

# 云计算中基于 M/Geom/C/∞ 排队系统的任务调度模型研究

甘娜, 黄裕锋, 陆晓梅

(广东工程职业技术学院 信息工程学院, 广州 510520)

**摘要:** 传统的排队模型已经难以满足云计算中任务调度系统的高效和低成本目标要求, 为缩短云任务调度过程任务等待时间及提高虚拟机任务调度系统的执行效率, 提出一种云环境下基于 M/Geom/C/∞ 排队系统的任务调度模型, 利用改进的排队模型提高云任务调度系统性能; 对该模型中系统的嵌入马尔可夫链性质、平衡条件、稳态分布和条件随机分解结果进行了分析, 给出该模型的稳态队长的随机分解和稳态等待时间等性能指标; 结合数值例子, 准确的找到服务率与期望队长、期望等待时间之间的关系及其它系统稳态性能指标; 通过云任务调度系统的仿真, 实验结果验证了该模型能够快速地完成云任务的调度, 提高了虚拟机资源的平均利用率。

**关键词:** 云任务调度; 排队系统; 期望队长; 等待时间

## Research on Task Scheduling Model Based on M/Geom/C/∞ Queuing System in Cloud Computing

Gan Na, Huang Yufeng, Lu Xiaomei

(Information Engineering Departments, Guangdong Engineering Polytechnic, Guangzhou 510520, China)

**Abstract:** The traditional queuing model has been difficult to meet the requirements of high efficiency and low cost of task scheduling system in cloud computing, in order to shorten task waiting time and improve the execution efficiency of task scheduling system in the cloud task scheduling, a task scheduling model based on M/Geom/C/∞ queuing system in cloud environment is proposed, the improved queuing model is used to improve the performance of cloud task scheduling system. The properties of the embedded Markov chain, the equilibrium condition, the steady-state distribution and the conditional random decomposition are analyzed, the performance index of the model, such as stochastic decomposition and steady-state waiting time of the steady-state queue length, are given. A numerical example is given to accurately find the relationship between the service rate and expected queue length, expected waiting time, and other steady-state performance index of the system. Through the simulation of the cloud task scheduling system, the experimental results show that the model can quickly complete the scheduling of cloud tasks, and improve the average utilization of virtual machine resources.

**Keywords:** cloud task scheduling; queuing system; expected queue length; waiting time

### 0 引言

云计算中任务调度主要研究的是如何将合适的虚拟机调度分配给用户提交的任务使用。用户提交的任务被分割成  $n$  个子任务, 形成  $n$  个子任务的排队队列, 队列中的任务通过调度器发送到  $m$  个虚拟机上执行计算<sup>[1]</sup>。云任务调度就是以实时掌握云环境下虚拟机的负载状态为前提, 根据云任务排队状态将虚拟机资源进行科学调度、分配, 均衡各虚拟机的负载, 达到云任务执行效率最高且成本低的目标。具体执行过程是: 根据执行效率高的虚拟机与执行效率低的虚拟机之间的负载状态及云任务排队状态, 均衡排队队列的队长, 优先使用执行效率高的虚拟机, 目的就是使得那些处于排队状态的任务能够得到最快响应并开始执行, 从而大大缩短任务执行时间和服务成本, 有效提高任务执行效率。云计算中的任务调度问题已被证明为

一类  $N_p$ -hard 问题, 传统调度算法在有效时间内得到满意解的效果并不理想, 但是, 通过对云任务调度问题进行建模, 分析影响虚拟机服务性能的因素, 获取该模型中的任务等待时间、系统资源利用率等性能指标来评估模型的综合性能。

目前, 一些学者从多方面入手来求解云计算任务调度问题, 但是, 在云任务调度模型分析与研究并不多。参考文献 [2] 根据 M/G/1 排队论, 给出了求解云任务调度的可靠性的均衡调度算法。参考文献 [3] 提出了一种基于 M/M/n/n+r 排队系统的云计算中心任务调度分析模型, 求解该模型的用户请求响应时间及 QoS 性能指标, 参考文献 [4] 确定了基于 M/M/n 排队模型的云计算资源调度模型, 以降低任务服务时间和能耗为调度目标。参考文献 [5] 给出了一种适用于云计算服务的有效 QoS 约束调度方案。

不同于已有的研究工作, 本文设计了一种基于 M/Geom/C/∞ 排队系统的云任务调度模型, 以虚拟机的服务可靠性为效用函数, 运用排队论分析任务在虚拟机上执行的任务调度策略。该调度策略保证了排队队长较短、任务等待时间较短、平均响应时间较快, 仿真结果证实了其有效性。

收稿日期: 2017-08-09; 修回日期: 2017-09-08。

基金项目: 广东省省级科技计划项目(2014A020217016); 广东省省级科技计划项目(2015A030303013)。

作者简介: 甘娜(1980-), 女, 江西萍乡人, 高级工程师, 硕士, 主要从事智能优化、云计算方向的研究。

## 1 云计算中的任务调度

云计算任务调度描述:

云计算中任务调度的实质是把用户提交的任务分解为  $n$  个独立的子任务集合  $T = \{t_1, t_2, \dots, t_n\}$ , 并通过调度器将子任务发送到  $m$  个虚拟机  $vm = \{vm_1, vm_2, \dots, vm_m\}$  上去执行, 任务调度的目标是尽量让排队队列中任务的等待时间较短、响应时间较快, 在虚拟机上执行时间相对较短。具体描述如下:

定义 1: 云任务调度系统中有  $k$  个用户, 其中  $U_i (1 \leq i \leq k)$  为第  $i$  个用户提交的任务, 用户  $U_i$  提交任务的平均速率  $\lambda_i$  服从负指数分布。

定义 2: 对用户提交的任务进行分割, 获得  $n$  个待调度的云任务队列  $T = \{t_1, t_2, \dots, t_n\}$ , 其中  $t_i (1 \leq i \leq n)$  为第  $i$  个子任务, 这  $n$  个子任务被调度器分配到虚拟机上执行, 子任务  $t_i$  并行发送, 每个子任务只能分配到一个虚拟机上执行。

定义 3:  $VM = \{vm_1, vm_2, \dots, vm_m\}$  是云环境中  $m$  个虚拟机的集合, 其中  $m < n, vm_j (1 \leq j \leq m)$  为第  $j$  个虚拟机。任务  $t_i$  在虚拟机  $vm_j$  上的平均执行速率  $\mu_j$ , 每一个虚拟机  $vm_j$  可以被看作一个 M/M/1 排队系统。

## 2 基于 M/Geom/C/∞ 排队系统的云任务调度算法

### 2.1 排队系统

排队问题可以描述为服务与被服务的关系, 由输入过程、排队规则和服务规则这 3 个部分组成。输入过程是指任务到达排队系统的过程, 考察的是任务到达排队系统的规律。可以用一定时间内任务到达数的概率分布或前后两个任务相继到达的间隔时间等来描述。排队规则则可采用最常见的先到者先服务模式。服务机构就是指排队系统中多个虚拟机。

假设请求服务是连续到达的, 服从参数为  $\lambda$  的负指数分布, 排队系统共有  $c$  个服务台, 请求服务到达间隔长为  $k$  时,  $c$  个服务台总共工作  $ck$  个时隙, 服务台的服务率  $\mu$  服从几何分布, 等价于任何时隙上的服务以概率  $\mu$  结束, 以概率  $\bar{\mu}$  将延至下一个时隙, 并且不同时隙上的服务可以看成相互独立的。引入下列条件随机变量:

$$L = \{L - c \mid L \geq c, J = 1\},$$

$$W = \{W \mid L \geq c, J = 1\} \quad (1)$$

其中:  $L$  是期望队长, 系统中排队等待的任务数。  $W$  是在服务台全忙时请求服务的期望等待时间, 根据上述性能指标确定服务台的平均利用率<sup>[6-7]</sup>。

### 2.2 排队模型描述

排队模型的选择对提高云任务调度的效率非常关键作用, 它还决定了调度系统对资源的利用率。所以要根据目标与现状, 加入一定的目标约束条件, 设计一种适合云任务调度系统的排队模型, 提高调度系统运行效能<sup>[8-10]</sup>。由于排队系统的特征, 排队模型可以有许许多多的组合, 本文采用 M/Geom/C/∞ 形式表示云任务调度中的排队模型。其中:  $M$  表示任务到达间隔时间服从负指数概率分布,  $Geom$  表示服务时间服从几何概率分布,  $C$  为虚拟机台数,  $\infty$  表示任务来源总体数无限。M/Geom/C/∞ 排队模型属于等待制排队模型, 模型如图 1 所示。

#### 2.2.1 仿真模型组成部分

1) 任务请求服务到达为固定速率  $\lambda$  的负指数分布, 负指

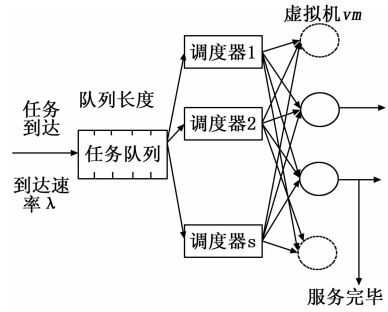


图 1 云任务调度中排队系统仿真模型

数分布用“ $M$ ”表示, 有无记忆性, 即 Markov 性。

2) 服务率  $\mu$ 、平均服务时间  $S_1$  相互独立并服从几何分布, 一般用“Geom”表示几何分布。  $b_k = P\{S_1 = k\} = \mu \bar{\mu}^{k-1}, k \geq 1$ ;

$$\bar{\mu} = 1 - \mu \quad (2)$$

3) 等待服务时间  $X$  服从参数  $p$  几何分布  $P\{X = k\} = (1 - p)p^k, 0 < p < 1$ ;

$$k = 0, 1, \dots \quad (3)$$

的数学期望和概率母函数分别是:

$$E(X) = \frac{1}{1-p}, X(z) = \frac{1-p}{1-zp}, |z| \leq 1 \quad (4)$$

#### 2.2.2 模型所使用到的数量指标

1) 单位时间内到达任务数的期望值, 即单位时间内的平均到达率, 记作  $\lambda$ 。而  $\frac{1}{\lambda}$  表示相邻两个任务到达的平均间隔时间;

2) 单位时间内被服务的任务数的期望值, 即单位时间内任务的平均离去率, 记作  $\mu$ 。而  $\frac{1}{\mu}$  表示每个任务的平均服务时间;

3) 在时刻  $t$  排队系统中恰有  $n$  个任务的概率为  $P_n(t)$ , 而  $P_0(t)$  为系统空闲率;

4) 排队队列中平均任务数称为队长, 记作  $L$ ;

5) 排队队列中等待服务的平均任务数称为等待队长, 记作  $L_q$ ;

6) 任务从进入排队队列到接受完服务后离开系统的平均时间称为平均逗留时间, 记作  $W$ ;

7) 任务在系统内排队等待服务的平均时间称为平均等待时间, 记作  $W_q$ 。

#### 2.2.3 云任务调度流程

1) 在 M/Geom/C/∞ 排队系统中, 任务到达间隔服从负指数分布, 以  $\{Q_n, n \geq 0\}$  表示时隙分点  $n$  处观察到的队长(队列中的任务数)过程。当时刻  $n$  落在某到达间隔之内时,  $Q_n$  的值取决于时刻  $n$  处剩余到达间隔的长度, 剩余到达间隔分布满足 Markov 性质;

2) 以  $\tau_n$  表示第  $n$  个到达时刻, 并约定  $\tau_n = 0$ , 计时原点  $n = 0$  是一个到达间隔的始点。 $\{\tau_n, n \geq 0\}$  是队长过程的再生点集, 设  $L_n^- = Q(\tau_n^-)$  表示第  $n$  个到达时刻前夕系统中任务数,  $\{L_n^-, n \geq 0\}$  是一个 MC, 称为嵌入在到达时刻的队长过程;

3) 选择到达时刻作为嵌入点。嵌入时刻不仅可落在忙期,

也可能落入假期。以  $L_n^-$  表示第  $n$  个任务到达前夕系统中任任务数;

4)  $k_j$  是一个到达间隔内恰好完成  $j$  个任务的概率,  $\{k_j, j \geq 0\}$  是一个概率分布,  $v_j$  是到达间隔大于休假时间  $V$ , 并且在休假结束后的剩余到达间隔内恰好完成  $j$  个任务的概率;

5) 虚拟机服务台寿命为  $X$ , 服从参数为  $\alpha (0 \leq \alpha < \infty)$  的负指数分布。服务台失效后立即进行修理, 其修理时间  $Y$  是任意分布。

### 2.3 模型分析

#### 2.3.1 嵌入马尔可夫链

马尔可夫链的特性就是事物将来的状态分布只取决于现在, 跟过去无任何关系, 也即“已知现在则将来与过去无关”。从概率上来马尔可夫链看是个随机的过程, 在这个随机过程中, 第  $n$  时刻的状态, 取决于  $n-1$  时刻的状态, 和  $n-2$  时刻的状态完全无关。对于 M/Geom/C/ $\infty$  排队系统, 由于服务时间是几何分布, 对任选的一个时刻  $t$ , 正在接受服务的任务可能还没有服务完。从时刻  $t$  起的剩余服务时间分布不再具有无记忆性的性质, 于是队长过程  $\{N(t), t \geq 0\}$  不再具有马尔可夫性质。但是, 假定  $T_n$  表示从  $t=0$  开始第  $n$  个任务的到达时刻,  $N_n^+$  表示第  $n$  个任务到达时看到系统中已有的任务, 即看到的任务数, 则  $\{N_n^+, J_n; n \geq 1\}$  是队长过程的嵌入马尔可夫链。

#### 2.3.2 系统平衡条件和系统转移率阵

下面分两种情况, 导出 MC 的转移概率:

1)  $0 \leq i < c, j \leq i+1 < c$ . 到达间隔开始时有  $i+1 < c$  个服务台工作。若到达间隔长为  $k$ , 从  $i$  转移到  $j$ , 表明这  $i+1$  个服务台上有  $j$  个在  $k$  的时隙上不发生离去, 其余  $i+1-j$  服务台上发生离去。

2)  $i \geq c, j \leq c$ . 到达间隔长为  $k$ , 则在这  $k$  个时隙上必有第  $r$  个时隙。该时隙开始时, 系统任务数大于  $c$ ; 而该时隙结束时系统有  $l$  个任务,  $j \leq l \leq c$ . 并且在该时隙上恰好完成了  $m$  个服务,  $c-l+1 \leq m \leq c$ .

注意到转移概率的时齐性, 嵌入马尔可夫链的初始概率分布为平稳分布, 因此该排队系统为平稳分布。

#### 2.3.3 系统稳态分布

下面应用矩阵几何技术来讨论该模型的平稳结果。

对于云任务调度的排队系统, 在系统达到统计平衡下, 任务到达时队长  $N^+$  可分解成独立的两部分之和, 即  $N^+ = N_1^+ + N_2^+$

而且

$$1) \text{ 当 } \theta \neq \mu(1-\delta) \text{ 时, 有}$$

$$P\{N_1^+ = j\} = \begin{cases} \frac{1-\delta}{(1-\delta) + \gamma[\delta - f(\theta)]}, & j = 0 \\ \frac{\gamma[\delta - f(\theta)]}{(1-\delta) + \gamma[\delta - f(\theta)]} (1-\delta)\delta^{j-1}, & j \geq 1 \end{cases} \quad (5)$$

$$P\{N_2^+ = j\} = [1 - f(\theta)] f^j(\theta), j \geq 0 \quad (6)$$

2) 当  $\theta = \mu(1-\delta)$  且  $0 < \delta \leq 1 - \frac{1}{e}$  时, 有

$$P\{N_1^+ = j\} =$$

$$\left\{ \begin{array}{l} 1 - \mu \int_0^\infty t e^{-\mu(1-\delta)t} dF(t), \quad j = 0, \\ (1-\delta)\delta^{j-1} \mu \int_0^\infty t e^{-\mu(1-\delta)t} dF(t), \quad j \geq 1, \end{array} \right\} \quad (7)$$

$$P\{N_2^+ = j\} = (1-\delta)\delta^j, j \geq 1, \quad (8)$$

其中:  $\gamma = \frac{\theta}{\theta - \mu[1 - f(\theta)]}$ ,

$f(\theta) = \int_0^\infty e^{-\theta t} dF(t), \delta (0 < \delta < 1)$  为方程  $z =$

$\int_0^\infty e^{-\mu(1-z)t} dF(t)$  在  $(0,1)$  内的唯一解。

在系统达到统计平衡下 ( $\rho < 1$ ), 任务到达时看到队长的平均数为

1) 当  $\theta \neq \mu(1-\delta)$  时, 有

$$E[N^+] = \frac{\gamma[\delta - f(\theta)]}{(1-\delta)^2 + \gamma[\delta - f(\theta)](1-\delta)} + \frac{f(\theta)}{1 - f(\theta)} \quad (9)$$

2) 当  $\theta = \mu(1-\delta)$  且  $0 < \delta \leq 1 - \frac{1}{e}$  时, 有

$$E[N^+] = \frac{\mu \int_0^\infty t e^{-\mu(1-\delta)t} dF(t)}{1-\delta} + \frac{\delta}{1-\delta} \quad (10)$$

#### 2.3.4 条件随机分解结果

对于一个多虚拟机服务台排队系统, 只有当所有服务台全忙时, 才能充分地反映出这个系统的本质特征。对应的无休假期经典 M/M/C 排队系统中, 同名条件随机变量记为  $L_0$  和  $W_0$ , 它们的 PGF 分别是

$$L_0(z) = \frac{1-\rho}{1-z\rho}, W_0(s) = \frac{c\mu(1-\rho)}{s + c\mu(1-\rho)} \quad (11)$$

当  $\rho < 1$  时, M/Geom/C/ $\infty$  排队中  $(L^-, J)$  的分布是

$$\left\{ \begin{array}{l} \pi_{k0} = \sigma(1-\alpha)H^k(\bar{\theta}), k \geq 0, \\ \pi_{k1} = \sigma(1-\alpha)\delta \sum_{j=0}^{k-1} \alpha^j H^{k-1-j}(\bar{\theta}), k \geq 1. \end{array} \right. \quad (12)$$

1) 当  $\rho < 1$  时, M/Geom/C/ $\infty$  排队中到达前夕的稳态队长  $L^-$  可分解成两个独立随机变量之和:

$$L^- = L_0^- + L_d, \quad (13)$$

其中:  $L_0^-$  是经典无休假排队中到达前夕的稳态队长, 服从参数  $\alpha$  的几何分布; 附加队长  $L_d$  有 PGF:

$$L_d(z) = \frac{\sigma(1-\alpha z + \alpha\delta)}{1 - zH(\bar{\theta})}. \quad (14)$$

由稳态队长的随机分解结构, 容易得出下列平均队长公式:

$$E(L_d) = \frac{\mu}{\theta} \frac{\sigma\delta}{1-H(\bar{\theta})}, E(L^-) = \frac{\alpha}{1-\alpha} + \frac{\mu}{\theta} \frac{\sigma\delta}{1-H(\bar{\theta})}. \quad (15)$$

另外, 在服务台忙的条件下, 讨论系统中排队等待任务数的分布。设:

$$Q^{(1)} = \{L^- - 1 \mid j = 1\} \quad (16)$$

由常数  $\sigma$  的定义, 有:

$$P\{J = 1\} = \frac{\delta}{1-\alpha + \delta} = \frac{\sigma\delta}{1-H(\bar{\theta})}. \quad (17)$$

根据定理可得出:

$$P\{Q^{(1)} = j\} = P\{L^- = j + 1 \mid J = 1\} = (1 - \alpha)(1 - H(\bar{\theta})) \sum_{r=0}^j \alpha^r H^{r-1}(\bar{\theta}), j \geq 0 \quad (18)$$

取 PGF, 有  $Q^{(1)}(z) = \frac{1 - \alpha}{1 - \alpha z} \frac{1 - H(\bar{\theta})}{1 - zH(\bar{\theta})}$ 。

已知系统处于忙期条件下, 排队等待任务数  $Q^{(1)}$  可分解成两个独立随机变量之和, 它们分别服从参数  $\alpha$  和  $H(\bar{\theta})$  的几何分布。

(2) 当  $\rho < 1$  时, M/Geom/C/∞ 系统中稳态等待时间  $W$  可分解为两个独立随机变量之和:

$$W = W_0 + W_d, \quad (19)$$

其中:  $W_0$  是对应经典无休假系统中的等待时间, 其分布由稳态等待时间给出; 附加延迟与休假时间  $V$  同分布, 有 PGF:

$$W_d(s) = \frac{\theta s}{1 - \theta s}. \quad (20)$$

### 3 算法仿真与结果分析

本节通过 Matlab 编程绘图, 以图表的形式展示基于 M/Geom/C/∞ 排队系统的云任务调度模型的一些稳态概率变化曲线图。

假定参数设置如下: 到达率 0.7 个/分, 服务率 (0.03, 0.5) 个/分, 平均服务时间 0.6 分/个, 期望服务水平 80%。

结合数值例子, 表 1 给出了期望队长和期望等待时间随着服务率变化的数据, 从表中也可以看出仿真系统的服务台平均利用率较高。

表 1 系统的稳态指标仿真值

服务率/ (个/分)	期望队长/ 个	期望等待 时间/秒	服务台平均 利用率
0.03	6.3613	140.3251	0.7562
0.07	6.1238	126.7325	0.7854
0.13	5.8231	108.6674	0.7635
0.17	5.7635	99.3365	0.6954
0.23	5.6238	78.9863	0.7346
0.27	5.3269	70.5638	0.6853
0.33	4.9546	65.3857	0.6236
0.37	4.2365	62.3429	0.6746
0.43	3.7685	56.3426	0.7364
0.47	3.3428	52.3246	0.7539

图 2 和图 3 分别给出了排队系统的期望队长、期望等待时间随着服务率变化的理论值与仿真值。

从图 2 和图 3 可以看出, 随着服务率的增大, 队长和等待时间均减小。并且从图中可以看出排队系统期望队长、期望等待时间的理论值与仿真值比较接近, 验证了本文中云任务调度模型的有效性、合理性。

### 4 结论

本文将排队论相关知识应用于云环境下的任务调度模型分析与研究中, 提出了一个基于 M/Geom/C/∞ 排队系统的云任务调度模型。结合排队论工具对云任务调度过程的分析, 利用嵌入马尔可夫链理论方法研究任务调度排队系统, 寻找到既可以缩短调度过程中任务排队等待的时间, 又能提高系统虚拟机

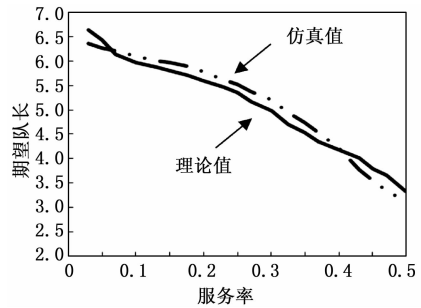


图 2 期望队长随服务率变化的趋势

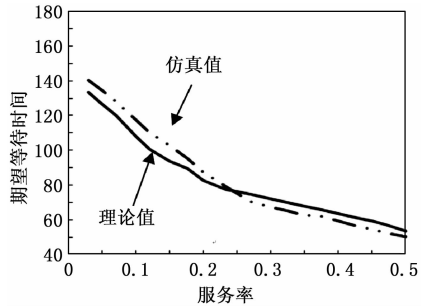


图 3 等待时间随服务率变化的趋势

资源利用率的调度方案。仿真实验得到了该模型的平稳分布及相关性能指标, 验证了其有效性。

#### 参考文献:

- [1] 胡德敏, 户 静, 余 星. 云计算环境下基于微粒群的虚拟机任务调度算法 [J]. 计算机测量与控制, 2014, 22 (4): 1189 - 1192.
- [2] 王 勇, 等. 云环境下基于可靠性的均衡任务调度算法研究 [J]. 计算机科学, 2015, 42 (6A): 325 - 330.
- [3] 何怀文, 等. 基于 M/M/n/n+r 排队模型的云计算中心服务性能分析 [J]. 计算机应用, 2014, 34 (7): 1843 - 1847.
- [4] 游卓浩. 基于 M/M/n 排队模型的云资源调度策略研究 [D]. 成都: 电子科技大学, 2014, 6.
- [5] Lee Liangteh, Liu Kangyuan, Chiang Mingjen. An Effective QoS-Constrained Scheduling Scheme for Cloud Computing Services [J]. Journal of Electronic Science and Technology, 2013, 11 (2): 161 - 168.
- [6] 汪 浩, 黄明和, 龙 浩. 基于 G/G/1-FCFS、M/G/1-PS 和 M/G/∞ 排队网络的 Web 服务组合性能分析 [J]. 计算机学报, 2013, 36 (1): 22 - 37.
- [7] 方 欢, 陆 阳, 黄镇谨, 等. 基于 CPN 仿真的排队系统建模及性能分析 [J]. 系统仿真学报, 2013, 25 (2): 228 - 234.
- [8] Pandey S, Nepal S. Cloud computing and scientific application: Big data, scalable analytics, and beyond [J]. Future Generation Computer Systems, 2013, 29 (7): 1774 - 1776.
- [9] Zhang Qi, Cheng Lu, Boutaba Raouf. Cloud computing: state-of-the-art and research challenges [J]. Internet Serv Appl, 2010, 6 (1): 7 - 18.
- [10] Gu J, Hu J, Zhao T, et al. A new resource scheduling strategy based on genetic algorithm in cloud computing environment [J]. Journal of Computers, 2012, 7 (1): 42 - 52.