云计算环境下的节能任务调度策略的随机 Petri 网分析

赵 彬 王 淖 王高才
(广西大学计算机与电子信息学院 南宁 530004)

摘 要 针对当前云计算异构服务器集群环境下的高能耗问题,提出一种最小能耗优先的任务调度策略(first scheduling with minimum energy)。该策略在调度任务时优先考虑处于运行状态的服务器,并以任务响应时间为约束,按照最小能耗原则将任务分配到相应的服务器上执行。当处于运行状态的服务器都不能满足任务对响应时间的要求时,则考虑处于休眠状态的服务器;同时,也基于最小能耗原则进行调度。采用随机 Petri 网工具对节能任务调度策略进行分析建模,考虑其能耗和相关性能指标。实验结果表明,该方法不仅能满足任务的 QoS 性能要求,而且具有较好的节能效果。

关键词 云计算,任务调度,节能,QoS,随机 Petri 网
中图法分类号 TP393 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2015.8.024

Analysis on Energy-saving Task Scheduling Strategy Based on Stochastic Petri Net for Cloud Computing

ZHAO Bin WANG Nao WANG Gao-cai
(School of Computer and Electronic Information, Guangxi University, Nanning 530004, China)

Abstract For the current high energy consumption problem in cloud computing, the paper proposed a task scheduling strategy, called first scheduling with the minimum energy (FSME). The strategy firstly considers working states of servers when the tasks are scheduled. And then the algorithm schedules the tasks to corresponding severs in terms of the minimum energy rules with respond time constraint. If all of the working servers can not meet the response time requirement of the current task, the algorithm will consider the idle servers and schedule the task to the lowest execution energy consumption server. Stochastic Petri net was used to model the algorithm and analyze the energy consumption and the performance. Simulation results show that the FSME can improve the energy efficiency while meeting the quality of service requirement.

Keywords Cloud computing, Task scheduling, Energy saving, QoS, Stochastic Petri net

1 引言

云计算作为一种新型的计算模式,自出现便受到了广泛的重视。目前,云计算的相关主题已成为学术界的研究热点。在产业界,国内外许多大型公司都建立了自己的云计算服务平台。国外的如 IBM 的"蓝云"计算平台、Google 的云计算平台、Amazon 的弹性计算云等[1];国内的如百度云存储、各大电商的云服务等。云计算在得到飞速发展的同时也带来了不少问题,统计结果显示:美国的数据中心在 2011 年全年的耗电量是 1000 亿千瓦时,而电费则高达 74 亿美元[4]。《The Wall Street Journal》指出,能耗费用占一个云计算数据中心运维成本的 40%。因此,云计算系统中的高能耗已经成为一个亟待解决的严重问题。

云计算系统产生高能耗的主要原因在于:1)大量运行的服务器处于空闲等待状态。用户提交任务的随机性导致云计算集群中的服务器会出现空闲情况,而当前的云计算服务提供商为了保证服务质量,通常在没有可执行任务的情况下也将服务器置于等待用户任务到来的状态,即空闲等待状态。但是服务器处于空闲等待状态时,其耗电功率会达到峰值功率的 50%~60%,造成极大的能源浪费。2)未充分考虑基于能耗因素的任务调度策略。研究结果表明,不同任务在异构计算机上的执行功耗不同。所以,只考虑性能因素而将任务调度到能耗大的服务器上执行,势必会造成不必要的能量浪费。

目前,关闭/休眠技术是能耗优化管理的重要技术之一,其主要方法是通过将空闲服务器置入休眠状态来降低能耗,优点是能最大程度地降低空闲能耗。Liao 等[5]提出了一种性能和能耗优化的调度方法,该方法尽可能地将任务安排在连续时间内执行,从而使得设备尽可能长时间地处于休眠状态,同时减少各状态之间切换的次数。RODERO I 等[8]提出了一种基于虚拟化计算平台的动态负载感知节能方法,通过关闭不需要的主机子系统来实现节能,该方法在保证服务质

到稿日期:2014-09-22 返修日期:2015-01-22 本文受国家自然科学基金(61262003),广西自然科学杰出青年基金(2013GXNSFGA019006)资助。

赵 彬(1989-),男,硕士生,主要研究领域为网络节能、随机 Petri 网应用;**王 淖**(1977-),女,硕士,讲师,主要研究领域为网络优化;**王高才**(1976-),男,博士,教授,博士生导师,CCF 高级会员,主要研究领域为计算机网络技术、系统性能评价和随机方法,E-mail:wanggcgx@163.com。

量的前提下可有效提高能耗效率。肖志娇等[9]针对集群环境下服务器状态能耗进行研究，提出基于状态管理的服务器能耗优化方法，即将部分服务器置于就绪或运行状态用以处理稀疏到达的用户请求，其余服务器置于休眠状态，同时根据请求到达的时间间隔动态设置就绪或运行服务器的数量。另外，为了解决抖动问题，还设置了唤醒阈值，当请求数量超过唤醒阈值时才唤醒休眠服务器。

在云计算集群服务器异构的环境下，通过合理的调度策略将用户任务调度到执行能耗较小的服务器上执行，可以有效地实现节能。谭一鸣等[10]对云计算系统的空闲能耗和"奢侈"能耗展开研究，分析了云计算系统平均响应时间和平均功率，并建立系统的能耗模型，提出一个满足性能约束的最小期望执行能耗调度算法 ME^3PC。该算法根据所有服务器的负载情况不同地调度策略，可在保证执行性能的情况下大幅降低云计算系统的能耗。李钰等[13]对 Min-Min 算法进行改进，提出一种节能任务调度算法 ECMM。ECMM 算法对异构集群中的每个节点应用相同的休眠状态控制策略，通过控制节点进入不同的休眠状态来减少能耗。何丽等[14]研究了云计算异构服务器集群环境下的节能问题，提出一种基于能耗优化的最早完成时间任务调度算法。该算法以经典 Min-Min 算法为基础，选取最短任务并优先考虑调度到拥有最小执行能耗的服务器上，如果该服务器满足任务对完成时间的要求则予以调度，否则采用 Min-Min 算法进行任务调度。

目前对于关闭/休眠技术大多是针对关闭/休眠时机的设定和预测进行研究，而在任务调度策略的研究上，通常是以提高系统性能为优化目标。而对于如何通过任务调度使更多的服务器空闲以便置入休眠状态，以及如何根据服务器的不同状态设置合理的任务调度策略从而实现节能的研究较少。

Petri 网[15]是一个应用较多的数学模型工具，适于描述分布的、并发的、异步的系统。它既有严格的数学表达方式和直观的图形表达方式，又有丰富的系统描述手段和系统行为分析技术。把随机时间参数引入 Petri 网中就形成了随机 Petri 网。由于随机 Petri 网的状态可达图同构于连续时间的马尔可夫链，因此可以借助于马尔可夫过程的相关理论分析系统的性能参数，这就为系统的分析求解带来了极大的方便。因此，本文使用随机 Petri 网工具对系统模型进行能耗和性能的分析。

本文针对云计算平台中的能耗问题展开研究，提出一种节能任务调度方法；使用随机 Petri 网工具对系统进行建模，并利用排队论及马尔可夫链等技术分析系统的能耗和性能；最后通过实验验证所提节能方法的有效性。

2 系统模型

2.1 任务模型

用户操作在时间上是随机的，导致用户任务到达云计算系统的时间也呈现出随机性。例如，在实际系统模型中，一般假定任务的到来是泊松分布的。另外，用户任务的到达具有一定的规律性。例如，某些资源在日常工作时间用户的请求比较多，而到了晚上或节假日时间请求会减少很多。而某个热门事件的出现引起大量用户的关注，会导致相应的用户任务在短时间内大量涌现，待热度下降之后相关的用户任务数又会迅速下降。此外，云计算系统提供给用户的服务是多种多样的，当前云计算服务可以分为 3 个层次：基础设施即服务(IaaS)、平台即服务(PaaS)、软件即服务(SaaS)。云计算服务的多样性使得用户提交任务的类型多种多样。不同类型任务的计算量、对资源的需求以及对 QoS 的要求等有所不同。本文假定同一类型任务的计算量、QoS 要求等是相同的。

定义 1 用户任务可以定义为一个三元组(T,RT,M)，其中，$T=\{t_i \mid 1\leqslant i\leqslant m\}$表示任务集合，$t_i$ 表示第 i 个任务，不同任务之间相互独立；$RT=\{RT_i \mid 1\leqslant i\leqslant m\}$表示用户任务对系统响应时间要求的集合，$RT_i$ 表示任务 t_i 对系统响应时间的要求；$M=\{m_i \mid 1\leqslant i\leqslant m\}$表示任务类型的集合，$m_i$ 表示任务 t_i 的类型，不同任务可以有相同的任务类型。本文不再对用户任务类型作另外的定义和公式化描述。

2.2 云计算系统模型

云计算系统的硬件基础设施是大量廉价服务器所组成的服务器集群，这些廉价服务器之间存在异构性。其原因在于，服务器是由不同厂商所生产提供的，具有不同的硬件配置，导致功能、性能以及耗电功率等有所不同。例如，不同服务器的峰值功率和空闲功率不同，不同任务在同一服务器上的执行功率不同。云计算服务器集群的异构性也为节能减排带来了更大的机会，我们可以利用这个特性合理地设计任务调度策略来达到降低能耗的目的。另外，本文所讨论的云计算环境中的服务器都具有休眠功能，可以通过设置或者在满足一定条件时进入休眠状态，当唤醒条件满足时自动完成透明的状态切换。

定义 2 云计算系统可以定义为一个四元组$(S,P^{peak},P^{idle},Z,P^{sleep})$，其中，

- $S=\{s_i \mid 1\leqslant i\leqslant n\}$表示云计算系统中的服务器集合，$s_i$ 表示第 i 个服务器。
- $P^{peak}=\{p_{ji}^{peak} \mid 1\leqslant i\leqslant n, 1\leqslant j\leqslant m\}$表示服务器执行任务的峰值功率集合，$p_{ji}^{peak}$ 表示服务器 s_i 执行任务 t_j 的峰值功率。
- $P^{idle}=\{p_i^{idle} \mid 1\leqslant i\leqslant n\}$表示服务器的空闲功率集合，$p_i^{idle}$ 表示服务器 s_i 的空闲功率。
- $Z=\{\lambda_{ji} \mid 1\leqslant i\leqslant n, 1\leqslant j\leqslant m\}$表示服务器 s_i 对任务的服务速率集合，λ_{ji} 表示服务器 s_i 对用户任务 t_j 的服务速率，这里 λ_{ji} 等于服务器 s_i 对该任务 t_j 所属类型任务的平均服务速率。
- $P^{sleep}=\{p_i^{sleep} \mid 1\leqslant i\leqslant n\}$表示服务器的休眠功率集合，$p_i^{sleep}$ 表示服务器 s_i 处于休眠状态时的功率。

云计算系统中用户任务的到达为泊松过程。系统接收用户提交来的任务，暂时将其放入任务缓冲队列中，全局的任务资源管理部件对系统资源进行统一的管理，如虚拟机的创建、销毁等，并根据所设定的任务调度策略统一将缓冲队列中的任务调度到相应服务器(虚拟机)上执行。服务器接收任务并放入本地缓冲队列，按照本地调度策略对任务进行服务。因此，得出的云计算系统模型如图 1 所示。

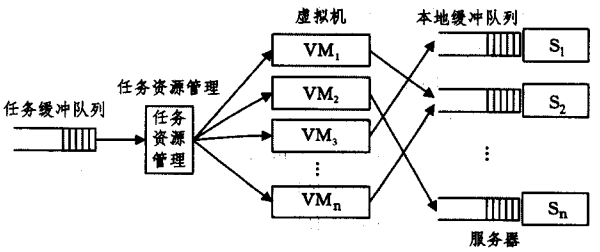

图 1 云计算系统模型

3 节能任务调度算法

根据前文分析，云计算系统产生高能耗的原因是大量服务器空闲等待和不合理的任务调度。因此本文从这两方面出发，来研究如何降低云计算系统的能耗。

当服务器本地任务队列中没有任务在等待时，服务器将处在空闲等待状态。大量服务器的空闲等待会产生极大的能源消耗，可将这些空闲服务器置于休眠状态，以减少其能耗。另外，使尽可能多的服务器空闲出来并休眠可以进一步减少系统的总能耗。对于将服务器切换到休眠状态的时机，目前有较多的研究。如根据任务到达时间间隔的历史记录预测下一任务到达的时间间隔，如果时间间隔足够大，就将服务器切换到休眠状态。这方面内容非本文研究点，为便于研究，对这方面内容做出相应假定，即当服务器本地任务队列中没有任务等待时，直接将服务器从空闲状态切换到休眠状态。

云计算系统中服务器的异构性使得同一任务在不同服务器上的执行能耗不同。将任务调度到执行能耗小的服务器上会获得较低的执行能耗。另一方面，处于不同状态的服务器，其状态之间的转换需要消耗一定的能量。频繁的状态转换会带来较高的能耗，因此考虑将任务优先调度到处于运行状态的服务器上，以减少状态转换频率。但是这样可能会出现一种极端的现象，即大量的任务都调度到同一台服务器上，使得该服务器负载过重，系统整体的服务质量下降。为了解决这一问题，在任务调度之前进行性能的判断，即检测将任务调度到当前服务器上能否满足任务对响应时间的要求，如果满足则予以调度，否则不予调度。

根据以上分析，本文设计了一种最小能耗优先调度算法FSME。算法的主要思想是：首先遍历所有服务器的运行状态，生成运行状态的服务器集合 Q_{run} 和休眠服务器集合 Q_{sleep}。任务调度时，根据任务 t_i 在集合 Q_{run} 中所有服务器上的执行能耗 w_{ji} 的大小，对 Q_{run} 中的服务器进行从小到大的排序，得到 Q'_{run}。检测 Q'_{run} 中的第一台服务器 $Q'_{run}[1]$ 能否满足任务 t_i 对响应时间 RT_i 的要求，如果满足则将任务调度到 $Q'_{run}[1]$ 上执行，否则继续检测序列中的第二台服务器 $Q'_{run}[2]$，依此类推。如果处在运行状态的所有服务器均不能满足任务 t_i 对响应时间的要求，则检测休眠服务器集合 Q_{sleep}，寻找能使当前任务执行能耗最小的服务器 $Q_{sleep}[i]$，并将其唤醒执行任务。算法的伪代码描述如下。

算法 1 FSME

```
1. if Q_global ≠ Ø do
2.     for j=1 to m do
3.         generate Q_run, Q_sleep;
4.     end for
5.     for k=1 to Q_global. num do
6.         if Q_run ≠ Ø do
7.             calculate w_ji;
8.             generate Q'_run;
9.             for i=1 to Q'_run. num do
10.                calculate RT_ki;
11.                if RT_ki ≤ RT_i do
12.                    schedule t_k → Q'_run[i];
13.                end if
14.            end for
15.        end if
16.        if Q'_run[1→num] > RT_i or Q_run = Ø do
17.            calculate w_ji;
18.            generate Q'_sleep;
19.        for i=1 to Q'_sleep. num do
20.            calculate RT_ki;
21.            if RT_ki ≤ RT_i do
22.                schedule t_k → Q'_sleep[i];
23.            end if
24.        end for
25.    end if
26. end for
27. end if
```

算法 FSME 的时间复杂度分析：算法在第 2 步—第 4 步，需要对所有服务器所处状态进行统计，并生成运行服务器队列和休眠服务器队列，总共有 m 个服务器，执行时间为 $O(m)$；在第 6 步—第 25 步，对服务器进行响应时间的检测并调度任务，最坏情况下的执行时间是 $O(m)$。因此，在第 2 步到第 25 步的执行时间是 $O(2m)$。系统环境中总共有 n 类任务，所以 FSME 算法的最坏时间复杂度为 $O(2mn)$。

4 系统能耗及性能分析

针对前面给出的云计算系统模型及节能任务调度算法，本节使用随机 Petri 网工具对系统进行建模，利用马尔可夫随机过程的性质及理论对系统各性能参数进行分析，并给出系统的能耗模型和性能模型。

4.1 系统 SPN 模型

任务集合 T 中的任务都有其所对应的任务类型，并且存在某两个或多个任务的类型相同的情况，因此任务类型数 h 小于等于任务数 m。进一步，用户任务的到达可以转换为各种类型任务的到达。在某一时刻，服务器的本地任务队列中可能存在多种不同类型的任务在等待，它们对处理器资源是竞争关系，这里采用随机任务调度策略对本地队列中的任务进行调度。基于这些描述对系统任务调度过程进行建模，得到如图 2 所示的系统 SPN 模型。

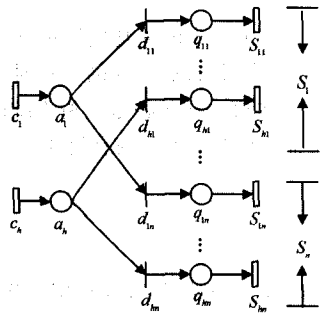

图 2 系统 SPN 模型

图 2 中各变迁和位置的含义如下（其中 $1 \leqslant i \leqslant n, 1 \leqslant j \leqslant h$）。

• 变迁

c_j：表示第 j 类用户任务的到达，其实施速率为 λ_j。

d_{ji}：瞬时变迁，表示任务资源管理部件根据设定的可实施谓词将任务调度到服务器 s_i 上。

S_{ji}：表示服务器 s_i 对第 j 类任务的执行，其实施速率为 λ_{ji}。

• 位置

a_j：表示系统对任务进行调度的位置。

q_{ji}：表示服务器 s_i 的本地任务缓冲队列。

在 SPN 模型中，根据节能任务调度算法对瞬时变迁 d_{ji} 设置相应的可实施谓词，这样就将调度算法应用到了云计算模型中。

4.2 系统能耗分析

云计算环境中的用户任务主要可以分为 CPU 密集型、I/O密集型和交互型[16]。在本文的研究中，主要讨论 CPU 密集型的任务。对于 CPU 密集型的系统，能量消耗主要集中在处理器对任务的执行上，而通信等产生的能耗较少，忽略不计。

首先，对系统 SPN 模型中的服务器 s_i 进行分析。根据随机 Petri 网的定义，云计算系统中的服务器 s_i 随运行情况而处于不同的状态，这些状态组成一个状态集合：$M_s=\{M_j \mid j=0,1,2,3,\cdots\}$。服务器在每个稳定状态下运行都会产生能耗。另外服务器在各个状态之间切换也需要消耗能量，本文中服务器的状态转换耗能主要是在休眠状态和运行状态之间进行切换所产生的能耗。假定两个状态之间的转换所带来的能耗相同。由此，服务器 s_i 的能耗可以表示为其在所有稳定状态下的能耗与状态转换能耗之和，用公式描述如下：

$$E_i=\sum_{M_j\in M_s}(p_j\times P(M_j))\times t+p_c(n_{p\text{-}s}+n_{s\text{-}p}) \tag{1}$$

其中，p_j 表示服务器 s_i 在状态 M_j 下的平均功耗，$P(M_j)$ 表示服务器 s_i 处于 M_j 状态的稳定状态概率，P_c 表示服务器 s_i 的状态转换能耗，$n_{p\text{-}s}$ 表示服务器 s_i 从运行状态转换到休眠状态的次数，$n_{s\text{-}p}$ 表示服务器 s_i 从休眠状态转换到运行状态的次数。因此，整个云计算系统的总能耗可以表示为：

$$E=\sum_{i=1}^{n}E_i=\sum_{i=1}^{n}\Big[\sum_{M_j\in M_s}(p_j\times P(M_j))\times t+p_c(n_{p\text{-}s}+n_{s\text{-}p})\Big] \tag{2}$$

下面介绍服务器 s_i 各状态的稳定状态概率 $P(M_j)$ 及平均功耗 p_j 的求解。对服务器 s_i 进行分析，由于状态空间比较大，对服务器的状态进行适当的简化，即把没有本地缓冲队列不为空的状态作为一个状态，把只有一个缓冲队列不为空的状态作为一个状态，把只有两个缓冲队列不为空的状态作为一个状态，依此类推。简化之后得到的可达状态集如表 1 所列(其中 $1\leqslant\omega$)。

表 1 服务器 s_i 的可达状态集

	q_{1i}	q_{2i}	q_{3i}	…	$q_{(l-1)i}$	q_{li}
M_0	0	0	0	…	0	0
M_1	ω	0	0	…	0	0
	0	ω	0		0	0
	⋮	⋮	⋮		⋮	⋮
	0	0	0		0	ω
M_2	ω	ω	0	…	0	0
	ω	0	ω		0	0
	⋮	⋮	⋮		⋮	⋮
	0	0	0		ω	ω
…	…	…	…	…	…	…

进一步简化求解过程，对 $P_j\times P(M_j)$ 进行求解，根据马尔可夫链稳态性质有：

$$p_0\times P(M_0)=P^{sleep}\times\prod_{k=1}^{l}P(M(q_{ki})=0) \tag{3}$$

$$p_1\times P(M_1)=\sum_{k=1}^{l}p_{ki}\cdot P(M(q_{1i})=0)\cdot P(M(q_{2i})=0)\cdot\cdots\cdot(1-P(M(q_{ki})=0))\cdots P(M(q_{li})=0) \tag{4}$$

$$p_2\times P(M_2)=\sum_{k=1}^{l-1}\sum_{j=k+1}^{l}\frac{1}{2}(p_{i1}+p_{j1})\cdot P(M(q_{1i})=0)\cdot P(M(q_{2i})=0)\cdot\cdots\cdot(1-P(M(q_{ki})=0))\cdot(1-P(M(q_{ji})=0))\cdot\cdots\cdot P(M(q_{li})=0) \tag{5}$$

其它情况依此类推。

对于服务器 s_i 处理第 j 类任务的子模型，利用马尔可夫链稳态性质计算出其各稳态状态概率如下：

$$P(M_{0i})=\frac{\lambda_{ji}}{\mu_{ji}+\lambda_{ji}} \tag{6}$$

$$P(M_{1i})=\frac{\mu_{ji}}{\mu_{ji}+\lambda_{ji}} \tag{7}$$

其中，μ_{ji} 表示第 j 类任务到达服务器 s_i 的速率。那么服务器在稳定状态下位置 q_{ji} 为空的概率为：

$$P(M(q_{ji})=0)=P(M_{0i})=\frac{\lambda_{ji}}{\mu_{ji}+\lambda_{ji}} \tag{8}$$

4.3 系统性能分析

吞吐量和响应时间是衡量系统性能的两个重要指标。吞吐量是指系统单位时间内完成的任务量，响应时间是指任务从提交到完成所经历的时间。本文使用系统吞吐量和系统平均响应时间进行系统性能的评价，并建立性能模型。在随机 Petri 网模型中，系统的各种评价指标可以根据模型的稳定状态概率求得。在 SPN 模型中，系统的总吞吐量可以表示为各服务器吞吐量之和，而每个服务器的吞吐量由代表服务器执行任务的变迁即 s_{ji} 的吞吐量组成。服务器 s_i 处理第 j 类任务的子模型的吞吐量 R_{ji} 可以描述为如下公式：

$$R_{ji}=P(M_{1i})\cdot\lambda_{ji}=\frac{\mu_{ji}}{\mu_{ji}+\lambda_{ji}}\cdot\lambda_{ji} \tag{9}$$

服务器 s_i 的吞吐量可以表示为：

$$R_i=\sum_{j=1}^{l}R_{ji}=\sum_{j=1}^{l}\frac{\mu_{ji}}{\mu_{ji}+\lambda_{ji}}\cdot\lambda_{ji} \tag{10}$$

则整个系统总的吞吐量为：

$$R=\sum_{i=1}^{n}R_i=\sum_{i=1}^{n}\sum_{j=1}^{l}\frac{\mu_{ji}}{\mu_{ji}+\lambda_{ji}}\cdot\lambda_{ji} \tag{11}$$

根据随机 Petri 网的性质，用 $M(q_{ji})$ 表示位置 q_{ji} 中的平均标记数，则服务器 s_i 处理第 j 类任务子模型的平均响应时间可以用如下公式得到：

$$RT_{ji}=\frac{M(q_{ji})}{R_{ji}} \tag{12}$$

稳定状态下任务 t_j 在服务器 s_i 上执行的响应时间 RT'_{ji} 包括执行本地任务队列中任务的时间及执行任务 t_j 本身的时间，用公式表示为：

$$RT'_{ji}=\sum_{k=1}^{h}\frac{M(q_{ki})}{\lambda_{ki}}+\frac{1}{\lambda_{ji}} \tag{13}$$

由式(12)可进一步得到整个系统的平均响应时间为：

$$RT=\frac{\sum_{i=1}^{n}\sum_{j=1}^{m}M(q_{ji})}{\sum_{i=1}^{n}\sum_{j=1}^{m}\lambda_{ji}} \tag{14}$$

5 实验数据和结果分析

本节中采用模拟仿真软件 SPNP6.0 对系统进行建模实

验。SPNP 是美国 Duke 大学 Trivedi 教授带领的团队研究和开发的，是一款比较成熟的随机 Petri 网建模分析工具。它能实现较好的系统模型描述，并可以运行得到模型的各种性能参数，如响应时间、吞吐量，及其它自定义参数。

为了说明 FSME 算法的有效性，本文将 FSME 算法与经典的 Min-Min 调度算法进行性能和能耗的比较。Min-Min 调度算法的思想是：对于所有的任务，优先将具有最小完成时间的最小任务调度到对应计算节点上。Min-Min 算法是一个动态调度算法，具有实现简单、可扩展性强、效率高等特点，被广泛应用于同构或异构分布式并行计算环境中。结合上节对系统能耗和性能的分析，本节从系统的总能耗、平均响应时间以及吞吐量 3 个方面对两个算法进行比较分析。

为了避免状态空间爆炸，不失一般性，这里仅考虑有两类到达任务和两个服务器的情况。假定两类任务 r_1 和 r_2 的到达服从相同参数的泊松分布，用 λ_1 和 λ_2 表示两类任务的平均到达速率，那么 $\lambda_1=\lambda_2$。另外，设置 λ_1 和 λ_2 从 25 请求/秒递增到 200 请求/秒，步长为 25 请求/秒，那么系统总的任务到达速率则从 50 请求/秒递增到 400 请求/秒，步长为 50 请求/秒。两个服务器对两类任务的服务时间服从参数为 $1/\lambda_{ji}$ 的负指数分布，并设定 λ_{ji} 的值分别为：$\lambda_{11}=300$，$\lambda_{21}=350$，$\lambda_{12}=250$，$\lambda_{22}=300$。实验结果如图 3－图 5 所示。

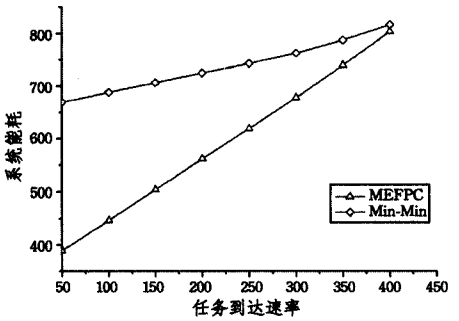

图 3　系统总能耗比较

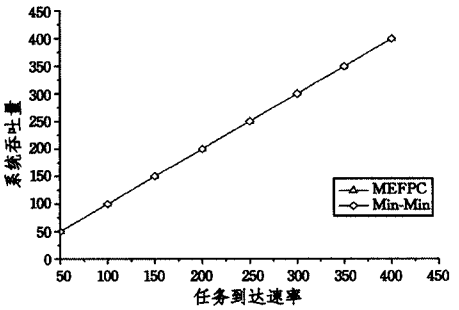

图 4　系统吞吐量比较

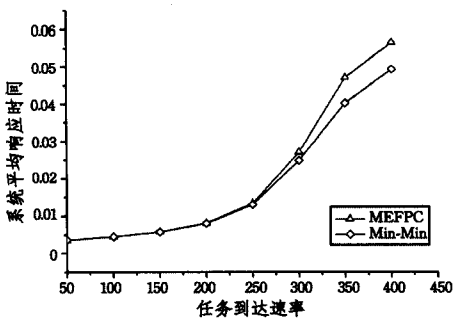

图 5　系统平均响应时间比较

从图 3 实验数据可以看出，在系统稳定状态下，FSME 调度算法的能耗始终低于 Min-Min 调度算法，并且任务到达速率越小效果越明显。其原因在于：(1)FSME 调度算法将任务优先调度到运行中的服务器上执行，将尽可能多的服务器空闲出来置于休眠状态，并减少了休眠状态和运行状态之间的转换次数；(2)FSME 调度算法以最小能耗原则将任务优先调度到执行能耗最小的服务器上执行。总的来说，FSME 调度算法在降低空闲能耗和执行能耗两个方面都进行了优化，因而能够达到较好的整体节能效果。而 Min-Min 调度算法只专注于任务完成时间，不考虑系统能耗因素，所以耗能较大。

图 4 和图 5 的实验数据表明，随着任务到达速率的增加，FSME 调度算法的吞吐量始终能与 Min-Min 调度算法保持一致，而系统平均响应时间则略大于 Min-Min 调度算法。这是因为 Min-Min 调度算法专注于系统的性能优化，关注的是系统响应时间的减少，因此算法性能较好。而 FSME 算法在关注系统性能的同时尽可能地降低系统能耗，在任务调度时将任务优先调度到运行中的服务器上，使得部分服务器负载加重，响应时间增加，进而导致了系统整体性能一定程度的下降。但是从实验数据来看，性能下降得并不明显，在可接受的范围之内。

综合来看，FSME 调度算法虽然性能上比 Min-Min 调度算法稍差，但是能带来较好的节能效果，因此总体性能要优于 Min-Min 调度算法。

结束语　针对云计算系统中的高能耗问题，通过对云计算任务调度模型的研究，提出了一种最小能耗优先的任务调度方法。采用随机 Petri 网工具建立系统模型，并利用马尔可夫链对状态转换过程进行分析，建立起系统的能耗模型和性能模型。该任务调度方法在满足任务对响应时间要求的前提下尽可能地将更多的服务器空闲出来并置于休眠状态，同时以最小能耗原则对任务进行调度。实验验证了本文所提节能调度策略的有效性。

云计算系统中虚拟机管理是一个重要方面，考虑能耗优化的虚拟机管理策略通过虚拟机迁移等技术将虚拟机聚拢，以空闲出更多的物理服务器并使之休眠来达到节能的目的。如何根据系统的运行状况将虚拟机进行合理的迁移，尽可能地减少系统能耗，将作为下一步研究的重点。

参 考 文 献

[1]　陈康，郑纬民. 云计算：系统实例与研究现状[J]. 软件学报，2009，20(5)：1337-1348
Chen Kang，Zheng Wei-min. Cloud Computing：System Instances and Current Research[J]. Journal of Software，2009，20(5)：1337-1348

[2]　绿色网络和绿色评价：节能机制、模型和评价[J]. 计算机学报，2011，34(4)：593-612
Lin Chuang，Tian Yuan，Yao Min. Green network and green evaluation：Mechanism，modeling and evaluation[J]. Chinese Journal of Computers，2011，34(4)：593-612

[3]　张小庆，贺忠堂，李春林，等. 云计算系统中数据中心的节能算法研究[J]. 计算机应用研究，2013，30(4)：961-964
Zhang Xiao-qing，He Zhong-tang，Li Chun-lin，et al. Research on energy saving algorithm of datacenter in cloud computing system [J]. Application Research of Computers，2013，30(4)：961-964

[4] Williams C. What is a green data center[OL]. 2011. http://www. mnn. com/-green-tech/computers/stories/what-is-a-green-data-center

[5] Liao W H, Yen Wen-ming. Power-saving Scheduling with a QoS Guarantee in a Mobile WiMAX System[J]. Journal of Network and Computer Applications, 2009, 32: 1144-1152

[6] Dong Ji-guang, Chen Wei-Wei, WU Hai-Jia, et al. Load balancing study in cloud storage based on dynamic replica technology [J]. Application Resea-rch of Computers, 2012, 29(9): 3422-3425

[7] 吴琦,熊光泽. 非平衡自相似业务下自适应动态功耗管理[J]. 软件学报,2005,16(8):1499-1505
Wu Qi, Xiong Guang-ze. Adaptive Dynamic Power Management for Non-Stationary Self-Similar Requests[J]. Journal of Software, 2005, 16(8): 1499-1505

[8] Rodero I, Jaramillo J, Quiroz A, et al. Energy-efficient application-aware online provisioning for virtualized clouds and data centers [C]// Proc of International Conference on Green Computing. Washington DC: IEEE Computer Society, 2010: 31-45

[9] 肖志娇,明仲,蔡树彬. 基于状态管理的服务器节能策略研究[J]. 计算机科学,2013,40(4):22-30
Xiao Zhi-jiao, Ming Zhong, Cai Shu-bin. Study on Energy Optimizaion of Servers Based on States Management[J]. Computer Science, 2013, 40(4): 22-30

[10] 谭一鸣,曾国荪,王伟. 随机任务在云计算平台中能耗的优化管理方法[J]. 软件学报,2012,23(2):266-278
Tan Yi-ming, Zeng Guo-sun, Wang Wei. Policy of Energy Optimal Management for Cloud Computing Platform with Stochastic Tasks[J]. Journal of Software, 2012, 23(2): 266-278

[11] Lin M H, Wierman A, Andrew L L, et al. Dynamic right-sizing for power-proportional data centers[J]. IEEE/ACM Trans. on Networking, 2013, 21(5): 1378-1391

[12] Beloglazov A, Abawajy J, Buyya R. Energy-aware resource allocation heuristics for efficient management of data centers for cloud computing [J]. Future Generation Computer Systems, 2012, 28(5): 755-768

[13] 李钰,张海文,吴名瑜,等. 一种面向异构机群系统的低功耗任务调度算法[EB/OL]. 2009-10-23. http://d. wanfangdata. c-om. cn/Conference_7242376. aspx
Li Yu, Zhang Hai-wen, Wu Ming-yu, et al. An Efficient, Energy-aware Heuristic Scheduling Algorithm for Heterogeneous Clusters[EB/OL]. 2009-10-23. http://d. wanfangdata. com. cn/Conference_72423-76. aspx

[14] 何丽,饶俊,赵富强. 一种基于能耗优化的云计算系统任务调度方法[J]. 计算机工程与应用,2013,49(20):19-22
He Li, Rao Jun, Zhao Fu-qiang. Task scheduling method based on energy optimization in cloud computing system[J]. Computer Engineering and Applications, 2013, 49(20): 19-22

[15] 林闯. 随机 Petri 网和系统性能评价(第二版)[M]. 北京:清华大学出版社,2005
Lin Chuang. Stochastic Petri net and Performance evaluation of systems(2nd edition)[M]. Beijing: Tsinghua University press, 2005

[16] 宋杰,李甜甜,闫振兴,等. 一种云计算环境下的能效模型和度量方法[J]. 软件学报,2012,23(2):200-214
Song Jie, Li Tian-tian, Yan Zhen-xing, et al. Energy-Efficiency Model and Measuring Approach for Cloud Computing[J]. Journal of Software, 2012, 23(2): 200-214

(上接第 94 页)

[4] Huang W W, Peng Y K, Wen J, et al. Energy-efficient mhop hier ar chical routing protocol for wireless sensor netwoks[J]. IEEE Computer Society, 2009, 35(2): 469-472

[5] Heinzelman W R, Chandrakasan A, Balakrishnan H. Energy-Efficient communication protocol for wireless microsensor networks[C]// Proc. of the Hawaii Int'l Conf. on System Sciences. San Francisco: IEEE Computer Society, 2000: 3005-3014

[6] Younis O, Fahmy S. Distributed clustering in ad-hoc sensor networks: A hybrid, energy-efficient approach[C]// Zhang Z S, Low S, eds. Proc. of the IEEE INFOCOM. San Francisco: IEEE Computer Society Press, 2004

[7] Manjeshwar A, Agrawal D P. TEEN: A routing protocol for enhanced efficiency in wireless sensor networks[C]// Proc. of the 15th Parallel and Distributed Processing Symp. San Francisco: IEEE Computer Society, 2001: 2009-2015

[8] Lindsey S, Raghavendra C S. PEGASIS: Power-Efficient gathering in sensor information systems[C]// Proc. of the IEEE Aerospace Conference Montana: IEEE Aerospace and Electronic Systems Society. 2002: 1125-1130

[9] 孙彦清,彭舰,刘唐,等. 基于动态分区的无线传感器网络非均匀成簇路由协议[J]. 通信学报,2014,35(1):198-206
Sun Yan-qing, Peng Jian, Liu Tang, et al. Uneven clustering routing protocol based on dynamic partition for wireless sensor network[J]. Journal on Communications, 2014, 35(1): 198-206

[10] 蒋畅江,石为人,唐贤伦,等. 能量均衡的无线传感器网络非均匀分簇路由协议[J]. 软件学报,2012,23(5):190-200
Jiang C J, Shi W R, Tang X L, et al. Energy-balanced unequal clustering routing protocol for wireless sensor networks[J]. Journal of Software, 2012, 23(5): 190-200

[11] 李朋飞,李志华,尹熙,等. 基于能量等级的分簇拓扑控制算法[J]. 计算机科学,2014,41(3):96-99
Li Peng-fei, Li Zhi-hua, Yin Xi, et al. Energy-level Based Clustering Network Topology Control Algorithm[J]. Computer Scie-nce, 2014, 41(3): 96-99

[12] 李洪兵,熊庆宇,石为人. 无线传感器网络非均匀等级分簇拓扑结构研究[J]. 计算机科学,2013,40(2):49-52,77
Li Hong-bing, Xiong Qing-yu, Shi Wei-ren. Study on Topology with Non-uniform Hierarchical Clustering for Wireless Sensor Networks[J]. Computer Science, 2013, 40(2): 49-52, 77

[13] 陈庆章,赵小敏,陈晓莹. 提高无线传感器网络能效的双轮成簇协议设计[J]. 软件学报,2010,21(11):2933-2943
Chen Q Z, Zhao X M, Chen X Y. Design of double roundsclustering protocol for improving energy efficient in wireless sensor networks[J]. Journal of Software, 2010, 21(11): 2933-2943

[14] 苏金树,郭文忠,余朝龙,等. 负载均衡感知的无线传感器网络容错分簇算法[J]. 计算机学报,2014,37(2):445-456
Su Jin-shu, Guo Wen-zhong, Yu Chao-long, et al. Fault-Tolerance Clustering Algorithm with Load-Balance Aware in Wireless Sensor Network[J]. Chinese journal of computers, 2014, 37(2): 445-456