

OpenStack 云环境中 KVM 虚拟机性能测试与优化

作者：李桂林 崔广章 李永宝

来源：《物联网技术》2016 年第 02 期

摘要：目前，虚拟化技术在一定程度上影响着云计算平台的性能，但 KVM 虚拟机上运行的操作系统（Guest OS）是针对物理机开发的操作系统，因而不适合虚拟化环境的因素。文章在 OpenStack 云环境下，对 KVM 虚拟机磁盘 I/O、网卡 I/O 和处理器的计算能力进行了测试，对比了 KVM 虚拟机与物理机的性能差距，并通过将虚拟机的处理器（VCPU）和物理机处理器（CPU）进行绑定，同时将宿主机的扩展页表（EPT）开启，从而实现了对 KVM 虚拟机性能的优化。

关键词：OpenStack；KVM；性能测试；优化

中图分类号：TP309 文献标识码：A 文章编号：2095-1302（2016）02-00-04

0 引言

随着计算机和互联网的快速发展，大数据和云计算应运而生，其中云计算发展尤为迅猛，现在比较成熟的云平台就有 Abiquo 公司的 AbiCloud、Apache 基金会的 Hadoop、微软的 Azure 平台、阿里巴巴的阿里云、Google 的 Google AppEngine、中国移动的 BigCloud-大云平台等，在现有这些云平台中，只有 OpenStack 是完全开源的，形势表明，开源云平台是未来云计算发展的一个趋势。

OpenStack 是一个由 NASA（美国国家航空航天局）和 Rackspace 合作研发并发起的，以 Apache 许可证授权的自由软件和开放源代码项目。OpenStack 包含两个主要模块：Nova 和 Swift，前者是 NASA 开发的虚拟服务器部署和业务计算模块；后者是 Rackspace 开发的分布式云存储模块，两者可以一起使用，也可以分开单独使用。OpenStack 除了得到以上两家公司的支持外还被 Dell、Citrix、Cisco、Canonical 等大公司支持，并且发展迅速。

1 虚拟化技术与虚拟机性能测试

1.1 KVM 技术

KVM（Kernel-based Virtual Machine）是一个开源的系统虚拟化模块，它需要硬件支持，如 Intel VT 技术或 AMD V 技术，是基于硬件的完全虚拟化，完全内置于 Linux。KVM 由两部分组成：（1）内核中管理虚拟化硬件的设备驱动，通过字符设备/dev/kvm 实现；（2）QEMU

用来模拟用户的 PC 机。KVM 与其他虚拟技术不同的是 KVM 是 Linux 内核的一部分，直接运行于硬件之上，虚拟机作为一个正常的 Linux 进程来实现，并由标准 Linux 调度程序调度，能受益于 Linux 内核的所有特征，它能够直接与硬件交互，不需要修改虚拟化操作系统，相比于 XEN 等其他虚拟技术 KVM 更加轻量级。KVM 虚拟机体系结构如图 1 所示。

KVM 虚拟机使用软件模拟的方式实现 I/O 设备的虚拟化，其实现方式是由内核中的 KVM 模块截获客户操作系统中的 I/O 请求，交给运行在宿主机 (Host OS) 上的 QEMU，QEMU 将这些请求转换为对宿主机操作系统的调用，通过宿主机操作系统的设备驱动访问物理硬件，实现对 I/O 设备的虚拟化。该方法依赖 QEMU 对设备的模拟，因此实现简洁，但是由于 I/O 处理流程中涉及多个环境，切换较多，其 I/O 性能很不理想。虽然 KVM 较新的版本中已经将一些关键设备的模拟进行了优化，但是主要的设备，例如磁盘和网卡虚拟化的性能开销仍旧较大。

1.2 KVM 虚拟机性能测试

在虚拟化的应用中，虚拟机的性能高低在生产一线是一个很严峻的问题。虚拟机性能主要由处理器 (CPU) 的计算能力、磁盘 I/O 和网卡 I/O 来进行综合评估。对 KVM 虚拟机的性能测试，主要是通过各种基准测试程序 (benchmark) 来比较 KVM 虚拟机和物理机的性能差距，根据对比结果分析 KVM 虚拟机的性能特征。

1.2.1 虚拟机的 CPU 性能测试

(1) 测试环境：曙光 A620r-G 服务器，1.8 T RAID5 硬盘阵列，8 核 E5-4670 Inter 至强 CPU，62 G 内存，云平台版本为 OpenStack Juno，KVM 版本为 KVM-83。

(2) 测试程序：基准测试程序 bonnie++-1.03e，Bonnie++ 是一个用来测试 UNIX 文件系统性能的测试工具，主要目的是为了找出系统的性能瓶颈，测试结果中包含的参数主要有字符读/写速度 (单位：k/sec)、块读/写速度 (单位：k/sec)、块的修改/重写 (k/sec)、各个操作的 CPU 利用率 (百分比)。

(3) 测试方法：在云平台上创建 10 台虚拟机，每个虚拟机分配 2 个虚拟内核 (VCPU)、2 G 内存、40 G 硬盘，让基准测试程序从在 1 个虚拟机上运行，依次增加直到在 10 个虚拟机上同时运行，将每次运行的结果与基准测试程序在物理服务器上运行的结果进行比较得出虚拟机磁盘字符读/写速度、块读/写速度、块的修改/重写速度和 CPU 的利用率与物理服务器的差距。

在测试过程中，虚拟机和物理服务器每一项测试都进行 10 次，最后对各项指标求平均值，然后以物理服务器的测试结果为准，将虚拟机的测试结果与物理服务器的测试结果进行对比和分析。每次测试结束之后都要重启虚拟机和物理服务器，从而保证每次测试结果都是互不影响的。

图 2 给出了从基准测试程序在 1 台虚拟机上运行到基准测试程序同时在 10 台虚拟机上运行时 CPU 利用率的对比结果，横坐标为同时运行基准测试程序虚拟机的个数，纵坐标为每种情况下虚拟机 CPU 利用率与基准测试程序在物理服务器上运行时 CPU 利用率的比值。

图 2 中给出了测试的 10 组 KVM 上虚拟机字符读/写、块读/写、块重写的 CPU 利用率和物理服务器的字符读/写、块读/写、块重写的 CPU 利用率的比值，通过对比可知，KVM 虚拟机与物理服务器的计算能力差距是很小的，损失率不到 10%。

1.2.2 虚拟机磁盘读写 I/O 的性能测试

图 3 给出了从基准程序在 1 台虚拟机上运行到基准测试程序在 10 台虚拟机上同时运行的对比结果，图中横坐标为同时运行基准测试程序虚拟机的个数，纵坐标为每种情况下基准测试运行在虚拟机上的字符读/写、块读/写、块重写的速度与基准测试程序运行在物理服务器上时字符读/写、块读/写、块重写的速度的比值。

分析图 3 可知，当虚拟机的数量为 1 时，虚拟机和物理服务器的磁盘读写 I/O 性能差距很小。但随着虚拟机数量的增加，虚拟机字符读/写、块读/写、块重写的速度下降比较明显，通过分析得出虚拟机的数量与磁盘读写性能存在线性关系。

1.2.3 虚拟机网络 I/O 性能测试

测试环境：曙光 A620r-G 服务器，1.8T RAID5 硬盘阵列，8 核 E5-4670 Inter 至强 CPU，62 G 内存，云平台版本为 OpenStack Juno，KVM 版本为 KVM-83。

测试程序：netperf-2.7.0 “netperf -H [部署 netserver 服务器的 IP]-t[网络连接类型]-l [测试的时间长度]（在该测试中时间统一为 60 s）”。

测试方法：在云平台上创建 10 台虚拟机，每个虚拟机分配 2 个虚拟内核、2 G 内存、40 G 硬盘，在每台虚拟机上安装一套 netperf-2.7.0 基准测试程序，让基准测试程序从在 1 个虚拟机上运行，依次增加直到在 10 个虚拟机上同时运行，让这些测试值与在服务器上运行 netperf-2.7.0 基准测试程序得出的试值进行比较。

Netperf 是一种网络性能的测量工具，主要针对基于 TCP 或 UDP 的传输。Netperf 根据应用的不同，可以进行不同模式的网络性能测试，即批量数据传输（bulk data transfer）模式和请求/应答（request/reponse）模式。Netperf 测试结果所反映的是一个系统能够以多快的速度向另外一个系统发送数据，以及另外一个系统能够以多快的速度接收数据。Netperf 工具以 client/server 方式工作。server 端是 netserver，用来侦听来自 client 端的连接，client 端是 netperf，用来向 server 发起网络测试。在 client 与 server 之间，首先建立一个控制连接，传递

有关测试配置的信息，以及测试结果；在控制连接建立并传递了测试配置信息后，client 与 server 之间会再建立一个测试连接，用来来回传递特殊的流量模式，以测试网络的性能。

在测试过程中，虚拟机和物理服务器每一项测试都进行 10 次，最后对各项指标求平均值，然后以物理服务器的测试结果为准，将虚拟机的测试结果与物理服务器的测试结果进行对比和分析。每次测试结束之后都要重启虚拟机和物理服务器，从而保证每次测试结果都是互不影响的。

图 4 给出了让基准测试程序 netperf-2.7.0 “netperf -H [部署 netserver 服务器的 IP]-t TCP_STREAM-l 120”在 1 台虚拟机上运行，依次增加直到在 10 台虚拟机上同时运行得出的测试值与基准测试程序 netperf-2.7.0 “netperf -H [部署 netserver 服务器的 IP]-t TCP_STREAM-l 120”在物理服务器上运行得出的测试值的比较。该比较是虚拟机网络 I/O 和物理服务器网络 I/O 在 TCP_STREAM 方面的比较，横坐标代表虚拟机的数量，纵坐标代表网卡的流量（单位为 MB）。

(TCP_STREAM)

分析图 4 可以看出，在 1 台虚拟机时，虚拟机 TCP_STREAM 方面的网络 I/O 与物理服务器 TCP_STREAM 方面的网络 I/O 性能几乎是一样的，随着虚拟机数量的增加，虚拟机 TCP_STREAM 方面的网络 I/O 大幅下降，下降的比例与虚拟机的数量成线性关系。

图 5 给出了让基准测试程序 netperf-2.7.0 “netperf-H[部署 netserver 服务器的 IP]-t TCP_STREAM-l 120”从在 1 台虚拟机上运行，依次增加直到在 10 台虚拟机上同时运行得出的测试值与基准测试程序 netperf-2.7.0 “netperf-H[部署 netserver 服务器的 IP]-t UDP_STREAM-l 120”在物理服务器上运行得出的测试值的比较。

(TCP_STREAM)

分析图 5 可以看出，在一台虚拟机的时候，虚拟机 UDP_STREAM 方面的网络 I/O 与物理服务器 UDP_STREAM 方面的网络 I/O 性能几乎是一样的，随着虚拟机数量的增加，虚拟机 UDP_STREAM 方面的网络 I/O 大幅下降，下降的幅度与虚拟机的数量成线性关系，与图 4 所描述的 TCP_STREAM 相类似，但是在下降幅度上比 TCP_STREAM 要大。

2 优化及优化效果

2.1 优化策略

基于上面的测试结果，发现 KVM 虚拟机的 CPU 和 I/O 都不是很理想，特别是磁盘读写 I/O，这是由于运行在虚拟机上的操作系统是针对物理机开发的，在虚拟机上运行时存在一些

不适应虚拟化环境的因素，影响虚拟机的性能。下面将从 CPU 亲和性、扩展页表等影响 KVM 性能的因素进行改进和优化。

2.2 CPU 亲和性

由 KVM 虚拟机的原理可知，每个虚拟机都是由 Linux 调度程序管理的标准进程。CPU 亲和性就是进程要在某个给定的 CPU 上尽量长时间地运行而不被迁移到其他处理器的倾向性。在多核的环境中，Linux 内核对进程的调度算法也是按照进程对处理器的亲和性设置的。开启进程与处理器的亲和性可以减少进程在多个 CPU 之间交换运行带来的缓存命中失效，但从一个进程运行的角度来看，如果能使一个进程一直运行在同一个 CPU 上，减少进程在多个处理器间频繁迁移，从而提高缓存的命中率，进而提高虚拟机的性能。

本文采用将进程绑定在固定 CPU 上的方法来减少 KVM 虚拟机进程在 CPU 间的频繁调度，从而优化虚拟机性能。具体要求如下：

- (1) 用户在创建虚拟机时，可以指定其在某一个 CPU 上运行；
- (2) 虚拟机运行在指定的核上时，整个运行期间不会发生迁移。

如图 6 所示，在 KVM 运行的过程中，通过 Virsh 工具将虚拟机 CPU 与物理 CPU 进行绑定，将虚拟机的 VCPU0 与物理机的 CPU0 绑定，并在某个时刻接触绑定关系，下一个时刻会绑定在物理机的 CPU1 上。在任意给定的时刻，VMCS 与物理 CPU 是一一对应的关系。

2.3 扩展页表 EPT

扩展页表 EPT (Extended Page Table) 是 Intel 在 VT-x 技术基础上增加的一种硬件辅助虚拟化技术。支持该虚拟化技术的处理器有根模式和非根模式两种工作模式。虚拟机监控

(VMM) 工作在根模式，虚拟机工作在非根模式。扩展页表只有在非根模式下才起作用，通过扩展页表的翻译可以将虚拟机物理地址转换成物理机的物理地址，这种转化是由硬件完成的，所用的扩展页表结构都是由虚拟机监控器创建、维护和更新。

2.4 优化效果

在本节中，将虚拟机操作系统中的 VCPU 与物理机 CPU 进行绑定，将宿主机的扩展页表 EPT 开启，测虚拟机的磁盘 I/O、网络 I/O 和 CPU 的利用率与在第 1 节中虚拟机操作系统中的 VCPU 与物理机 CPU 未进行绑定和宿主机的扩展页表未开启时，测得虚拟机的磁盘 I/O、网络 I/O 和 CPU 利用率并进行对比的结果如图 7 所示。

- (a) KVM 虚拟机 CPU 利用率优化前后对比

(b) KVM 虚拟机磁盘读写 I/O 优化前后对比

(c) KVM 虚拟网卡 I/O (TCP_STREAM) 优化前后对比

(d) KVM 虚拟网卡 I/O (UDP_STREAM) 优化前后对比

图 7 虚拟机磁盘 I/O、网络 I/O 和

CPU 利用率优化前后的对比结果

由图 7 分析可知, 图 7 (a) 中 KVM 虚拟机 CPU 利用率优化前后对比表明优化之后虚拟机的计算能力提高 3% ~ 5% ; 图 7 (b) 中 KVM 虚拟机磁盘读写 I/O 优化前后对比表明优化之后虚拟机的磁盘读写提高 7% ~ 10% ; 图 7 (c) 中 KVM 虚拟网卡 I/O (TCP_STREAM) 优化前后对比和图 7 (d) KVM 虚拟网卡 I/O (UDP_STREAM) 优化前后对比表明, 优化前后的网络 I/O 几乎没有变化。

3 结 语

虚拟化技术在云计算领域所起的作用越来越突出, 但在虚拟机的处理器 (VCPU)、I/O 方面还存在一些问题, 需要进一步优化。本文通过一些基准测试程序对 KVM 虚拟机的处理器 (VCPU)、磁盘 I/O 和网络 I/O 进行了测试, 对在 OpenStack 云环境下虚拟机和物理机的性能差异有了初步认识, 接着根据 KVM 虚拟机的虚拟化原理分析了影响虚拟机性能的因素, 并通过将虚拟机的处理器 (VCPU) 与物理机的处理器 (CPU) 进行绑定和开启宿主机的扩展页表 (EPT) 来对 KVM 虚拟机进行优化。优化后虚拟机的计算能力提高 3% ~ 5%, 磁盘 I/O 提高 7% ~ 10%, 而网络 I/O 没有变化。因此, 可以得出结论: 在 OpenStack 私有云环境中, KVM 虚拟机比较适合执行 CPU 密集型和网络 I/O 密集型的任务, 但是当 KVM 虚拟机在执行磁盘 I/O 密集型的任务时, 虚拟机的性能会大幅度下降, 并且在增加虚拟机数量的时候需要着重考虑网卡 I/O 和硬盘 I/O 性能的下降, 否则 OpenStack 云平台的性能会很不理想。

参考文献

[1]汤儒, 李勤伟.Openstack 云环境中的 KVM 虚拟机性能分析[J].微型机与应用, 2013, 32 (23) : 94-96.

[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]崔泽永, 赵会群.基于 KVM 的虚拟化研究及应用[J].计算机技术与发展, 2011, 21 (6) : 108-111.

[4]张彬彬, 汪小林, 杨亮, 等.修改客户操作系统优化 KVM 虚拟机的 I/O 性能[J].计算机学报, 2010, 33 (12) : 2312-2320.

[5]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.

[6]于波.Linux KVM 的虚拟化性能[J].软件世界, 2007 (11) : 49-50.

[7]兰雨晴, 宋潇豫, 马立克, 等.系统虚拟化技性能评测[J].电信科学, 2010 (S1) : 19-24.