



云计算数据中心网络研究综述^{*}

罗 萱¹, 叶 通², 金耀辉^{1,2}

(1. 上海交通大学网络信息中心 上海 200240;

2. 上海交通大学区域光纤通信网与新型光通信系统国家重点实验室 上海 200240)

摘 要: 云计算的兴起给数据中心网络提出了新的需求。针对数据中心网络在规模持续增长的情况下, 面临租户隔离和服务保证等挑战这一问题, 从物理网络、网络虚拟化和性能保证等方面介绍了当前数据中心网络的研究热点。

关键词: 数据中心网络; 可扩展; 网络虚拟化; 性能保证

doi: 10.3969/j.issn.1000-0801.2014.02.013

Survey on Data Center Network for Cloud Computing

Luo Xuan¹, Ye Tong², Jin Yaohui^{1,2}

(1. Network & Information Center of Shanghai Jiao Tong University, Shanghai 200240, China;

2. State Key Lab of Advanced Optical Communication System and Networks, Shanghai Jiao Tong University, Shanghai 200240, China)

Abstract: Data center network has faced several challenges from the arising of cloud computing. Along with the continuous expansive bandwidth requirements, data center network also should provide network isolation and service guarantee for each tenant. Current researches focused on the physical network, network virtualization and service guarantee were introduced.

Key words: data center network, scalability, network virtualization, performance guarantee

1 数据中心网络需求

云计算是一种起源于互联网的计算服务, 可以按照需求通过网络将共享的软硬件资源和信息提供给使用者。云计算可以分为 3 个层面: SaaS (software as a service, 软件即服务)、PaaS (platform as a service, 平台即服务) 和 IaaS (infrastructure as a service, 基础设施即服务)。从最基本的层面, IaaS 通过虚拟化技术以共享资源池的形式为用户提供计算、存储和网络等 IT 资源, 用户可以部署、控制和运行操作系统及应用软件, 并只需按照实际使用量付费, 极大地降低了建设和运营成本。同时, 云计算提高了 IT 业务

的灵活性和敏捷性, 能够对大规模的服务器资源进行灵活的调配, 快速地支持大量的并发请求或作业。国外很多公司已经将自己的业务迁移到云计算平台之上, 新兴的互联网公司通常选择将业务放在 Amazon 或 Rackspace 等提供的公有云上, 如 Netflix、Dropbox 等; 而传统的大型公司如 AT&T、思科等则选择在自己的数据中心建设私有云。

云计算的快速发展, 伴随着移动互联网和大数据的发展, 对数据中心网络提出了更高的要求。

1.1 大规模

业务的增长导致带宽的需求持续增加。云计算业务、多媒体业务以及移动设备的普及使得人们对数据的消费

^{*} 国家自然科学基金资助项目 (No.61371084), 国家重点基础研究发展计划 (“973”计划) 基金资助项目 (No.2010CB328205)



加速增长。举例而言,NetFlix 作为美国最大的网络视频服务提供商,其业务流量已经达到全美国流量的 30%,其用户仍在快速增长。Amazon 提供的 S3(simple storage server)云存储服务以及 Dropbox 提供的云存储服务在近几年来一直维持了指数增长。根据 2012 年 6 月的统计数字,Amazon S3 已经存储了 1 万亿对象(每个对象的容量上限为 1 TB)。这样的增长对互联网的带宽提出了很高的要求。同时,它对数据中心网络的带宽要求更为强烈。其原因在于数据中心中不仅要处理前端用户对这些海量数据的访问,也要进行后端程序对此类数据的处理、索引、同步和备份等操作。

1.2 租户隔离

数据中心需要能够同时支持多个租户。出于数据隔离和隐私保障的需求,租户申请的资源,包括计算、存储和网络,应该与其他租户的资源隔离开,以保护租户的数据在任何地方都不被泄漏。在云计算环境中,租户的隔离通常采用虚拟化技术,对应在网络中就是网络虚拟化。数据中心为每个租户或业务分配一个虚拟网络,租户在虚拟网络内可以独享自己定制的网络资源,而不关心物理网络的细节。虚拟网络之间的相互访问需要通过租户定制的策略许可。对于大型的数据中心,其服务的租户数量也非常大,例如 Amazon 数据中心在全球范围内已拥有数十万计的企业租户,并且还在不断增长。根据数据中心运营商 Digital Realty 的统计,近几年租用数据中心的企业数量按每年 7% 增长,这要求网络虚拟化技术具有一定的可扩展性。

1.3 服务保障

租户的业务呈现多样性。不同的业务对于网络的服务质量要求是不一样的,这些需求包括分组丢失率、传输时延、可靠性等。例如,金融业务对实时性、可靠性要求就非常严格,股票交易时间抢点可能会导致租户收益的显著差别;相反,电子邮件业务则可忍受较高的传输时延和秒级以上的系统故障。当租户提出不同的需求时,网络需要能够自动、迅速地给出资源分配方案。

此外,数据中心业务的动态性呈现出急剧增长。云计算的一大特性就是能够动态地为用户提供弹性服务,在很短的时间内提供大量资源。很多业务需要端到端的大容量通道进行数据传输与备份,而其特殊性在于此类通道的需求具有突发性,而且并不会持续很长时间。一个典型的例子是 MapReduce,shuffle 过程中需要的带宽远大于其他过程中需要的带宽,并且通常不会持续太久。移动设备的普

及也对网络的动态资源分配提出了迫切需求。在某个特定的时间与地点,大量的移动用户请求会产生巨大的网络流量。例如大型体育比赛中,很多用户会通过移动设备获取相关资料,也会通过文字、图片以及音频、视频分享现场实况。可以预见的是,随着视频业务的不断发展,在类似的情况下数据消耗总量将会出现巨大增长,数据中心网络需要能够保证这类应用的弹性带宽需求。

为了满足大量租户巨大的带宽需求,数据中心网络在数据层面必须拥有一个可扩展的物理网络,即在不显著增加系统代价(如能耗、硬件成本)的前提下,提升数据中心的带宽供应能力,满足不同租户的宽带传输需求;为对多个租户的资源定制需求做出快速响应,在控制平面需要具有较强的网络虚拟化能力,从大量复杂的物理资源中为每个租户虚拟出一个私有的子网络;为满足不同租户对可靠性的要求,需要同时从两个平面出发设计一个可感知服务质量的可靠性机制,既能满足可靠性需求又不会造成物理资源的过度冗余。下文将围绕物理网络、网络虚拟化以及性能保证这 3 个方面介绍当前的研究热点。

2 物理网络

对于控制层面而言,理想的数据平面应该提供一个简单而又无阻塞的物理层网络,即物理层面可以快速地为服务器之间或者虚拟机之间的业务分配一组带宽足够的通路。为达到这个目标,业界相继在交换机结构设计、路由算法等方面提出各种不同的解决方案。

2.1 大容量、可扩展的交换机

租户数量的增加迫使数据中心不断扩容。2012 年 Amazon 单个数据中心平均拥有的服务器数已超过 6 万台,Google 单个数据中心的服务器数超过了 4.5 万台,微软公司正在打造一个包含 30 万台服务器的数据中心。

把这些服务器互联起来,需要大容量的核心交换机作为网络枢纽。例如,微软拟建成的超大型数据中心包含 30 万台服务器,按照一个边缘交换机可以挂接 50 台服务器计算,共需要约 6 000 个边缘交换机。目前,单台服务器所需要的带宽为 10 Gbit/s,在微软的数据中心中甚至可以达到 40 Gbit/s,据统计有 20%~40% 的流量会流出边缘交换机^[1]。换言之,如果按照 30% 的比例来算,边缘交换机的出口流量将达到 600 Gbit/s。基于业界尽量使用扁平化拓扑的共识,此处假设采用边缘交换机和核心交换机全互联的二层 fat-tree 拓扑结构,那么边缘交换机之间的流量将全部通过

核心交换机中转。由此估算,超大型数据中心的核心层所需的吞吐量将高达 3.6 Pbit/s。这个数字还是按照目前可以看到的数据中心规模和每台服务器的通信需求计算的,如果考虑到数据业务的爆炸式增长,不久的将来所需的核心交换容量将更大。目前业界最先进的数据中心级交换平台为思科的 Nexus 7000,其容量达到 15 Tbit/s,然而这一款交换机的占地面积达 0.40 m²,功耗超过 18 kW。如果按照简单的加法,用它来构造核心层至少需要 240 台交换机(实际情况会更多),将占地 200 m²(考虑机架间距),功耗达到 4.32 MW(不包括冷却功耗)。以此算来,仅核心节点就要消耗掉一个大型电厂的电力输出。

为了缓解核心层交换的功耗压力,2011 年美国的 Rice 大学和 UCSD 分别演示了 C-Through^[2]和 Helios 平台^[3]。这两个平台的共同特点是,采用电交换机和光交叉连接设备共同构成核心交换层,以分别负责传送不同容量或速率的数据流。据估算,这种光电混合方案可将能耗降低 9 倍左右。但是光交叉连接设备的重构时间很慢。例如,Helios 平台采用光微机电系统(MEMS)搭建光交换矩阵,建立光路时间长达 96 ms。为补偿建立光路的开销,光路建立好之后的持续时间最好超过 1 s 以便能一次传送大量数据。为此,C-Through 和 Helios 设置了缓存量阈值。只有当发送端缓存数据量积累超过阈值,才会通知控制平面建立光路集中发送数据。由此可见,对于通过光路传输的数据流来说,通信时延至少包括缓存时延和光路建立的时延。因此,如何提供时延保证的服务依然是这类光电混合核心交换平台的一个挑战。

2.2 带宽保证的路由算法

数据中心通过出租计算、存储为租户提供服务。这些计算和存储资源可以分布在不同位置,它们需要通过具有一定容量的通路连接起来协同工作。在具备大吞吐量交换机和宽带链路的条件下,路由算法是给租户提供通路的关键手段。好的算法不但能满足不同租户包括分组丢失率、时延等参数在内的通信要求,而且具有低的计算复杂度。

等代价多路径(ECMP)算法是一种 RFC 标准化的分组路由技术,ECMP 同时使用预先计算出的多条代价相等的路径,以增加服务器之间的传输带宽。但 ECMP 在分配路径时并不考虑实际的流量负载,而是根据头部信息静态散列到其中一条路径上,容易造成局部拥塞。为了克服这个问题,2009 年微软研究院在其演示的 VL2 网络架构中提出利用 ECMP 技术实现负载均衡(VLB)机制,以处理动

态的流量负载。VL2 以每个流为单位,将数据流随机分配一组等代价路径。尽管如此,遇到大数据流时,VL2 仍然无法避免链路拥塞问题^[4]。

2010 年 UCSD 提出一种集中式数据流调度器 Hedera^[5]。这种调度器需要实时监测网络流量从而识别出大数据流,进而根据全局网络负载信息为每个大数据流计算出优化路径,最后对每个大数据流进行重路由以避免拥塞。显然,这种全局优化算法的扩展性受制于活动的大数据流数。随着网络规模的增大,实时检测并维护整个网络中的全局负载以及所有活动的大数据流信息对集中控制器将是一个巨大的挑战。

同年惠普实验室提出了一种智能路径分配 (SPAIN) 算法^[6]。SPAIN 预先为每对边缘交换机计算出多条不相交的路径,并映射为不同的 VLAN,以充分利用网络的多路径资源。但 SPAIN 要求每个交换机的转发表为所有目的地地址和 VLAN 创建多个条目,比一般商用交换机复杂。

2011 年美国 Wisconsin-Madison 大学和 IBM 演示的 CloudNaaS 平台考虑了区分服务质量的路由问题。此平台将虚拟机间的通信业务分为 3 个等级,并分别采用不同的算法计算路由。为了减少路由表的条目,CloudNaaS 只把数据流分发到有限几条路径上,网络链路利用不高,容易导致局部拥塞^[7]。

由此可见,现有的算法仍然存在各种问题,要么计算复杂度高,要么会带来拥塞无法保证服务质量。如何找到一个既能保证带宽又能简单高效的路由方法,仍然需要开展更多的研究。

3 网络虚拟化

数据中心租户,尤其是企业租户,希望内部数据局限在自己的虚拟网络中而不为他人所知。网络虚拟化需要做到在不同的虚拟网络之间进行隔离,以保证租户的隐私和安全。虚拟网络的隔离主要体现在两方面:第一,通过有效的封装技术区分不同的虚拟网络;第二,允许每个租户在自己的虚拟网络内自由定制地址空间。常见的网络虚拟化技术主要包括 VLAN、网络重叠(overlay)和白名单等。

3.1 VLAN

VLAN 是最早用来实现网络虚拟化的技术。在计算机网络中,一个二层网络可以被划分为多个不同的广播域,这些广播域是相互隔离的,每个广播域就是一个 VLAN。根据 IEEE 802.1q 协议,VLAN 仅有 12 bit 用于区分不同



子网,因此租户数不能超过 4 096 个;另外,VLAN 隧道需要手工配置,基于最小生成树的路由协议限制了网络规模的扩展。

2010 年,Bell Labs 提出将物理网络划分成核心区域和边缘区域,每个边缘区域用若干转发元素与核心区域连通,只用 VLAN 区分一个边缘区域内的不同租户,在一定程度上改善了租户数量的可扩展性。但同一边缘区域的租户数仍然不能超过 4 096 个;每个发往不同边缘区域的数据分组都增加了一个外部 MAC 地址,因此需要交换机支持超大以太网帧^[8]。

3.2 网络重叠

网络重叠是工业界采用的网络虚拟化主流技术。IETF 成立了网络虚拟覆盖(NVO3)技术工作组,讨论各种重叠隧道技术标准用于支持网络虚拟化,比如思科等提出的 VXLAN (virtual extensible LAN)、微软等提出的 NVGRE (network virtualization using generic routing encapsulation)和 Nicira 等提出的 STT(stateless transport tunneling)等。网络重叠通过建立端到端的隧道,在三层网络上保证了二层虚拟网络的全网可达性,具体技术见表 1。

VXLAN 采用一种类似于 VLAN 的封装技术,把基于 MAC 的二层以太网帧封装到三层 UDP 分组中。每一个 VXLAN 划分是通过 24 bit 的 VNI(VXLAN identifier)来标识的。VXLAN 通过多播来实现地址学习,这不仅要求物理设备支持多播,还会因为多播树技术的限制导致 VXLAN 的可扩展性受到影响。

NVGRE 与 VXLAN 类似,两者实际都使用封装策略创建大量的 VLAN 子网,NVGRE 使用通用路由封装创建一个独立的虚拟二层网络。NVGRE 与 VXLAN 不同的是地址学习是通过控制平面实现的,但是,目前 NVGRE 还没有地址学习的具体方案。NVGRE 相对 VXLAN 在负载均衡方面有天生的缺陷。对于传统的三层网络设备,基本都可以实

现基于四层端口的负载均衡,但是却无法基于 GRE key 实现负载均衡。此外,由于建立的是端到端的隧道,因此隧道的数量随终端数量增加以平方速率上升,导致维持隧道的开销增大。

STT、VXLAN 及 NVGRE 使用了类似的封装技术,区别是把数据帧先进行分割再封装,充分利用了网卡的硬件加速功能提高效率。由于 STT 技术将原 STT 分组伪装成了 TCP/IP 分组,而在 TCP 分组头中又没有维护 TCP 状态信息,如果 IP 分组在传输过程中发生了分组丢失,STT 终端不会进行重传,这使得建立的隧道是不可靠通信信道。

3.3 白名单

白名单技术利用类似防火墙的机制来实现虚拟网络间的隔离。白名单通常由用户自定义,维护在网络边缘侧(例如 Hypervisor 或者虚拟机)。所有收到的数据分组要先经过白名单检验,符合白名单中的规则才允许虚拟机之间的通信。白名单技术主要由做安全出身的互联网公司在推动,如 M2Mi。相对于 VLAN 和网络重叠,白名单可以提供更好的灵活性,并且不依赖于底层的网络技术。其局限性也很明显,当虚拟网络规模很大时,白名单的规模会迅速增长;每一次虚拟网络成员以及规则的变化都会导致白名单的更新,维护开销可能成为制约其扩展性的瓶颈。

4 性能保证

云计算环境中给租户提供性能可保证的资源是很大的挑战。本质在于云计算是建立在共享资源池上的复用技术,租户之间的资源竞争导致了云计算服务提供商很难为租户提供性能可保证的云计算服务。

微软研究院根据多篇公开测量报道对以 Amazon 云主机服务 EC2 为主的云计算环境下的 CPU、磁盘特别是网络性能做了分析比较^[9],图 1 显示了云计算环境下虚拟机间的网络带宽测量百分位统计(1%-25%-50%-75%-100%)。

表 1 网络重叠技术

方法	全称	支持厂商	手段
DOVE	distributed overlay virtual ethernet	IBM	利用 OTV 和 VXLAN
EVI	ethernet virtual interconnect	HP	类似 OTV、GRE
NVGRE	network virtualization using generic routing encapsulation	Arista, Broadcom, Dell, Emulex, HP, Intel, Microsoft	24 bit 租户网络标识(TNI)
OTV	overlay transport virtualization	Cisco	MAC-in-IP 封装
STT	stateless transport tunneling	Nicira (VMware)	64 bit context ID, 类 TCP 头利用硬件 offloading 能力
VXLAN	virtual extensible local area network	Arista, Broadcom, Cisco, Citrix, Red Hat,VMware	24 bit 虚拟网络标识(VNI)

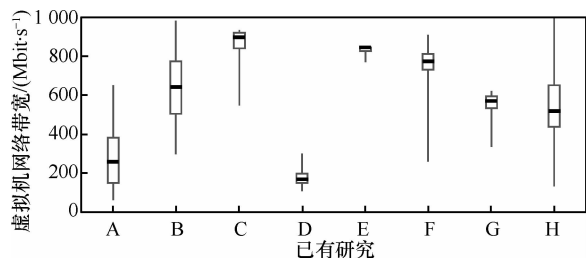


图1 过去研究工作对云计算环境下虚拟机间网络带宽性能的测量百分位统计^[9]

从图中可以看出,虚拟机间网络带宽变化非常大,差异达到5倍甚至更高。

云计算环境网络性能无法保证或者无法预见,将给租户和云计算提供商带来很多负面影响。

首先,限制了云计算的应用领域。例如,2~3年前,能否在公有云上实现 Top 500 的高性能计算(HPC)虚拟集群是很热的话题,但大量研究表明,如果不能解决虚拟机及虚拟网络的性能劣化和变化问题,HPC 应用很难在 IaaS 公有云上运行。

其次,增加了云计算的使用风险。USENIX HotCloud 2012 组织了一个专题讨论云计算风险,云计算环境中由于多租户相互干扰,同一租户的虚拟机间网络性能不稳定、变化大,造成了应用性能无法预知,用户体验下降,租户不敢将他们的应用向云计算环境迁移。无论是对于面向事务处理的电子商务 Web 应用还是 MapReduce 型的数据密集型应用都存在这个问题。

最后,增加了云计算的租用成本。因为云计算按照租户占用虚拟机的时间以及网络流量收费,而任务完成的时间受网络性能影响很大,租户无法估计任务结束时间,分组丢失造成的重传会增加网络流量,租户将潜在支出额外的费用。

Amazon 为了支持网络性能可保证的高端应用,采用了万兆专有网络且无超额认购(over subscription)的高代价方案。因此,云计算环境中更加有效的虚拟数据中心(VDC)需求被提出;VDC 的定义是一个用户可以自定义 IP 地址并具有完整业务等级规约(SLA)的虚拟机集群,SLA 不仅包括计算和存储的需求,也包括虚拟机间通信网络即虚拟网络的带宽需求^[10]。对于租户来说,VDC 在使用上应该与专有数据中心没有差别,因此要求虚拟网络具有良好的隔离性;同时 VDC 比专有数据中心具有更大的吸引力是由于其具有弹性,即可以根据租户的需求动态调整容量。

性能可保证、可预见的虚拟网络是 VDC 技术的关键,可以把这个问题和电信网络中流量工程做一个类比,可以采用带宽预留和呼叫允许控制机制。但是在云计算环境中,性能可预见、可保证的虚拟网络面临的挑战更多。首先,可扩展性要求更高,带宽预留的机制应该能够扩展到超大规模云计算环境的百万台虚拟机规模;其次,实现开销要求更小,电信网络中采用的多协议标签交换(MPLS)并不适用于云计算环境,云计算中引入了虚拟化,网络的边界从接入交换机推到了服务器的虚拟交换机,虚拟交换机实现 MPLS 开销太大,一般采用三层到二层的协议封装技术;最后,云计算具有高动态特点,工作负载弹性可变,尽管工作负载迁移可以缓解网络拥塞,频繁调整工作负载可能给整个云计算环境的稳定带来灾难性影响。

为解决云计算环境中虚拟网络性能可预见、可保证问题,学术界提出了以下几种思路。

(1) 通过预约带宽,保证用户请求的网络服务质量

参考文献[10]提出了一种能提供端到端带宽保证的数据中心网络架构。首先,这个方案需要全网交换设备支持新协议的寻址功能,且布线工作量远高于目前的树型拓扑结构,实际部署较难;其次,端到端的带宽保证带来了可扩展性的问题,当虚拟机数量和租户数量增加时,其对应的管理复杂度将大为提升。再次,在传输流量前,虚拟机需要显式地预约流量,这种静态的流量预约方式不适合高动态的数据中心业务。参考文献[9]在不改变物理拓扑的前提下,使用虚拟拓扑映射构建了一些具有服务质量保证的虚拟子网,但还是无法满足高动态网络负载的需求。

(2) 动态调节虚拟网络的带宽限制或者路由

参考文献[11]提出了一种改进的时间交织的虚拟集群(TIVC)模型来抽象网络带宽需求随时间变化的特性。与固定带宽模型相比,TIVC 模型尽管提高了整个数据中心的效率,但依然不能应对云计算突发业务的实时需求。另外一种思路是大象流被认为是数据中心网络中影响性能的关键流量,因此可以只调整部分大象流的路由来解决可扩展问题^[12]。

(3) 引入反馈控制机制

参考文献[13]通过虚拟交换机的动态流量限制,确保了虚拟机在边界的接入带宽可控。参考文献[14]在每对虚拟机宿主(Hypervisor)间增加了一条拥塞控制的端到端的逻辑隧道。然而,反馈控制方式的时滞较大,简单的域值控



制还可能导致整个系统的震荡不稳定。另外,单纯的边界控制保证虚拟机端到端带宽的能力有限,还应当结合核心交换网络来避免拥塞出现。

综上,一个更加合理的思路是采用预测控制的方法,对于云计算环境中的业务流量、工作负载和链路带宽等进行观测和分析,得出一定的预测规律,从而实现稳定的实时控制,为租户提供性能可预见的虚拟网络。与性能可保证相比,性能可预见在满足租户 SLA 的同时,进一步提高了网络资源利用率。

事实上,已经有一些针对数据中心资源预测的相关研究,这些研究针对数据中心虚拟化的 CPU 资源使用情况进行了预测和控制,使得各虚拟机实际所得的 CPU 资源可以根据工作负载的变化而动态分配。另外,也有研究者使用了预测模型进行云计算环境中工作负载自动扩展的控制。这些预测大都着重计算和存储资源的使用和分配问题,而与 CPU、内存、硬盘等资源不同,网络资源并非从属于某个虚拟设备,而是连接虚拟设备之间的资源,网络流量的分析更为复杂。在数据中心网络的流量测量和分析方面已经有一些研究工作^[15],但基本都是基于网络踪迹(network trace)的离线分析,预测控制需要基于实时或准实时的网络流量分析。

由上可见,在云计算环境中实现性能可预见、可保证的虚拟网络,提供虚拟数据中心业务,具有广泛的应用前景。目前对它的研究尚处于起步阶段,在扩展性、稳定性、实现及部署的低开销和实时响应等方面还有大量的问题有待解决。

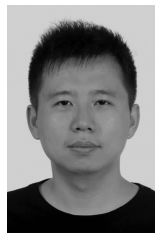
参考文献

- 1 Benson T, Anand A, Akella A, *et al.* Understanding data center traffic characteristics. *SIGCOMM Computer Communication Review*, 2010,40(1):92~99
- 2 Wang G, Andersen D G, Kaminsky M, *et al.* C-through: part-time optics in data centers. *Proceedings of ACM SIGCOMM 2010*, New Delhi, India, 2010:327~338
- 3 Farrington N, Porter G, Radhakrishnan S, *et al.* Helios: a hybrid electrical/optical switch architecture for modular data centers. *Proceedings of ACM SIGCOMM 2010*, New Delhi, India, 2010: 339~350
- 4 Chen K, Hu C, Zhang X, *et al.* Survey on routing in data centers: insights and future directions. *IEEE Network*, 2011, 25(4): 6~10
- 5 Mohammad A F, Radhakrishnan S, Raghavan B, *et al.* Hedera: dynamic flow scheduling for data center networks. *Proceedings of*

NSDI 2010, San Jose, USA, 2010

- 6 Mudigonda J, Yalagandula P, Al-Fares M, *et al.* SPAIN: COTS data-center ethernet for multipathing over arbitrary topologies. *Proceedings of ACM USENIX NSDI 2010*, San Jose, USA, 2010
- 7 Bari M, Boutaba R, Esteves R, *et al.* Data center network virtualization: a survey. *IEEE Communications Surveys and Tutorials*, 2012(2)
- 8 Hao F, Lakshman T, Mukherjee S, *et al.* Secure cloud computing with a virtualized network infrastructure. *Proceedings of USENIX HotCloud 2010*, Boston, USA, 2010
- 9 Hitesh B, Costaet P, Karagiannis T, *et al.* Towards predictable datacenter networks. *Proceedings of ACM SIGCOMM 2011*, Toronto, Canada, 2011
- 10 Guo C X, Lu G, Wang H J, *et al.* Secondnet: a data center network virtualization architecture with bandwidth guarantees. *Proceedings of ACM Co-Next 2010*, Philadelphia, USA, 2010
- 11 Xie D, Ding N, Hu Y C. The only constant is change: incorporating time-varying network reservations in data centers. *ACM SIGCOMM Computer Communication Review*, 2012,42(4): 199~210
- 12 Curtis A R, Mogul J C, Tourrilhes J. DevoFlow: scaling flow management for high-performance networks. *SIGCOMM-Computer Communication Review*, 2011,41(4)
- 13 Rodrigues H, Santoset J R. Gatekeeper: supporting bandwidth guarantees for multi-tenant datacenter networks. *Proceedings of USENIX WIOV '11*, Portland, 2011
- 14 Alan S, Kandula S, Greenberg A. Sharing the data center network. *Proceedings of USENIX NSDI 2011*, Boston, USA, 2011
- 15 Benson T, Akella A, Maltz D A. Network traffic characteristics of data centers in the wild. *Proceedings of IMC 2010*, Melbourne, Australia, 2010

[作者简介]



罗莹,男,博士,现就职于上海交通大学网络信息中心,负责学校数据中心网络的规划建设、云计算平台的开发和建设,主要研究方向为数据中心网络、云计算、网络虚拟化和数据挖掘。

叶通,男,博士,上海交通大学副教授,主要研究方向为宽带光核心交换网络、光接入网、数据中心网络。

金耀辉,男,博士,上海交通大学教授、网络信息中心副主任,主要研究方向为软件定义网络、数据中心网络、网络数据分析。

(收稿日期:2014-01-22)