

# 光突发交换网中的智能组包研究

李彦君<sup>1</sup>, 洪小斌<sup>2</sup>, 郭宏翔<sup>1</sup>, 伍 剑<sup>1</sup>, 林金桐<sup>1</sup>

(1. 北京邮电大学光通信中心, 北京 100876; 2. 北方烽火科技有限公司, 北京 100020)

**摘 要:** 光突发交换(OBS) 相对传统的电路交换方式具有更高的传输效率和更短的网络时延, 而相对于未来网络的发展趋势光分组交换, 光突发交换具备更现实的可行性. 在 OBS 中, 突发汇聚是一个很关键的技术, 它对 OBS 网络的性能有着重要的影响, 在本文中, 我们对突发汇聚机制进行了较深入的研究, 并提出了一种新的突发汇聚算法, 它对抑制网络流量的自相似性, 提高突发传输效率和避免在不同节点由于突发同步所引起的带宽资源竞争都有较好的结果.

**关键词:** 光突发交换; 汇聚; 门限; 自相似性; ON/OFF 源

**中图分类号:** TN929.1 **文献标识码:** A **文章编号:** 0372-2112 (2005) 11-2040-04

## Study on Intelligent Burst Assembly Mechanism in Optical Burst Switch Networks

LI Yan jun, HONG Xiao bin, GUO Hong xiang, WU Jian, LIN Jin tong

(1. Lab. of Optical Fiber Communications, Beijing University of Posts and Telecommunications, Beijing 100876, China;

2. Beijing Northern Fiberhome Technologies CO. LTD, Beijing 100020, China)

**Abstract:** As a promising switching paradigm, Optical Burst Switch (OBS) was proposed for optical networks showing higher efficiency and shorter delay than traditional circuit switching, and it is more feasible than optical packet switching. In OBS, burst assembly is one of the critical technologies, and will significantly affect the performance of networks. In this paper, we focus on the study of burst assembly mechanism, and propose a new burst assembly algorithm, which reduces the self similarity of IP traffic, enhances the data burst utilization and avoids the bandwidth contention due to sending burst data synchronously at different ingress nodes.

**Key words:** optical burst switching; assembly; threshold; self similarity; ON / OFF sources

## 1 引言

由于 IP 业务的急剧增长, IP over WDM 方式成为下一代核心光网络架构设计的主要特征. 做为一种简单灵活的交换方式, 光突发交换 (OBS) 通过结合电处理和光传输各自的优势, 为未来光网络架构发展提供了一种很有前景的方案<sup>[1~5]</sup>.

OBS 的主要思想是将比较小的 IP 进行组装成一个大的突发包, 使得突发数据通过光开关的时间为毫秒级, 现在的光器件可以满足这一要求. 相对传统的电路交换而言, OBS 网络具有更大的灵活性和带宽利用效率, 相对于目前的技术研究热点光分组交换而言, 它具有更简单的可实现性.

同时, OBS 网络也面临着一些技术难点有待解决, 其中边缘节点处突发包的组装是 OBS 的一项很关键的技术, 直接影响到 OBS 网络的综合性能. 本文中, 我们针对网络业务的突发特性, 提出了一种具备流量变化自适应调整的智能组包算法, 并运用文献[6]中提出的 ON/OFF 模型用仿真模拟实际的网络流量, 与文献[7]中提出的基于时间门限和文献[4]中提出的基于时间和定长双门限的组装算法进行了综合比较和性能评估.

文中第 2 部分概述了 OBS 网络架构及边缘节点汇聚的大致流程, 第 3 部分讨论和分析了组装算法的研究情况, 并提出了流量自适应的组包算法流程, 第 4 部分为仿真结果, 第 5 部

分为结论.

## 2 OBS 网络架构及突发数据汇聚

OBS 网络由边缘节点和核心节点构成, 它采取了将控制信息和数据信息剥离的方法, 提前发送控制分组, 沿路为数据信息预约节点资源, 数据信息在 OBS 网络的边缘接入节点进行组装, 突发包从源节点到目的节点的传输, 始终落在全光域内. 而控制信息则在每个节点处要进行 O/E/O 的转换, 如图 1 所示:

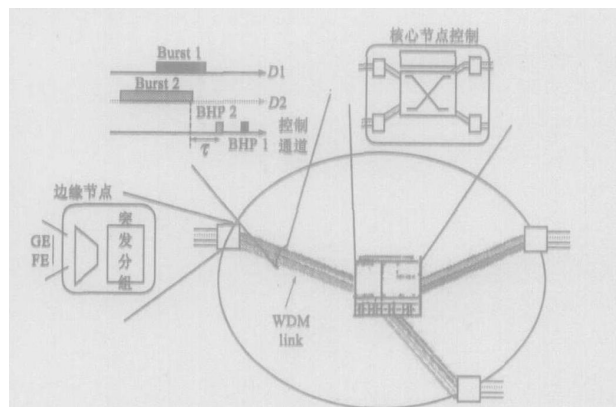


图 1 OBS 网络基本模型

边缘节点对 IP 业务按目的地址和业务类型进行分类汇聚,根据业务流量的特性进行突发组装和偏置时间的计算,在核心节点,通过对突发分组控制包(BHP)携带的路由信息、偏置时间等信息的处理,转发节点完成对交换矩阵的驱动设置,为预约成功的突发数据在相应的时间段保留带宽资源。

进入边缘汇聚节点的 IP 业务,在节点通过汇聚模块进行地址和业务类型的分类识别,同一业务类型和不同 IP 目的地址但却有相同的边缘节点地址的数据流入同一个突发组装器,即缓存队列,汇聚组装算法一般通过对组装时间和队列长度的静态或动态设置来完成突发包的汇聚机制,当突发包在组装器中形成后,即进入调度模块排队处理,对位于队首的突发包,调度器通过一定的算法生成偏置时间,同时发送对应的控制分组,该分组承载了突发数据的数据信道,偏置时间信息, QOS 类别等信息,而突发数据则进入边缘节点发送缓冲区,待偏置时间到达后调制成帧通过 WDM 层发送。

### 3 组装机理研究

在 OBS 网络边缘节点的汇聚过程中,包长控制是组装机理中一个很关键的因素,它对偏置时间生成、突发调度、拥塞控制和流量突发性抑制等相关的网络性能有着重要的影响。

好的组装算法应当具有高效的网络资源利用效率、良好的时延特性和拥塞控制性能,随着近年来对 OBS 网络系统的深入研究,一些突发组装算法相继提出,如 An GE 和 Franco Callegati 提出的一种基于时间门限控制的组装算法(Fixed Assembly Time, 以下一律简称 FAT 算法)<sup>[7]</sup>,它通过对每个组装缓存器设置一定的组装时间计数器,当组装时间到达门限时,创建新的突发包,并复位计数器,开始下一个突发组装。仿真结果表明,时间门限的组装算法对流量具有一定的平滑作用,但这种简单的算法存在一些问题,比如在网络流量高峰期它需要大容量的缓存队列,而在低网络负载的情况下,又会导致整体网络时延过长,突发组装效率降低。为了解决前一个问题,Xiong 在文献[4]中提出了一种基于时间和定长双门限的组装算法(Fixed Assembly Time & Length, 以下一律简称 FAT-L 算法)。

然而,这两种组装算法,在网络流量低负荷的情况下,均会存在突发长度相对突发门限的较大的变化范围,这将导致网络整体时延增加,造成突发组装效率降低,同时,文[7]中的 FAT 算法,由于突发产生的时间是一个固定的时间量,当不同的边缘节点突发产生时间发生同步时,必然将引起不同边缘节点处发出的突发数据包在核心节点连续的资源竞争,也就是说,在突发包形成的同时发出的控制包在核心节点处预约资源时的持续竞争。在边缘节点数量较多,网络资源相对紧张的情况下,这种算法会造成极大的阻塞率。这些缺陷在文献[4]中 FAT-L 算法中,同样存在。

针对此,在本文中我们提出这种可以根据业务流量状况进行自适应调整,并且有利于破坏各边缘节点组装同步性的智能组装算法(Fixed Assembly Time & Variable Length, 以下简称

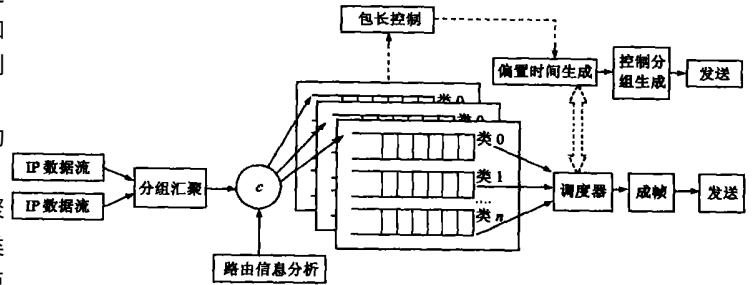


图2 OBS网络边缘节点汇聚流程图

FAT-VL 算法):

我们在前面算法的基础上再引入了一个突发包长限制门限  $Q_{th}$ , 以及调节步长  $L_{step}$ , 调节容限  $L_{td}$ , 和流量计数器  $N_{um}$ , 同时, 为了防止拥塞和提高传输效率, 突发包长应有最长突发  $BS_{max}$ , 和最短突发  $BS_{min}$  的限制<sup>[4]</sup>。

我们算法的基本思想是设立一个动态可变的包长组装门限, 它通过监测流量负荷的变化来进行调节, 这种流量自适应的调节方式, 在低流量时会降低组包门限, 从而缩短业务的等待时延, 提高突发组装效率, 同时, 门限的可调性, 将使得低流量时, 不同流量特性的边缘节点突发仍然保持异步, 这对破坏各节点的突发同步竞争, 改善网络性能很有意义, 在后面的仿真中我们验证了这一结论。

近年来, 对现有网络中业务的大量统计表明, 网络流量具有自相关性<sup>[8,9]</sup>, 它意味着当前流量统计特性和历史是长相关的, 其统计特性表明流量高峰(低谷)期已持续的时间越长, 它还将持续的时间就越长, 利用这一特点, 我们通过设置一个流量计数器  $N_{um}$  对近时间段的网络流量特性进行统计, 便可实现对门限的动态滞后调整。

每次组装完成后, 我们通过比较当前的突发长度门限和按算法生成的突发包长度, 当突发包长度超出当前设置的门限值, 对计数器  $N_{um}$  进行递增操作, 如突发长度低于门限值, 且两者之差超出了调节容限  $L_{td}$ , 则计数器  $N_{um}$  进行递减操作, 当  $N_{um}$  到达一个规定的上界后(这意味着网络流量出现持续高峰/低谷, 需要调节门限), 按调节步长增加组装门限值, 反之, 当  $N_{um}$  到达下界



图3 算法组装简图

后, 按调节步长减少组装门限值, 如图3所示, 设立调节容限  $L_{td}$  是为了尽量降低突发门限的变化量, 这有利于生成变化量较小的突发包长, 资源调度时降低阻塞率。

具体算法的步骤简述如下:

(1) 初始化包长门限  $Q_{th} = (BS_{max} + BS_{min})/2$ , 初始化步长  $L_{step} = (BS_{max} - BS_{min})/N$ , 初始化时间门限  $T_{th}$ , 初始化流量计数器  $N_{um}$ , 其中  $N_{um}$  和  $N$  的大小决定了门限调节的速度。

(2) 当第一个 IP 包进入组装队列时启动计时器, 如果计时器到达时间门限  $T_{th}$  或者队列长度超出或等于包长门限  $Q_{th}$ , 报告生成新的突发包。

(3) 比较上一个生成的突发包与包长门限的大小, 如果生成的突发包大于或者等于突发门限, 计数器  $N_{um}$  加一, 如果

小于突发门限, 且与门限长度差大于  $Ltol$ , 那么计数器  $Num$  减一。

(4) 比较计数器  $Num$  和规定的上界和下界, 如果  $Num$  超出上界, 表明网络流量统计处于高负荷, 将门限提升至  $Qth + Lstep$ , 如果  $Num$  低于下界, 表明网络流量统计处于低负荷, 将门限下调为  $Qtr - Lstep$ , 同时保证  $Qth + Lstep < BSmax$  (最大突发包长限制) 和  $Qtr - Lstep > BSmin$  (最小突发包长度)。

(5) 计时器清零, 重新开始下一个突发组装。在这种算法控制下, 突发产生时间和突发包长都是随流量变化的, 有利于破坏各节点突发发射的同步性引起的持续阻塞, 网络流量低负荷时, 门限下调, 从而组包时长也相应缩短, 相对于定时门限和定长定时双门限的组包方式具有更低的网络整体时延和更高的突发包传输利用率。

#### 4 仿真模型及结果分析

在理论上研究自相似性流量的统计复用性能, 难度非常大, 本文依据文献[6]的结论, 采用 40 个独立 Pareto 分布的 ON/OFF 源进行叠加的模型来模拟一路网络的自相似性流量特性, 对上述算法和分析进行仿真验证。

仿真程序中, 最大突发包长度  $BSmax$  定为 500KByte, 最小突发包长度  $BSmin$  定为 15KByte, 我们对独立 Pareto 分布的 ON/OFF 源分别取 Shape 参数  $\alpha$  为 1.2, 1.4, 1.6, 自相似性参数  $H = (3 - \alpha)/2$ 。

$H$  越大, 表明网络流量相关性越强, 自相似性越大。

其余参数设置分别为:  $Ltol = 48$  5KByte,  $Lstep = (BSmax - BSmin)/10 = 48$  5KByte。

仿真采用的是如图 4 所示典型的星形拓扑结构图:

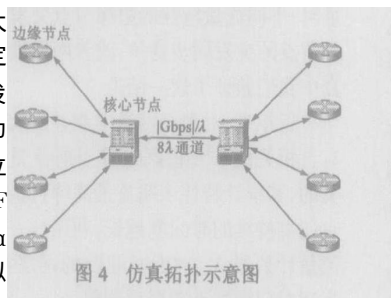


图4 仿真拓扑示意图

流量归一化负荷使用  $\rho = \sum_{i=1}^N \frac{\overline{ON}_i}{\overline{ON}_i + \overline{OFF}_i} \cdot B_{channel}$  计算,

其中  $\overline{ON}_i$  为  $i$  号分布源的平均发包时间,  $\overline{OFF}_i$  为  $i$  号分布源的平均静默时间,  $B_{channel}$  为单个边缘节点链路带宽, 控制平均发包时间, 便可得到不同的流量负荷。

考虑到对时延敏感性业务的 QoS 保证, 组包时间门限设定为 8 至 10 毫秒, 而语音和视频业务的传输最大延迟容忍一般可以达 100 毫秒左右。

图 5(a)、(b) 分别给出了自相似参数为 0.7、0.8、0.9 时, FAT(8ms) 算法、FAT-L(500Kbyte) 算法以及我们提出的 FAT-VL 算法, 三者随负载流量变化的平均突发时间对比。FAT 算法的突发时间不随流量负荷而改变, 极易引起突发同步造成的持续资源竞争, 同时, 其包长变化量很大, 在边缘节点处需要大容量的队列缓存, 且整体平均时延远高于 FAT-L 算法, 但后者在网络负荷量较低的情况下, 同样存在上述问题。而 FAT-VL 算法在低流量负荷情况下, 由于长度控制门限下调, 因而它的整体平均时延要优于前两种算法。

图 5(c)、(d) 示出了入口流量负荷为 0.1, Hurst 参数为 0.9, 时间门限 10ms 下的 FAT-VL 算法和 FAT-L 算法的突发时间对比图(仿真持续时间为 2 秒), 显然在平均流量低负荷情况下, FAT-L 算法仍然会出现突发同步现象, 即多个节点由于组装时间完成的一致性, 而引起的资源持续竞争问题, 而我们的 FAT-VL 算法彻底避免了这一缺陷, 在平均流量低负荷的情况下, 突发时间也会随流量大小的变化而产生变化。

图 5(e)、(f) 是对这一推论的仿真检验, 利用图四所示拓扑, 设定时间门限为 10ms, 在两种计数器界差(即  $Num$  的上界减去下界)下的 FAT-VL 算法和性能相对较好的 FAT-L 算法进行了较全面的比较, 显然, 我们的算法通过流量反馈式的智能调节, 不仅拥有更低的全网阻塞率, 而且获得更短的时延特性。结果还表明, 计数器界差较小, 调节更灵活迅速, 性能优化越明显, 当然这需要付出更多的硬件代价。

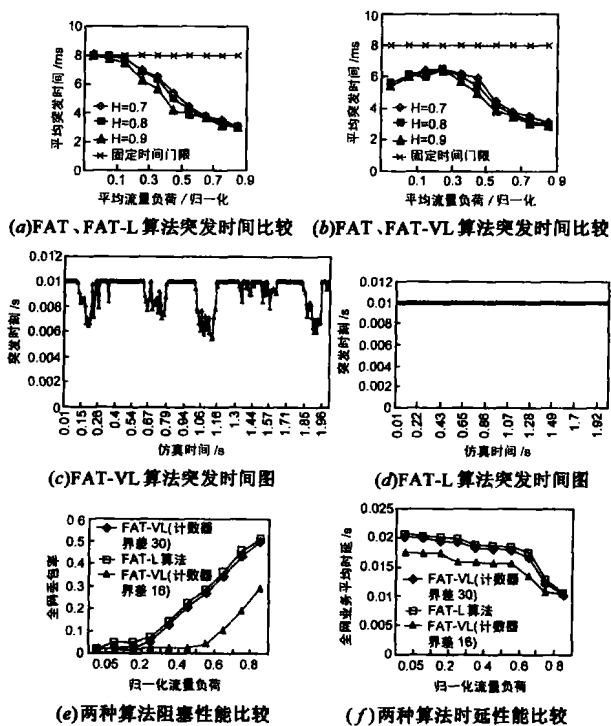


图5

#### 5 结论

仿真结果表明, 我们在文献[4, 7]两种组包算法的基础上提出的变长定时双门限智能组包(FAT-VL)算法, 具有随网络流量自适应变化的特点, 在低流量时, 算法可自动降低组包包长门限, 从而降低了组包等待时延, 提高了突发组装效率, 同时, 突发组装时间随流量的自适应变化可以消除不同边缘节点突发同步引起的持续阻塞问题, 有效降低 OBS 网络的阻塞概率。

参考文献:

- [1] J Turner. Terabit burst switching[J]. High Speed Networks, 1999, 8(4): 3-16.

- [ 2 ] M Yoo, C Qiao, S Dixit, QoS performance of optical burst switching in IP over WDM networks[ J]. IEEE J Select Areas Commun, 2000, 18( 10): 2062– 2071.
- [ 3 ] F Callegati, H C Cankaya, Y Xiong, M Vandenhoute. Design issues of optical IP routers for Internet backbone applications [ J]. IEEE Commun Mag, 1999, 37( 12): 124– 128.
- [ 4 ] Y Xiong, M Vandenhoute, H Cankaya. Control architecture in optical burst switched WDM networks[ J]. IEEE J Select Areas Commun, 2000, 18( 10): 1838– 1851.
- [ 5 ] C Qiao. Labeled optical burst switching for IP over WDM integration[ J]. IEEE Commun. Mag, 2000, 38( 9): 104– 114.
- [ 6 ] Adler R J, Feldman R E, Taqqu M S eds. A Practical Guide To Heavy Tails: Statistical Techniques and Applications[ M]. Boston: Birkhauser, 1998. 27– 53.
- [ 7 ] A Ge, F Callegati, L Tamil. On optical burst switching and self similar traffic[ J]. IEEE Commun Lett, 2000, 4( 3): 98– 100.
- [ 8 ] V Paxson, S Floyd. Wide area traffic: The failure of poisson modeling[ R]. SIGCOMM' 94, 1994. 257– 268.
- [ 9 ] S Manthorpe, I Norros, J Y Lboudet. The self similarity of TCP traffic[ R]. 瑞士 EPFL 技术报告, 1996. 7.

#### 作者简介:

李彦君 男, 1979 年 1 月出生于湖南省娄底市, 目前在北京邮电大学攻读光通信博士学位, 研究方向主要为网络性能与流量分析, internet 中的 QOS 保证及光突发交换. E-mail: gold007@ 263. net.

洪小斌 男, 2000 年于北京邮电大学光通信中心获博士学位, 2000~ 2001 年, 在朗讯科技从事网络规划研究; 目前在武汉邮电科学院北京分部从事网络规划、高速大容量光纤通信系统与宽带网络技术的研究.