



面向多租户数据中心资源回收利用的能效激励机制

周知¹, 刘方明^{2*}

1. 中山大学数据科学与计算机学院, 广州 510006

2. 华中科技大学计算机科学与技术学院, 武汉 430074

* 通信作者. E-mail: fmliu@hust.edu.cn

收稿日期: 2019-05-13; 接受日期: 2019-08-13; 网络出版日期: 2021-04-23

国家重点研发计划 (批准号: 2017YFB1001703)、国家自然科学基金 (批准号: 61722206, 61761136014, 61520106005, 61802449) 和国家高层次人才特殊支持计划资助项目

摘要 多租户公有云是云计算的一种重要组成形式. 近年来, 多租户公有云数据中心在如火如荼发展的同时, 其过低资源使用率所导致的巨大能耗浪费也引起了社会的高度关注. 然而, 由于多租户公有云的资源通常以“包年包月”的租赁形式预留给租户, 云服务提供商无法通过关闭空闲服务器等常用能耗管理方法来降低数据中心能耗. 针对这一难题, 本文提出了面向多租户数据中心资源回收利用的激励机制来提升多租户数据中心能效. 该机制以经济激励的形式来回收利用租户空闲预留型资源, 即云服务提供商向租户支付合适的经济补偿, 从而回收租户空闲预留型资源, 并充分利用其运行其他云计算服务 (如电商和搜索服务). 基于 Lyapunov 方法, 本文设计并分析了一种面向租户预留资源回收利用的动态定价算法, 该算法无需预测系统未来信息便可作出在线定价决策, 并且在保障系统稳定性的同时使得云服务提供商长期成本无限接近于理论最低值. 通过严格的数学证明以及由真实数据驱动的仿真实验, 本文验证了所提出的激励机制在回收利用租户空闲资源方面的有效性.

关键词 多租户数据中心, 能耗管理, 资源回收, 激励机制, 在线算法

1 引言

当前, 云计算产业已在全球范围内高度普及, 日益繁荣. 作为新一代信息技术的重要代表, 云计算技术正深刻改变着信息产业的发展格局, 推动着社会生产生活方式发生革命性的巨变. 稳定可靠的云计算服务为广大企业和个人用户提供了日益丰富的高性能网络应用, 包括在线办公、智能助理、在线支付、社交网络、在线存储, 以及虚拟/增强现实等, 它们极大地改善了人们生活质量, 丰富了社会生产方式, 提升了社会生产效率. 在新一代网络通信技术使得云计算服务触手可得、深度学习技术助力大数据价值不断凸显的时代背景下, 未来的云计算产业必将呈现出持续爆发之势.

引用格式: 周知, 刘方明. 面向多租户数据中心资源回收利用的能效激励机制. 中国科学: 信息科学, 2021, 51: 735-749, doi: 10.1360/SSI-2019-0098
Zhou Z, Liu F M. An energy-efficient incentive mechanism for resource recycling in multi-tenant datacenters (in Chinese). Sci Sin Inform, 2021, 51: 735-749, doi: 10.1360/SSI-2019-0098

然而, 在云计算高速发展的同时, 其能耗和碳排放问题也变得日趋严峻, 成为制约其可持续发展的重大问题. 咨询公司 Gartner 的研究数据显示, 2012 年全球数据中心耗电量占全球总能耗的 1.3%, 达到 7203 亿千瓦时. 同样据估测, 到 2020 年, 全球数据中心的耗电量将占全球总能耗的 8%, 攀升至 20000 亿千瓦时, 这将超过了法国、西班牙、德国和巴西这 4 个工业大国的总能耗^[1]. 而同年, 全球数据中心的碳排放将高达 14 亿吨, 占全球总碳排放 2.6%, 并超过德国这一工业强国的全年碳排放^[2]. 在中国, 2011 年全国数据中心总耗电量就已经达到 700 亿千瓦时, 相当于整个天津市的用电量^[3]. 而 2015 年, 全国数据中心用电量已高达 1000 亿千瓦时, 占到全社会用电量的 1.5%, 超过了 2014 年整个三峡水电站一年的发电量 988 亿千瓦时^[4].

多租户公有云 (multi-tenant public cloud)^[5] 是云计算的一种重要组成形式, 其所提供的“基础设施即服务” (infrastructure-as-a-service, IaaS) 旨在让大规模的租户 (tenants, 包括企业、政府和众多个人用户等) 通过先进的虚拟化技术来高效共享大型数据中心中数以万计的计算能力、存储空间和网络资源, 从而实现“规模经济效益”. 目前, 国内外主要的 IT 企业, 例如国外的 Google, Amazon, Microsoft 和 IBM, 以及国内的华为、阿里巴巴、腾讯、百度、中国电信、奇虎 360 等, 均投入大量资本参与到多租户公有云市场并向海量用户提供 IaaS 云服务. 近年来, 多租户公有云市场在国内外发展迅猛, 规模与日俱增. 据 Gartner 报道, 2017 年 IaaS 市场全球营收达到 347 亿美元, 相比于 2016 年增长 36.6%. 在中国, 2017 年上半年 IaaS 云计算整体规模超过 10 亿美元, 相比 2016 年同期增长近 70%. 其中, 阿里巴巴旗下的 IaaS 云服务阿里云增长率则高达 127%^[1].

然而, 在多租户公有云数据中心如火如荼发展的同时, 其过低资源使用率所导致的巨大能耗浪费同样不容忽视^[6]. 对多租户公有云数据中心而言, 其资源通常以按需型和预留型这两种主流模式租赁给用户 (即租户). 其中, 按需型资源采取“按量付费”的定价模式, 即租户根据当前时段 (通常为一小时) 实际使用的资源数量付费. 预留型资源则采取“包年包月型”付费模式, 即租户预先一次性支付较长时间段 (通常为数月至数年) 的租赁费用, 从而获取该时间段资源的对应使用权. 受经济学中规模经济效应²⁾的影响, 预留型资源的平均价格通常远低于同等配置的预留型资源, 例如阿里云通用型计算实例 ecs.g5.large 五年期预留的平均价格仅为按需型的 $\frac{1}{5}$. 因此, 为了降低成本, 企业级租户通常租用大量平均价格更低的预留型资源来支撑其业务. 然而, 由于租户业务的实际资源需求量通常剧烈波动, 因此预留型实例的平均资源利用率较为低下, 这也导致了 IaaS 公有云数据中心整体资源利用率低下^[7,8]. 进一步, 物理资源的长期空转又将造成巨大的能耗浪费与碳排放^[9].

对单租户数据中心而言, 为了提升资源利用率从而改善服务器空转耗电, 最为直接有效的方法为将负载整合到少数服务器运行, 并关闭空闲服务器 (或调整到低功耗的睡眠状态)^[10~12]. 然而, 对本文所研究的多租户 IaaS 数据中心而言, 上述方法并不直接适用. 这是因为, 在租户一次性支付预留型资源的费用后, 云服务提供商必须在接下来的租赁期内保证预留给租户的资源的高可用性. 否则一旦发生服务不可用事件, 按照租户与云服务提供商之间签订的服务等级协议 (service level agreement, SLA), 后者需要向租户赔偿由于资源不可用所带来的高额损失. 由此可见, 这一付费模式使得即使预留给租户的资源利用率过低, 服务提供商也不能关闭空闲服务器 (或调整到低功耗的睡眠状态). 此外, 在一次性付费的模式下, 租户和云计算服务提供商均无关闭空闲服务器从而达到节能目的的动机. 首先, 对租户而言, 关闭空闲服务器无法降低成本, 但可能带来损害服务质量的风险. 其次, 对于云服务提供商而言, 由于预留型资源的费用已一次性收取, 因此资源利用率过低也不会为自身带来经济损失.

1) Gartner forecasts worldwide public cloud services revenue to reach \$260 Billion in 2017. <https://www.gartner.com/en/newsroom/press-releases/>.

2) Economies of scale. https://en.wikipedia.org/wiki/Economies_of_scale.

显然,若要通过关闭空闲服务器来达到降低多租户公有云数据中心能耗的目的,则亟需高效的经济激励机制来强化租户与云服务提供商关闭空闲服务器的动机。

针对上述租户与云服务提供商双方节能减排动机不足这一难题,本文将提出一种面向多租户公有云数据中心资源回收利用的能效激励机制。借助经济补偿的方式,该机制同时激励租户共享和云服务提供商回收利用空闲预留资源,从而降低租户和云服务提供商双方运营成本,达到双赢的目的。具体而言,现实中大量云计算服务提供商除提供 IaaS 服务外,通常还提供其他的云计算服务,例如阿里巴巴和 Amazon 还提供电商服务,而 Google 和 Microsoft 均提供搜索服务。那么,如果能够将部分这类非 IaaS 服务迁移到租户的空闲预留资源上运行,云服务提供商便可以关闭部分原本用于服务这类非 IaaS 服务的服务器(下称后备资源),从而达到降低数据中心能耗的目的。为了促成上述空闲预留资源回收利用的局面,云服务提供商需要向租户提供一定的经济补偿来降低租户成本,以激励租户产生向云服务提供商“出租”空闲预留资源的动机。同时,对于云服务提供商而言,同样可以通过回收利用租户空闲资源来达到节约成本的目的。

在实际实施过程中,上述能效激励机制所面临的关键挑战在于如何制定合适的资源回收价格来同时激励租户和云计算服务提供商。若价格过低,则难以回收足够数量的空闲预留资源。反之若价格过高,则会加大云服务提供商所支付的资源回收成本,不利于提升其回收利用租户空闲资源的积极性。此外,若进一步考虑数据中心的动态性(如非 IaaS 服务的资源需求和后备资源价格的动态性),那么资源定价问题的复杂度进一步加剧,即云服务提供商需要根据系统动态性来实时调整资源回收价格,以便最小化自身成本。针对上述挑战,本文基于 Lyapunov 优化方法设计并分析了一种面向高效能多租户公有云数据中心的动态资源回收定价机制。该机制通过动态优化资源回收价格和后备资源使用量来自适应地平衡优化云服务提供商长期成本和系统稳定性这两大目标。该机制的优点在于,其无需预测系统未来信息便可作出在线优化决策,并且在保障系统稳定性的同时使得云服务提供商长期成本无限接近于理论最低值。最后,本文通过严格的数学证明以及由真实的电价和数据中心负载历史数据驱动的仿真实验,验证了所提出的动态资源回收定价机制在回收利用租户空闲资源方面的有效性。

2 研究背景与动机

本文考虑某云计算服务提供商向多用户提供基础设施即服务(IaaS)云计算服务。在这种 IaaS 云计算服务模式,云计算服务提供商将把客户所需的硬件资源(CPU、内存、网络、存储等)打包封装,并以“实例(instance)”(亦称“云主机”,这里云主机可以是一台独立物理主机,但更多的情况是虚拟机)的形式提供给用户。由于用户本身对计算实例不具有所有权,只是通过向云计算服务提供商付费租赁实例来获取使用权,因此这里的用户通常被称为租户,而提供这种 IaaS 服务的数据中心则被称为多租户公有云数据中心。

目前,多租户数据中心的资源租赁方式主要有两种:预留型和按需型。其中,预留型也称为“包年包月型”,顾名思义,租户预先向云计算服务提供商一次性支付较长时间段(通常为数月至数年)的租赁费用。在接下来的租赁期内,无论租户使用与否,云计算服务提供商都将为该租户预留和维护对应的资源实例。对应地,按需型(on-demand)也称为按量付费型,即租户根据当前实际租用的实例数量付费即可。相比于预留型长达数月甚至数年的付费周期,按需型的付费周期通常不超过一小时,甚至可以达到秒级。这种更细粒度的付费机制使得按需型实例具有更大的灵活性:用户无需按照峰值需求支付高昂的一次性费用,只需根据需求的变化动态调整配置的实例数量即可。然而,针对计算需求较大的企业用户而言,完全依靠灵活性更好的按需型实例并非最优选择。这是因为,就单位时间租赁成

本而言, 按需型实例远高于预留型实例. 以阿里云为例³⁾, 其通用型计算实例 ecs.g5.large 的按需价为 0.66 元/小时, 包月预留的平均价格为 0.25 元/小时, 包年预留的平均价格为 0.20 元/小时, 3 年预留的平均价格为 0.12 元/小时, 5 年预留的平均价格则低至 0.08 元/小时, 与按需型实例相差 7.25 倍. 因此, 为了达到节约成本的目的, 企业级租户通常购买大量的预留型实例来满足基本计算需求, 而在需求高峰时期则自适应地配置按需型实例来满足额外的计算需求.

然而, 预留型实例的大量应用带来了新的能耗问题. 具体而言, 由于租户的负载通常动态变化, 而预留型实例的数量则固定不变, 因此, 预留型实例的资源利用率波动较大, 导致整体利用率低下, 能耗浪费严重. 第三方机构评估表明, Amazon EC2 云平台的资源利用率仅为 3%~17%^[8]. 同时 Gartner 的调研也表明租户平均超买的资源比例高达 55%, 由此导致的公有云成本浪费也高达 100 亿美金⁴⁾. 针对过低资源利用率所导致的成本和能耗浪费问题, 在私有云、SaaS (software-as-a-service) 公有云数据中心或 PaaS (platform-as-a-service) 公有云数据中心中, 云计算服务提供商可以将计算负载整合到部分服务器, 并关闭空闲服务器从而消除服务器空转能耗 (idle power) 来达到节能减排的目的^[10~12]. 然而, 这种关闭空闲服务器的方法并不能直接应用于多租户 IaaS 数据中心. 这是因为, 当租户租赁预留型实例后, 服务提供商必须根据服务协议 SLA 保证实例在租赁期内的高可用性.

考虑到云计算服务提供商除提供 IaaS 服务外, 通常还提供其他的云计算服务, 例如阿里巴巴和 Amazon 还提供电商服务, 而 Google 和 Microsoft 均提供搜索服务. 那么, 面向多租户数据中心资源利用率提升与节能减排的一种有效方法为, 将云计算服务提供商的其他非 IaaS 计算服务迁移到租户的空闲实例上运行. 这种资源共享方法使得云计算服务提供商可以关闭部分原本用于运行非 IaaS 计算服务的服务器, 从而降低数据中心整体能耗. 值得注意的是, 这种资源共享方法在实践中面临着如何激励租户共享空闲预留资源的问题. 具体而言, 租户本身没有动机将已付费的空闲实例共享, 除非给予其一定的经济激励 (即经济补贴). 经济激励作为激励租户共享空闲资源从而促成云计算服务提供商和租户达成双赢的局面的关键措施, 其一方面能降低用户成本, 提升用户参与资源共享的积极性, 另一方面能降低数据中心能耗成本. 显然, 如何制定最优的经济激励力度直接影响着云计算服务提供商的效用: 若激励力度不足, 将无法提升租户参与程度和降低能耗成本; 若激励力度过大, 将大大增加服务提供商需要向租户支付的激励成本. 针对这一难题, 本文接下来将从云计算服务提供商的角度出发, 研究如何动态优化资源回收价格, 从而在激励租户参与资源共享的同时最小化云计算服务提供商的长期成本.

3 面向资源回收利用的定价模型与问题建模

3.1 系统概览

为了更好地刻画数据中心的动态性 (例如动态变化的租户资源利用率以及服务提供商非 IaaS 服务的动态负载), 本文所提出的资源回收定价机制以离散的时间片 (slot) 方式运行, 其中每个长度固定时间片对应一个定价周期. 在每个时间片的开头, 数据中心环境变量如租户的资源利用率和服务提供商非 IaaS 服务负载发生改变, 因此服务提供商动态更新预留资源回收价格从而平衡可回收预留资源的供给与需求. 而在每个时间片内, 资源回收价格保持不变. 在典型的数据中心场景中, 一个长度固定的时间片长度通常对应数分钟到数小时^[11~14]. 例如, 对阿里云所提供的按需型实例而言, 其定价周期

3) 阿里云定价. <https://www.aliyun.com/price/>.

4) How and why \$10 billion in public cloud costs are unwarranted losses. <https://weekly-geekly.github.io/articles/359338/index.html>.

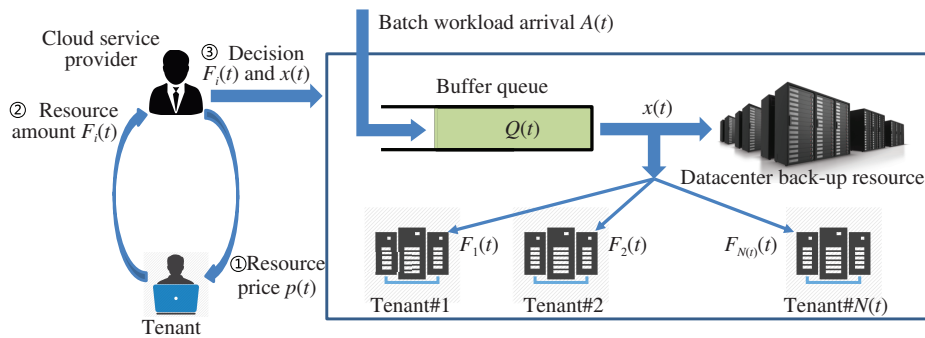


图 1 (网络版彩图) 本文所提出的面向多租户数据中心资源回收利用的能效激励机制示意图

Figure 1 (Color online) An illustration of the proposed incentive mechanism for resource recycling in multi-tenant datacenters

为 1 小时 (即一个时间片长度为 1 小时)。

数据中心处理的作业负载通常可以分为两类. 其中一类为延迟敏感型交互式负载, 这类负载通常由前端应用用户产生, 具有高度交互性, 因此要求能够快速返回计算结果. 常见的交互式负载包含网络搜索、即时通讯和在线购物等^[11,12]. 除了交互式负载, 数据中心还运行着大量的批处理负载, 批处理负载通常由数据中心后台产生, 由于计算结果不被即刻使用, 因此批处理负载通常具有延迟容忍性. 常见的批处理负载包含机器学习训练、数据挖掘、统计分析和网页排序等^[13,14]. 实证研究表明, 批处理负载在数据中心总负载中占据的比例高达 70%^[14]. 批处理负载对延迟的容忍性使得其对可用性波动不稳定的回收资源具有更好的适应性, 即若当前时间片的回收的资源数量不足以覆盖批处理负载的资源需求量, 服务提供商可将批处理负载延迟处理. 考虑到批处理负载的延迟容忍性这一良好特性, 本文接下来的资源定价模型将专注于优化利用回收资源服务数据中心批处理负载.

本文所提出的面向多租户公有云数据中心资源回收利用的能效激励机制如图 1 所示. 在每个时间片 t , 数据中心新产生的批处理负载对资源的需求量为 $A(t)$ (单位为实例个数), 这些新产生的批处理负载首先被存储在缓冲队列中, 以便等待资源价格降低后被调度处理. 考虑到随着时间的推移, 租户会动态进入或离开系统, 因此本文假设 t 时刻系统中的租用预留型实例的租户数量为 $N(t)$, 并使用集合 $\mathcal{N}(t) = \{1, 2, \dots, N(t)\}$ 来表示这 $N(t)$ 个租户. 当云服务提供商公布资源回收价格 $p(t)$ (即回收一单位预留型资源的价格为 $p(t)$) 后, 每个租户根据 $p(t)$ 的大小来决策被回收的资源数量, 本文用 $F_i(t)$ 来表示 t 时刻第 $i \in \mathcal{N}(t)$ 个用户提供的回收资源数量. 除了利用回收的预留型资源处理缓冲队列中的批处理负载外, 云服务提供商还可以使用未对外租用的后备资源来服务批处理负载, 本文接下来使用 $x(t)$ 来表示第 t 个时刻用于服务批处理负载的后备资源数量.

3.2 租户对资源回收价格的响应模型

在每个时刻 t , 租户 $i \in \mathcal{N}(t)$ 所愿意提供的回收资源数量 $F_i(t)$ 与资源回收价格 $p(t)$ 相关. 资源回收价格 $p(t)$ 越高, 租户所愿意提供的回收资源数量 $F_i(t)$ 越大, 反之则 $F_i(t)$ 越小. 同时值得注意的是, 虽然资源回收能够降低租户的运营成本, 但是资源回收同样可能造成服务质量下降. 具体而言, 若资源回收后租户的任务请求剧烈增加, 那么则有可能导致服务延迟增加甚至部分任务请求无法被服务的弊端. 因此, 在决策回收资源数量 $F_i(t)$ 时, 租户会权衡考虑资源回收所带来的成本降低机会与服务质量下降风险. 不失一般性, 在给定资源回收价格 $p(t)$ 后, 租户 i 对该价格的响应, 即所愿意提供的回收资源数量 $F_i(t)$ 可以表示为 $p(t)$ 的函数, 亦即 $F_i(t) = F_i(p(t), t)$. 不难发现, 为了通过调整 $p(t)$ 从而平

衡资源回收的供给与需求并降低资源回收总成本, 云计算服务提供商需要对租户响应函数 $F_i(p(t), t)$ 进行精确建模. 在实际数据中心系统中, 云服务提供商可以通过机器学习等方法对租户响应函数进行建模, 即通过调整价格 $p(t)$ 来学习响应函数 $F_i(p(t), t)$. 具体而言, 云服务提供商设置不同的价格 $p(t)$ 并记录对应的资源供应数量 $F_i(t)$, 然后使用非线性最小二乘法拟合或支持向量机等方法来学习函数 $F_i(p(t), t)$ 中的关键参数^[15]. 由于篇幅限制, 本文不再介绍获取租户响应函数 $F_i(p(t), t)$ 的技术细节, 假设 $F_i(p(t), t)$ 已知. 值得注意的是, 本文不对 $F_i(p(t), t)$ 的具体形式作任何假设, 后文所提出的方法同样适用于非线性甚至离散函数 $F_i(p(t), t)$ (例如若资源回收的单位为虚拟机个数, 那么 $F_i(p(t), t)$ 则为离散函数).

3.3 云服务提供商成本模型

对云服务提供商而言, 其期望通过动态优化资源回收价格 $p(t)$ 来最小化服务批处理负载的长期资源成本. 服务批处理负载的资源成本包含两部分: 租户资源回收成本和数据中心后备资源使用成本. 其中, 租户资源回收成本即为回收租户预留型实例时支付给租户的经济激励. 在每个时刻 t , 给定第 $i \in \mathcal{N}(t)$ 个租户的资源回收量 $F_i(p(t), t)$, 云服务提供商需要支付给当前时刻 $\mathcal{N}(t)$ 个租户的总的激励成本可以表示为 $p(t) \sum_{i \in \mathcal{N}(t)} F_i(p(t), t)$. 后备资源使用成本即为使用数据中心后备资源服务缓冲队列中的批处理负载所引发的运营成本, 在实际数据中心中, 这部分成本主要为后备服务器运行所引发的数据中心能耗成本^[16]. 若给定 t 时刻的后备资源使用量 $x(t)$ 和单位后备资源价格 $R(t)$, 那么该时刻后备资源使用总成本为 $R(t)x(t)$. 不难发现, 资源价格 $R(t)$ 主要由当前电网电价决定, 在主流数据中心广泛参与的实时电力市场上, 实时电价通常随时间动态变化. 综合考虑上述两方面成本, 每个时刻 t 处理批处理负载的总成本可以表示为

$$\text{Cost}(t) = \sum_{i \in \mathcal{N}(t)} p(t)F_i(p(t), t) + R(t)x(t).$$

进一步考虑数据中心的整个运行周期, 则云服务提供商的优化目标可以表示为下式所示的长期平均成本:

$$\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} E\{\text{Cost}(t)\}.$$

3.4 批处理负载缓冲队列稳定性模型

云计算服务提供商在优化资源回收价格 $p(t)$ 与后备资源使用量 $x(t)$ 来最小化长期平均成本的同时, 还必须保障存储批处理负载的缓冲队列的稳定性, 即避免缓冲队列中存储的批处理负载无限度的增长, 从而导致部分负载无法在有限时间内被服务^[17]. 值得注意的是, 队列稳定性与资源成本这两个目标之间存在折中关系: 降低资源成本要求云服务提供商降低资源回收价格或者延迟批处理负载的处理, 然而这会损害队列稳定性; 反之, 提高队列稳定性要求缓存的批处理负载能够被快速处理, 显然, 这将导致资源回收价格或资源成本上升.

为了保障缓冲队列稳定性, 本文接下来建立缓冲队列稳定性的数学模型. 本文使用 $Q(t)$ 来表示在 t 时刻开始时缓冲队列的长度, 即缓冲队列中缓存的批处理负载对计算资源的需求总量. 不失一般性, 本文假设队列初始长度为 0, 即 $Q(0) = 0$. 综合考虑 t 时刻数据中心新产生的批处理负载数量 $A(t)$ 和被服务的负载数量 $\sum_{i \in \mathcal{N}(t)} F_i(p(t), t) + x(t)$, 那么队列长度 $Q(t)$ 的动态变化关系可以表示为

$$Q(t+1) = Q(t) - \sum_{i \in \mathcal{N}(t)} F_i(p(t), t) - x(t) + A(t).$$

为了避免出现资源过度分配或缓冲队列长度为负数的情况, 本文进一步限制在每个时间片 t 为批处理负载所分配的资源总量不得超过当前缓冲队列长度 $Q(t)$:

$$\sum_{i \in \mathcal{N}(t)} F_i(p(t), t) + x(t) \leq Q(t).$$

基于上述队列长度动态演化关系, 本文进一步定义队列长期平均长度为

$$\bar{Q} = \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} E\{Q(t)\}.$$

进一步, 队列的稳定性即队列长期平均长度有界的要求可以建模为下式所示的约束条件^[17]:

$$\bar{Q} < \infty.$$

可以证明^[17], 若上式所示的稳定性约束条件被满足, 那么 $\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} E\{\sum_{i \in \mathcal{N}(t)} F_i(p(t), t) + x(t)\} \geq \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} E\{A(t)\}$, 即表明长远来看, 为批处理负载分配的平均资源数量不低于其所需求的平均资源数量.

3.5 资源回收定价问题建模

基于上述云服务提供商成本模型与队列稳定性模型, 本文所研究的优化问题可以描述为: 如何动态优化资源回收价格 $p(t)$ 与后备资源使用数量 $x(t)$ 来最小化数据中心服务批处理负载的长期平均成本, 同时保障缓冲队列的稳定性. 数学上, 上述优化问题可以形式化表达如下:

$$\min \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} E\{\text{Cost}(t)\} \quad (1)$$

$$\text{s.t. } \bar{Q} < \infty, \quad (2)$$

$$\sum_{i \in \mathcal{N}(t)} F_i(p(t), t) + x(t) \leq Q(t), \quad (3)$$

$$0 \leq x(t) \leq C(t), \quad (4)$$

$$p(t) \geq 0,$$

其中, 式 (2) 为队列稳定性约束, 即确保队列长期平均长度有界. 式 (3) 为批处理负载可用性约束, 即确保被服务的批处理负载不超过当前队列长度. 式 (4) 为后备资源容量约束, 其中 $C(t)$ 为 t 时刻数据中心可用后备资源的总数量.

不难发现, 在上述模型中, 队列稳定性约束 (2) 将不同时刻的决策变量耦合在一起, 从而使得当前时刻 t 的决策直接影响未来决策的最优性, 这类问题在数学上被称之为随机优化问题^[17]. 在理想情况即式中参数如未来负载资源需求量 $A(t)$ 与后备资源价格 $R(t)$ 已知的情况下, 上述优化问题可以基于现有的优化工具 (如 CVX) 准确求解. 然而, 在实际数据中心系统中, 系统参数如负载资源需求量 $A(t)$ 和后备资源价格 $R(t)$ 高度动态化且难以预测. 那么, 在未来参数未知的情况下, 如何才能确保当前的优化决策既能够最小化长期平均成本, 同时又能够保障队列稳定性呢?

4 资源回收定价的在线优化算法

针对前述随机优化问题中未来参数如负载数量和后备资源价格无法精确预测这一难题, 本节借助 Lyapunov 优化方法^[17] 来设计并分析一种资源定价在线优化算法. 其基本思想为将上述原始随机优

化问题中的长期优化目标函数 (1) 和约束条件 (2) 解耦成一系列可以在每个时间片独立求解、无需未来信息的实时优化问题. 该方法的优点在于, 其能够在无需预测未来参数的条件下动态优化决策变量, 使得长期平均成本能够无限接近理论最优成本, 同时保障缓冲队列稳定性.

4.1 基于 Lyapunov 优化方法的问题转化

Lyapunov 优化方法为控制论中解决队列稳定性控制问题的经典方法, 遵循其基本原理, 本文首先定义二次 Lyapunov 函数:

$$L(Q(t)) = \frac{1}{2}Q^2(t).$$

上述 Lyapunov 函数 $L(Q(t))$ 是对缓冲队列 $Q(t)$ 拥塞程度的一种度量. 为了确保队列稳定性, 所设计的优化算法应该使得 $L(Q(t))$ 朝着低拥塞程度的方向发展. 为了达到这一目的, 进一步定义如下所示的单步条件 Lyapunov 漂移 (one-step conditional Lyapunov drift):

$$\Delta(Q(t)) = E\{L(Q(t+1)) - L(Q(t))|Q(t)\}.$$

Lyapunov 漂移 $\Delta(Q(t))$ 度量了在两个连续的时间片之间 Lyapunov 函数 $L(Q(t))$ 的变化量. 不难发现, 通过在每个时间片 t 最小化 Lyapunov 漂移 $\Delta(Q(t))$, 可以达到阻止缓冲队列的长度无限制增长, 从而维护队列稳定性的目的.

至此, 通过引入 Lyapunov 优化方法, 原始的最小化长期平均成本同时保障队列稳定性的优化问题可以转化为同时最小化期望成本 $E\{\text{Cost}(t)|Q(t)\}$ 与 Lyapunov 漂移 $\Delta(Q(t))$ 这样的双重目标. 为了平衡优化这两个目标, 本文选择最小化如下二者加权之和:

$$\Delta(Q(t)) + V E\{\text{Cost}(t)|Q(t)\}, \quad (5)$$

其中, 控制参数 ($V \geq 0$) 为平衡优化长期平均成本与缓冲队列稳定性的关键之所在, 其代表了最小化长期平均成本与最小化 Lyapunov 漂移即保持队列稳定性权重的比较. 即 V 越大, 上述优化问题越强调最小化当前期望成本, 从而使得长期平均成本越小, 但缓冲队列的稳定性越差. 反之, V 越小, 缓冲队列稳定性越强, 但长期平均成本越高. 通过调整控制参数 V 的取值, 云计算服务提供商在不同的成本与队列稳定性平衡优化点间做出灵活选择. 此外, 进一步观察上式不难发现, 该转化后的优化问题仅含有当前时间片 t 相关的参数, 而不包含不确定的未来参数. 这使得云计算服务提供商可以基于当前系统状态信息动态决策资源回收价格与后备资源使用量, 而无需参考未来信息.

4.2 在线优化算法的设计

通过引入 Lyapunov 优化方法, 本文将原始的含有未来不确定信息的随机优化问题 (1) 转化为一些可在线求解、无需未来信息的实时优化子问题 (5). 然而, 由于涉及到非显式的难以处理的 $\max\{x, 0\}$ 表达式, 问题 (5) 难以直接求解. 因此, 本文选择最小化问题 (5) 的一个上确界, 后文严格的理论分析与证明表明, 这种近似优化方法所带来的误差可控并且可以无限接近于 0.

为了推导问题 (5) 的下确界, 如下所示的引理 1 首先给出 Lyapunov 漂移 $\Delta(Q(t))$ 的一个上确界.

引理1 在任意时间片 t , 对于任意队列长度 $Q(t)$, 该时间片对应的 Lyapunov 漂移 $\Delta(Q(t))$ 满足下列不等式:

$$\Delta(Q(t)) \leq B - Q(t) E \left\{ \sum_{i \in \mathcal{N}(t)} F_i(p(t), t) + x(t) - A(t) | Q(t) \right\}, \quad (6)$$

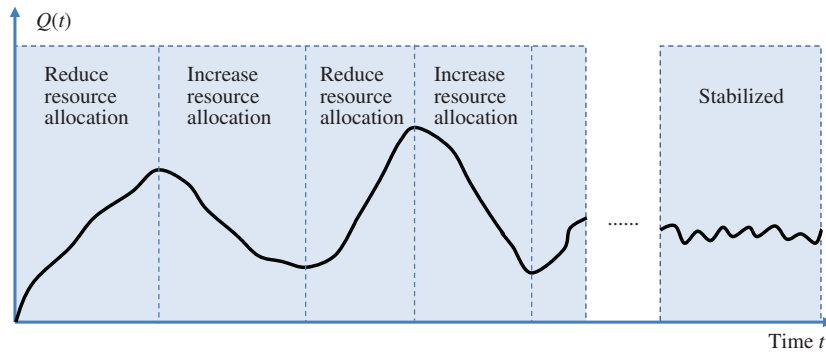


图 2 (网络版彩图) 本文所提出在线算法基于当前队列长度 $Q(t)$ 动态自适应调整资源分配量的示意图

Figure 2 (Color online) An illustration of the proposed online algorithm adaptively adjusting resource allocation based on the queue backlog $Q(t)$

其中, $B \triangleq \frac{1}{2}(M_{\max}^2 + C_{\max}^2 + A_{\max}^2)$ 为一给定常量参数, 其中 $A_{\max} = \max_t A(t)$ 批处理负载最大资源需求量, $C_{\max} = \max_t C(t)$ 为数据中心后备资源最大数量, $M_{\max} = \max_t \sum_{i \in \mathcal{N}(t)} F_i(p(t), t)$ 为所有租户所能提供的回收资源之和的上限.

受篇幅限制, 笔者通过在线技术报告^[18]给出了上述引理 1 的证明过程. 基于引理 1, 在不等式 (6) 两边同时加上加权成本 $VE\{\text{Cost}(t)|Q(t)\}$, 即可得到每个时间片 t 对应的成本与 Lyapunov 漂移的一个有限上确界, 如下式所示:

$$\Delta(Q(t)) + VE\{\text{Cost}(t)|Q(t)\} \leq B + VE\{\text{Cost}(t)|Q(t)\} - Q(t)E \left\{ \sum_{i \in \mathcal{N}(t)} F_i(p(t), t) + x(t) - A(t) | Q(t) \right\}. \quad (7)$$

最小化上述不等式 (7) 右边所示的上确界等价于忽略其中的常数项 B 和 $Q(t)A(t)$, 进而在每个时间片最小化如下式所示的目标函数:

$$V \times \text{Cost}(t) - Q(t) \left[\sum_{i \in \mathcal{N}(t)} F_i(p(t), t) + x(t) \right]. \quad (8)$$

观察上述目标函数结构, 不难发现其蕴含着丰富的经济学含义. 具体而言, $Q(t)$ 为当前时间片队列稳定性的度量, 而 $F_i(p(t), t) + x(t)$ 则为当前时间片分配给批处理负载的总的资源数量. 显然, 从维持队列稳定性的角度而言, $F_i(p(t), t) + x(t)$ 期望被最大化从而提升队列稳定性, 而当前队列长度 $Q(t)$ 则体现了系统对提升系统稳定性的迫切程度. 从这个视角来看, $Q(t)[\sum_{i \in \mathcal{N}(t)} F_i(p(t), t) + x(t)]$ 可以看成是为批处理负载分配资源所带来的收益或效用, 上述目标函数 (8) 可以理解为资源分配的净成本 (即成本与收益之差), $Q(t)$ 则可以理解为当前时刻单位资源分配所带来的收益. 显然, 若 $Q(t)$ 较小, 即资源分配带来的收益较低 (同时也意味着队列稳定性较佳), 那么此时成本项 $V \times \text{Cost}(t)$ 在上述目标函数 (8) 中占据主导地位, 系统便减少资源分配量从而降低净成本. 资源分配量减少后, 队列长度 $Q(t)$ 会随着新的批处理负载的产生而增大, 使得队列稳定性变差, 资源分配带来的收益增加, 那么此时收益项 $Q(t)[\sum_{i \in \mathcal{N}(t)} F_i(p(t), t) + x(t)]$ 在上述目标函数 (8) 中占据主导地位, 系统进而增加资源分配量从而增加收益, 提升队列稳定性. 上述资源分配的动态自适应调整过程可以借助图 2 来直观反映. 不难发现, 队列长度 $Q(t)$ 正是动态自适应地平衡优化队列稳定性与长期成本的调节器.

进一步将 3.3 小节中的表达式 $\text{Cost}(t)$ 代入到上述目标函数 (8) 并考虑约束条件, 可以得到每个时间片 t 对应的如下所示的优化问题:

$$\begin{aligned} \min \quad & \sum_{i \in \mathcal{N}(t)} F_i(p(t), t)[Vp(t) - Q(t)] + x(t)[VR(t) - Q(t)], \\ \text{s.t.} \quad & \sum_{i \in \mathcal{N}(t)} F_i(p(t), t) + x(t) \leq Q(t), \\ & 0 \leq x(t) \leq C(t), \quad p(t) \geq 0. \end{aligned} \quad (9)$$

不难发现, 上述实时优化问题 (9) 的复杂度远低于原始随机优化问题 (1). 虽然上述实时优化问题 (9) 对决策变量 $p(t)$ 而言可能为非凸函数 (例如, 表达式 $F_i(p(t), t)p(t)$ 非凸), 但这种非凸性挑战在实际问题求解过程中相对容易解决. 这是因为: (i) 上述实时优化问题 (9) 仅含有两个变量 $p(t)$ 和 $x(t)$, 这表明该优化问题维度低, 复杂度不受租户数量影响; (ii) 上述实时优化问题 (9) 在每个时间片仅需求解一次, 即若单个时间片长度为 1 小时, 上述优化问题仅需每小时求解一次.

基于前述问题转化与化简, 本文接下来具体表述所提出的面向多租户公有云数据中心资源回收定价的在线优化算法, 如算法 1 所示.

Algorithm 1 Online pricing algorithm for resource recycling in multi-tenant datacenters

- 1: At the beginning of each time slot t , observing the current queue backlog $Q(t)$, backup resource price $R(t)$, backup resource capacity $C(t)$ and the response function $F_i(p(t), t)$ for each tenant i ;
- 2: Determine the optimal resource recycling $p(t)$ and backup resource usage $x(t)$, by solving the optimization problem (9);
- 3: At the end of each time slot t , observe the amount of newly arrived batch workload $A(t)$ and update the queue backlog $Q(t+1)$ according to the following queuing dynamics:

$$Q(t+1) = Q(t) - \sum_{i \in \mathcal{N}(t)} F_i(p(t), t) - x(t) + A(t).$$

4.3 算法性能分析

本文接下来定量分析并证明所提出的算法 1 在优化长期平均成本和维持缓冲队列稳定性等两方面的性能表现. 下述定理 1 表明, 所提出的在线优化算法 1 能够确保缓冲队列 $Q(t)$ 稳定, 并使得所求得的长期平均成本与理论最优值之差不超过 $\frac{B}{V}$.

定理1 假设批处理负载的产生速率 $A(t)$ 满足独立同分布 (i.i.d) 规律, 那么对于任意控制参数 $V > 0$, 上述在线优化算法 1 可以提供如下性能保障:

- (1) 所求得的长期平均成本与理论最小值之差有界且不超过 $\frac{B}{V}$, 即

$$\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}\{\text{Cost}(t)\} \leq \text{Cost}^{\text{opt}} + \frac{B}{V}. \quad (10)$$

- (2) 缓冲队列能够维持稳定性, 即缓冲队列长期平均长度有界,

$$\bar{Q} = \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}\{Q(t)\} \leq \frac{B + V \times \text{Cost}^{\text{opt}}}{\epsilon}. \quad (11)$$

这里 Cost^{opt} 为原始随机优化问题 (1) 的理论最优值, 即在离线情况下 (亦即获知未来所有信息精确值) 所能求得的最小长期平均成本. $\epsilon > 0$ 为一非非常数, 其代表着所提出在线算法对应的平均资源分配数量与某种静态控制算法所取得的实际平均资源分配数量之间的差距. B 为引理 1 中定义的常数.

定理 1 对所提出在线优化算法在优化长期平均成本与维护缓冲队列稳定性两方面的性能作出了定量分析. 可以发现, 随着控制参数 V 的增长, 所取得长期平均成本与理论最优值之间的差距越来越小, 且以 $O(1/V)$ 的速率递减. 这表明, 通过调节控制参数 V 使其取值足够大, 可以使得长期平均成本无限接近于理论最优值. 然而, 随着控制参数 V 的增长, 队列长期平均长度则越来越大, 呈 $O(V)$ 线性增长规律. 定理 1 所揭示的平均成本与平均队列稳定性之间的 $[O(1/V), O(V)]$ 折中关系, 定量地验证了前面中所定性分析的结论, 即控制参数 V 扮演着调节平均成本与平均队列稳定性之间折中关系的关键作用. 此外, 定理 1 同样表明所提出算法在最差情况下所取得的平均成本与平均队列长度均随着常数 B 的增长而变大. 从引理 1 中常数 B 的定义不难发现, B 随着批处理负载产生速率 $A(t)$ 的最大值 $A_{\max} = \max_t A(t)$ 单调递增. 这进一步表明, 若批处理负载产生速率 $A(t)$ 对应的数学分布越均匀 (即最大值与平均值越接近), 那么所提出的算法在优化平均成本与队列稳定性方面的效果越佳. 反之若 $A(t)$ 对应的数学分布越不均匀, 则所提出的算法效果越差. 受篇幅限制, 笔者通过在线技术报告 [18] 给出了上述定理 1 的证明过程.

5 仿真实验

本节将通过由真实数据中心批处理负载以及电价数据驱动的仿真实验来验证本文所提出的能效激励机制的有效性. 为了利用电价和负载的动态性, 同时避免频繁的服务器开关机操作, 本文参考现有的相关工作 [12, 13], 将每个定价周期即时间片的长度设置为 15 分钟. 此外, 现有公有云平台提供的按需型计算实例的计价周期通常处于 1 秒 (如 Google 云平台和 Amazon EC2 云平台) 至 1 小时 (如华为云和阿里云) 的范围内, 因此本文所设置的 15 分钟的定价周期恰好符合公有云平台按需型计算实例计价周期的整体平均水平.

5.1 实验设置

数据中心批处理负载. 为了模拟真实数据中心批处理负载对资源的动态需求, 本文从 Google 公布的一份集群利用率数据中抽取出为期一周的资源利用率数据⁵⁾. 该集群主要作业为 Hadoop 等大数据处理类批处理作业, 其归一化后的资源利用率如图 3(a) 所示. 容易发现, 图中所示的资源利用率呈现了较为明显地以日为周期的周期性规律. 实验将上述利用率数据等比例放大 1000 倍, 从而生成数据中心批处理负载的资源需求 $A(t)$ (量纲为单位资源).

后备资源价格. 考虑到数据中心的主要运营成本来自于电力消耗开支, 因此本文利用电价数据的波动性来模拟数据中心资源价格的动态性. 为达到这一目的, 本文从美国纽约州电力系统独立运营公司 (New York Independent System Operator, NYISO) 网站下载了 2018 年 10 月 10 日至 16 日纽约州为期一周的区域实时电价数据 (locational marginal price, LMP)⁶⁾, 如图 3(b) 所示. 容易发现, 电价随着时间的推移而剧烈波动. 后文实验将采取上述动态电价作为单位后备资源的使用价格 $R(t)$.

租户响应函数. 实验假设数据中心中参与预留型资源回收的租户数量 $N(t)$ 不变, 且 $N(t) = 10$. 参考现有的建模云计算用户对资源价格敏感程度的文献 [13], 实验假设每个租户对资源回收价格 $p(t)$

5) Borg cluster traces from Google. <https://github.com/google/cluster-data>.

6) NYISO pricing data. <https://www.nyiso.com/energy-market-operational-data>.

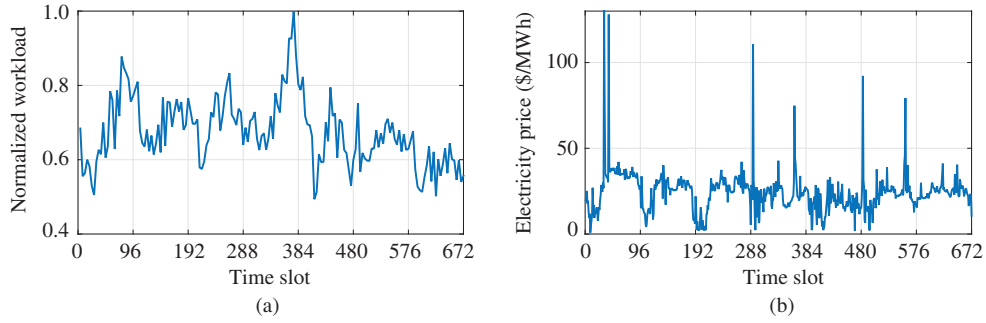


图 3 (网络版彩图) Google 服务器集群资源利用率历史数据 (a) 和纽约市实时电价历史数据 (b)

Figure 3 (Color online) The workload data of a Google cluster (a) and the real-time electricity price data of New York city (b)

的响应函数形式如下:

$$F_i(p(t)) = \max\{\alpha_i(t) \log(1 + p(t)), E_i(t)\}, \quad (12)$$

上式中 $E_i(t)$ 为租户 $I \in \mathcal{N}(t)$ 在 t 时刻最多所能提供的被回收资源数量. 此外, 式中 \log 形式的响应函数反映了租户对回收价格的敏感程度随价格上升而降低, 这一现象遵从经济学中的“边际效用递减”定律⁷⁾. 实验设置 $E_i(t)$ 为 200~500 之间符合均匀分布的随机数, 租户敏感性参数 $\alpha_i(t)$ 为 30~50 之间符合均匀分布的随机数.

比对实验设置. 为了论证所提出的动态定价机制 (实验中简称为 DPS) 的高效性, 本文将其与另外两种常用策略相比较: (1) 仅调度后备资源使用量的策略 BRO, 该策略仅利用后备资源来服务批处理负载. 基于 Lyapunov 方法, 该策略根据当前队列长度和后备资源价格决策后备资源使用量^[19]. (2) 仅动态定价的策略 DPO, 该策略忽略批处理负载的可延迟性, 每个时刻批处理负载产生后立即被全部服务. 该策略基于当前负载产生量和后备资源价格决策当前资源回收价格, 对未能被回收资源处理的批处理负载, 全部由后备资源在当前时间片处理^[20].

5.2 实验结果

所提出动态定价机制能够有效降低云服务提供商成本. 图 4 展示了不同策略下长期平均成本与 Lyapunov 控制参数 V 的关系. 可以发现: (1) 随着控制参数 V 的增长, 所提出动态定价机制 DPS 对应的长期平均成本显著下降, 但下降速度越来越慢并且逐渐收敛到一个最低水平. 这一现象定量地验证了定理 1 所表明的云服务提供商长期平均成本 $O(1/V)$ 的下降速率 (参考不等式 (10)). (2) 作为对比策略, 可以发现 BRO 和 DPO 这两种策略对应的长期平均成本均显著高于本文提出的动态定价机制 DPS, 这表明了本文所提出动态定价机制的高效性. (3) 在控制参数 V 足够大时, BRO 和 DPS 所对应的长期平均成本高度接近. 这是因为当控制参数 V 足够大时, 这两种策略均激进地延迟服务批处理负载, 从而有利于探索利用未来后备资源更低的价格.

所提出动态定价机制能够有效稳定缓冲队列. 图 5 展示了 BRO 和 DPS 这两种策略下缓冲队列平均长度 (即累积的资源需求数量) 与 Lyapunov 控制参数 V 的关系. 可以发现: (1) 随着控制参数 V 的增长, 队列长度次线性增长, 即随着 V 的增长, 队列长度增长速率减缓, 这使得缓冲队列的稳定性能被保持. (2) 进一步考虑图 4 所展示的长期平均成本随 V 的增加而降低的现象, 二者共同揭示了前文所讨论的云服务提供商长期平均成本与缓冲队列稳定性之间的折中关系. (3) 相比于仅利用后

7) Law of diminishing marginal utility. https://en.wikipedia.org/wiki/Marginal_utility#Diminishing_marginal_utility.

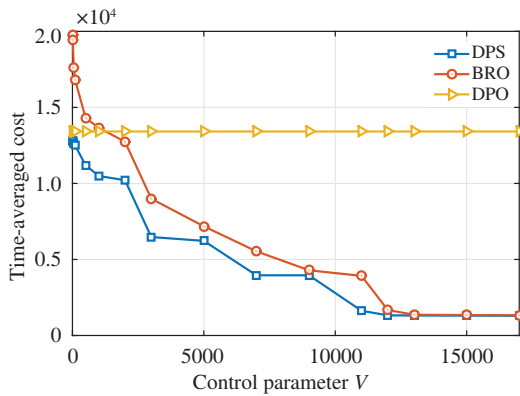


图 4 (网络版彩图) 长期平均成本 vs. 控制参数 V
Figure 4 (Color online) Long-term time-averaged cost vs. control parameter V

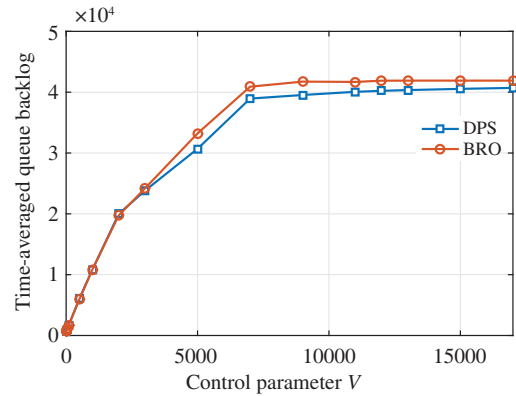


图 5 (网络版彩图) 长期队列平均长度 vs. 控制参数 V
Figure 5 (Color online) Long-term time-averaged queue backlog vs. control parameter V

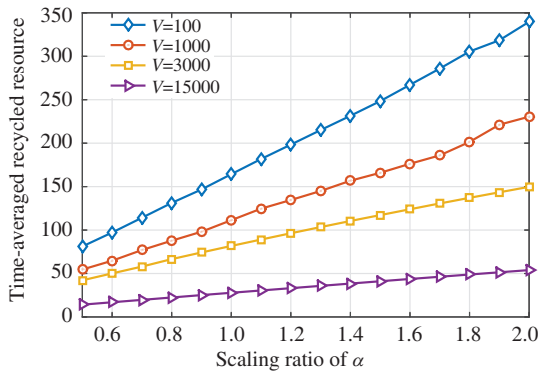


图 6 (网络版彩图) 长期平均资源回收数量 vs. 租户敏感性参数 α 缩放比例
Figure 6 (Color online) Long-term time-averaged recycled resource vs. scaling ratio of α

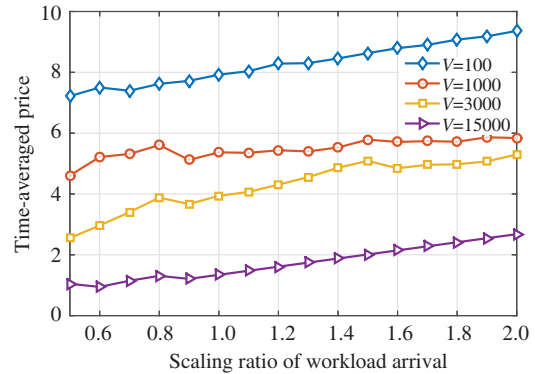


图 7 (网络版彩图) 长期平均资源回收价格 vs. 负载产生率 $A(t)$ 缩放比例
Figure 7 (Color online) Long-term time-averaged resource recycling price vs. scaling ratio of $A(t)$

备资源的 BRO 策略, 本文所提出的动态定价机制 DPS 不仅能够有效降低长期平均成本, 还能够降低长期平均队列长度从而增强队列稳定性.

租户敏感性参数对资源回收数量的影响. 实验继续探讨式中租户响应函数 (12) 中的敏感性参数 α_i 对云服务提供商所回收资源数量的影响. 为了达到这一目的, 实验对租户敏感性参数 α_i 放大或缩小一定的比例. 图 6 展示了在不同的控制参数 V 下, 云服务提供商所回收的资源数量与参数 α_i 缩放比例的关系. 可以发现, 随时缩放比例的增大 (亦即 α_i 的增大), 云服务提供商所回收的资源数量上升. 这表明, 租户对资源回收价格敏感性越高, 越有利于云服务提供商回收租户空闲资源. 此外, 在相同的缩放比例下, 长期平均资源回收数量随控制参数 V 的增大而降低. 这是因为, 随着 V 的增大, 所提出的机制更加倾向于延迟处理批处理负载, 而非回收资源来处理批处理负载.

批处理负载产生率对资源回收价格的影响. 实验继续探讨批处理负载产生率 $A(t)$ 对云服务提供商资源回收定价 $p(t)$ 的影响. 为了达到这一目的, 实验对批处理负载产生率 $A(t)$ 放大或缩小一定的比例. 图 7 展示了在不同的控制参数 V 下, 资源回收价格 $p(t)$ 与参数 $A(t)$ 缩放比例的关系. 可以发现,

随时缩放比例的增大 (亦即 $A(t)$ 的增大), 资源回收价格 $p(t)$ 上升. 这是因为, $A(t)$ 增大后, 缓冲队列长度增大, 稳定性降低, 为了维护缓冲队列稳定性, 云服务提供商需要回收更多的资源来降低队列长度, 因此需要提升资源回收价格. 此外, 在相同的缩放比例下, 资源回收定价 $p(t)$ 随控制参数 V 的增大而降低. 这是因为, 随着 V 的增大, 所提出的动态定价机制更加倾向于延迟处理批处理负载, 而非回收资源来处理批处理负载, 因此资源回收定价 $p(t)$ 降低.

6 结论

多租户公有云数据中心的资源通常以“包年包月”的租赁形式长期预留给租户, 租户负载的动态性使得这些预留型资源的利用率低下, 并造成了巨大的能耗浪费. 针对这一愈发严重的问题, 本文提出了面向多租户公有云数据中心资源回收利用的能效激励机制. 该机制通过经济补贴的方式回收租户空闲预留型资源, 并利用其处理云服务提供商所运营的非 IaaS 类业务, 从而提升数据中心整体资源利用率和能效. 基于 Lyapunov 方法, 本文设计并分析了一种面向租户预留资源回收利用的动态定价算法, 该算法无需预测系统未来信息便可作出在线定价决策, 并且在保障系统稳定性的同时使得云服务提供商长期成本无限接近于理论最低值. 通过严格的数学证明以及由真实数据中心负载和电价数据驱动的仿真实验, 本文验证了所提出的激励机制在回收利用租户空闲资源方面的有效性.

参考文献

- 1 Deng W, Liu F M, Jin H, et al. Harnessing renewable energy in cloud datacenters: opportunities and challenges. *IEEE Network*, 2014, 28: 48–55
- 2 Gao P X, Curtis A R, Wong B, et al. It's not easy being green. In *Proceedings of ACM SIGCOMM Conference on Applications, technologies, architectures, and protocols for computer communication*, Helsinki, 2012. 211–222
- 3 Zhou Z. On the energy efficiency of green geo-distributed datacenters. Dissertation for Ph.D. Degree. Wuhan: Huazhong University of Science and Technology, 2017 [周知. 面向绿色跨域数据中心的能效管理研究. 博士学位论文. 武汉: 华中科技大学, 2017]
- 4 Song J, Sun Z Z, Liu H, et al. Research advance on energy consumption optimization of hyper-powered data center. *Chinese J Comput*, 2018, 41: 2670–2678 [宋杰, 孙宗哲, 刘慧, 等. 混合供电数据中心能耗优化研究进展. *计算机学报*, 2018, 41: 2670–2678]
- 5 Armbrust M, Fox A, Griffith R, et al. A view of cloud computing. *Commun ACM*, 2010, 53: 50–58
- 6 Xu Z W, Li C D. Low-entropy cloud computing systems. *Sci Sin Inform*, 2017, 47: 1149–1163 [徐志伟, 李春典. 低熵云计算系统. *中国科学: 信息科学*, 2017, 47: 1149–1163]
- 7 Ma J Y, Sui X F, Sun N H, et al. Supporting differentiated services in computers via programmable architecture for resourcing-on-demand (PARD). *SIGPLAN Not*, 2015, 50: 131–143
- 8 Delimitrou C, Kozyrakis C. Quasar: resource-efficient and QoS-aware cluster management. In: *Proceedings of the 19th International Conference on Architectural Support for Programming Languages and Operating Systems*, 2014. 127–144
- 9 Sun X, Ansari N, Wang R. Optimizing resource utilization of a data center. *IEEE Commun Surv Tut*, 2016, 18: 2822–2846
- 10 Lin C, Tian Y, Yao M. Green network and green evaluation: mechanism, modeling and evaluation. *Chin J Comput*, 2011, 34: 593–612 [林闯, 田源, 姚敏. 绿色网络和绿色评价: 节能机制、模型和评价. *计算机学报*, 2011, 34: 593–612]
- 11 Lin M H, Wierman A, Andrew L L, et al. Dynamic right-sizing for power-proportional data centers. *IEEE/ACM Trans Netw (TON)*, 2015, 21: 1378–1391
- 12 Zhou Z, Liu F M, Zou R L, et al. Carbon-aware online control of geo-distributed cloud services. *IEEE Trans Parallel Distrib Syst*, 2016, 27: 2506–2519
- 13 Ren S L, van der Schaar M. Dynamic scheduling and pricing in wireless cloud computing. *IEEE Trans Mobile Comput*, 2014, 13: 2283–2292

- 14 Feng C, Xu H, Li B. An alternating direction method approach to cloud traffic management. *IEEE Trans Parallel Distrib Syst*, 2017, 28: 2145–2158
- 15 Islam M A, Mahmud H, Ren S, et al. Paying to save: reducing cost of colocation data center via rewards. In: *Proceedings of the 21st International Symposium on High Performance Computer Architecture (HPCA)*, San Francisco, 2015. 235–245
- 16 Greenberg A, Hamilton J, Maltz D A, et al. The cost of a cloud. *SIGCOMM Comput Commun Rev*, 2008, 39: 68–73
- 17 Neely M J. Stochastic network optimization with application to communication and queueing systems. *Synth Lectures Commun Netw*, 2010, 3: 1–211
- 18 Zhou Z, Liu F M. An Energy-Efficient Incentive Mechanism for Resource Recycling in Multi-Tenant Datacenters. Technical Report, 2019. https://1drv.ms/b/s!Ar9mS_s-frkZgd0FE2XOinBcZ7A1Yg
- 19 Yao Y, Huang L B, Sharma A B, et al. Power cost reduction in distributed data centers: a two-time-scale approach for delay tolerant workloads. *IEEE Trans Parallel Distrib Syst*, 2014, 25: 200–211
- 20 Roh H, Jung C, Lee W, et al. Resource pricing game in geo-distributed clouds. In: *Proceedings IEEE International Conference on Computer Communications*, Torino, 2013. 1519–1527

An energy-efficient incentive mechanism for resource recycling in multi-tenant datacenters

Zhi ZHOU¹ & Fangming LIU^{2*}

1. *School of Data and Computer Science, Sun Yat-sen University, Guangzhou 510006, China;*

2. *School of Computer Science and Technology, Huazhong University of Science and Technology, Wuhan 430074, China*

* Corresponding author. E-mail: fmliu@hust.edu.cn

Abstract Multi-tenant public cloud is an important sector of cloud computing. While it has witnessed an unprecedented proliferation in the past years, its low resource utilization that incurs enormous energy waste has received great attention. Since the resource of multi-tenant datacenters is typically rented to the tenants in a reserved manner, the cloud provider is unable to reduce the energy consumption or improve resource utilization via turning down idle servers. To address this challenge, this paper proposes an efficient incentive mechanism to stimulate the recycling of reserved idle resources. Specifically, by economically compensating the tenants, the reserved resources that are idle can be recycled by the cloud providers to host other cloud applications, such as e-commerce and search engine. Taking advantage of the Lyapunov optimization method, this paper rigorously designs and analyzes a dynamic pricing scheme for the recycled reserved resources. Without requiring the future information a priori, the proposed pricing scheme can make online decisions to optimize the long-term operational cost of the cloud provider, while still maintaining the stability of the system. Both rigorous theoretical analysis and extensive trace-driven simulations verify the efficacy of the proposed incentive mechanism.

Keywords multi-tenant datacenter, energy management, resource recycling, incentive mechanism, online algorithm



Zhi ZHOU was born in 1993. He received his B.S., M.E. and Ph.D. degrees in 2012, 2014 and 2017, respectively, all from School of Computer Science and Technology, Huazhong University of Science and Technology (HUST), Wuhan, China. He is currently a research fellow in School of Data and Computer Science, Sun Yat-sen University, Guangzhou, China. In 2016, he has been a visiting scholar at University

of Gottingen. His research interests include cloud computing, edge computing, and distributed systems.



Fangming LIU was born in 1982. He received his B.Eng. degree from Tsinghua University, Beijing, and the Ph.D. degree from the Hong Kong University of Science and Technology, Hong Kong. He is currently a full professor with Huazhong University of Science and Technology, Wuhan, China. His research interests include cloud computing and edge computing, datacenter and green computing, SDN/NFV/5G and

applied ML/AI.