

基于博弈论的空间数据中心私有云资源分配管理分析

翟永 刘津 刘磊 陈杰

国家基础地理信息中心 北京 100830

(zhaiyong@ngcc.cn)



摘要 针对空间数据中心私有云资源使用过程中存在的浪费和低效问题,采用算法博弈论的数学方法研究分析了用户资源占有的驱动动机,得出资源在用户之间均分时,在人人相互制衡的前提下全局满意度最大的结论。在此基础上,进一步分析了资源使用在个人优先前提下和集体优先前提下的使用特点,得出为保持全局满意度最大且能够可持续使用资源,采用集体优先前提下的资源分配模式更优的结论。基于上述两个结论,构建了集体优先前提下的以用户自治、IT管理部门支持为特征的资源分配与管理博弈模型,并给出了资源分配决策、用户行为分析和用户满意度评价的数学方法。通过结合空间数据中心资源使用实际数据进行计算,验证了所提出的资源分配与管理博弈模型和满意度评价方法的适用性。该算法对解决空间数据中心私有云资源利用率低下等问题具有参考价值。

关键词: 算法博弈论;私有云资源;个人优先前提;集体优先前提;自治分配;纳什均衡

中图法分类号 O225

Analysis of Private Cloud Resource Allocation Management Based on Game Theory in Spatial Data Center

ZHAI Yong, LIU Jin, LIU Lei and CHEN Jie

National Geomatics Center of China, Beijing 100830, China

Abstract In view of the problems of waste and inefficiency with the use of private cloud resources in spatial data centers, the driving motivation of user resource possession is analyzed by the mathematical method of algorithmic game theory. It is concluded that when resources are shared equally among users, the global satisfaction is the largest under the premise that everyone checks and balances each other. Based on the above conclusion, the characteristics of resource use under the premise of individual priority and collective priority are further analyzed. It is concluded that it is better to adopt a resource allocation model under the collective priority, which not only can maintain maximum global satisfaction, but also sustainable use of resources. Based on the above two conclusions, resource allocation and management game model is constructed, which features are user autonomy and IT management department support under the premise of collective priority, and the mathematical methods of resource allocation decision-making and user behavior analysis and user satisfaction evaluation are given. Then, the applicability of the proposed resource allocation and management game model and satisfaction evaluation method are verified by the calculation and verification of the actual data in the spatial data center. This algorithm has reference value for solving the problem of the low utilization rate of private cloud resources in the spatial data center.

Keywords Game theory, Private cloud resource, Individual priority, Collective priority, Autonomy allocation, Nash Equilibrium

1 引言

随着空间数据获取技术与信息技术的融合发展,地图、影像、地形等空间数据的精度和分辨率不断提高,种类不断丰富,覆盖范围不断扩展,数据量呈指数级增长^[1-4]。以国家基础地理信息中心为例,2009年管理的在线空间数据量为12TB,2019年底则达到了5.6PB,十年间数据量增长了478倍。数据量的持续增长对空间数据中心的私有云计算和存储能力提出了严峻的挑战。在空间数据中心内部,私有云计算和存储资源常常处于短缺状态。一方面,用户时常抱怨资源不足,影响业务运行;另一方面,IT管理部门则认为用户使用不够科学合理,存在低效浪费现象。在空间数据中心内部,上演着一场又一场私有云资源使用与管理的博弈^[5]。

在空间数据中心内部,计算和存储等私有云资源在申请和分配方面与公有云存在本质不同。私有云资源长期免费使用,公有云资源需付费使用;私有云资源的提供者、管理者和用户常处于同一数据中心内部,而公有云这3类人员多属于不同机构。目前,空间数据中心私有云资源分配多考虑效率因素,主要采用主管领导集中决策模式。长期来看,该模式难以排除主观因素的影响。

目前,学术界对私有云资源分配的研究主要集中在虚拟化技术下的动态分配理论。在私有云环境下,计算、存储、网络等均可作为虚拟资源池向用户提供服务。为保障用户业务的正常运行,动态分配理论为每位用户设置了资源使用基线阈值,并动态监测用户应用的实时情况,来弹性优化资源分配策略,实时增加或回收资源。动态分配理论与私有云的技术

特点紧密结合,能够灵活分配资源。但是,动态分配理论在用户资源使用基线阈值的设定、不同应用对资源的个性要求、资源增减与业务未来发展的匹配等方面,难以采用动态、自动手段,同时该理论对用户使用行为缺乏分析,对不合规行为缺少反馈。动态分配理论得益于技术手段的进步,但是在资源分配的公平与效率的结合上依旧存在矛盾,没有将用户的利益需求与资源分配的满意度相结合进行考虑。因此,在利益需求和实现层面,针对空间数据中心内部私有云资源使用与管理的博弈现实,引入博弈论或者算法博弈论进行分析,将效率与公平紧密结合,以实现用户间博弈的均衡,从而达到帕累托最优;在技术实现层面,可借助私有云资源的动态分配技术进行操作。

本文基于算法博弈论对空间数据中心私有云资源分配管理模式进行研究与应用分析,探究用户资源占有的驱动动机和公用资源使用的特点,构建以用户自治、IT 管理部门支持为特征的资源分配与管理博弈模型,给出用户满意度的评价方法,经验证该模型能够提高资源利用率,提升用户的满意度。

2006 年以来兴起的算法博弈论^[6]主要关注算法的可计算性,主要应用于以下研究领域:1)研究均衡的计算复杂性问题;2)从博弈论的观点研究计算机学科中的问题;3)算法机制设计领域,包括网络结构及性能方面的研究、在线拍卖和在线交易、在线广告、搜索结果页面排序及分布式应用;4)计算性社会选择等问题^[7]。在计算机学科的具体研究方向上,博弈论主要应用于网络、计算机安全和公有云资源管理等领域^[8-10],如异构无线网络接纳控制和负载均衡^[11]、无线自组网络能耗控制^[12]、网络对抗^[13]、主动防御^[14]、计算信任、病毒防御^[15]以及云资源分配^[16-18]等。目前,基于算法博弈论的私有云资源分配与管理的研究文献很少。

2 资源占有的驱动动机分析

本文从算法博弈论的视角来讨论私有云资源占有的用户驱动动机,探讨如何分配才能使所有用户的满意度最高。用户满意度分析是博弈论衡量分配公平性的重要指标。

设空间数据中心可用来分配的私有云资源总量为 S,用户数量为 n。私有云资源分配的前提包括:1)资源全部分配完毕;2)用户是理性的,地位是平等的,能够相互监督、相互制约;3)一个用户得到的资源越多,满意度就越高;反之,满意度越低。

设任意一个用户获得的资源数量为 $x_i, x_i \in (0, S)$,令其满意度函数 $v_i(x_i, x_{-i}) = \ln(x_i)$,由前提 3)可知用户的满意度与 $\ln(x_i)$ 成正比。

所有人的满意度之和,即满意度函数之和 $w(x_1, x_2, \dots, x_n) = \sum_{i=1}^n v_i(x_i, x_{-i}) = \sum_{i=1}^n \ln(x_i)$ 。从所有用户的角度分析,资源分配满意度最大就是所有人最满意的分配方案。资源分配满意度达到最大即为求解 $Max(w(x_1, x_2, \dots, x_n))$ 。

设 $\sum_{i=1}^n x_i = S$,且 $x_i \in (0, S)$, $Max(w(x_1, x_2, \dots, x_n)) = Max(\sum_{i=1}^n \ln(x_i)) = Max \ln(\prod_{i=1}^n x_i)$,则资源分配满意度最大值求解等价于求解 $Max(\prod_{i=1}^n x_i)$ 。

由 $\sum_{i=1}^n x_i = S$,得:
 $x_n = S - x_1 - x_2 - \dots - x_{n-1}$

$$\begin{aligned} \prod_{i=1}^n x_i &= x_1 x_2 \cdots x_{n-1} (S - x_1 - x_2 - \cdots - x_{n-1}) \\ &= \prod_{i=1}^{n-1} x_i (S - x_1 - x_2 - \cdots - x_{n-1}) \\ &= S \prod_{i=1}^{n-1} x_i - x_1 \prod_{i=1}^{n-1} x_i - x_2 \prod_{i=1}^{n-1} x_i - \cdots - x_{n-1} \prod_{i=1}^{n-1} x_i \end{aligned}$$

令 $m = \prod_{i=1}^n x_i$,当 m 存在极值时,有:

$$\begin{cases} \frac{\partial m}{\partial x_1} = 0 \\ \frac{\partial m}{\partial x_2} = 0 \\ \vdots \\ \frac{\partial m}{\partial x_{n-1}} = 0 \end{cases}$$

$$\frac{\partial m}{\partial x_1} = S \prod_{i=2}^{n-1} x_i - 2 \prod_{i=1}^{n-1} x_i - x_2 \prod_{i=2}^{n-1} x_i - \cdots - x_{n-1} \prod_{i=2}^{n-1} x_i = 0$$

整理上式,得:
 $S - 2x_1 - x_2 - \cdots - x_{n-1} = 0$ (1)

同理, $\frac{\partial m}{\partial x_2} = 0$,计算得:
 $S - x_1 - 2x_2 - \cdots - x_{n-1} = 0$ (2)

由 $\frac{\partial m}{\partial x_i} = 0$,得:
 $S - x_1 - x_2 - \cdots - x_{i-1} - 2x_i - x_{i+1} - \cdots - x_{n-1} = 0$ (3)

由 $\frac{\partial m}{\partial x_{n-1}} = 0$,得:
 $S - x_1 - x_2 - \cdots - 2x_{n-1} = 0$ (4)

由上文得方程组:
$$\begin{cases} S - 2x_1 - x_2 - \cdots - x_{n-1} = 0 \\ S - x_1 - 2x_2 - \cdots - x_{n-1} = 0 \\ \vdots \\ S - x_1 - x_2 - \cdots - x_{i-1} - 2x_i - x_{i+1} - \cdots - x_{n-1} = 0 \\ \vdots \\ S - x_1 - x_2 - \cdots - 2x_{n-1} = 0 \end{cases}$$
 (5)

求解方程组(5),得:
 $x_1 = x_2 = \cdots = x_{n-1} = \frac{S}{n}$ (6)

则 m 的驻点为 $(x_1, x_2, \dots, x_{n-1}) = (\frac{S}{n}, \frac{S}{n}, \dots, \frac{S}{n})$,此时 $x_n = S - \sum_{i=1}^{n-1} x_i = \frac{S}{n}$ 。下面判断 m 在驻点的取值是否为最大值。

m 的 Hessian 矩阵为:

$$\begin{aligned} H &= \begin{bmatrix} \frac{\partial^2 m}{\partial x_1^2} & \frac{\partial^2 m}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 m}{\partial x_1 \partial x_{n-1}} \\ \frac{\partial^2 m}{\partial x_2 \partial x_1} & \frac{\partial^2 m}{\partial x_2^2} & \cdots & \frac{\partial^2 m}{\partial x_2 \partial x_{n-1}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 m}{\partial x_{n-1} \partial x_1} & \frac{\partial^2 m}{\partial x_{n-1} \partial x_2} & \cdots & \frac{\partial^2 m}{\partial x_{n-1}^2} \end{bmatrix} \\ &= (\frac{S}{n})^{n-2} \begin{bmatrix} -2 & -1 & -1 & \cdots & -1 \\ -1 & -2 & -1 & \cdots & -1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -1 & -1 & -1 & \cdots & -2 \end{bmatrix} \end{aligned}$$

求解 H 的特征值,有:

$\det(\lambda I - H) = 0$

则

$$\lambda_1 = \lambda_2 = \cdots = \lambda_{n-2} = -\left(\frac{s}{n}\right)^{(n-2)}$$

$$\lambda_{n-1} = -n\left(\frac{s}{n}\right)^{(n-2)}$$

则 m 的特征值 $\lambda_i < 0, 1 \leq i \leq n-1$ 。因此, m 的 Hessian 矩阵为负定矩阵, 则其在驻点取得最大值, 得:

$$\text{Max}(m) = \text{Max}\left(\prod_{i=1}^n x_i\right) = \left(\frac{s}{n}\right)^n$$

$$\text{Max}(w(x_1, x_2, \cdots, x_n)) = \ln(\text{Max}\left(\prod_{i=1}^n x_i\right)) = \ln\left(\left(\frac{s}{n}\right)^n\right) = n \ln\left(\frac{s}{n}\right)$$

这说明资源在用户之间均分时, 全局满意度最大。这也是纳什均衡解。因此, 私有云资源的分配者应以平均分配为公平保障的根本手段, 应视为基本原则。

针对资源的分配, 中国历史上存在着“均贫富”“毋患寡而患不均”的观念, 《礼记·礼运》篇认为大同世界即大道之世, 是人人有德、无处不均匀的理想社会。这说明了平均主义分配方式从古代就是人类的美好愿望, 主要在于能够实现人人满意、全局满意度最大这个最优结果。

因此, 可以认为, 在计算机资源分配中, 用户将平均分配作为资源占有的驱动力。

为维持资源的可持续性使用, 更贴近实际情况, 下文将考虑资源的剩余因素影响。并在上文提到的遵循平均主义分配原则的基础上, 分为个人优先前提、集体优先前提两种情况对资源分配进行分析。

3 资源使用的特点分析

在资源使用免费的场景下, 空间数据中心私有云资源等同于公共资源, 用户对这类资源的使用常常具有只占不退、多多益善的特点, 伴随而生的是使用过程中常存在的浪费、低效现象。

在空间数据中心内部, 设有 n 个用户进行数据处理工作, 需要私有云资源提供支撑。该数据中心的私有云资源总量为 S , 设每一个用户 i 为维持其业务工作必须占有的私有云资源为 $x_i > 0$, n 个用户使用后剩下的私有云资源为 $S - \sum_{i=1}^n x_i$ 。占有资源 x_i 给用户 i 带来的满意度为 $\ln(x_i)$, 没有其他用户从用户 i 的选择中获得满意度。每一个用户都具有享受剩余私有云资源的权利, 这为每个用户带来的满意度均为 $\ln(S - \sum_{i=1}^n x_i)$ 。因此, 用户 i 得到满意度的计算公式如下:

$$v_i(x_i, x_{-i}) = \ln(x_i) + \ln(S - \sum_{i=1}^n x_i)$$

3.1 个人优先前提下的纳什均衡求解

个人优先前提下的分配模式更重视用户满意度的提升, 使得个人满意度达到最大值。

为了求出纳什均衡解, 需要分别求解每一个用户的最优反应函数以计算其最优满意度, 从而找出满足全部最优反应函数条件的策略剖面。 n 个用户的纳什均衡解满足:

$$\begin{cases} \frac{\partial v_1(x_1, x_{-1})}{\partial x_1} = 0 \\ \frac{\partial v_2(x_2, x_{-2})}{\partial x_2} = 0 \\ \vdots \\ \frac{\partial v_n(x_n, x_{-n})}{\partial x_n} = 0 \end{cases}$$

该方程组中, 用户 i 的最优反应函数计算如下:

$$\frac{\partial v_i(x_i, x_{-i})}{\partial x_i} = \frac{1}{x_i} - \frac{1}{S - \sum_{i=1}^n x_i} = 0$$

用户 i 的最优反应函数为:

$$BR_i(x_{-i}) = x_i = \frac{S - \sum_{j \neq i} x_j}{2}$$

得到如下方程组:

$$\begin{cases} BR_1(x_{-1}) = x_1 = \frac{S - \sum_{j \neq 1} x_j}{2} \\ BR_2(x_{-2}) = x_2 = \frac{S - \sum_{j \neq 2} x_j}{2} \\ \vdots \\ BR_n(x_{-n}) = x_n = \frac{S - \sum_{j \neq n} x_j}{2} \end{cases} \quad (7)$$

对式(7)化简, 得:

$$\begin{cases} 2x_1 + \sum_{j \neq 1} x_j = S \\ 2x_2 + \sum_{j \neq 2} x_j = S \\ \vdots \\ 2x_n + \sum_{j \neq n} x_j = S \end{cases} \quad (8)$$

求解方程组(8), 得:

$$x_1 = x_2 = \cdots = x_n = \frac{S}{n+1}$$

$$\sum_{i=1}^n x_i = \frac{n}{n+1} S$$

剩余资源为:

$$S - \sum_{i=1}^n x_i = \frac{1}{n+1} S$$

全局满意度为:

$$2n \ln \frac{1}{n+1} S$$

个人优先前提分配模式下, 每位用户占有的资源相同, 均为 $\frac{S}{n+1}$, 满足平均主义的分配原则, 同时剩余资源为 $\frac{1}{n+1} S$ 。

3.2 集体优先前提下的纳什均衡求解

集体优先前提下的分配模式在重视资源的可持续使用的同时, 使得全局满意度达到最大值。

下文将求解全局满意度函数的最大值。

$$\text{令 } w(x_1, x_2, \cdots, x_n) = \sum_{i=1}^n v_i(x_i, x_{-i}) = \sum_{i=1}^n \ln(x_i) + n \ln(S - \sum_{i=1}^n x_i)$$

当 $w(x_1, x_2, \cdots, x_n)$ 存在最大值时, 有:

$$\begin{cases} \frac{\partial w(x_1, x_2, \cdots, x_n)}{\partial x_1} = 0 \\ \frac{\partial w(x_1, x_2, \cdots, x_n)}{\partial x_2} = 0 \\ \vdots \\ \frac{\partial w(x_1, x_2, \cdots, x_n)}{\partial x_n} = 0 \end{cases} \quad (9)$$

求解方程组(9), 得:

$$x_1 = x_2 = \cdots = x_n = \frac{S}{2n}$$

$$\sum_{i=1}^n x_i = \frac{1}{2} S$$

剩余资源为:

$$S-\sum_{i=1}^n x_i=\frac{1}{2} S$$

则全局满意度为：

$$n \ln \frac{S^2}{4 n}$$

比较个人优先前提下和集体优先前提下的分配模式的纳什均衡解,可得出如下结论：

- 1)两者均遵循平均主义的分配原则；
- 2)集体优先前提下,用户*i*的满意度小于个人优先前提,即 $\frac{S}{2 n}<\frac{S}{n+1}, n>1$ ；
- 3)集体优先前提下,满意度总和大于个人优先前提,即 $n \ln \frac{S^2}{4 n}>2 n \ln \frac{1}{n+1} S, n>1$ ；
- 4)集体优先前提下,资源剩余总和大于个人优先前提,即 $\frac{S}{2}>\frac{1}{n+1} S, n>1$ 。

以上分析表明,在个人优先前提下的分配模式中,用户对资源占有欲的约束低于集体优先前提,用户平均占有更多的资源,而剩余资源相对较少,可持续使用的能力不足。研究表明,在私有云资源的分配实际中,个人优先前提下的分配模式易导致“公地悲剧”^[19],对私有云资源的使用效率低于集体优先,易导致资源浪费。

避免或有效降低资源浪费,实现全局满意度最大,且保持可持续使用,采用集体优先前提下的资源分配模式更优。但是,在实际资源分配过程中,用户的各自情况不尽相同,绝对的平均主义是不合理的,加权平均主义是可行的选择。因此,如何合理决策资源分配方案?是通过外部权威来命令,抑或通过用户内部博弈来达到目的。这是下文需要解决的问题。

4 私有云资源的分配管理

4.1 “自治”分配

目前,在私有云资源的分配中,只占不退、资源浪费等问题困扰着 IT 管理部门。确保私有云资源分配合理,是摆在 IT 管理部门面前需要解决的重要问题。

本文首先引入双方资源分配博弈模型进行辅助分析。设甲乙两类用户共同申请使用计算机资源,如何分配才能保障公平?易知双方资源分配的纯策略纳什均衡解是一人获得资源决策权,另一人获得资源分配权,此为最优策略。如果多人参与资源分配,如何解决?可以采用用户推举的方式,分别确定资源决策人和分配人。为确保公平,同一人不能同时承担资源决策和分配职责。

上述资源分配策略是否比引入第三方(如权威)来负责资源分配更好?实践表明,权威的引入从长期看难以确保公平,且成本较高^[20]。基于集体优先前提,采用双方资源分配博弈模型,可持续使用资源的同时,用户之间通过博弈实现“自治”管理的方式更有效率。

私有云资源分配采用“自治”模式,分配成功与否与用户的申请和使用行为有关。本文将用户从自身利益出发,虚报需求(如夸大任务参与人数、数据量等)、只占不退等视为异常使用方式;反之,将实报需求、合规使用等视为正常使用方式。

因此,将用户划分为正常使用、异常使用两个群体,二者在私有云资源使用上存在博弈关系。本文以正常使用、异常使用两个群体在私有云资源分配中的博弈为例,来说明资源使用“自治”管理的有效性。

设用户的策略集 $\sigma=(\sigma_1, \sigma_2)=(\text{正常使用}, \text{异常使用})$,双方正常使用的满意度为 $v=(v_1, v_2)$,异常使用的满意度为 $g=(g_1, g_2)$,正常使用的用户因其他用户异常使用而遭受的损失为 $l=(l_1, l_2)$,异常使用被发现的概率为 $x=(x_1, x_2)$,异常使用被发现后受到的惩罚为 $d=(d_1, d_2)$ 。私有云资源使用满意度矩阵如图 1 所示。

		使用部门2	
		正常使用	异常使用
使用部门1	正常使用	v_1, v_2	$v_1-l_1, g_2-x_2 d_2$
	异常使用	$g_1-x_1 d_1, v_2-l_2$	$g_1-x_1 d_1-l_1, g_2-x_2 d_2-l_2$

图 1 私有云资源使用满意度矩阵

Fig. 1 Private cloud resource usage satisfaction matrix

用户比较异常使用、正常使用分别获得的满意度来决定使用行为。下面分析不同策略下的用户满意度纳什均衡解。

- （1）纯策略纳什均衡求解过程
- 1)如果 $g_1-x_1 d_1>v_1$,且 $g_2-x_2 d_2>v_2$,纯策略用户满意度纳什均衡解为：
 $(g_1-x_1 d_1-l_1, g_2-x_2 d_2-l_2)$
即用户全部采用异常使用方式。这是私有云资源分配应避免的策略选择。
- 2)如果 $g_1-x_1 d_1\leq v_1$,且 $g_2-x_2 d_2\leq v_2$,纯策略用户满意度纳什均衡解为：
 (v_1, v_2)
即用户全部采用正常使用方式。这是私有云资源分配应采用的最优策略。

- （2）混合策略纳什均衡求解过程
- 设用户 1 采用正常使用行为的概率为 p ,采用异常使用的概率为 $1-p$;用户 2 采用正常使用行为的概率为 q ,采用异常使用的概率为 $1-q$ 。本文令 nu 代表正常使用, anu 代表异常使用。

用户 1 从其两个纯行动中的每一个获得的期望支付如下：

$$\begin{aligned} v_1(nu, q) &= q \times v_1 + (1-q) \times (v_1-l_1) \\ &= v_1-l_1(1-q) \\ v_1(anu, q) &= q \times (g_1-x_1 d_1) + (1-q) \times (g_1-x_1 d_1-l_1) \\ &= g_1-x_1 d_1-l_1(1-q) \end{aligned}$$

- 当且仅当 $v_1(nu, q) \geq v_1(anu, q)$,即：
 $v_1-l_1(1-q) \geq g_1-x_1 d_1-l_1(1-q)$
 $v_1 \geq g_1-x_1 d_1$
此时,用户 1 偏好正常使用行为胜过异常使用。
反之,当 $v_1(nu, q) < v_1(anu, q)$,即：
 $v_1-l_1(1-q) < g_1-x_1 d_1-l_1(1-q)$
 $v_1 < g_1-x_1 d_1$
此时,用户 1 偏好异常使用行为胜过正常使用。

同理,当且仅当 $v_2(nu, p) \geq v_2(anu, p)$,即 $v_2 \geq g_2-x_2 d_2$,用户 2 偏好正常使用行为胜过异常使用;反之,当 $v_2(nu, p) < v_2(anu, p)$,即 $v_2 < g_2-x_2 d_2$,用户 2 偏好异常使用行为胜过正常使用。

在私有云资源分配决策中,在利益的驱使下,用户往往采纳满意度最大的使用行为。通过用户之间的相互监督,发现资源分配中的问题并协商形成妥协或惩罚措施,可促使用户选择正常的使用策略。例如异常使用被发现后,可以对该用户现在以及未来资源使用进行惩罚,降低用户通过异常行为

获得的满意度来进行约束。如果异常使用难以被发现或发现后的惩罚力度不够,用户可能选择异常使用策略来追求自己的最大利益,则个人优先分配模式将逐渐占据主流,资源难以可持续使用。

本文提出的解决方案与目前业界常规基于权威的分配策略明显不同,用户的资源占用不需要权威来决策。在空间数据中心内部,目前的实际情况多是IT管理部门既负责私有云资源的运维,又负责资源的分配,承担着资源分配者和维护者的双重角色,这是问题的症结所在。事实上,IT管理部门不应该是“资源分配者”(不是权威),而应为资源的保管者、维护者。本文认为资源应由用户自己按照集体优先的前提进行内部博弈,提出加权平均主义分配方案。

在具体操作上,IT管理部门作为资源的保管者、运维者,应将空间数据中心可分配的私有云资源的有关信息公示给用户,实现完全信息博弈。用户通过推举生成召集人,根据各自的实际需求,集体讨论形成分配方案。IT管理部门负责在技术层面落实上述分配方案,实现私有云资源的分配策略。同时,在私有云资源的使用中,IT管理部门负责将各用户的资源使用情况等统计信息动态公布,使用户相互监督。因此,用户和IT管理部门之间的博弈转移为用户之间的博弈,实现了用户“自治”。

IT管理部门从私有云资源的分配决策中撤出,用户有权按照集体优先前提进行决策,同时,空间数据中心的管理者保留对用户提出的私有云资源分配策略的监督权,并将资源分配、使用情况与用户的经济利益或者荣誉挂钩。通过上述方法,能够较好地避免个人优先前提下私有云资源只占不退、过度占用这类公地悲剧问题,实现相对公平(加权平均主义)。

4.2 量化决策

私有云资源分配采用用户“自治”管理,需要量化指标支持,否则易陷入无休止的争吵。量化指标应依据用户的任务经费、完成时间以及参与人数、数据量、已占有资源及其历史合理度等进行专家打分,作为资源分配决策的参考数据。

设用户 i 申请新增资源为 ΔS ,其任务经费为 F (以万元为单位),任务完成时间为 t (以年为单位),任务参与人数为 n ,数据总量为 D (以TB为单位),算力总量为 E (以CPU核数为单位),该用户使用资源的历史合理度为 k , $k \in (0, 1)$ 。

则用户 i 申请新增存储资源 ΔS_1 的评分为:

$$R_{\Delta 1}^i = \frac{(\log F) \times n \times D}{t} \times k \quad (10)$$

用户 i 申请新增计算资源 ΔS_2 的评分为:

$$R_{\Delta 2}^i = \frac{(\log F) \times n \times E}{t} \times k \quad (11)$$

任务经费 F 以万元为单位,并且取以10为底的对数,主要是减小经费对评分的影响,避免经费决定一切。另外,对数函数能够保持原函数的性质不变,不影响评分结论。使用资源的历史合理度 k 为用户一段时间内实际有效利用资源数与其占有资源数之比。

IT管理部门根据用户提交的任务材料,按照式(10)或式(11)进行计算,将所得分值进行公示,可作为各用户新增资源决策的基线参考。

5 私有云资源分配满意度评价方法

在实际工作中,私有云资源分配难以实现绝对的平均主义。用户的业务存在差异,彼此对私有云资源的需求也不尽

相同,因此应在集体优先前提下实行加权平均主义,以实现相对公平。但是,从用户角度出发,平均主义的资源分配驱动动机依旧是满意度的基线衡量指标。因此,计算用户满意度与平均主义分配模式下的最优满意度(纳什均衡解)之间的差值,并进行用户间差值的横向比较,为判断分配是否公平提供评价依据,以避免用户间矛盾的滋生。

以空间数据中心为例,设有 n 个用户需要私有云资源。每一个用户 i 为维持其业务工作,实际占有的私有云资源设为 $y_i, y_i > 0$ 。占有资源 y_i 给用户 i 带来的满意度为 $\ln y_i$,没有其他用户从 i 的选择中得到满意度。设空间数据中心已分配的私有云资源的总量为 S ,则 $S = \sum_{i=1}^n y_i$ 。由于用户从剩余资源中获得的满意度相同,因此本节不考虑剩余资源对满意度评价方法的影响。

用户 i 从资源占有集合 $y = (y_1, y_2, \dots, y_n)$ 中得到的满意度等于:

$$v_i(y_i, y_{-i}) = \ln y_i$$

评价单个用户的满意度,需计算用户 i 的满意度偏差,计算式如下:

$$\Delta w_i = \frac{\ln y_i - \ln\left(\frac{S}{2n}\right)}{\ln\left(\frac{S}{2n}\right)} = \frac{\ln\left(\frac{2n y_i}{S}\right)}{\ln\left(\frac{S}{2n}\right)}$$

Δw_i 越大,说明用户 i 的满意度偏差越高;反之,用户 i 的满意度偏差越低。

评价全局满意度,需进行满意度方差的计算,计算式如下:

$$\begin{aligned} D(w) &= \frac{1}{n-1} \sqrt{\sum_{i=1}^n \left(\ln y_i - \ln \frac{S}{2n} \right)^2} \\ &= \frac{1}{n-1} \sqrt{\sum_{i=1}^n \left(\ln \frac{2n y_i}{S} \right)^2} \end{aligned}$$

$D(w)$ 越大,整体看用户满意度与最优满意度差距越大;反之,越接近理想状态。 $D(w)$ 可作为评价私有云资源分配满意度的指标。在实际工作中,可从一个较长的时间周期(如5年)来综合考查 $D(w)$ 。

6 案例分析

针对待分配计算机资源,依据用户申请新增资源评分,开展以下工作:

- (1)公示待分配计算机资源以及用户申请新增资源评分;
- (2)用户间进行资源占有博弈,实现自治管理;
- (3)根据用户博弈结果,计算最优满意度、满意度和满意度评价;

(4)管理部门根据满意度评价,分析用户博弈结果。如果用户满意度偏差较大,可要求用户重新进行博弈;

(5)根据用户博弈结果,资源运维部门分配用户资源。

通过上述工作流程,用户获取资源,完成资源分配工作。

本文考察了某国家级空间数据中心3个主要的用户 D_1 , D_2 和 D_3 服务器资源的分配情况,并进行说明。

首先以2018年为例说明用户资源分配的自治管理。该年度待分配的服务器资源(按CPU核数统计)为2070颗,用户 D_1 , D_2 和 D_3 的申请新增资源评分分别为63, 17, 29, 27, 144, 42。

用户第一轮自治博弈的结果为 380,212,1478,最优满意度为 5.84,满意度 D_1 为 5.94, D_2 为 5.36, D_3 为 7.30,满意度评价 D_1 为 0.017, D_2 为 -0.082, D_3 为 0.25。各用户满意度偏差绝对值不大于 0.5,空间数据中心的管理者予以认可,责成资源运维部门实施分配。

另外,综合 2012—2018 年间服务器资源的分配情况说明长周期下的满意度计算及其评价。

表 1 列出了 2012—2018 年间 3 个用户服务器资源(按 CPU 核数统计)分配统计情况。图 2 直观描述了 2012—2018 年间 3 个用户服务器资源的分配情况,其中 D_1 、 D_3 服务器资源的占有量呈增长趋势, D_2 服务器资源的占有量呈减少趋势。

表 1 用户服务器资源分配统计表

Table 1 Statistics of user server resource allocation

statistics of user server resource allocation							
	2012	2013	2014	2015	2016	2017	2018
D_1	68	68	100	60	64	380	380
D_2	156	156	474	486	240	144	212
D_3	156	156	624	596	592	942	1478
Total	380	380	1198	1142	896	1466	2070

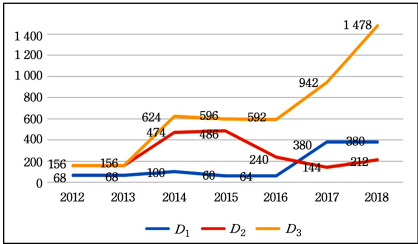


图 2 用户服务器资源分配统计

Fig. 2 Statistics of user server resource allocation

表 2 采用集体优先前提下分配模式的纳什均衡解进行用户最优满意度计算(见 3.2 节),得到基于集体优先前提下的用户服务器资源分配最优满意度统计。根据 3.2 节中集体优先前提下的纳什均衡求解,各用户服务器资源分配最优满意度相等,另外从 2012—2018 年的总体变化趋势看最优满意度呈现出增长的趋势。但数据表明 2014—2016 年间 3 个用户的满意度出现较小幅度的下滑,调研发现主要原因为部分服务器资产集中报废,导致可用服务器资源减少,满意度随之减小。

表 2 基于集体优先的用户服务器资源分配最优满意度统计表

Table 2 Statistics of user server resource allocation optimal satisfaction

Statistics of User server resource allocation optimal satisfaction							
	2012	2013	2014	2015	2016	2017	2018
D_1	4.15	4.15	5.30	5.25	5.01	5.50	5.84
D_2	4.15	4.15	5.30	5.25	5.01	5.50	5.84
D_3	4.15	4.15	5.30	5.25	5.01	5.50	5.84

表 3 列出了 2012—2018 年间 3 个用户服务器资源分配满意度统计情况。图 3 直观描述了 3 个用户在实际工作中 2012—2018 年间的服务器资源分配满意度,其中 D_1 、 D_3 的满意度呈增长趋势, D_2 的满意度呈减少趋势。同时,该图也描述了 3 个用户满意度与基于集体优先的最优满意度的对比情况, D_1 、 D_2 、 D_3 的满意度与最优满意度的线条越相近,表示满意度越高;反之,表示实际工作中的满意度越低。

表 3 用户服务器资源分配实际满意度统计表

Table 3 Statistics of user server resource allocation actual satisfaction

statistics of user server resource allocation satisfaction							
	2012	2013	2014	2015	2016	2017	2018
D_1	4.22	4.22	4.61	4.09	4.16	5.94	5.94
D_2	5.05	5.05	6.16	6.19	5.48	4.97	5.36
D_3	5.05	5.05	6.44	6.39	6.38	6.85	7.30

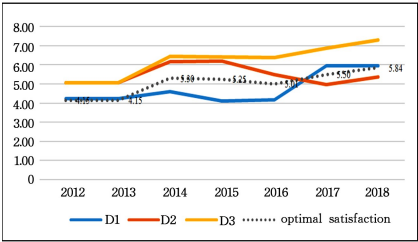


图 3 用户服务器资源分配满意度统计

Fig. 3 Statistics of user server resource allocation satisfaction

图 4 描述了 3 个用户 2012—2018 年间服务器资源分配满意度方差的变化情况。2015 年升高的原因主要是 2014—2016 年间 3 个用户的部分服务器资产集中报废,导致可用服务器资源减少,资源分配与最优满意度要求相比出现了较大的偏离,满意度随之减小。总体看,由于实现了资源分配均衡,兼顾了效率和公平,3 个用户在 2012—2018 年对服务器资源分配较为满意,未出现矛盾。

表 4 用户服务器资源分配满意度方差统计表

Table 4 Statistics of user server resource allocation satisfaction variance

statistics of user server resource allocation satisfaction variance							
	2012	2013	2014	2015	2016	2017	2018
Satisfaction variance	0.64	0.64	0.79	0.94	0.84	0.76	0.77

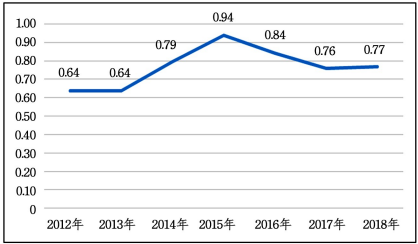


图 4 用户服务器资源分配满意度方差统计

Fig. 4 Statistics of user server resource allocation satisfaction variance

结束语 本文采用算法博弈论的数学方法分析了用户资源占用中存在的平均主义驱动动机和资源使用特点,在兼顾资源分配公平、提高满意度、减少浪费、保持可持续使用的原则下,提出在集体优先模式下以用户自治为特征的资源分配与管理模型,并给出资源量化决策、用户行为分析和用户满意度衡量的数学方法,结合资源使用的实际数据进行了验证。

本文在理论上将空间数据中心的多个用户分为两类进行博弈,采用双方博弈的数学模型研究资源占用的用户行为分析和资源分配等问题,虽然适用于多个用户的情形,但未考虑博弈方为 3 类以上的多方情形。在未来的研究工作中,应引入多方博弈数学模型,对资源使用过程中的复杂情形进行深

入分析,以更好地优化空间数据中心私有云资源的使用效率,促进资源分配更加公平。

在实际工作中,博弈论的方法应与基于虚拟化的动态分配理论相结合,前者解决用户资源占用博弈的均衡,以达到公平;后者解决用户资源需求的实现效率。

参 考 文 献

[1] LI D R,WANG S L,SHI W Z,et al. On Spatial Data Mining and Knowledge Discovery (SDMKD)[J]. Geomatics and Information Science of Wuhan University,2001,26(6):491-499.

[2] LI S,DRAGICEVIC S,CASTRO F A,et al. Geospatial big data handling theory and methods:A review and research challenges [J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2015,115:119-133.

[3] LE P,WU Z Y,SHANGGUAN B Y. Design and implementation of a distributed geospatial data storage structure based on spark [J]. Geomatics and Information Science of Wuhan University, 2018,43(12):2295-2302.

[4] WANG Y,WANG S. Research and implementation on spatial data storage and operation based on hadoop platform[C]//2010 Second IITA International Conference on Geoscience and Remote Sensing. IEEE,2010:275-278.

[5] WEI G,VASILAKOS A V,ZHENG Y,et al. A game-theoretic method of fair resource allocation for cloud computing services [J]. The Journal of Supercomputing,2010,54(2):252-269.

[6] ROUGHGARDEN T. Twenty lectures on algorithmic game theory[M]. Cambridge University Press,2016.

[7] ROUGHGARDEN T. Algorithmic game theory[J]. Communications of the ACM,2010,53(7):78-86.

[8] XU X,YU H. A game theory approach to fair and efficient resource allocation in cloud computing[J]. Mathematical Problems in Engineering,2014,2014.

[9] NEZARAT A,DASTGHAIBIFARD G H. Efficient nash equilibrium resource allocation based on game theory mechanism in cloud computing by using auction[C]//2015 1st International Conference on Next Generation Computing Technologies (NGCT). IEEE,2015:1-5.

[10] MA Y,LI B,ZHANG Y,et al. Efficient auction mechanism with group price for resource allocation in clouds[C]//2014 Second International Conference on Advanced Cloud and Big Data. IEEE,2014:85-92.

[11] CHANG C J, TSAI T L, CHEN Y H. Utility and game-theory

based network selection scheme in heterogeneous wireless networks[C]//2009 IEEE Wireless Communications and Networking Conference. IEEE,2009:1-5.

[12] LI P,TAO Y,XU X Y,et al. Energy consumption balance clustering protocol in wireless sensor network based on game theory [J]. Computer Engineering,2018,44(12):156-162.

[13] GU Q Y,SUN Y L,GAO F. Research on the Game Theory based network attack-defense and its application[J]. Network information security,2013(1):52-54.

[14] JIANG W,FANG B X,TIAN Z H,et al. Evaluating network security and optimal active defense based on attack defense game model[J]. Chinese Journal of Computers,2009,32(4):817-827.

[15] LIANG X,XIAO Y. Game theory for network security [J]. IEEE Communications Surveys & Tutorials,2012,15(1):472-486.

[16] JIE B,YANG Y,ZHONG Z W. A virtual computing resource allocation modal of cloud computing based on evolutionary game theory[J]. Journal of Beijing Jiaotong University,2013,37(5):75-79.

[17] MENG C J,ZHANG F Q,LUO Y F. Nash equilibrium algorithm of cloud resource allocation under game model[J]. Hydro-mechatronics Engineering,2018,46(6):184-192.

[18] XU F,WANG S C,YANG W X. Cloud Resource Scheduling Algorithm Based on Game Theory[J]. Computer Science,2019,46(S1):295-299.

[19] HARDIN G. The tragedy of the commons[J]. Science,1968,162(3859):1243-1248.

[20] OSTROM E. Governing the commons[M]. Cambridge University Press,2015.



ZHAI Yong, born in 1969, master, professor. His main research interests include design and integration of spatial data storage system and network security.



LIU Jin, born in 1989. Her main research interests include the design and maintenance of spatial databases and geographic information systems.