即使新增的 title 降低了负载应用程序的性能.

每次的得分按以下方法计算:

- ① 每个负载应用程序的得分通过一个统一的参考

系统进行基准化;

- ② 计算每个 title 的总分: 对 title 内各负载应用程序的基准化得分计算几何平均值;

- ③ 计算总得分: 将每片的总分相加.

整个测试过程较长, 在测试期间每隔 1 分钟会周期性地检测系统是否达到稳定状态. 一旦到达满负荷的稳定状态时, 将重新执行三次满负荷测试以保证结果的可重复性, 三次得分的平均值作为最后结果^[27].

5 基于SPECvirt_sc2010基准评估法^[3]

5.1 SPEC 简介

SPEC(Standard Performance Evaluation Corporation, 标准性能评估机构)是一个全球性的、权威的第三方非赢利性组织, 旨在联合许多高科技领域的顶尖级企业共同建立公平的标准, 以评测各种技术的性能. SPEC 于 2008 年成立了虚拟化测试组, 包括威睿、AMD、惠普、IBM、甲骨文、Red Hat、英特尔等知名软硬件公司. 2010 年 SPEC 发布了权威的工业标准虚拟化服务器整合基准测试工具 SPECvirt_sc2010^[28, 29].

5.2 SPECvirt_sc2010 基准评估指标

SPECvirt_sc2010 结构、加压方法、评分方法都与 VMmark 1.x 类似. 每个负载单元 title 包括 6 个虚拟机分别运行不同的负载应用服务器: 网页服务器、邮件服务器、应用程序服务器、数据库服务器、基础服务器(infrastructure server)、闲置服务器.

表 6 SPECvirt_sc2010 工作负载

负载类型	负载应用程序
网页服务器	SPECweb2005
邮件服务器	SPECmail2009
应用程序服务器	SPECjAppServer2004
数据库服务器	PostgreSQL
基础服务器	Apache/PHP
闲置服务器	无负载程序

5.3 SPECvirt_sc2010 基准评估过程

SPECvirt_sc2010 结构、加压方法、评分方法都与 VMmark 类似, 最主要的区别是^[3, 4]:

- (1) 解决了移植性问题. 对硬件架构和操作系统没有要求, 所有负载应用程序可以通过各类开源解决方案来实现.

- (2) 引入了客户虚拟机间的网络通讯. 每个 title 内的客户虚拟机间可以相互通讯. 例如数据库服务器

为应用程序服务器的提供后端数据库; 基础服务器为网页服务器提供网络共享文件.

面向企业的虚拟系统性能评估需要综合考虑计算机软硬件、虚拟化方案以及系统的一些特殊应用场合, 有针对性地开展性能评测, 表 7 分析比较传统评估方法与上述四种方法各自的特点和局限性.

6 评估方法比较分析及应用价值

6.1 评估方法比较分析

表 7 服务器虚拟化性能评估方法比较

评估方法	特点	局限
传统的评估方法	过程单一、针对性强: 如有针对性的就 CPU、内存、I/O 等虚拟化的性能进行评估测试.	单一负载: 测试一台服务器上单一负载的性能, 无法评估支持多个虚拟系统或工作量状况下的系统性能.
基于 Capacity Planner 容量规划法	(1)易实施: 通过在单独机器上安装容量规划软件监测所有待评估系统, 获取数据, 上传后并最终生成评估报告; (2)实效性: 对在线应用系统运行状况进行实时监测和评估; (3)周期短: 可自定义评估周期、灵活、方便管理.	(1)受限因素: 评估无法覆盖所有数据参数, 例如核心交换机、冷却系统、网络带宽等条件能否满足虚拟化环境需求; (2)软件许可: 在服务器虚拟化后可能需要购买新的许可或者不同的授权模式; (3)周期性: 数据性质易受评估周期、业务周期影响, 且只能评估当前阶段的系统状况; 若新增应用时, 需要重新评估.
基于 VConsolidate 基准评估法	测试基准程序最主要的任务是针对新场景来设计相应的负载, 不会影响原有在线业务系统, 风险较低. 两者区别: (1)负载加压方式: VConsolidate 监控 CPU 的利用率, 不断增加 CSU 数量直到 CPU 利用率达到最高, 并记录不同 CSU 数量情况下的总分及其 CPU 利用率; (2)最终结果: VConsolidate 的测试结果由两部分组成, 分别是最大总分及对应的 CPU 利用率;	(1)额外资源: 为每个虚拟机增加额外的计算机用来运行基准测试程序将会使得整个测试场景中客户计算机过多, 不便于统一维护和管理; (2)依赖性: 依赖于一个针对邮件服务器的基准测试程序 LoadSim, 该程序只能配合微软公司的 Exchange Server 使用; (3)宏观性: 测试采用固定的几种服务器, 测试结果是宏观的, 无法解释产生性能差异的原因; (4)隔离性: 多虚拟机测试时, 虚拟机之间可能相互干扰.
基于 VMmark 基准评估法	(3)测试过程: VConsolidate 测试过程并非完全自动化, 测试中需要用户参与交互; (4)可移植性: VMmark 主要针对的是 VMware 虚拟化环境, 而 VConsolidate 则支持多种虚拟机管理程序.	(1)与基于 VConsolidate 基准评估法类似; (2)可移植性: 只支持 VMware 公司自己的虚拟机管理器, 因此只能用于比较不同硬件平台的服务器整合性能差异.
基于 SPEC 基准评估法	(1)风险低: 测试基准程序最主要的任务是针对新场景来设计相应的负载, 不会影响原有在线业务系统, 风险低; (2)可移植: 解决了移植性的问题, 对硬件架构和操作系统没有要求, 所有负载应用程序可以通过各类开源解决方案来实现; (3)网络通讯: 引入了客户虚拟机间的网络通讯, 例如基础服务器为网页服务器提供网络共享文件; 数据库服务器为应用服务器的提供后端数据库; (4)成本低: SPEC 是一个全球性的、权威的第三方非赢利性组织, 为客户节省了评估费用.	(1)额外资源: 为每个虚拟机增加额外的计算机用来运行基准测试程序将会使得整个测试场景中客户计算机过多, 不便于统一维护和管理; (2)宏观性: 测试采用固定的几种服务器, 测试结果是宏观的, 无法解释产生性能差异的原因; (3)隔离性: 多虚拟机测试时, 虚拟机之间可能相互干扰.

6.2 基于容量规划法的企业应用价值

(1) 分析资源需求

通过收集当前正在运行的 IT 设备清单、资源使用情况、工作负载在数据中心内的分布状态等数据, 帮助 IT 人员全面掌控数据中心环境, 进而更积极主动的管理决策, 实现项目投资效益最大化. 例如, 在企业需要采购新服务器来满足新的工作负载时, 可以先利用 Capacity Planner 进行假设分析, 帮助 IT 人员在不影

响当前系统运行的前提下进行环境模拟, 评估资源需求, 基于这些分析去更好的选择服务器硬件;

(2) 规划虚拟化环境

很多时候, 企业在实施服务器虚拟化时, IT 人员往往凭借主观经验, 对在线的应用系统或服务器进行整合. 一方面无法充分的利用宿主机的硬件资源, 造成资源浪费. 另一方面在一台宿主机的用户和虚拟机如果过分集中, 容易造成虚拟机性能下降, 不能满

足负载均衡的需求. 通过 Capacity Planner 对可用数据信息和资源清单进行收集和全面分析, IT 人员可以对如何规划和部署虚拟服务器作出更为精确的假设和评估, 选择最佳整合率合理部署应用, 以满足负载均衡和容灾的需求.

6.3 基于 VConsolidate、VMmark、SPECvirt_sc2010 基准评估方法的企业应用价值

(1) 提高资源利用率

近年来, 多处理器、多线程服务器成为在企业中应用的主流. 但是在大多情况下, 这些物理服务器的性能都不能得到充分的利用, 即便是通过虚拟化技术能够在同一台物理服务器内创建多个独立的虚拟机, 由于对系统性能缺乏一个量化的了解, 没有办法给出实际负载能力, 导致在同一台物理服务器内的虚拟系统个数过份集中, 从而影响了虚拟机的整体性能, 又或者建立的虚拟机数量太少而不能充分的利用物理资源. 虚拟机基准测试可以对物理服务器的性能加以检验, 产生一个易于理解的指标, 帮助了解物理服务器的可建虚拟机数量上限, 指导 IT 人员更好的建立虚拟机.

(2) 指导物理服务器的采购

在应用虚拟化技术过程中, 客户需要一种有效的方法来确定如何配置其硬件才能够充分利用这一普遍存在的技术. 基于基准的性能评估方法可以帮助打算购买虚拟基础架构的客户做出准确的决策. 例如 VMmark 是一款可以衡量异构虚拟化工作量的可扩展性并提供一种统一的方法用于不同虚拟化平台上比较基准结果的基准系统, 可以在不同的虚拟化软件和不同的硬件平台上进行比较, 知道服务器上虚拟机的整体效能, 方便使用者去选择他们所需的硬件资源, 来应付虚拟化后的工作. 因此, 企业可以利用这些评估方法选择合适的硬件, 还可以比较不同虚拟化平台上的系统性能和可扩展性.

7 总结与展望

本文针对四种服务器虚拟化性能评估方法进行简要综述, 通过对比分析列举了这四种评估方法的特点与局限性. 为企业在建设服务器虚拟化的过程中给出以下参考意见: (1)通过基于 VConsolidate、VMmark、SPECvirt_sc2010 等方法评估不同硬件平台的服务器虚拟化性能差异, 指导软硬件设施采购; (2)通过基于 Capacity Planner 的容量规划方法对现有的在线系统进

行评估, 帮助和指导企业在进行数据中心服务器虚拟化整合时作出合理的规划, 使 IT 资源被充分利用, 虚拟系统工作效率达到最大化.

随着云服务普及和大规模数据中心的不断涌现, 企业在部署实施服务器虚拟化时考虑因素是多方面的. 如何搭建一整套完备的、自动的系统性能评估体系可以涵盖到虚拟机管理器性能、可伸缩性、整合性、迁移性和隔离性等各方面综合性能评估是未来发展中亟待解决和深入研究的课题.

参考文献

- 1 陈青茶. 虚拟化计算系统性能评测管理工具的研究[硕士学位论文]. 武汉: 华中科技大学, 2008.
- 2 怀进鹏, 李沁, 胡春明. 基于虚拟机的虚拟计算环境研究与设计. 软件学报, 2007, 18(8): 2016-2026.
- 3 黄达伟, 叶可江等. 虚拟计算系统性能评测. 中国计算机学会通讯, 2011. 10(10).
- 4 冯运辉, 兰雨晴. 云计算环境中虚拟机性能评测工具的研究与设计. 信息安全与技术, 2012, 3(11): 39-44.
- 5 Younge AJ, Henschel R, Brown JT, et al. Analysis of virtualization technologies for high performance computing environments. 2011 IEEE International Conference on Cloud Computing (CLOUD). IEEE. 2011. 9-16.
- 6 Bugnion E, Devine S, Rosenblum M, et al. Bringing virtualization to the x86 architecture with the original vmware workstation. ACM Trans. on Computer Systems (TOCS), 2012, 30(4): 12.
- 7 Varrette S, Guzek M, Plugaru V, et al. HPC Performance and energy-efficiency of Xen, KVM and VMware hyper-visors. 2013 25th International Symposium on Computer Architecture and High Performance Computing (SBAC-PAD). IEEE. 2013. 89-96.
- 8 Shirinbab S, Lundberg L, Ilie D. Performance comparison of KVM, VMware and XenServer using a large telecommunication application. The 5th International Conference on Cloud Computing, GRIDs, and Virtualization. 2014. 114-122.
- 9 Ye K, Jiang X, He Q, et al. Evaluate the performance and scalability of image deployment in virtual data center. Network and Parallel Computing. Springer Berlin Heidelberg. 2010. 390-401.

- 10 Ferreto TC, Netto MAS, Calheiros RN, et al. Server consolidation with migration control for virtualized data centers. *Future Generation Computer Systems*, 2011, 27(8): 1027–1034.
- 11 Uddin M, Rahman AA. Server consolidation: An approach to make data centers energy efficient and green. *arXiv preprint arXiv*. 2010. 1010.5037.
- 12 Ye K, Jiang X, Huang D, et al. Live migration of multiple virtual machines with resource reservation in cloud computing environments. *2011 IEEE International Conference on Cloud Computing (CLOUD)*. IEEE. 2011. 267–274.
- 13 Jo C, Gustafsson E, Son J, et al. Efficient live migration of virtual machines using shared storage. *ACM SIGPLAN Notices*. ACM, 2013, 48(7): 41–50.
- 14 Wang X, Sang Y, Liu Y, et al. Considerations on security and trust measurement for virtualized environment. *Convergence*, 2011, 2(2):19–24.
- 15 Zhang W, Zhu M, Gong T, et al. Performance degradation-aware virtual machine live migration in virtualized servers. *2012 13th International Conference on Parallel and Distributed Computing, Applications and Technologies (PDCAT)*. IEEE. 2012. 429–435.
- 16 McVoy LW, Staelin C. Imbench: Portable tools for performance analysis. *USENIX Annual Technical Conference* 1996. 279–294.
- 17 Ruan L, Peng J, Xiao L, et al. Distributed virtual machine monitor for distributed cloud computing nodes integration. *Grid and Pervasive Computing*. Springer Berlin Heidelberg, 2013: 23–31.
- 18 Jiang Q. Virtual machine performance comparison of public IaaS providers in China. *Cloud Computing Congress (APCloudCC), 2012 IEEE Asia Pacific*. IEEE. 2012. 16–19.
- 19 Jin H. Desktop virtualization: Path to pervasive computing in cloud computing era. *2011 6th International Conference*. Port Elizabeth, South Africa. 2011. 26–28.
- 20 Jones R. Netperf benchmark. Accessed March, 2012. <http://www.netperf.org/>.
- 21 Norcott W, Capps D. IOzone Filesystem Benchmark. 2008. <http://www.iozone.org/>.
- 22 Getting Started with Capacity Planner. <http://www.vmware.com/support/pubs>.
- 23 Installation Guide. <http://www.vmware.com/support/pubs>.
- 24 Troubleshooting Guide. <http://www.vmware.com/support/pubs>.
- 25 Theurer A, Rister K, Dobbstein S. A survey of virtualization workloads. *Proc of the 2008 Linux Symposium*. 2008, 2. 215–225.
- 26 van Surksum K. Release: VMware VMmark 2.5. Red, 2014.
- 27 VMware VMmark Benchmarking Guide. VMark 2.5.1, 2013. <http://www.vmware.com/support>.
- 28 SPECvirt_sc2010 Design Overview Version 1.01. http://www.spec.org/virt_sc2010/docs/SPECvirt_Design_Overview.html#VmsTiles.
- 29 White Paper Performance Report PRIMERGY RX600 S4. Version: 2.2b, November 2009.49.