

云环境下基于改进 NSGA II 的虚拟机调度算法

殷小龙,李 君,万明祥

(南京邮电大学 计算机学院,江苏 南京 210003)

摘 要:在云环境中,如何将大量的虚拟机调度到物理节点上是一个基本且复杂的问题。文中首先对虚拟机的调度建立装箱问题模型,将该模型的求解转化一个多目标优化问题,目标分别为负载均衡、提高任务执行效率和降低能耗;接着对基于非支配排序的遗传算法(Non-dominated Sorting Genetic Algorithm, NSGA II)进行改进,利用回溯法中的剪枝函数确定最优初始种群,引入正态分布密度函数限制优秀精英。仿真结果表明,基于改进 NSGA II 的虚拟机调度算法在任务执行时间、负载均衡和能量消耗三个方面优于其他一些常用算法。

关键词:云计算;虚拟资源调度;装箱问题;多目标优化;NSGA II

中图分类号:TP301.6

文献标识码:A

文章编号:1673-629X(2014)08-0071-05

doi:10.3969/j.issn.1673-629X.2014.08.017

Virtual Machines Scheduling Algorithm Based on Improved NSGA II in Cloud Environment

YIN Xiao-long, LI Jun, WAN Ming-xiang

(College of Computer, Nanjing University of Posts and Telecommunications,
Nanjing 210003, China)

Abstract: In cloud environment, how to schedule the large number of virtual machines to the physical nodes is a fundamental and difficult problem. Firstly, establish bin packing problem model based on virtual machines scheduling and solves the model through transforming it into a multi-objective optimization problem. The objectives respectively are load-balancing, to improve task-execution efficiency and to reduce energy consumption. Then, non-dominated sorting genetic algorithm is improved. The pruning function in the backtracking is used to confirm the optimal initial population. The normal distribution density function is introduced to restrict elite. The results of simulation show that the virtual machines scheduling algorithm based on improved NSGA II on the aspects of task execution time, load-balancing and energy consumption is better than other commonly used algorithms.

Key words: cloud computing; virtual resources scheduling; bin packing problem; multi-objective optimization; NSGA II

0 引 言

近年来,云计算已得到工业界和学术界的普遍认可,已逐渐由研究转向应用,大量商业机构建立了自身的云服务平台^[1]。对于这些云服务提供商来说,其核心技术是如何合理地将用户申请的虚拟机映射到相应的物理节点上,即虚拟机调度。而由于用户需求的实时动态变化很难准确预测,以及需要考虑系统性能和成本因素使得虚拟机调度问题变得非常复杂^[2]。由此可见,虚拟机调度策略的优劣直接影响到系统的整体资源利用率、能量消耗以及服务质量。针对虚拟机调度,业内各大机构和学者已经做了大量的研究工作,采

用的方法可分为简单调度策略和基于启发式算法进行全局寻优策略^[3]。

目前,业内各大 IT 公司建立的云计算数据中心基本采用一些简单的调度策略来实现虚拟机的调度。Amazon EC2 的调度目标是负载均衡,调度中心采用简单的轮转(Round Robin)方式将所有虚拟机调度到物理节点上^[4]。HP 的调度目标是公平分配,采用先来先服务(FIFO)的方式为优先被选择的虚拟机调度至物理节点^[5]。简单的虚拟机调度策略已经很难满足日益增长的云计算的规模,因此不少学者开始采用启发式方法来进行全局优化搜索最佳虚拟机调度方案。文献[6]利用博弈论的方法来解决云计算资源的

收稿日期:2013-09-24

修回日期:2013-12-27

网络出版时间:2014-04-24

基金项目:江苏省科技支撑计划(BE2012849);江苏省研究生科研创新计划项目(CXLX12_0481)

作者简介:殷小龙(1989-),男,江苏盐城人,硕士研究生,研究方向为资源管理、云计算。

网络出版地址: <http://www.cnki.net/kcms/detail/61.1450.TP.20140424.0824.066.html>

负载均衡问题,提出了一种基于博弈论的虚拟机调度算法,给出一种较好的折衷虚拟机调度方法来保证系统的负载均衡,但是算法的时间复杂度较高,使得用户 QoS 需求得不到充分保证。文献[7]提出一种基于免疫克隆的云资源调度优化算法,对用户应用偏好和多维 QoS 中的用户效用进行了量化,给出了多维 QoS 优化的目标函数,采用免疫克隆算法进行寻优,该算法充分考虑了用户的 QoS,却忽略了系统的能量消耗。文献[8]提出了一种最佳配合启发式(Best Fit Heuristic, BFH)算法,依次将虚拟机调度至当前物理节点,直到节点的资源不足,然后选择下一新的节点。此算法的不足在于忽略了资源过载导致对任务执行效率的影响。文献[9]提出一种基于内存感知的服务器整合算法,将虚拟主机聚合在少数物理节点上以达到节能的目的,但是该文献没有充分考虑服务器间的负载均衡,因为物理服务器的性能会随着负载的不均衡而降低。文献[10]在研究现有云计算服务调度算法的基础上,设计了基于 QoS 服务的调度算法。该算法兼顾用户需求 and 系统整体性能,但是其能耗消耗过大。文献[11]提出一种基于多目标进化算法的虚拟资源分配策略,使用 NSGA II 求解该优化问题,但是该文献仅解决的是云数据中心的能耗问题,只是从多个角度考虑能耗而已。

综上所述,现有的虚拟机调度策略研究主要存在以下两个问题:

(1)简单的调度策略易于实现,但是对于云计算环境下的大量物理节点,执行效率则较低;

(2)基于启发式算法的调度策略只能从用户需求、系统性能和成本因素的单个方面考虑,或者考虑的只是单个方面的多个目标而已。

因此,针对上述问题,文中提出一种基于改进 NSGA II 的虚拟机调度算法,通过对云环境下虚拟机调度的分析,建立云计算中虚拟机调度的装箱模型,将求解装箱问题的贪心算法与多目标优化的 NSGA II 算法有效地结合来解决虚拟机的多目标调度。改进的 NSGA II 算法利用回溯法中的剪枝函数确定最优初始种群,对精英保留策略的改进使得算法避免陷入局部最优。仿真结果表明,与其他虚拟机调度算法相比,在相同目标的情况下文中提出的算法具有相对较快的任务执行速度和较少的能量消耗。

1 虚拟机调度的建模

虚拟机的调度是将虚拟机(物品)合理地装载到物理机(箱子)中,其中的载入方式对应一种调度方案,因此可描述为一个多维装箱问题,装入的物品为虚拟机,虚拟机所用资源是物品的大小,箱子是物理节

点,箱子容量是节点资源的使用阈值。文中考虑的资源包括 CPU、内存、网络带宽三种,为此建立虚拟机调度的三维装箱模型。装箱问题的目标是要确定一系列可行的装载方案,使得在满足给定装载容量约束的条件下,每个箱子中所装物品的体积尽可能的大。但是由于装箱问题的复杂性,使其在寻优方面的能力较弱,且针对虚拟资源调度的问题,需要配合一种多目标寻优算法才能取得比较好的效果,下面首先对其多个目标进行数学描述。

目标 1:任务执行时间最少,即 $\min \sum_i \sum_j T_{ij}$ 。

目标 2:产生的能耗最小,即 $\min \sum_i \sum_j E_{ij}$ 。

目标 3:负载均衡系数最小,即 $\min LB$ 。

约束条件:物理资源足够,即 $C_{jk} \leq 1$ 。

其中,下标 i 为虚拟机编号, $i \in \{1, 2, \dots, m\}$; j 为物理机编号, $j \in \{1, 2, \dots, n\}$; k 为资源编号, $k \in \{1, 2, \dots, s\}$, 分别包括 CPU、内存、磁盘 I/O、网络带宽等资源。 T_{ij} 表示虚拟机 i 上的任务在物理机 j 上的执行时间,由公式(1)计算得到。 E_{ij} 表示虚拟机 i 上的任务在物理机 j 上执行所产生的能耗,参照文献[12]的能耗模型可由公式(2)计算得到。 LB 表示虚拟机调度方案的系统负载均衡系数,参照文献[13]的负载均衡模型可由公式(3)计算得到。 C_{jk} 表示物理机 j 上的资源 k 的使用率。

$$T_{ij} = I_i \times S_j \quad (1)$$

其中, I_i 表示任务的指令条数(单位为 MI); S_j 为处理器速度(单位为 MIPS)。

$$E_{ij} = E_j^{\text{cpu}} + E_j^{\text{mem}} + E_j^{\text{net}} \quad (2)$$

其中,静态能耗由 CPU 能耗、内存能耗和网络能耗组成。

$$LB = \sqrt{\frac{\sum_{j=1}^n (\text{Load}(j) - \bar{L})^2}{n}} \quad (3)$$

其中, $\text{Load}(j)$ 表示物理机 j 的综合负载,可由公

式(4)计算得到; n 为物理机的台数; $\bar{L} = \frac{\sum_{j=1}^n \text{Load}(j)}{n}$ 为各物理机综合负载的均值。

$$\text{Load}(j) = \sum_{k=1}^s A_k L_k \quad (4)$$

其中, A_k ($\sum A_k = 1$) 为各项负载的权重,用来表示其重要程度; L_k 为各资源的使用率,包括 CPU、内存和带宽。

2 基于改进 NSGA II 的虚拟机调度算法

NSGA II 是多目标优化算法中的经典算法,以非

支配个体排序和精英保留机制为其主要特点^[14],文中将 NSGA II 算法与贪心法相结合来解决装箱问题,其基本思想是:使用贪心法生成一系列的装箱模式作为初始解,然后采用 NSGA II 算法作全局最优点的探索以寻求满足多目标的虚拟资源调度方案。但是 NSGA II 算法存在缺陷,因此,文中对 NSGA II 算法进行改进,利用回溯法中的剪枝函数确定最优初始种群,并对精英保留策略进行改进使得算法避免陷入局部最优。

2.1 染色体编码

常见的染色体编码方式有二进制编码、字符编码和实数编码等。文中采用实数编码方式来表示装箱问题的染色体。染色体的长度为当前需调度的虚拟机数量,染色体的每一位为虚拟机的编号,用数据结构 (i, l, w, h, x, y, z) 表示,其中 i 为箱子的编号,对应物理机的编号, l, w, h 为物品的长、宽、高,表示虚拟机的 CPU、内存和带宽, x, y, z 为物品在箱子中的相对坐标。一个染色体代表了一种调度策略,用 X_k 表示。

2.2 初始种群确定

NSGA II 算法对初始群体依赖性较大,在初始群体的确定过程中,一般都是采用随机选取,致使初始群体优越性差距很大,直接影响到最优结果。因此,文中对初始群体确定的过程建立状态空间树并使用限界函数进行剪枝,提出一种基于回溯法的个体寻优算法,不仅保证了个体的优秀,同时也在时间复杂度上远远优于穷举法。

回溯法本质是一种以深度优先方式,逐一生成状态空间树的节点并检测答案节点的方法。文中所构建的状态空间树中的节点表示物理机,每一条枝表示调度到该节点的虚拟机,根节点到叶子节点的一条全路径则表示一种可行调度方案。文中提出的基于回溯法的个体寻优算法以负载均衡系数为限界函数进行剪枝,如若该节点上部署虚拟机后负载均衡系数变高则剪去该枝,继续从根节点开始检测下一个调度方案。最后将所有根节点到叶子节点的全路径对应的调度方案作为初始种群。

下面以 3 台虚拟机为例进行说明。

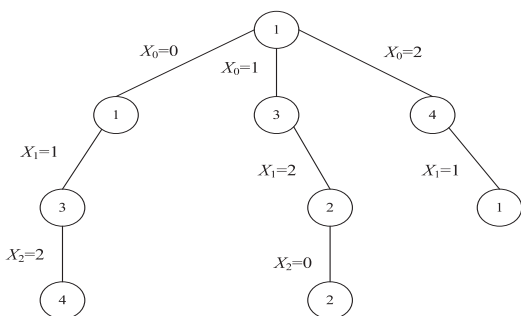


图1 3 台虚拟机调度问题的状态空间树

如图1所示,有3台虚拟机需要调度,其中每一条

根节点到子节点的路径代表正在构造的调度方案的一部分,约束条件为物理机的负载均衡系数要得到降低,文中采用此约束函数剪去负载均衡系数提升的状态子树,从而保留优异的种群,图中只有2条为根节点到叶子节点的全路径,即初始种群有两个染色体。

2.3 精英保留策略的改进

传统的 NSGA II 算法采用隐性精英保留策略,可以明显地提高算法的收敛速度。但是这种精英保留策略可能会造成进化过程中大部分非支配解都处于第一层的非支配曲面,会造成种群的提前收敛或收敛于局部最优。为了确保算法最终能收敛到较优的 Pareto 曲面和得到均匀分布的解,文中设计了一个改进的精英保留策略。该策略通过引入一个分布函数来限制精英解的数量,并且不能够遗弃所有的非精英解,该分布函数要能随着支配等级的增加平滑的递减。由于一维标准正态分布的密度函数 $f(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$ 在 $x>0$ 时具有平滑递减的特性,因此文中对此函数加以改进,如公式(5)所示,随着非支配等级 i 的递增 $e^{-\frac{i^2}{2}}$ 从1开始平滑的递减,从而限制精英解的数量且不遗弃所有的非精英解。

$$n_i = F_i \cdot e^{-\frac{i^2}{2}} \quad (5)$$

其中, $i(i \geq 1)$ 为非支配等级; F_i 表示第 i 层非支配解集的初始数量; n_i 表示在第 i 层非支配解集 F_i 上选取的个体的数量。

2.4 基于改进 NSGA II 的虚拟机调度流程

首先对由启发式方法生成一系列的装箱模式进行编码,初始解集为 P_0 ,种群规模为 N ,最大迭代次数为 t_{\max} ,算法伪代码如下:

```

Qi = RBacktrack(Pi)    // 确定初始种群
Ri = Pi ∪ Qi
F = QuickSort(Ri)      // 非支配排序分层
for(每一层级 Fi) {      // 精英保留
    Congestion[i]        // 计算 Fi 中个体的拥挤度
    IndividNum[i]        // 根据公式(5)计算 Fi 中个体数量
}
Pt+1 = TSelection()      // 锦标赛选择
Temp = MCrossover(Pt+1) // 矩阵交叉
Qt+1 = RReplacement(Temp) // 随机替换变异
t = t + 1
t > tmax 算法终止
Output(Pareto 最优解集)
  
```

3 仿真实验与性能分析

3.1 实验的仿真环境

文中使用 CloudSim2.1.1 建立仿真环境进行实

验,开发环境为 Eclipse。实验的机器配置为:CPU Intel Core2 双核 T5750 @ 2.0 GHz;内存 2 GB;硬盘 160 GB;操作系统 Windows 7 旗舰版 32 位 SP1。

仿真中创建了 1 个数据中心,包括 100 台物理主机,然后按照 1:5 的比例生成虚拟机,实验中的任务随机生成,大小为 1 000 ~ 4 000 MI,由于文中研究的是虚拟机到物理机的调度,故将任务与虚拟机通过轮转(Round Robin,RR)的方法依次绑定。为了体现云计算的异构特性,物理机和虚拟机的各个属性随机生成。

3.2 虚拟资源调度的仿真结果和性能对比分析

下面分别对文中提出的调度算法、基于 NSGA II 的虚拟机调度算法、最佳配合启发式虚拟机调度算法^[9]和虚拟资源分配优化算法^[12]的任务完成总时间、负载均衡系数及其任务执行能耗进行对比。

表 1 列举了任务数为 500 时,算法最终的 10 个 Pareto 最优解,从表中可以看出 INSGA II 算法通过一次执行可获取多个不同的解。因此,对应于文中的三个目标:任务执行时间最短、负载均衡和任务执行能耗最低,实际使用中可以根据不同的需要进行选择,当需要协调这三个目标时,可以选取三个目标值之和最小的方案。

表 1 Pareto 最优解集

方案	任务执行时间/s	负载均衡系数	能耗/kJ
方案 1	217.78	0.19	396
方案 2	212.34	0.21	412
方案 3	200.05	0.35	480
方案 4	204.75	0.32	455
方案 5	220.32	0.14	409
方案 6	209.86	0.28	515
方案 7	204.33	0.15	511
方案 8	212.25	0.11	583
方案 9	209.73	0.15	443
方案 10	255.35	0.15	477

由于各目标函数之间的矛盾性,使它们同时得到最优解的可能性较小,因此在实际应用时,决策者应该根据优化重点从 Pareto 解集中选择最优的方案配置,而简单的单目标算法则无法达到这样的要求,比如:以任务执行时间最小为主要目标函数时,选择方案 3;以负载均衡为主要目标函数时,选择方案 8;以最低能耗为主时,选择方案 1;均衡考虑这三个目标,选择较优配置时,可以选择方案 9。下面给出文中的 INSGA II 算法与其他算法的性能比较。

任务的执行时间反映了虚拟资源调度算法的执行效率。文中比较了拥有 100 个物理主机 500 个虚拟机的数据中心,在任务从 100 到 1 000 变化时,INSGA 算法、NSGA 算法、MVRA 算法和 BFH 算法的完成时间。

仿真结果如图 2 所示。

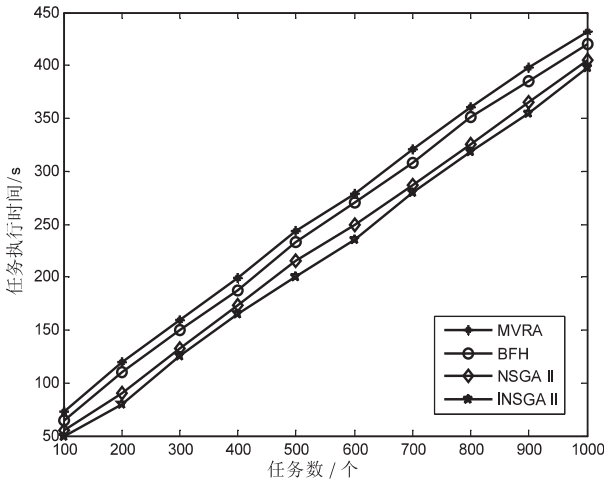


图 2 不同任务数量下执行时间对比

从图 2 可以看出,随着任务数量的不断增长,虚拟资源与物理节点的匹配更加复杂,四种算法的任务执行时间也随之增长且较为接近。MVRA 和 BFH 算法的任务完成时间始终较高,这是由于 MVRA 算法是通过降低 CPU 的频率和供应电压来达到节能的目的,因此延长了任务的执行时间。BFH 算法是将当前虚拟机调度至具有最小剩余资源量的可行物理节点上,多个任务共同竞争较少的资源,直接导致任务总执行时间增大。文中的 INSGA 算法任务执行时间低于 NSGA 算法,这是由于 INSGA 算法对初始种群的生成进行了改进,解决了原算法遗失精英个体的问题,使其个体更加优秀,从而缩短了任务的执行时间。

负载均衡反映了系统的稳定情况。文中比较了拥有 100 个物理主机 500 个虚拟机的数据中心,在任务从 100 到 1 000 变化时,INSGA 算法、NSGA 算法、MVRA 算法和 BFH 算法的负载均衡系数,仿真结果如图 3 所示。

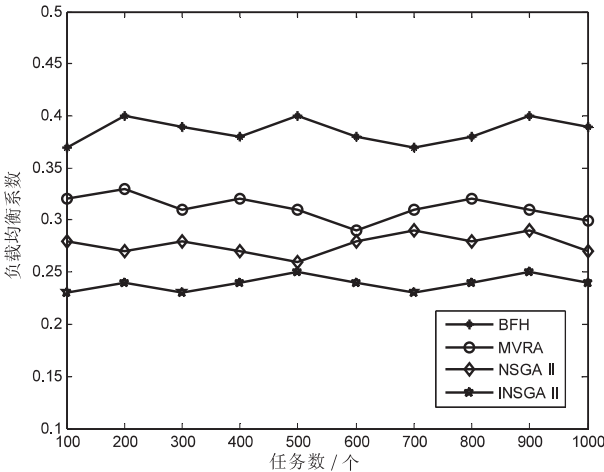


图 3 不同任务数量下负载均衡系数对比

从图 3 可以看出,随着任务数的变化,BFH 算法的负载均衡系数最大,这是由于虚拟机被调度到较少的

物理机上,导致系统的负载严重不均衡。相比之下 MVRA 算法的均衡系数较之略低,这是由于关闭了空闲的物理机,工作中的物理机间的负载相对均衡。NSGA 算法的均衡系数低于 MVRA 算法是由于将负载均衡作为目标之一,种群不断迭代最终在 Pareto 解集中选择了最优方案。文中的 INSGA 算法相比于改进前的 NSGA 算法可获得更加优秀的个体,因此系统的负载均衡系数最低。

系统的能耗能够反映节能效果,文中比较了拥有 100 个物理主机 500 个虚拟机的数据中心,在任务从 100 到 1 000 变化时,INSGA 算法、NSGA 算法、MVRA 算法和 BFH 算法的系统能耗,仿真结果如图 4 所示。

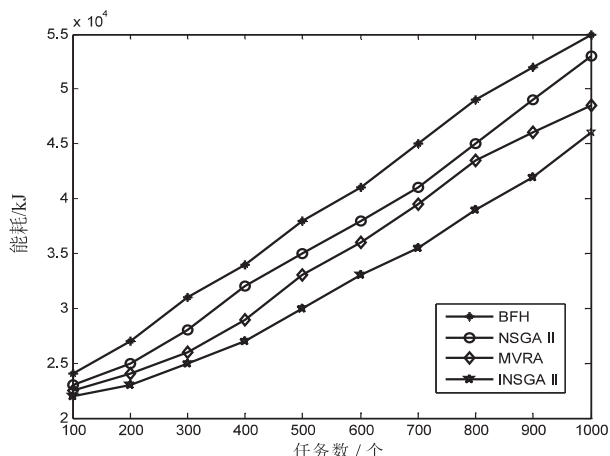


图4 不同任务数量下系统能耗对比

从图4可以看出,BFH算法的系统能耗最高是由于处理能力较快的节点上资源不足造成任务执行时间增加且处理能力较慢的节点有可能处于闲置状态。MVRA算法的系统能耗低于NSGA II算法,是由于将虚拟机调度至尽可能少的物理机上且关闭空闲的物理机,该算法还通过动态调压调频技术(Dynamic Voltage and Frequency Scaling, DVFS)调节CPU速率已达到节能。文中提出的INSGA II算法将能耗最小设为目标之一,在精英保留和交叉操作两方面对NSGA算法做了改进,不仅解决了原算法提前收敛的问题,而且其选取的精英更为优秀,从而能耗低于其他三种算法。

由仿真结果可以得出结论,INSGA算法不仅在执行效率高于BFH、MVRA和NSGA算法;在节能方面也远好于其余三种算法,关键是管理者针对不同的目标可以选择最佳的方案。可见,文中提出的虚拟资源调度算法更适用于云环境。

4 结束语

文中对云计算中的虚拟机调度进行了研究,提出了一种基于改进NSGA II的虚拟机调度算法,在用户需求、系统性能和成本因素三个方面为决策者提供最

优的选择,以便在任务执行时间最短、负载均衡和任务执行能耗最低三个目标间权衡。仿真结果表明,该算法和目前云计算常用的一些算法相比在任务执行效率、负载均衡和节约能耗方面都有明显的优势。如何从文中的最终Pareto解集中选出各目标的均衡决策方案是下一步的工作重点。

参考文献:

- [1] Armbrust M, Fox A, Griffith R, et al. A view of cloud computing[J]. Communications of the ACM, 2010, 53(4): 50-58.
- [2] 张建成, 宋丽华, 鹿全礼, 等. 云计算方案分析研究[J]. 计算机技术与发展, 2012, 22(1): 165-167.
- [3] Khazael H, Misic J, Misic V B, et al. Analysis of a pool management scheme for cloud computing centers[J]. IEEE Transactions on Parallel and Distributed Systems, 2013, 24(5): 849-861.
- [4] 高邦仁. “云”中的亚马逊[J]. 互联网周刊, 2011(17): 8-10.
- [5] Dukaric R, Juric M B. Towards a unified taxonomy and architecture of cloud frameworks[J]. Future Generation Computer System, 2013, 29(5): 1196-1210.
- [6] Teng F, Magoules F. A new game theoretical resource allocation algorithm for cloud computing[C]//Proc of advances in grid and pervasive computing. [s. l.]: [s. n.], 2010: 321-330.
- [7] 孙大为, 常桂然, 李凤云, 等. 一种基于免疫克隆的偏好多维QoS云资源调度优化算法[J]. 电子学报, 2011, 39(8): 1824-1831.
- [8] 李强, 郝沁汾, 肖利民, 等. 云计算中虚拟机放置的自适应管理与多目标优化[J]. 计算机学报, 2011, 32(12): 2253-2264.
- [9] Jang Jae-Wan, Jeon M, Kim Hyo-Sil, et al. Energy reduction in consolidated servers through memory-aware virtual machine scheduling[J]. IEEE Transactions on Computers, 2011, 60(4): 552-564.
- [10] 郭怡, 茅苏. 云计算下基于CRP算法的资源提供策略[J]. 计算机技术与发展, 2012, 22(2): 80-84.
- [11] 许力, 曾智斌, 姚川. 云计算环境中虚拟资源分配优化策略研究[J]. 通信学报, 2012, 33(Z1): 9-16.
- [12] Lee Y C, Zomaya A Y. Energy efficient utilization of resources in cloud computing systems[J]. The Journal of Supercomputing, 2012, 60(2): 268-280.
- [13] Randles M, Lamb D, Taleb-Bendiab A. A comparative study into distributed load balancing algorithms for cloud computing[C]//Proc of 2010 IEEE 24th international conference on advanced information networking and applications workshops. Perth, WA: IEEE, 2010: 551-556.
- [14] 许波, 彭志平, 余建平, 等. 基于云模型的NSGA-II算法改进[J]. 小型微型计算机系统, 2012, 33(7): 1599-1602.

云环境下基于改进NSGA II的虚拟机调度算法

作者：[殷小龙](#)，[李君](#)，[万明祥](#)，[YIN Xiao-long](#)，[LI Jun](#)，[WAN Ming-xiang](#)

作者单位：[南京邮电大学 计算机学院](#)，[江苏 南京](#)，[210003](#)

刊名：[计算机技术与发展](#)

英文刊名：[Computer Technology and Development](#)

年，卷(期)：[2014\(8\)](#)

引用本文格式：[殷小龙](#)，[李君](#)，[万明祥](#)，[YIN Xiao-long](#)，[LI Jun](#)，[WAN Ming-xiang](#) [云环境下基于改进NSGA II的虚拟机调度算法](#)[期刊论文]-[计算机技术与发展](#) 2014(8)