

面向云计算环境的动态公平性度量方法

卢笛¹, 马建峰², 王一川¹, 习宁¹, 孟宪佳¹

(1. 西安电子科技大学 计算机学院, 陕西 西安 710071; 2. 通信信息控制和安全技术重点实验室, 浙江 嘉兴 314033)

摘要: 针对云计算环境资源动态性特点, 给出了一种资源分配算法公平性动态度量模型——DFE (dynamic fairness evaluation)。模型引入了公平、效率和时间参数, 通过将资源分配结果进行量化处理, 定量地描述不同分配算法的公平性。首先考虑了云环境下的2种典型动态因素: 节点资源需求动态变化和虚拟节点动态变化, 并分别对这2种场景建立子模型, 通过这2种模型导出了DFE。分析了2种子模型以及DFE, 通过DFE对几种典型资源分配算法进行度量, 验证了在资源需求以及虚拟节点动态变化的情况下, DFE能够有效度量分配算法公平性。

关键词: 资源分配; 公平性评估; 云计算

中图分类号: TP301

文献标识码: A

文章编号: 1000-436X(2014)07-0140-11

Dynamic fairness evaluation framework for cloud computing

LU Di¹, MA Jian-feng², WANG Yi-chuan¹, XI Ning¹, MENG Xian-jia¹

(1. School of Computer Science, Xidian University, Xi'an 710071, China;

2. Science and Technology on Communication Information Security Control Laboratory, Jiaxing 314033, China)

Abstract: Dynamic fairness evaluation framework (DFE) was given according to the dynamic features of resources in cloud environment. In DFE, the factors of fairness, efficiency and time were introduced, and the allocation results were processed into numeric values, which could measure the fairness of allocation algorithms quantitatively. Two typical dynamic factors in cloud environment are considered, dynamic resource demands of virtual nodes and dynamic quantity of virtual node number. Then, two sub-models are established for both of the dynamic factors, with which DFE is finally derived. Both of the sub-models and DFE are analyzed, and evaluated with several typical resource allocation algorithms. The experimental results prove that DFE can effectively provide a measurement for allocation fairness evaluation in the case of both dynamic resource demands and node number.

Key words: resource allocation; fairness evaluation; cloud computing

1 引言

云计算是对资源的高度整合, 提高了资源的利用率, 为应用程序和服务的迁移提供了便捷的途径。而如何“有效”地进行资源的分配调度, 最大限度地满足服务的资源需求, 是云计算平台调度系统需要解决的重要问题。“有效”调度这里包含两

层含义: 1) 资源分配的公平性; 2) 资源利用率最高。前者保证了资源需求能够最大程度被满足, 后者确保资源能够最大程度被利用。因此, 在云计算资源调度系统设计过程中, 需要同时考虑资源分配的公平性和利用率, 并使两者尽可能达到最大值, 即分配问题的最优解。

为解决分配的公平性或资源利用率问题, 目

收稿日期: 2014-04-02; 修回日期: 2014-05-14

基金项目: 中央高校基本科研业务费基金资助项目 (JY0900120301); 国家自然科学基金重点资助项目 (U1135002); 长江学者和创新团队发展计划基金资助项目 (IRT1078)

Foundation Items: The Fundamental Research Funds for the Central Universities (JY0900120301); The Key Project of the National Natural Science Foundation of China (U1135002); Supported by Program for Changjiang Scholars and Innovative Research Team in University (IRT1078)

前，研究工作采用了基于多种优化方法的调度算法。在公平性方面，文献[1~3]寻求最小资源需求的最大化，即“max-min fairness”，保证多数用户的资源需求得到满足。文献[4,5]在2个冲突利益之间寻求资源分配的平衡，即“proportional fairness”。文献[6~11]提出了多资源公平性分配算法，文献[12]则提出了基于“dominant share”的多资源公平分配方法。同时考虑分配的公平性和资源利用率的实质是从需求和资源利用最大化2个角度考虑分配的问题，并通过优化方法得到问题的最优解或近似最优解。文献[13]提出了一种在公平和效率之间进行折中的模型，即“ α -fairness”。文献[14]提出一种基于“Jain's fairness index”和利用率的折中模型。文献[15~17]基于遗传算法进行资源分配的优化，但遗传算法随着问题规模的扩大和计算的深入，其复杂程度呈指数级增长，为得到问题最优解而大量消耗系统计算资源，增加了系统的负荷，不适用于云计算环境下大规模虚拟化资源调度问题。文献[18]提出并讨论了基于贪心和非贪心的资源调度算法，通过与批调度(batch scheduling)算法的比较，说明文中提出的算法在调度质量方面有所提高。文献[19]采用合作博弈论对资源效用分配进行了研究，基于费用最小化的MIN_COST算法得出资源的最优化分配方案。文献[20]则提出了一种基于预测的资源重配置算法FRRA。

尽管现有工作对资源分配的最优化问题进行了深入而广泛的研究，提出了基于多种优化理论的算法，但对于如何评估分配结果的公平性却少有涉及。而公平性评估方法对分配算法的选择、性能评价都具有重要的意义。目前，公平性评估方法主要通过构建评估函数。文献[21]提出了基于5条基本公理的资源分配策略，并构建了基于单一资源的分配公平性评估框架，为公平性的度量提供了准则。文献[22]在文献[21]的基础上结合“DRF (dominant resource fairness)”的核心思想，提出了一种基于“dominant share”的统一公平性度量框架，为多资源分配算法提供了定量的评估方法。该框架同时兼顾效率在评估中的作用，引入了公平性参数 β ，效率参数 λ ，通过调整参数而调整评估函数特性，确定评估指标的重点(公平性/效率)。本文将基于该框架，针对云计算平台动态资源需求以及虚拟机节点数量的动态特性，提出一种适合于云计算的资源分配

度量方法，为资源分配算法的确定提供定量的依据。

2 问题描述

2.1 问题提出

由前述可知，资源分配算法需要处理公平性和效率的问题，达到总体的最优解或近似最优解。而云计算平台中各应用服务由虚拟机提供，根据服务类型的不同，虚拟机对平台资源的需求也呈现出变化的特性。譬如，当服务对计算量要求高时，此时CPU和内存的需求较大；当服务程序需要进行数据传输时，带宽将成为资源消耗的重点。因此，虚拟机节点的资源请求是由其上运行的服务程序决定的。另一方面，平台虚拟机节点也呈现变化的特性。当新任务开启时，新节点被创建，而任务结束后，节点将被销毁，释放占用资源。因此，平台虚拟机节点随时间变化，进一步来看，由于虚拟机节点的这种变化，实质上导致了平台可用资源也随时间变化的特性。由此，公平性度量函数需要充分考虑云计算平台的这种特性，在需求和资源动态变化的情况下，对资源分配算法进行有效的评估。

2.2 数学模型

公平性原则基于如何满足虚拟计算节点对不同资源的需求，并得到该问题的优化解；而动态性原则是一个基于时间的概率函数，因此定义动态公平性评估函数为

$$T = \mathcal{F}_{\beta, \lambda, t}(\mathbf{x}) \quad (1)$$

其中，函数 \mathcal{F} 为动态公平性函数， β 描述公平性类型， λ 强调资源利用率， t 描述时间的变化， \mathbf{x} 给出节点计算任务集合。可以看出，如何选择恰当的公平性函数 \mathcal{F} ，对资源动态性进行描述是解决本文提出问题的关键。Ali Ghodsi等在文献[14]中提出一种基于“优势份额(dominant share)”的公平性度量方法“dominant resource fairness (DRF)”。针对多用户使用多资源的情况，给出了用户当前优势份额的计算方法，并在保证优势份额得到满足的情况下，对所有用户的资源分配是公平的。Carlee Joe-Wong等基于文献[21]中单资源评估模型，在文献[22]中给出了多种资源分配公平性统一评估框架，通过引进DRF^[14]的思想，提出了面向优势份额的公平性度量模型

$$\mathcal{F}_{\beta,\lambda} = \text{sgn}(1-\beta) \left(\sum_{j=1}^n \left(\frac{\mu_j x_j}{\sum_{k=1}^n \mu_k x_k} \right)^{1-\beta} \right)^{\frac{1}{\beta}} \left(\sum_{j=1}^n \mu_j x_j \right)^{\lambda} \quad (2)$$

式(2)中引入了参数 $\beta \in \mathbb{R}, \lambda \in \mathbb{R}$ 。 β 对公平性进行度量,不同 β 值反映出不同的公平性类型,其大小决定了公平的强度。而 λ 强调资源利用率,该参数表明资源是否被最大化利用。 λ 越大则说明模型更加注重资源的利用率。 $\lambda \rightarrow 0$ 则说明可忽略资源利用率的因素。 μ_j 为用户 j 的优势份额 (dominant share), x_j 为用户 j 的任务数。其中,

$$\mu_j = \max_i \left\{ \frac{d_{ij}}{R_i} \right\} \quad (3)$$

d_{ij} 为用户 j 对资源 i 的请求量, R_i 为资源 i 的总量。通过对参数 β 、 λ 值的调整,可以得到不同性质的公平性度量函数。例如,当 $\beta \in (0,1)$, $\mathcal{F}_{\beta,\lambda}$ 为基于 α -fairness 的度量函数;当 $\beta \rightarrow \infty$, 且 $\lambda = \frac{1-\beta}{\beta}$ 时, $\mathcal{F}_{\beta,\lambda}$ 转化为基于优势份额的 max-min fairness 度量函数。若 $\mu_j x_j$ 表示用户 j 的优势份额,且满足约束 $\sum_{j=1}^n d_{ij} x_j \leq R_i$, 则使 $\min\{\mu_1 x_1, \mu_2 x_2, \dots, \mu_n x_n\}$ 最大化即得到 DRF。因此,式(2)可以作为一种资源分配公平性度量的通用框架。不失一般性,假设用户 j 仅执行 1 个计算任务,则式(2)可以简化为

$$\mathcal{F}_{\beta,\lambda} = \text{sgn}(1-\beta) \left(\sum_{j=1}^n \left(\frac{\mu_j}{\sum_{k=1}^n \mu_k} \right)^{1-\beta} \right)^{\frac{1}{\beta}} \left(\sum_{j=1}^n \mu_j \right)^{\lambda} \quad (4)$$

由于云计算平台下,计算节点的数量、各计算节点对资源的需求呈现动态变化的特点,式(4)虽能够较好地刻画多资源分配公平性问题,但并未充分考虑资源需求以及节点数量变化的动态性。

3 资源分配公平性动态评估方法

为了适应云计算环境下虚拟化资源需求以及计算节点数量动态变化的特性,资源分配评估模型还必须考虑资源需求以及虚拟节点随时间变化的因素,即评估函数应具有式(1)的形式。本文将基于

式(1)提出动态公平性评估框架 (DFE, dynamic fairness evaluation)。

3.1 节点资源需求动态模型

假设平台 P 中存在 m 种资源,每种资源总量为 R_i , 节点 j 对资源 i 的请求为 d_{ij} , 则得到节点 j 的优势份额如式(3)所示。可以看出,对于节点 j , 假设总量 R_i 恒定的情况下,其优势份额是由当前最大 d_{ij} 决定的。然而,节点的计算任务随时间变化,其优势份额对应资源类型也发生变化,因此节点的优势份额应是时间的函数,定义为

$$\mu_j(t) = \max_i \left\{ \frac{d_{ij}(t)}{R_i} \right\} \quad (5)$$

假设节点 j 的资源请求可以表示为: $D_j = (d_{1j}, d_{2j}, \dots, d_{mj})$, 系统对其分配的份额可能并不能完全满足其资源请求,即实际分配额 $a_{ij} \leq d_{ij}$ 。因此,定义系数 v_j , $0 < v_j \leq 1$, 则 $a_j = v_j D_j$ 。由于资源需求和分配结果同为时间的函数,进一步可得

$$a_j(t) = D_j(t) v_j(t) \quad (6)$$

根据式(6),对于任意资源 i , 节点 j 的配额 $a_{ij}(t)$ 可表示为 $a_{ij}(t) = d_{ij}(t) v_j(t)$ 。若资源 i 总量为 R_i , 则任意时刻 t_p , 系统对 i 的分配需满足

$$\sum_j^n a_{ij}(t_p) = \sum_j^n d_{ij}(t_p) v_j(t_p) \leq R_i \quad (7)$$

因此,式(6)给出动态需求下节点资源配额模型,并且服从式(7)的约束。

3.2 虚拟节点数量动态模型

随着时间推移,完成计算任务的节点会释放所占用的资源,而执行新任务的计算节点会产生并占用相应的资源。考虑到不同时刻计算节点的产生和销毁,系统可用资源总量也会受其影响而变化。考虑任意时刻 t_p , 任意资源 i , 销毁的计算节点集合为 $\Phi(t_p) = \{\text{node}_j \mid \hat{v}_j(t_p) = 0\}$, 新产生的计算节点集合为

$$\Omega(t_p) = \{\text{node}'_k \mid v'_k(t_p) \neq 0 \wedge v'_k(t_p) \notin v_j(t_p)\}$$

考虑计算节点数量随时间变化的特性,记 t_p 时刻节点总量为 $n(t_p)$ 。则 $\Omega(t_p)$ 中的节点满足约束

$$\sum_k^n d'_{ik}(t_p) v'_k(t_p) + \sum_j^n d_{ij}(t_p) v_j(t_p) \leq R_i \quad (8)$$

同时考虑系统中计算节点的产生和销毁，则任意资源 i 在 t_p 时刻的可用量可以表示为： $U_i(t_p) = U_i(t'_p) - C_i^\Omega(t_p) + C_i^\Phi(t_p)$ 。其中， t'_p 表示上一时刻， $C_i^\Omega(t_p)$ 表示为新节点分配的资源 i 的总量， $C_i^\Phi(t_p)$ 表示销毁节点释放资源 i 的总量。同理，定义 t_p 时刻系统被占用的资源量 $O_i(t_p)$ ，有 $O_i(t_p) = O_i(t'_p) + C_i^\Omega(t_p) - C_i^\Phi(t_p)$ ，可得 $R_i = U_i(t_p) + O_i(t_p)$ 。由于 R_i 为常数，当系统空闲资源 $U_i(t)$ 增加时，占用资源 $O_i(t)$ 必然减少，反之，当 $O_i(t)$ 增加时，导致减少。特别地，当 $U_i(t) = 0$ 时，则资源 i 在 t_p 时刻达到饱和。

假设同一时刻系统产生新计算节点的概率为 p_ω ，而任务结束销毁计算节点的概率为 p_ϕ ，定义函数 $g_1(t)$ ， $g_2(t)$ 满足

$$g_1(t) = \begin{cases} 1, & \text{in } p_\omega \\ 0, & \text{in } 1 - p_\omega \end{cases}, \quad g_2(t) = \begin{cases} 1, & \text{in } p_\phi \\ 0, & \text{in } 1 - p_\phi \end{cases} \quad (9)$$

资源 i 在 t_p 时刻的占用量可以重写为

$$O_i(t_p) = O_i(t'_p) + g_1(t_p) \cdot \sum_k^{\|\Omega(t_p)\|} d'_{ik}(t_p) v'_k(t_p) - g_2(t_p) \cdot \sum_k^{\|\Phi(t_p)\|} \hat{d}_{ik}(t_p) \hat{v}_k(t_p) \quad (10)$$

$O_i(t'_p)$ 为上一时刻可用资源，由式(6)可得

$$O_i(t'_p) = \sum_j^{n(t'_p)} d_{ij}(t'_p) v_j(t'_p), \text{ 因此, 式(10)可重写为}$$

$$O_i(t_p) = \sum_j^{n(t'_p)} d_{ij}(t'_p) v_j(t'_p) + g_1(t_p) \cdot \sum_k^{\|\Omega(t_p)\|} v'_k(t_p) d'_{ik}(t_p) - g_2(t_p) \sum_k^{\|\Phi(t_p)\|} \hat{v}_k(t_p) \hat{d}_{ik}(t_p) \quad (11)$$

式(11)给出资源 i 占用量的“时间—概率”动态模型，能够较为真实地反映在计算节点动态变化的情况下，系统被占用资源的变化特性。同理，由于系统中节点的优势份额为其资源需求向量中最大的元素，因此，节点优势份额总量同 $O_i(t_p)$ 具有相似的形式，可以表示为

$$M(t_p) = \sum_j^{n(t'_p)} \mu_j(t'_p) v_j(t'_p) + g_1(t_p) \sum_k^{\|\Omega(t_p)\|} \mu'_k(t_p) v'_k(t_p) - g_2(t_p) \sum_k^{\|\Phi(t_p)\|} \hat{\mu}'_k(t_p) \hat{v}'_k(t_p) \quad (12)$$

由 $\mu_j(t_p)$ 的定义可得 $M(t_p) < \sum_i^m O_i(t_p)$ ，结合式

(4)的公平模型以及式(12)的动态模型，可以导出动态公平性度量模型 (DFE)。

3.3 DFE 模型

由 $\mu_j(t)$ 的定义可知，节点 j 优势份额(dominant share)的实际量为 $\mu_j(t)v_j(t)$ 。则式(4)的公平模型可以表示为

$$\mathcal{F}_{\beta,\lambda}(t) = \text{sgn}(1 - \beta) \left(\sum_{j=1}^{n(t)} \left(\frac{\mu_j(t)v_j(t)}{\sum_{k=1}^{n(t)} \mu_k(t)v_k(t)} \right)^{1-\beta} \right)^{\frac{1}{\beta}} \left(\sum_{j=1}^{n(t)} \mu_j(t)v_j(t) \right)^\lambda$$

进一步考虑，上式中 $\sum_{k=1}^{n(t)} \mu_k(t)v_k(t)$ 为系统中所有节点优势份额之和，考虑平台中计算节点数量动态变化特点，由式(12)可得

$$\mathcal{F}_{\beta,\lambda}(t) = \text{sgn}(1 - \beta) \left(\sum_{j=1}^{\eta(t)} \left(\frac{\mu_j(t)v_j(t)}{M(t)} \right)^{1-\beta} \right)^{\frac{1}{\beta}} (M(t))^\lambda \quad (13)$$

式(13)为动态公平性评估框架(DFE)。其中， $\eta(t)$ 表示 t 时刻系统中的节点数量，有 $\eta(t) = n(t) + \|\Omega(t)\| - \|\Phi(t)\|$ 。由式(13)可以看出，随着时间的推移，系统中节点数量及其优势份额也随之发生变化。式(13)体现出这种变化对公平性的影响，即动态特性对 $\mathcal{F}_{\beta,\lambda}(t)$ 函数的影响。

性质 1 DFE 不会改变原公平性度量函数的性质。

证明 对任意时刻 t_p ，在该时刻可认为系统计算节点数量瞬时恒定，假设为 n_p ，则节点优势资源总

和恒定，即 $M(t_p) = \sum_j^{n_p} \mu_j(t_p)v_j(t_p)$ 。又 $\mu_j(t_p), v_j(t_p)$

在该时刻已确定，有 $M(t_p) = \sum_j^{n_p} \mu_j v_j$ ，因此有

$$\mathcal{F}_{\beta,\lambda}(t_p) = \text{sgn}(1 - \beta) \left(\sum_{j=1}^{n_p} \left(\frac{\mu_j v_j}{\sum_j \mu_j v_j} \right)^{1-\beta} \right)^{\frac{1}{\beta}} \left(\sum_{j=1}^{n_p} \mu_j v_j \right)^\lambda$$

即式(13)退化为式(4)的形式（考虑优势份额的实际分配量，因此乘以系数 v_j ）。综上，对于式(13)所述模型，在确定时刻可转化为原公平性模型。另一方面，式(4)也可认为是在确定时刻下的 DFE 的特殊情况。

4 分析评估

本节对资源请求动态模型和虚拟节点变化模型进行分析，并对 DFE 模型进行验证。

4.1 节点资源请求动态模型分析

由式(6)可得，对任意资源 i ，其分配总量为 $s_i(t) = \sum_j^{n(t)} d_{ij}(t)v_j(t)$, ($s_i \leq R_i$)。进一步表示为

$$\mathcal{L}_i : \sum_j^{n(t)} d_{ij}(t)v_j(t) - s_i(t) = 0 \quad (14)$$

式(14)实质上确定一个超平面 \mathcal{L}_i ，法向量 $d_i(t) = (d_{i1}(t), d_{i2}(t), \dots, d_{im(t)}(t))$ 垂直于资源分配向量 $v(t)$ ，即垂直于 $v(t)$ 所在超平面 \mathcal{L}_i 。 $d_i(t)$ 与 \mathcal{L}_i 的交点即为对应时刻 t 的资源分配结果，记为 v^* 。而公平分配算法的实质是为满足某种公平性需求（例如，max-min fairness, proportional fairness 等），构建一个超平面 \mathcal{L}_i ，并求出存在于 \mathcal{L}_i 上的分配结果 v^* 。针对不同时刻 $t, t \in [0, \infty)$ ，存在对任意资源 i 的请求向量族 $D_i = (d_i(t_1), d_i(t_2), \dots, d_i(t_n))$ ，其中每一个元素为一个针对资源 i 的需求向量，且服从约束： $\sum d_i(t_p) \leq R_i$ 。而资源分配的本质为在某一时刻，通过资源分配算法提供的公平性准则，找到该时刻法向量 $d_i(t_p)$ 所垂直的超平面，最终得到与该平面的交点，即分配问题的解 $v^*(t_p)$ 。以 2 个节点、3 种资源的情形为例，假设其需求向量分别为：

$d_1 = (\frac{1}{4}, \frac{2}{3}, 1)$ 和 $d_2 = (\frac{1}{2}, \frac{1}{3}, 0)$ ，分配结果为 $v = (v_1, v_2)$ ，则下式成立

$$\begin{cases} \frac{1}{4}v_1 + \frac{1}{2}v_2 \leq 1: L_1 \\ \frac{2}{3}v_1 + \frac{1}{3}v_2 \leq 1: L_2 \\ v_1 \leq 1: L_3 \end{cases} \quad (15)$$

式(15)定义了 3 个直线方程： L_1 、 L_2 和 L_3 。3 条直线围成一个封闭区域，如图 1 所示。

图 1 中灰色区域为 3 条直线围城的区域，即在资源需求约束下的解区域 $Q(t)$ 。通过采用公平性算法求得在 $Q(t)$ 区域里的某一点 (v_1^*, v_2^*) ，该点为分配问题的解。采用不同的公平性算法，得到的解也不相同，但解集一定在 $Q(t)$ 区域内。由于不同时刻需求向量不同，因而 L_1 、 L_2 和 L_3 所表示

直线及其包络区域 $Q(t)$ 也不同，而根据算法得到的解仍位于对应时刻的解区域 $Q(t)$ 内。若分配算法确定， t 时刻需求向量确定，则能够得到确定的区域 $Q(t)$ 和解。

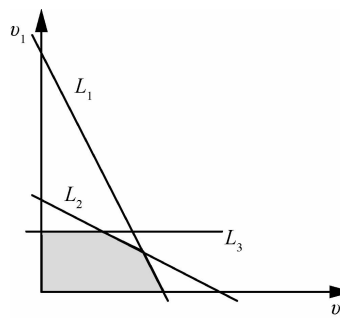


图 1 d_1 、 d_2 下的解空间

考虑 n 个用户 m 个资源的情况，上例中的直线演变为 n 维超平面，对于某种资源的请求向量则为该超平面的法向量，而解空间为 m 个 n 维超平面围成的空间 $S(t)$ ，由分配算法得到的解向量位于 $S(t)$ 内。由于资源需求随时间变化的特性，包含解向量的空间大小也是时间的函数。又因为不同任务阶段虚拟节点的数量也在发生变化，因此，虚拟节点数量也是时间的函数。虚拟节点数量的变化导致了上述超平面维度的变化，因而得到不同的解空间 $S(t)$ 。下面将对虚拟节点数量的动态特性进行分析。

4.2 虚拟节点数量变化模型分析

由式(11)可知，资源 i 在任意时刻的使用量 $O_i(t)$ 由 3 部分构成：上一时刻的使用量 $O_i(t')$ 、当前时刻新增使用量 $C_i^\Omega(t)$ 和当前时刻销毁资源量 $C_i^\Phi(t)$ 。由于当前时刻所销毁资源是上一时刻占用的资源，可得 $O_i(t') \geq C_i^\Phi(t)$ 。若等号成立，则系统在 t 时刻销毁了前一时刻的所有占用资源，有 $O_i(t) \geq 0$ 。同理，任意时刻资源 i 空闲资源量 $U_i(t)$ 由上一时刻空闲资源量 $U_i(t')$ 、当前新增资源量 $C_i^\Omega(t)$ 以及当前销毁资源量 $C_i^\Phi(t)$ 构成。若 $U_i(t) = 0$ ，则系统在 t 时刻已无可用资源，且此时 $O_i(t)$ 达到最大值。反之，若 $O_i(t) = 0$ ，则 $U_i(t)$ 达到最大值。

为进一步讨论 $O_i(t)$ 与 $U_i(t)$ 的关系，对资源 i 总量 R_i 做归一化处理，即 $R_i = 1$ ，由 3.2 节可得 $O_i(t) + U_i(t) = 1$ 。例如，在 t' 时刻 $O_i(t') = 0.3$ ，则 $U_i(t') = 0.7$ ，如图 2 所示点 $p_{t'}$ 。从图中可以看出，

$O_i(t), U_i(t)$ 的取值由直线 L 确定, 即直线上任意一点为某时刻可能的 $O_i(t), U_i(t)$ 的值。定义函数 f 为经过原点的直线与 L 相交, 交点集合即为 $O_i(t), U_i(t)$ 的取值集合。定义 $k = f' = \frac{O_i(t)}{U_i(t)}$ 可以得到以下关系。

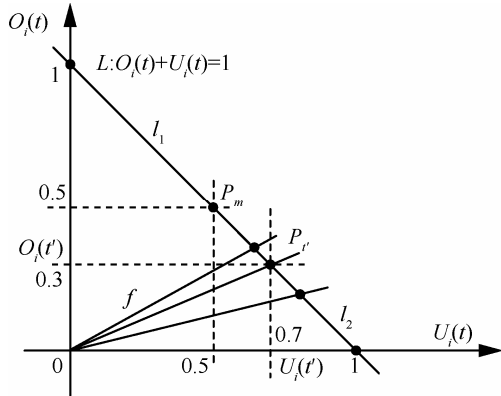


图 2 系统占用资源 $O_i(t)$ 与可用资源 $U_i(t)$ 关系

1) $k \rightarrow 0$: 有 $O_i(t) \rightarrow 0$, 系统可用资源增加至最大值 $U_i^{\max} = R_i$ 。

2) $k \rightarrow \infty$: 有 $U_i(t) \rightarrow 0$, 系统占用资源增加至最大值 $O_i^{\max} = R_i$ 。

3) $k = 1$: 有 $U_i(t) = O_i(t)$, 系统占用资源与可用资源相等, 即图 2 中 p_m 点。

$k = 1$ 为系统的特殊情况, 此时 L 与 f 相交于 p_m 点, 当 $O_i(t), U_i(t)$ 取值为该点的下半段时 ($0 \leq k < 1$), 资源 i 的可用量大于占用量, 此时资源 i 能够满足未来的需求; 当 $O_i(t), U_i(t)$ 取值为 p_m 点的上半段时 ($1 \leq k < \infty$), 资源 i 的占用资源将大于可用资源, 随着 k 值的继续增加, 资源 i 可能成为系统的瓶颈资源, 此时需要考虑调整分配算法, 保证资源不会过早饱和。

由前述可知, $O_i(t), U_i(t)$ 的值实质是由 $C_i^{\Omega}(t)$ 和 $C_i^{\Phi}(t)$ 确定的, 因此 k 值也由这 2 个值间接确定, 因此能够得到如下关系。

1) k 增加: 由请求资源较多而释放资源较少引起, 即在大多数时刻有 $C_i^{\Omega}(t) > C_i^{\Phi}(t)$ 。

2) k 降低: 由请求资源较少而释放资源较多引起, 即在大多数时刻有 $C_i^{\Omega}(t) < C_i^{\Phi}(t)$ 。

综上, 为控制系统资源不过快成为瓶颈, 调度器应根据 k 值控制 $C_i^{\Omega}(t)$ 和 $C_i^{\Phi}(t)$, 保证系统资源的可用性。

4.3 DFE 模型分析实验

实验将对 DFE 模型 (式(13)) 在不同分配算法情况下的有效性进行验证。实验分为 2 个部分, 首先采用 DFE 对 DRF、全平均分配 (Avg) 和 Max-min 分配算法的动态公平性进行度量并对比分析, 其次, 以 α -fairness 公平调整算法为例对基于效用的公平性算法进行评估。

4.3.1 DRF、Avg 与 Max-min 公平性评估实验

在 DRF、Avg 和 Max-min 比较实验中, 主要参数分别取 $\beta \in (0, 20)$, $\lambda = \frac{1-\beta}{\beta}$, 新节点产生概率 $p_{\omega} = 0.6$, 销毁计算节点概率 $p_{\phi} = 0.7$, 资源总量 $R = \{300, 400, 250\}$, 初始节点数量 $n = 20$, 节点资源需求 $0 \leq d_{ij} \leq 30$, 时间 $t \in [0, 50]$, 不同 t 时刻的需求向量 $D_i(t), (i = 1, 2, \dots, n)$ 构成类似图 1 的 n 维解空间。

图 3 显示基于时间 t 的 DFE 模型, 图中“时间 (实验点)”轴表示时间变化, S1 为采用 DRF 算法时公平性度量值的变化情况, S2 为采用 Avg 算法时公平值的变化情况, S3 为采用 Max-min 算法时公平值的变化情况。直观上, S1 将 S2 和 S3 完全覆盖, 说明采用 DRF 算法的公平性大于 Avg 和 Max-min。下面通过结合这 3 种分配算法的特点, 进一步讨论 3 种算法公平性差异。

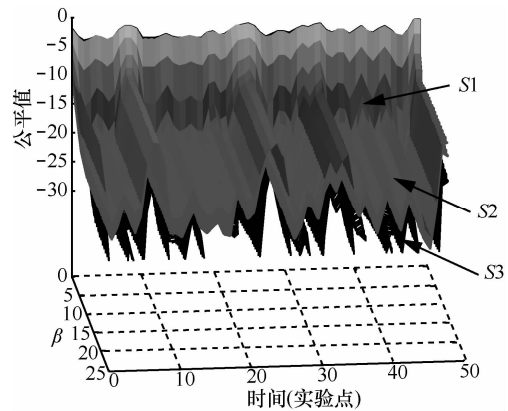


图 3 基于 DFE 的公平性度量 (以 DRF、Avg 和 Max-min 分配算法为例)

结论 1 在 DFE 评估模型下, DRF 能够取得较 Avg 和 Max-min 高的公平性。

证明 在采用 DRF 算法的情况下, 节点 i 的 dominant share 可以表示为: $\mu_i = \max_k \left\{ \frac{d_{ik}}{R_k} \right\}$, 节点资源分配量由 μ_i 决定。而 Avg 算法是对资源的平均分配, 各节点分得资源份额相等。因此 dominant

资源不一定能够被满足（只有当各节点资源需求近似平均时 Avg 算法将接近 DRF）。定义采用 Avg 算法的优势份额为 $c_i = \frac{1}{n}$ （ n 为节点数量），当 $k > 1$ 时（多于一种资源），有 $\mu_i > c_i$ 。考虑任一时刻 t_p ，由性质 1 可知，式(13)可以表示为

$$\mathcal{F}_{\beta,\lambda}(t_p) = \text{sgn}(1-\beta) \left(\sum_{j=1}^{n_p} \left(\frac{\mu_j v_j}{\sum_j \mu_j v_j} \right)^{1-\beta} \right)^{\frac{1}{\beta}} \left(\sum_{j=1}^{n_p} \mu_j v_j \right)^\lambda$$

当 $\beta > 0$ ， $\lambda = \frac{1-\beta}{\beta}$ 时，由文献[22]，上式可以化简为

$$\mathcal{F}_{\beta,\lambda}(t_p) = \text{sgn}(1-\beta) \left(\sum_{j=1}^{n_p} (\mu_j v_j)^{1-\beta} \right)^{\frac{1}{\beta}}$$

将 μ_i 、 c_i 分别代入上式可得

$$\mathcal{F}_{\beta,\lambda}^{\text{Avg}}(t_p) = \text{sgn}(1-\beta) \left(\sum_{j=1}^{n_p} (c_j v_j)^{1-\beta} \right)^{\frac{1}{\beta}}$$

$$\mathcal{F}_{\beta,\lambda}^{\text{DRF}}(t_p) = \text{sgn}(1-\beta) \left(\sum_{j=1}^{n_p} (\mu_j v_j)^{1-\beta} \right)^{\frac{1}{\beta}}$$

定义 $\gamma = \frac{\mathcal{F}_{\beta,\lambda}^{\text{Avg}}(t_p)}{\mathcal{F}_{\beta,\lambda}^{\text{DRF}}(t_p)}$ ，有

$$\gamma = \frac{\text{sgn}(1-\beta) \left(\sum_{j=1}^{n_p} (c_j v_j)^{1-\beta} \right)^{\frac{1}{\beta}}}{\text{sgn}(1-\beta) \left(\sum_{j=1}^{n_p} (\mu_j v_j)^{1-\beta} \right)^{\frac{1}{\beta}}} = \sqrt[\beta]{\frac{\sum_{j=1}^{n_p} (c_j v_j)^{1-\beta}}{\sum_{j=1}^{n_p} (\mu_j v_j)^{1-\beta}}}$$

$\because c_j v_j \leq \mu_j v_j$ ，当 $\beta > 1$ 时，有

$$(c_j v_j)^{1-\beta} \geq (\mu_j v_j)^{1-\beta}$$

则 $\gamma \geq 1$ ，可得： $|\mathcal{F}_{\beta,\lambda}^{\text{Avg}}| \geq |\mathcal{F}_{\beta,\lambda}^{\text{DRF}}|$ 。又 $\text{sgn}(1-\beta) < 0$ ，可得： $\mathcal{F}_{\beta,\lambda}^{\text{Avg}} < \mathcal{F}_{\beta,\lambda}^{\text{DRF}}$ 。采用 Max-min 算法时，分配总是从节点的最小资源需求开始，节点 Dominant 资源需求不一定能够被满足。定义采用 Max-min 算法 Dominant 资源份额为 m_i ，有 $\mu_i \geq m_i$ ，同理可得 $\mathcal{F}_{\beta,\lambda}^{\text{MM}} < \mathcal{F}_{\beta,\lambda}^{\text{DRF}}$ （ $\mathcal{F}_{\beta,\lambda}^{\text{MM}}$ 为代入 m_i 后的评估函数）。因

此，采用 DRF 在 $\beta > 1$ 时能够取得较 Avg、Max-min 更高的公平性。证毕。

进一步讨论 Max-min 和 Avg 算法。首先给出 DFE 模型的另一性质。

性质 2 在 t 确定的情况下，当 $\beta > 1$ ， $\lambda = \frac{1-\beta}{\beta}$ ，DFE 是单调递减函数。

证明 当 t 确定时，由性质 1，DFE 退化为式(4)。当 $\beta > 0$ ， $\lambda = \frac{1-\beta}{\beta}$ 时，由文献[22]，式(4)可变为

$$\mathcal{F}_{\beta,\lambda}(t_p) = \text{sgn}(1-\beta) \left(\sum_{j=1}^{n_p} (\mu_j v_j)^{1-\beta} \right)^{\frac{1}{\beta}}$$

$\because 0 < v_j < 1, 0 < \mu_j < 1$ ，定义 $\phi_j = \mu_j v_j$ ，有

$0 < \phi_j < 1$ 。定义 $g(\beta) = \left(\sum_{j=1}^n \phi_j^{1-\beta} \right)^{\frac{1}{\beta}}$ ，令

$h(\beta) = \ln \left(\sum_{j=1}^n \phi_j^{1-\beta} \right)^{\frac{1}{\beta}}$ ， $h(\beta)$ 未改变 $g(\beta)$ 的单调性。

由文献[22]可知， $g(\beta)$ 连续，则 $h(\beta)$ 连续。对 $h(\beta)$ 求导

$$\frac{dh}{d\beta} = \frac{-\beta \sum_{j=1}^n (\phi_j^{1-\beta} \cdot \ln \phi_j) - \sum_{j=1}^n \phi_j^{1-\beta} \cdot \ln \left(\sum_{j=1}^n \phi_j^{1-\beta} \right)}{\beta^2 \sum_{j=1}^n \phi_j^{1-\beta}}$$

定义

$$\gamma = -\beta \sum_{j=1}^n (\phi_j^{1-\beta} \cdot \ln \phi_j) = \sum_{j=1}^n (\ln \phi_j^{-\beta} \phi_j^{1-\beta})$$

$$\psi = \sum_{j=1}^n \phi_j^{1-\beta} \cdot \ln \left(\sum_{j=1}^n \phi_j^{1-\beta} \right) = \ln \left(\sum_{j=1}^n \phi_j^{1-\beta} \right)^{\sum_{j=1}^n \phi_j^{1-\beta}}$$

定义 $\phi_m = \min_j \{ \phi_j \}$ ，将 ϕ_m 分别代入 γ, ψ 有

$$\gamma' = \ln \left(\frac{1}{\phi_m^{n\beta\phi_m^{1-\beta}}} \right), \quad \psi' = \ln \left(n\phi_m^{1-\beta} \right)^{n\phi_m^{1-\beta}}$$

有 $\psi' - \gamma' = \ln \left[(n\phi_m)^{n\phi_m^{1-\beta}} \right]$ 。 $\because \phi_m < \frac{1}{n}$ ， $\therefore n\phi_m < 1$ 。 $\because n\phi_m^{1-\beta} > n$ ， $\therefore (n\phi_m)^{n\phi_m^{1-\beta}} < 1$ ， $\therefore \ln \left[(n\phi_m)^{n\phi_m^{1-\beta}} \right] < 0$ 。 $\therefore \gamma' > \psi'$ ， $\therefore \gamma > \psi$ 。又 \because

$\beta^2 \sum_{j=1}^n \phi_j^{1-\beta} > 0, \therefore \frac{dh}{d\beta} > 0, \therefore h(\beta)$ 单调递增, $\therefore g(\beta)$ 单调递增。当 $\beta > 1$ 时, $\mathcal{F}_{\beta,\lambda}(t_p)$ 单调递减。证毕。

采用 Avg 算法时, 由优势份额定义可知, 任一节点的优势份额为 $\mu_j = \max_i \left\{ \frac{d_{ij}}{R_i} \right\} = \frac{1}{n}$ 。代入 $\mathcal{F}_{\beta,\lambda}(t_p)$ 可得 $\mathcal{F}_{\beta,\lambda}(t_p) = -n$ 。因此, 在采用 Avg 算法时, 公平值为常数。另由结论 1 可知, DRF 始终能够取得较 Avg 更高的公平性, 而 Avg 的公平值为常数, 随着 β 的增大, DRF 的公平值虽然也会递减 (性质 2), 但不会低于 Avg。

采用 Max-min 算法时, 节点优势份额 (m_j) 存在 2 种情况: 1) $m_j < \frac{1}{n}$; 2) $m_j > \frac{1}{n}$ ($m_j = \frac{1}{n}$ 时与 Avg 相同)。若 $m_j < \frac{1}{n}$, 参照结论 1 的证明过程, 此时有 $\mathcal{F}_{\beta,\lambda}^{MM} < \mathcal{F}_{\beta,\lambda}^{Avg}$ 。若 $m_j > \frac{1}{n}$, 有 $\mathcal{F}_{\beta,\lambda}^{MM} > \mathcal{F}_{\beta,\lambda}^{Avg}$ 。由性质 2 可知, 评估函数 $\mathcal{F}_{\beta,\lambda}$ 在 t 确定, 且 $\beta > 1$, $\lambda = \frac{1-\beta}{\beta}$ 时, 为单调递减函数。因此, 当 β 递增时, $\mathcal{F}_{\beta,\lambda}^{MM}$ 值不断下降, 又 $\mathcal{F}_{\beta,\lambda}^{Avg}$ 值为常量, 总有 $\mathcal{F}_{\beta,\lambda}^{MM} < \mathcal{F}_{\beta,\lambda}^{Avg}$ 。

图 4 显示沿“时间 (实验点)”轴方向对图 3 进行采样切片, 横坐标为实验次数, 即时间变化纵坐标为公平性值 (FV), 采样点选取较为均匀的分段 $\beta = 5, 10, 15, 20$ 。可以看出, 通过 DFE 模型

评估, 采用 DRF 算法的曲线值始终大于或近似等于采用 Avg 和 Max-min 算法的曲线。若节点资源需求较为平均, 即 $d_{ij}/R_j, (j=1,2,\dots,n)$ 的值较为接近, 且不存在资源 k , 使得 $d_{ik}/R_k \gg d_{ij}/R_j, k \neq j$ 。此时, 优势份额将不会凸显。同时, 当 β 值较小时, 评估函数递减不明显, 因此会出现曲线重叠或差别较小的现象, 如图 4(a)所示。随着 β 值的增大, $\mathcal{F}_{\beta,\lambda}^{MM}$ 的值逐渐递减, 与 DRF、Avg 算法的公平值逐渐拉开距离, 如图 4(b)~图 4(d)所示。

图 5 给出沿 β 轴方向对图 3 进行切片, 横坐标为公平性系数 $\beta \in (0, 20)$, 采样点选取较为均匀的分段 $t = 6, 18, 24, 48$ 。每一个采样点内节点数量呈现概率变化的特性 (p_ω, p_ϕ), 因此, 节点的 dominant share 总量同样呈现动态的特征。随着公平性系数 β 的变化, DRF 始终能够获得较 Avg、Max-min 算法更高的公平性。当 β 值逐渐增大时, DRF、Avg 算法的 FV 曲线在同等条件下变化较缓, 而 Max-min 的 FV 曲线相对下降较快, 反映出 DRF、Avg 分配算法较 Max-min 有更高的稳定性。

可以看出, 图 5(a)~图 5(d)中, 在 β 值较小时, Max-min 的公平值略大于 Avg 算法的公平值。这是因为在这 3 个测试点采用 Max-min 算法时的优势份额 $m_j > \frac{1}{n}$, 根据前面的讨论, 此时有 $\mathcal{F}_{\beta,\lambda}^{MM} > \mathcal{F}_{\beta,\lambda}^{Avg}$ 。随着 β 值的增加, $\mathcal{F}_{\beta,\lambda}^{MM}$ 逐渐减小, 因此当 β 较大时, 总有 $\mathcal{F}_{\beta,\lambda}^{MM} < \mathcal{F}_{\beta,\lambda}^{Avg}$ 。图 5(c)中, 起始时有 $\mathcal{F}_{\beta,\lambda}^{MM} < \mathcal{F}_{\beta,\lambda}^{Avg}$, 这是因为该测试点采用 Max-min 算法时的优势份额

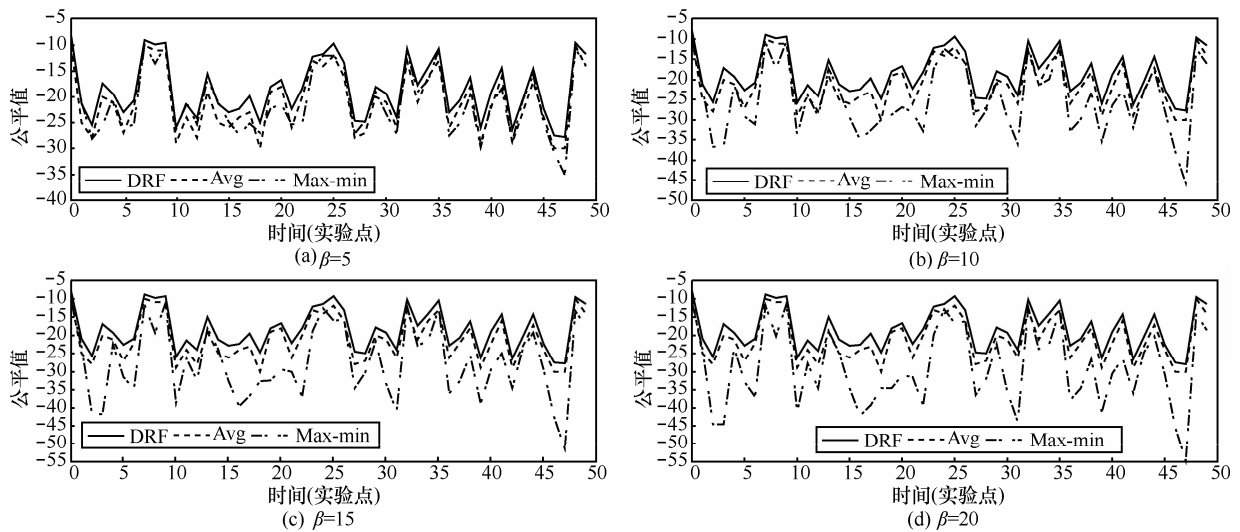


图 4 $\beta = 5, 10, 15, 20$, DFE 基于 β 的公平值变化

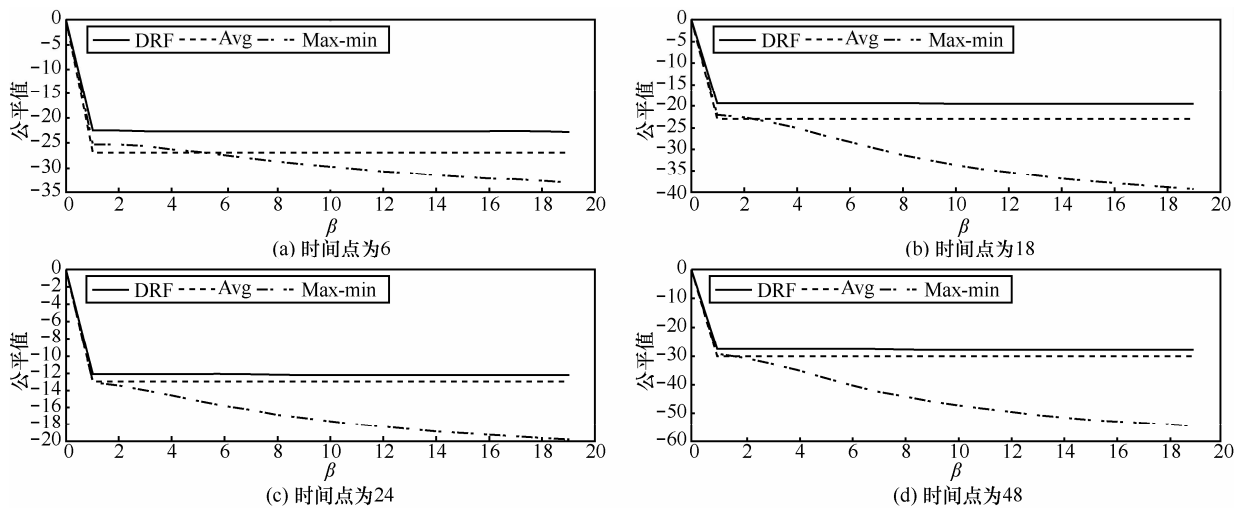


图 5 $t=6、18、24、48$, DFE 基于 t 的公平值变化

$m_j < \frac{1}{n}$, 因此始终有 $\mathcal{F}_{\beta,\lambda}^{MM} < \mathcal{F}_{\beta,\lambda}^{Avg}$ 。

从实验可以看出,通过 DFE 模型的评估,DRF、Avg 算法能够获得较高的公平性值,且公平值稳定。而采用 Max-min 算法的公平值随 β 参数的增加而递减。然而,在实际应用中,Avg 算法是对资源的完全分配,并不考虑节点的实际资源需求,虽然该算法保证了“公平”,但全平均的分配会导致一些节点配额过量,而另一些节点资源配额不足。Max-min 算法的公平性虽然随 β 的增大有较快的递减,但 Max-min 算法考虑了节点的资源需求,因此,不会出现 Avg 算法带来的问题。而 DRF 从优势份额的角度考虑资源的分配,既兼顾到节点的资源需求,又保证了节点尽可能多地获得资源。因此,DRF 是一种兼顾公平和效率的算法。

4.3.2 基于效用的公平性算法评估实验

在本实验中,通过 DFE 模型对基于效用的 α -fairness 算法进行评估。 α -fairness 实质是一种公平性调整算法,通过对参数 α 的调整,在资源分配公平性和利用率之间寻找均衡点 (equilibrium point) [13]。

实验主要参数分别取 $\beta = 0.5$, $\lambda = \frac{1-\beta}{\beta}$, 新节点产生概率 $p_\omega = 0.6$, 销毁计算节点概率 $p_\phi = 0.7$, 资源总量 $R = \{600\}$ (α -fairness 为单资源算法), 初始节点数量 $n = 10$, 节点资源需求 $5 \leq d_{ij} \leq 25$, 时间 $t \in [0, 50]$ 。采用 Max-min 进行资源分配,通过改变 α 参数对公平值进行调整。

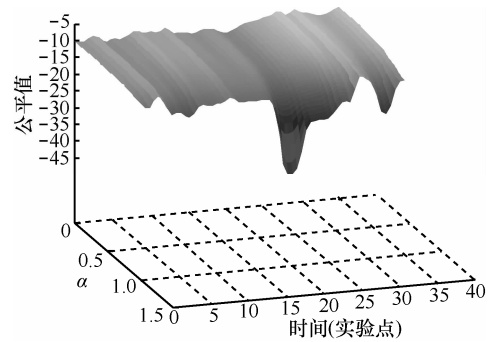


图 6 基于 DFE 的公平性度量 (以 α -fairness 算法为例)

图 6 给出了多次实验中,通过改变 α 的值,分配公平性的变化情况。由于节点数量随时间变化,因此,图中去掉了节点数量为 0 时的测试点。可以看出,随着 α 值的增大,公平值逐渐减小。然而,在不同分配结果的情况下,调整 α 带来的公平值衰减速率也不同。

图 7 显示了在 2 种不同资源分配结果下,调整 α 值后引起的公平值变化。如图 7(a)所示, $t = 1$, 此时原始分配结果为 $A_1 = \{38, 78, 58, 62, 50, 70, 58, 46, 90, 50\}$ 。

此时资源占用总量为 $O_1 = \sum A_1 = 600$ 。由 α -fairness 可知,当 $\alpha = 0$ 时,达到最大公平,此时 $A_1^{\alpha=0} = \{a_i | a_i = 60\}$, 即 $\frac{600}{10}$ 。而 A_1 实质是 $\alpha = 1$ 时的分配结果。在 α 从 0 增长到 1 的过程中,公平值是不下降的。从图 7(a)中可以看出,公平值的下降幅度较小,为 $-10 \sim -11.2$ 。这是由于 A_1 中各分量值相差不大,因此,调整 α 不会引起公平值的较大变化。

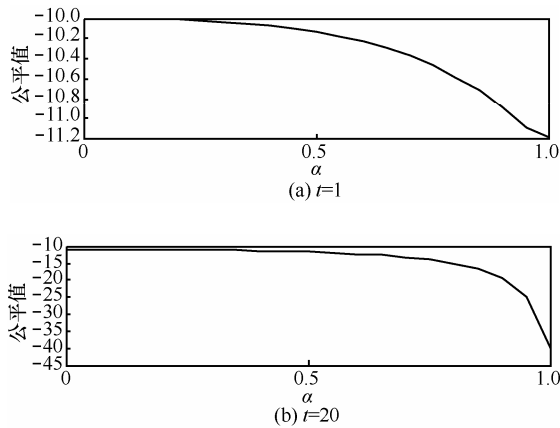


图 7 不同分配方案下的公平性变化

如图 7(b)所示, $t = 20$, 此时资源分配结果为: $A_{20} = \{90, 55.75, 45.94, 83.53, 80, 28, 104.51, 44.51, 46.88, 6, 14.88\}$ 资源占用总量为 $O_{20} = \sum A_{20} = 600$ 。同理, 当 $\alpha = 0$ 时, 公平值达到最大, 此时 $A_{20}^{\alpha=0} = \{a_i | a_i \approx 54.55\}$, 即 $\frac{600}{11}$ 。而 A_{20} 为 $\alpha = 1$ 时的分配结果。在 α 增长的过程中, 公平值的下降幅度较大, 这是由于 A_{20} 中各分量值相差较大 ($-10 \sim -40$), 分配极为不均, 调整 α 参数会引起公平值较大幅度的变化。在实际应用中, 虽然 $\alpha = 0$ 时能够获得最大公平性, 但此时是对资源的绝对平均划分 (等同于 Avg 方法), 并未考虑节点的资源需求。因此, 采用 α -fairness 对公平性进行调整时, 既要避免公平值有较大的损失, 又需要充分考虑节点的资源需求, 在这两者之间寻求均衡点。

综上, 通过分别采用分配算法 (DRF、Avg、Max-min) 和基于效用的公平性调整算法 α -fairness, 验证了 DFE 模型能够有效地评估不同算法的公平性, 为云计算资源分配算法的选取提供重要的参考。

5 结束语

本文提出了一种基于优势份额的动态公平性度量模型 DFE。模型针对云计算平台虚拟化计算节点数量以及资源需求变化的特点, 引入了时间和概率参数, 建立了 2 个子模型: 1) 节点资源需求动态模型; 2) 节点数量动态模型。基于这 2 个基本模型提出了 DFE, 对云计算资源分配算法公平性进行定量的度量。文章分别采用资源分配算法 (以 DRF、Avg、Max-min 为例) 和公平性调整算法 (以 α -fairness 为例) 对 DFE 模型进行了验证, 进一步

说明该模型能够在资源变化的情况下有效地反映分配算法的公平性变化, 为分配算法的选择提供了重要的参考和依据。

参考文献:

- [1] BERTSEKAS D, GALLAGER R. Data Networks[M]. Prentice Hall, 1992.
- [2] CHARNY A, CLARK D, JAIN R. Congestion control with explicit rate indication[A]. Proceedings of IEEE ICC[C]. 1995.1954-1963.
- [3] TAN L S, PUGH A C, YIN M. Rate-based congestion control in ATM switching networks using a recursive digital filter[J]. Control Engineering Practice, 2003, 11(10):1171-1181.
- [4] KELLY F P. Charging and rate control for elastic traffic[J]. European Transaction on Telecommunications, 1997,(8):33-37.
- [5] MASSOULIÉ L, ROBERTS J. Bandwidth sharing: objectives and algorithms[A]. INFOCOM[C]. New York, 1999.1395-1403.
- [6] BARUAH S K, JOHANNES G, GREG P C. Fast scheduling of periodic tasks on multiple resources[A]. Proceedings of the 9th International Conference on Parallel Processing Symposium (IPPS1995)[C]. Santa Barbara, CA, 1995. 280-288.
- [7] BARUAH S K, COHEN N K, PLAXTON C G, et al. Proportionate progress: a notion of fairness in resource allocation[J]. Algorithmica, 1996, 15(6): 600-625.
- [8] ZHU D, MOSSÉ D, MELHEM R G. Multiple-resource periodic scheduling problem: how much fairness is necessary?[A]. RTSS, IEEE Computer Society[C]. 2003.142-151.
- [9] BLANQUER J M, ÖZDEN B. Fair queuing for aggregated multiple links[A]. SIGCOMM[C]. 2001. 189-197.
- [10] KLEINBERG J M, RABANI Y, TARDOS É. Fairness in routing and load balancing[J]. Journal of Computer System Science, 2001, 63(1): 2-20.
- [11] LIU Y, KNIGHTLY E W. Opportunistic fair scheduling over multiple wireless channels[A]. INFOCOM[C]. San Francisco, CA, 2003. 1106-1115.
- [12] GHODSI A, ZAHARIA M, HINDMAN B, et al. Dominant resource fairness: fair allocation of multiple resource types[A]. Proceedings of the 8th USENIX Conference on Networked Systems Design and Implementation[C]. Berkeley, CA, 2011.24.
- [13] MOSHE Z, LIANSHENG T, HANWU W, et al. Efficiency-fairness tradeoff in telecommunications networks[J]. IEEE Communications Letters, 2005,7, 9(7):643-645.
- [14] SEDIQ A B, GOHARY H, YANIKOMEROGLU H. Optimal tradeoff between efficiency and Jain's fairness index in resource allocation[A]. IEEE PIMRC[C]. Sydney, NSW, 2012.577-583.
- [15] CAMPEGIANI P A. Genetic algorithm to solve the virtual machines resources allocation problem in multi-tier distributed systems[A]. The 2nd International Workshop on Virtualization Performances: Analysis, Characterization and Tools Boston[C]. Massachusetts, 2009.
- [16] GU J H, HU J H, ZHAO T H. A new resource scheduling strategy based on genetic algorithm in cloud computing environment[J]. Journal of Computers, 2012, 7(1):42-52.
- [17] 郑伟伟, 邹华, 林荣恒. 云虚拟环境下资源分配的研究与实现[J].

软件, 2012,33(1):46-48.

ZHENG W W, ZOU H, LIN R H. Research of resource allocation for cloud computing environment[J]. Software, 2012, 33(1):46-48.

[18] STILLWELL M, VIVIEN F, CASANOVA H. Dynamic fractional resource scheduling for HPC workloads[A]. Proceedings of the 24th IEEE International Parallel and Distributed Processing Symposium[C]. Atlanta, GA, 2010.1-12.

[19] 张小庆, 李春林, 钱琼芬. 基于合作博弈的虚拟化资源效用分配策略. 计算机科学[J]. 2012,39(6):51-53.

ZHANG X Q, LI C L, QIAN Q F. Utility allocation strategy for virtualized resource based on cooperative game[J]. Computer Science, 2012, 39(6):51-53.

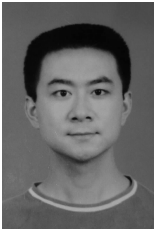
[20] 张顺利, 邱雪松, 潘亚莲. 网络虚拟化环境下基于预测的资源重配置算法[J]. 通信学报, 2011, 32(7):57-63.

ZHANG S L, QIU X S, PAN Y L. Forecast-based resource reconfiguration algorithm for network virtualization[J]. Journal on Communications, 2011, 32(7):57-63.

[21] LAN T, CHIANG M. An axiomatic theory of fairness in resource allocation[A]. Proceedings of the 29th Conference on Information Communications[C]. San Diego, CA, 2010.1343-1351.

[22] WONG C J, SEN S, LAN T, *et al.* Multi-resource allocation: fairness-efficiency tradeoffs in a unifying framework[A]. INFOCOM, IEEE[C]. Orlando, FL, 2012.1206-1204.

作者简介:



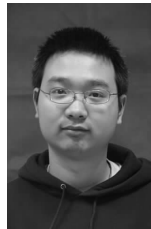
卢笛 (1983-), 男, 陕西安康人, 西安电子科技大学博士生, 主要研究方向为虚拟化技术、存储系统、操作系统安全、虚拟化资源调度。



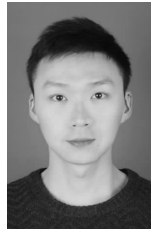
马建峰 (1963-), 男, 陕西西安人, 博士, 西安电子科技大学教授、博士生导师, 主要研究方向为计算机安全、密码学、移动与无线网络安全。



王一川 (1983-), 男, 河南开封人, 西安电子科技大学博士生, 主要研究方向为云计算环境下密钥管理、虚拟节点间安全通信。



习宁 (1986-), 男, 陕西渭南人, 西安电子科技大学博士生, 主要研究方向为服务计算、信息流安全、异构网络融合安全。



孟宪佳 (1985-), 男, 陕西西安人, 西安电子科技大学博士生, 主要研究方向为面向服务的计算、信任管理、网络结构。