

突发大数据在存储辅助光电路交换网络中的传输

李超, 林霄, 孙卫强, 胡卫生
上海交通大学, 上海 200240

摘要

电路交换网络交换粒度较粗, 在应对突发式业务时阻塞率较高。针对这一问题, 提出在光电路交换网络节点引入存储, 从而在突发阶段暂存“时延不敏感”的业务数据并错峰传输。基于时移多层图研究了存储容量与业务突发度的关系, 并通过仿真验证了对网络性能的改善程度。研究结论包括: 突发业务对光电路交换网络性能有较大影响; 引入存储可以有效平滑突发业务, 降低阻塞率; 在固定的网络负载下, 所需的存储容量随着业务突发度的增大而增大。

关键词

大数据传输; 辅助存储; 突发业务; 光电路交换

中图分类号: TN915.02

文献标识码: A

doi: 10.11959/j.issn.2096-0271.2016041

Transmission of burst bulk data in optical circuit-switched networks with assistive storage

LI Chao, LIN Xiao, SUN Weiqiang, HU Weisheng

Shanghai Jiao Tong University, Shanghai 200240, China

Abstract

With coarse switching granularity, circuit switching suffers from high blocking probability, especially when the traffic is bursty. How massive intermediate storage may alleviate traffic blocking in OCS (optical circuit switching) networks was studied. The relationship between the storage capacity and the traffic burst was introduced, and the improvement of the network performance was verified by the simulation results. The results suggest that burstiness in traffic is indeed harmful to the OCS blocking performance, and can be effectively handled with massive intermediate storage. It also indicates that with different levels of storage support, the network may exhibit different levels of burst smoothing functions and achieve different blocking performance.

Key words

big data transmission, assistive storage, burst traffic, optical circuit switching

1 引言

近年来,随着信息技术的不断发展和“数据密集型”应用的不断出现,互联网上的数据量正在呈爆炸式增长。宽带化的发展也使得人均网络接入带宽和流量迅速提升。全球新产生数据年增40%,即信息总量每两年就可以翻一番,并且这一趋势将不断持续^[1],这表明大数据时代已经到来。在大数据时代,传输和处理如此大量的数据对网络设施提出了很高的要求^①。

研究显示,在大数据时代的网络中,巨块数据占据和消耗了大量的网络带宽^[2]。例如,为了保证服务质量(QoS),提高访问容错率,云服务提供商通常建立多个地理位置分散的数据中心,并且定期同步和复制它们之间的内容。这些数据中心之间经常需要传输和同步巨量的备份数据^[3]。这类巨块数据的传输有以下两个特点。

- 传输的数据流量巨大。单个业务传输的数据量可以达到GB甚至TB级别。根据某些机构的观察和预测,数据中心之间的流量在2016年将达到1.3 ZB,之后每年以31%的速度增长,在2019年将达到2.8 ZB^[2]。

- 对业务的实时性要求较低。由于没有实时业务的交互性要求,上述业务通常可以容忍数小时甚至数天的时延^[4]。

在传输流量如此巨大的数据业务时,“端到端”的电路交换网络比“逐个分组转发、逐跳处理”的分组交换网络有更大的优势^[5],而且以光路作为传输媒介的光电路交换(optical circuit switching, OCS)技术具有低能耗、大带宽等优点。但是,在电路交换网络中,通信双方在通信之前需要建立一条双方独占的电路。在建立电路之后、释放线路之前,整个线路不允许其他业务使用。这将导致较低的网络资

源利用率。特别当业务表现出一定的突发性时,集中到来的业务请求可能因找不到足够的带宽资源而被拒绝,从而造成阻塞率过高的问题^[6]。大数据传输场景中的突发性业务的问题逐渐引起了人们的关注和研究^[7,8]。

由于大数据传输业务大都能够能够在一定程度上容忍时延,因此可以在电路交换网络中引入存储队列来暂存业务数据。当网络忙碌时,找不到足够带宽资源的业务数据暂存在存储中,等待未来时刻带宽资源充足时再进行错峰传输,这提高了网络带宽利用率,降低了网络中业务请求的阻塞率^[9,10]。

当存储引入网络后,如何为新到业务找到一条既包含带宽资源(空间链路)又包含存储资源(时域链路)的可用通路,是需要两个维度上调度的复杂问题。为了应对这样的复杂调度问题,时移多层图作为一种通用的路由框架被提出^[11]。时移多层图的基本思想是将网络状态随时间的变化记录下来,并用代表存储资源使用状况的时域链路将这些快照连接起来。使用时移多层图进行路由调度,可以统一管理网络中的带宽资源和存储资源,极大地简化了路由过程。时移多层图提供的路由框架可以有效地研究存储电路交换网络传输突发业务的问题。

2 大数据传输环境中的突发业务

在大数据传输网络的经典场景——数据中心之间的通信过程中,业务特性表现出了强烈的不规律性和波动性^[2]。Padgavankar M H等人把易变性作为大数据业务的三大特性之一^[12]。通常情况下,这种波动性的突发业务对网络性能的影响很大,会带来较高的阻塞率^[5],从而使网络性能恶化。

为了改善网络性能,一种可以利用网

①

<http://sloanreview.mit.edu/article/the-storage-and-transfer-challenges-of-big-data/>

②

http://www.cisco.com/c/en/us/solutions/collateral/service-provider/global-cloud-index-gci/Cloud_Index_White_Paper.html

络空闲时的带宽传输大数据业务的策略被提出^②,这样可以在不增加网络成本的情况下完成大量数据的传输。由于网络的空闲和忙碌状态通常表现出以“日”为单位的周期性交替现象^{②[2,4]},把网络忙碌时产生的“时延不敏感”业务存储等待至网络空闲时错峰传输的过程可以看作一个应对周期性突发业务的存储—转发策略。在研究更为一般的随机性非周期性突发业务时,需要一个通用的业务模型和研究方法。

针对于大数据传输场景中的突发业务,有研究者关注了数据中心的微突发业务(micro-burst traffic),基于广泛应用在以太网交换机上的“动态门限值”策略提出了改进的“增强动态门限值”策略来降低微突发业务对网络性能的负面影响^[7]。另一种应对突发性业务的方法是提前预测突发业务的产生,有研究工作提出了一种基于熵的预测方法^[8]。通过仿真验证,这种预测方法的准确度可以达到85%以上。

上述关于突发业务在大数据传输环境中的研究,主要集中在如何应对或预测产生在以太网交换机上的突发性业务。相比于对以太网交换机存储队列造成的影响,突发业务对存储辅助的光交换网络造成的影响将更加明显。这是因为电路交换网络在应对突发业务时的灵活性比分组交换网络差。除此以外,突发业务还可能造成存储使用的溢出,这对光存储技术还不成熟的光网络带来的困难则更为棘手^[6]。因此,深入研究突发大数据业务在存储辅助光电路交换网络中的传输问题,对理解和设计此类网络有重要的实际意义。

3 存储辅助网络研究工具: 时移多层图

在存储辅助的电路交换网络中,当新

的业务请求到达时,网络调度者需要判断是否存在足够的带宽资源可以马上建立电路传输数据,是否需要暂存业务数据,暂存的数据将在何时经由哪条路径传送至目的节点,需要占用的带宽大小和传输时间等问题^[5]。这些决策要求调度者充分了解未来时刻的网络带宽使用状态,并为暂存的业务提前预约未来时刻将要占用的带宽资源。除此以外,当调度者在未来时刻的网络状态中查找路径时,需要同时兼顾各条链路上的带宽资源和各个网络节点存储资源在多个未来时刻的使用情况。这极大地增加了存储辅助的电路交换网络的控制复杂度。时移多层图^[5]可以有效地应对这些情况。以下介绍时移多层图的构成和更新演变过程。

(1) 时移多层图的构成

时移多层图由一系列离散时间点上的网络状态快照组成。在时移多层图中,每一层都是一个时间点上的网络状态快照,随着网络状态的更新不断生长。当前时刻的网络状态快照记录在时移多层图的最上层,下面的层中记录的是未来时刻的网络状态。层与层之间由代表存储使用状态的时域链路相连。**图1**是一个3层的时移多层图示例。

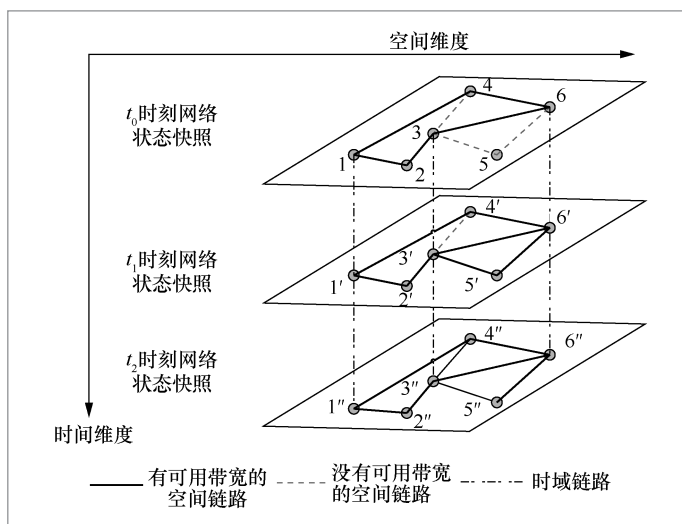


图1 时移多层图示例

时移多层图统一了时间维度和空间维度的调度过程,在时移多层图中使用传统的路由策略(比如最短路径算法),不仅可以像通常在网络的多个节点之间寻找通路,而且可以兼顾多个时刻的网络状态。通过一次路由算法,可以在多个层之间找到所需的传输路径。

如图2所示,新到业务请求从节点1到节点6传输数据。但是,在时移多层图的最上层找不到一条带宽资源足够的可用路径。这意味着当前时刻无法建立电路并传输数据。在时移多层图的所有层中使用最短路径算法后,可以找到一条经过3个层的通路,如图2中黑色虚线所示。业务的传输将在第三层图对应的时刻完成。存储资源和带宽资源统一对待,传统的最短路径算法即可完成该过程。

在时移多层图中为业务查找可用路径时,计算复杂度会随着层数的变多而增大。在实际应用时,可以按照网络对阻塞率的要求和网络控制系统的计算能力,合理限制查找路径的最大层数,从而获取网络性能和计算复杂度的折中^[11]。

(2) 时移多层图的更新演变

在存储网络中,业务的到达或离开会

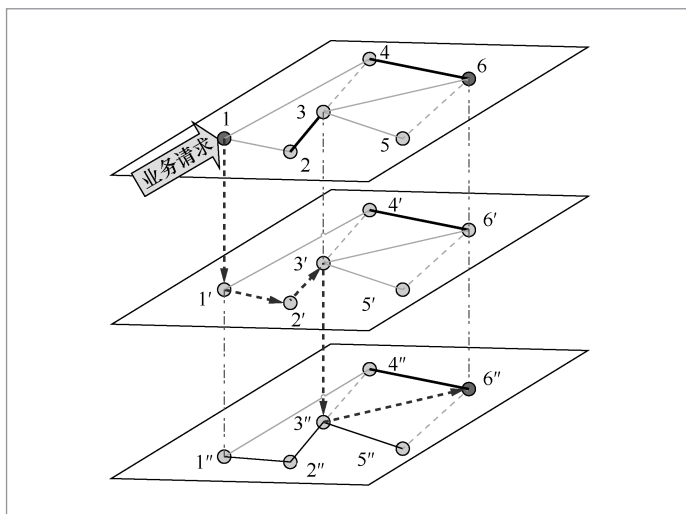


图2 使用时移多层图为业务选路

导致网络状态变化,从而使时移多层图更新^[10]。当新业务请求到达网络时,网络调度者首先根据当前时刻的时移多层图为该业务查找当前或未来时刻可用的传输路径,并在找到后更新时移多层图,从对应时刻的网络状态中减去被占用的带宽资源和存储资源。当有业务完成传输时,该业务占用的带宽资源和存储资源被归还到时移多层图中对应时刻的网络状态中。存储资源的使用情况和带宽资源的使用情况在时移多层图中得到了统一更新和管理^[5,11]。

4 突发业务模型

突发业务广泛存在于当前时代的互联网中,其显著特性是业务传输请求到达网络的速率随时间的变化而随机变化。研究网络中的突发业务时,可以用多种模型对突发业务建模^[6,7,13-15]。利用马尔可夫调制的泊松过程(Markov-modulated Poisson process, MMPP)为突发业务建模可以准确地描述并记录突发业务随时间变化的到达速率^[13]。实际应用中,通常使用只有两个状态的一阶马尔可夫过程来调制突发业务的到达速率状态,即on-off过程。on状态下业务以较高的速率到达网络,off状态下业务的到达速率为零。本文采用上述马尔可夫调制的on-off过程对突发业务建模。

不同研究环境中的突发度的定义并不相同,但都应从统计概念的角度表示出突发阶段和突发间隔阶段业务强度的差异。在马尔可夫调制的on-off过程中,on阶段为突发区间,持续时间服从均值为 θ_{on} 的指数分布;off区间为突发间隔区间,持续时间服从均值为 θ_{off} 的指数分布。

定义突发度为on区间占空比的倒数,即 $(\theta_{on} + \theta_{off}) / \theta_{on}$ 。在off区间,业务的到达率

为0。为了保证突发度不同时的平均业务负载不变,突发区间的业务负载应调整为突发度乘以平均业务负载。例如,突发度为2表示 $\theta_{on}=\theta_{off}$,即突发区间和突发间隔区间的平均长度相同,且突发区间的业务负载为平均业务负载的2倍;突发度为1表示off区间的持续时间为0,即业务没有突发性。

图3是突发度为1和5时,在1 000个单位的仿真时间内随机生成的平均负载相同的业务分布示意。通过对比可以看出,随着突发度的增大,突发区间的持续时间减少,突发间隔的持续时间增大;在突发区间,业务强度变大。

5 结果与分析

针对14个节点的NSFNet进行研究。假设网络的每个节点都装备存储资源。网络每条链路为双向链路,每个方向上各有一个波长。表征平均业务负载大小的单位为Erl,其值为业务的平均到达率(单位时间到达网络的业务数量均值)和平均服务率(单位时间网络可以传输完成的业务数量均值)的比值。存储容量为业务平均负载为1 Erl时,单位时间内到达的业务数据的平均大小归一化的无量纲单位。在上述前提下,从以下3个方面研究了突发业务在存储辅助光交换网络中的传输。

(1) 业务突发性对网络性能的影响

笔者研究了在传统的无存储电路交换网络中,突发业务是如何影响网络性能的。图4为平均业务负载为1~5 Erl时,在不同突发度下网络阻塞率的变化趋势。从图4可以看出,突发业务在不同的负载下对网络性能都有较大的影响,即使在较轻的负载(如1 Erl)下,当突发度从1变为5时,业务的突发性使得阻塞率提高了接近

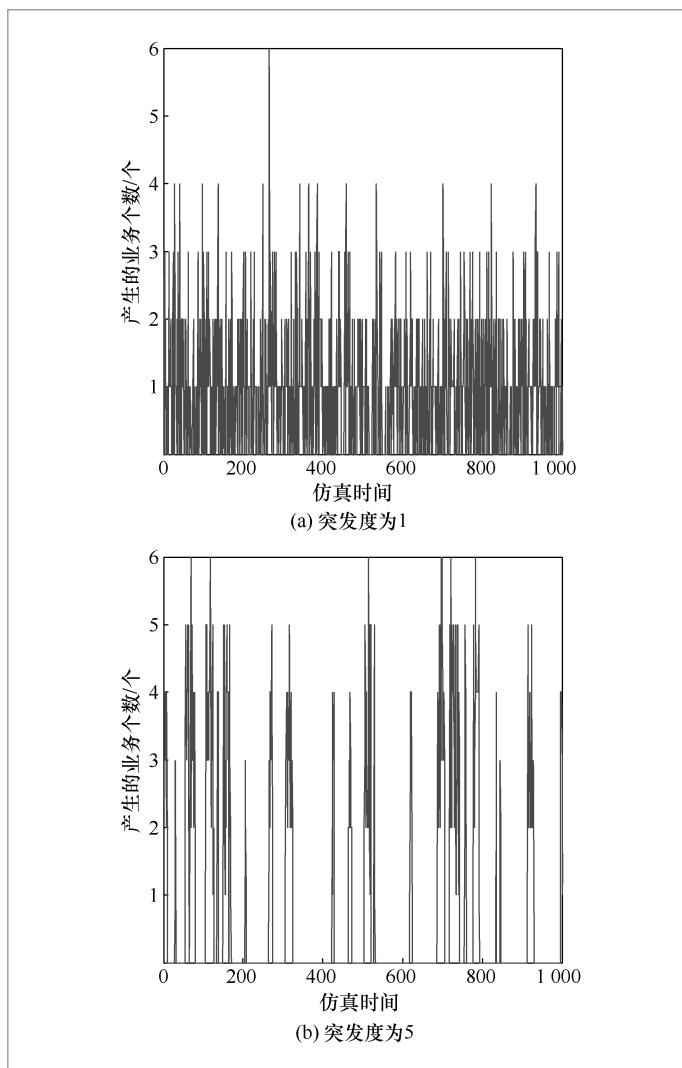


图3 突发度为1和5时的业务分布对比

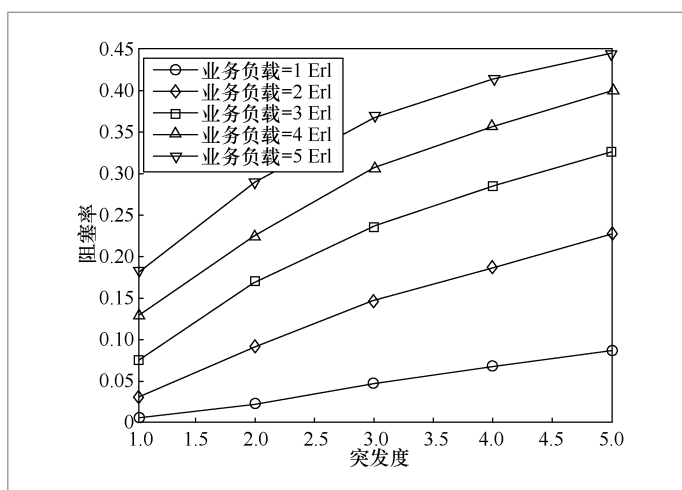


图4 突发度对网络性能的影响

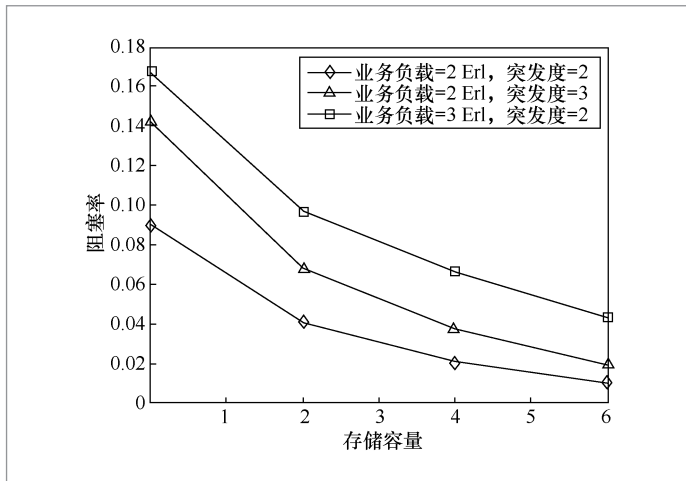


图5 引入存储对网络阻塞率的影响

10倍。

(2) 引入存储对网络阻塞率的影响

存储的作用是可以暂存突发业务在网络空闲时错峰传输，从而降低网络的阻塞率。笔者研究了几种业务负载和突发度下，网络节点上的存储容量对网络性能的影响，如图5所示。可以看出，随着存储的从无到有，从小到大，阻塞率逐渐降低，网络性能逐渐改善。这种改善的效果对于重的负载更为明显。因此，引入存储可以有效地平滑突发业务，降低网络的阻塞率。

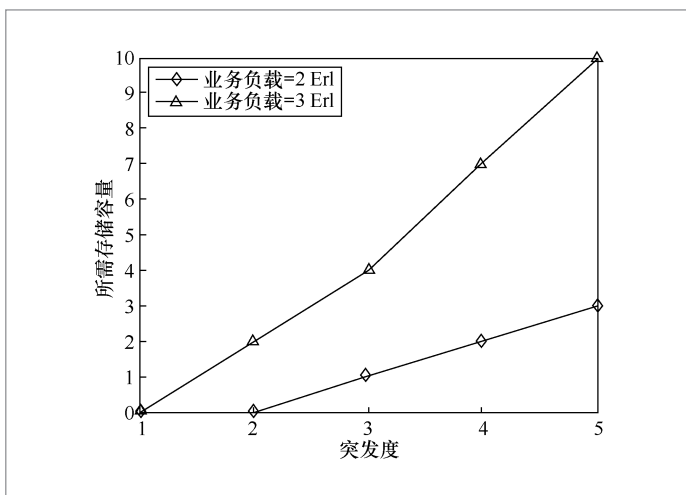


图6 使阻塞率小于10%所需存储容量与突发度的关系

(3) 存储容量与业务突发度的关系

笔者研究了在业务的平均负载分别为2 Erl和3 Erl时，为了使阻塞率小于10%而需在节点上配备的存储容量大小。突发度的增高使得突发业务的到达更加集中，需要更多的存储来暂存业务以保证网络性能。从图6可以看出，在固定负载下，为使网络阻塞率达到一定的标准(<10%)，所需的存储容量随着业务平均负载和突发度的增大而增大。在一定的负载范围内，二者之间的关系近似于线性关系。

6 结束语

在大数据传输场景中，电路交换网络能提供保证服务质量的带宽资源，并且控制开销更低。引入存储暂存“时延不敏感”的大数据业务至网络空闲时错峰传输，可以缓解电路交换在应对突发性业务时灵活性的问题。

本文基于时移多层图研究了突发业务的突发度对存储辅助电路交换网络的影响以及不同大小的存储容量在不同突发度情况下，对网络性能的改善程度。仿真结果显示：突发业务对电路交换网络性能的影响显著，即使在业务负载较轻时，突发度对阻塞率的影响仍然很高，网络性能恶化严重；引入存储后，存储网络在传输业务负载和突发度不同时性能均有所改善，对重负载业务的改善程度更大；在特定的网络负载下，随着突发度的增加，网络为了维持一定的性能，所需的存储资源也在加大，在一定的负载范围内，两者之间近似于线性关系。

参考文献:

[1] 郭贺铨. 大数据时代的机遇与挑战[J]. 求是,

- 2013(4): 47–49.
- WU H Q. Opportunities and challenges in the era of big data[J]. Qiushi, 2013(4): 47–49.
- [2] CHEN Y, JAIN S, ADHIKARI V K, et al. A first look at inter-data center traffic characteristics via Yahoo! datasets[C]// The 30th IEEE International Conference on Computer Communications, Jointly with the Conference of the IEEE Computer and Communications Societies, April 10–15, 2011, Shanghai, China. New Jersey: IEEE Press, 2011: 1620–1628.
- [3] LAOUTARIS N, SIRIVIANOS M, YANG X, et al. Inter-datacenter bulk transfers with netstitcher[C]// The ACM SIGCOMM 2011 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications, August 15–19, 2011, Toronto, Canada. New York: ACM Press, 2011: 74–85.
- [4] LAOUTARIS N, SMARAGDAKIS G, STANOJEVIC R, et al. Delay-tolerant bulk data transfers on the Internet[J]. IEEE/ACM Transactions on Networking, 2009, 21(6): 1852–1865.
- [5] 孙卫强, 胡卫生. 大数据时代的数据传输网[J]. 大数据, 2015, 1(2): 2015019.
- SUN W Q, HU W S. Networking challenges in the big data era[J]. Big Data Research, 2015, 1(2): 2015019.
- [6] CAI Y, LIU Y, GONG W, et al. Impact of arrival burstiness on queue length: an infinitesimal perturbation analysis[C]// The 48th IEEE Conference on Decision and Control, Jointly with the 28th Chinese Control Conference, December 15–18, 2009, Shanghai, China. New Jersey: IEEE Press, 2009: 7068–7073.
- [7] SHAN D, JIANG W, REN F. Absorbing micro-burst traffic by enhancing dynamic threshold policy of data center switches[C]//2015 IEEE Conference on Computer Communications (INFOCOM), April 26–May 1, 2015, Hong Kong, China. New Jersey: IEEE Press, 2015: 118–126.
- [8] WU H, TAHERI J, ZOMAYA A. B-alarm: an entropy based burst traffic prediction approach for ethernet switches in data centers[C]// 2013 IEEE 10th International Conference on High Performance Computing and Communications & 2013 IEEE International Conference on Embedded and Ubiquitous Computing (HPCC_EUC), November 13–15, 2013, Zhangjiajie, China. New Jersey: IEEE Press, 2013: 1221–1226.
- [9] PATEL A, TACCA M, JUE J P. Time-shift circuit switching[C]//Optical Fiber Communication/National Fiber Optic Engineers Conference, February 24–28, San Diego, CA, USA. New Jersey: IEEE Press, 2008: 1–3.
- [10] PATEL A, ZHU Y, SHE Q, et al. Routing and scheduling for time-shift advance reservation[C]//The 18th International Conference on Computer Communications and Networks, August 3–6, 2009, San Francisco, CA, USA. New Jersey: IEEE Press, 2009: 1–6.
- [11] LIN X, SUN W, VEERARAGHAVAN M, et al. Time-shifted multilayer graph: a routing framework for bulk data transfer in optical circuit-switched networks with assistive storage[J]. IEEE/OSA Journal of Optical Communications and Networking Year, 2016, 8(3): 162–174.
- [12] PADGAVANKAR M H, GUPTA S R. Big data storage and challenges[J]. International Journal of Computer Science & Information Technologies, 2014, 5(2): 2218–2223.
- [13] TIAN W. An exact and efficient analytical model for the on/off burst traffic[C]//The 40th Annual Conference on Information

Sciences and Systems, March 22-24, 2006, Princeton, NJ, USA. New Jersey: IEEE Press, 2006: 889-893.

- [14] ŽVIRONIENĖ A, NAVICKAS Z, RINDZEVIČIUS R. Bursty traffic simulation by on - off model[J]. Elektronika Ir Elektrotechnika, 2006, 4(3):145-147.

- [15] WANG X, BAO Y, HU C, et al. Impact of traffic burstiness on optimal batching policy for energy-efficient video-on-demand services[C]//The 6th International Symposium on Communications, Control and Signal Processing (ISCCSP), May 21-23, 2014, Athens, Greece. New Jersey: IEEE Press, 2014: 326-330.

作者简介



李超 (1988-), 男, 上海交通大学硕士生, 主要研究方向为大数据网络。



林霄 (1988-), 男, 上海交通大学博士生, 主要研究方向为大数据网络和网络优化。



孙卫强 (1976-), 男, 博士, 上海交通大学教授、博士生导师, 主要研究方向为大数据网络、信息通信网、网络优化、网络性能评估等。



胡卫生 (1964-), 男, 博士, 上海交通大学教授、博士生导师, 国家杰出青年基金获得者, 主要研究方向为下一代光接入网、光交换、光网络等。

收稿日期: 2016-06-20

基金项目: 国家自然科学基金资助项目 (No.61433009, No.61271217, No. 61471238)

Foundation Items: The National Natural Science Foundation of China(No.61433009, No.61271217, No. 61471238)