

一种适用于光突发交换网络的背景流量估计模型

牛大伟 王 海 于卫波 米智超 郭 晓

(解放军理工大学通信工程学院, 江苏 南京 210007)

摘要 平面背景流量的强度是光突发交换网络中的一个重要参数。该参数直接影响网络的性能并约束网络边缘节点的组装算法。指出了在配置波长转换器的情况下,核心节点控制平面的服务速率是与背景流量密切相关的随机变量而不是确定性变量。构建了 M/D/1 和一种特殊的 M/G/1 排队模型分别描述不同配置的核心节点的行为,并以剩余偏置时间为观测变量,构建了核心节点的背景流量估计模型。仿真和数值分析表明,该模型能够实时反映出核心节点控制平面的背景流量强度。

关键词 光通信;光网络;光突发交换;排队模型;网络测量

中图分类号 TN915.26 **文献标识码** A **doi**: 10.3788/AOS201232.1106002

A Cross Traffic Estimate Model for Optical Burst Switching Networks

Niu Dawei Wang Hai Yu Weibo Mi Zhichao Guo Xiao

(*Institute of Communication Engineering, PLA University of Science and Technology,
Nanjing, Jiangsu 210007, China*)

Abstract The cross traffic intensity of the optical burst switching networks is a key parameter in the network. This parameter restricts the operation of the assembly algorithm in edge nodes. It is revealed that the service time of the control plane in core nodes, closely related with the cross traffic, is a random variable instead of a deterministic variable in the wavelength conversion devices. Two distinct queuing models named M/D/1 and M/G/1 are constructed to describe the behavior of the core node. An estimate model to infer the cross traffic in the control plane of core nodes is proposed based on the observing variable of residual time of the burst. Numerical simulation results verify that the model can reflect the intensity of cross traffic on the control plane of core nodes in time.

Key words optical communications; optical networks; optical burst switching; queuing model; network measurement

OCIS codes 060.1155; 060.2330; 060.4259; 060.4264

1 引 言

光突发交换(OBS)网络被认为是向全光交换网络演进过程中的一种较为可行的光电交换方案^[1]。在 OBS 网络中,核心节点为实现数据突发的全光透明转发,对控制突发进行光电转换和电域的逻辑处理并为延迟到达的数据突发预留波长资源。控制平面的处理能力被认为是 OBS 网络性能的瓶颈所在^[2,3]。其引入的处理时延和排队时延在一定条件下可能会导致较大的早到丢弃(IOT)率。在 OBS 网络中,当网络加载于控制平面的流量较大时,控制报头处理和再生模块会在输出端口处排队,从而引入较大的排队时延。目前,一些文献^[4~6]提出了

以限制控制平面流量为优化参数的边缘节点组装算法。然而,上述算法均假设控制平面的背景流量先验已知,并在此基础上设计算法。事实上,网络中的背景流量是一个动态变化的随机过程,其相对稳态值需要通过特定的测量技术获取。目前一些端到端网络测量技术^[7]主要采用注入探测报文的主动测量方法,这类方法会在网络中引入过多的探测报文,开销较大。文献^[5]的测量方法以端到端突发丢失率为观测变量,在总流量较小的情况下需要较大的采样样本,测量周期较长。提出了一种以出口节点控制突发剩余偏置时间为依据的网络核心节点控制平面背景流量估计模型。应用该模型的测量算法可以

收稿日期: 2012-04-23; 收到修改稿日期: 2012-05-22

基金项目: 国家 973 计划(2009CB320402)和国家自然科学基金重点项目(61032005)资助课题。

作者简介: 牛大伟(1979—),男,讲师,主要从事光网络性能分析、无线自组织网络的测量和认知无线电等方面的研究。

E-mail: flyndw@gmail.com

无需发送探测报文,不引入额外开销。该方法针对核心节点波长转换器的配置情况,分别构建不同的估计模型以提高模型的适用性和针对性。

在 OBS 网络中,每个控制突发都携带一个偏置时间字段。该字段用以指示与该控制突发相关联的数据突发的到达时间并且在每个节点的控制平面进行更新并逐渐减少(源于电域的时间耗费)。出口节点收到控制突发后将检测到偏执时间字段值(后文称为剩余偏置时间)。该偏置时间可以用来采样并计算控制突发在网络中的端到端时延。该时延即为控制突发在瓶颈核心节点中的逗留时间(包括排队时间和处理时间)。控制突发在队列中的逗留时间与核心节点的背景流量密切相关^[8]。若能构建逗留时间到背景流量的反推模型,则可实现网络背景流量的被动估计。构建了两种不同的排队模型,分别描述了核心节点在配置和未配置波长转换器条件下的节点控制平面行为,而后分别针对这两种不同的排队模型开发了背景流量估计模型。仿真和数值分析表明,两种背景流量估计模型能够反映核心节点背景流量的稳态值。本文讨论了两种排队模型和背景流量估计模型的构建依据和构建过程;通过仿真和数值分析给出估计模型的性能评价和正确性验证;并提出下一步研究工作的设想。

2 背景流量估计模型

一般而言,核心节点的排队行为与其负载和服务时间的分布有密切关系。根据文献[9]论述,虽然 OBS 网络边缘节点整形并未改变突发流量的长程相关性,但是在一定时间尺度之内(例如组装时延 T_B),从突发到达间隔的角度分析,整形后流量仍近似服从泊松分布。因此,假设网络中控制平面的流量符合泊松分布。另外,假设一对入口和出口边缘节点之间的通信数据为一个流。每个流在转发路径中至多存在一个瓶颈核心节点(即控制突发至多仅在一个核心节点排队)。当网络中存在多个瓶颈核心节点且数据流相对于背景流量可忽略时,可以将多个瓶颈核心节点中的队列合并为单一队列考虑。一般认为,OBS 控制平面的服务时间为

$$T_S = T_R + T_{OXC}. \quad (1)$$

式中 T_R 为核心节点的资源预留时间, T_{OXC} 为光交换矩阵的配置时间^[10]。在现有的工艺和器件水平下,一旦设备选定,其配置时间 T_{OXC} 一般为确定的。由于资源预留过程可能较为复杂,核心节点资源预留时间 T_R 一般为随机分布的变量,其具体的分布

特点与核心节点所采用资源预留算法有关。因此,资源预留时间 T_R 决定了核心节点处理时间 T_S 的统计特性并进而决定了控制平面的排队特性。决定 T_R 分布特性的资源预留算法的选用主要取决于核心节点是否配置了波长转换器。一般而言,如果核心节点配置了波长转换器,资源预留算法将在若干个符合条件的波长信道间进行仲裁,从而选出最优的波长信道。这类算法的波长资源利用率较高但时间复杂度也较高;在无波长转换器配置的核心节点中,资源预留算法仅进行简单的存储和转发,时间复杂度低,但波长资源利用率也较低。本节将分别分析有波长转换器和无波长转换器的核心节点控制平面的排队模型,并分别给出控制突发在两种模型中逗留时间的分布,进而构建通过逗留时间的观测值估计核心节点背景流量的估计模型。

2.1 无波长转换器的核心节点模型

当核心节点未配置波长转换器^[11]时,其控制平面对突发仅做一次比较运算,即判断该突发是否能在入口节点指定的波长上转发。这个操作较为简单,硬件实现的运算时间几乎为确定的。由(1)式可知,此时核心节点控制平面的处理时间为确定性分布。在泊松流量的负载激励下,可将核心节点建模为 M/D/1 排队系统。假设核心节点队列的输入负载每秒有 λ 个突发,核心节点控制平面的单次查找比对时间为 τ (此处的 τ 为预留算法的一次比较时间),突发在队列中的逗留时间的概率密度函数为 $w(t)$,节点处理时间的概率密度函数为 $b(t)$ 并且队列的归一化负载为 ρ 。由于核心节点无波长转换器,则在输出链路中必须选用与输入波长一致的波长信道进行传输,无需在多个可用输出波长中选择最优信道,其复杂度可以认为是 $O(1)$,此时节点控制平面的资源预留时间 $T_R = \tau$,则逗留时间概率密度函数 $w(t)$ 的拉普拉斯变换为^[8]

$$W(s) = \frac{s(1-\rho)B(s)}{s-\lambda+\lambda B(s)}, \quad (2)$$

式中 $B(s)$ 为 $b(t)$ 的拉普拉斯变换。如前所述, $b(t)$ 为确定性分布且其值为 τ ,所以其拉普拉斯变换为

$$B(s) = \int_{-\infty}^{\infty} \exp(-st)dB(t) = \int_{-\infty}^{\infty} \exp(-st)\delta(t-\tau)dt = \exp(-s\tau). \quad (3)$$

对 $B(s)$ 进行 2 阶泰勒展开并代入(2)式并化简,则突发逗留时间的拉普拉斯变换可写为

$$W(s) = \frac{s(1-\rho)\exp(-s\tau)}{s + \lambda s\tau(s\tau/2 - 1)}. \quad (4)$$

对(4)式进行拉普拉斯逆变换,则可以得到 M/D/1 模型中,突发逗留时间的概率分布密度函数为

$$\omega(t) = a(\lambda)\exp[-a(\lambda)(t-\tau)], \quad (5)$$

式中

$$a(\lambda) = \frac{2\mu(\mu-\lambda)}{\lambda}, \quad (6)$$

式中 $\mu=1/\tau$ 为控制平面的平均处理速率。

在光突发交换网络中,出口节点每收到一个突发,就根据采样检测到的剩余偏置时间计算该突发的端到端时延 t_i (第 i 个突发的端到端时延),该时延值即为控制突发在瓶颈核心节点中的逗留时间。出口节点可根据历史上采样到的突发逗留时间序列 t' , 即 $\{t_1, \dots, t_n\}$, 以及(5)式和(6)式构造核心节点背景流量估计的似然函数,即

$$M(\lambda) = \frac{\sum_{i=1}^n \ln \omega(t_i | \lambda)}{n} = \ln a(\lambda) - a(\lambda)(\bar{t} - 1/\mu), \quad (7)$$

其中符号 \bar{t} 代表突发逗留序列 t' 的采样平均值。背景流量的估计值 λ' 即为使似然函数 $M(\lambda)$ 最大化的值,该估计值可以通过求取微分方程 $dM(\lambda)/d\lambda = 0$ 获取。背景流量估计值 λ' 的求取结果化简为

$$\lambda' = \frac{2\mu(\bar{t}\mu - 1)}{2\bar{t}\mu - 1} \text{ or } \rho' = \frac{2(\bar{t}\mu - 1)}{2\bar{t}\mu - 1}. \quad (8)$$

由(8)式可知,在未配置波长转换器的情况下,出口节点可以通过采样突发端到端时延的平均值估计瓶颈核心节点的背景流量。该流量的估计值可以为边缘节点的动态组装算法、动态偏置时间指派算法或者拥塞避免算法提供优化依据。

2.2 配置波长转换器的核心节点模型

当核心节点配置了波长转换器时,资源预留算法需要在多个可用波长信道中选择一个最优转发信道。此时突发在控制平面的逗留时间取决于资源预留算法的时间复杂度。目前较高效的资源预留算法的时间复杂度为 $O[\lg(N)]^{[8]}$, N 为调度时刻的波长空隙数目。一般认为, N 是符合一定分布的随机变量并且与核心节点背景流量的强度密切相关。如果突发到达时刻,输出链路各波长的空隙总和为 N , 则核心节点资源预约的时间耗费为 $N\tau$, 即共需执行 N 次比较操作,每次比较操作耗费时间常量 τ 。综上,此时的核心节点处理时间是一个符合一定分布的随机变量而非确定值。在泊松流量的激励下,此时的核心节点排队模型可以建立 M/G/1 模型。文

献[8]给出了

$$P(t = N\tau) = \frac{(\lambda D/2)^N}{N!} \exp(-\lambda D/2) \quad (9)$$

所示的核心节点处理时间的分布,即该分布的基本依据为:每个控制突发的到达并成功预约将为输出链路的波长资源增加一个资源空隙,即波长资源的空隙产生分布恰恰就是背景流量的强度分布。在泊松分布的背景流量条件下,资源预留时间的分布可以用离散的泊松分布来表示,其预留时间为单次比较操作时间 τ 的倍数。假设 D 为入口边缘节点指定的最大偏执时间,则上游核心节点对偏置时间的消耗作用导致瓶颈核心节点检测到的偏置时间为区间 $[0, D]$ 内均匀分布的随机变量,其均值为 $D/2$ 。

由(9)式可知,资源预留时间的分布与背景流量的强度密切相关,因此 M/G/1 系统的突发逗留时间 $\omega(t)$ 的解析形式较难获取。无法通过构建类似(5)式的似然函数来估计背景流量。在这种情况下,边缘节点可以利用距估计对核心节点背景流量进行估计。在 M/G/1 排队模型中,突发的平均逗留时间为^[12]

$$W = E(t) + \lambda \frac{1 + C_b^2}{2(1-\rho)}, \quad (10)$$

其中控制突发的平均处理时间 $E(t) = \tau E(N)$, 输出链路的平均波长资源空隙数目 $E(N) = D\tau/2$, C_b^2 为资源预留算法的方差系数,在泊松分布的情况下, $C_b^2 = \frac{V^2(t)}{[E(t)]^2} = \frac{2}{\lambda D}$ 。(10)式中值得注意的是:归一化负载 $\rho = \lambda E(t) = \lambda^2 D\tau/2$ 而不是 $\lambda\tau$, 其原因在于此时的核心节点处理需要执行多次比较操作而非确定分布。将上述变量代入(10)式,则 M/G/1 模型中的突发平均逗留时间可以重写为

$$W(\lambda) = \frac{1}{2}\lambda D\tau + \frac{\lambda D + 2}{D(2 - \lambda^2 D\tau)}. \quad (11)$$

为便于处理,假设偏置时间的分布上限 $D = K\tau$, 即偏置时间的分布上限为离散的 K 倍查找时间,并且核心节点的单次查找操作时间 τ 为单位时间,则平均突发逗留时间为

$$W(\lambda) = \frac{1}{2}\lambda K + \frac{\lambda K + 2}{K(2 - \lambda^2 K)}. \quad (12)$$

由(12)式,假设变量 K 先验已知,则出口节点在获取了采样端到端时延序列 t' 的滑动平均值 \bar{t} 后,可利用方程 $W(\lambda') = \bar{t}$ 获取背景流量的估计值 λ' 。

综上所述, OBS 网络可以采用如图 1 的方法对核心节点背景流量进行估计。图中的出口节点根据网络中波长转换器的配置情况,选择(8)式或(12)式

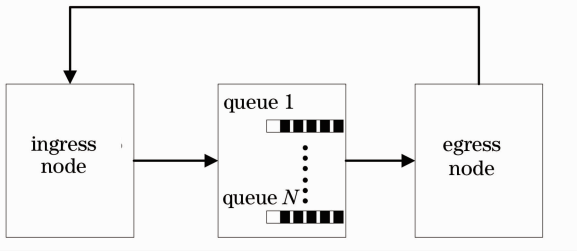


图 1 OBS 网络核心节点流量估计方法示意图
Fig. 1 Diagram of traffic estimation method in the OBS network

对核心瓶颈节点的背景流量进行估计,并以一定的策略为触发条件,将感知结果反馈给入口节点,供入口节点的动态偏置时间和动态组装算法使用。

3 仿真和性能分析

利用仿真和数值分析验证上述背景流量估计模型的有效性。其中仿真利用计算机模拟实现单个瓶颈核心节点的排队行为并依据图 1 所示的方法实现流量感知算法。入口边缘节点的流量采用离散事件驱动模型模拟泊松分布的离散事件。仿真中核心节点的一次查表比较时间取为 $10 \mu\text{s}$,即 $\tau = 10 \mu\text{s}$ 。偏置时间的最大值 $D = K\tau$,其中 K 在仿真中依次取 1~5 的整数。仿真中的核心节点配置为 1 个出端口,波长数为 64,单波长的带宽为 10 Gbit/s。实际上,相对于高速的信道带宽,核心节点的处理能力(每秒有 10^5 个突出)成为瓶颈,并且成为直接影响控制突发端到端时延和背景流量估计的唯一因素。从后面的仿真结论可以看到,核心节点出端口的波长速率和波长的数目并不影响背景流量的估计和控制突发端到端的时延,而真正能够反映背景流量变化的是端到端时延与核心节点处理能力 τ 的相对值的大小。

图 2 为核心节点未配置波长转换器时出口节点检测的端到端平均时延与核心瓶颈节点的实时归一化负载的对比关系图。图中横轴为端到端时延(即突发在系统逗留时间)。纵轴为归一化负载,无量纲。图中的实线曲线为依据(8)式计算的数值结果,方形散点为仿真过程中实时采集的核心节点背景流量及相同时刻的实时端到端平均时延。可以看到,图中实时采集的仿真数据与模型的数值结果十分相近,验证了在核心节点在未配置波长转换器的条件下,流量估计模型的准确性。

图 3 为核心节点配置了波长转换器时的背景流量估计模型的仿真和数值分析结果。图中横轴为核心节点输入的绝对负载,以核心节点资源预留算法

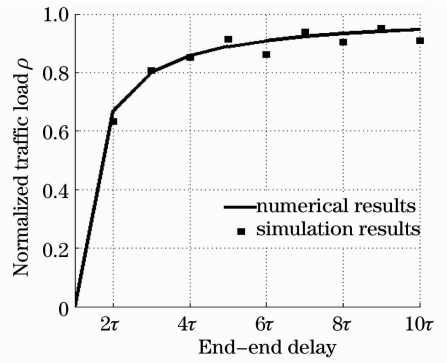


图 2 无波长转换器模型仿真与数值结果
Fig. 2 Numerical and simulation results without wavelength converter equipped in core nodes

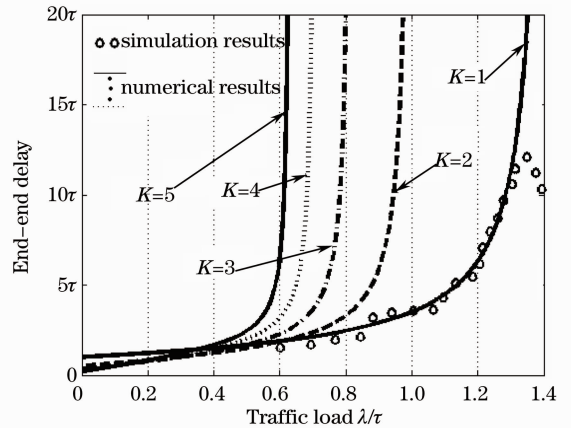


图 3 配置波长转换器模型仿真与数值结果
Fig. 3 Numerical and simulation results with wavelength converter equipped in core nodes

的单次查找时间为单位时间。图中所考察的核心节点背景流量的负载范围从 0 到 1.4。纵轴为端到端时延,其单位为资源预留单次查找时间 τ 的倍数。图中的 5 条曲线为依据(12)式计算的数值结果,分别表示 $K=1$ 到 $K=5$ 情况下的数值输出,限于图 3 的幅面,仅给出 $K=1$ 条件下的仿真结果, K 取其他数值时的仿真结果均能与数值结果相匹配。可以看到, K 越大(即偏置时间的分布范围越广),网络核心节点能够承载的绝对背景流量的极限越小。这是因为在配置波长转换器的情况下,偏置时间分布范围越大,则节点在进行波长资源预约时需要执行的查表操作越复杂(D 增大)。这种复杂的查表操作带来了较大的平均突发处理时延,导致了核心节点所能承载的绝对负载降低。图中可以看到,在 $K=1$ 时核心节点能够承载的最大绝对负载接近 1.4,而当 $K=5$ 时就仅能承载 0.6。另外,在大部分流量负载的情况下仿真数据和数值曲线基本重合。当流量较大时,仿真数据的端到端时延会维持在有限值,而

数值结果趋向于无穷大。这是由于数值模型采用的是 M/G/1 模型,假设核心节点的缓存为无限大,而仿真中的缓存为有限值。文献[13]认为,从吞吐量的角度衡量,当核心节点中的缓存数目大于等于 6 个突发时,其吞吐量性能接近 M/G/1 模型的数值结果。所以,在实际应用中,只要核心节点控制平面的缓存空间大于等于 6 个控制突发,则采用的 M/G/1 模型与实际网络性能趋于一致。但是从端到端时延的角度而不是吞吐量的角度衡量,且当负载接近核心节点所能承受的极限时(如图 4 中归一化负载接近于 1 时),背景流量的估计可能会有一定的误差。这种误差来源于负载极限情况下 M/G/1 模型和有限缓存空间模型的端到端时延差距,此时的核心节点背景流量估计可以采用辅助参考丢失率的方法^[7]。

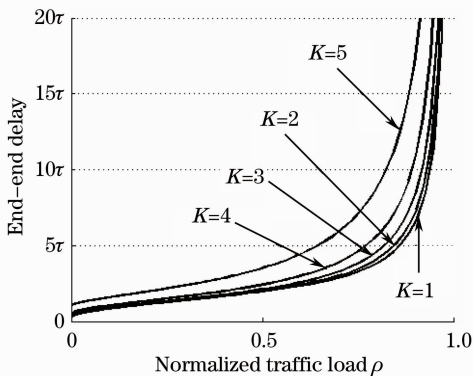


图 4 配置波长转换器条件下归一化负载与端到端时延关系

Fig. 4 Normalized cross traffic versus end-end delay with wavelength converter equipped in the core nodes

图 3 中的背景流量是核心节点背景流量的绝对值。在核心节点配置波长转换器的情况下, K 值的不同将导致核心节点处理能力的不同[即核心节点对控制突发的处理速度 $\mu = \mu(K)$]。在某些情况下,边缘节点可能需要通过端到端时延来估计核心节点的归一化背景流量[即 $\rho = \frac{\lambda_c}{\mu(K)}$, 其中 λ_c 为核心节点背景流量的绝对值]。图 4 为图 3 负载的归一化变换结果, 相比于图 3, 其横轴为背景流量对核心节点处理速度的归一化。综合图 3 和图 4 可知, K 值仅影响核心节点所能承载的绝对背景流量极限, 而不影响其相对流量极限。 K 取不同值时, 核心节点所能承载的归一化负载极限均接近于 1。

4 结 论

本文研究了 OBS 网络中的核心节点控制平面的背景流量估计模型。该模型在区分核心节点波长转

换器配置情况的前提下, 依据不同的排队模型, 构建了利用边缘节点估计核心节点背景流量的模型。仿真和数值分析表明, 该模型能够较准确地估计出核心节点控制平面的背景流量。下一步的研究方向将是依据背景流量估计的结论, 构建高效地边缘节点组算法, 以提高网络整体的吞吐量, 降低碰撞概率。

参 考 文 献

- Guan Aihong, Wang Boyun, Zhang Yuan *et al.*. Optical buffer mechanism based on priority and burst segmentation in optical burst switching networks [J]. *Laser & Optoelectronics Progress*, 2011, **48**(6): 060601
管爱红, 张波云, 张元等. 光突发交换网络中基于优先级与突发分割的光缓存方法[J]. *激光与光电子学进展*, 2011, **48**(6): 060601
- N. Barakat, T. E. Darcie. Control plane congestion in optical burst switched networks[J]. *IEEE/OSA. J. Optical Commun. and Networking*, 2009, **1**(3): B98~B110
- L. K. Pagaria, T. Venkatesh, C. S. R. Murthy. Modeling the impact of control plane losses on the performance of optical burst switched core node[J]. *IEEE Commun. Lett.*, 2010, **14**(11): 1071~1073
- M. A. Toksoz, N. Akar. Dynamic threshold-based assembly algorithms for optical burst switching networks subject to burst rate constraints[J]. *Photo. Netw. Commun.*, 2010, **20**(2): 120~130
- C. Yuan, Z. Zhang, Z. Li *et al.*. A unified study of burst assembly in optical burst switching networks[J]. *Photo. Netw. Commun.*, 2011, **21**(5): 228~237
- A. K. Garg. A novel hybrid approach for efficient network utilization of OBS[J]. *International J. Software Engineering and Its Applications*, 2012, **6**(1): 47~60
- Niu Dawei, Peng Laixian, Yu Weibo *et al.*. Control plane measurement based dynamic offset time algorithm in optical burst switching networks [J]. *J. Electronics and Information Technology*, 2012, **34**(4): 776~781
牛大伟, 彭来献, 于卫波等. 一种基于控制平面测量的光突发交换网络动态偏置时间算法[J]. *电子与信息学报*, 2012, **34**(4): 776~781
- L. de Pedro, J. Aracil, J. A. Hernandez *et al.*. Analysis of the processing and sojourn times of burst control packets in optical burst swithes[C]. *International Conference on Optical Networks Design and Modeling*, 2008. 1~3
- M. Izal, J. Aracil. On the influence of self-similarity on optical burst switching traffic [C]. *IEEE Global Telecommunication Conference*, 2002, **3**: 2308~2312
- Le Zichun, Chen Jun, Fu Minglei *et al.*. Optical cross connection: novel architecture and performance analysis[J]. *Acta Optica Sinica*, 2011, **31**(3): 0306005
乐孜纯, 陈君, 付明磊等. 一种新型结构光交叉连接节点及其联网性能分析[J]. *光学学报*, 2011, **31**(3): 0306005
- Hu Weisheng, Sun Weiqiang, He Hao *et al.*. Time structure and space architecture of optical switching [J]. *Laser & Optoelectronics Progress*, 2012, **49**(1): 010001
胡卫生, 孙卫强, 何浩等. 光交换的时间及空间结构分析[J]. *激光与光电子学进展*, 2012, **49**(1): 010001
- L. Kleinrock, R. Gail. *Queuing Systems: Problems and Solutions* [M]. New York: Wiley-Interscience Publicatoin, 1996. 119~179
- M. Hassan, R. Sarker, M. Atiquzaman. Modeling IP-ATM gateway using M/G/1/N queue [C]. *IEEE Global Telecommunication Conference*, 1998, **1**: 465~470