

## 光突发交换网络中基于波长分集的 QoS 算法

王雄 王晟 谭伟

(电子科技大学宽带光纤传输和通信技术重点实验室 成都 610054)

**摘要** 光突发交换(Optical Burst Switching, OBS)被公认为是构建下一代光网络的有效交换技术。在 OBS 网络中有效地支持 QoS (Quality of Service)是一个很重要的问题。该文从波长分集的思想出发,提出了几种适合 OBS 的 QoS 算法。这些算法可以根据各个优先级业务的变化情况,动态地调整各个优先级的业务使用的波长数目。通过仿真,并和已有算法相比,说明所提出的算法可以更好地提供区分服务,同时有效提高信道的利用率,降低整体的丢失率。

**关键词** 光突发交换, 波长分集, QoS, 偏移时间

中图分类号: TN915.63

文献标识码: A

文章编号: 1009-5896(2006)12-2374-04

## QoS Algorithms Based on Wavelength Grouping in Optical Burst Switching Networks

Wang Xiong Wang Sheng Tan Wei

(Key Lab of Optical Fiber Transmission and Communication Networks,

University of Electronic Science and Technology of China, Chengdu 610054, China)

**Abstract** Optical Burst Switching (OBS) is considered as an efficient switching technique for the next generation optical networks. QoS (Quality of Service) support is an important issue in OBS networks. This paper focuses mainly on the QoS algorithms based on wavelength grouping schemes, and proposes several new QoS algorithms. These algorithms can adjust the wavelengths used in each class dynamically, according to the variation of its traffic. Simulation results and comparisons with the existing algorithms show that these algorithms can provide better differentiated service, decrease the burst loss probability and improve the channel utilization.

**Key words** Optical Burst Switching(OBS), Wavelength grouping, QoS, Offset-time

### 1 引言

急剧增长的因特网业务要求未来的骨干传输网络具有极高的传输容量。密集波分复用(DWDM)技术被认为是满足骨干网络巨大传输带宽要求的最好方法。DWDM网络中单个波长的传输速率可以达到 10~40Gbps。传统的波长路由技术不能充分地利用网络带宽,造成大量网络资源的浪费。而光分组交换虽然能有效地利用带宽,但全光的分组交换在实现上存在若干难点,许多关键技术的突破仍有待于光器件技术的成熟(如灵活有效的光逻辑器件、光存储器件以及光域的同步技术等)。因此突发交换(Optical Burst Switching, OBS)作为一种光分组交换和波长路由的中间技术被提了出来<sup>[1]</sup>。在 OBS 网络中,突发数据(Burst)和控制分组BHP(Burst Header Packet)是在不同的物理通道上传输的, BHP作为控制分组会在电域进行处理,而Burst则是全光传输的。当多个IP包在边缘节点汇聚成一个Burst后,它相应的BHP就会比Burst提前一个偏移时间(offset-time)发往核心节点进行资源的预约。

如同分组交换一样, OBS仅仅是尽最大努力转发Burst而并不确保Burst的转发成功。当一个Burst进入核心节点后或许会与其它Burst冲突从而在某一个核心节点被丢弃,尤其在

业务量比较大的时候,丢包的现象会更明显。虽然可以通过改善调度算法来改善丢包性能,但性能的改善增加了调度算法的复杂性,而且到一定程度后,调度算法对性能的改善变得非常有限。这就要求通过其它途径来进一步降低突发丢失率,如只保证一部分突发数据有较低的丢失率。另一方面,由于IP网络不能很好地支持QoS(Quality of Service),不能满足实时性较高的业务,因此提出了两种QoS模型,即Interserv和Diffserv,这两种模型已被深入研究,有些还形成了标准。但是传统Internet的QoS方案都是基于电缓存的,如FQ(Fair Queuing),WFQ(Weighted Fair queuing)等不能移植到OBS网络中。在OBS网络中,除了依靠上层的IP或MPLS协议来区分不同业务以提供QoS外,还希望直接在光层提供基本的QoS,用以补充和提高IP层的QoS。综上所述,如何在OBS网络中有效地支持QoS是一个很重要的问题<sup>[2]</sup>。

目前,在光突发交换网络中提供QoS的方法主要有:(1)基于额外偏移时间的QoS<sup>[2]</sup>,通过给高优先级分配较大的偏移时间,可以使高优先级不受低优先级的影响,以更大的概率成功预约资源。(2)基于主动丢弃的QoS<sup>[3]</sup>,根据高优先级的丢失情况决定是否对低优先级的突发数据分配资源,如果高优先级的丢失率超过一定限度时,将主动提前把低优先级

突发数据丢弃。(3)基于混合突发汇聚的QoS<sup>[4]</sup>,在汇聚时将不同优先级的IP包放到一个突发数据中,并且将低优先级的IP包置于易丢失的部位,如果在核心节点产生冲突,则通过分段尽量保留高优先级部分,而抛弃低优先级部分。(4)基于DiffServ的突发汇聚和调度的QoS<sup>[5]</sup>。该方案采用了IETF的标准,将业务分为Expedited Forward, Assured Forward和Best Effort 3类,并在核心节点通过WRR (Weighted Round Robin)等传统的算法进行调度。本文主要从波长分集出发,提出了几种基于波长分集的QoS算法。

## 2 基于波长分集的 QoS 算法

在波长路由中,为了提供区分服务,提出了波长分集。波长分集的主要思想是按业务的优先级来分配波长,高优先级业务分配的波长比低优先级的多,从而获得较低的网络阻塞率。文献[3]把波长分集引入了光突发交换,提出了 SWG 和 DWG 两种算法。SWG 算法中每个优先级分配固定的波长;DWG 算法则只在一个时段内限定每个优先级使用的波长个数,而不限制具体使用的波长。尽管 SWG 和 DWG 算法可以通过设定不同数量的波长来提供区分服务,但是还有不足之处,因为给各个优先级分配波长后,所有的波长将分别被各个优先级独享,那么当高优先级的波长不够用时,即使低优先级的波长空闲,也不能为高优先级使用;对高优先级业务而言,波长的利用率将比较低。为此,我们进一步提出几个基于波长分集的 QoS 算法,如 FLPL(Fixed Low Priority Limitation), DLPQ(Dynamic Low Priority Quota)和 IDLPQ(Improved Dynamic Low Priority Quota)。

### 2.1 FLPL 算法

FLPL 算法的主要思想是允许高优先级在所分配的波长不能满足要求时可使用低优先级的波长,而低优先级任何时候都不能使用高优先级的波长。

设交换阵每个端口上总的波长数目为  $M$ , 波长按编号分别记为  $w_j$  ( $j=1,2,\dots,M$ ); 总的优先级数为  $N$  ( $i=0,1,\dots,N-1$ ),  $0$  为最高优先级;  $t_n$  表示时刻;  $\Delta$  为观察时间长度;  $W_{t_0,i}$  为初始分配给优先级  $i$  的可用波长集合;  $W_{m,i}$  为时间  $(t_n, t_n + \Delta)$  内分配给优先级  $i$  的可用波长集合;  $\tilde{w}_{m,i}$  为时间  $(t_n, t_n + \Delta)$  内优先级  $i$  所需的波长数目;  $N_i$  为时间  $(0, t_n)$  内优先级  $i$  的 BHP 数目;  $N_{m,i}$  为时间  $(t_n - \Delta, t_n)$  内优先级  $i$  的 BHP 数目;  $\lambda_i$  为时间  $(0, t_n)$  内优先级  $i$  的 BHP 到达率,  $\lambda_i = N_i / t_n$ ;  $\lambda_{m,i}$  为时间  $(t_n - \Delta, t_n)$  内优先级  $i$  的 BHP 到达率, 即  $\lambda_{m,i} = N_{m,i} / \Delta$ 。具体算法如下:

(1) 给优先级  $i$  ( $i=0,1,\dots,N-1$ ) 预先分配不同的可用波长集合  $W_{t_0,i}$ ,  $W_{t_0,i}$  满足以下两个条件: (a)  $W_{t_0,i} \cap W_{t_0,j} = \Phi$ ,  $i \neq j$ ; (b)  $\cup_j W_{t_0,i} = \cup_j w_j$  ( $j=1,2,\dots,M$ ), 即所有的波长都可用。

(2) 等待 BHP 到达。如果有 BHP 到达, 转(3)。

(3) 根据 BHP 的优先级  $i$ , 首先在  $W_{t_0,i}$  中采用调度算法

查找可用波长; 如果未找到可用波长, 再根据  $j$  ( $j=i, i+1, \dots, N-1$ ) 的增序在  $W_{t_0,j}$  集合中查找, 如果仍然没有找到合适的波长, 则丢弃该 BHP 及其相应的突发数据, 否则调度成功。统计调度成功或失败的数据长度, 并返回(2)。

需要说明的是, 本算法第(3)步中采用的调度算法没有特殊的要求, 如传统的 LAUC<sup>[6]</sup>等算法即可。

### 2.2 DLPQ 算法

FLPL 算法中, 虽然能够充分利用低优先级的波长资源为高优先级服务, 但如何为各个优先级分配波长成为一个关键问题, 分配不同的波长数将导致不同的性能; 而且当高优先级的业务量较小、低优先级的业务量较大时, 高优先级的波长资源将被浪费。因此, 提出了改进的 DLPQ 算法, 该算法将采用动态分配波长的方法, 在保证高优先级服务质量的同时进一步提高波长的利用率, 从而降低低优先级的数据丢失, 提高总体性能。DLPQ 算法具体如下, 其中  $k$  是一个常量 ( $0 < k < 1$ ), 可以控制波长变化的灵敏度。

(1) 给优先级  $i$  ( $i=0,1,\dots,N-1$ ) 预先分配不同的可用波长集合  $W_{t_0,i}$ ,  $W_{t_0,i}$  满足以下两个条件: (a)  $W_{t_0,i} \cap W_{t_0,j} = \Phi$ ,  $i \neq j$ ; (b)  $\cup_j W_{t_0,i} = \cup_j w_j$  ( $j=1,2,\dots,M$ ), 即所有的波长都可用。设置观察时间长度  $\Delta$ , 置观察时间  $t=0$ ,  $n=0$ 。

(2) 启动定时器, 有 BHP 到达时, 统计当前时间段内到达的各个优先级的 BHP 个数  $N_{m,i}$ , 并转(3)。

(3) 根据 BHP 的优先级  $i$ , 首先在  $W_{m,i}$  中采用调度算法查找可用波长; 如果未找到可用波长, 再根据  $j$  ( $j=i, i+1, \dots, N-1$ ) 递增的顺序在  $W_{t_0,j}$  集合中查找, 如果仍然没有找到合适的波长, 则丢弃该 BHP 及其相应的突发数据, 否则调度成功。统计调度成功或失败的数据长度。

(4) 检查观察时间  $t$  是否等于观察时间长度  $\Delta$ , 如果是则转(5), 否则转(2)。

(5) 根据(2)中的统计值计算  $\lambda_i$  和  $\lambda_{m,i}$ 。

(6) 计算  $\tilde{w}_{m,i}$ : 如果  $\lambda_{m,i} / \lambda_i \geq 1$ , 则  $\tilde{w}_{m,i} = |W_{t_0,i}| + \min \left\{ \left[ (\lambda_{m,i} / \lambda_i - 1) \times |W_{t_0,i}| \right], k \times |W_{t_0,i}| \right\}$ ; 如果  $\lambda_{m,i} / \lambda_i < 1$ , 则  $\tilde{w}_{m,i} = |W_{t_0,i}| + \max \left\{ \left[ (\lambda_{m,i} / \lambda_i - 1) \times |W_{t_0,i}| \right], -k \times |W_{t_0,i}| \right\}$ 。

(7) 根据  $j$  ( $j=i, i+1, \dots, N-1$ ) 递增的顺序依次为各个优先级  $i$  ( $i=0,1,\dots,N-1$ ) 重新分配可用波长集合  $W_{m,i}$ 。观察时间  $t=0$ ,  $n=n+1$  并转(2)。

需要说明的是, (7)中因为是按优先级从高到低的顺序重新分配波长, 所以有可能出现分配给低优先级的可用波长的数目  $|W_{m,i}|$  小于所需求的波长数目  $\tilde{w}_{m,i}$ 。

举例说明 DLPQ 算法。设交换阵的每个交换端口有  $M=8$  (波长号为  $0 \sim 7$ ) 个波长, 优先级数为  $N=3$  ( $i=0, 1, 2$ ),  $0$  的优先级最高。初始分配给  $0, 1$  和  $2, 3$  个优先级的可用波长集合分别为  $W_{t_0,0} = \{0,1,2,3\}$ ,  $W_{t_0,1} = \{4,5,6\}$ ,  $W_{t_0,2} = \{7\}$ 。经过一段时间, 设在时刻  $t_n$ ,  $3$  个不同优先级所需的波长数目分别为  $\tilde{w}_{m,0} = 5$ ,  $\tilde{w}_{m,1} = 4$ ,  $\tilde{w}_{m,2} = 1$ 。但由于是按优先级

从高到低的顺序来重新分配波长，所以重新分配的可用波长集合分别为  $W_{m,0} = \{0,1,2,3,4\}$ ， $W_{m,1} = \{5,6,7\}$ ， $W_{m,2} = \Phi$ 。

### 2.3 IDLPQ 算法

在 IDLPQ 算法中，根据上一个时间段内业务变化情况，对各个优先级业务在下一个时间段内使用的波长数目进行调整。但这种业务的变化趋势不一定有相关性，而且有可能由于业务的增长或减少，在上一个时段内已经造成 Burst 大量丢弃或波长利用率降低。所以我们对 IDLPQ 算法进行改进，把观察时间  $\Delta$  内收到的 BHP 暂时放在一个队列中，等到完成波长调整后再调度。由于调度是在完成波长调整后开始的，所以可以避免由于预测误差带来的丢包或波长利用率降低的问题。

(1) 同 DLPQ 算法第(1)步。

(2) 启动定时器，如有 BHP 到达，将 BHP 放入队列  $Q$ ，并统计当前时间段内到达的各个优先级的 BHP 个数  $N_{m,i}$  并转(3)。

(3) 检查观察时间  $t$  是否等于观察时间长度  $\Delta$ ，如果是则转(4)，否则转(2)。

(4) 根据(2)中的统计值计算  $\lambda_i$  和  $\lambda_{m,i}$ 。

(5) 同 DLPQ 算法第(6)步。

(6) 同 DLPQ 算法第(7)步。

(7) 依次从队列  $Q$  中取出 BHP 用 DLPQ 算法第(3)所描述的方法调度。调度完队列中的所有 BHP 后，将观察时间置零  $t = 0$ ， $n = n + 1$ ，转至(2)。

### 3 仿真及结果分析

仿真实验基于 OPNET 网络仿真平台，采用如图 1 所示的一个单核心节点和 4 个边缘节点的试验网络模型，核心节点的光交换结构有 4 个输入/输出端口，每个端口 8 波长，交换模块采用  $1 \times 32 \times 32 + \text{TWC}$  结构，没有配置 FDL。假设 IP 包到达边缘节点的间隔时间服从负指数分布，均值为  $\lambda$ 。边缘节点采用 MAPMBL<sup>[7]</sup> 汇聚算法，突发包的最大长度为  $130\mu\text{s}$ 。边缘节点产生两种不同优先级的突发数据，分别为 class0 和 class1(class0 为高优先级)。class0 的业务量在总业务量中的比重为  $p$ ，并设  $p$  在  $(0.5, 0.9)$  的区间内服从均匀分布。BHP 的偏移时间在  $(300\mu\text{s}, 400\mu\text{s})$  的区间内服从均匀分布。核心节点每个端口的业务量强度相同。仿真中  $k$  取 0.5，DLPQ 和 IDLPQ 算法中观测时间  $\Delta$  均为  $200\mu\text{s}$ 。

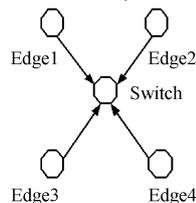


图 1 试验网络模型

Fig.1 Experimental network

图 2 给出了 FLPL 算法中各优先级的数据丢失曲线，并

与文献[3]中的 SWG 算法进行了比较。可见 FLPL 算法中高低优先级(分别以 FLPL-class0 和 FLPL-class1 表示)之间性能的差别非常明显。同时与 SWG 相比，FLPL 中高优先级的丢失率更小，而由于在 FLPL 算法中，高优先级业务可以在波长不够用的条件下使用分配给低优先级的波长，所以 FLPL 算法中低优先级业务的丢失率稍高；但是从平均丢失率来看，当业务量强度大于 0.6 后，FLPL 的平均丢失率要低于 SWG 的平均丢失率，说明总体性能趋好。

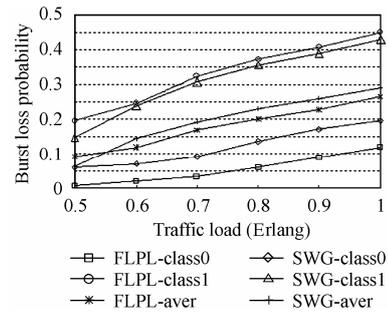


图 2 FLPL 和 SWG 算法的丢失性能

Fig.2 Burst loss performance for FLPL and SWG QoS schemes

图 3 为使用 DLPQ 算法时，在一段时间内低优先级(class1)的业务所使用波长数目的变化情况。可见低优先级的业务所使用的波长数目是随业务量的大小变化的。初始分配了 3 个波长给低优先级使用，低优先级使用的平均波长数为 3.2。

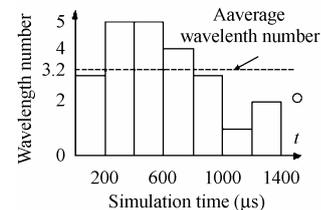


图 3 DLPQ 算法时低优先级所使用波长的变化情况

Fig.3 Wavelength number used by class1 of DLPQ QoS scheme

FLPL, DLPQ, IDLPQ 算法的高、低优先级和平均丢失性能如图 4-图 6 所示。与 FLPL 算法相比，DLPQ 和 IDLPQ 算法中的高优先级性能更好；而且业务量强度越大时，效果越明显；这是因为业务量强度较大时，DLPQ 和 IDLPQ 算法中的高优先级业务可以分配到更多的独占波长，而 FLPL 算法中的高优先级业务只是在波长不够用时才使用低优先级的波长。由于 DLPQ 和 IDLPQ 算法在高优先级业务量小时可以增加低优先级所使用的波长，所以 DLPQ 和 IDLPQ 算法中的低优先级和平均性能明显要好于 FLPL 算法。这说明采用动态的波长分集不仅能改善高优先级的性能，还能改善低优先级以及总体的性能。值得注意的是，DLPQ 算法中一个时段内波长的动态分配主要是根据上一个时段内业务的变化情况，实际上涉及了业务的预测问题，这里没有采用自适应等复杂的预测算法，仅采用了相对简单可行的控制算法以说明 DLPQ 算法的主要思想，在实际当中可以调整  $k$  和  $\Delta$  的取值来有效控制波长分配变化的灵敏度和频率。而 IDLPQ

是根据业务变化对各个优先级使用的波长进行了调整才调度队列中的 BHP, 调整时采取了高优先级优先的策略, 所以 IDLPQ 算法中高优先级有更小的突发丢失率。

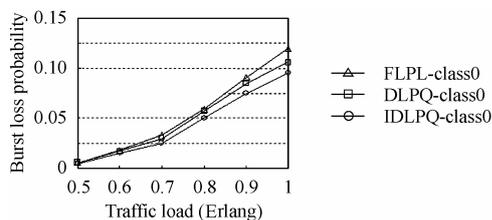


图 4 FLPL, DLPQ 和 IDLPQ 算法 class0 的丢失性能

Fig.4 Class0 loss probability for FLPL, DLPQ and IDLPQ QoS schemes

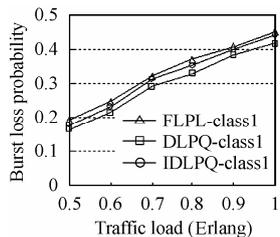


图 5 FLPL, DLPQ 和 IDLPQ 算法 class1 的丢失性能

Fig.5 Class1 loss probability for FLPL, DLPQ and IDLPQ QoS schemes

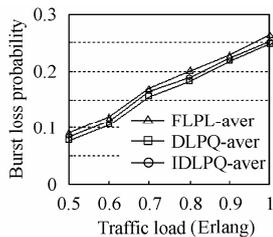


图 6 FLPL, DLPQ 和 IDLPQ 算法平均丢失性能

Fig.6 Aaverage loss probability for FLPL, DLPQ and IDLPQ QoS schemes

#### 4 结束语

本文对现有 OBS 网络中支持 QoS 策略进行了总结, 并从波长分集的思想出发, 提出了几种适合 OBS 的 QoS 算法。这几种算法是通过为不同优先级的业务分配不同的波长的思想来区分服务的, 而且可以根据各个优先级业务的变化情况, 动态地调整各个优先级的业务使用的波长数目。通过仿真, 进行了性能分析和比较。说明所提出的算法可以更好地

失率。

#### 参 考 文 献

- [1] Qiao C, Yoo M. Optical burst switching (OBS)—A new paradigm for an optical Internet. *Journal of High Speed Networks*, 1999, 8(1): 69–84.
- [2] Qiao C, Yoo M. Supporting multiple classes of services in IP over WDM networks. *Global Telecommunications Conference, Brazil*, 1999, 1(b): 1023–1027.
- [3] Qiong Zhang, Vokkarane V M, Biao Chen, *et al.* Early drop scheme for providing absolute QoS differentiation in optical burst-switched networks. *Workshop on High Performance Switching and Routing (HPSR)*, 2003, 6(1): 153–157.
- [4] Vokkarane V M, Zhang Q, Chen B. Generalized burst assembly and scheduling techniques for QoS support to optical burst-switched networks. *Global Telecommunications Conference, Taipei*, 2002, 3(1): 2747–2751.
- [5] Long Keping, Rodney S T, Wang Chonggang. A new framework and burst assembly for IP DiffServ over optical burst switching networks. *Global Telecommunications Conference, San Francisco*, 2003, 3(1): 3159–3164.
- [6] Turner J S. Terabit burst switching. *Journal of High Speed Networks*, 1999, 8(1): 3–16.
- [7] Hashiguchi T, Wang X, Morikawa H, *et al.* Burst assembly mechanism with delay reduction for OBS networks. *Conference on the Optical Internet (COIN)*, Melbourne, 2003, 2(1): 664–666.

王 雄: 男, 1980 年生, 硕士生, 研究方向为光通信网络.

王 晟: 男, 1971 年生, 教授, 主要研究方向为光通信网络.

谭 伟: 男, 1975 年生, 博士, 研究方向为 OBS 网络.

提供区分服务, 同时有效提高信道的利用率, 降低整体的丢