

基于能耗感知的虚拟机迁移管理软件*

Research of Optimization Models for Energy based on Live-migration of Virtual Machine

张汉林, 王泽鹏, 陈宁江**

ZHANG Hanlin, WANG Zepeng, CHEN Ningjiang

(广西大学计算机与电子信息学院, 广西南宁 530004)

(School of Computer, Electronics and Information in Guangxi University, Nanning, Guangxi, 530004, China)

摘要:【目的】在云计算环境中,通过虚拟机在线迁移可以实现服务器的节能优化。【方法】对多种虚拟机迁移模型进行综合分析,研发一个虚拟机在线迁移管理软件,实现云环境的监控与虚拟机迁移模拟,并采用 cloud-sim 作为云仿真平台对虚拟机在线迁移管理软件(Energy-aware Live-migration of Virtual Machine Manager Software, ELVMan)的应用进行测试。【结果】该软件可有效指导降低能源损耗。【结论】通过该软件可以为现实的云环境集群选择最优的迁移策略,达到最大化节能。

关键词:云计算 虚拟机迁移 节能 模拟

中图分类号: TP31 **文献标识码:** A **文章编号:** 1002-7378(2017)01-0048-05

Abstract:【Objective】In cloud computing environment, live-migration of virtual machine can achieve the goal of energy optimization on the server. 【Methods】In this paper, by comprehensive analysis and modeling of VM migration model, a virtual machine live-migration management software has been developed to achieve cloud environment monitoring and virtual machine migration simulation. And the CloudSim is used as a cloud simulation platform to test the application of ELVMan software. 【Results】The experimental results show that it can provide effective guidance for reducing energy consumption. 【Conclusion】The software can be used to select the optimal migration strategy for the real cloud cluster to achieve maximum energy saving.

Key words: cloud computing, virtual machine migration, energy saving, simulation

0 引言

【研究意义】云计算^[1]以虚拟化^[2]为载体,可以

有效规约异构资源,达到计算能力、网络、存储的重新调度与共享化。利用虚拟机在线迁移技术^[3],可以在几乎不影响用户使用的情况下进行虚拟机在物理主机间的转移,实现物理机与虚拟机的重映射;通过关闭多余的物理机,可以降低能源消耗,提高资源利用率。【前人研究进展】已有不少研究工作关注虚拟机迁移过程中的资源调度与能耗状况,例如 Younge 等^[4]在多层次虚拟机迁移调度中考虑运行/结束时间、硬件故障以及虚拟机间的争夺问题;Beloglazov 等^[5]利用多个维度来计算资源相互平衡的虚拟机重新放置算法,以减少资源碎片和提高设备利用率;Miller 等^[6]用云计算框架研究能量感知的

收稿日期: 2016-11-16

修回日期: 2016-12-07

作者简介: 张汉林(1993—),男,硕士研究生,主要从事云计算技术方向研究。

* 国家自然科学基金项目(61063012, 61363003)和国家科技支撑计划课题(2015BAH55F02)资助。

** 通信作者: 陈宁江(1976—),男,博士,教授,主要从事软件、网络分布计算和云计算方向研究, E-mail: chnj@gxu.edu.cn.

虚拟机调度和适应性的资源管理方式;Semeraro等^[7]设计一种节能的虚拟机资源管理系统,综合考虑节点间的拓扑与物理机间的热量分布;等等。【本研究切入点】主流的云计算产品安装繁琐、功能冗余,不方便独立研究虚拟机迁移现象。因此本文研发一个虚拟机在线迁移管理软件(Energy-aware Live-migration of Virtual Machine Manager Software,简称ELVMan),实现或者改进虚拟机迁移策略,动态收集物理机与虚拟机运行信息,提供合适的节能评价。【拟解决的关键问题】在云计算管理中,通过云环境的监控与虚拟机迁移模拟,对实际环境中的虚拟机资源调度进行有效评价,帮助更好地实现节能。

1 软件架构

ELVMan 软件主要由两部分组成,一部分是由服务器集群为用户提供按需分配的虚拟机服务,一部分采用 Web 客户端提供虚拟机的管理方式,总体架构图如图 1 所示。采用 Kernel-based Virtual Machine(KVM)虚拟机作为虚拟机监控器,使用 NFS 网络文件系统服务实现虚拟机的硬盘共享,使用 VNC 远程桌面功能为虚拟机提供访问功能。在虚拟机迁移方式中采用基于共享存储的在线迁移机制,并使用内存预拷贝技术,即通过内存反复迭代的方式快速地将所有内存内容传输到目的主机。在 Web 端主要使用 Struts2 及相关 Java web 技术等,实现远程桌面控制。

在虚拟机管理模块,主要采用 libvirt 为软件接口,实现对虚拟机的管理操作(图 2)。管理员使用软件进行虚拟机在线迁移管理,在用户使用虚拟机服务不间断的情况下实现虚拟机在主机之间进行迁移。在虚拟机创建功能中,有自定义模板创建与按照已有模板创建两种方式。虚拟机的暂停/激活功能可以帮助切换虚拟机运行状态与运行状态静态存储两个状态。存储池管理模块通过管理共享文件系统的方式,可以管理虚拟机的共享存储信息。日志管理模块通过记录使用者在系统中的操作情况,分析出虚拟机出现的错误。CPU 监控模块远程获取 CPU 利用率,包括物理主机的 CPU 利用率与虚拟机主机的 CPU 利用率。

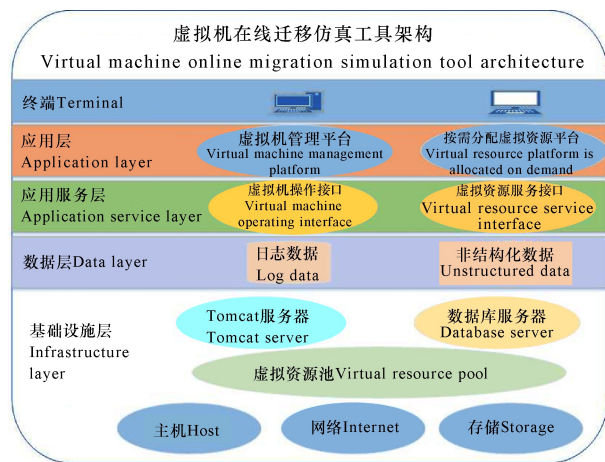


图 1 ELVMan 软件架构

Fig. 1 ELVMan's software architecture



图 2 虚拟机迁移管理操作

Fig. 2 Virtual machine migration management operations view

2 ELVMan 对虚拟机动态迁移的支持

ELVMan 软件能够收集物理机的 CPU 使用率与实际虚拟机的 CPU 使用率,对多个迁移检测算法进行模拟仿真,选择出最优的虚拟机在线迁移策略。按照迁移策略对虚拟机进行动态迁移,关闭多余的物理主机,达到减少资源碎片、减低能源消耗的目的。

2.1 迁移性能评价模型

虚拟机迁移虽然可以帮助实现虚拟机集群的性能优化,但是迁移过程仍会损失在迁移时间段内的性能,并且一般平均虚拟机迁移所带来的损耗为 CPU 消耗的 10%^[8]。因此,我们在该软件中,设置有针对性对虚拟机迁移所产生的性能损耗设计相关的评价模型。

虚拟机预拷贝算法因为停机速度短、迁移过程对用户的影响较低的优点,成为主要的迁移算法。在虚拟机预拷贝过程中,网络会反复传送内存某一时刻的状态,而在传输的这段时间里,内存的部分页面会发生改变,被称为内存页面变脏。其中, K 为预拷贝算法中变脏速率/网络传输速率。

公式中标识符的含义: Memory, 内存; bandwidth, 带宽速率大小; n , 迭代传输的次数; u_j , 计算 CPU 的函数; N , 服务器个数; K , 内存的变脏速率与网络传输内存速率的比值; M , 虚拟机个数; T_{si} , 迁移过程中实际延迟时间; T_{ai} , 协议等级时间; C_{di} , 计算性能总数实际退化性能; C_{rj} , 计算性能总数。

虚拟机迁移时间利用式(1)计算:

$$T_{m_j} = \frac{\text{Memory} * \frac{1 - K^{n-1}}{1 - K}}{\text{bandwidth}} \quad (1)$$

在迁移过程损耗的计算使用量:

$$U_{d_j} = 0.1 * \int_{t_0}^{t_0 + T_{m_j}} u_j(t) dt \quad (2)$$

迁移损耗为迁移过程中所降低的 CPU 计算用量的比例,0.1 被假定在运行过程中 10% 的性能被损耗。

平均每台主机延误服务等级协议时间(SLA Violation Time per Active Host)(该数值越大表明对服务等级协议违背越多):

$$\text{SLATAH} = \frac{1}{N} \sum_{i=1}^N \frac{T_{si}}{T_{ai}} \quad (3)$$

平均迁移产生的性能退化(Performance Degradation due to Migration, PDM)等于实际退化性能/计算性能总数(该数值越大,代表性能损耗越大):

$$\text{PDM} = \frac{1}{M} \sum_{i=1}^N \frac{C_{di}}{C_{rj}} \quad (4)$$

对 SLA Violation(SLAV)的定义为上述两个性能指标的乘积(该数值越大,说明对于 SLA 的服务质量越难保证):

$$\text{SLAV} = \text{SLATAH} * \text{PDM} \quad (5)$$

对于 SLAV 的计算是评价一个迁移算法的质量是否好的标准,亦可以分析得出,对于一些配置较高的虚拟机,如内存较大或者 CPU 计算量较大的虚拟机都会使服务质量降低。

2.2 虚拟机迁移仿真

对于虚拟机迁移,主要是考虑选择需要迁出的虚拟机和迁入目的主机的算法,将虚拟机从负载过大的主机上转移。判断负载的方法除静态阈值法,还有一些统计学上的方法可以动态判断阈值的范围。ELVMan 软件实现的过载函数有中位数绝对偏差(Median Absolute Deviation, MAD)^[9]、四分位距法(Interquartile Range, IQR)、鲁棒性局部回归(Robust Local Regression, LRR)等。

我们对线性回归(Local Regression, LR)方法进行改进,设计一种局部加权线性回归法,其中局部回归^[10]是一种通过给拟合数据增加权值的方法,使拟合函数更加接近真实。局部加权线性回归法的主要思想在于,拟合数据越接近真实函数的大小,所分配的权值就越大,就越接近真实的拟合数据;拟合数据与真实函数的差距越大,就说明数据的可靠性越小,拟合权值就越小,通过这种情况拟合的数据可以比较真实地把过载主机选择出来。加权函数如下:

$$w_i(x) = (1 - \frac{(x_k - x_i)^3}{x_k - x_1})^3 \quad (6)$$

为找到最小二乘意义上的最佳拟合,利用 a 和 b 两个参数对下面的式子做最小化处理:

$$\min \sum_{i=1}^n w_i(x) (y_i - a - bx_i)^2 \quad (7)$$

另一方面,过载主机需要对其上的多个虚拟机进行选择迁移。ELVMan 主要实现多种虚拟机迁移策略,包括选择最小迁移时间^[11]、随机选择^[12]、虚拟机最大关联^[13]、虚拟机最小利用率 Minimum Utilization(MU)^[14]等。

3 仿真测试

3.1 测试方法

我们对 ELVMan 软件的应用进行测试。在测试中,采用 CloudSim 作为云仿真平台,利用 ELV-

Man 将已有的虚拟机利用率数据进行多种算法的模拟。表 1 给出多种类型虚拟机模拟的性能参数。

实验模拟的硬件环境为异构的两种物理服务器,其配置为 HP ProLiant ML110 G4 (1 x [Xeon 3040 1860 MHz, 2 cores], 4 GB); HP ProLiant ML110 G5 (1 x [Xeon 3075 2660 MHz, 2 cores], 4 GB)。

表 1 模拟的物理机性能参数

Table 1 The physical machine performance parameters in simulation

类型 Type	单核计算 能力 Single core computing power (MIPS)	核数 Auditing	内存 RAM (MB)	内网带宽 Intranet bandwidth (Mib/s)	虚拟机存 储大小 Virtual machine storage size(GB)
1	2 500	1	870	100	2.5
2	2 000	1	1 740	100	2.5
3	1 000	1	1 740	100	2.5
4	500	1	613	100	2.5

两台物理机的耗电指标分别按照 SPEC 测试的标准进行耗电模拟,数据分别为 { 86, 89.4, 92.6, 96, 99.5, 102, 106, 108, 112, 114, 117 }; { 93.7, 97, 101, 105, 110, 116, 121, 125, 129, 133, 135 }, 单位为瓦(W)。

该实验主要通过不同规模的物理机与虚拟机对算法进行模拟。模拟时间为 1 h, 每 5 min 对集群内的虚拟机进行过载检测, 并将负载不足的虚拟机进行迁移算法模拟。

由于云计算对外租用服务时不清楚用户实时消耗物理机资源的情况(CPU 利用率), 所以采集真实的虚拟机利用率数据, 来模拟用户对虚拟机的使用行为, 每 5 min 从云任务文件取一回数据。

3.2 测试结果分析

在 10 台物理主机, 10 个云任务下虚拟机各个算法的性能指标测试(图 3)。其中, Dvdfs 表示为在传统的动态电压频率调整的方式进行节能省电的情况。可以看到, 相比绝大多数虚拟机迁移所采用的性能指标, 基于云计算的迁移算法能节省大量电能, 确实能够明显减少能耗损失, 在该规模的模拟下可以得出静态阈值的过载函数与最小迁移的选择函数最为节能。

在集群为 10 的规模下, 前面已经测试出使用静态阈值-最小迁移函数可以达到最好的节能效果。为验证静态阈值-最小迁移函数是否为任何集群规模下的最佳策略, 再以 800 个物理服务器与 1 052 个云任务进行模拟(图 4)。通过实验仿真, 可以在

不同集群规模下选择性能最好的算法, 然后进行迁移策略, 以达到最少的耗能。

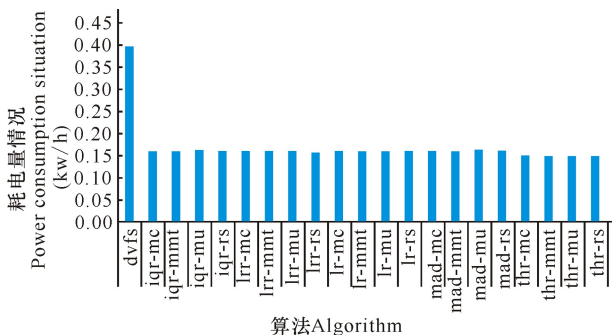


图 3 10-10 的服务规模下与传统节能方式耗能情况

Fig. 3 10-10 scale of services with traditional energy-saving mode of energy consumption

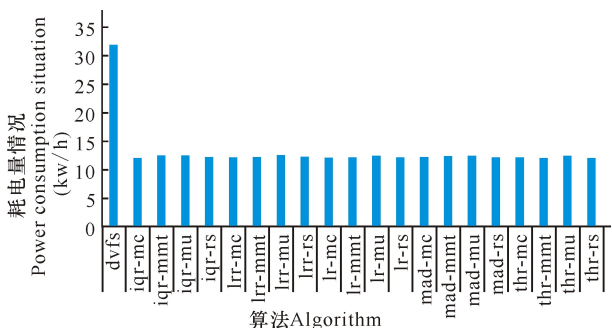


图 4 800-1052 的服务规模下与传统节能方式耗能情况

Fig. 4 800-1052 scale of services with the traditional energy-saving mode of energy consumption

从图 5 可以看到, 在规模为 1 000 的集群范围里, 使用局部加权的线性回归函数和随机选择的组合策略, 都能很好降低集群的能耗, 对比 CloudSim 自带的算法, 局部加权的线性回归函数在能源优化和服务等级上都比 CloudSim 有优势。

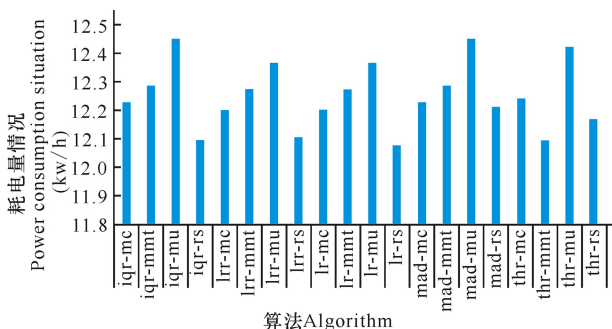


图 5 800-1052 的服务规模下不同算法耗能情况

Fig. 5 Energy consumption in different algorithms under 800-1052 service scale

从服务等级协议的延迟上来看(图 6), 改进的线性回归方法与最优的策略差距并不大, 而且优于部分策略结果(四分位法, 鲁棒性的线性回归)。并

且从该图可以看到,在选择策略上最小迁移时间的方法较好。

以上结果说明:不同规模下的虚拟机集群最优的迁移策略会不同,通过软件模拟的方式可以找出最优的迁移策略,并通过主动迁移,减低能源消耗;在800-1052的规模下改进的局部加权线性回归算法可以较好地节省能源,优于原有的检测算法;在选择主机算法中最小迁移时间的策略最优。

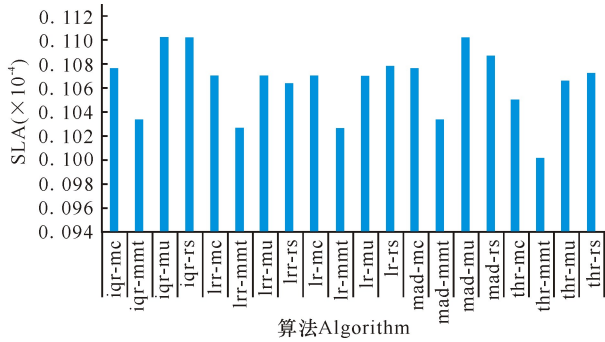


图6 800-1052的服务规模下SLA服务下降水平

Fig.6 SLA service decline under 800-1052 service scale

4 结论

本研究通过设计虚拟机在线迁移管理软件,对服务器的虚拟机服务进行管理,帮助用户和管理员更好地使用与维护虚拟机服务。通过对虚拟机迁移算法的研究建模并模拟实现不同规模下的能源优化策略。实验证明在较大的规模下,优化的线性回归函数可以实现更好的能源优化性能。

参考文献:

- [1] 黄晓雯. 云计算体系架构与关键技术[J]. 中国新通信, 2014(13):29.
HUANG X W. Cloud computing architecture and key technologies [J]. China New Communications, 2014 (13):29.
- [2] 刘海坤. 虚拟机在线迁移性能优化关键技术研究[D]. 武汉:华中科技大学,2012.
LIU H K. Key technologies of performance optimization for virtual machine live migration [D]. Wuhan: Huazhong University of Science and Technology, 2012.
- [3] HUANG W, LI X, QIAN Z Z. An energy efficient virtual machine placement algorithm with balanced resource utilization [C]//Proceedings of the seventh international conference on innovative mobile and internet services in ubiquitous computing. Taichung: IEEE, 2013:313-319.
- [4] YOUNGE A J, VON LASZEWSKI G, WANG L Z, et al. Efficient resource management for cloud computing environments [C]//Proceedings of 2010 international

- green computing conference. Chicago: IEEE, 2010: 357-364.
- [5] BELOGLAZOV A, BUYYA R. Energy efficient allocation of virtual machines in cloud data centers [C]// International conference on cluster, cloud and grid computing. Melbourne, Victoria, Australia, 2010.
- [6] MILLER F P, VANDOME A F, MCBREWSTER J. Median absolute deviation [M]. Dunfermline: Alphascript Publishing, 2010.
- [7] SEMERARO G, MAGKLIS G, BALASUBRAMONIAN R, et al. Energy-efficient processor design using multiple clock domains with dynamic voltage and frequency scaling [C]//Proceedings of the eighth international symposium on high-performance computer architecture. Cambridge, MA: IEEE, 2002: 29-40.
- [8] BELOGLAZOV A, BUYYA R. Optimal online deterministic algorithms and adaptive heuristics for energy and performance efficient dynamic consolidation of virtual machines in cloud data centers [J]. Concurrency and Computation Practice and Experience, 2012, 24 (13): 1397-1420.
- [9] LOADER C. Local regression and likelihood [M]. New York: Springer Science & Business Media, 1999.
- [10] 王舜. 基于最小迁移代价的云计算资源负载均衡调度算法研究 [D]. 成都: 电子科技大学, 2013.
WANG S. Research on resource load balancing scheduling algorithm based on minimum migration cost in cloud computing environment [D]. Chengdu: University of Electronic Science and Technology of China, 2013.
- [11] KIM K H, BUYYA R, KIM J. Power aware scheduling of bag-of-tasks applications with deadline constraints on DVS-enabled clusters [C]//Proceedings of the seventh IEEE international symposium on cluster computing and the grid. Rio De Janeiro: IEEE, 2007, 7: 541-548.
- [12] 杨星, 马自堂, 孙磊. 云环境下基于改进蚁群算法的虚拟机批量部署研究 [J]. 计算机科学, 2012, 39(9): 33-37.
YANG X, MA Z T, SUN L. Research on extended ant colony optimization based virtual machine deployment in infrastructure clouds [J]. Computer Science, 2012, 39(9): 33-37.
- [13] 曲晓雅. 负载感知的虚拟机初始化放置和迁移时机判决机制的研究 [D]. 北京: 北京交通大学, 2016.
QU X Y. Research on workload-aware virtual machines initial placement and migration decision [D]. Beijing: Beijing Jiaotong University, 2016.
- [14] 罗刚毅, 钱柱中, 陆桑璐. 一种基于网络感知的虚拟机再调度算法 [J]. 计算机学报, 2015, 38(5): 932-943.
LUO G Y, QIAN Z Z, LU S L. A network-aware VM re-scheduling algorithm [J]. Chinese Journal of Computers, 2015, 38(5): 932-943.