

一种新型的 OBS 网络单向资源预留方法

袁 巍 李津生 洪佩琳 杜 平

(中国科学技术大学信息网络实验室 合肥 230027)

摘要: 光突发交换(OBS)是实现 IP/DWDM 网络的一种极有前途的方案。该文着重讨论了光空发交换(OBS)网络中单向资源预留的方式。文中首先介绍了著名的 JET(Just Enough Time), 提出了一种新的单项资源预留方式 Just Little Time (JLT), 它不但能使用偏转路由, 还能有效利用 Fiter Detag Line (FDL) 延时来弥补偏移时间的不足, 只需使用很小的初始偏移时间, 降低了数据在网络中的延时。随后, 文章讨论了使用偏移时间增量来实现区分服务的方法。计算机仿真结果证明了所提方法的有效性。

关键词: OBS, JET, JLT, 数据信道调度算法, 偏转路由

中图分类号: TP393.14

文献标识码: A

文章编号: 1009-5896(2005)12-1978-05

A Novel One-Way Delayed Reservation Scheme for OBS Networks

Yuan Wei Li Jin-sheng Hong Pei-lin Du Ping

(InfoNet Lab, University of Science and Technology of China, Hefei 230027, China)

Abstract Optical Burst Switching (OBS) is a promising solution for realizing IP/DWDM. This paper focuses on the one-way delayed reservation scheme. As a well-known scheme, JET (Just Enough Time) is described first. Then a new scheme called Just Little Time(JLT) is introduced in this paper. JLT is compatible with deflection routing and has no offset redundancy. To support service differentiation, offset time increments are adopted by JLT. JLT proves to be effective via computer simulations.

Key words Optical Burst Switching (OBS), Just Enough Time (JET), Just Little Time (JLT), Data channel scheduling algorithm, Deflection routing

1 引言

随着 DWDM(Dense Wavelength division Multiplexing)技术的发展, 光纤的传输容量有了极大地提高。现在的骨干网络采用了 IP / ATM / SDH(SONET) / WDM 的网络结构模型。这种结构设备成本高、开销大, 不便于网络扩展, 传输速率受电处理的限制。为了有效利用光纤的巨大传输带宽, 人们提出全光传输。下一代光传输网络应采用 IP / WDM 结构, 即无需光电间的转换, 直接在光层传输 IP 分组。光传输网络主要有 3 种交换方式: 波长路由(Wavelength routing), 光分组交换(Optical Packet Switching, OPS)和光突发交换(Optical Burst Switching, OBS)。光突发交换兼具电路交换和分组交换的优点, 受到了最广泛的关注。

光突发(Optical Burst)就是具有相同目的地址和特性(如 QoS 参数等)的多个 IP 分组的组合, 它分为数据突发单元(Data Burst, DB)和控制头(Burst Head Packet, BHP)两部分。DB 是

Burst 的数据信息, 而 BHP 是 Burst 的控制信息, 它包含 Burst 的源和目的地址、DB 的长度、偏移时间(Offset time)和 QoS 参数等信息。在 OBS 网络边缘, BHP 比 DB 提前一段时间或同时出发, 在核心路由器上进行处理(在电域中进行)并对光交换矩阵进行配置(即对 DB 所需资源进行预留)以转发 DB。转发完毕, 核心路由器释放 DB 使用的资源^[1]。这里资源是指可使用的输出数据信道和使用的的时间。

OBS 网络一般使用单向资源预留(One-way reservation)方式(比如 JET 或 TAG), 它具有较小的端到端延时和较高的信道利用率^[2]。JET 方式一直是研究的热点, 与 TAG(Tell And Go)相比, 它具有无需输入 FDL(Fiber Delay Line)缓存等优点^[3]。但是, 当使用偏转路由(Deflection routing)^[4]时, JET 的偏移时间难以确定; 另外, JET 偏移时间往往有富余, 因而带来不必要延时。为避免在使用偏转路由时偏移时间难以确定的问题并减小 DB 在 OBS 网络中的延时, 本文提出了一

种新的单向资源预留方式 JLT(Just Little Time)。文章接下来的部分首先简要介绍 JET 的基本原理, 然后讨论 JLT, 最后给出网络仿真的结果并进行性能评价。

2 JET 简介

OBS 网络由光核心路由器、边缘路由器及光链路组成。在骨干网络边缘, 来自接入网的 IP 分组在边缘路由器中被汇聚(Assemble)成突发单元(Burst), 通过核心路由器的转发在 OBS 骨干网络中传输, 再在目的端的边缘路由器中拆分(Disassemble), 恢复成一个个的 IP 分组进入对方接入网。Burst 的控制信息(BHP)与数据信息(DB)分别在控制信道和数据信道(每个信道对应一个波长)上传输。对于基于 JET 的 OBS 网络, BHP 比 DB 提前一段时间(偏移时间)从边缘路由器发出, 光核心路由器收到 BHP 后, 首先依据其携带的地址信息进行路由计算, 然后运行数据信道调度算法^[1,5]来选择 DB 的输出数据信道。为解决信道竞争问题, 核心路由器使用 FDL 阵列对 DB 进行缓存^[7]。为提高带宽利用率, 基于 JET 的 OBS 网络使用资源延时预留(Delayed reservation)的方法, 即仅当 DB 到达时才开始资源的预留^[3]。当 DB 转发结束, 释放资源给其他 DB 使用。

图 1 简单示出了核心路由器的结构, 它由 BHP 处理器, FDL 阵列和光交换矩阵以及波长转换器(图中未画)等部件组成。在 BHP 处理器中 BHP 首先经过光电转换变成电信号, 然后进入电缓存器等待进行路由查找和输出信道调度处理, 最后再转换成光信号从输出控制信道输出。控制器根据 BHP 处理结果在 DB 到来前完成对 FDL 阵列以及光交换矩阵进行配置。当 DB 交换完毕, 此配置失效。值得一提的是, 控制信道也需要调度。对于 BHP 的处理细节的讨论超出了本文的范围, 故不再赘述。

FDL 阵列有多种结构, 比如环结构, 线结构和多级结构^[2,7]。本文只讨论最简单的等差延时的线结构 FDL 阵列。这种 FDL 阵列提供的延时可为 $D, 2D, \dots, ND$, 其中 D 为最小延时单元(单位微秒), N 为阵列中 FDL 数目。

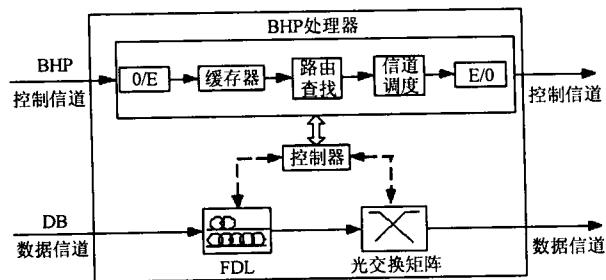


图 1 核心路由器结构示意图

信道调度算法也是当前研究的一大热点。所谓数据信道调度就是根据 BHP 携带的关于 DB 的信息(如 DB 长度和偏移时间等), 选择一条在 DB 到达光交换矩阵时或经过 FDL 阵列

的延时(文中称为 FDL 延时)之后可用的数据信道进行输出; 当无可用信道时, 将 Burst 丢弃。本文将不经过 FDL 阵列延时的情况视为 FDL 延时时间为 0。

对于 JET, DB 比 BHP 滞后一段时间从边缘节点发出(由此导致的延时称为预传输延时), 文中将这段时间值称为偏移时间 T 。

在 DB 到达每个节点前, BHP 必须有足够时间进行处理并为 DB 预留资源。也就是说 BHP 到达节点时 DB 的滞后时间不能小于 BHP 在节点上的延时(包括排队时间、光电转换、路由查找和信道调度等时间), 否则 Burst 被丢弃(文中称为早到丢弃)。为保证这一点, 偏移时间必须不小于 BHP 在除源边缘节点外的所有节点上延时时间之和。假设 Burst 在 OBS 网络中所经节点数为 H , 在每一个节点上延时时间为 P_i , 则偏移时间 $T \geq \sum_{i=1}^H P_i$ 。实际应用中, 取 $T \geq HP_{\max}$, P_{\max} 为 BHP 在核心节点上延时时间的最大值^[8]。图 2 示出了每一节点上 FDL 延时 D_i 都为 0 的情形。图中, 在 Burst 的传输过程中, DB 的滞后时间不断缩小, 直到离开目的边缘节点时, 其值几乎减为零。图 3 显示了当某个节点上 FDL 延时不为 0 的情形。图中, 节点 1 上的 FDL 延时不为零, 导致 DB 离开目的边缘节点时仍滞后 BHP 较长时间。文中称这种现象为延时冗余, 它导致 Burst 的不必要延时。为减小 Burst 的网络延时, 偏移时间可以取得更小一些。但是, 边缘节点无法预知 Burst 在网络中所经过每一节点上 FDL 的使用情况, 因此无法避免延时冗余。JET 的另一个问题在于, 当使用了偏转路由, 边缘节点无法预知 Burst 在网络中经过的路由和节点数, 从而难以准确选取 T 的大小。如图 4 所示, 路由 1 是工作路由, 路由 2 是偏转路由, 路由 2 比路由 1 的节点数多。源边缘节点根据路由 1 来确定偏移时间的大小, 当由于某种原因改用路由 2 后, 原来的偏移时间可能不够大, 比如 DB 在离开节点 5 时的滞后时间已变为 0, 从而导致 Burst 的早到丢弃^[1,8]。

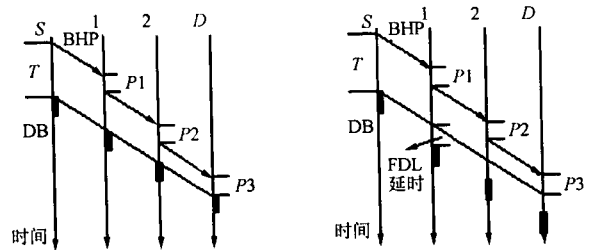


图 2 JET(FDL 延时等于 0) 图 3 JET(FDL 延时不等于 0)

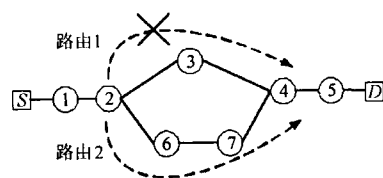


图 4 偏转路由示意图

3 JLT 的基本原理

要解决以上两个问题，偏移时间的选取应该在途经的每个节点上进行而不是在源边缘节点上一次性选定。基于这一考虑，本文提出了 JLT(Just Little Time)方法。JET 中的偏移时间是一种全局偏移时间，它用来补偿 BHP 在所有节点(源边缘节点除外)上的延时时间；JLT 使用的是局部偏移时间，它仅仅用于补偿 BHP 在一个节点上的延时时间。核心节点的调度算法不仅要解决信道竞争问题，而且要保证下一跳节点的偏移时间足够大以防止早到丢弃，因此 JLT 消除了由于使用偏转路由导致的偏移时间难以选取的问题；JLT 使用很小的初始偏移时间，且节点上的 FDL 延时可用于补充下一跳节点的偏移时间，因此 JLT 有效地减小了 Burst 在网络中的延时。此外，为提高带宽利用率，JLT 同样使用延时预留的方法。

3.1 局部偏移时间

JET 使用 FDL 延时纯粹是用于解决输出信道竞争问题；而 JLT 使用的 FDL 延时有两个作用，除解决输出信道竞争，还用于补充偏移时间的不足。如图 5 所示，设每一节点上的偏移时间为 $T_i (i \geq 1)$ ，BHP 每经过一个节点，都必须更新其携带的偏移时间值(见式(1))作为下一跳节点的偏移时间。

$$T_{i+1} = T_i - P_i + D_i \quad (1)$$

其中 D_i 是 DB 在节点 i 上经历的 FDL 延时大小， P_i 是 BH 在节点 i 上的延时， T_{i+1} 是下一跳节点的偏移时间。

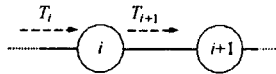


图 5 偏移时间示意图

由于 JLT 的偏移时间用于补偿 BHP 在节点上的延时，为避免早到丢弃，每一节点的偏移时间至少为 P_{max} (条件(a))。通常，Burst 途径的第一个核心节点上的偏移时间 T_1 (初始偏移时间)取为 P_{max} 。

为满足条件(a)，信道调度算法不仅要考虑数据信道冲突问题，还要考虑下一跳节点的偏移时间大小，当下一跳节点的偏移时间不足时，必须对其补充。因此，JLT 的信道调度算法首先从 FDL 阵列的所有延时值中选择一个合适的范围(延时集合)，然后才从此范围中选择 FDL 延时值用于对 DB 延时。正是延时集合的选取保证了条件(a)得到满足。

假设 FDL 阵列提供的延时值为 $0, D, 2D, \dots, ND$ ，且 $(M-1)D < P_{max} \leq MD$ ， M 为自然数，则延时集合的选择方法如表 1 所示。其说明如下：

(1) 当 $T_i \geq 2P_{max}$ 时，有 $T_i - P_i \geq P_{max}$ 。这意味着即使 FDL 延时为 0，下一跳节点的偏移时间 T_{i+1} 也足够大。这时，可供调度算法使用的延时集合为 $\{0, D, 2D, 3D, \dots, ND\}$ 。这种情况下，FDL 延时纯粹用于解决输出信道竞争问题。

表 1 延时集合的选取

偏移时间 T_i 满足的条件	延时集合的选择
(a) $T_i \geq 2P_{max}$	$\{0, D, 2D, \dots, ND\}$, 即全集
(b) $P_{max} \leq T_i < 2P_{max}$	$\{MD, (M+1)D, \dots, ND\}$

(2) 当 $P_{max} \leq T_i < 2P_{max}$ 时，由式(1)可知，为满足条件(a)，可使 $D_i \geq P_{max}$ 。为此，调度算法选择延时集合为 $\{MD, (M+1)D, \dots, ND\}$ ，其中 M 为自然数。这种情况下 FDL 延时不仅用于解决信道竞争问题，还用于补充偏移时间的不足。

3.2 数据信道调度算法

如上所述，JLT 的数据信道调度算法不同于 JET，它在选择可用信道之前要先根据 BHP 携带的偏移时间值选择延时集合。以下示出了一种用于 JLT 的，基于 FF(First Fit)算法^[1]的数据信道调度算法。其中函数 $channel_search(x, L)$ 用于为长度为 L 的 DB 选择在 $[x, x+L]$ 时间内可用的输出数据信道，当无可用信道时返回 -1，否则返回可用信道号。需要指出的是，对于目的边缘节点，由于没有下一跳节点，无需选择延时集合。

Begin { FF-VF-JLT 调度算法 }

$n = 0; x =$ 到达时间;

if ($P_{max} \leq T_i < 2P_{max}$)

$n = n + M;$

while ($n \leq N$) {

$result_num = channel_search(x, L);$

 if ($result_num \neq -1$)

 {

 报告选择数据信道 $result_num$ 输出，选择

 FDL 阵列的延时为 $n * D;$

 break;

 }

$n = n + 1;$

$x =$ 到达时间 + $n * D;$

 }

if ($n > N$)

 报告无可用数据信道;

End { FF-VF-JLT 调度算法 }

3.3 性能评价

JLT 使用局部偏移时间，核心节点保证下一跳节点的偏移时间足够大，它没有由于引入偏转路由导致的偏移时间难以选取的问题。

OBS 网络中 Burst 的端到端延时可以看成预传输延时、FDL 延时和光纤中传输延时等延时之和，其中传输延时是难以改变的。JLT 的预传输延时远小于 JET，并用 DB 在核心节

点上的 FDL 延时弥补偏移时间的不足，从而有效减少了端到端延时。随着 FDL 延时时间的增加，JLT 在延时性能方面的优势会更明显。

JLT 必须使用 FDL 阵列，而 JET 只是为了改善数据信道冲突情况才使用 FDL 阵列。与不使用 FDL 阵列的 JET 方式相比，JLT 具有更低的突发数据丢失率，但它的节点复杂度更大，成本有所增加。另外，当 $P_{max} \leq T_i < 2P_{max}$ 时，调度算法选择延时集合为 $\{MD, (M+1)D, \dots, ND\}$ 。此时，JLT 解决信道竞争问题的能力略小于拥有相同 FDL 阵列的 JET。随着 FDL 延时时间的增加，这种差距将逐渐减少。

4 区分服务的实现

在 OBS 网络中，如果网络资源不足以满足突发性的高业务量就会导致拥塞。为保证关键业务的服务质量，要求 OBS 网络提供类似 IP 网络的区分服务功能。JET 的一个极大优点是能利用偏移时间增量(Extra offset time)方便地在 WDM 层实现区分服务^[3]。这种方法同样适用于 JLT。下面简单讨论在 JLT 中引入偏移时间增量 Δ 以实现 QoS 的实现方法。

引入 Δ 的 JLT 使用扩展偏移时间 $T_E = T + \Delta$ ， Δ 是对偏移时间的一个微小增量， T 是偏移时间，其大小仍满足条件(a)。对于不同优先级的各类数据，其 Δ 取值与优先级大小成正比。假设 OBS 网络中有 n 类业务，分别为 $C_0, C_1, \dots, C_j, \dots, C_n$ ，其优先级依次递增，则有 $\Delta_j = j \times \Delta_B$ ，其中 Δ_B 是偏移时间增量的最小单位， $0 < \Delta_B \ll T$ 。对于第 j 类数据，其扩展偏移时间 $T_{Ej} = T + \Delta_j$ 。通常取 $T = P_{max}$ ，则有 $T_{Ej} = P_{max} + \Delta_j$ ， $0 \leq \Delta_j < T$ 。为简单起见，下面以网络中仅有两类业务(C_0 和 C_1)的情况为例进行说明。如图 6 所示，0 类业务的 DB 的到达时间尽管早于 1 类业务的 DB，但由于 1 类业务具有更大的偏移时间增量，亦即 1 类业务的 BHP 可以比 0 类业务的 BHP 到达更早，因此，1 类业务的 DB 尽管晚到却更早地得到调度，从而抢占了资源，0 类业务的 DB 只能经过 FDL 阵列的延时等待可用资源或被丢弃。

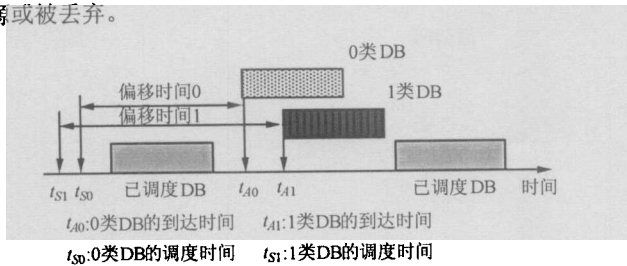


图 6 区分服务的实现原理

引入 Δ 的 JLT 的延时集合选择方法如表 2 所示，其数据信道调度算法与 JLT 类似，由于篇幅限制，不再赘述。

表 2 引入偏移时间增量后 j 类业务的延时集合的选取

偏移时间 T_i 满足的条件	延时集合的选择
$T_i \geq 2P_{max} + \Delta$	$\{0, D, 2D, \dots, ND\}$, 即全集
$P_{max} + \Delta \leq T_i < 2P_{max} + \Delta$	$\{MD, (M+1)D, \dots, ND\}$

注： $(M-1)D < P_{max} + \Delta_j \leq MD$

5 仿真

在这一部分，我们根据使用网络仿真软件 NS 仿真的结果分析 JLT 的性能。仿真使用的网络拓扑借用了美国 NSFNET 网络，如图 7 所示。

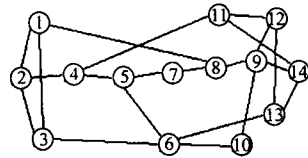


图 7 NSFNET 示意图

最新研究表明，因特网业务具有自相似(self-similar)特性^[9]。为模拟具有自相似性的 IP 业务，仿真中将一组 ON / OFF 源模型(ON/OFF sources model)的输出数据复合为边缘路由器的输入。这种模型交替处于两种状态：ON 状态和 OFF 状态。处于 ON 状态时，业务模型持续产生分组；处于 OFF 状态时，业务模型停止产生分组。两种状态的保持时间 T_{ON} 和 T_{OFF} 服从 Pareto heavy-tail 分布。自相似特性用 Hurst 参数来衡量，最小值为 0.5。在仿真过程中，取 $H=0.85$ 。

OBS 网络中的 DB 平均长度 L_b (微秒)必须满足：(1) $L_b \geq (K-k)L_h/k$ ，(2) $L_b \geq \theta/\epsilon$ 。其中 L_h 为 BHP 的长度， K 和 k 分别为信道总数和控制信道数， θ 为 DB 的保护时间， ϵ 为 Burst 开销的最大值^[1]。仿真中，光纤传输速率取为 10Gb/s，信道数为 32，控制信道数为 2， L_h 为 $0.0512 \mu s$ (即 BHP 分组长 64 byte)，保护时间为 $0.1 \mu s$ ， ϵ 取 0.05，则由(1)(2)得 L_b 至少为 $2 \mu s$ 。 L_b 的实际值可用下式估算：

$$L_b = r(K-k)/\lambda \tag{2}$$

其中 r 为数据信道利用率， λ 为每根光纤的 Burst 到达率^[1,8]。仿真过程中， L_b 值范围是 $5 - 12 \mu s$ 。

P_{max} 的值与 BHP 处理器的处理速度和 Burst 的到达率等因素有关。仿真中取值 $5 \mu s$ ^[8]。

仿真中假定每个核心节点的 FDL 阵列相同，最小延时单元 D 为 $2.5 \mu s$ ， N 为 8。

图 8 显示了与 JET 相比，JLT 在延时性能方面的改善。图中横坐标为网络中 Burst 丢失率(丢弃比特数/所有比特数)，纵坐标是 DB 在网络中平均延时(包括预传输延时和 FDL 延时，单位为微秒)。计算表明，使用 JLT 的 Burst 延时比 JET 平均减少了 14% 左右，最大值达到了 26% 左右。另外，从图中可

以看到, JLT 的 Burst 延时随网络负载的增加而变化不大, 而 JET 的延时性能受网络负载影响较大。这是由于当网络负载增大时, 往往导致 FDL 延时变大, 对于 JLT 而言, 由于 FDL 延时能用于弥补偏移时间, 因此 JLT 能够较好地缓解网络负载增大导致的 Burst 延时变大的问题。

利用偏移时间增量实现区分服务的仿真结果如图 9 所示。为简化起见, 仿真中使用了 3 类业务(类 0-2), 其优先级依次增加, 偏移时间增量取为 $0.5 \mu\text{s}$ 。如图所示, 纵坐标表示各类业务的 Burst 丢失率, 横坐标为信道利用率。图中, 0, 1 和 2 类业务的数据丢失率各不相同, 优先级越大, 丢失率越低。这说明高优先级数据的传输质量比低优先级数据高, 实现了区分服务。

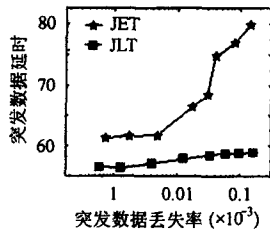


图 8 JET 和 JLT 的延时比较

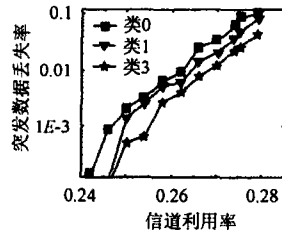


图 9 不同优先级数据的丢失率

6 结束语

本文简要地介绍了 OBS 的基本原理, 详细讨论了单向资源预留方式。由于现有的 JET 有延时富余和难以使用偏转路由的缺点, 本文提出了一种新方法 JLT。这种方法使用很小的初始偏移时间, 有效地利用节点上的 FDL 延时来弥补偏移时间不足, 从而减小了 Burst 在 OBS 网络中的延时。由于 JLT 使用的是局部偏移时间, 因此, 在使用偏转路由时不会有初始偏移时间难以确定的问题。通过在 JLT 中引入偏移时间增量, 本文还讨论了在 WDM 层实现区分服务的方法。文章的最后给出并分析了仿真的结果, 并较好地证明了以上结论。

参考文献

- [1] Xiong Y, Cankaya H C. Control architecture in optical burst-switched WDM network. *IEEE J. on Selected Areas in Communications*, 2000, 18(10): 1838 - 1851.
- [2] Liu L, Wan P J, Frieder O. Optical burst switching: the next IT revolution worth multiple billions dollars? 21st Century Military Communications Conference Proceedings, 2000, 2: 881 - 885.
- [3] Yoo M, Qiao C, Dixit S. Optical burst switching for service differentiation in the next-generation optical internet. *IEEE Communications Magazine*, 2001, 39(2): 98 - 104.
- [4] Hsu C F, Liu T L, Huang N F. On the deflection routing in QoS supported optical burst-switched networks. *IEEE International Conference*, 2002, 5: 2786 - 2790.
- [5] Tancevski L, Yegnanarayanan S, Castanon G. Optical routing of asynchronous, variable length packets. *IEEE J. on Selected Areas in Communications*, 2000, 18(10): 2084 - 2093.
- [6] Detti A, Eramo V, Listanti M. Performance evaluation of a new technique for IP support in a WDM optical network: optical composite burst switching (OCBS). *IEEE Journal of Lightwave Technology*, 2002, 20(2): 154 - 165.
- [7] Yoo M, Qiao C, Dixit S. The Effect of limited fiber delay Lines on QoS performance of optical burst switched WDM networks. *IEEE International Conference*, 2000, 2: 974 - 979.
- [8] Cankaya H C, Xiong Y, Vandenhoute M. Design issues of optical IP routers for internet backbone applications. *IEEE Communications Magazine*, 1999, 37(12): 124 - 128.
- [9] Ge A, Callegati F, Tamil L S. On optical burst switching and self-similar traffic. *IEEE Communications Letters*, 2000, 4(3): 98 - 100.

袁 巍: 男, 1978年生, 博士生, 研究方向为下一代网络体系结构。

李津生: 男, 1937年生, 教授, 博士生导师, 研究方向为信息通信网。

洪佩琳: 女, 1961年生, 教授, 博士生导师, 研究方向为信息通信网。