

# 虚拟化中存储性能瓶颈研究

何 鹏

(四川旅游学院信息与工程学院, 四川 成都 610100)

**摘 要:** 随着信息技术的发展,越来越多的行业都向信息化、智能化发展,随之而来,业务系统也不断增多,那么就需要更多的服务器,更多的存储空间来满足业务系统日益增多的需求,传统的增加物理服务器方式难以满足日益扩增的数据存储需求,数据存储也逐渐由单一的硬盘管理向数据中心发展,虚拟化技术的诞生对IT行业产生了重大的影响,也对存储性能提出了更高的要求,尤其是大量并发用户在同时访问数据中心时,更是要求存储具有较高性能,虚拟环境下面临着各种各样的难题,其中最主要的就是存储性能瓶颈。

**关键词:** 虚拟化; 数据存储; 性能瓶颈; 数据中心

**中图分类号:** TP399 **文献标识码:** A

## Research on Storage Performance Bottleneck in Virtualization

HE Peng

(School of Information and Engineering, Sichuan Tourism University, Chengdu 610100, China)

**Abstract:** With the development of information technology, more and more industries are developing towards information and intelligence. As a result, more servers and storage space are needed to satisfy the requirements of the increasing business systems. However, the traditional way of adding physical servers is difficult to meet the ever-increasing demand for data storage. Data storage has gradually evolved from a single hard disk management to a data center. The birth of virtualization technology has had a major impact on the IT industry, imposing higher requirements for storage performance. Especially, when a large number of concurrent users access the data center at the same time, the storage is required to have higher performance, and the virtual environment is faced with various problems, the most important of which is the storage performance bottleneck.

**Keywords:** virtualization; storage; performance bottleneck; data center

### 1 引言(Introduction)

近年来,虚拟化技术得到普遍应用,由此也开启了存储的任意门,面对不断增长的业务系统和庞大的数据量,普通的单块或多块物理硬盘,已经不再能满足需求,由此必须不断扩充存储空间,但同时还需要兼顾不影响业务的正常运作,尤其是大量并发用户在同时访问数据中心时,对存储性能更是提出了新的挑战。因此存储架构也需要随机应变,存储虚拟化技术中,有多种架构,每种架构都有其独特的优势,但也存在一定的局限性。

本文的研究内容主要基于VMware vSphere虚拟化方式,首先简要介绍了虚拟化和存储虚拟化的基本概念,研究了虚拟化中数据的存储实现方式,其次分别剖析了虚拟化技术在服务器虚拟化、桌面虚拟化,以及应用虚拟化几种虚拟化环境下,遇到的存储性能方面的瓶颈,最后得出结论,要想突破此性能瓶颈,必须同时满足高读、写IOPS,高吞吐

量,低延迟,以及更快部署和简单的网络架构。

### 2 虚拟化简介(Virtualization profile)

什么是虚拟化,通俗的讲,虚拟化就是将一台物理计算机分割成多台逻辑计算机。实质上虚拟化是指计算机元件在虚拟的基础上而不是真实的基础上运行,它是一个为了简化管理,优化资源的解决方案<sup>[1]</sup>。CPU的虚拟化技术可以单CPU模拟多CPU并行,允许一个平台同时运行多个操作系统,并且应用程序都可以在相互独立的空间内运行而互不影响,从而显著提高计算机的工作效率。虚拟化技术可以扩大硬件的容量,简化软件的重新配置过程。计算机中,Virtualization是一种资源管理技术,是将计算机的各种实体资源,如服务器、网络、内存及存储等,予以抽象、转换后呈现出来,打破实体结构间的不可切割的障碍,使用户可以比原本的组态更好的方式来应用这些资源。这些资源的新虚拟部分是不受现有资源的架设方式,地域或物理组态所限

制。一般所指的虚拟化资源包括计算能力和资料存储。

实际生产环境中，虚拟化技术主要用来解决高性能的物理硬件产能过剩的老的旧的硬件产能过低的重组重用，透明化底层物理硬件，从而最大化的利用物理硬件。

虚拟化技术中，可以同时运行多个操作系统，而且每一个操作系统中都有多个程序运行，每一个操作系统都运行在一个虚拟的CPU或者是虚拟主机上。

虚拟化技术与VMware Workstation等同样能达到虚拟效果的软件不同，是一个巨大的技术进步，具体表现在减少软件虚拟机相关开销和支持更广泛的操作系统方面。

### 3 存储虚拟化(Storage virtualization)

存储虚拟化<sup>[2]</sup>是通过虚拟卷映射、流数据定位、数据快照、虚拟机等技术实现异构存储设备的统一管理，以及存储位置无关性而提出的,目的在于屏蔽存储管理中的一系列复杂问题而向用户提供简单透明统一的存储访问模式。简单来说，就是对物理硬盘等存储资源进行抽象化表示，即将实体存储空间(如硬盘)分隔成多个不同的逻辑存储空间。具体来讲，就是将可用的物理硬盘等存储空间分成几个或若干个虚拟的卷，使得这些磁盘卷不再受磁盘阵列，磁盘驱动器等实际存储元件的物理布局或拓扑结构的限制。通常情况下，虚拟的磁盘卷展现在操作系统面前的是一种物理磁盘的抽象，使操作系统像使用硬盘一样使用。图1就是存储虚拟化示意图。不同的应用系统使用逻辑上不同的存储空间(虚拟磁盘)，但是实质上都是使用的一个物理存储空间，即一块或多块物理硬盘构成的同一存储池。

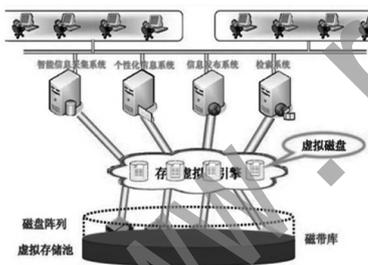


图1 存储虚拟化示意图

Fig.1 Storage virtualization schematic

虚拟化连接存储方式<sup>[3]</sup>多种多样，实现方式不同，性能也不尽相同，比如有基于iscsi的，有基于NFS的，有基于NAS网络的，有基于光纤通道(FC)的等，如图2存储虚拟化实现方式所示。

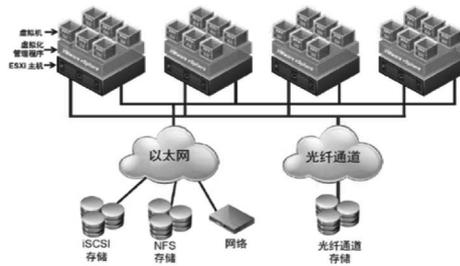


图2 存储虚拟化实现方式

Fig.2 Storage virtualization implementations

虽然随着计算机的普及，虚拟化技术的出现，颠覆了IT的传统架构，使得桌面、应用、数据快速的转移到数据中心，让IT变得更加高效、便捷、可扩展、易维护，同时降低了IT建设的成本。然而，在虚拟化技术大幅提升服务器CPU利用率的同时，传统服务器存储受其机械原理的局限，无法满足虚拟化环境下的共享存储资源的性能要求。存储成为制约服务器整体效率的瓶颈。传统改善的方式，依靠扩展SAN、NAS等网络存储，一方面部署成本高昂，另一方面使得IT架构和维护变得更加复杂，且最终也无法较好的解决性能瓶颈问题。

### 4 服务器虚拟化的困境(The plight of server virtualization)

服务器虚拟化<sup>[4]</sup>是将服务器物理资源抽象成一种逻辑资源，使得传统的服务器变成几个或者是几十个逻辑上相互隔离的虚拟服务器，使得系统资源不再受到服务器物理上的限制。把CPU、内存、硬盘、I/O等硬件集中起来变成可以动态调配和管理的统一资源，并将“资源池”中的资源动态分配给需要工作的负载，实现服务器的整合，达到充分利用服务器硬件资源，提升其利用率的目的。

随着信息技术的发展和学校信息化建设的步伐，我校已经初步实现服务器虚拟化。服务器虚拟化的部署大大提升了服务器CPU、内存等核心资源的利用率，Windows Server或者Linux Server等虚拟机可以动态的共享服务器的核心资源，确保资源利用最大化的同时，减少服务器的采购、维护及各种电力开销等成本。服务器虚拟化将服务器CPU的平均利用率从15%提升到75%左右，而存储性能的低下及用户并发、业务、数据的爆发增长，不得不使用RAID或网络存储的方式来缓解，这使得服务器虚拟化整体拥有成本(TCO)依旧保留在一个较高的水平。运行在服务器上的虚拟机，部署着企业的众多关键应用或服务，如数据库系统、邮件系统、网站系统、共享下载业务等。这些应用与服务大部分都是用户并发量多的IO密集型业务，因此在小块数据随机写、大块顺序读，或者混合读写等方面，对存储性能提出了更高的要求，而存储系统本身在虚拟化前仅能够满足单系统应用的小并发需求。在虚拟化后，不同虚拟机及虚拟机上应用系统对存储的性能争用，当前业务、数据并发的爆发性增长，都使得原有的存储系统无法满足服务器虚拟化对存储性能的需求。

而简单的磁盘阵列方式的本地存储扩展，或者NAS、SAN的网络存储扩展，在耗费大量的IT成本支出的同时，只能很有限的缓解服务器虚拟化对存储性能带来的挑战，无法从根本上解决虚拟化后存储的性能瓶颈，也无法全面提升服务器虚拟化的整体效率。

### 5 桌面虚拟化的挑战(Desktop virtualization challenges)

VDI<sup>[5]</sup>英文全称Virtual Desktop Infrastructure，即虚拟

桌面基础架构,是近年来虚拟化技术由服务器虚拟化向桌面虚拟化延伸的一个技术名称。虚拟桌面基础架构采用“集中计算,分布显示”的原则,通过虚拟化技术,将所有客户端的运算合为一体,在企业数据中心内进行集中处理,而桌面用户仅负责输入输出与界面显示,不参与任何计算和应用。VDI(虚拟桌面基础架构),带来了IT桌面交付的巨大变革,在降低IT成本的同时,带来了更大的灵活性,然而也带来了全新的挑战,它依托于服务器虚拟化技术,将用户桌面迁移到数据中心,提升了服务器的利用率,降低终端设备成本,同时在IT管控(桌面及应用管理、病毒防护、补丁升级、数据安全等)及IT维护(终端采购、故障处理、系统部署等)方面,带来突破性的变革,在企业桌面、应用快速增长时,确保IT可以高速及高效的运转。

VDI(虚拟桌面基础架构)虽然存在巨大的优势,但同时也引入全新的挑战,除了对网络的依赖及远程显示协议固有缺陷之外,还有其他问题:如启动、关闭风暴,当有大量并发用户同时登录、启动或者关闭虚拟桌面时,对存储性能进行争抢,又比如杀毒风暴,当多个虚拟机桌面用户同时间进行并发杀毒或者扫描漏洞等,也会造成存储性能的争抢,还有一种最常见的就是更新风暴,Windows系统补丁集中进行更新等,这些问题都会给部分用户带来启动桌面数分钟,甚至几十分钟的延迟,桌面操作卡顿,极大的降低了用户体验,原因说到底就是我们的存储性能低下,读、写速度低所致。虚拟桌面通过部署SAN或NAS来试图解决以上问题,却又带来新的挑战,比如IT架构改造,无法利用现有服务器环境部署虚拟桌面,需要重新架构网络和存储,又比如整体成本提升,需要额外众多光纤或以太存储及交换设备。

## 6 应用虚拟化的难题(Problems in applying virtualization)

至今为止,大家对应用虚拟化<sup>[6]</sup>的解释都不尽相同。其中Vmware的解释是通过应用虚拟化技术,可以将应用程序的部署、执行、升级等操作均在一个隔离环境中进行,避免了对本地操作系统配置的更改,确保了本地操作系统和其他应用程序的完整性。由此应用虚拟化带来了很多有利的价值,诸如可以让用户不再需要在本地下载,安装应用,用户可以通过共享应用的方式,简化应用的部署和使用,而企业也可以更好的来管控用户在企业内部对应用的使用权限,提升IT安全。

在应用虚拟化的场景中,客户端无需部署和安装应用程序,通过远程共享应用模式,即众多的用户通过网络远程共享服务器上的同一应用。这一应用虚拟化方式给部署应用的服务器带来了极大的IO考验。众多并发的应用访问使得传统的存储IO难以应对,IT不得不拆分并发量、同时部署额外的

服务器来满足这一需求。

而另一种应用虚拟化场景,是通过沙盒应用的模式,将应用程序沙盒化,存储于共享应用服务器,用户在有授权,有需要或者登录时刻,前往共享应用服务器下载应用,在用户终端本地或者虚拟桌面无需安装即可使用这些应用程序。这一场景对存储系统的吞吐量或者带宽提出了更大的需求,不得不增加部署共享服务器,以满足峰值或者大并发时刻的存储带宽需求。

应用虚拟化的出现,颠覆了应用的使用方式,同时也给企业对提升IT安全管控带来了新思路。应用虚拟化技术可有效解决应用软件数量增多带来的管理压力、应用软件配置冲突、部署困难等问题,然而虚拟化应用的最终载体—存储系统如何去满足大并发的存储带宽及高IOPS需求,却是一个很大的难题。

## 7 结论(Conclusion)

那么,我们究竟如何解决遇到的这些问题呢?各种虚拟化方式及应用都给用户带来了各种便利和价值,然后存储性能低下、扩展代价高却成为制约虚拟化发展的瓶颈。要想突破各种虚拟化方式和应用虚拟化的存储性能瓶颈,我们就必须要突破虚拟化在存储系统性能方面的瓶颈,必须同时满足高读、写IOPS,高吞吐量,低延迟,以及更快部署和简单的架构。在此基础之上,我们才可以进行深层次的虚拟化性能优化,全面的提升单虚拟机/桌面的IOPS和吞吐量,在不改变用户当前虚拟化架构,以及不进行存储的复杂扩展的基础上,彻底解决虚拟化在存储方面的性能瓶颈,获得最佳用户体验,同时整体降低用户虚拟化部署的成本。

## 参考文献(References)

- [1] 邱震,贺春林,王洪静.存储虚拟化技术研究[J].软件导刊,2013(01):35-39.
- [2] 谭生龙.存储虚拟化技术的研究[J].微计算机应用,2010(01):33-38.
- [3] 常征.存储虚拟化技术的研究与比较[Z].2011年安徽省智能电网技术论坛,2011(16):142-144.
- [4] 杨照岩,滕红丽,谷小青,等.服务器虚拟化技术在高校数据中心建设中的应用[J].农业网络信息,2013(11):107-109.
- [5] 张庆萍.虚拟桌面基础架构(VDI)安全研究[J].计算机安全,2011(04):72-74.
- [6] 陈靖,黄聪会,孙璐,等.应用虚拟化技术研究进展[J].空军工程大学学报(自然科学版),2013,14(06):53-58.

## 作者简介:

何 鹏(1986-),男,硕士,助教.研究领域:计算机网络,虚拟化方向,大数据.