



云数据中心的虚拟机放置问题

樊勇兵¹, 陈 天¹, 陈 楠¹, 黄志兰¹, 吕翠娥²

(1. 中国电信股份有限公司广州研究院 广州 510630; 2. 浪潮集团有限公司 济南 250101)

摘 要:总结了虚拟机放置所面临的算法、优化目标、算法评估、基准模型及验证工具等关键问题,阐述了虚拟机放置问题的典型算法,深入分析了骨干互联网中虚拟机放置问题的特点及对云数据中心工程实践的借鉴意义,并从算法和工程相互适配的角度,对云数据中心设计、运营中所涉及的监控和计量、业务模型、SLA 和资费设计、云资源池设计、资源池规模和架构、业务分区和迁移控制的设计、资源池的资源均衡性、逻辑网络和物理网络基础功能设计及可能的创新等关键问题提出了若干重要原则。

关键词:云数据中心;虚拟机放置;虚拟网络映射;云计算;软件定义网络

doi: 10.11959/j.issn.1000-0801.2015023

Research on Virtual Machine Placement in Cloud Data Center

Fan Yongbing¹, Chen Tian¹, Chen Nan¹, Huang Zhilan¹, Lv Cui'e²

(1. Guangzhou Research Institute of China Telecom Co., Ltd., Guangzhou 510630, China;

2. Inspur Group Co., Ltd., Ji'nan 250101, China)

Abstract: The key issues of virtual machine placement and the related algorithms, optimization goals, algorithms evaluation, benchmark models, verification tools were summarized. The typical algorithms of virtual machine placement were introduced. The characteristics of virtual machine placement of the backbone internet and the related reference values to the engineering practice of the cloud datacenter were analyzed. From the perspective of the adaptation of algorithm and engineering, several important principles for the key problems of the design and operation of cloud datacenter were proposed systematically, such as monitoring and measurement, service model, SLA and tariff design, cloud resource pool design, size and architecture of resource pool, service zone, migration control design, resource balance of resource pool, function design of logical network and physical network, and possible innovation.

Key words: cloud data center, virtual machine placement, virtual network mapping, cloud computing, software defined networking

1 虚拟机放置问题的来源

在云数据中心(cloud data center, CDC),计算资源和网络资源被多个租户动态共享,其中计算资源共享方式主要是虚拟化,即物理服务器(又称物理机)被虚拟化为多台

虚拟服务器(又称虚拟机)。当一个或多个租户的资源请求动态到达并动态变化时,如何有效地为其分配(或调整)虚拟机和网络连接且满足一定的服务等级协议(service level agreement, SLA)和资源约束,称为虚拟机放置(virtual machine placement, VMP)问题。

与 VMP 问题很类似的一个问题是虚拟网络映射 (virtual network mapping, VNM) 问题: 在一个物理网络 (又称为基础网络) 上有多个虚拟网络 (virtual network, VN), 这些 VN 在一定的 SLA 和资源约束下动态共享物理网络中的物理节点 (即物理网络设备) 资源、物理链路资源。

如果将 VNM 中的节点理解为 (物理的或虚拟的) 服务器, 并且将节点限定为接入节点 (即直接接入通信终端设备的节点), 那么 VMP 问题可以理解为带位置约束的 VNM 问题。本文的讨论对象是 VMP。因为 VNM 的算法大多都可以直接或间接应用于 VMP, 所以在不会引起歧义的情况下, 以下将不加区分地使用 VMP 和 VNM。

VMP 问题是云 IDC 的核心问题之一, 在数学上它是一个 NP 难的问题^[1], 这使得 VMP 研究既具实用价值又具学术价值, 从而成为云 IDC 的研究热点。

2 VMP 问题分析

大部分虚拟机放置算法都包含节点映射和链路映射两个阶段, 二者都是组合最优化问题, 也都是 NP 难的问题, 所以虚拟机放置面临的首要问题是算法。此外还有优化目标的综合权衡、算法评估、算法比较、模拟与仿真等。

在规模足够大时求解 NP 难问题只能采用启发式算法。所谓启发式算法, 是相对于最优算法提出的, 可以这样定义^[2]: 一个基于直观或经验构造的算法, 在可接受的成本下给出待解决问题每一个实例的一个可行解, 该解与最优解的偏离程度不一定可以事先预计。启发式算法有诸多好处, 但也有明显不足: 不能保证求得最优解; 表现不稳定; 算法的好坏依赖于实际问题 and 设计者的经验, 很难总结规律, 同时使不同算法之间难以比较。启发式算法可以大致分为如下几类: 一步法; 迭代法; 数学规划法; 解空间松弛法; 现代优化算法, 如模拟退火、遗传算法等, 它们的共同目标是求 NP 难问题的全局最优解, 但 NP 难问题的特性又使它们只能以启发式的算法去求解; 其他方法, 一类算法是根据实际问题而产生的, 另一类算法是集成诸多启发式算法的算法。在虚拟机放置问题中, 算法的选择、集成、改进或适配、算法的并行化、新算法的研究、经验的整理和应用都是重要内容。

在不同场合, 虚拟机放置问题有不同优化目标, 如资源利用率、资源利用均衡性、SLA 保障 (节点保障、链路保障、服务保障、容灾备份、SLA 违规成本等)、节能、收益成本比、使用节点最少 (可能要求更多虚拟机迁移)、降低虚

拟机迁移代价 (可能增加物理服务器数量) 等。其中, 有些目标是互斥的, 如追求资源利用的均衡性则不能保证节能, 追求资源利用率可能导致大量的 SLA 违规, 即使是相容的目标也需要考虑如何转换为综合的单一目标。因此, 优化目标的综合权衡也是虚拟机放置问题不容忽视的方面。

启发式算法有诸多缺陷, 虚拟机放置问题的场景、输入、约束、优化目标复杂多样, 所以对算法的评估成为虚拟机放置问题的重要内容, 评估参数主要包括: 算法的收敛速度、算法适用的问题规模、优化目标的满足程度以及对多优化目标的适用性、映射成功率、对不同请求模型和物理网络拓扑的适用性。

与算法评估紧密相关的是基准模型。为了比较不同算法的特点和优劣势, 必须建立各种算法可以共享的比较基准。参考文献[3]提出了一个标准输入和评估度量, 以对不同算法进行客观比较。

设计完成的算法最终是需要生产环境中使用的, 但在此之前该算法必须要得到验证, 云计算模拟和仿真工具应运而生。

CloudSim 是云计算环境下的现代模拟框架和工具套件。相比于其他工具套件 (如 SimGrid、GangSim), CloudSim 允许虚拟化环境的建模, 支持按需资源指配和管理, 并扩展到支持能源感知的模拟, 还集成了动态负载下的业务应用模拟能力。其他如面向能耗感知的 GreenCloud、面向应用性价比建模的 iCanCloud、面向社交网络工具评估的 CloudAnalyst 等都是各具特色的云环境模拟工具。

此外, 还有用于网格计算和 VNM 环境的平台, 如 GridSim^[4]、GangSim^[5]、Netbed^[6]、PlanetLab^[7]等。

3 VMP 算法概述

表 1 是对 VNM 相关算法的一个分类^[8], 这些算法基本都可以直接或间接地应用于 VMP 问题。表 1 的一个明显缺陷是分类维度不一致。

关于 VMP 算法或 VNM 算法的参考文献很多, 下面补充介绍几个具有显著特色的算法。

参考文献[18]利用了云 IDC 归属于单一管理实体的特点, 提出了在多租户的云中以虚拟数据中心 (virtual data center, VDC) 为资源分配单位的方法, 设计了一种数据中心网络虚拟化架构, 称之为 SecondNet, 将虚拟机到物理机的映射、路由、带宽预留状态等功能都分布在服务器的虚拟机管理器 (virtual machine manager, VMM) 上, 从而达到



表 1 虚拟网络映射算法分类

虚拟网络映射算法	同时考虑节点和链路映射	基础网络资源分配方式	虚拟网络请求处理方式	VNM 的计算方式	虚拟节点与链路映射顺序
面向负载均衡 ^[9]	否	动态	在线	集中式	节点优先
基于流量约束 ^[10]	否	静态	离线	集中式	节点优先
利用路径分割与迁移 ^[11]	是	静态	在线	集中式	节点优先
协同节点与链路映射 ^[12]	是	静态	在线	集中式	节点优先
基于子图同构检测 ^[13]	是	静态	在线	集中式	同时映射
面向 QoS ^[14]	否	静态	离线	集中式	同时映射
分布式 ^[15]	否	静态	离线	分布式	节点优先
动态自适应 ^[16]	否	动态	/	分布式	/
拓扑感知 ^[17]	是	静态	在线	集中式	二种均有

可扩展性;其基于源路由的端口交换(port-switching based source routing, PSSR)特性进一步使 SecondNet 适用于任意网络拓扑;定义了 3 种服务模型;虚拟机分配采用启发式算法。为此,参考文献[18]将相邻的服务器组合成一个集群,如 ToR 集群、Pod 集群或 n -hop 集群,一个服务器可以同时属于一个 ToR 集群和一个 Pod 集群,甚至整个物理服务器集群。当分配一个 VDC 时,仅需要搜索集群而不是整个物理网络。使用单路径的最小成本流算法将虚拟机映射到物理机上,然后通过最短路径算法为虚拟机对分配路径。

仿真结果显示,在一个 10 万物理服务器的数据中心里,该方法可以在平均 493 s 的时间内分配 5 000 个虚拟机,速度非常快;缺点是不能处理多路径问题(可能导致乱序)、不能处理特殊情况(如某些虚拟机要单独部署)、不能处理服务器失效等。

参考文献[19]提出的服务器整合方法充分利用了虚拟机间的通信关系和应用特征的相容性。服务器整合时需要迁移虚拟机,可能使本来距离很近的需要相互通信的虚拟机变得很远。大部分研究者都是基于 CPU 利用率进行虚拟机迁移的,这可能导致同一物理服务器上运行多台有同样资源需求的虚拟机,使该服务器的某类资源(如 CPU 或内存)迅速耗尽,最终结果是需要开启更多的服务器以容纳其他虚拟机,从而提高了总体能耗和流量成本。该参考文献给一对虚拟机定义了一个流量权、一个通信成本(如跳数),并对所有虚拟机及其通信关系建模得到一个图。顶点是虚拟机,边是通信关系,边的权值是流量权。对边的权值从小到大排序,将权值最小的边删除,得到若干子图。如

此递归,最终所有子图都只包含一个顶点。整个过程可以用树型结构描述,最终同一层次的兄弟叶子节点部署在靠近的物理机上(但不是同一台物理机)。

参考文献[20]将商业运营的云 IDC 因 SLA 违规导致的成本作为算法的重要考虑因素,并得到很多富于启发意义的结论。提出一种基于历史数据以自适应预测过载主机的算法,然后根据 3 种策略选择被迁移虚拟机:策略 1——根据所使用内存决定迁移时间最小的虚拟机;策略 2——随机选择;策略 3——选择和其他虚拟机的 CPU 利用率有最大相关性的虚拟机进行迁移。用 Power Aware Best Fit Decreasing 方法^[21]找到目的主机:将虚拟机按 CPU 利用率降序排列,目的主机应提供最小的能耗增加;将所有过载主机处理完之后,选择负载最低的主机,将其上虚拟机迁移到其他主机且使目的主机不过载。若该主机上所有虚拟机被迁移,则关闭该主机。如此循环,遍历所有主机。

参考文献[20]得出的结论是:动态整合算法显著优于静态分配策略;启发式动态整合算法显著优于优化在线确定性算法;上述虚拟机迁移策略的策略 1 显著优于策略 2 和策略 3;基于本地回归的动态整合优于基于阈值和自适应阈值的算法,因为前者更好地预见过载,从而降低了 SLA 违规和迁移次数;基于本地回归的算法优于顽健的本地回归算法,这可以解释为对于实验中的模拟负载,响应峰值负载比平滑这些峰值更重要。

4 骨干互联网的“虚拟机”放置问题

直观上看,骨干互联网也存在 VMP 问题。数以亿计的

接入用户(如宽带用户、专线用户、移动用户、IDC用户、它网用户)通过互联网进行通信,犹如数以亿计的虚拟机通过云 IDC 的网络进行通信。无论是从规模还是从复杂性,似乎骨干互联网的问题严重性都远大于云 IDC 网络,但从实践上看前者解决 VMP 问题的方法是粗糙而有效的,不需要非常精致:以历史流量流向为基础,结合宏观规划或预测(如 GDP、城市规划、大的技术升级或新的应用部署),通过简单的拟合外推等手段,辅以仿真和局部调整,总是能得到基本满意的网络方案。但这并不意味着看似规模和复杂性都更小的云 IDC 的 VMP 问题也同样如此。

骨干互联网的 VMP 具有以下几个特点。

- 只需要考虑单个“虚拟机”的出口带宽,不需要考虑多个虚拟机承载在一个物理机上的情况,也不需要考虑“虚拟机”的 CPU、内存等其他节点约束。
- 只需要考虑独立“虚拟机”的出口带宽,不需要考虑有组网关系和通信需求的多个“虚拟机”的通信约束,其根本原因并不在于“虚拟机”间真的没有通信约束,而是由于骨干互联网以 B/S 架构为主的通信模式、分层汇聚结构、巨大的网络容量和“虚拟机”数量、可预知的单“虚拟机”出口带宽和相对稳定的通信需求、内置的负载均衡、粗颗粒的调度、“尽力而为”的通信保障机制,共同导致了流量模型的统计稳定性,使得“虚拟机”间的通信约束湮没其中,可以忽略不计。
- “虚拟机”数量和总容量是预知且有限的,不需要考虑未知且无限的请求序列;“虚拟机”的位置分布是几乎固定的(除了极少量的移动用户),不需要考虑因节点容量、SLA 等因素的约束而导致的位置选择和迁移,因而简化了链路映射。

总之,不管从实践上还是从理论上,相对云 IDC 来说,骨干互联网的 VMP 问题都是经过高度简化的,简化主要来源于以下 3 个方面。

- 问题特性,如不需要考虑过多的节点约束、链路约束。
- 技术上的“不作为”,如不考虑多个“虚拟机”间的通信约束(事实上也难以考虑)、不大规模地部署细颗粒度的流量工程。
- 技术上的“作为”,如网络架构、内置的负载均衡、分布式处理、粗颗粒的 QoS 机制等。这种简化一方

面带来了操作上的便利性、有效性;另一方面也间接地说明了在大型公众互联网上,细颗粒度、多目标、逐跳预留的流量工程没有被大规模成功部署的原因。

骨干互联网的 VMP 问题为云 IDC 的工程实践提供了重要借鉴。

- 业务特性决定网络特性。云 IDC 网络与骨干互联网本质上是两种不同的网络,这使得前者既存在创新的可能和必要,也存在借鉴和甄别传统技术的可能和必要。
- 虚拟机通信是云 IDC 网络承载的基本业务,可以考虑将 VMP 问题的复杂性部分卸载到网络层或商业层解决(缓解),而不是仅由算法层承担。
- 看似完美但已经证明是徒劳无功的做法(如上述“理想的”流量工程)不应再次成为新网络的负担。

5 VMP 算法和骨干互联网规划为云 IDC 带来的工程启示

VMP 问题是 NP 难问题,一定规模下只能采用启发式算法,而启发式算法的有效性高度依赖于算法设计者的经验——既包括工程经验(对具体问题的理解)也包括算法经验(基于对算法的理解,将工程经验转化为算法设计)。一方面,工程经验可以指导云 IDC 的规划设计,并转化为算法经验内置在运营平台中;另一方面,一个既有的平台可以为算法设计者和云 IDC 设计者反馈经验。因此,经验的获取、积累、应用、反馈、改进以及工程经验和算法经验的互相转化,成为一个重要问题。显然,算法和工程是一种相互适配、相得益彰的关系。

5.1 监控和计量

为了积累、复用基础数据和经验,并将其用于验证或修正已有方法,必须对运营平台进行实时、连续、全面的监控和计量,包括客户请求信息和到达情况、平台和算法运行情况、资源占用情况、SLA 违规情况、流量流向数据、能耗数据等。运营者宜事先筹划并不断完善监控和计量的对象、参数、工具。例如,算法的历史数据既可以作为比对和改进的基础,也可以留待将来有类似需求时参照使用。参考文献[22]针对中心辐射拓扑结构的虚拟网络请求,提出了一种个性化的节点映射算法。推而广之,如果能根据历史数据总结各种请求模型并分别得到较优化的算法,可以极大提高运营效率和效益。



5.2 业务模型、SLA 和资费设计

通过完善的业务模型、SLA 和资费设计,既可以有效发掘和引导用户需求,提供差异化增值服务,又可以使得用户请求的内在关联显性化,量化程度和可预见性也大大增强,从而有利于部分卸载算法层的复杂性和运营管理的迭代改善。例如,现在绝大部分公有云平台所能接受的用户请求参数基本限于虚拟机规格和数量,并不涉及虚拟机间的通信拓扑和带宽、时延等参数,仅有的几个 SLA 指标也是由云服务提供商单方面强加给用户的,这显然不利于业务的精细开展和资源的高效使用。一种可供借鉴的选择是以 VDC 为分配单位(或者考虑虚拟机之间的通信密度),并尽量将有通信需求的虚拟机分配在邻近的节点,不仅能节省资源,还能改善 SLA。参考文献[19]提出的算法可以部分满足要求,前提是要通过某种手段(技术手段或非技术手段)获取用户的通信模型。但从激励用户或提供增值服务的角度看,非技术手段(完善的业务模型、SLA 和资费设计)才是用户可以直接感知的,因此显然是更合适的。

5.3 云资源池的设计

资源池设计的根本依据是业务和技术要求。设计一旦确定,它对算法的效能将起到决定性作用。所以,从算法出发反推,可以给出另一个审视资源池设计的角度,期望设计既不违反通用原则,又有利于算法高效运行。事实上,既存在通用设计原则,也存在针对性设计原则,可以使资源池与算法更为匹配,从而在一定程度上卸载算法复杂性,提高总体效能。

5.3.1 资源池规模和架构

如果资源池中存在明显的稀缺资源,容易导致映射失败^[23]。骨干互联网的“虚拟机”放置问题之所以简单,一个重要原因是规模带来的流量流向的统计稳定性。而传统 IDC 网络的分层汇聚架构恰好在骨干链路形成瓶颈,因此需要规模更大的、稠密的扁平网络。这里的“规模”是相对于资源池中典型的单用户虚拟机通信容量而言的,规模越大,则资源越丰富,流量流向的统计稳定性越强,有利于消除启发式算法内在的不稳定性;“稠密”指的是网络节点的“度”(degree)数要达到一定水平,也就是网络的链路数量要足够多;“扁平”意指去掉不必要的汇聚层次,并使网络在一定范围内(以下称为“域”)形成对称结构,消除局部的资源稀缺。当某些情况下只能采用特殊的网络架构,则算法改变成为必然。

5.3.2 业务分区和迁移控制的设计

严格说,业务分区和迁移控制(尤其是迁移网络)的设计也是资源池架构的一部分。

与上述“扁平”所强调的“连通性”不同,“分区”强调的是“隔离”,即按照一定分类标准(如用户类型、业务类型、生命周期等),每个类别形成一个独立的域,域之间往往通过 3 层进行隔离。分区对算法的意义在于:降低计算复杂度,有利于算法的并行化和隔离,将域特定的经验应用于算法,从而形成定制算法,算法的不稳定性或非全局最优性不会无限累加,也不会域间互相渗透。此外,参考文献[18]认为,VDC 分配过程中能否为 VDC 选择到合适的集群比在确定的集群中能否找到一个好的分配对算法的影响更大(这里的“集群”可以粗略理解为“分区”)。

对于高频率使用的大规模单一应用(如 Hadoop 计算),可使用物理上独立的业务分区。该应用如果必须与其他应用共享物理基础设施,算法上则可将该应用处理为已经成功映射的虚拟机放置请求。

关于云 IDC 网络业务分区的设计,可以借鉴参考文献[24]提出的方法。

云 IDC 区别于传统 IDC 的一个重要特点就是虚拟机迁移。在手工管理或小规模自动化管理的情况下,虚拟机迁移也许不是大事,但在大规模自动化管理环境中,系统往往借助它完成优化部署任务。由于虚拟机迁移对环境敏感,所受限制较多,资源占用厉害,所以规模越大,自动化程度越高,迁移网络相对于业务网络的隔离越重要,迁移的启动控制和实施控制也应该越精细。参考文献[20]提出了具体的迁移控制建议,包括迁移决策方法、迁移依据、迁移与 SLA 违规的关系等。

5.3.3 资源池的资源均衡性

包括同质资源的均衡性(如扁平对称网络、链路带宽和成本、物理服务器的配置和功能)、异质资源之间的均衡性(如网络容量和计算容量之间的均衡性、CPU 和 I/O 之间的均衡性)。这种均衡性本质上是以规模为基础,消除某些资源类别的稀缺性(从资源配比的角度),从而达到物理资源消耗的统计稳定性。

5.3.4 逻辑网络和物理网络基础功能的设计

首先是负载均衡功能,这可以极大缓解因算法内在缺陷导致的对资源池的负面影响,提升资源池对算法的支持能力。通常,网络都具备负载均衡能力。云 IDC 的特殊之处在于,网络已经延伸到物理服务器内部,并且为了尽量不

改变原有基础网络,当前倾向于采用叠加(Overlay)组网方式(在当前实现中,基础网络和叠加网络几乎无法相互感知、适配),隧道端点既可以在物理服务器上,也可以在 ToR 交换机上。如果是前者,受服务器处理能力的限制,一般采用 TCP/UDP 之上的隧道封装,以便很多协议处理可以卸载到网卡上,因此对网卡的要求也成为整体网络功能设计的一部分;如果是后者,为了避免“大象流”对网络的冲击,也需要采用 TCP/UDP 之上的隧道封装,以便形成足够细颗粒度的负载均衡能力,而不是传统基于五元组的负载均衡。另一个特殊之处在于,如果 VMP 算法使用路径分割^[11]方法,负载均衡可能导致分组乱序;此外,某些云 IDC 网络拓扑需要成千上万的等价路径,而传统网络一般只支持 16 条或更少等价路径;最后,业务链(service chain)的引入破坏了传统端到端的负载均衡机制,使原来纯粹基于网络的负载均衡与业务编排、网络服务节点的部署紧耦合。关于路径分割、海量等价路径和业务链导致的负载均衡问题,目前并无成本可行的、统一的商用解决方案,需要进一步深入研究。

其次是 QoS 功能,它同样可以有效卸载算法复杂性,并提供增值服务。云 IDC 网络在数学模型上是一个比骨干互联网复杂得多的网络,如果连后者都从来没有大规模成功部署过细颗粒度、多目标、逐跳预留的流量工程,那在设计云 IDC 的 VMP 算法时更应避免沦为繁琐的流量工程工具,而应借鉴骨干互联网,在物理承载网络上采用区分服务 QoS 模型。

5.4 广阔的创新空间

云 IDC 网络与骨干互联网存在本质不同,并且云资源池大多归属于单一管理实体,这为相关的创新提供了广阔空间。例如,参考文献[18]提出的基于 VMM 的源路由,在公共互联网上就很难实现,因为网络运营商不能控制通信终端,所以路由不能由终端决定,只能由网络决定;Google 提出的 B4 网络^[25]是云 IDC 网络整体架构上的创新,这种短时间内整体架构的创新在公共互联网上是难以想象的,而其技术上的成功又是因为该架构与 Google 新型业务模型的高度适配。

6 结束语

本文总结了虚拟机放置相关的算法、目标及算法评估等关键问题,概述了虚拟机放置问题的主要算法,深入分析了骨干互联网中虚拟机放置问题的特点和对云数据中

心工程实践的借鉴意义,并从算法和工程相互适配的角度,提出了云数据中心设计、运营的若干重要原则,对未来云数据中心的建设和发展具有重大的借鉴价值。

随着云服务和软件定义网络的兴起与发展,VMP 算法作为云管理平台核心组件的重要性越来越突出。虽然算法涉及的基础理论研究大部分都比较成熟,但在该领域的应用研究现状并不尽如人意:没有完整的通用算法框架;算法经验不丰富,更没有系统性;欠缺算法与云资源池、需求模型的适配研究;算法评估和基准比较不完善;有些重要算法(如神经网络)的应用研究还只是非常零星地被发现。这些问题应该是 VMP 算法下一步的重点研究方向。

参考文献

- Andersen D G. Theoretical approaches to node assignment (unpublished manuscript). <http://nms.lcs.mit.edu/papers/andersen-assign.ps>, 2002
- 邢文训, 谢金星. 现代优化计算方法. 北京: 清华大学出版社, 1999
- Xing W X, Xie J X. Modern Optimization Computing Method. Beijing: Tsinghua University Press, 1999
- Zhu J. Benchmarking virtual network mapping algorithms. University of Massachusetts, 2012
- Buyya R, Murshed M. GridSim: a toolkit for the modeling and simulation of distributed resource management and scheduling for grid computing. Journal of Concurrency and Computation: Practice and Experience, 2002, 14(13~15): 1175~1220
- GangSim: a simulator for grid scheduling studies with support for uSLAs. <http://people.cs.uchicago.edu/~cldumitr/GangSim>, 2006
- White B, Lepreau J, Stoller L, *et al.* An integrated experimental environment for distributed systems and networks. Proceedings of the 5th Symposium on Operating Systems Design and Implementation, Boston, MA, USA, 2002: 255~270
- Peterson L, Anderson T, Culler D, *et al.* A blueprint for introducing disruptive technology into the internet. Proceedings of HotNets-I, Princeton, NJ, USA, 2002
- 程祥, 张忠宝, 苏森等. 虚拟网络映射问题研究综述. 通信学报, 2011, 32(10)
- Cheng X, Zhang Z B, Su S, *et al.* Survey of virtual network embedding problem. Journal on Communication, 2011, 32(10)
- Zhu Y, Ammar M. Algorithms for assigning substrate network resources to virtual network components. Proceedings of IEEE INFOCOM, Barcelona, Spain, 2006:1~12



- 10 Lu J, Turner J. Efficient Mapping of Virtual Networks onto a Shared Substrate. Technical Report WUCSE-2006-35, Washington University, 2006
- 11 Yu M, Yi Y, Rexford J, *et al.* Rethinking virtual network embedding: substrate support for path splitting and migration. ACM SIGCOMM Computer Communication Review, 2008, 38(2): 17~29
- 12 Chowdhury N, Rahman M, Boutaba R. Virtual network embedding with coordinated node and link mapping. Proceedings of IEEE INFOCOM, Rio de Janeiro, Brazil, 2009: 783~791
- 13 Lischka J, Karl H. A virtual network mapping algorithm based on subgraph isomorphism detection. Proceedings of the 1st ACM Workshop on Virtualized Infrastructure Systems and Architectures, Barcelona, Spain, 2009: 81~88
- 14 Shamsi J, Brockmeyer M. QoSMap: QoS aware mapping of virtual networks for resiliency and efficiency. Proceedings of IEEE GLOBECOM Workshop, Washington, IX, USA, 2007: 1~6
- 15 Houidi I, Louati W, Zeghlache D. A distributed virtual network mapping algorithm. Proceedings of IEEE ICC, Beijing, China, 2008: 5634~5640
- 16 He J, Zhang S R, Li Y, *et al.* Davinci: dynamically adaptive virtual networks for a customized internet. Proceedings of the ACM CoNEXT Conference, Madrid, Spain, 2008: 1~12
- 17 Cheng X, Su S, Zhang Z, *et al.* Virtual network embedding through topology-aware node ranking. ACM SIGCOMM Computer Communication Review, 2011, 41(2): 39~47
- 18 Guo C X, Lu G H, Wang H J, *et al.* SecondNet: a data center network virtualization architecture with bandwidth guarantees. ACM CoNEXT 2010, Philadelphia, USA, 2010
- 19 Vu H T, Hwang S. A traffic and power-aware algorithm for virtual machine placement in cloud data center. International Journal of Grid and Distributed Computing, 2014, 7(1): 21~32
- 20 Beloglazov A, Buyya R. Optimal online deterministic algorithms and adaptive heuristics for energy and performance efficient dynamic consolidation of virtual machines in cloud data centers. Concurrency and Computation: Practice and experience, 2012(24): 1397~1420
- 21 Beloglazov A, Abawajy J, Buyya R. Energy-aware resource allocation heuristics for efficient management of datacenters for cloud computing. Future Generation Computer Systems, 2011
- 22 Yu M L, Yi Y, Rexford J, *et al.* Rethinking virtual network embedding: substrate support for path splitting and migration. ACM SIGCOMM Computer Communication Review, 2008, 38(2)
- 23 Ricci R, Alfeld C, Lepreau J. A solver for the network testbed mapping problem. ACM SIGCOMM Computer Communications Review, 2003
- 24 樊勇兵, 丁圣勇, 陈楠. 基于业务交换机的大规模云数据中心通用网络架构. 电信科学, 2013, 29(10): 1~4
- Fan Y B, Ding S Y, Chen N. A general network architecture design of large-scale cloud data center based on service switch. Telecommunications Science, 2013, 29(10): 1~4
- 25 Jain S, Kumar A, Mandal S, *et al.* B4: experience with a globally-deployed software defined WAN. Proceedings of SIGCOMM'13, Hong Kong, China, 2013

[作者简介]



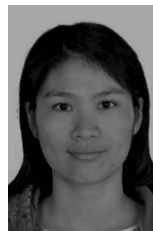
樊勇兵,男,博士,中国电信股份有限公司广州研究院云计算研究所副所长、高级工程师,数据中心联盟IT设备工作组组长,主要研究方向为云计算、IDC、互联网等。



陈天,女,现就职于中国电信股份有限公司广州研究院,主要研究方向为互联网、虚拟化、云计算、IDC。



陈楠,男,博士,现就职于中国电信股份有限公司广州研究院,主要研究方向为SDN、NFV、服务器虚拟化、数据中心网络、管理平台技术等。



黄志兰,女,现就职于中国电信股份有限公司广州研究院,主要研究方向为SDN、云计算、IPv6。



吕翠娥,女,浪潮集团有限公司通信行业部产品总监,主要研究方向为云计算(专注行业云)、大数据。