

# 云计算与虚拟化

肖臻  
北京大学

关键词：云计算 虚拟化 资源配置 网格计算

随着网络的普及，有人预测未来有很大一部分日常事务都会在网上完成。这为云计算模型提供了广阔的应用空间。

云计算的应用范围很广，包括提供复杂商业服务的网上银行、网上商店、电子商务、网络游戏、社交网站等。互联网应用程序通常需要查询分布式数据库来支持金融交易，因而对云计算中资源的有效利用和优化提出了新的挑战。云计算上运行的网络应用展现在无限多的潜在客户面前，从而引发资源供应的两难问题：如果按照系统的平均负荷来配置资源，则不足以应付繁忙期的高负荷量；如果按照系统可能出现的最高负载来配置资源，那么大部分时间资源将会闲置。比如，一家网上书店通过云计算模式来租用所需资源，如果按照其大多数时候的客户需求量来配置服务器、带宽等资源，则不足以应付像《哈利波特》这样的畅销书发行时带来的抢购狂潮。如果该网站总是按照《哈利波特》的标准来配置资源，那么将导致其资源利用率低下。人们从前所说的效用计算模型和现在兴起的云计算模型力图将计算资源像水、电之类的公用服务一样来提供。实现这种服务模式需要加强现有计算资源的易用性、稳定性、可靠性和流动性。

此问题以前在网格计算中研究过，但是传统的网格计算主要是针对批处理的科学计算程序，其做法基本上是将一个大的计算任务分解成几个子任务，然后将子任务分发到各计算节点，每个节点接受到计算任务后一般会运行很长一段时间（比如几个小时、几天或更长的时

间），在任务完成之后再行新分配的任务。与之相比，云计算应用具有一系列不同的特点。很多使用云计算的应用系统需要与用户进行实时交互，比如商业网站。与传统的科学计算程序相比，用户每一次访问要求的计算量一般不大，但是可能同时收到大量的用户访问，并且必须把反应时间控制在可以接受的范围之内（秒级的响应速度甚至更快）。此外，这些应用通常需要一天24小时不间断地运行，因此资源的重新配置（比如网站的复制、转移等）必须不能打扰其正常运行。相比之下，批处理的科学计算程序一旦把某项计算任务分配到某个节点，通常会令其在该节点运行结束后再进行新的资源配置。

北京大学云计算课题组针对云计算模式下资源的有效管理和优化进行了深入的研究。图1显示了课题组开发的虚拟集群系统的动态迁移过程[3]。该系统在基于开源的Xen[1]虚拟机管理器和Usher[2]虚拟集群管理系统架构下加入了动态资源调配算法。云计算用户只需将其应用安装在云中的任何一台服务器上，系统就能够根据用户需求自动地将该应用部署到相关的一系列服务器上，并根据用户的地理位置和服务器的负载来决定由哪一台服务器来完成用户的要求。为了实现这个目标，云计算应用所需的资源必须根据该应用的需求进行实时配置。例如，如果一个应用突然成为热门应用而收到大量的用户请求，系统将自动拨出更多的服务器以满足该应用的需求。过一段时间后，这个应用可能不再热门，收到的用户请求有明

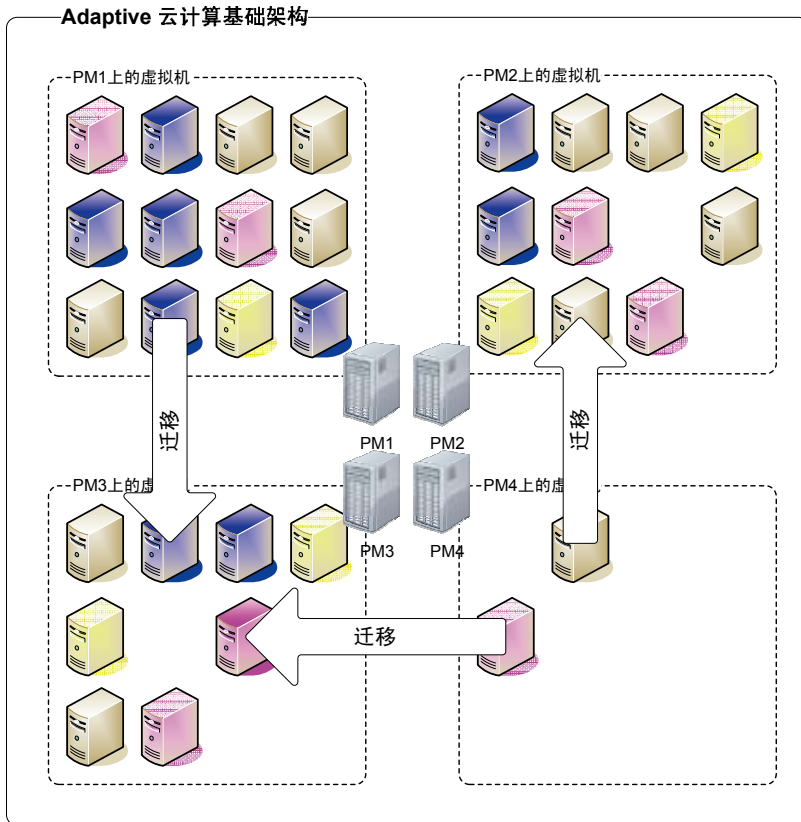


图1 动态虚拟集群架构

显的下降，那么这些服务器可以被重新分配给其它应用。相比之下，现有的很多数据中心服务器的平均负载往往很低，而当用户请求猛增时某些服务器有可能出现过载，这是因为没有能够实现资源动态调配所致。

虚拟技术的应用对于提高网络计算资源的效率有很大帮助。由于各应用对计算环境有不同的要求，在某台计算机上执行的应用程序往往很难转移到其它计算机上执行。比如有些应用是在Windows环境下开发的，而有些应用是在Unix环境下开发的。即使同样是Unix大环境，其操作系统的具体类型也有很多种，如Solaris、BSD、Linux等等。Linux也有许多不同版本。除了操作系统以外，各机器所安装的开发库（Library）和辅助执行程序往往也有区别。这种现象不利于充分利用计算资源。比如，一个计算中心有些服务器负载可能过大，

而其它服务器负载很低。在这种情况下，很自然的做法是将那些负载过大的服务器上的一些应用转移到低负载的服务器上。但是运行环境的差别给应用的迁移带来很大困难。这个问题的解决办法是运用虚拟机技术对物理机的硬件细节进行屏蔽，为其上的应用提供一个统一的执行平台。可以把虚拟技术看成对其下计算资源的一种抽象，这种抽象通过牺牲部分性能来降低计算环境的异构性。一个具有高度一致性的虚拟计算环境对于

提高计算资源的可维护性也有很多好处。与不断下降的硬件价格相比，IT产业的劳动力成本已经成为相关企业的一项大开支。一个熟练的系统管理员的工资远远超过一组高端电脑的价格。IT企业为培养和保留这样一支专业队伍要花费大量的财力，即总拥有成本（Total Cost of Ownership, TCO）。虚拟化技术带来的同构环境大大降低了系统的复杂性，甚至可以将应用程序与其执行环境捆绑在一起（即所谓的虚拟应用，Virtual Appliance），然后共同迁移。这是因为一个大型操作系统的代码中往往只有一小部分是其上应用程序所真正需要的。虚拟应用的概念就是将一个复杂系统中真正要用到的部分抽取出来与其上应用结合在一起形成一个可以独立执行的整体，从而大大提高资源利用的效率。

对于许多应用来说，服务器往往不是单机使

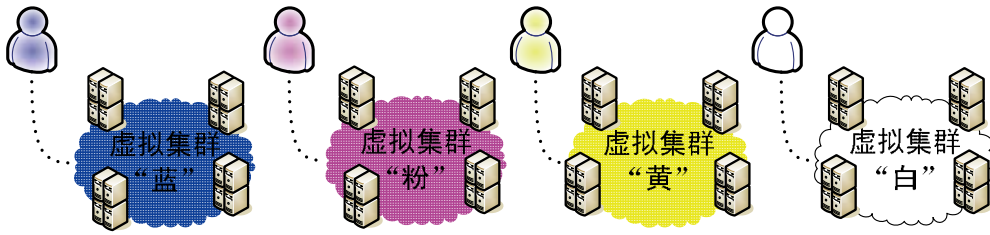


图2 不同应用所构建的虚拟集群与用户的交互

用的，而是将一组服务器组成一个集群来提供服务，使之具有良好的容错性、可靠性，并负载均衡。机群的前端可以使用一个负载均衡交换机（Load Balancing Switch）作为与用户交互的界面。该交换机有外部的IP（Internet Protocol）地址，而集群内的其它服务器则使用内部IP地址。市场上有很多集群管理软件。但是集群软件重新配置的时间往往很长。比如在资源重新配置中，需要把一台服务器从原有的机群中抽调出来，加入到另一个机群中。这需要关闭该服务器上原来运行的应用，配置新的应用并启动该应用。这个过程可能带来很长时间的延迟，尤其对于一些复杂的商业应用，各种参数配置很复杂，其关闭、启动的时间本来就很长，在有些情况下需要重启操作系统，所花的时间就更长了。我们曾对市场上几种最流行的集群软件做过测试，发现在这方面的开销都很大，反映出有软件没有针对这方面进行优化。有的软件片面依靠“心跳”信息的超时来检测集群中服务器的变化，这对于快速资源配

置来说是远远不够的。

虚拟技术为此提供了很好的解决方案。我们可以构造由虚拟机组成的虚拟集群（Virtual Cluster），如图2所示。由于一台物理服务器上可以运行多个虚拟机，每个虚拟机可以加入虚拟集群并对应于一个应用。一台服务器实际上可以同时加入对应于不同应用的集群。

系统可以根据各应用的优先级对在其上运行的各虚拟机进行动态的资源分配。如果某应用有剩余资源，可以将其虚拟集群中的部分虚拟机挂起（Suspend），即将当前状态保留到硬盘，而让出CPU、内存等资源给其它虚拟机使用。也可以让各虚拟机保持运行状态，但是用“气球”技术将一台虚拟机的内存资源动态地调配给另一台虚拟机，或者让它们以不同的比例分享CPU资源，等等。这样避免了前面提到的从集群中直接转移物理服务器所带来的开销。图3显示了两个应用通过虚拟化技术共享五台物理机的一个实例。

缩短资源重新配置的时间有很强的实用意

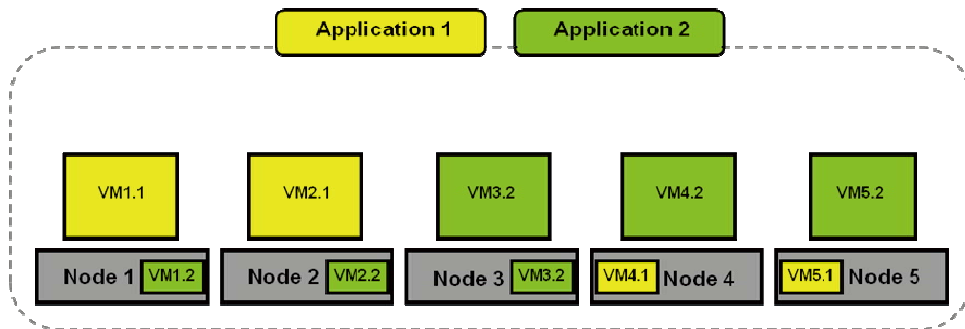


图3 虚拟集群的动态资源调配

义。如果进行一次资源配置需要几分钟时间，那么重新配置资源的决定就只能在更长的间隔上（如每隔几个小时）做一次。从这个意义上讲，加快资源配置的过程可以让我们能够更及时地、灵活地响应系统中资源的动态变化，真正体现出“随需而变”的资源理念。用挂起和恢复（Resume）的方法进行资源调配的效率与磁盘I/O的性能有关。用独立冗余磁盘阵列（Redundant Array of Independent Disk, RAID）等并行读写技术可以提高磁盘的吞吐量，也可以在恢复时先调入部分内存然后在运行中按需取页。■



肖臻

北京大学信息科学技术学院研究员，博士生导师。主要研究方向为分布式系统、云计算、虚拟化、容错性计算等。zhen.xiao@gmail.com

## 参考文献

- [1] Paul Barham et al. "Xen and the art of virtualization", Proc. of the 19th ACM symposium on Operating systems principles (SOSP'03), Oct 2003.
- [2] Marvin McNett, Diwaker Gupta, Amin Vahdat, and Geoffrey M. Voelker. "Usher: An Extensible Framework for Managing Clusters of Virtual Machines", Proc. of the 21st Large Installation System Administration Conference (LISA), Nov 2007.
- [3] Christopher Clark et al. "Live migration of virtual machines". Proc. of the Symposium on Networked Systems Design and Implementation (NSDI'05), May 2005.