

# 光突发与光电路混合交换边缘节点的设计与实现\*

郭宏韬 吴龟灵 左 羽 李新碗 陈建平

(上海交通大学区域光纤通信网与新型光通信系统国家重点实验室, 上海 200030)

**摘 要** 提出了一种将光电路交换和光突发交换融合于同一网络的边缘节点设计方案. 根据此方案, 网络边缘节点能根据业务需求灵活选择交换方式. 该方案在实际实验系统中进行了验证. 结果表明, 采用光电路交换方式, 网络传输速率相对较快, 而采用光突发交换方式, 网络的带宽利用率更好. 这两种交换方式共存于同一网络, 一定程度上能提高网络的灵活性和适应性.

**关键词** 光突发交换; 光电路交换; 混合交换; 边缘节点

**中图分类号** TN929.11 **文献标识码** A

## 0 引言

为了提供对数据业务的高效支持, 基于密集波分复用(DWDM)的光网络需要有快速、灵活的交换和处理方式来对光纤带宽进行有效、合理的分配和管理. 由于电子速率“瓶颈”的限制, 传统光/电/光转换的交换方式无法适应网络发展的需求, 光交换是光通信发展的必然趋势. 目前主要的光交换方式有三种, 分别为: 光电路交换(Optical Circuit Switching, OCS), 光突发交换(Optical Burst Switching, OBS), 光分组交换(Optical Packet Switching, OPS). 其中, 光分组交换所需的光逻辑处理器件目前还很不成熟<sup>[1]</sup>, 具备处理最小分组能力的高速光控制器件还很少, 短期内难以实现. 光突发交换以突发包(burst)为传输和交换的基本单位, 具有中等交换粒度. 其数据分组(BDP-Burst Data Packet)和控制分组(BHP-Burst Header Packet)是分离的, 并采用单向资源预留方式. OBS的上述特点使其能充分发挥电处理的成熟和灵活的优点以及光传输大容量的优势, 降低了对光子器件的要求和中间交换节点的复杂度, 减小了建立通道的等待时延, 并很好地实现对突发业务的支持<sup>[2]</sup>. 光电路交换/波长路由技术已较为成熟. 它以波长为选路通道, 采用双向资源预留方式, 不需要光缓存, 可提供有保证的服务. 在连接持续时间较长且业务量稳定的情况下, 效率较高<sup>[3]</sup>. 由上可见, 光突发交换和光电路交换具有不同的业务支持能力. 如果在网络中采用同时具有光突发交换和光电路交换能力的混合交换方式, 则可以根据业务的特点和服

务需求, 灵活地选择更为有效的交换方式, 以获得更高的网络带宽利用率.

本文提出了一种将光电路交换和光突发交换融合于同一网络的边缘节点设计方案, 根据业务的要求通过管理配置选择不同的交换方式, 并在实际的光交换系统中验证了这种方案的可行性.

## 1 混合交换边缘节点的功能结构

图 1 为混合交换边缘节点的功能结构图, 分为发送和接收两部分. 发送部分包括输入线卡、协议选择、OBS 发送和 OCS 发送. 输入线卡完成上层业务的解帧, 协议选择模块根据不同业务需求通过管理配置选择 OCS 或 OBS 交换方式.

OBS 模块的发送部分包括组装、调度及输出几个模块<sup>[4]</sup>. 组装模块实现接入子网地址和业务等级到 OBS 网络地址和等级的映射, 并将用户数据放入相应的组装队列组装 BDP<sup>[5]</sup>. 调度模块根据组装队列中 BDP 的类型和等级及其出口路由器的地址, 确定该 BDP 的偏置时间, 并为其分配传输的数据通道波长和时隙, 产生相应的 BHP 并为其分配控制信道. 发送模块根据调度的结果, 发送 BHP 和 BDP. 对应的 OBS 接收部分包括接受模块、拆卸模块, 完成突发包的接受并将其还原成用户数据.

OCS 的发送和接收过程复杂度相对较低, 其发送过程主要有连接状态判断模块, 调度模块和发送模块, 根据不同的连接状态生成对应的控制包, 并在“连接建立”的状态下调用发送模块发送数据包. 调度模块为控制包和数据包分配信道, 发送模块根据调度结果发送控制包和数据包. OCS 接收部分负责将接收到的数据包直接向上层传递, 并根据所接受到的控制包的类型进行对应的处理.

接收部分的用户数据调度和管理模块、用户数据转发引擎及适应于不同子网的线卡执行传统路由器的功能, 为 OBS 和 OCS 共用.

\*国家自然科学基金(90204006, 60377013), 国家教育部(20030248035, 2003034258)和上海市科委资助(036105009)

Tel: 021-62933302 Email: frankhunt7@etang.com

收稿日期: 2005-06-14

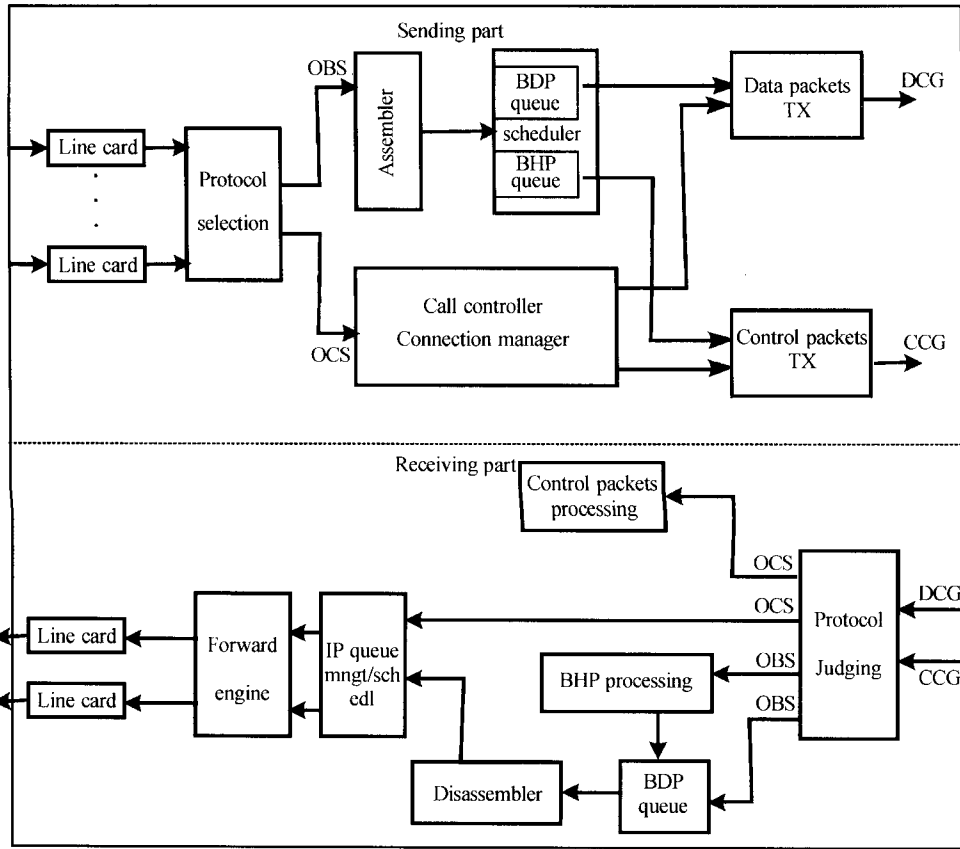


图 1 混合交换网络边缘节点功能结构图  
Fig. 1 Functional architecture of hybrid switching edge routers

## 2 基于 Linux 系统的混合交换网络边缘节点软件的实现

通过修改 Linux 系统内核<sup>[6]</sup>,在 IP 层和 MAC 层之间加入了 OBS 协议软件包、OCS 协议软件包和相关控制程序,实现了基于 Linux 系统的混合交换边缘节点。

图 2 为混合边缘节点总的软件结构示意图,其中箭头所示为数据流的传递方向。协议选择单元根据业务的要求选择 OCS 或 OBS 发送处理模块。OBS 发送处理模块完成组装、调度等 OBS 协议相关的发送处理。OCS 发送处理模块则是完成基于 OCS 协议的发送处理。协议识别单元根据所接收

的控制包的类型选择相应协议的控制包接收处理单元。设计中 OBS 数据和 OCS 数据均采用同样的以太帧封装,接受端对 OBS 数据和 OCS 数据采用同一个数据包接收处理模块。

### 2.1 边缘节点 OBS 软件包的实现

图 3 为 OBS 协议发送部分的软件结构。在发送过程中,组装进程将进入 OBS 层的 IP 包分类组装成突发包并放入组装队列,每个突发包结构都包括一个控制包结构和一个数据包结构。调度进程采

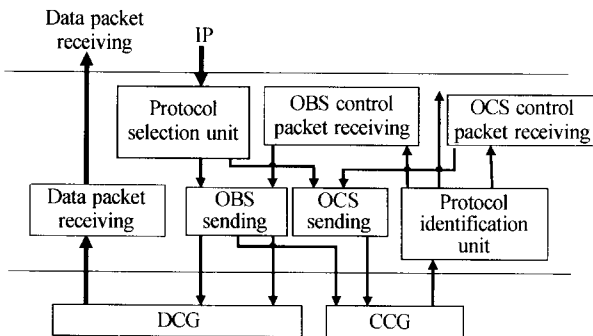


图 2 混合边缘节点软件结构  
Fig. 2 Software structure of hybrid switching edge routers

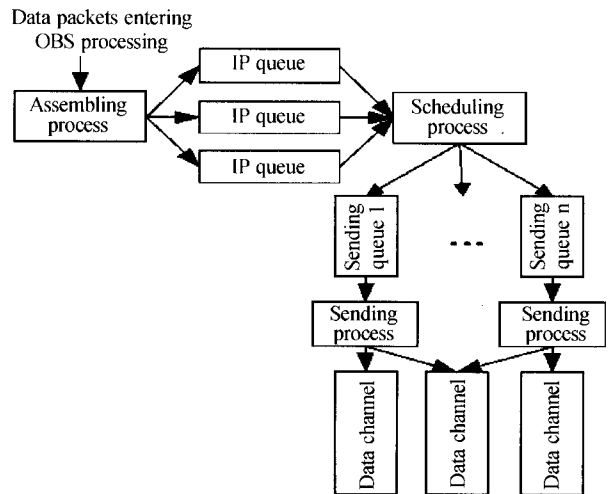


图 3 OBS 发送部分软件结构  
Fig. 3 Software structure of OBS sending part

用基于优先级的最短信道分配机制,将已组装好的突发包调度到某信道的发送队列. 每个数据信道都有一个发送进程,用于从信道的发送队列中提取突发包,从而实现单个突发包的发送.

控制包接收和处理过程的流程见图 4. 硬件设备接收到以太帧后,产生一个硬件中断. 系统调用接受中断处理程序,将接受到的数据放入系统接受队列. 然后系统通过软中断从接受队列中读取数据,并将它传给上层进行处理. 处理函数根据接受

到的包的协议类型来判断接受到的是数据包(IP包)还是 OBS 控制包或 OCS 控制包. OBS 和 OCS 控制包的协议类型在发送时被定义. 如果是数据包,则直接向上传送. 如果是 OBS 控制包则读取控制包的类型域;当控制包为 BHP 时,根据其中的信息产生对数据接收端口的控制信号,当接受到释放包(BEP-Burst End Packet)时,产生应答包释放确认(BEP\_ACK),并发回源端. 当接受到 BEP\_ACK 时,清除对应的 BEP 定时器.

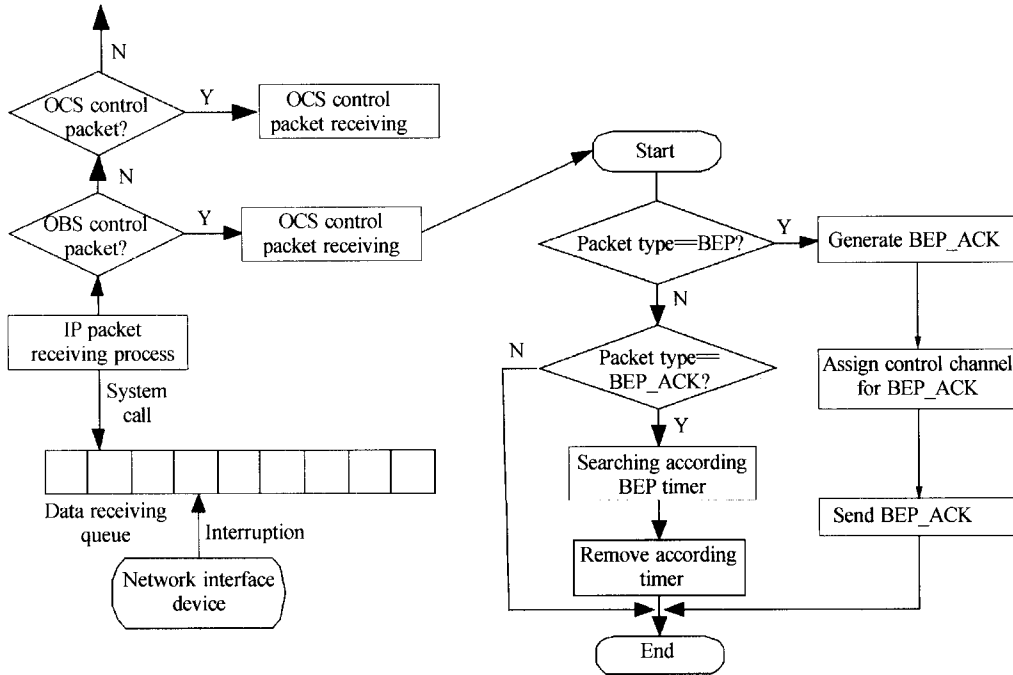


图 4 OBS 接收过程  
Fig. 4 OBS receiving process

2.2 边缘节点 OCS 软件包的实现

OCS 协议与 OBS 协议的相似之处在于控制分组/呼叫分组和数据分组都是分离的,所不同的是 OCS 有一个等待连接确认的过程,并且不需要将数据组装成突发包.

图 5 为 OCS 发送部分的软件结构. 在数据发送过程中,需要建立一系列 IP 队列,建立的原则是:将具有相同目的地址的 IP 包放入同一队列中. OCS 层发送原理为:首先接收上层的 IP 包进入 OCS 层,然后判断通往目的节点的连接是否已经建立,若已建立,则将 IP 数据放入对应链路的队列中,等待发送. 反之,则再判断是否已向目的节点发送连接请求,若是,则将数据加入呼叫队列;若否,则立即调用发送函数通过控制信道向目的节点发送连接请求控制包,并建立一个新的呼叫队列. 连接管理进程调用发送函数将各链路 IP 队列里的数据传送到下层,并在一个链路上的所有会话都结束后发送释放分组,以释放链路.

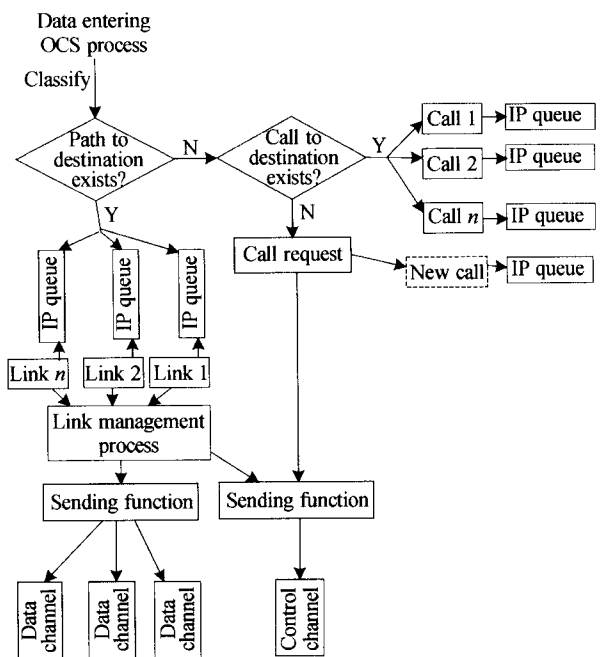


图 5 OCS 发送部分软件结构  
Fig. 5 Software structure of OCS sending part

在接受过程中,若接收到的为数据帧,则直接传至上层处理,若为 OCS 控制帧,则根据帧的类型进行对应的处理,具体为过程如下:

Switch (控制包类型){

Case ‘连接请求’: 分配信道,若成功,则生成连接确认包,发送至目的端;若失败,则返回“连接失败”信息;

Case ‘连接确认’: 建立新的链路,其的 IP 队列指向对应呼叫队列,将新建链路加入连接管理进程的链路队列中;

Case ‘连接失败’: 重新发送连接请求包至目的端,或放弃发送,释放对应的呼叫队列;

Case ‘释放请求’: 释放信道,向源端返回“释放确认”;

Case ‘释放确认’: 释放链路;

}

### 3 实验结果与分析

为了验证本文所提出方案的可行性和有效性,在实际系统中进行了测试. 系统由三个边缘节点和一个核心节点构成,如图6. 其中边缘节点为

Dell PE2650 服务器,运行 Linux 系统,包括数据信道接口卡和控制信道接口卡. 数据信道速率为 1.25 Gb/s;控制信道速率为 100 Mb/s. 核心节点由一台 PC 机和一个自行设计的光交换矩阵组成<sup>[7]</sup>,PC 机负责接收和处理突发控制包,并根据突发控制包的信息设置光交换矩阵,为数据包预留信道.

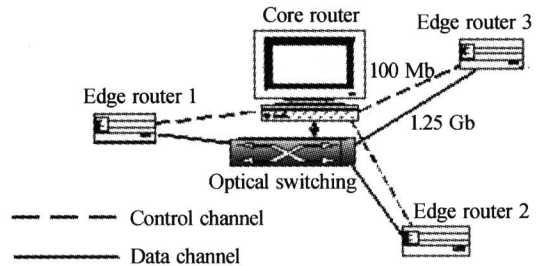


图6 实验系统

Fig. 6 Architecture of experimental system

边缘节点1同时做成FTP服务器和VOD点播服务器,边缘节点2和3作为客户端,演示FTP文件传输以及VOD视频点播. 服务器端通过管理配置程序来选择OBS方式或者OCS方式. 对六种典型情况进行了测试,结果列于表1.

表1 测试结果

边缘节点序号 实验方式	2号边缘节点	3号边缘节点
OBS方式,2号节点和3号节点同时演示FTP业务	传输速率:674 kb/s	传输速率:668 kb/s
OBS方式,2号节点和3号节点同时演示VOD业务	效果良好,无画面模糊及中断	效果良好,无画面模糊及中断
OBS方式,2号节点演示FTP业务,3号节点演示VOD业务	传输速率:728 kb/s	效果良好,无画面模糊及中断
OCS方式,2号节点和3号节点同时演示FTP业务	传输速率:16189 kb/s	无法建立连接
OCS方式,2号节点和3号节点同时演示VOD业务	无法建立连接	效果良好,无画面模糊及中断
OCS方式,2号节点演示FTP业务,3号节点演示VOD业务	传输速率:15446 kb/s	无法建立连接

由表中数据可以看到,采用OBS方式,两台客户机可以同时与服务正常通信,但相对传输速率较慢,这从FTP业务中可以看出. 而在VOD点播中,由于保证画面连续流畅的传输速率只需200 kb/s,因此从中体现不出速率差别;采用OCS方式,只有率先发送连接请求的客户机才能与服务通信,但是在已连接的状态下,数据传输速率与OBS方式相比要快许多.

基于两种交换方式的原理,不难对上述现象做出解释:OCS方式在数据传送期间信道带宽始终由先发出连接请求的客户机所占用,只有在数据传输结束后,由服务器端发送释放信息以释放带宽,这样客户机才能重新对带宽进行竞争;而OBS方式带宽

始终由两台客户机所共享,每台客户机在发送BDP之前需要向核心节点发送BHP以申请带宽,在每一个BDP发送完毕后随即向核心节点发送BEP以释放带宽. 这样在只有一条信道的情况下,OCS方式每次只能满足一台客户机与服务通信,不能像OBS方式那样实现带宽的复用,但由于省去了传输过程中申请带宽和释放带宽所需的时间,故而传输速率相对较快.

### 4 结论

本文提出了一种将OBS和OCS两种交换方式融合于同一光交换网络的边缘节点设计方案,给出了基于Linux系统的实现方法以及在实际系统中的

实验结果. 由实验结果可以看出,这两种交换方式能互相取长补短,所设计的混合交换边缘节点可以根据不同业务的要求选择不同的交换方式,将一定程度上提高网络的灵活性和适应性.

#### 参考文献

- 1 Zhao Huandong, Chi Hao, Zeng Qingji, *et al.* A novel optical buffer configuration optimization scheme for slotted optical packet switching. *Acta Photonica Sinica*, 2004, **33**(4):452~455
- 2 White J, Zukerman M, Hai L V. A Framework for optical burst switching network design. *IEEE Communications Letters*, 2002, **6**(6):268~270
- 3 Veeraraghavan M, Lee H, Zheng X. File transfers across optical circuit-switched networks. PFLDnet2003, <http://datatag.web.cern.ch/datatag/pfldnet2003/program.html>, Geneva, Switzerland, Feb. 3~4, 2003
- 4 吴龟灵,李新碗,陈俊峰,等. 光突发交换边缘路由器性能分析. *光子学报*, 2005, **34**(3):412~415
- 5 Zhang Zhizhong, Luo Jiangtao, Zeng Qingji, *et al.* Hybrid assembly and improved head dropping scheme for service differentiation in optical burst switching. *Acta Photonica Sinica*, 2004, **33**(10):1200~1203
- 6 李善平,刘文峰,李程远,等. *Linux 内核 2.4 版源代码分析大全*. 北京:机械工业出版社, 2001. 433~677
- 7 Li S P, Liu W F, Li C Y, *et al.* Linux kernel 2.4 Source Codes analysis. Beijing: China Machine Press, 2001. 433~677
- 8 Li Xinwan, Chen Jianping, Wu Guiling, *et al.* An experimental study on optical burst switching network based on wavelength selective optical switches. *IEEE Communication Magazine*, 2005, **43**(5):S3~S10

## Edge Router Design for Hybrid Optical Burst Switching and Optical Circuit Switching

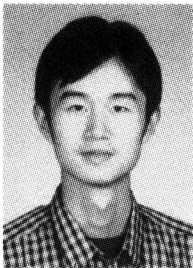
Guo Hongtao, Wu Guiling, Zuo Yu, Li Xinwan, Chen Jianping

*State Key Laboratory on Fiber-Optic Local Area Communication Networks and Advanced Optical Communication Systems, Shanghai Jiaotong University, Shanghai 200030*

Received date: 2005-06-14

**Abstract** An edge router for hybrid optical switching network which mixed OBS and OCS together was designed and implemented in the experimental system. The switching method (OBS or OCS) in the hybrid network can be selected by the service requirement. The result of the experiment shows that the data rate is relatively higher in OCS way, while the network bandwidth is better utilized in OBS way. By mixing these two switching ways together, the flexibility and adaptability of the network can be improved to a certain extent.

**Keywords** Optical burst switching (OBS); Optical circuit switching (OCS); Hybrid switching; Edge router



**Guo Hongtao** received his B. S. Degree from Zhejiang University in China in 2003. Now he is a graduate student in the State Key Laboratory on Fiber-optic Local Area Communication Networks and Advanced Optical Communication Systems, Shanghai Jiaotong University, China. His research interest focuses on optical switching.