

# 采用部分共享光缓存和共享波长转换器的 光分组交换节点性能分析

吴龟灵, 陈建平, 李新碗

(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, 上海华山路 1954 号浩然大厦 1709 室, 上海 200030)

**摘 要:** 本文研究了采用部分共享缓存 (PSB) 和共享波长转换器 (SWC) 的光分组交换节点在自相似流量下分组丢失率和时延与输出光缓存深度、循环共享光缓存的深度和组数以及波长转换器数量之间的关系, 并讨论了相关的尺度和规模问题。结果表明, 部分共享光缓存和共享波长转换器的光分组交换节点交换结构可以达到很高的性能; 采用共享方式可显著减少波长转换器的数量; 在保持共享缓存总的深度不变的情况下, 适当增加循环共享光缓存组数可以使交换达到更好的性能。

**关键词:** 光分组交换; 部分共享光缓存; 共享波长转换器; IP over WDM; 光 Internet

**中图分类号:** TN929. 11 **文献标识码:** A **文章编号:** 0372-2112 (2003) 121626-04

## Performance Analysis of Optical Packet Switching Employing Partially Shared Buffering and Shared Wavelength Converter

WU Guiling, CHEN Jianping, LI Xinwan

(The State Key Laboratory on Fiber Optic Local Area Communication Networks and Advanced Optical Communication Systems, Shanghai Jiaotong University, Shanghai 200030, China)

**Abstract:** The performance of optical packet switching employing partially shared buffering (PSB) and shared wavelength converter (SWC) for self-similar traffics are studied. The effects of output buffers, the PSBs and SWC on the packet loss probability and delay are analyzed with consideration on the issues, such as dimension of switching matrix, buffering size, etc. The simulation results show that optical packet switching nodes with PSB and SWC are of good performance. The scheme of SWC can obviously reduce the number of wavelength converters in comparison with the scheme employing wavelength converters at each input wavelength path. For fixed depth, grouping PSB into several subgroups can improve the performance of optical packet switching nodes.

**Key words:** optical packet switching; partially shared buffering; shared wavelength converter; IP over WDM; optical internet

### 1 引言

网络业务的迅猛增长要求未来的主干网能提供更大的带宽。尽管 DWDM 技术使光纤的承载能力大为提高, 然而光通信巨大带宽的充分利用还需要通过快速、灵活的交换方式才能得以实现。采用光/电/光方式的交换节点存在电子/瓶颈。克服这种电子/瓶颈的有效办法是直接进行光交换。在所有的光交换方式中, 光分组交换能以更细的颗粒度快速、灵活地对 DWDM 的巨大带宽进行有效分配和利用, 具有交换速率高、数据速率/格式透明、支持 ATM、IP 等不同的交换方式等优点。因此, 利用 DWDM 技术设计模块化、可扩展、高透明性的光分组交换节点已成为现在工业界和学术界非常重视的研究课题<sup>[1]</sup>。

在光分组交换节点的设计中, 如何解决分组竞争以降低分组丢失率是要考虑的重点之一。光缓存技术和波长转换技术是交换节点解决分组竞争的两种行之有效的办法。光缓存技术用来解决时域竞争。其缓存结构有输入缓存、输出缓存、循环共享缓存。输入缓存存在头阻塞问题(HOL), 吞吐量有限。输出缓存可达到较高的吞吐量, 因而在光分组交换中得到了广泛的应用。但由于目前尚没有成熟的光 RAM, 只能采用光纤延迟线(FDLs)来代替。而 FDLs 只能实现有限的缓存, 单纯的 FDLs 输出光缓存技术很难获得令人满意的性能, 大量的 FDLs 还使节点结构庞大。波长转换技术通过波长域来解决竞争。目前有非共享式和共享式两种方案。非共享式的方案需要大量昂贵的可调谐波长转换器, 成本高, 限制了交换结构的可扩展性。考虑到在一个时隙的交换中并不是所有的光分组都

收稿日期: 20021008; 修回日期: 20030701

基金项目: 国家自然科学基金(No. 90204006); 国家 863 计划(No. 2001AA122074, No. 2002AA122022); 上海市科委(No. 012261022, No. 022261001); 电子科技大学宽带光纤传输与通信系统技术国家重点实验室开放基金

需要波长转换的,提出了共享波长转换器(SWC)光分组交换结构<sup>[2,3]</sup>。研究表明,这种结构能在达到同样的性能的同时显著减少波长转换器数。不过当共享波长转换器达到一定的数量,分组丢失率达到饱和。这是因为当所有的波长都被同时占用时,波长转换就无能为力了。此时只有从时域着手,增加缓存才能进一步降低分组丢失率。文献[4]提出了一种循环共享缓存方案。在没有波长转换器的情况下,可达到较好的性能。但需要一个复杂的控制算法,且其光域实现还有待验证。J. Diao 等人提出了控制算法较为简单的部分共享缓存(PSB)技术<sup>[5]</sup>。我们曾经提出了基于 DWDM 的带有多组部分共享光缓存的光分组交换结构模型和采用部分共享缓存和共享波长转换器的光分组交换结构模型,并在贝努利流量下对其性能作了分析,结果表明这种节点结构可达到较高的性能<sup>[6,7]</sup>。

## 2 交换结构模型

基于 DWDM 的带有多组循环共享光缓存和共享波长转换器的光分组交换结构模型如图 1 所示。它基于一个  $(N+M) \times (N+M)$  无阻塞的空分交换矩阵,具有  $N$  根输入、输出光纤,每根光纤可支持  $K$  个不同波长。节点的实际交换能力为  $(N \times K) \times (N \times K)$ 。每个输出端口都有一个深度为  $P$  的输出缓存(OB),最大能容纳  $K \times P$  个分组。此外,还有为所有信道共享的  $M$  组深度为  $Q$  的 PSB 和  $W$  个 SWC。PSB 能容纳的最大分组数为  $M \times Q \times K$  个,其输出经分波器再输入到交换矩阵。OB 和 PSB 用来处理时域竞争,SWC 用来解决波长域的竞争。在分组交换的每个时隙内,系统提取该时隙所有到达分组的分组头信息,并根据这些信息确定分组的端口/光纤,设置光交换矩阵,使得分组交换到正确的目的端口/光纤,改写分组头,并使分组头和分组一同输出到其目的光纤,完成该时隙的分组交换。当一个时隙内有两个或两个以上相同波长上的分组需要交换到同一目的光纤,即产生竞争时,只有一个分组能直接输出到目的光纤,其余分组压入 OB。若 OB 已满,则将剩余的冲突分组通过交换矩阵输出到 PSB 对应的端口,使其进入 PSB。如果 PSB 也已满,则丢弃冲突分组。进入 PSB 的

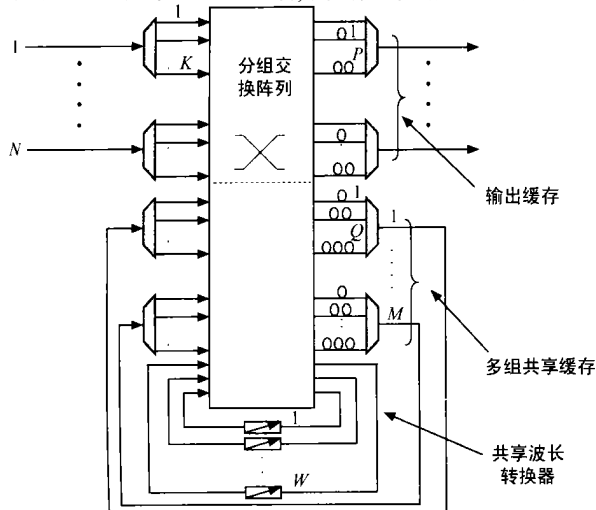


图 1 多组 PSB+ SWC 光分组交换节点结构示意图

分组经过一段延迟后,回到交换矩阵的输入端,重新参与信道的分配。SWC 的作用是改变冲突分组的承载波长,使冲突分组的数量降到最低程度,减少分组丢失率。此外,采用 SWC 使得压入缓存的分组数量下降,从而使冲突分组经历较少的时间即可离开交换节点,降低了分组的平均时延。

分组的控制和交换是按时隙进行。设某一时隙内,PSB 中要交换的分组的集合为  $CI = \{CI_{ij} | 1 \leq i \leq N, 1 \leq j \leq k\}$ ,  $CI_{ij}$  表示波长为  $K_j$ 、目的光纤为  $i$  的所有分组的集合;输入信道中该时隙内的分组集合为  $I = \{I_{ij} | 1 \leq i \leq N, 1 \leq j \leq k\}$ ,  $I_{ij}$  是其中波长为  $K_j$ 、目的光纤为  $i$  的分组集合。在进行交换时考虑到时序问题并防止多次循环,PSB 中的分组在 OB 和 WC 的分配上具有较高的优先级,而在 PSB 的分配上优先级较低。这是通过在 OB 和 WC 的分配中先处理 PSB 中的分组,而 PSB 分配中先处理新输入的分组来实现的。另外,还要考虑公平性和有效性。

一个时隙中的控制/调度过程分为:OB 和 WC 的分配(包括直通输出信道的分配,相当于经过零缓存)以及 PSB 的分配两个过程。首先是 OB 和 WC 的分配,尽可能地将分组交换到直通的目的端口或对应的 OB。为了保证 PSB 中的分组在 OB 和 WC 的分配上的高优先级,首先处理 PSB 分组集合  $CI$  中的分组。它包括三个步骤:第一步,直通信道(波长与分组的承载波长一致的输出信道)的分配。从目的端口有空闲直通信道的分组集合中随机取出一个分组,直接发送;第二步,为没有直通信道,但有其它波长的空闲信道的分组,分配 SWC。如果成功,则将其转换到另一个可用波长上发送出去;第三步,OB 的分配,即将剩余的未发送的分组移入对应目的端口、对应承载波长的 OB,直至没有分组或对应的 OB 已满;如果  $CI$  中仍有分组未处理且分组的端口口的其它波长的 OB 仍未满,则分配未用完的 SWC,将输入承载波长对应的 OB 已满的分组转换到其他波长对应的 OB 中。若  $CI$  中仍有剩余的分组,则等到 PSB 分配时处理。

在完成了对 PSB 中分组的 OB 和 WC 的分配处理后,再为输入信道中的新输入分组分配 OB 和 WC。其过程同上。

在整个 OB 和 WC 的分配过程完成后,若仍存在冲突分组未处理,则进入 PSB 分配过程,以将  $I$  和  $CI$  中的剩余冲突分组压入 PSB。此时,由于 PSB 中分组的优先级较低,所以输入信道中的剩余分组先得到处理。其过程如下:若  $I$  中仍有非空的分组集合,从  $I$  中随机选取一个非空的  $I_{mn}$ ,然后从  $I_{mn}$  中随机取出一个分组存入队列长度最短的  $PSB_{kn}$ ( $PSB_{kn}$  表示第  $k$  组 PSB 中用来存放承载波长为  $n$  的分组的缓存)中。如果该分组对应的所有 PSB 都满了,则丢弃  $I_{mn}$  中所有剩余的冲突分组。重复上述过程直至  $I$  中没有分组。随后,按同样的过程处理  $CI$  中的分组。

## 3 模拟结果与分析

针对分组丢失率以及分组平均时延这两个重要的性能指标,模拟分析了不同参数设置下交换结构的性能,探讨了不同器件和参数对交换性能的影响。假定输入/输出光纤数为  $N$ ;每个输入信道承载具有自相似的流量,其流量负载为  $Q$ 。每一个分组都按均匀随机概率(即,  $1/N$ ) 选取其目的光纤/端口。

311 自相似流量模型

研究表明, Internet 业务是长程相关的, 可以用自相似模型来描述<sup>[8]</sup>. 自相似流量可由无限多个独立具有 Pareto 分布的 ON/OFF 信号源合成<sup>[9]</sup>. Pareto 分布的累积分布函数为:

$$F(x) = 1 - (k/x)^A, \quad 1 < A \leq 2, \quad x \geq k \quad (1)$$

其中, A 表示重尾分布的程度, k 为 x 的最小值. 合成的自相似通信量的 Hurst 参数为  $(3 - a) / 2$  ( $0.5 \leq H < 1$ ).

在实际应用中, 只要符合 Pareto 分布的 ON/OFF 信号源的数量足够多, 便可合成具有近似自相似性的流量. 我们在模拟分析过程中采用了 128 个这样的信号源来合成所需要的通信量. 每个 ON/OFF 信号源的 ON 周期时隙的个数符合最小值为 1 的 Pareto 分布, OFF 周期的长度符合连续 Pareto 分布. OFF 周期的最小长度由负载和 ON 周期的平均长度确定<sup>[9]</sup>. 考虑到时隙交换, 在最终合成的流量中, OFF 间隔的长度也被转化成固定时隙的整数倍. 图 2 是我们合成的负载为 0.6 的流量的 R/S 图. A 分别为 112, 114, 116. 线性拟合表明合成流量对应的 Hurst 参数分别为 0.905, 0.81, 0.703, 与理论预期结果吻合的非常好.

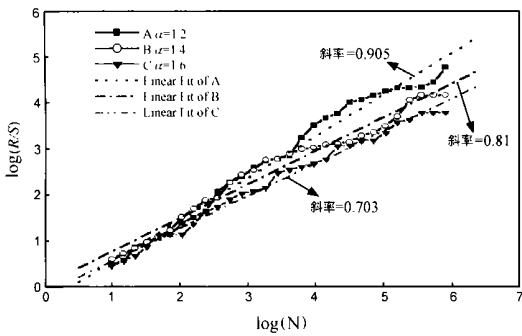


图 2 合成的自相似流量的 R/S 图

312 模拟结果

模拟的光分组交换节点有 16 个输入/输出光纤, 每纤 4 个通道/波长. 输入的自相似流量由 Pareto 参数为:  $A_n = 112$ ,  $A_n = 114$  的 128 个 ON/OFF 源合成. 未加特别说明时, 波长转换器的数目为 4, 输出缓存的深度为 2, 负载为 0.8.

由图 3 可见, 随着输入负载的增加丢包率和平均时延均增加; 同时可见, PSB 的使用使丢包率明显下降, 平均时延增加, 特别是在负载较低时 ( $< 0.8$ ); 另外, 模拟结果表明当共享缓存总的深度相同时增加 PSB 的组数可以降低丢包率并减少时延. 但随着 PSB 组数的增加改善的幅度下降, 并且随着负载的增加不同组数 PSB 节点间性能的差异减小. 这是因为在总

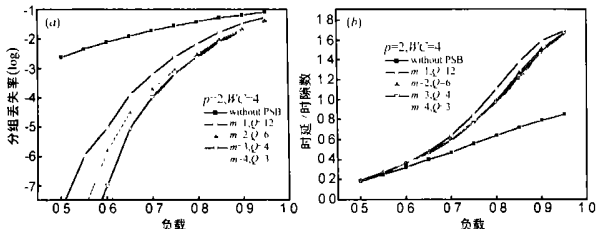


图 3 不同 PSB 配置下光分组交换节点的丢包率 (a) 和平均时延 (b) 与负载间的关系

容量一定的情况下, PSB 组数越多, PSB 的深度越小, 进入 PSB 的冲突分组在再次调度时, 再次发生冲突的可能也就越大. PSB 组数越多, 这种作用越明显. 此外, PSB 的组数越多要求的交换矩阵的端口也越多 (即矩阵规模越大).

由图 4 可见, 随着输出缓存深度的增加丢包率下降, 而平均时延先增加到到达一定值后保持不变; 由不同 PSB 配置的结果可见, 采用 PSB 可以使丢包率明显下降, 特别是当输出缓存有一定深度时, 而且此时的平均时延与没有 PSB 时的几乎相同. 如当  $P = 7$  时, 不同 PSB 配置下的平均时延的几乎相同 ( $\sim 1120$ ), 有一组 PSB 时的丢包率约为  $10^{-5}$ , 而没有 PSB 时的丢包率越为  $10^{-2.75}$ . 对总的深度相同而组数不同的 PSB 配置可以看到与图 3 类似的结果, 即 PSB 组数的增加可以降低丢包率并减少时延, 但随着 PSB 组数的增加不同组数 PSB 节点间性能的差异减小; 此外可见, 随着输出缓存深度的增加不同组数 PSB 节点间性能的差异增加.

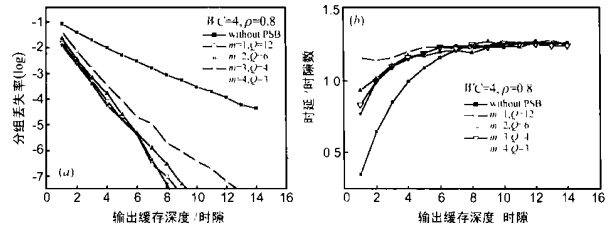


图 4 不同 PSB 配置下光分组交换节点的丢包率 (a) 和平均时延 (b) 与输出缓存间的关系

图 5 给出了波长转换器对节点丢包率 (a) 和时延 (b) 的影响. 由图 5 可见, 一开始, 随着波长转换器数的增加丢包率和时延下降, 最后趋向饱和. 这是因为波长转换器可以通过波长域解决竞争问题, 降低了丢包率, 同时减小了队列的长度, 改善了时延特性. 但当所有的波长都被占用时, 再增加波长转换器也无能为力了. 我们可以看到达到饱和的共享波长转换器的数量 ( $\sim 11$ ) 要远低于输入端口数 ( $16 @ 4$ ). 这表明, 与每个输入波长均采用转换器的方式相比, 在同样的丢包率和时延等性能下, 采用共享波长转换器结构可以显著减少波长转换器数量; 由图还可见, 有 PSB 时节点的丢包率得到明显改善, 但时延特性有所下降. 多组 PSB 配置间的性能比较有与图 3 和图 4 类似的结果.

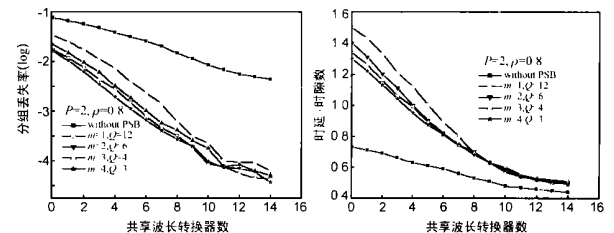


图 5 不同 PSB 配置下光分组交换节点的丢包率 (a) 和平均时延 (b) 与共享波长转换器数间的关系

4 结论

本文研究了采用部分共享缓存 (PSB) 和共享波长转换器 (SWC) 光分组交换节点的性能, 给出了交换和分配算法原则.

模拟分析了自相似流量下输出光缓存深度、循环共享光缓存的深度和组数以及波长转换器数量对节点性能(丢包率和时延)的影响,并讨论了相关的尺度和规模问题。结果表明,PSB 可以显著改善节点的丢包率,同时降低时延特性;在一定范围内,多组 PSB 能在不增加共享缓存总的深度的情况下降低丢包率和平均时延,但交换矩阵的规模随之增加,算法也更加复杂性,并且随着 PSB 组数增加改善的幅度下降,尤其是在高负载和多 SWC 时,实际使用时需要折中考虑。SWC 可同时改善节点的丢包率和时延特性。采用共享方式可显著减少波长转换器的数量,降低成本。其代价是增加了控制的复杂性。部分共享光缓存和共享波长转换器的结合则可进一步提高光分组交换节点的性能。通过优化设计可以在丢包率、时延及成本方面都达到较高的性能。

#### 参考文献:

- [ 1 ] S Yao, B Mukherjee. Advances in photonic packet switching an overview[ J]. IEEE Communication Magazine, 2000, 38( 2): 84- 94.
- [ 2 ] V Eramo, M Listanti. Packet loss in a bufferless optical WDM switch employing shared tunable wavelength converters[ J]. J Lightwave Technol, 2000, 18( 12): 1818- 1833.
- [ 3 ] V Eramo, M Listanti. Wavelength converter sharing in a WDM optical packet switch: dimensioning and performance issues [ J]. Computer Networks, 2000, 32( 5): 633- 651.
- [ 4 ] M J Karol. A shared memory optical packet (ATM) switch [ A]. Proceedings of the 6th IEEE Workshop [ C]. Local and Metro Area Networks, 1993. 205- 211.
- [ 5 ] J Diao, P L Chu. Analysis of partially shared buffering for WDM optical packet switching [ J]. J Lightwave Technol, 1999, 17( 12): 2461 - 2469.

- [ 6 ] X Li, Jianping Chen, et al. Performance analysis of photonic packet switching employing multiple groups of partially shared buffering[ J]. Journal of Optical Commun, 2003, 24( 2): 68- 70.
- [ 7 ] X Li, Jianping Chen, et al. Improvement of optical packet switching performance by employing partially shared buffering with shared wavelength converters[ A]. Proceedings of SPIE APOC 2001[ C]. Beijing, China, 2001. 4582. 139- 146.
- [ 8 ] V Paxson, S Floyd. Wide area traffic: the failure of poisson modeling [ J]. IEEE/ ACM Trans. Networking, 1995, 3( 6): 226- 244.
- [ 9 ] M S Taqqu, W Willinger, R Sherman. Proof of a fundamental result in self similar traffic[ J]. Computer Commun. Rev, 1997, 27( 2): 5- 23.

#### 作者简介:



吴龟灵 男,1971 年 11 月生于安徽芜湖,博士,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从事光突发交换和光分组交换方面的研究与开发工作,在国内外期刊上发表论文 20 多篇。



陈建平 男,1962 年 9 月生于浙江金华,博士、教授,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副主任,上海交通大学电子工程系副主任,一直从事密集波分复用(DWDM)光纤通信器件、系统和网络的研究与开发工作,承担完成国家自然科学基金项目、国家 863 计划项目等多项,在国内外期刊上发表论文

40 多篇。