

# 控制与决策

Control and Decision

## 异构云环境下AHP定权的多目标强化学习作业调度方法

袁景凌, 陈旻骋, 江涛, 李超

引用本文:

袁景凌, 陈旻骋, 江涛, 等. 异构云环境下AHP定权的多目标强化学习作业调度方法[J]. *控制与决策*, 2022, 37(2): 379–386.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2020.0911>

---

### 您可能感兴趣的其他文章

Articles you may be interested in

#### 复杂人力资源约束下的抢占式维修工序调度

Preemptive maintenance process scheduling under complex human resource constraints

*控制与决策*. 2022, 37(2): 393–400 <https://doi.org/10.13195/j.kzyjc.2020.1250>

#### 面向建材装备集团制造的分布式多项目资源调度

Distributed multi-project resource scheduling oriented to manufacturing of building materials equipment group

*控制与决策*. 2021, 36(9): 2133–2142 <https://doi.org/10.13195/j.kzyjc.2019.1802>

#### 基于正态云模型的状态转移算法求解多目标柔性作业车间调度问题

State transition algorithm based on normal cloud model for solving multi-objective flexible job shop scheduling problem

*控制与决策*. 2021, 36(5): 1181–1190 <https://doi.org/10.13195/j.kzyjc.2019.1233>

#### 铁路集装箱中心站资源分配与作业调度联合优化

Integrating optimization of resource allocation and handling scheduling in railway container terminal

*控制与决策*. 2021, 36(12): 3063–3073 <https://doi.org/10.13195/j.kzyjc.2020.0597>

#### 基于深度强化学习与迭代贪婪的流水车间调度优化

Scheduling optimization for flow-shop based on deep reinforcement learning and iterative greedy method

*控制与决策*. 2021, 36(11): 2609–2617 <https://doi.org/10.13195/j.kzyjc.2020.0608>

# 异构云环境下 AHP 定权的多目标强化学习作业调度方法

袁景凌<sup>1†</sup>, 陈旻骋<sup>1</sup>, 江涛<sup>1</sup>, 李超<sup>2</sup>

(1. 武汉理工大学 计算机科学与技术学院, 武汉 430070;  
2. 上海交通大学 电子信息与电气工程学院, 上海 200240)

**摘要:** 随着新型基础设施建设(新基建)的加速,云计算将获得新的发展契机. 数据中心作为云计算的基础设施,其内部服务器不断升级换代,这造成计算资源的异构化. 如何在异构云环境下,对作业进行高效调度是当前的研究热点之一. 针对异构云环境多目标优化调度问题,设计一种 AHP 定权的多目标强化学习作业调度方法. 首先定义执行时间、平台运行能耗、成本等多个目标,其中定义服务延迟成本以描述用户对服务质量的满意程度;然后设计面向异构资源的多目标调度综合评价方法,利用层次分析法(analytic hierarchy process, AHP)确定各个目标的权重;最后将该方法引入 Q-learning 的奖励值计算,使其能反映异构云环境下作业的总体执行情况,并对后续抵达的作业起到良好的经验学习作用. 实验结果表明,所提出的方法优于大部分对比方法,能够较好地优化作业执行效率和保障用户及服务提供商的利益.

**关键词:** 强化学习; 多目标; 作业调度; 异构资源; 服务延迟成本

**中图分类号:** TP311      **文献标志码:** A

**DOI:** 10.13195/j.kzyjc.2020.0911

**引用格式:** 袁景凌,陈旻骋,江涛,等. 异构云环境下 AHP 定权的多目标强化学习作业调度方法[J]. 控制与决策, 2022, 37(2): 379-386.

## Multi-objective reinforcement learning job scheduling method using AHP fixed weight in heterogeneous cloud environment

YUAN Jing-ling<sup>1†</sup>, CHEN Min-cheng<sup>1</sup>, JIANG Tao<sup>1</sup>, LI Chao<sup>2</sup>

(1. School of Computer Science and Technology, Wuhan University of Technology, Wuhan 430070, China; 2. School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China)

**Abstract:** With the acceleration of new infrastructure, cloud computing will be given an entirely new opportunity to develop. Data centers, as the infrastructure of cloud computing, their internal servers are continuously updated, which leads to the heterogeneity of computing resources. How to efficiently schedule jobs in heterogeneous cloud environment has become an increasingly popular research. This paper designs a multi-objective reinforcement learning job scheduling method using AHP fixed weight in heterogeneous cloud environment. First, we define the execution time, energy consumption, execution cost, etc. Service delay cost is used to describe the user satisfaction for service. Then, a comprehensive evaluation method for multi-objective scheduling is designed. The weight coefficient of each object is determined by the analytic hierarchy process (AHP). The method is introduced into the calculation of rewards, which can reflect the overall situation and serve as an excellent learning tool for the following jobs. The experimental results show that the proposed method can better optimize the execution efficiency, while ensuring the interests of users and service providers.

**Keywords:** reinforcement learning; multi-objective; job scheduling; heterogeneous resources; service delay cost

## 0 引言

随着新基建建设的提速,数据中心建设步伐明显加快. 普通数据中心一般由 5~10 万台服务器组成,但国内外著名互联网公司的数据中心规模远不止如

此,谷歌全球数据中心的服务器总量已超百万台,为用户提供搜索服务;亚马逊公司仅弹性计算云(elastic compute cloud, EC2)就部署了约 50 万台服务器;国内腾讯公司天津数据中心在 2018 年服务器总数突破 10

收稿日期: 2020-07-06; 录用日期: 2020-12-03.

基金项目: 国家自然科学基金项目(61303029); 湖北省创新团队项目(2015CFA069); 湖北省技术创新专项重大项目(2017AAA122).

责任编辑: 胡清华.

<sup>†</sup>通讯作者. E-mail: yuanjingling@126.com.

万台;国内阿里巴巴将在南通建立云计算数据中心,计划容纳30万台服务器.数据中心规模越来越大,造成计算资源的异构化和规模化,给大规模数据中心的作业调度带来一定的挑战.云计算数据中心作为商业计算服务平台,首先,需要满足各种不同的用户需求.例如,有些用户需要快速的服务响应时间,有些用户需要服务的稳定性和可靠性,有些更倾向于低廉的计算成本,与此同时还要保障用户的服务体验<sup>[1]</sup>.再者,云计算的商业性决定了调度策略需保障服务供应商的利益,减少能耗及相应开销.传统的调度方法已不能适应其复杂多变的调度环境,因此,研究云计算异构环境下的多目标优化调度策略具有相当重要的意义.

针对异构云环境多目标优化调度问题,本文定义了执行时间、服务延迟成本、平台运行能耗、成本等多个目标,其中服务延迟成本用以描述用户对服务质量的满意程度;设计AHP定权的多目标调度综合评价方法,并将该方法引入强化学习Q-learning的奖励值计算,有利于后续作业的经验学习;利用MultiRECloudSim平台进行仿真实验,实验结果表明所提出方法优于大部分对比方法,能够较好地改善作业执行效率并保证用户及服务提供商的利益.

## 1 相关工作

云数据中心作业调度主要分为3类:传统作业调度、多目标优化作业调度和基于机器学习的作业调度.

### 1.1 传统作业调度方法

传统作业调度方法一般优先考虑性能,被广泛应用的算法有先来先服务调度算法、max-min调度算法、min-min调度算法、公平调度算法以及单目标启发式调度算法等.这类算法一般不考虑资源需求及利用率,只为查找满足计算需求的节点来执行作业.但是,实际的作业调度问题往往有多个目标,例如最短执行时间、最佳服务质量、最少计算成本等,因此这类方法实用性较为欠缺.

### 1.2 多目标优化作业调度方法

考虑到云环境下的多个优化目标,并将作业映射至合适的云计算资源,多目标优化作业调度方法较传统的作业调度方法更具实用性<sup>[2-3]</sup>.近年来,随着云服务快速普及,多目标优化作业调度方法受到广泛关注,并已取得一定成果. Akbari等<sup>[4]</sup>设计了基于布谷鸟搜索的多目标作业调度算法,该算法不但减少了异构系统中的作业执行时间,而且最大程度保障了作业并行化.胡成玉等<sup>[5]</sup>建立了以利润、碳排放以及

QoS为目标的数据中心负载优化调度模型,并设计出一种非支配排序遗传算法,保证算法能快速收敛. Zhu等<sup>[6]</sup>以作业完成时间和成本作为调度目标,设计了基于演化算法的多目标调度算法以优化IaaS (infrastructure as a service)平台的工作流调度.苏命峰等<sup>[7]</sup>结合能耗、成本和系统可用性等指标构建了多维QoS云资源调度模型,并提出结合多重贪心算法的云资源调度算法.上述方法取得了比传统作业调度方法更好的解,但同时也存在一些问题.现有的调度策略研究较少兼顾用户和服务供应商双方的利益,如何使资源供给与消费双方互利共赢是当前云作业调度需考虑的问题.

### 1.3 基于机器学习的作业调度方法

基于机器学习的作业调度方法学习数据中心资源消耗、性能变化与负载波动之间的关系,探索高效执行作业时所需资源.在实际执行调度算法时,根据学习结果与实时监控状况调整作业分配方案<sup>[8]</sup>. Islam等<sup>[9]</sup>利用神经网络和线性回归设计了资源预测和供应策略,实现了云作业资源的按需分配. Chen等<sup>[10]</sup>设计了基于动态加权随机森林的负载预测方法,并根据预测结果进行自适应作业调度.上述方法通过预测负载波动调节计算资源规模,为减少处于活跃状态的服务器,需频繁地进行虚拟机迁移,这会对云作业调度造成一定程度的影响.

此外,基于强化学习的作业调度方法也是研究热点. Yi等<sup>[11]</sup>利用LSTM网络构建系统状态预测模型,并使用该模型训练基于深度强化学习的作业分配器,有效地为数据中心节省近10%的计算能力. Yang等<sup>[12]</sup>提出了基于强化学习的资源分配方法,并通过优化长期奖励以改善工作流调度效率.

本文介绍的方法充分考虑到节能增效、用户体验、执行成本等多个目标之间的平衡,兼顾用户和服务供应商双方利益,着眼于云计算资源的异构性及作业多样化,利用强化算法对云数据中心作业进行智能自主管理.

## 2 多目标与约束定义

**定义1(异构资源约束)** 云数据中心由众多异构服务器构成,这些服务器被虚拟化成虚拟机,虚拟机 $V_i$ 有 $d$ 个维度资源,可表示为 $V_i = \{V_i^1, V_i^2, \dots, V_i^d\}$ .同理,定义作业 $T_j$ 所需的资源为 $T_j = \{T_j^1, T_j^2, \dots, T_j^k\}$ .虚拟机与作业之间的放置关系可用调度矩阵 $X = [x_{ij}]$ 表示,当 $T_j$ 作业分配到虚拟机 $V_i$ 时, $x_{ij} = 1$ ,否则 $x_{ij} = 0$ .通过上述对异构资源的分配关系,可以获得异构资源环境下作业调度的资源约束目标为

$$\sum_{j=1}^m T_j^k \cdot x_{ij} \leq V_i^k. \quad (1)$$

式(1)反映了表示作业执行所消耗的全部资源量应小于等于平台的总资源量。

**定义2** (执行时间) 执行时间是作业调度中最为普遍的优化指标, 作业 $T_j$ 的执行时间 $t_e(T_j, V_i)$ 可计算为

$$t_e(T_j, V_i) = \frac{\text{len}(T_j)}{V_i^{\text{CPU}}}. \quad (2)$$

其中: $\text{len}(T_j)$ 为作业 $T_j$ 的长度; $V_i^{\text{CPU}}$ 为虚拟机 $V_i$ 中的CPU执行能力. 假定每个任务只能在一台虚拟机上执行, 云平台的总执行时间 $t_{\text{all}}$ 为

$$t_{\text{all}} = \max \left\{ \sum_{j=1}^{m_1} t_e(T_j, V_1), \dots, \sum_{j=1}^{m_n} t_e(T_j, V_n) \right\}, \quad (3)$$

其中 $m_i$ 为全部被虚拟机 $V_i$ 成功执行的作业数目,  $\sum_{i=1}^n m_i = m$ . 对于云平台而言, 总执行时间越少, 用户的满意程度越高.

**定义3** (服务延迟成本) 对于云计算平台管理者而言, 更关注平台的整体性能, 如资源负载均衡、能耗等. 对于提交作业的用户而言, 更关心的是服务执行效果. 用户访问数据中心, 不仅需要考虑服务价格, 同时也要考虑服务质量. 本文利用服务延迟成本描述用户对服务质量的满意程度. 作业等待响应的时间越长, 延迟成本越大, 服务体验就相对较差. 服务延迟成本可表示为作业等待响应时间的函数. 作业 $T_j$ 的等待响应时间 $t_w(T_j)$ 可计算为

$$t_w(T_j) = t_{\text{est}}(T_j) - t_{\text{ar}}(T_j) + t_s(t_w). \quad (4)$$

其中: $t_{\text{est}}(T_j)$ 为作业开始执行的时间; $t_{\text{ar}}(T_j)$ 为作业到达时间; $t_s(t_w)$ 为调度算法 $t_w$ 执行时长. 作业 $T_j$ 的服务延迟成本 $\text{cost}_d(T_j)$ 可计算为

$$\text{cost}_d(T_j) = \begin{cases} 0, & t_w(T_j) < \psi t_{\text{rt}}; \\ L(t_w(T_j)), & t_w(T_j) \geq \psi t_{\text{rt}}. \end{cases} \quad (5)$$

$$L(t_w(T_j)) = \text{cost}_{d/u} \cdot (t_w(T_j) - \psi \cdot t_{\text{rt}}). \quad (6)$$

其中: $t_{\text{rt}}$ 为用户的可容忍等待时间; $\psi$ 为自适应调节因子, 用户可容忍等待时间在不同时间段是不同的; $\text{cost}_{d/u}$ 为单位时间内服务延迟成本.  $\text{cost}_d(T_j)$ 越小, 表明用户体验越好. 总服务延迟成本 $\text{cost}_d^{\text{all}}$ 可计算为

$$\text{cost}_d^{\text{all}} = \sum_{j=1}^m \text{cost}_d(T_j). \quad (7)$$

**定义4** (平台运行能耗) 利用CPU利用率-功率函数<sup>[13]</sup>计算数据中心运行能耗 $\text{energy}_{\text{all}}$ . 服务供应商不仅需考虑服务器能耗, 还要考虑制冷、照明等能耗. 因此数据中心能耗利用能源效率指标(power usage effectiveness, PUE)进行估算<sup>[14]</sup>, 具体计算为

$$\text{energy}_{\text{all}} = t_{\text{all}} \cdot \sum_{i=1}^p P_i(U_i). \quad (8)$$

其中: $p$ 为数据中心内活跃的服务器数量; $P_i(\cdot)$ 为第 $i$ 台服务器的CPU利用率-功率函数; $U_i$ 为第 $i$ 台服务器的CPU利用率, 其计算公式为

$$U_i = \frac{1}{c} \cdot \sum_{k=1}^c \text{VM}_k^{\text{CPU}}. \quad (9)$$

其中: $c$ 为部署在第 $i$ 台服务器上的虚拟数量; $\text{VM}_k^{\text{CPU}}$ 为该虚拟机CPU利用率.

**定义5** (成本) 总成本 $\text{cost}_{\text{all}}$ 包括用户成本 $\text{cost}_u$ 和服务供应商成本 $\text{cost}_p$ . 用户成本主要为作业执行成本, 服务供应商成本主要为能耗成本. 因此总成本的具体计算公式为

$$\begin{aligned} \text{cost}_{\text{all}} &= \text{cost}_u + \text{cost}_p = \\ & \sum_{j=1}^m \text{cost}_v(V_i) \cdot t_e(T_j, V_i) + \text{energy}_{\text{all}} \cdot \text{cost}_{\text{ep}}. \end{aligned} \quad (10)$$

其中: $\text{cost}_v(V_i)$ 为虚拟机 $V_i$ 的单位计算成本,  $\text{cost}_{\text{ep}}$ 表示单位电力成本.

定义作业调度过程中的多个目标, 其调度目标如下:

$$\min \{ t_{\text{all}}, \text{energy}_{\text{all}}, \text{cost}_{\text{all}}, \text{cost}_d^{\text{all}} \}. \quad (11)$$

需要满足的约束条件如下:

$$\sum_{j=1}^m T_j^k \cdot x_{ij} \leq V_i^k. \quad (12)$$

### 3 AHP定权的多目标强化学习作业调度

#### 3.1 马尔可夫决策过程

根据排队模型, 动态到达的作业调度过程可视为马尔可夫决策过程, 因此构建马尔可夫决策过程的四元组 $E = \langle X, A, P, R \rangle$ . 作业调度的状态空间为 $X = \{ \text{schedule}(T_j), \text{queue}(V_i), \text{excute} \}$ , 包括3类状态: 调度状态 $\text{schedule}(T_j)$ 、在虚拟机前队列等待 $\text{queue}(V_i)$ 和执行状态 $\text{excute}$ .  $A$ 为各个作业 $T_j$ 的动作; $P(x, x', a)$ 为作业执行动作 $a$ 后到达状态 $x'$ 的概率; $R$ 为奖励函数, 具体为

$$\begin{aligned} R_1 : A \times X &= \text{schedule}(T_j) \rightarrow \text{queue}(V_i), \\ R_2 : A \times X &= \text{queue}(V_i) \rightarrow \text{excute}. \end{aligned} \quad (13)$$

当作业从调度状态转换为等待状态, 获得奖励 $R_1$ ; 当作业变成执行状态时, 则获得奖励 $R_2$ .

可以发现, 状态空间 $X$ 的规模为作业数量 $m$ 和虚拟机数量 $n$ 的乘积. 奖励 $R_2$ 需要很长的反馈过程, 且计算较为复杂, 本文研究重点在调度过程, 因此在后续研究中, 奖励 $R$ 即因调度动作而产生的奖励 $R_1$ . 奖励 $R$ 是根据虚拟机前的作业等待队列长度确定的, 即根据作业被分配至虚拟机 $V_i$ 上的数量, 可计

算为

$$R_{x \rightarrow x'}^a = k \cdot \text{size}(\text{queue}(V_i)). \quad (14)$$

其中:  $R_{x \rightarrow x'}^a$  为执行动作  $a$  后, 状态从  $x$  转变成  $x'$  所获得的奖励;  $k$  为常数, 用于调节  $k$  的奖励  $R$  值大小. 但仅依据虚拟机前作业等待队列长度确定奖励值  $R$ , 则忽略了作业调度的多个优化目标, 为此将面向异构资源的多目标调度综合评价引入奖励  $R$  的计算策略.

### 3.2 AHP 定权的多目标调度综合评价

首先需构建面向异构资源的多目标调度综合评价体系, 该体系是简化复杂调度问题的关键. 该评价体系分为3层: 最上层为目标层; 中间层为对象层, 包括用户与服务供应商; 最下层为指标层, 包括用户所关注的指标(作业执行时间、成本(作业执行成本)和服务延迟成本)以及服务供应商所关注的指标(成本(平台能耗成本)和平台运行能耗). 该评价体系的层次结构详见图1.

在建立多目标调度综合评价体系之后, 还需确定评价体系中各指标的权重. 对此, 本文利用层次分析法设置其初始权重, 具体步骤如下:

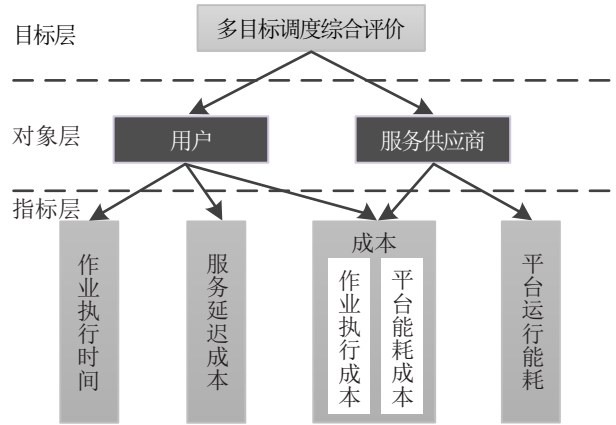


图1 异构资源的多目标调度综合评价体系结构

1) 建立判断矩阵.

假设有  $n$  个指标  $I_1, I_2, \dots, I_n$  对上一层要素有影响, 利用 T. L. Saaty 的9分位标度<sup>[15]</sup>, 两两比较同层指标, 获得相对于上层要素的判断矩阵  $A = (a_{ij})_{m \times n}$ . 其中:  $a_{ij} > 0, a_{ij} = 1/a_{ji}, a_{ii} = 1$ .

利用9分位标度<sup>[15]</sup>方法, 将本文所建立的多目标调度综合评价体系结构中的对象层和指标层内的各个指标进行两两比较, 比较结果如表1所示.

表1 各指标对比结果

指标对比	标度值	选取理由
用户 ( $I_1$ ) → 服务供应商 ( $I_2$ )	$a_{12} = 2(A_{11})$	当前云计算越来越关注用户的服务体验, 因此在对象层中, 用户稍微重要于服务供应商
作业执行时间 ( $I_1$ ) → 服务延迟成本 ( $I_2$ )	$a_{12} = 5(A_{21})$	作业执行时间越短, 服务延迟成本一般就相对越低, 因此作业执行时间比服务延迟成本明显重要
作业执行时间 ( $I_1$ ) → 成本 ( $I_3$ )	$a_{13} = 2(A_{21})$	在尽早完成作业的前提下, 用户希望开销的成本较少, 因此作业执行时间比成本(作业执行成本)次稍微重要
成本 ( $I_3$ ) → 服务延迟成本 ( $I_2$ )	$a_{32} = 3(A_{21})$	成本(作业执行成本)越高, 服务供应商收取用户的费用就越多. 若超过用户的心理价位, 将会流失用户, 因此成本比服务延迟成本稍微重要
成本 ( $I_1$ ) → 平台运行能耗 ( $I_2$ )	$a_{12} = 1(A_{22})$	对于服务供应商而言, 成本(平台能耗成本)和平台运行能耗同等重要

根据表1得到如下3个判断矩阵:

$$A_1 = \begin{bmatrix} 1 & 2 \\ 1/2 & 1 \end{bmatrix}, \quad (15)$$

$$A_{21} = \begin{bmatrix} 1 & 5 & 2 \\ 1/5 & 1 & 1/3 \\ 1/2 & 3 & 1 \end{bmatrix}, \quad (16)$$

$$A_{22} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}. \quad (17)$$

其中: 矩阵  $A_1$  为对象层的判断矩阵, 矩阵  $A_{21}$ 、 $A_{22}$  分别为指标层中针对用户和服务供应商的判断矩阵.

2) 一致性检验.

判断矩阵客观地反映出不同指标间的相对重要性, 但无法保证每个判断矩阵都是一致的. 例如, 判断

指标  $I_1$  比  $I_2$  重要,  $I_2$  比  $I_3$  重要, 那么  $I_3$  比  $I_1$  重要就明显不合理. 本文利用一致性指标  $CI$  衡量判断矩阵是否拥有完全一致性,  $CI$  具体计算公式为

$$CI = \frac{\lambda_{\max} - n}{n - 1}. \quad (18)$$

其中:  $n$  为判断矩阵的阶数;  $\lambda_{\max}$  为判断矩阵的最大特征值. 若  $CI = 0$ , 则表示判断矩阵拥有完全一致性; 反之, 则需计算随机一致性比率  $CR = CI/RI$ ,  $RI$  为平均随机一致性指标, 其取值见表2. 当  $CR < 0.1$  时, 判断矩阵的一致性可以被接受, 否则判断矩阵需进行调整. 本文判断矩阵  $A_1$ 、 $A_{21}$ 、 $A_{22}$  一致性检测如下:

$A_1$  的  $\lambda_{\max} = 2, CI = (\lambda_{\max} - n)/(n - 1) = (2 - 2)/(2 - 1) = 0$ , 所以  $A_1$  是完全一致的.

$A_{21}$  的  $\lambda_{\max} = 3.0037, CI = (\lambda_{\max} - n)/(n - 1) =$

$(3.0037 - 3)/(3 - 1) = 0.00185 \neq 0$ , 因此还需计算CR,  $CR = CI/RI = 0.00185/0.52 = 0.0036 < 0.1$ , 所以  $A_{21}$  的一致性可以被接受.

$A_{22}$  的  $\lambda_{\max} = 2$ ,  $CI = (\lambda_{\max} - n)/(n - 1) = (2 - 2)/(2 - 1) = 0$ , 所以  $A_{22}$  是完全一致的.

表2 平均随机一致性指标

$n$	1	2	3	4	5	6	7	8
RI	0	0	0.52	0.89	1.12	1.24	1.36	1.41

3) 计算相对权重.

计算各指标相对权重, 即求各判断矩阵的特征向量, 也就是计算符合  $A \cdot W = \lambda_{\max} \cdot W$  的特征向量, 并将  $W_{\lambda_{\max}}$  归一化, 即得到其相对权重. 文中判断矩阵  $A_1$  的相对权重为 (0.67, 0.33),  $A_{21}$  的相对权重为 (0.58, 0.11, 0.31),  $A_{22}$  的相对权重为 (0.5, 0.5).

4) 计算综合权重.

计算各判断矩阵的相对权重之后, 还需计算确定指标层各指标相对于总目标的权重, 计算方法为

$$w_i = \sum_{j=1}^m w_{ij}^{(3)} \cdot w_j^{(2)}. \quad (19)$$

其中:  $w_{ij}^{(3)}$  表示指标层第  $i$  个指标相对于对象层中第  $j$  个指标的相对权重;  $w_j^{(2)}$  表示对象层第  $j$  个指标相对于总目标的权重. 根据式 (19) 可得指标层各指标相对于总目标的综合权重为 (0.39, 0.07, 0.375, 0.165). 图2为综合权重计算过程.

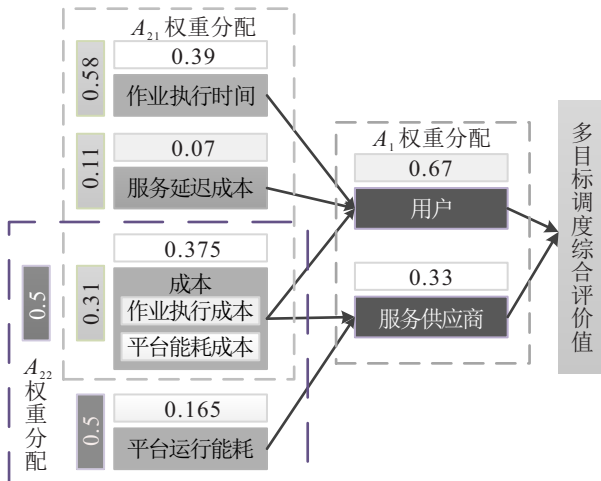


图2 综合权重计算过程

最后, 结合综合权重可得异构资源的多目标调度综合评价值计算公式, 即

$$T_{\text{value}} = 0.39 \cdot t_{\text{all}} + 0.07 \cdot \text{cost}_d^{\text{all}} + 0.375 \cdot \text{cost}_{\text{all}} + 0.165 \cdot \text{energy}_{\text{all}}. \quad (20)$$

需要注意的是, 在实际应用中各指标需进行归一化处理. 将异构资源的多目标调度综合评价值  $T_{\text{value}}$

引入多目标调度奖励值  $R$  的计算策略, 得出奖励值计算公式, 即

$$R_{x \rightarrow x'}^a = k \cdot \text{size}(\text{queue}(V_i)) \cdot T_{\text{value}}. \quad (21)$$

利用式 (21) 计算出的多目标调度奖励值  $R$  能够反映出异构云环境下作业的执行时间、成本、能耗等情况, 对后续待分配作业起到良好的经验学习作用.

### 3.3 基于强化学习的调度方案

强化学习 Q-learning 算法的基本思想是根据状态和动作构造一张  $Q$  值表保存  $Q$  值, 然后从  $Q$  值表找出使得作业调度收益最大的动作. 本文 Q-learning 算法在对下一步状态  $x'$  进行探索时采用了  $\epsilon$  贪心策略, 即有  $\epsilon$  的概率随机挑选一个动作, 余下  $1 - \epsilon$  的概率挑选效用值  $Q$  最大的动作.  $\epsilon$  一般取 0.01 ~ 0.1. 具体的  $Q$  值更新公式如下:

$$Q_{t+1}(x, a) = Q_t(x, a) + \alpha \cdot (R_{x \rightarrow x'}^a + \gamma \cdot Q_t(x', a') - Q_t(x, a)). \quad (22)$$

其中:  $\alpha$  为更新步长,  $\alpha$  越大表明越靠后的累计奖励越重要;  $\gamma$  为折扣奖赏权重,  $\gamma$  值越大表明之前的经验越重要. AHP 定权的 Q-learning 多目标作业调度的具体步骤详见算法 1.

**算法 1** AHP 定权的 Q-learning 多目标作业调度.

输入: 第  $w$  批次队列 List  $T$ , 虚拟机队列 List  $V$ ,  $Q$  值表;

输出: 第  $w$  批的调度方案 assignedList.

1. begin
2. 构建状态空间  $X$ ;
3. 选择第一个作业  $T_1$  作为开始
4. for  $T_j$  in List  $T$  do
5. if (random  $< 1 - \epsilon$ ) then
6. 在  $Q$  值表中选择最大  $Q$  值所对应的动作  $a$ ;
7. else
8. 为作业  $T_j$  随机选择一个动作  $a$ ;
9. end if
10. 执行动作  $a$ , 并获取下阶段状态  $x'$  和奖励  $R$ ;
11. 根据下阶段状态, 选择最大  $Q$  值所对应的动作  $a$ ;
12. 根据式 (22) 更新当前状态-动作的  $Q$  值;
13. 将作业与虚拟机的映射关系放入 assignedList;
14. 令下一个状态  $x'$  变为  $x$ ;
15. end for
16. 输出 assignedList;
17. end

算法 1 在初始化  $Q$  值表过程中, 时间复杂度和状态数量、动作空间密切相关, 该步骤的时间复杂度为  $O(n \cdot m)$ ; 在利用  $\epsilon$  贪心策略进行动作选择时, 时间复

杂度与动作空间大小有关, 单次复杂度为  $O(n)$ ; 最后在更新  $Q$  值表时, 时间复杂度为  $O(n \cdot \log n)$ .

### 3.4 AHP 定权的多目标强化学习作业调度框架

图3是AHP定权的多目标强化学习作业调度总体框架图, 反映出多目标调度的总体运行状况及控制机制, 其主要由4个部分组成: 1) 用户层; 2) 计算资源层; 3) 监控层; 4) 决策层.

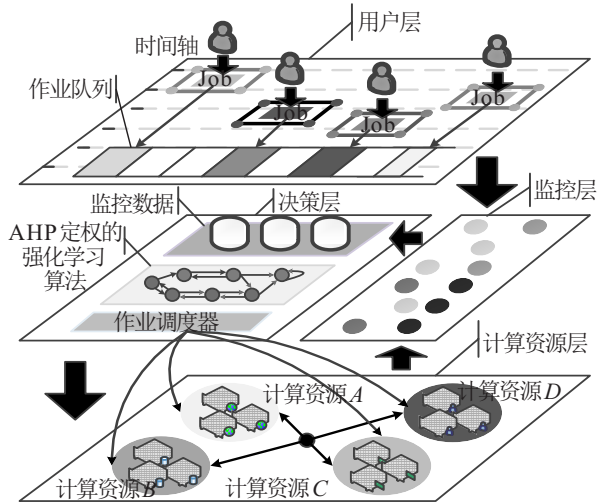


图3 AHP定权的多目标强化学习作业调度框架

1) 用户层: 用户在不同时刻提交所需处理的作业, 然后这些作业进入作业队列等待处理. 这些作业种类不同, 而且对各类资源的需求也各异.

2) 计算资源层: 将异构服务器的实体资源抽象成虚拟资源, 便于对计算资源进行高效管理, 最大程度地利用数据中心内的可用计算资源.

3) 监控层: 作业的执行使得虚拟机以及物理服务器的资源利用率、耗电功率等发生变化. 因此不仅需要采集与作业执行相关的参数, 还需要采集计算资源的变化参数. 最后, 这部分采集的监控数据将作为反馈信息传送至决策层, 以供决策之需.

4) 决策层: 计算资源层和用户层的信息被实时更新至决策层. 而决策层利用AHP定权的强化学习算法对异构环境进行实时学习, 并确定当前状态下最优作业调度方案. 最后将该方案交给作业调度器执行. 随着作业部署到虚拟机, 计算资源层和用户层的变化信息被监控层传递给决策层. 这4个部分通力合作, 周而复始, 最终实现作业和异构资源的合理分配.

## 4 实验设计与结果分析

### 4.1 实验平台介绍

MultiRECloudSim<sup>[16]</sup> 是基于 CloudSim3.0.3 拓展开发的云计算仿真平台, 在原有的 CloudSim 平台增加了作业的多种资源需求模拟、预取资源管理模块以及多资源运行的能耗计算等功能.

### 4.2 异构资源平台设置

本文在 MultiRECloudSim<sup>[16]</sup> 构建了6款异构服务器, 具体型号及功率映射函数如表3所示. 其中: 功率映射函数根据 SPEC 官网提供的数据计算得到,  $x$  为CPU利用率,  $y$  为耗电功率<sup>[17]</sup>.

表3 各类型主机的资源详情

No.	服务器型号	功率映射函数
1	Fujitsu TX150 S5	$y = 89.5 + 42.9x$
2	Acer GR160 F1	$y = 100.7 + 103.6x - 8.7x^2$
3	DEPO Race X340H	$y = 81.1 + 73.2x - 23.6x^2$
4	Acer R520	$y = 152.9 + 115.9x$
5	Sun Netra X4250	$y = 225.5 + 99.9x - 31.4x^2$
6	Hitachi HA8000	$y = 102.2 + 73.4x + 159.1x^2$

为便于计算, 在每台服务器上部署一台虚拟机 (VM), 虚拟机的计算能力在 600~1000 MIPS 之间, 依据阿里云的收费标准, 计算出1核CPU计费为 0.0675 ¥/h, 1GB内存计费为 0.0375 ¥/h, 即虚拟机  $V_i$  单位时间内的计费  $\text{cost}_v(V_i)$  为

$$\text{cost}_v(V_i) = 0.0675n + 0.0375m. \quad (23)$$

其中:  $n$  为  $V_i$  的CPU核数;  $m$  为内存容量. 本文模拟的作业不考虑硬盘需求, 因此设定虚拟机计费标准时不考虑硬盘.

### 4.3 实验作业流

本文共设置5个作业流, 作业分不同批次到达, 各批次作业的到达时间按泊松分布设定. 各作业流的到达时间和数量如图4所示, 具体参数设置如表4所示.

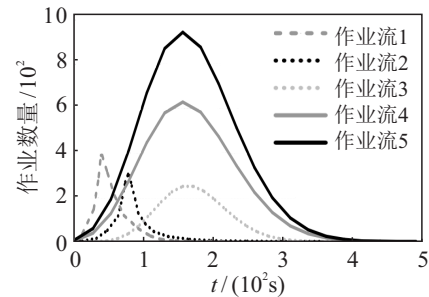


图4 各批次作业到达时刻与数量

表4 各作业流的参数设置

批次数	间隔/s	最后到达时刻/s	作业总数	活跃VM数
15	15	210	2000	40
20	18	342	2000	40
25	15	360	2000	40
20	25	475	4000	80
20	25	475	6000	120

### 4.4 仿真实验

在仿真调度实验中, 设置如下基准作业调度算法:

1) 先来先服务调度算法(FCFS): 根据作业到达数据中心的先后顺序, 依次分配至空闲虚拟机。

2) 基于蚁群算法的调度算法(ACO): 以执行时间和平台能耗作为优化目标进行作业调度。

3) 基于遗传算法的主资源负载均衡调度算法(LVS): 通过对各指标加权设计适应度函数, 利用遗传操作挑选染色体, 当满足阈值时输出调度方案。

4) 贪心调度算法(GA): 将每个作业分配至当前最优虚拟机。

FCFS 和 GA 属于传统作业调度算法, ACO 和 LVS 属于多目标作业调度算法。这些算法与本文算法(下文中用 AHP-MQ 表示)在执行时间、服务延迟成本、能耗、成本几个方面进行对比。

图5为作业总执行时间。由图5可见, GA 和 AHP-MQ 在作业总执行时间上表现较好。AHP-MQ 比 FCFS、ACO、LVS 分别平均减少 42.13%、19.18%、29.76%。

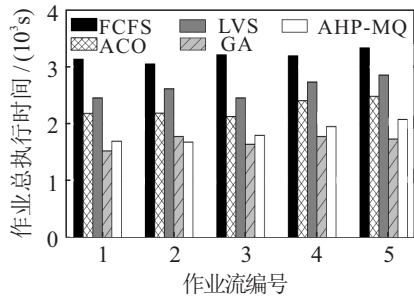


图5 作业总执行时间

图6为服务延迟成本。由图6可见, FCFS 和 LVS 的服务延迟成本较大, 主要原因是 FCFS 按到达顺序分配作业, 并未考虑优化策略, 而 LVS 在调度时关注负载均衡, 不让过多作业集中到性能高的虚拟机上, 因此有较多作业处于等待状态。AHP-MQ 算法在服务延迟成本这个调度目标上仅次于 GA。

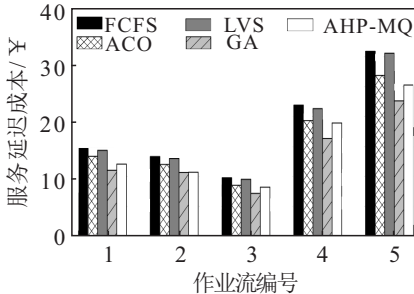


图6 服务延迟成本

图7为平台运行能耗。由图7可见, 在平台运行能耗方面, AHP-MQ 算法和 GA 表现相当, 虽然 AHP-MQ 和 GA 在高性能虚拟机上部署较多作业导致单位时间内功率偏高, 但是 AHP-MQ 和 GA 的总执行时间远低于 FCFS、ACO 和 LVS, 所以总体而言能耗更低。

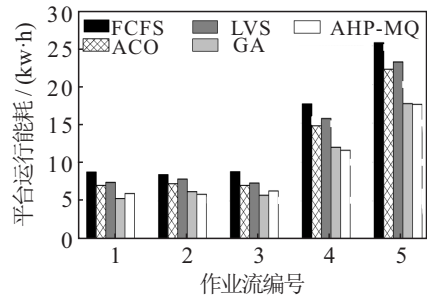


图7 平台运行能耗

图8为总成本。由图8可见, GA 和 AHP-MQ 的总成本比较低, 这与能耗较低的原因相似, 均为这些调度算法总执行时间少, 此外 AHP-MQ 在计算多目标调度综合评价时, 成本的权重占较大比例。

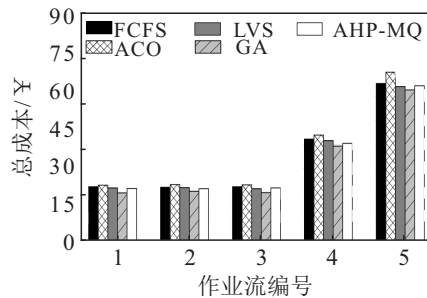


图8 总成本

在仿真平台内的实验结果表明, 所提出的 AHP 定权多目标强化学习作业调度算法在作业总执行时间、平台运行能耗方面表现较好, 但在服务延迟成本方面比 GA 表现劣势。此外, AHP-MQ 在总成本控制方面也表现得较好。所提出的 AHP-MQ 算法能够发挥强化学习的能力, 较好地优化调度方案。

### 5 结论

随着云计算的快速发展, 越来越多的作业提交至云数据中心。在海量的异构资源的云数据中心进行作业调度是一个 NP 难问题。本文针对云环境下多目标优化调度问题, 定义执行时间、服务延迟成本、平台运行能耗、成本等调度目标, 其中服务延迟成本描述用户对服务质量的满意程度; 然后设计 AHP 定权的多目标强化学习作业调度算法并将面向异构资源的多目标调度综合评价方法引入 Q-learning 奖励值的计算, 有利于后续作业的学习; 最后利用 MultiRECloudSim 仿真平台进行仿真实验, 结果表明所提出方法优于大部分对比方法, 能够较好地保障用户及服务提供商的利益。在今后的研究中, 将考虑负载均衡, 优化奖励值的计算, 使该算法更加全面合理。

### 参考文献(References)

[1] Guo J, Chang Z H, Wang S, et al. Who limits the resource efficiency of my datacenter: An analysis of alibaba

- datacenter traces[C]. IEEE/ACM 27th International Symposium on Quality of Service (IWQoS). Piscataway: IEEE, 2019: 1-10.
- [2] Gabaldon E, Vila S, Guirado F, et al. Energy efficient scheduling on heterogeneous federated clusters using a fuzzy multi-objective meta-heuristic[C]. IEEE International Conference on Fuzzy Systems (FUZZ-IEEE). Piscataway: IEEE, 2017: 1-6.
- [3] 李智勇, 陈少淼, 杨波, 等. 异构云环境多目标Memetic优化任务调度方法[J]. 计算机学报, 2016, 39(2): 377-390.  
(Li Z Y, Chen S M, Yang B, et al. Multi-objective memetic algorithm for task scheduling on heterogeneous cloud[J]. Chinese Journal of Computers, 2016, 39(2): 377-390.)
- [4] Akbari M, Rashidi H. A multi-objectives scheduling algorithm based on cuckoo optimization for task allocation problem at compile time in heterogeneous systems[J]. Expert Systems with Applications, 2016, 60: 234-248.
- [5] 胡成玉, 余果, 颜雪松, 等. 基于改进多目标优化算法的分布式数据中心负载调度[J]. 控制与决策, 2021, 36(1): 159-165.  
(Hu C Y, Yu G, Yan X S, et al. Modified multi-objective optimization algorithm for load scheduling of distributed data centers[J]. Control and Decision, 2021, 36(1): 159-165.)
- [6] Zhu Z M, Zhang G X, Li M Q, et al. Evolutionary multi-objective workflow scheduling in cloud[J]. IEEE Transactions on Parallel and Distributed Systems, 2016, 27(5): 1344-1357.
- [7] 苏命峰, 王国军, 李仁发. 基于利益相关视角的多维QoS云资源调度方法[J]. 通信学报, 2019, 40(6): 102-115.  
(Su M F, Wang G J, Li R F. Multidimensional QoS cloud computing resource scheduling method based on stakeholder perspective[J]. Journal on Communications, 2019, 40(6): 102-115.)
- [8] Farahnakian F, Pahikkala T, Liljeberg P, et al. Energy-aware VM consolidation in cloud data centers using utilization prediction model[J]. IEEE Transactions on Cloud Computing, 2019, 7(2): 524-536.
- [9] Islam S, Keung J, Lee K, et al. Empirical prediction models for adaptive resource provisioning in the cloud[J]. Future Generation Computer Systems, 2012, 28(1): 155-162.
- [10] Chen M C, Yuan J L, Liu D L, et al. An adaption scheduling based on dynamic weighted random forests for load demand forecasting[J]. Journal of Supercomputing, 2020, 76(3): 1735-1753.
- [11] Yi D L, Zhou X, Wen Y G, et al. Toward efficient compute-intensive job allocation for green data centers: A deep reinforcement learning approach[C]. IEEE 39th International Conference on Distributed Computing Systems(ICDCS). Piscataway: IEEE, 2019: 634-644.
- [12] Yang Z, Nguyen P, Jin H M, et al. MIRAS: model-based reinforcement learning for microservice resource allocation over scientific workflows[C]. IEEE 39th International Conference on Distributed Computing Systems (ICDCS). Piscataway: IEEE, 2019: 122-132.
- [13] 罗亮, 吴文峻, 张飞. 面向云计算数据中心的能耗建模方法[J]. 软件学报, 2014, 25(7): 1371-1387.  
(Luo L, Wu W J, Zhang F. Energy modeling based on cloud data center[J]. Journal of Software, 2014, 25(7): 1371-1387.)
- [14] 李翔, 姜晓红, 吴朝晖, 等. 绿色数据中心的能量管理方法研究[J]. 计算机学报, 2015, 38(10): 1976-1996.  
(Li X, Jiang X H, Wu Z H, et al. Research on thermal management methods for green data centers[J]. Chinese Journal of Computers, 2015, 38(10): 1976-1996.)
- [15] 李蕊, 李跃, 徐浩, 等. 基于层次分析法和专家经验的重要电力用户典型供电模式评估[J]. 电网技术, 2014, 38(9): 2336-2341.  
(Li R, Li Y, Xu H, et al. Assessment on typical power supply mode for important power consumers based on analytical hierarchy process and expert experience[J]. Power System Technology, 2014, 38(9): 2336-2341.)
- [16] Lin W W, Xu S Y, He L G, et al. Multi-resource scheduling and power simulation for cloud computing[J]. Information Sciences, 2017, 397/398: 168-186.
- [17] SPEC. All published SPECpower\_ssj2008 results[EB/OL].(2020-08-12)[2020-09-13]. [http://www.spec.org/power\\_ssj2008/results/power\\_ssj2008.html](http://www.spec.org/power_ssj2008/results/power_ssj2008.html).

## 作者简介

袁景凌(1975—),女,教授,博士生导师,从事云计算、绿色计算和机器学习等研究, E-mail: yuanjingling@126.com;

陈旻骋(1990—),男,博士生,从事云计算、分布式计算的研究, E-mail: wester589@163.com;

江涛(1993—),男,硕士生,从事云计算、绿色计算的研究, E-mail: jany@whut.edu.cn;

李超(1986—),男,博士,从事计算机体系结构的研究, E-mail: lichao@cs.sjtu.edu.cn.

(责任编辑:魏冰)