



## 云计算环境下虚拟机资源均衡调度方法研究<sup>\*</sup>

李济汉<sup>1</sup>, 李素粉<sup>2</sup>, 张云勇<sup>2</sup>, 房秉毅<sup>2</sup>

(1. 北京邮电大学 北京 100876; 2. 中国联通集团研究院 北京 100048)

**摘要:** 虚拟机资源是云计算环境下的一类主要云服务资源, 有效的资源调度方法对于改善负载均衡和提高资源利用率具有重要意义。针对虚拟机资源的调度问题进行分析, 利用任务到达触发任务分配, 当任务到达时, 首先判断云服务资源池的运行状态, 分为饱和与不饱和两种; 然后针对当前任务请求下的云服务资源池状态, 分别给出资源调度策略, 并通过算例分析验证所提方法的有效性。所提虚拟机资源调度方法将有助于云服务资源池中虚拟机资源的均衡利用。

**关键词:** 云计算; 云服务资源池; 虚拟机; 资源调度

**doi:** 10.3969/j.issn.1000-0801.2013.04.016

## Research on the Resources Scheduling Method of Virtual Machines Under the Cloud Computing Environments

Li Jihan<sup>1</sup>, Li Sufen<sup>2</sup>, Zhang Yunyong<sup>2</sup>, Fang Bingyi<sup>2</sup>

(1. Beijing University of Posts and Telecommunications, Beijing 100876, China;

2. China Unicom Research Institute, Beijing 100048, China)

**Abstract:** As one kind of the main resources of cloud computing, virtual resource sustains sorts of cloud services, and an effective resource scheduling mechanism helps to balance the load and improve the resource utility rate. Focusing on virtual resource of cloud, a resource scheduling mechanism to achieve load balancing was proposed. Triggered by an incoming job, the scheduler checked the instant state of resource pool first, and then took interrelated strategies.

**Key words:** cloud computing, cloud services pool, virtual machine, resource scheduling

### 1 引言

云计算环境下, 一些大型企业纷纷由传统业务向云服务转型, 组建企业云, 形成云服务资源池。例如, 亚马逊建立专门的数据中心, 以云服务的形式向用户和开发者提供存储服务器、带宽和CPU资源等。云计算在企业的深度和广度应用, 必将催生大量云服务。以电信行业为例, 国内外电信运营企业纷纷开展云计算相关的研究和应用, 逐渐形成基础设施即服务(infrastructure as a service, IaaS)、平台

即服务(platform as a service, PaaS)和软件即服务(software as a service, SaaS)3个层次的服务资源池, 服务资源将呈现规模化和商业化特点<sup>[1,2]</sup>。这种情况下, 面对规模日益庞大的云服务资源池, 有效的云服务资源调度方法对于提高服务资源的整体利用率具有重要意义。

当虚拟化成为云计算的主要支撑技术后, 资源虚拟化对数据中心的资源复用、关联、动态管理等方面都提出了新的挑战。云计算数据中心虚拟机资源利用率问题已经成为学术界的研究热点之一。目前常用的云服务资源调度规

<sup>\*</sup> 国家自然科学基金资助项目(No.71172134)

则主要有轮转法、最小负载优化法和散列法等<sup>[3]</sup>。轮转法通常是预先设定好一个轮转周期(如物理服务器个数),依次将用户需求的虚拟机分配给不同的物理服务器,一个轮转周期结束后重新开始新一轮转;最小负载优化法是每次分配虚拟机时,检查前一段时间内有物理服务器的负载情况,总是选择当前一段时间内负载最小的物理服务器分配虚拟机;散列法预先设置一个散列函数,用于将反映用户需求的虚拟机映射到相应的物理服务器上。

基于上述基本调度规则,相关学者对云计算数据中心的虚拟机调度问题进行了大量研究。参考文献[4]针对云环境下,数据中心虚拟化资源的一致共享和多租户应用异构性之间的矛盾所带来的,服务器内部多种资源间分配不均衡问题,研究并提出了基于多维协同聚合的虚拟机调度机制。参考文献[5]针对云计算中物理服务器间的负载不均衡问题,提出一种基于多属性层次分析的虚拟机部署与调度策略,该算法将虚拟机按照资源的需求特点进行分类,根据该虚拟机量化后的权向量以及服务器资源的使用记录,对各个服务器进行预测评价,以期选择最佳服务器进行部署。参考文献[6]针对传统虚拟机调度方法中,虚拟机申请时调用相应负载均衡算法将虚拟机调度到相应的物理服务器上,不能达到整体效益最优的问题,提出了一种实时满足运营商资源最高效益的动态调度方法。基于该方法,每次新加入用户时,新加入的用户资源和先前分配的资源一起重新在所有服务器上规划,进行全局虚拟机调度,迁移那些能给系统带来足够效益的虚拟机,完成全局最优的虚拟机部署。

由于服务资源在处理业务请求时需要消耗一定的时间,随着访问服务资源次数的增加,服务请求会产生排队现象,导致等待时间的产生。在实际的业务服务选择中,排队等待时间往往是云服务资源调度的一个关键决策因素。然而现有文献鲜有提及。因此,本文提出了一种服务运行状态和服务等待时间的调度方法,并设计具体的计算规则,提高服务资源的均衡利用率。

## 2 问题描述与符号定义

本文主要针对云计算环境下虚拟机资源的调度问题进行分析。虚拟机资源是云计算环境下的一类主要云服务资源,载体是物理服务器资源。基于如图1所示的网络架构,用户向云服务资源池提出服务请求,云服务资源池根据物理服务器和虚拟机的当前运行情况,依据一定的策略将合适的虚拟机分配给当前用户,在满足用户需求的同时,使得服务器资源的总体利用率最高。

为有效提高资源池中虚拟机的平均利用率,本文针对饱和与不饱和两种资源池运行状态,分别设计虚拟机资源调度策略。为便于描述虚拟机资源调度算法,给出表1所示符号定义。

## 3 资源调度实现算法

假设云服务资源池中有 $I$ 个物理服务器 $PS_i, i=1, 2, 3, \dots, I, I$ 为正整数,每个物理服务器 $PS_i$ 上承载 $J_i$ 个虚拟服务器 $S_{ij}$ (以下简称虚拟机), $i=1, 2, 3, \dots, I, j=1, 2, 3, \dots, J_i, J_i$ 是正整数,每个虚拟机可同时并行运行 $pN_{ij}$ 个任务, $pN_{ij}>0$ ,是

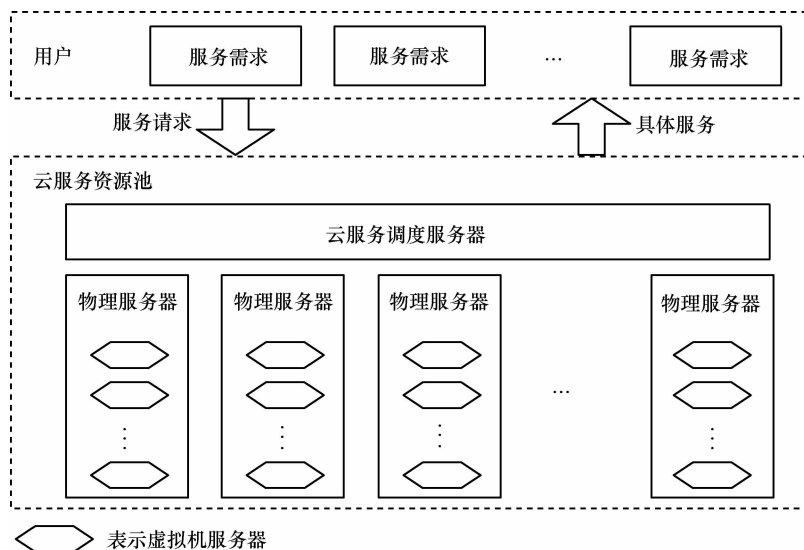


图1 云服务资源调度网络架构示意



表 1 基本符号定义

| 符号            | 意义及说明                                                                            |
|---------------|----------------------------------------------------------------------------------|
| $I$           | 云服务资源池中的物理服务器编号, $i=1, 2, 3, \dots, I$ , $I$ 是正整数                                |
| $J_i$         | 物理服务器 $i$ 上虚拟机的编号, $J_i=1, 2, 3, \dots, J_i$ , $J_i$ 是正整数, $i=1, 2, 3, \dots, I$ |
| $PS_i$        | 物理服务器                                                                            |
| $S_{i,j}$     | 运行在物理服务器 $PS_i$ 上、编号为 $J_i$ 的虚拟机                                                 |
| $pN_{i,j}$    | 每个虚拟机可同时并行运行的最大任务数, $pN_{i,j}>0$ , 是正整数                                          |
| $pn_{i,j}$    | 虚拟机 $S_{i,j}$ 上当前正在运行的任务数, $pn_{i,j} \geq 0$ , 是正整数                              |
| $PMRate_i$    | 物理服务器 $PS_i$ 的利用率                                                                |
| $queue_{i,j}$ | 当前虚拟机 $S_{i,j}$ 的任务排队数量, 是正整数                                                    |
| $t_{i,j}$     | 虚拟机 $S_{i,j}$ 运行完一个任务所需要的时间, $t_{i,j}>0$                                         |
| $wt_{i,j}$    | 虚拟机 $S_{i,j}$ 的资源等待时间                                                            |
| $st_{i,j}$    | 虚拟机 $S_{i,j}$ 上进度最快的任务的执行进度, $st_{i,j} \in (0, 1)$                               |
| $TN$          | 当前到达云服务资源池的新任务数                                                                  |

正整数。

利用任务到达触发资源分配, 当任务到达时, 基于资源调度规则, 选择一个虚拟机资源, 执行该任务。本文提出的虚拟机资源调度步骤如下。

**步骤 1** 根据资源池中物理服务器的运行状态, 判断云服务资源池的运行状态  $PoolStatus$ , 具体方法如下。

(1) 通过图 1 所示的云服务资源监控与调度服务器, 对于每个物理服务器  $PS_i$ , 获取其承载的每个虚拟机  $S_{i,j}$  上正在运行的任务数  $pn_{i,j}$ 。

(2) 计算物理服务器  $PS_i$  的利用率  $PMRate_i$ ,  $i=1, 2, 3, \dots, I$ , 计算规则如式(1):

$$PMRate_i = \frac{\sum_{j=1}^{J_i} pn_{i,j}}{\sum_{j=1}^{J_i} pN_{i,j}}, i=1, 2, 3, \dots, I \quad (1)$$

其中, 分母表示物理服务器  $PS_i$  可以并行运行的最大任务数, 分子表示当前运行的任务数。 $PMRate_i=1$ , 表示物理服务器  $PS_i$  处于饱和运行状态;  $PMRate_i=0$ , 表示物理服务器  $i$  上没有虚拟机在执行任务, 本文假设这种状态下, 物理服务器  $i$  处于未开启状态。

(3) 判定云服务资源池的运行状态  $PoolStatus$ 。对于所有物理服务器  $PS_i$ , 如果均有  $PMRate_i=1$ ,  $i=1, 2, 3, \dots, I$ , 即所有可以执行当前任务的物理服务器都开启且处于饱和运行状态, 则将云服务资源池的状态记为饱和状态, 即  $PoolStatus=1$ , 转步骤 3; 否则记为不饱和状态, 即

$PoolStatus=0$ , 转步骤 2。

**步骤 2** 对于不饱和运行的云服务资源池, 选择一个虚拟机资源执行当前任务, 具体实现步骤如下。

(1) 根据物理服务器  $PS_i$  的利用率  $PMRate_i$ , 判断是否所有物理服务器均未开启, 即当前时刻该资源池没有运行任何任务。判断方法为: 对于  $i=1, 2, 3, \dots, I$ , 均有物理服务器  $PS_i$  的利用率  $PMRate_i=0$ , 则表明所有物理服务器均未开启, 转步骤(2), 否则, 转步骤(3)。

(2) 开启第一台物理服务器  $PS_1$ , 并将当前任务安排在虚拟机  $S_{1,1}$  上运行。

(3) 在开启的物理服务器中, 根据步骤 1 中的(2)计算物理服务器利用率  $PMRate_i$ , 选择利用率最小的一台物理服务器, 转步骤(4)。存在两种特殊情况: 若同时存在多台物理服务器  $PS_i$  的利用率相同且最小, 则选择值较小的一台物理服务器, 转步骤(4); 若利用率最小的物理服务器  $PS_i$  的利用率  $PMRate_i=1$ , 也就是说所有已开启的物理服务器均饱和, 则开启下一台物理服务器, 并选择该物理服务器上的第一个虚拟机执行当前任务, 转步骤 4。

(4) 在步骤(3)选定的物理服务器上, 根据每个虚拟机  $S_{i,j}$  上正在运行的任务数  $pn_{i,j}$ , 选择  $pn_{i,j}$  最小的一台虚拟机执行当前任务, 如果同时存在多个虚拟机的  $pn_{i,j}$  相同且最小, 则选择  $j$  值较小的一个虚拟机执行当前任务。

**步骤 3** 对于饱和运行的云服务资源池, 计算所有虚拟机资源的等待时间, 选择等待时间最短的虚拟机执行当前任务; 对于候选虚拟机资源  $S_{i,j}$ , 如果当前运行实例数  $pn_{i,j}$  等于最大并行实例数  $pN_{i,j}$ , 则新的任务请求需要排队等待。对于饱和运行状态的云服务资源池, 每个虚拟机资源均处于饱和运行状态, 虚拟机资源等待时间的计算方法如下。

(1) 由调度服务器检测虚拟机资源  $S_{i,j}$  的运行状态并自动获取相应参数值, 包括虚拟机  $S_{i,j}$  当前运行实例数  $pn_{i,j}$ 、虚拟机  $S_{i,j}$  上进度最快的任务的执行进度  $st_{i,j}$ 、当前虚拟机  $S_{i,j}$  的任务排队数量  $queue_{i,j}$ 、虚拟机  $S_{i,j}$  运行完一个任务所需要的时间  $t_{i,j}$ 。

(2) 计算虚拟机  $S_{i,j}$  资源等待时间  $wt_{i,j}$ , 计算方法如式(2)所示:

$$wt_{i,j} = \left\lceil \text{int} \left( \frac{queue_{i,j}}{pN_{i,j}} \right) + (1 - st_{i,j}) \right\rceil \times t_{i,j} \quad (2)$$

(3) 选择等待时间  $wt_{i,j}$  最短的虚拟机执行当前任务, 若同时存在多个虚拟机的等待时间相同且最小, 则选择  $i, j$

值较小的虚拟机资源。

步骤4 结束。

## 4 算例分析

为了评测算法性能,采用现有的策略与本方法进行比较。最小负载优化法是目前常用的一种均衡调度方法,即每次分配虚拟机时,检查前一段时间内有物理服务器的负载情况,总是选择当前一段时间内负载最小的物理服务器分配虚拟机。本文所提的调度方法实质上是对最小负载优化方法的一种改进方法。本节通过算例分析,将本文所提的虚拟机资源均衡调度方法与最小负载优化法进行比较分析。

假设云服务资源池包含  $I$  个物理服务器节点,令  $I=20$ ,不失一般性,假设  $I$  个物理服务器上的虚拟机数量相同且  $J=5, i=1,2,3,\dots,I$ 。对于云服务资源池的初始状态,采用随机方法设置。根据第3节步骤3的(1),描述云服务资源池当前运行状态的参数包括:虚拟机  $S_{ij}$  当前运行实例数  $pn_{ij}$ 、虚拟机  $S_{ij}$  上进度最快的任务的执行进度  $st_{ij}$  以及当前虚拟机  $S_{ij}$  的任务排队数量  $queue_{ij}$ 。若  $\sum_{j=1}^J pn_{ij}=0$ ,表示物理服务器  $PS_i$  处于关闭或者待机状态。对于当前需求情况,设计4种不同的当前任务数,分别是  $TN=1, 5, 20, 50$ 。

计算指标分为两种:物理服务器的平均利用率  $\mu$  和物理服务器的利用率偏差  $\sigma$ ,前者计算方法如式(3)描述,反映物理资源的利用率情况;后者采用均方差的计算方法,如式(4)描述,计算结果反映一个数据集的离散程度,在这里用于描述各个物理服务器资源的均衡利用程度。

$$\mu = \frac{1}{I-I_0} \sum_{i=1}^I PMRate_i \quad (3)$$

其中,分母  $I_0$  表示处于关闭状态的物理服务器数量,即  $PMRate_i=0$  时的物理服务器数量。

$$\sigma = \sqrt{\frac{1}{I-I_0} \sum_{i=1}^I Y_i (PMRate_i - \mu)^2} \quad (4)$$

其中,  $Y_i \in \{0, 1\}$ , 当  $PMRate_i=0$  时,  $Y_i=0$ , 否则  $Y_i=1$ 。

分别采用最小负载优化法和本文所提的调度方法,针对4种不同的当前任务需求,分别对上述算例进行计算,两种计算指标计算结果分别如表2和图2、图3所示。

表2和图2、图3显示,对于不同的当前任务请求数量,的值均优于或不劣于,且的值均优于或不劣于。表明本

表2 两种调度策略下平均利用率  $\mu$  和偏差  $\sigma$  计算结果

| 当前需求       | $TN=1$ | $TN=5$ | $TN=20$ | $TN=50$ |
|------------|--------|--------|---------|---------|
| $\mu_1$    | 0.62   | 0.58   | 0.67    | 0.71    |
| $\mu_2$    | 0.62   | 0.65   | 0.73    | 0.82    |
| $\sigma_1$ | 0.051  | 0.049  | 0.042   | 0.039   |
| $\sigma_2$ | 0.051  | 0.043  | 0.037   | 0.030   |

注:  $\mu_1$  和  $\sigma_1$  表示采用最小负载优化方法时的计算结果;  $\mu_2$  和  $\sigma_2$  表示采用本文所提调度方法时的计算结果。

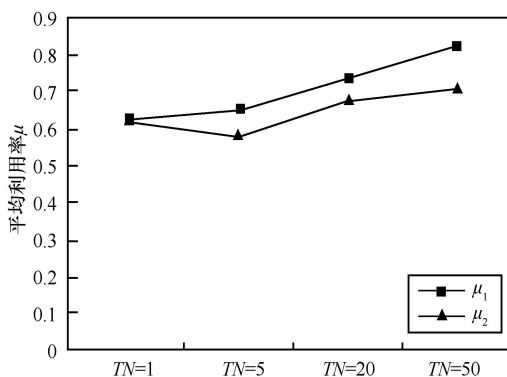


图2 两种调度策略下平均利用率  $\mu$  曲线

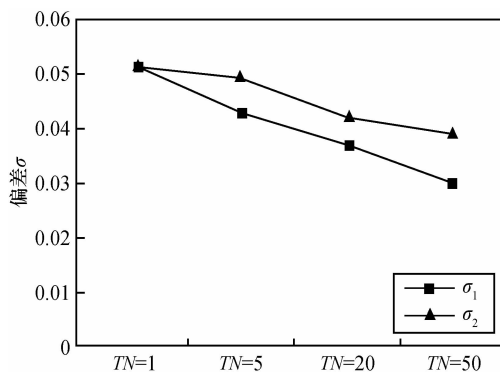


图3 两种调度策略下偏差  $\sigma$  计算结果曲线

文所提的调度方法优于一般的最小负载优化方法,且随着当前请求任务数的增多,优势愈加明显。本文所提调度方法实质上是对最小负载优化方法的一种改进。通过该算例分析,进一步验证本文所提的虚拟机资源调度方法相对一般的最小负载优化方法具有改进效果。

## 5 结束语

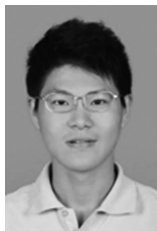
针对云计算环境下的虚拟机资源调度问题,基于云服务资源池的虚拟机资源利用率和等待时间等状态参数,设计一种面向虚拟机资源的均衡调度算法,算例分析表明所提方法的有效性。该算法对于提高虚拟机资源的均衡利用具有重要作用。



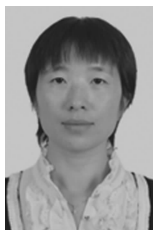
## 参考文献

- 1 童晓渝, 张云勇, 房秉毅等. 电信运营商实施云计算的策略建议. 信息技术, 2012(1)
- 2 袁文成, 朱怡安, 陆伟. 面向虚拟资源的云计算资源管理机制. 西北工业大学学报, 2010, 28(5)
- 3 田文洪, 赵勇. 云计算——资源调度管理. 北京: 国防工业出版社, 2011
- 4 孙鑫. 面向云环境数据中心的高效资源调度机制研究. 北京邮电大学博士学位论文, 2012
- 5 庄威, 桂小林, 林建材等. 云环境下基于多属性层次分析的虚拟机部署与调度策略. 西安交通大学学报, 2013(2)
- 6 尹红军, 李京, 宋许等. 云计算中运营商效益最优的资源分配机制. 华中科技大学学报(自然科学版), 2011(S1)

## [作者简介]



**李济汉**, 男, 北京邮电大学本科在读, 主要研究方向为人工智能、机器学习和云计算等。



**李素粉**, 女, 博士, 现就职于中国联通集团研究院, 主要研究方向为云计算相关技术。



**张云勇**, 男, 博士后, 中国联通集团研究院高级工程师、平台与云计算研究中心主任, 主要研究方向为云计算、下一代开放网络、固定移动融合核心网、移动互联网及业务、公共运算。



**房秉毅**, 男, 博士, 中国联通集团研究院高级工程师, 主要从事云计算、核心网新技术的研究工作。

(收稿日期: 2013-04-01)

(上接第 77 页)

同步 HARQ 和异步 HARQ 过程。设计出一种求得 HARQ RTT 的方案, 仿真分析了不同配置和 HARQ 进程利用率对 HARQ RTT 的影响。本文将有助于对下行 HARQ 的研究和理解, 并且可根据对比仿真图中的 HARQ RTT, 为网络侧进行不同业务时的上下行配置提供参考。

## 参考文献

- 1 3GPP TS 36.211 V9.0.1. Evolved Universal Terrestrial Radio Access (E-UTRA); Physical Channels and Modulations, 2009
- 2 Dahlman E, Parkvall S, Skold J, *et al.* 3G Evolution: HSPA and LTE for Mobile Broadband, 2nd Ed. Academic Press, 2008
- 3 3GPP TS 36.321 V9.1.0. Evolved Universal Terrestrial Radio Access (E-UTRA) Medium Access Control (MAC) Protocol Specification, 2009
- 4 3GPP TS 36.213 V9.0.1. Evolved Universal Terrestrial Radio Access (E-UTRA); Physical Layer Procedures, 2009
- 5 Susitaival R, Meyer M. LTE coverage improvement by TTI bundling. Proceedings of VTC 2009-Spring, Spain, 2009

## [作者简介]



**陈发堂**, 男, 硕士, 重庆邮电大学副教授、硕士生导师, 重庆邮电大学“重庆移动通信工程研究中心”算法研究部副主任, 主要研究方向为 TD-SCDMA/GSM 双模终端开发项目的物理层算法研究。



**万翠**, 女, 重庆邮电大学硕士研究生, 主要研究方向为 LTE 高层协议栈开发。

**樊玺**, 女, 电子科技大学硕士研究生, 主要研究方向为通信网络技术研究。

(收稿日期: 2012-12-30)