

基于 KVM 的虚拟桌面基础架构设计与优化

唐红梅¹ 郑 刚²

(安徽工程大学现代教育技术中心 芜湖 241000)¹ (安徽工程大学计算机与信息学院 芜湖 241000)²

摘 要 随着云计算的不断发展,虚拟桌面基础架构(Virtual Desktop Infrastructure, VDI)解决方案日益成熟。VDI 建立在虚拟化技术的基础上,突破了时间和空间的限制,有效地解决了传统个人计算机使用过程中存在的诸多问题,是当前桌面云解决方案中主流的架构与部署方式。充分利用 VDI 的优势,结合目前流行的 KVM 虚拟化技术,探讨解析虚拟桌面架构并进行实际部署,详细设计了平台优化方案,最后进行了测试并记录了性能表现,验证了系统的正确性和可用性。结果表明,VDI 为现代机房带来了移动计算、方便管理和降低运维成本等诸多改进,为高校实际安装部署虚拟化平台提供了现实的指导意义。

关键词 虚拟桌面基础架构,虚拟化技术,KVM,优化方案

中图法分类号 TP311.1 文献标识码 A

Design and Optimization on Virtual Desktop Infrastructure Based on KVM

TANG Hong-mei¹ ZHENG Gang²

(Modern Education Technology Center, Anhui Polytechnic University, Wuhu 241000, China)¹

(School of Computer and Information Technology, Anhui Polytechnic University, Wuhu 241000, China)²

Abstract With the continuous development of cloud computing, the virtual desktop infrastructure(VDI) solution is becoming more and more mature. VDI is established on the basis of the virtualization technology, break through the limit of time and space, effectively solve the problems in the process of using traditional personal computer, and is the mainstream in current desktop cloud solution architecture and deployment. In this paper, making full use of the advantages of VDI, combining with the epidemic KVM virtualization technology, virtual desktop resolution architecture was discussed and made actual deployment, and the optimization platform was detailed designed. Finally we did the test and recorded performance results, which verify the correctness and availability of the system. The results show that the VDI brings many conveniences for mobile computing, management and reducing operational costs of modern computer room, and provides real guidance for colleges deploying virtualization platform.

Keywords Virtual desktop infrastructure, Virtualization technology, KVM, Optimization scheme

1 引言

虚拟化技术是云计算为用户提供灵活增减 IT 资源、按需付费等特色服务的基础核心,在虚拟化环境下,云计算成为可能,因此虚拟化技术无疑是当今研究的热点。随着高校办学规模的不断扩大,中心机房也将承载着更多的教学、实践与考试任务,为了满足这些需求,需要构建更新颖、更复杂的软件环境,而传统的 PC 桌面管理模式面临着数据安全、管理成本、设备功耗等诸多方面的挑战^[1]。VDI 通过虚拟化技术将用户的桌面管理都迁移至数据中心,终端用户通过虚拟桌面传输协议访问桌面系统,以“集中管理,分布显示”的原则将用户的桌面应用环境与其显示终端进行解耦合,使得用户可以通过网络接入的方式随时随地获取属于自己的计算机教学实验环境,极大地提高了系统的安全性和硬件系统的利用率,有效弥补了传统机房管理和运维的不足。

2 虚拟化技术分析

2.1 虚拟桌面

虚拟化技术种类很多,主要包含了虚拟服务器技术、虚拟

桌面技术和虚拟应用技术等多个方面,虚拟化技术从实现结构上可以分为 Hypervisor 型虚拟化、宿主型虚拟化、混合型虚拟化。虚拟桌面技术是将平台虚拟化,在硬件、软件环境与用户间增加虚拟层,把计算机的实际物理特性隐藏起来,建立起一个虚拟的计算环境(称为虚拟桌面或虚拟机)。由于具有成本低,更可控、更安全的桌面环境以及面向 PC 和其他终端应用集中等诸多优势,虚拟桌面技术无疑成为当今研究的热点。现阶段,虚拟桌面技术的解决方案较多,如:VMware Viewer, Citrix XenDesktop 和 KVM^[2](Kernel-based Virtual Machine)。本文以 KVM 技术构建的虚拟化桌面为讨论对象,采用当前兼容性最好的 Spice 作为桌面传输协议,讨论其在机房环境应用下带来的改进。

2.2 KVM 技术

KVM 虚拟化技术是 x86 硬件平台上 Linux 操作系统完全原生的全虚拟化解决方案^[3],具有开源的特性以及出色的表现,是目前最流行的虚拟化产品,适用于包括 Windows 以及 Linux 操作系统的各种变种的多种操作系统。

KVM 技术采用宿主型虚拟化模型,是第一个基于 Linux

本文受安徽工程大学教学研究项目(2015jcxzz01)资助。

唐红梅 高级工程师,主要研究领域为计算机软件技术、云计算及其应用;郑 刚 副教授,硕士生导师,主要研究领域为云计算、分布式计算。

内核的 Hypervisor^[4](虚拟机监视器,也称 VMM),使 KVM 平台上运行的虚拟机能够使用 Linux 内核提供的内存管理和进程管理等机制,从而让虚拟机获得与宿主机(Host)相当或者比 Host 更好的性能。KVM 并不是一个完整的模拟器,具体的模拟器工作需要借助 QEMU^[5](虚拟化模拟器)来完成,形成可控制 KVM 内核模块的用户空间工具 QEMU-KVM。由于 QEMU 工具效率仍然不高,本文采用 QEMU-KVM^[6]和 Libvirt(KVM 虚拟机进行管理的工具和应用编程接口 API)相结合的方式,提供对虚拟机客户机(Guest)的管理以及对虚拟化网络和存储的管理。

3 虚拟桌面架构设计

基于 Libvirt 及 QEMU-KVM 的虚拟桌面架构是面向智能移动终端解决方案中的服务端实现部分,系统架构如图 1 所示。

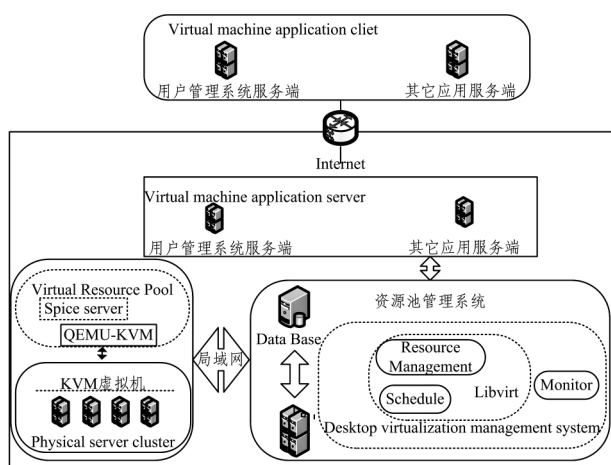


图 1 虚拟桌面系统架构

在解决方案中,虚拟化资源池(Virtualization Resource Pool)借助 Libvirt 工具库实现且构建于 QEMU-KVM 之上的机器集群,利用 Spice 协议实施虚拟桌面向终端用户的交付,通过 Spice server 以 Spice 协议^[7]的方式与终端用户上的 Spice client 完成通信以实现虚拟桌面传输,client 通过 Spice 协议显示并操作虚拟桌面。资源池管理系统基于 QEMU-KVM 完成,通过 TCP/IP 协议连接到这些服务器节点的 Libvirt 守护进程,借助 Libvirt 库实现对物理资源和虚拟资源的监控和管理。资源池管理系统和虚拟化资源池共同构成了混合型虚拟桌面服务端,并通过统一的标准化 API 为上层应用提供底层服务,上层应用的服务端通过调用 API 获得底层服务^[8]。

4 系统优化方案

本系统采用的 KVM 虚拟化技术虽然已有很高的性能,但由于运行在虚拟机上的操作系统是针对物理机开发的,其在虚拟机上运行时会出现一些不适应虚拟化环境的因素,造成系统在使用过程中的一些瓶颈。因此,通过优化资源的解决方案可以进一步提高系统的整体性能,本文主要探讨虚拟机环境下的内存优化策略。

4.1 内存虚拟化的实现

实现内存虚拟化的核心是虚拟地址到物理地址的转换,KVM 可以采用影子页表(Shadow page)或扩展页表(Extended Page Table, EPT)技术来实现^[9]。

由于 Guest 维护的虚拟 Guest 内存页表地址和 Host 维护的物理内存页表地址在开始时并没有形成映射关系,影子页表技术的核心思想是利用页故障的方式产生一个 Page fault int 的异常,然后跳出 KVM 的执行流程,使得 KVM 异常执行程序可以捕获并且处理该异常,同时根据 Guest 的 Page table 填充影子页表,即可使得在影子页表中保存客户机操作系统(Guest OS)中的虚拟地址到宿主机操作系统(Host OS)中物理地址的转换映射。

影子页表的使用带来了地址转换过程的简化,但是也有一些明显的缺点,如 Guest OS 在读写 CR3 或 Guest 页表不完整时都会导致 VM exit 的产生,使内存虚拟化的效率明显降低。为了改善这一状况,KVM 会使用 EPT 这种高效的虚拟机-主机内存管理方法来实现内存虚拟化。

如图 2 所示,EPT 页表结构定义了虚拟机物理地址到主机物理地址之间的映射关系,所有的虚拟机物理地址都必须经过 EPT 页表结构转换后,用户才能访问主机物理内存。在采用 EPT 实现内存虚拟化之后,虚拟机读写 CR3、执行影子页表虚拟化指令和 Guest 页表结构自身导致的页故障都不会产生 VM exit,可以大大提高内存虚拟化的实现效率。

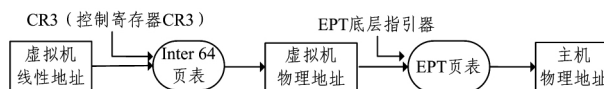


图 2 EPT 内存虚拟化工作原理

4.2 内存的过载使用

KVM 中内存允许过载使用(over-commit),KVM 能够让分配给 Guest 的内存总数大于实际可用的物理内存总数,这样就可以达到节省内存开销的目的。由于 Guest OS 及其应用程序并非一直百分之百地利用其分配到的内存,而且 Host 上的多个 Guest 一般也不会同时达到 100% 的内存使用率,因此内存过载分配是可行的,实现内存过载使用的方式有 3 种:1)内存交换(swapping),用交换空间(swap space)来弥补内存的不足;2)气球(ballooning),通过 virio_balloon 驱动实现 Host Hypervisor 和 Guest 之间的协作来完成;3)页共享(page sharing),通过 KSM(Kernel samepage merging)合并多个 Guest 进程使用的相同内存页。其中,swapping 方式最为成熟,在目前被广泛使用。

若用 swapping 的方式让内存过载使用,则要求有足够的 swap space 来满足所有的 Guest 进程和 Host 中其他进程所需内存。可用的物理内存空间和交换空间的大小之和应该等于或大于配置给所有 Guest 的内存总和,否则在各个 Guest 内存同时达到较高使用比率时可能会有 Guest 被强制关闭(因内存不足)。

若服务器有 32GB 的物理内存,目标是在其上运行 64 个内存配置为 1GB 的 Guest。在 Host 中,大约需要 4GB 大小的内存用于满足系统进程、驱动、磁盘缓存及其他应用程序所需(不包括 Guest 进程所需内存)。计算过程如下:Guest 所需交换分区为 $64 \times 1\text{GB} + 4\text{GB} + 32\text{GB} = 36\text{GB}$,而根据 Redhat 的建议,对于 32GB 物理内存的 RHEL 系统,推荐使用 8GB 的交换分区。因此,在 Host 中总共就需要建立 44GB($36\text{GB} + 8\text{GB}$)的交换分区来安全地实现 Guest 内存的过载使用,以达到节省内存的目的。

4.3 KSM 技术

KSM(内核相同页面合并)^[10]是一个 Linux 内核模块,允

许不同的进程和 KVM 虚拟机共享匿名内存。虚拟化环境中,多个虚拟机一般都使用相同的操作系统,因此存在大量的相同页,KSM 的主要任务就是在系统中找到相同的页面,它使用两个 rbtree,一个是稳定树,另一个是不稳定树。稳定树只包含已经共享的匿名页,不稳定树只包含由 KSM 监测但未共享的页面。使用两棵树是一种优化方式,对最可能适合使用共享树的虚拟内存区域使用两棵树会增加页面共享的概率,同时降低不稳定树的不稳定性。

KSM 页面合并过程包含两个功能:page_wrprotect()和 replace_page()。前者写保护所有页表映射的页面;后者合并两个页面并相应地更新页表,然后释放已合并并且没有页表映射的匿名页面。在 page_wrprotect()之后,通过 memcmp()确定两个页面的内容完全相同,之后用 replace_page 合并页。

通常 x86 架构包括的大多数体系结构都是 4096 个字节,找到一个相等的页面成本是 memcmp()的成本乘以树的级数,这得益于 rbtree 的特性。所有的计算时间复杂度为 $O(\log(N))$, N 是扫描的 KSM 的页面总数。因此即使在不理想的情况下,前 4092 个字节的所有页面是相等的,只有最后 4 个字节不同,KSM 树算法的性能也不会降得太多。

5 实施与测试

基于 KVM 的虚拟桌面架构的实施主要包括硬件方案及系统配置。1)硬件方面,KVM 虚拟化技术对计算机配置要求较高,在具体选择时,可以结合实际需要,若规模较小,则可以采用共享服务器进行配置,在硬件的选择上可以考虑虚拟化服务器及云终端两种,要选择能够支持 Spice 协议的云终端。2)系统配置方面,在服务器配置上,要求服务器可以运行 QEMU-KVM 软件以支持服务器虚拟化功能^[11],并安装 Libvirt 软件以响应远程桌面的管理需求;在网络管理平台安装过程中,可以通过相关开发环境及框架,采取具有开源性的 Nginx 作为服务器;在虚拟创建方面,虚拟化平台的核心是管理虚拟机,在项目设计过程中,虚拟机平台的设计要兼顾多个方面(如系统类型等),以满足系统运行的具体要求。

对系统进行性能测试,本文关注的重点是内存的使用率,测试的方法是通过增加固定数量的虚拟机来增加 Host 的内存使用率^[12]。测试结果如图 3 所示,从涨跌柱线可以看出,优化后的内存性能有了很明显的改善,优化后的内存使用率维持在非常低的水平,因此在内存不足的环境下使用 VDI 机房管理平台是可行的。

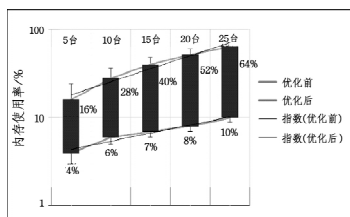


图3 优化前后内存性能测试结果图

为了更进一步地对内存性能进行测试,将 KVM 虚拟机系统与目前使用率很高的 VMware 虚拟机系统进行对比测试,在保证系统正常运行且用户体验相当的情况下反映出两者能力值的差异。如图 4 所示,在相同的虚拟机数量下,两种虚拟化环境下的具体的微观性能指标可以通过内存使用率来表现。

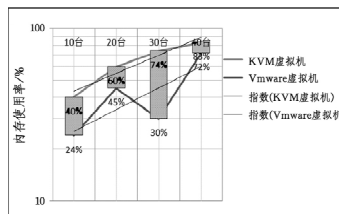


图4 虚拟机系统内存性能测试对比图

从测试结果可以看出,VMware 虚拟机的性能优于 KVM 虚拟机,性能优势平均在 10%~20% 之间,但基于 KVM 虚拟化平台开发的机房管理系统的优势体现在:1)KVM 属于开源软件而 VMware 不开源,这就决定了 KVM 具有更大的灵活性和可扩展性,同时开源也带来了创新性;2)KVM 内存使用效率比较高,CPU 利用量稳定,也比较低;3)KVM 是内核的一部分,因此可以充分利用内核本身的优化和改进将标准内核转换成一个系统管理程序,因此它的优势不言而喻。

结束语 近年来,云计算技术得到了飞速的发展,随着网络技术与虚拟化技术的发展,虚拟桌面技术越发被广泛接受与使用,通过与虚拟服务器技术、虚拟应用技术和数据计算中心的灵活整合,可以提供比传统计算方式更便捷、更有可扩展性的管理方式^[13]。KVM 技术是虚拟桌面技术演变过程中的最新发展,具有其他虚拟化解决方案所不及的独特优势,不仅可以提高机房的共享度,减少重复建设,降低维护成本,而且可以实现绿色环保和低碳节能的规划目标。随着高校教育改革和教育信息化的推进,搭建校级统一虚拟化教学环境平台将成为高校信息化建设的热点之一。

参考文献

- [1] 陈益全. 基于桌面云的计算机基础实验室建设模式研究[J]. 兰州教育学院学报, 2015, 31(6): 104-105.
- [2] 黄秋兰, 李莎, 程耀东, 等. 高能物理计算环境中 KVM 虚拟机的性能优化与应用[J]. 计算机科学, 2015, 42(1): 67-70.
- [3] 孙琳程. 虚拟机 KVM 和 XEN 的性能分析[J]. 电脑知识与技术, 2013, 9(10): 2364-2366.
- [4] 肖涛, 何怀文, 梁瑞仕. 桌面云在高校计算机实验室中的应用[J]. 实验室科学, 2015, 18(1): 82-88.
- [5] 王森, 朱常鹏, 韩博. 一种基于 qemu 的动态迁移模型[J]. 计算机科学, 2015, 42(6A): 337-340.
- [6] 车翔. QEMU-KVM 设备虚拟化研究与改进[D]. 成都: 成都理工大学, 2012.
- [7] 金彪, 郑小建, 姚志强, 等. 桌面虚拟化与计算机实验室管理[J]. 实验技术与管理, 2014, 31(2): 85-88.
- [8] 黄金敢. 高校教学环境中桌面云架构研究与实现[J]. 计算机技术与发展, 2013, 23(12): 222-225.
- [9] 李思扬. 云计算平台中虚拟化环境的研究与实现[D]. 北京: 北京邮电大学, 2012.
- [10] 任永杰, 单海涛. KVM 虚拟化技术实战与原理析[M]. 北京: 机械工业出版社, 2013: 48-197.
- [11] 宋蕾. 虚拟技术在高校计算机机房实验室中的应用[J]. 实验室研究与探索, 2013, 32(6): 274-278.
- [12] 白伟, 李凤英. 桌面虚拟化软件数据处理能力试与分析[J]. 宁夏师范学院学报, 2013, 34(6): 74-76.
- [13] 吴坤喜, 冯广. 虚拟云桌面的系统架构研究与设计[J]. 广东工业大学学报, 2015(1): 113-115.