

·光电工程系统技术·

光分组交换网络中光纤延迟线缓存技术

刘焕淋^{1,2}, 庞俊宇¹

(1. 重庆邮电大学通信学院, 重庆 400065; 2. 重庆大学光电技术及系统教育部重点实验室, 重庆 400044)

摘要:光分组交换网是全光网络发展的必然趋势. 然而, 光分组交换网络发展的瓶颈是光缓存技术. 目前, 在光域比较现实的还是采用光纤延迟线(FDL)作光缓存. 重点研究了光纤延迟线光缓存技术, 对 FDL 光缓存技术进行了深入的分析 and 归纳, 并对每一种光纤延迟线光缓存调度策略的优缺点都进行了细致的分析. 最后指出了光纤延迟线光缓存技术的未来研究重点和发展方向.

关键词:光分组交换; 缓存调度算法; 光纤延迟线

中图分类号: TN212

文献标识码: A

文章编号: 1673-1255(2009)01-0001-04

Fiber Delay Line Buffer Technology in Optical Packet Switching Network

LIU Huan-lin^{1,2}, PANG Jun-yu¹

(1. School of Telecommunication, Chongqing University of posts and telecommunications, Chongqing 400065, China;

2. Key Lab. of Optoelectronic Technology and System, Ministry of Education, Chongqing University, Chongqing 400044, China)

Abstract: The optical packet switching network is the necessary tendency in the development of all optical networks. But the bottleneck in developing the optical packet switching network is the optical buffer technology. At the present time, the most realistic approach is to use fiber delay line (FDL) as optical buffer. The FDL optical buffer technology is emphatically studied, analyzed and generalized further. The merits and faults of each FDL optical buffer technology are analyzed in detail. And the key issues and directions of future study in this area are pointed out.

Key words: optical packet switching; buffer scheduling algorithm; fiber delay line

新业务的不断涌现和 IP 业务的不断发展, 对传送网和交换系统容量的需求正以前所未有的速度增加, 波分复用技术的应用为信息提供了巨大的传送带宽, 网络瓶颈从传输环节移动到了交换环节^[1]. 光分组交换(optical packet switching, OPS)技术直接在光层上实现小颗粒的光分组交换, 具有较少的开销、高的带宽利用率, 是解决光通信“电子瓶颈”, 实现高速光传输的有效途径, 是光交换技术发展的方向和理想模式^[2]. 由于光分组交换中光分组作为最小的交换颗粒, 带宽可以按照用户的需求在

光分组级上进行灵活的分配, 所以它能够有效地支持多粒度的灵活 IP 业务, 提高带宽的利用率.

根据光分组到达时刻与长度大小的不同, OPS 网络可以划分为同步 OPS 网络和异步 OPS 网络^[3]. 在同步 OPS 网络中, 光分组只出现在固定长度的时隙中, 网络节点要对分组进行时间和相位上的同步校准, 边缘网络节点还要对分组进行封装、分拆和重组. 在异步 OPS 网络中, 分组可变化长度, 可以出现在任意时刻, 节点不必对其分拆和重组, 也不必实施分组级的同步. 显然异步 OPS 网络更灵活,

收稿日期: 2009-01-14

基金项目: 重庆邮电大学自然科学基金(A2008-61); 重庆市自然科学基金(KJ060513, KJ060508, KJ070513, KJ090511)

作者简介: 刘焕淋(1970-), 女, 重庆人, 副教授, 主要从事光交换和光网络研究工作; 庞俊宇(1984-), 男, 硕士研究生, 研究方向为光交换及全光网络.

但其流量特性比较复杂,由于光分组到达的时刻随意性较强,所以光分组在交换节点中发生冲突的几率也更大,交换节点需要对光分组进行复杂的调度,以提高交换网的性能.

如何解决光分组竞争是 OPS 网络一个至关重要的问题,对网络性能也将产生很大的影响^[4].在光随机存取存储器没实现前,解决光分组竞争有 3 种方法:(1)光纤延迟线(FDL)解决时间上的竞争;(2)偏射路由解决空间上的竞争;(3)波长转换解决波长上的竞争.其中,FDL 是光缓存的一种可选和最容易实现的方案,它在一定程度上能够减少分组/突发的丢包率,当 2 个分组竞争同一条输出链路时,一个分组被传输,另一个被送入一圈光纤经过充分的延迟以解决竞争问题.对 FDL 光缓存的结构和配置进行了介绍,同时对部分光分组的调度算法和缓存管理方法进行了分析和比较.

1 FDL 光缓存的结构和配置

1.1 光缓存的结构

在光分组交换网络中,根据 FDL 光缓存在 OPS 交换节点中的配置位置及功能,通常可分为以下 3 种结构^[5]:

(1) 输出光缓存

输出光缓存结构,如图 1 所示.在每个输出端配

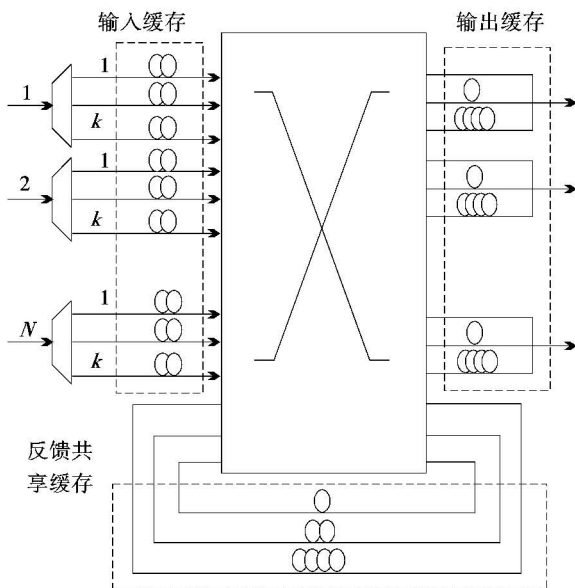


图 1 输出光缓存结构

置光缓存器.这种结构在光分组包进入缓存前,即需确定好延迟时间,其缺点是在控制分组的缓存时间长度时,灵活性较差.当同时到达的分组数大于缓存总容量时,则多余的分组就会被丢弃,造成丢包率增大.

(2) 输入光缓存

输入光缓存结构是在输入链路中引入缓存器,其优点是可以较为灵活地配置分组包的缓存时间,提高缓存器的工作效率.这一结构存在的最大缺点是存在分组头阻塞的问题.

(3) 反馈共享式 FDL 光缓存

共享缓存与输入/输出缓存不同,缓存资源配置于公共区域,交换机的多个输入/输出端口可通过交换模块接入到该区域缓存冲突的数据,缓存资源对各端口不是独占而是共享的,有利于减小 FDL 体积和提高 FDL 利用率.其缺点是分组在循环缓存过程中,会有放大噪声的积累.

1.2 光缓存的配置

在同一种类型的 FDL 光缓存中,FDL 又可以有不同的排列规律或配置方法.总体而言,可以将缓存配置分为 2 类:固定长度光纤延迟线(Fixed-FDL)和可变长度光纤延迟线(Variable-FDL)^[13].如图 2 所示,其中图 2a 配置中,所有的 FDL 都是等长的,一个进入缓存的数据包如果想获得足够长的缓存时间,那么这个数据包就需要循环多次.而图 2b 配置中,FDL 长度是逐步增加的,其根据需要遵循一定的规律.在 I-FDL 配置中,比较常见的有 3 种:简并式、双简并式和递增简并式^[6].

(1) 简并式(DB):如图 3a 所示,本配置类型中,FDL 光缓存中 B 根 FDL 的长度以 FDL 的粒度 D 为步长逐根递增.覆盖了从 D 到 BD 的所有范围.

(2) 双简并式(DDB):如图 3b 所示,本配置类型中,FDL 光缓存中 B 根 FDL 被分为奇数段和偶数段 2 部分,在这 2 部分中,FDL 的排队规律都是简并型式.第 i 根 FDL 的缓存长度为 $\lceil \frac{i}{2} \rceil \times D$, $i \in (1, 2, \dots, B)$, D 为 FDL 的粒度,表示对 x 向上取整.

(3) 递增简并式(IDB):如图 3c 所示,本配置类型中,长度为 D 的 FDL 的数目为 1,长度为 $2D$ 的 FDL 的数目为 2,依此类推,长度为 iD 的 FDL 的数目为 i ,其中, D 为 FDL 的粒度, $i \in (1, 2, \dots, n)$.

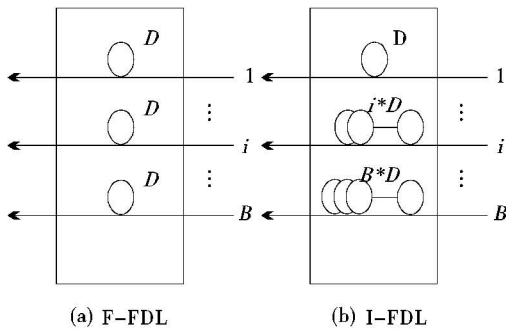


图 2 F-FDL 和 I-FDL

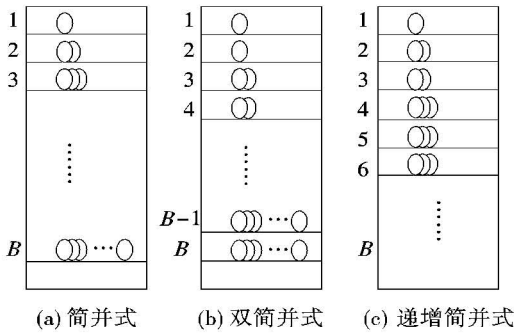


图 3 3 种 FDL 配置方法

2 FDL 光缓存典型调度算法的思想和优缺点

虽然 FDL 光缓存技术的原理简单,但 FDL 光缓存构建和分组包调度算法却至关重要,因为它们直接影响核心节点数据包的交换效率以及 OPS 网络的性能.FDL 光缓存调度算法可以说是 OPS 网络核心节点数据包交换的重点.

下面介绍几种典型的基于 I-FDL 配置的 FDL 调度算法:

(1) 抢占式短包优先调度算法

在此算法^[7]中,可变长 OPS 抢先式交换体系结构包括 N 个输入/输出光线接口,输入队列包括 N 个虚拟输出队列 (VOQ) 和一个按先入先出 (FIFO) 服务的短包队列. 当一个分组到达输入端口时,长度较长的分组分配到相应的 VOQ k (k 大于等于 0 小于等于 $N-1$) 队列中,长度较短 (小于 64 字节) 的分组被分配到 FIFO-S 中排队等待,在光的随机存储器没有应用的条件下,可以用光纤延迟线 (FDL)

存储光信号. 交换机调度器在每一个交换时隙内,根据输入队列请求和输出端口状态,按照短包优先的 PSPF 算法配置长分组和短分组的交换请求,通知获得确认的分组在下一时隙内传输分组到指定输出端口. 当分组到达输出端口时,短包存储在输出 FIFO-S 队列中,长包存储在输出 OQ-L 队列中,调度器首先输出 FIFO-S 队列中的短包,然后输出长队列中分组到光线上,在分组从输入端口输出到输出端口期间,调度保证分组传输的连续性.

优点:该算法简单、灵活、有效,非常适用于对分组传输时延有严格要求的实时 IP 业务,大大提高可变长 OPS 交换的性能.

(2) 共享反馈型缓存调度算法

Hiroaki Harai 提出的共享反馈型缓存调度算法^[8]用于交换处理异步可变长度数据包. 该算法又进一步细分为算法 1 和算法 2,其中算法 1 是作者思想的主要体现,算法 2 在算法 1 的基础上稍作改变. 该算法将进入核心交换节点的数据包归类为排队型、非排队型、排队型抢占和排队型竞争 4 种. 其中排队型竞争只在算法 2 中出现. 这里只对算法 1 进行详细介绍.

第 1 步:寻找可用外部输出端口. 如果数据包来自任意外部输入端口或者是数据包是排队型,同时外部输出端口 n 空闲 ($u_n < t_n$); 立刻将数据包传送到输出端口 n ,同时更新 $u_n = t_n + l_n$. 如果数据包来自任意外部输入端口,更新 q_n ,令 $q_n = u_n$. 同时,以端口 n 为目的地的排序型数据包变成了排队抢占型数据包.

第 2 步:为排队型数据包寻找可用缓存. 如果数据包在第 1 步中没有被传输至输出端口且可用缓存延迟线为 C ,进行这一步操作;数据包进入较小延迟的那条光纤延迟,此时 $q_n = t_n + l_n + d_i$, $s_i = t_n + l_n$. 缓存的数据包是排队型数据包.

第 3 步:为非排队型数据包寻找可用缓存. 如果数据包在第 2 步中没有变成排队型,且可用缓存延迟线为 C 进行这一步操作;数据包进入较大延迟的那条光纤延迟线. 这个数据包被看作是非排队型数据包.

第 4 步:如果在前 3 步后仍没被传输的数据包,直接被丢弃.

该算法可以尽最大能力缓存冲突数据包. 通过不同的条件下分类调度,减小了数据丢包率. 但是,难以避免经循环缓存后的分组包与新到达分组包再

次产生竞争冲突的可能性。

(3) 可变长度分组部分缓存和改写算法

文献[9]在部分共享缓存算法(partial buffer sharing scheme, PBS)^[10]的基础上提出了可变长度分组的部分缓存和改写法(vPBSO)。该算法基于优先级思想为不同类别的分组提供差异化缓存服务。论文中讨论2类优先级不同的数据,其中数据1的优先级比数据2的优先级高,在vPBSO情况下,当缓存占用期低于阀门时间时,缓存管理器允许所有到达的数据进入缓存器。当缓存占用期等于或大于阀门时间时,一个到达的类别为2的数据包跟在队列的尾部。另一方面,一个已经到达的类别为1的数据包被允许占用下一个位置,这个位置是为处在阀门值后的那个类别为1的数据所准备的。如果类别2数据在同一个位置等待,那么它们将被改写。被改写的数据意味着被丢弃。当优先级高的数据包到达率低于优先级低的数据包的到达率时,这个算法更有效。但是该算法当优先级高的数据包到达率高于优先级低的数据包的到达率时,vPBSO算法性能并不比PBS算法具有明显优越性。

(4) FDLs 光缓存快速树搜索调度算法

本算法^[11]适应于全光共享缓存 OPS,算法首先将数据包调度作为一种树形搜索问题对待,将时间域上每个时隙看成一个节点,每个时隙的下一时隙被看作为本时隙节点的子节点,以此类推,形象为一个树图。通过在底层每个节点安装一个处理器,实现由树尖向下搜索与由树根向上反馈信息的结合。通过众多次级处理器同时工作,进入交换节点数据包调度的理想路径将会在很短的时间内找到。本算法的优点是使得整个缓存时间复杂度大大降低。但是,由于需要装置许多处理器,交换节点结构的装置成本大大提高。

(5) 其他算法

Wang 和 Ningning Hu 等人研究了不区分长度的公平分组调度算法将导致 TCP 业务的性能下降^[12-13],文献[14]研究了基于服务优先级区分的高优先级分组在业务负载小于 0.4 时可以获得 10^{-12} 的分组丢失率,但业务负载增加时,OPS 性能劣化严重,而对 TCP 类实时业务的短包没有特别考虑,使短包的等待时间较长。而采用改进的分组调度算法可以减少分组的平均等待时间,提高网络性能。最大权重匹配(MWM)算法根据队列长度或分组等待时间分配权重匹配值^[15],目的是使长队列中

的分组或等待时间长的分组先得到调度,以减少分组的平均等待时间及提高分组的通过率。基于队列长度权重配置的算法没有考虑分组获得大权值时对其他短队列中大量短分组形成阻塞,致使长队列饿死,短队列或年龄大的信元阻塞实时性强的年龄小的短信元,从而使分组平均等待时间较长。

3 结 论

随着网络化时代的到来,人们对信息的需求与日俱增。IP 业务和多媒体数据业务在全球范围突飞猛进的增加,越来越要求建设高效、大容量、高质量、高带宽的光纤网络,也即全光网络。光交换的全光网络是通信网络发展的必然趋势,光数字分组交换是全光网络的最终选择。然而光分组交换网络的发展瓶颈是光缓存技术。本文主要是针对光分组交换网络中光缓存技术展开的。

光缓存器和存储器是组成光分组交换网的关键器件之一,受到广泛重视,已经成为当前的热点课题。目前全光缓存器的研究和应用,仍然处于百花齐放的状态,而目前在光域比较现实的还是采用 FDL 作光缓存,即利用 FDL 对光信号进行延迟的方法构造光缓存,对于光分组交换来说,一个合适的 FDL 缓存调度算法将极大地提高数据包在核心节点的交换成功率。但 FDL 无法实现随机的光信号存储操作,仅能提供离散的缓存时间,因此,FDL 光缓存的构造、配置及调度需要进行更为细致深入的研究。

参考文献

- [1] MUKHERJEE B. WDM optical communication networks: progress and challenge [J]. IEEE Journal on Selected Areas in Communications, 2000, 18(10): 1810 - 1824.
- [2] RANGARAJAN S, POULSEN H N, BLUMENTHAL D J. All-optical packet compression of variable length packets from 40 to 1500 B using a gated fiber loop[J]. IEEE Photonics Technology Letters, 2006, 18(2): 322 - 324.
- [3] TUCKER R S, WEN Z D. Photonic Packet Switching: an Overview [J]. IEICE Trans. On Commun, 1999, E82 - C(2): 254 - 264.
- [4] LIU Huan-lin, PAN Ying-jun, ZHANG Zhi-zhong. An improved first-fit-void-filling algorithm in asynchronous

(下转第 35 页)

相似黏度迅速下降,在实际使用时可能发生流失现象,污染光学系统,影响传感器的性能和指标,从而无法实现光电设备长期工作的要求。

(2) 较低的启动转矩

启动转矩是指润滑脂阻滞低速滚动轴承转动的程度^[3]。启动力矩值越小,则对稳定精度影响越小,启动功率消耗也越小,即润滑脂的启动性能越好。某些脂在低温下的启动力矩值较大,甚至出现卡住现象,即脂的低温性能差,影响运动平稳性,造成稳定精度下降,不适合在低温下应用。启动转矩是低温用润滑脂和宽温度范围用润滑脂在机载光电稳定平台应用的重要性能指标。

(3) 较宽的温度使用范围

在低温环境下,使用不合适的润滑脂最直观的结果是使又稠又软的油脂冻结,进而使得光电设备起动力矩大起启动困难,运动时产生较高的摩擦,功率损耗增大。

润滑脂受热后性质可能发生许多变化,从不流动状态变为流动状态,出现蒸发/氧化/析油等,这些改变都会影响润滑脂在高温下的使用性能。

(4) 严格的杂质含量

润滑脂内存在了机械杂质,使用过程中就会带入机械摩擦部位,它不但会降低润滑脂的减摩作用,而且加剧了摩擦交点和工作面的磨损,并能造成摩擦面擦伤等等,致使被润滑的滚珠轴承迅速丧失精度,降低运动精度,从而缩短使用寿命。

此外,机载光电稳定平台对润滑脂还要求一定的安定性,防止在与空气的接触中氧化;一定的抗水性,在与水及水蒸气接触后不会生成乳化体;一定的机械稳定性,在受机械作用后其稠度改变很小。

对于所用润滑脂的选择,除了以上各方面的阐述之外,还有其他许多制约条件,应对整个系统有一定了解,根据使用寿命、载机飞行高度、速度、使用场合是陆地还是海上等因素进行综合选择,使机械润滑更加合理完善。

参考文献

- [1] 颜志光. 润滑材料与润滑技术[M]. 北京:中国石化出版社,2000.
- [2] 徐灏. 机械设计手册[M]. 北京:机械工业出版社,1991.
- [3] 王焯军,曹小平,张有智. 军事装备润滑剂应用技术[M]. 北京:国防工业出版社,2005.
- [4] 季伟,张民,叶培大. 光分组交换网络中的光缓存技术研究[J]. 光通信技术,2004,28(10):26-28.
- [5] 杨俊杰,肖石林. 光缓存配置研究[J]. 光通信技术,2005(12):15-18.
- [6] 刘焕淋,陈前斌. 可变长光分组交换的短包优先调度[J]. 重庆邮电大学学报(自然科学版),2007(1):76-80.
- [7] Hiroaki Harai, Motoshi Suzuki and Takeshi Ozeki. Buffer Management for Shared Feedback Buffer-Type Optical Packet Switches[C]//National Institute of Information and Communications Technology,2006.
- [8] Hiroaki Harai, Masayuki Murata. Optical Fiber-Delay-Line Buffer Management in Output-Buffered Photonic Packet Switch to Support Service Differentiation[J]. IEEE journal on selected areas in communication,2006,24(8):108-116.
- [9] Lin Y and Silverster J A. Priority queueing strategies and buffer allocation protocols for traffic control at an ATM integrated broadband switching system [J]. IEEE J. Sel. Areas Commun,1991,9(12):1524-1536.
- [10] 颜志光. 润滑材料与润滑技术[M]. 北京:中国石化出版社,2000.
- [11] 徐灏. 机械设计手册[M]. 北京:机械工业出版社,1991.
- [12] 王焯军,曹小平,张有智. 军事装备润滑剂应用技术[M]. 北京:国防工业出版社,2005.
- [13] Soung Y Liew, Siong F Law. A fast scheduling algorithm for all-optical shared-buffer packet switches[C]// Faculty of Information and Communication Technology, 2008.
- [14] WANG R, PAU G, YAMADA K, et al. TCP startup Performance in Large bandwidth delay networks [EB/OL]. (2004-12-19)[2006-11-12]. Http://ieeexplore.ieee.org/iel5/9369/29790/01356968.pdf.
- [15] NINGNING H, STEENKISTE P. Improving TCP startup performance using active measurements algorithm and evaluation [EB/OL]. (2003-11-13)[2006-11-10]. http://ieeexplore.ieee.org/iel5/8838/27972/01249761.pdf? arnumber=1249761.
- [16] II G Y, ABDULLAH M, DIMYATI K. Preemptive Hybrid Scheduling Scheme for Optical Packet Switch with Class Differentiation [EB/OL]. (2006-08-20) [200611-15]. Http://ieeexplore.ieee.org/iel5/10826/34121/01625805.pdf.
- [17] MCKEOWN N, ANNANTHARAM V, WALRANDJ, et al. Achieving 100% throughput in an input queued switch[J]. IEEE Trans. Commun,1999,47(8):1260-1267.

(上接第 4 页)