

BM-VF-SBD: 一种支持 QoS 的光突发交换数据信道调度算法¹

阳小龙*** 张敏** 李乐民*

*(电子科技大学宽带光纤传输与通信网络技术重点实验室 成都 610054)

** (重庆邮电学院光通信研究所 重庆 400065)

摘要: 在光突发交换 (OBS) 网络中, 数据信道的调度算法是一个关键问题。然而, 当前的调度算法大多只强调带宽利用效率, 而忽略了 QoS 支持。该文提出了一个算法 BM-VF-SBD, 其基本思想为: 若所有信道上没有一个 Void 能容纳新突发, 则搬移一些突发到别的信道后, 再为新突发分配信道资源; 若还失败, 则再选择性地丢弃一些低优先级的突发, 重复前面操作。它是利用 BM, VF 和 SBD 3 种机制减少带宽碎片, 支持 QoS。若以平衡二叉树组织 Void 和突发相关信息, 它的计算复杂度与 LAUC-VF 和 ODBR 接近, 小于 $O((2w+1)\log w)$ 。仿真表明它在带宽碎片率和突发损失率 (包括总的和各个优先级的) 上优于 LAUC-VF 和 ODBR。

关键词: 信道调度, 波长变换, 光突发交换, 服务质量 (QoS)

中图分类号: TN919.2 **文献标识码:** A **文章编号:** 1009-5896(2004)10-1534-06

BM-VF-SBD: A Data Channel Scheduling Algorithm to Support QoS for Optical Burst Switching Networks

Yang Xiao-long*** Zhang Min** Li Le-min*

*(Key Lab of Broadband Optical Fiber Transmission and Communication Networks, UEST of China, Chengdu 610054, China)

** (Optical Comm. Lab, Chongqing Univ. of Posts & Telecomm., Chongqing 400065, China)

Abstract In the Optical Burst Switching (OBS) network, the data channel scheduling is one of key issues. Bandwidth efficiency and QoS support are the two concerned focuses. However, the existing algorithms pay more attentions to bandwidth efficiency. In this paper, BM-VF-SBD, an efficient data channel scheduling algorithm is developed. It effectively integrates void filling, burst migration and selective burst discard to reduce the bandwidth fragment and support QoS. Its basic idea is in that a new burst is scheduled by migrating some bursts to other channels if none of voids in any channels can accommodate it; otherwise repeating the above processes after selectively dropping some bursts. Meanwhile under the balanced binary search tree data structure, its computational complexity will be $O((2w+1)\log w)$ at most, and be close to that of LAUC-VF and ODBR. In the proposed algorithm, burst migration plays a key role in the improvement of bandwidth efficiency while selective burst discard has great effects on the two sides. The simulation results show that it performs much better than LAUC-VF and ODBR in burst loss probability (overall or individual) and bandwidth fragment ratio.

Key words Channel scheduling, Wavelength conversion, Optical Burst Switching(OBS), QoS

¹ 2003-06-05 收到, 2003-09-03 改回
国家 863 计划项目 (2002AA122021) 资助课题

1 前言

在光突发交换 (OBS) 网络中, 光突发由突发控制分组 (BCP) 和相应的突发分组 (BP) 组成, 它们分别在相互独立的控制信道 (CC) 和数据信道 (DC) 上传输^[1]. 但 BCP 先于 BP 一定偏移时间传输, 以便能在 OBS 中间节点完成 BCP 处理和 BP 资源预留. 这里, BP 资源预留须通过数据信道调度来实现. 然而光突发同 IP 分组一样, 突发性地注入网络^[2], 因此在 BP 调度后, 数据信道上将留下许多带宽碎片, 即“空隙” (Void). 如何有效利用它, 提高带宽利用率, 则是调度算法关心的首要问题. 此外, 如何提供 QoS 保证, 则是算法关心的另一个重要问题.

现有算法中, 与非 VF(Void Filling) 类算法 (如: Horizon^[3]) 相比, VF 类算法 (如: LAUC-VF^[1] 和 Best-fit^[4] 等) 因可有效利用 Void, 提高带宽利用率, 其使用更广泛. 然而, 它们在某些情形下会失效, 如图 1 所示: 当新 BP 与 BP4 和 BP7 部分重叠时, 即使 DC0 和 DC3 上的 Void 与它长度相当, 新 BP 也无法利用这些碎片, 它只好被丢弃. 为此, 文献 [5] 提出了两种基于 BR(Burst Rescheduling) 的算法 ODBR(On-Demand Burst Rescheduling) 和 ABR (Aggressive Burst Rescheduling). 然而它们仅对最近突发 (如图 1 的 BP9 和 BP10) 操作, 而非 Void 前后的任何突发. 因此, 对图 1 而言, 它们同样会失效. 为此, 本文在 VF 和 BR 基础上提出了 BM-VF(Burst Migration with Void Filling) 机制, 使之突发搬迁与 Void 填充并重: BP7 搬迁到 DC3 后, DC0 上的 Void 就可以分配给新 BP.

然而 BM-VF 在某些情形同样会失效, 如图 2 所示: BP4 和 BP8 无法搬移到其它位置. 为此, 文献 [6] 提出的 BS (Burst Segmentation) 解决办法. 但是它的额外开销较大; 另外, 它同其它算法一样, 只强调带宽利用效率, 而忽略 QoS 保证. 为此本文提出了一种简单的 SBD(Selective Burst Discard) 机制: 丢弃低优先级突发, 以调整相应的 Void 间隔, 减少高优先级突发丢弃概率.

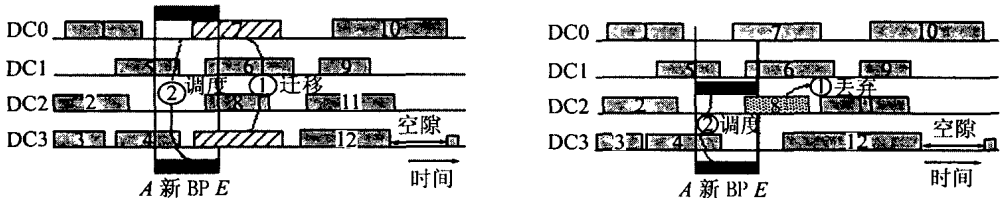


图 1 VF 类算法失效的情形 (通过 BM-VF 机制改进) 图 2 BM-VF 机制失效的情形 (通过 SBD 机制改进)

本文调度算法 BM-VF-SBD (Burst Migration with Void Filling based on Selective Burst Discard) 集成了 VF, BM 和 SBD 3 种机制. 本文根据计算几何^[4] 提出了一种简单的计算方法, 就能快速而准确地为突发分配到资源. 若采用平衡二叉树, 该算法的复杂度小于 $O((2w+1)\log w)$. 最后本文从带宽碎片率和突发损失率 (包括总的和各个优先级的) 两个方面对该算法进行了仿真, 结果表明: 该算法能充分利用带宽, 减少带宽碎片, 同时支持区分服务.

2 数据信道调度算法 BM-VF-SBD

2.1 调度问题的描述及 Void 映射平面

根据计算几何方法, 文献 [4] 将 Void 映射成二维坐标平面的点 (s_i, e_i) , x 坐标 s_i 表示 Void 开始时刻, y 坐标 e_i 表示 Void 结束时刻, 如图 3 和图 4 所示. 由于 Void 的结束时刻肯定不小于开始时刻, 因此 Void 映射点不会落在 $s = e$ 下面, 即区域 V. 新突发也可映射成一个点

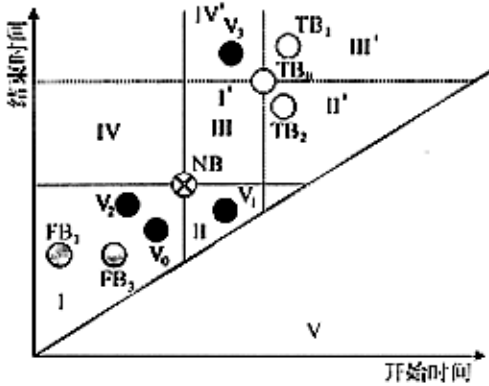


图 3 图 1 对应的映射平面及突发搬迁操作

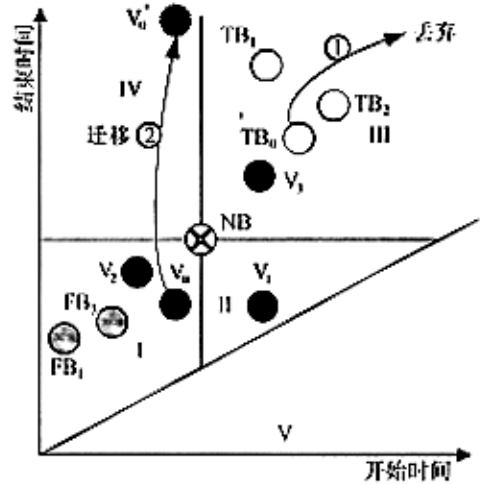


图 4 图 2 对应的映射平面及突发丢弃操作

(A, E)，其中 A 和 E 分别代表突发的开始和结束时刻，并且 $E = A + L_{burst}$ ， L_{burst} 为突发长度。该调度问题就转化为在 $s = e$ 的上半部分区域内搜索满足“ $s \leq A$ 且 $e \geq E$ ”的 Void 映射点问题。

点 (A, E) 坐标线将 $s = e$ 的上半部分划分成 4 个区域 I，II，III 和 IV。区域 IV 内的任何一个映射点均满足“ $s \leq A$ 且 $e \geq E$ ”，因此都是该问题的解，称它为绝对可调度区域。然而区域 I 内的点虽 $s \leq A$ ，但 $e \leq E$ ，即表明某一已调度突发必定与新突发的尾部重叠，称该已调度突发为尾叠突发 TB(Tail-overlapping Burst)。同样，区域 III 内的点虽 $e \geq E$ ，但 $s \geq A$ ，即表明某一已调度突发必定与新突发的头部重叠，称该已调度突发为前叠突发(Front-overlapping Burst, FB)。而区域 II 内的点不但 $s \geq A$ ，而且 $e \leq E$ ，这表明新突发与其他突发既头部重叠，又尾部重叠。这 3 个区域的映射点可通过一定机制转化为调度问题的解，因此称它们为条件可调度区域。这样，该问题就可用 Void 映射点的变换操作来描述，调度算法的设计重心就是如何高效地将 Void 映射点从条件可调度区域调整变换到绝对可调度区域。

2.2 BM-VF-SBD 算法

不同于当前一些算法 (如 Best-fit^[4], ODBR^[5] 和 SFMOC-VF^[6] 等)，BM-VF-SBD 算法将突发丢弃率、信道利用率和 QoS 支持等性能一并考虑。它的基本思想比较简单：若所有信道上没有一个 Void 能容纳新突发，则搬移一些突发到别的信道，再为新突发分配信道资源；若仍然失败，则再选择性地丢弃低优先级的突发后，重复前面操作。在该算法中，带宽利用效率是通过 BM-VF 来改善的，而 OBS 网络 QoS 支持则是通过 SBD 来实现。

根据该算法的基本思想，它的操作可分为 3 个阶段，各个阶段与 Void 映射点分布有关。阶段 1 以 Void 填充为标志，阶段 2 以突发搬迁为标志，阶段 3 以选择性突发丢弃为标志，其具体操作如下所述。

阶段 1 搜索区域 IV 是否有 Void 映射点，若有，则按一定的选择规则 (如：Min-EV, Min-SV 等^[4]) 从中选择一个 Void 分配给新突发。若没有，则进行阶段 2 的搜索和搬迁两个操作。搜索是在重叠突发 FB_i 或 TB_i 坐标线重新划分的区域 IV 中进行。若搜索到一个 Void 映射点 V_j ，则将 FB_i 或 TB_i 搬迁到信道 DC_j ($i \neq j$) 上。通过这搬迁操作，Void 映射点的区域归属得以重新调整，如 V_i 调整到新的位置，然后再返回阶段 1。例如：在图 3 中，只有 V_3 落在 FB_0 或 FB_2 新划分的区域 IV 中，这意味着图 1 中 BP7 或 BP8 能搬迁到信道 DC_3 ，这样就能将 DC_0 或 DC_2 分配给新突发。若没能在重新划分的区域 IV 中搜索到任何 Void 映射点，则进行阶段 3

完成两个操作, 即突发优先级比较、及选择性丢弃突发。若新突发的优先级低于重叠突发 FB_j 或 TB_j , 则直接丢弃新突发, 算法结束; 否则丢弃 FB_j 或 TB_j 。通过这丢弃操作, Void 映射点的区域归属也得以重新调整, 之后再返回阶段 1。例如: 在图 4 中, 丢弃 TB_0 (即图 2 中的 BP7) 后, V_0 移动到了区域 IV 里 V'_0 位置, 这样就能将 DC0 分配给新突发。

在丢弃操作时, 若候选者为多个优先级相同的重叠突发 FB_j 或 TB_j , 则按如下的规则选择: 在被选中突发丢弃前它占用的带宽资源与丢弃后被重新利用的带宽资源之差最小, 即因被选中突发丢弃而造成的带宽资源浪费最小。这里, 带宽资源浪费程度可用传输效率 η_{trans} 和带宽碎片率 η_{frag} 来衡量, 它们可简单地定义如下:

$$\eta_{trans} = L_{burst}/(L_{burst} + L_{void}) \text{ 和 } \eta_{frag} = 1 - \eta_{trans}$$

其中 L_{burst} 和 L_{void} 分别表示重叠突发长度和 Void 大小。例如: 假定图 2 中的 BP4, BP7 和 BP8 具有相同的优先级, 均比新突发的低, 按照上述定义, 丢弃它们造成的资源浪费程度如表 1 所示。显然, 丢弃 BP8 比较合适, 因为丢弃前后 η_{trans} 减少最小, 而且 η_{frag} 增加也最小。

表 1 图 2 中不同重叠突发丢弃前后所造成的资源浪费程度 ($L_{newburst} = 5$ 单元)

数据信道	L_{burst}	L_{void}	丢弃前		丢弃后	
			η_{trans}	η_{frag}	η_{trans}	η_{frag}
DC0	$L_{TB_0} = 7$	$L_{V_0} = 4$	7/11	4/11	5/11	6/11
DC2	$L_{TB_2} = 5$	$L_{V_2} = 5$	5/10	5/10	5/10	5/10
DC3	$L_{TB_3} = 6$	$L_{V_3} = 5$	6/11	5/11	5/11	6/11

2.3 算法的计算复杂度

在该算法中, 所涉及的信息较多, 有新突发、以及各信道上与之相关的 Void、重叠突发的开始和结束时刻信息。很显然, 信息的数据结构会对算法的计算复杂度影响较大。在最坏情形下, 计算复杂度可达 $O(w^4)$, 其中 w 为 OBS 网络中数据信道总数目。若采用文献 [4] 提出的平衡二叉树, 那么该算法中一项主要操作, 即搜索, 它的复杂度将从 $O(w)$ 减为 $O(\lg w)$, 这样整个算法大幅度地降为 $O((2w+1)\lg w)$ 。然而, LAUC-VF 和 ODBR 的复杂度, 根据文献 [1, 5] 在没有光纤延迟线的情况下, 最大为 $O(w^2)$ 。因此, BM-VF-SBD 在一定的数据结构下, 其计算复杂度与 LAUC-VF 和 ODBR 的非常接近。

3 性能仿真

这里, 将通过仿真来评估算法 BM-VF-SBD。为简单起见, 以边缘节点出口信道的调度为仿真对象, 如图 5 所示。当然也可以是其他 OBS 节点的出口信道, 这不影响仿真结论。假定该出口链路由 8 个数据信道和 1 个控制信道组成, 带宽均为 1 Gbps。有 3 个优先级的突发, 最高优先级为 1。在边缘节点, 突发以速率为 λ 的 Poisson 过程进行装配注入 OBS 网络, 而且突发长度为指数分布, 均值为 L 。在总的输入负载中, 各类突发所占比例相等, 均为 1/3。所有突发 BP 和 BCP 间的偏移时间 τ 相同。为消除其它因素对突发丢失概率 (Burst Loss Probability, BLP) 的影响, 要求 L/τ 大于数据信道数与控制信道数之比, 并假定突发丢失只是数据信道拥塞造成。

本文以总的、不同优先级的 BLP 和带宽利用率作为衡量指标, 将 BM-VF-SBD 与 ODBR 和 LAUC-VF 进行性能比较。图 6 给出了这 3 种调度算法总的 BLP 随输入负载的变化情况。很明显, 3 种算法总的 BLP 都是随负载的增加而增加, 原因在于当负载越重时, 为新突发搜索并分配可用信道将变得很困难, 新突发就很可能被丢弃。然而总体上, BM-VF-SBD 的总的 BLP 优于 ODBR 和 LAUC-VF。而且随着负载的增加, 这种趋势越明显, 因为 BM-VF-SBD

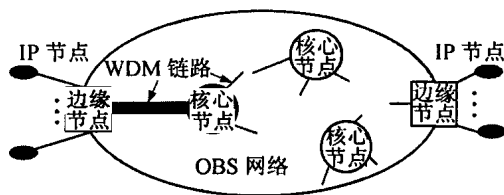


图 5 以边缘节点与核心节点间的信道调度为仿真对象

通过 Void 填充和突发搬迁, 使得信道上传输的突发更加紧凑, 信道的可用带宽间隙比较宽, 从而为后续突发留下更多的被成功调度的机会。

接下来研究在不同负载下, 这 3 种算法在每类突发的 BLP 上表现。在这次仿真实验中, 观察到: ODBR 和 LAUC-VF 算法的各类突发的 BLP 随负载的变化情况基本上相同, 因此为了简洁, 图 7 只给出了最高优先级突发的 BLP, 然而 BM-VF-SBD 则不一样, 各类突发的 BLP 随负载的变化情况各不相同, 且差异较大, 如图 8 所示。这表明本文提出的算法具有 QoS 支持能力, 能按照突发类别提供有区分的承载能力, 相反 ODBR 和 LAUC-VF 就不具有这些能力。

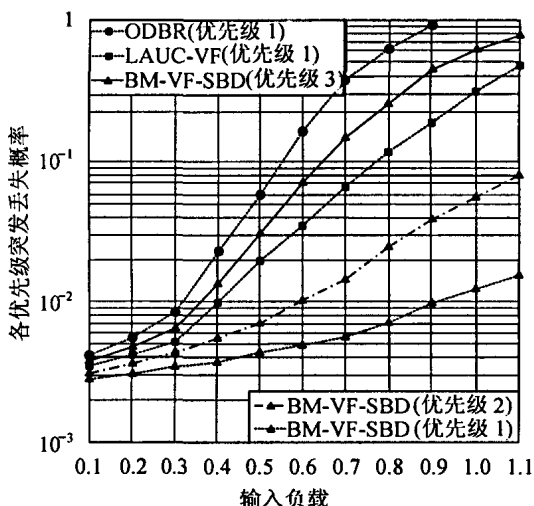
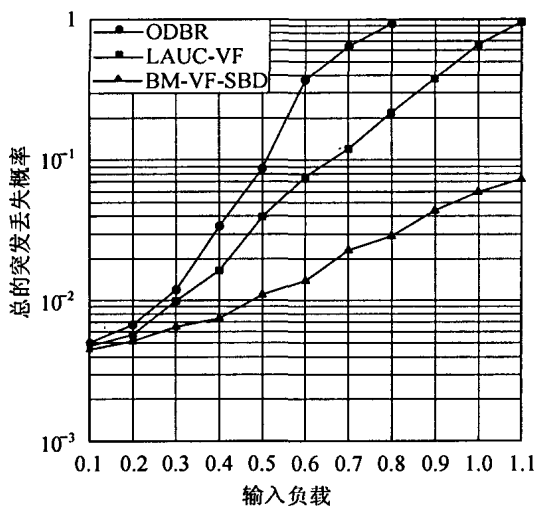


图 6 不同负载下总的突发丢失概率比较 (假定 $L/\tau=10$)

图 7 不同负载下每类突发的丢失概率比较 (假定 $L/\tau=10$)

此外, 从图 7 还可看到: BM-VF-SBD 算法的优先级 2 的 BLP 比 ODBR 和 LAUC-VF 算法的优先级 1 的 BLP 低得多; 而且甚至 BM-VF-SBD 算法的优先级 3 的 BLP 也比 ODBR 算法的优先级 1 的 BLP 低。BM-VF-SBD 这种良好的表现得益于 SBD 机制。因为在数据信道拥塞时, 将按照优先级顺序和一定选择规则丢弃一些突发。这样就有一部分占用资源被释放, 通过 Void 填充和突发搬迁的重新整合后, 其可用带宽间隙变宽, 因此能以较高的成功概率再次分配给新突发。

现在我们来考察该算法的带宽利用效率, 可以用 η_{trans} 或者 η_{frag} 来衡量, 两者等效。这里, 我们用 η_{frag} 。根据前面定义, 带宽碎片率 η_{frag} 可按下式测量得到:

$$\eta_{frag} = (nT - \sum_m L_{burst_m}) / (nT)$$

其中 n 表示数据信道数目, m 表示在观察期间 T 内链路上正在传输的突发数目。上式的 η_{frag} 从时间轴上描述了数据信道资源的浪费程度。假定观察时间为 10 倍突发长度。如图 8 所示, 当

输入负载比较轻 (低于 0.4) 时, 3 种算法的 η_{frag} 比较接近。但是当负载高于 0.5 时, 它们各自 η_{frag} 的差异越来越明显。总体上讲, BM-VF-SBD 算法的 η_{frag} 随负载的增长速率小于其他两

种算法。即在同等条件下, 它成功调度的突发比其他两种算法的要多。BM-VF-SBD 较低的 η_{frag} 得益于 BM 和 SBD 操作后 Void 映射点区域归属的重新调整。

通过前面 3 个性能衡量指标的比较, 在这 3 种信道调度算法中, BM-VF-SBD 算法性能较好, 它既能有效而充分地利用数据信道的传输能力, 又根据不同优先级突发的传输要求给予有区分的承载服务。根据图 6~图 8 仿真结果的分析, 我们可知: 在该算法中, 带宽利用率提高是通过 BM 和 VF 对 Void 映射点的区域归属的重新调整来实现的, 而 OBS 网络 QoS 支持则通过 SBD 来增强的。

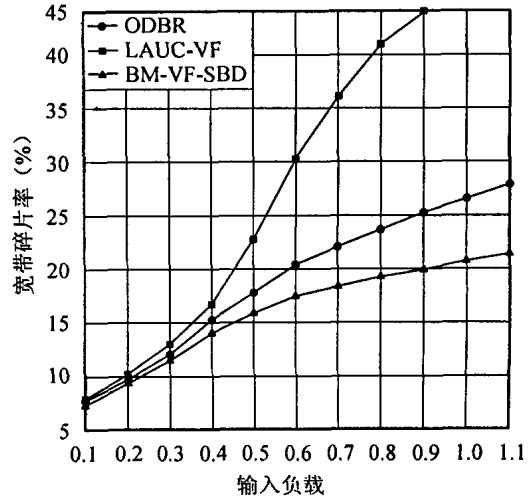


图 8 不同负载下带宽碎片率的比较
($L/\tau = T/L = 10$)

4 结束语

本文在借鉴了文献 [1, 4, 5] 算法思想的基础上, 提出了一种的数据信道调度算法, 即 BM-VF-SBD。它的基本思想比较简单: 若所有信道上没有一个 Void 能容纳新到光突发, 则搬移一些突发到别的信道后, 再为新到光突发分配信道资源; 若仍然失败, 则再选择性地丢弃一些低优先级的突发后, 重复前面操作。它将突发搬移、空隙填充和选择性突发丢弃 3 种机制有机地融为一体, 既能有效减少数据信道带宽碎片、提高信道带宽利用率, 又能支持 QoS。经仿真比较, 它在带宽碎片率和突发损失率 (包括总的和各个优先级的) 上均明显地优于 LAUC-VF 和 ODBR。

参 考 文 献

- [1] Xiong Y, Vandenhoute M, Cankaya H. Control architecture in optical burst-switched WDM networks. *IEEE J. on Selected Areas in Communications*, 2000, 18(10): 1838-1851.
- [2] Yu X, Chen Y, Qiao C. A study of traffic statistics of assembled burst traffic in optical burst switched networks. Proc. of SPIE Opticomm'02, Boston, July 2002: 149-159.
- [3] Turner J. Terabit burst switching. *Journal of High Speed Networks*, 1999, 8(1): 3-16.
- [4] Xu J, Qiao C, Li J, Xu G. Efficient channel scheduling algorithms in optical burst switched networks. Proc. of IEEE INFOCOM'2003, San Francisco, USA, April 2003, vol.3: 2268-2278.
- [5] Tan S K, Mohan G, Chua K C. Algorithms for burst rescheduling in WDM optical burst switching networks. *Computer Networks*, 2003, 41(1): 41-55.
- [6] Vokkarane V, et al.. Burst segmentation: An approach for reducing packet loss in optical burst-switched networks. Proc. of IEEE ICC'2002, New York City, April 2002, vol.5: 2673-2677.

阳小龙: 男, 1970 年生, 博士, 主要研究方向为 WDM 全光网络。

张敏: 女, 1972 年生, 硕士生, 主要研究方向计算机网络及信息安全。

李乐民: 男, 1932 年生, 中国工程院院士, 教授, 博士生导师, 主要研究方向是信息传输与通信网。