

## 摘要

随着业务需求的增长和光传输技术的高速发展，光网络将逐渐由光传输、电交换向光传输、光交换系统演变。光突发交换（OBS）作为一种过渡性交换技术，以其良好的数据单元支持性和可实现性受到广泛关注。目前的光纤城域网大都采用网络结构确定、硬件实现简单的环形拓扑，随着 WDM 技术的日趋成熟，应用 OBS 思想的 WDM 城域光纤环网成为研究热点之一。

针对这一研究热点，本文首先提出了一种新型光突发环网的设计方案—虚同步光突发环（VS-OBRR）。该方案将数据信道划分时隙，采用弹性偏移时间控制机制（ROTC）实现控制信道中突发控制标签（BCL）与数据信道中数据突发时隙（DBS）之间的持续映射和有效控制；其主要优点是支持光纤链路空间重用、信道接入快速灵活、支持可抢占的优先级，并且既不要求在结点配置光纤延迟线也不要求全网结点严格同步。文中详细介绍了 OPNET Modeler 10.5 仿真软件下 VS-OBRR 的整体设计、结点的模型建立和网络中传输的各种分组格式，旨在构建一种灵活、扩展性强的通用 VS-OBRR 仿真模型。最后，根据大量仿真数据，分析了 VS-OBRR 的最大吞吐量、丢失率等关键性能；结果证明 VS-OBRR 较已有环网控制方案有更优良的性能。

**关键词：**光突发交换 光突发环网 突发控制标签 数据突发时隙 虚同步光突发环

## Abstract

With the increment of traffic requirements and the development of optical transmitting technology, the current optical network is evolving into an all-optical system with optical transmitting and optical switching capabilities, while it was an optical transmitting and electrical switching system. Optical Burst Switching (OBS), regarded as a transitional switching technology, attracts more attention for its burst-based transporting and technological feasibility. At the same time, WDM ring metropolitan networks are usually adopted thanking for mature WDM technology and simple ring topology. Now, a WDM ring metropolitan optical network with OBS technology has become one of the heated studying points.

Firstly, a new method of controlling an optical burst ring network is presented here, called Virtual Synchronized-Optical Burst Ring. The presented scheme splits all data channels into small slots and uses an effective resilient off-time control mechanism (ROTC) to keep the offset time large enough between the burst control label (BCL) in the control channel and the data burst slot (DBS) in the data channels; its main advantages are its supporting the space reuse of fiber links, fast and flexible bandwidth accessing, permitting preemptive priority and requiring neither FDLs in each node nor strict clock synchronization through the network. Then, the simulation modules of initialization node, source node, Virtual Synchronized-Optical Burst Add-Drop Multiplexer (VS-OBADM) node and all packet formats presented in this system are described in detail under OPNET Modeler 10.5 software. The aim of constructing these modules is to build a generalized VS-OBR simulation model with flexibility and good expansibility. Finally, the simulation results are analyzed and show the good performances of VS-OBR in maximal throughput, loss probabilities, end to-end delay and so on.

**Keywords:** Optical Burst Switching   Optical Burst Ring Networks   Burst Control Label   Data Burst Slot   Virtual Synchronized-Optical Burst Ring

## 独创性声明

本人声明所呈交的论文是我个人在导师指导下进行的研究工作及取得的研究成果。尽我所知，除了文中特别加以标注和致谢中所罗列的内容以外，论文中不包含其他人已经发表或撰写过的研究成果；也不包含为获得西安电子科技大学或其它教育机构的学位或证书而使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中做了明确的说明并表示了谢意。

申请学位论文与资料若有不实之处，本人承担一切相关责任。

本人签名：李波

日期：2007.3.5

## 关于论文使用授权的说明

本人完全了解西安电子科技大学有关保留和使用学位论文的规定，即：研究生在校攻读学位期间论文工作的知识产权单位属西安电子科技大学。本人保证毕业后离校后，发表论文或使用论文工作成果时署名单位仍然为西安电子科技大学。学校有权保留送交论文的复印件，允许查阅和借阅论文；学校可以公布论文的全部或部分内容，可以允许采用影印、缩印或其它复制手段保存论文。（保密的论文在解密后遵守此规定）

本人签名：李波

日期：2007.3.5

导师签名：刘爱军

日期：2007.3.5

## 绪论

随着信息社会的发展,各种新业务不断涌现,近年来 Internet 上通信业务量呈爆炸性增长。现在加入 Internet 的用户已超过 2 亿,并以每年 165% 的增长率在全球扩展;Internet 上的数据量每 100 天就翻一番,每 30 分钟就需增加新的网络连接骨干网,每 6~9 个月带宽就得涨一倍。传统的通信网络越来越难以适应这些需要。随着光电子技术的迅猛发展,传统通信网络正在向新一代网络演变。

1966 年英籍学者 C.K.Kao (高锟) 和 C.A.Hockham 发表了关于传输介质新概念的论文,指出了利用光纤传输信息的可能性和技术途径,奠定了现代光纤通信的基础。继美国人梅曼发明了第一台红宝石激光器之后,氦-氖激光器、二氧化碳激光器的发明和完善极大地促进了光通信的发展。光通信已经由点到点的传输系统迅速演变为大规模的复杂网络。在当前光通信发展进入这样一个承上启下的新阶段,人们热衷于两方面的研究:一方面在传输领域追求大容量、高速率和长距离;另一方面开始研究新的光交换和组网技术,寻找更加经济、灵活、可靠的途径将光纤通信的巨大传输容量提供给用户使用。

掺铒光纤放大器(EDFA)和波分复用(WDM)技术先后成熟并商用,使光通信传输领域不断取得突破。在 2005 年度的光纤通信会议 OFC 和 OCEC 上报道了单纤 10 波长,每波长信道 320Gb/s 的 WDM 实验系统<sup>[1]</sup>;而全部使用拉曼放大器的 40×40Gb/s 的海底光缆传输系统传输距离超过 10 000 公里<sup>[2]</sup>。目前,采用 WDM+光放大器的实验系统可实现 159×40Gb/s 无电中继传输 6 100km<sup>[3]</sup>, 373×10Gb/s 传输 11 000km<sup>[4]</sup>。

在网络方面,以 IP 为基础的网络智能化技术和以 WDM 为基础的光层组网技术结合成为当前发展的一种趋势。最新提出的 IP 分组直接在 WDM 层上传输的网络模型(IP over WDM)较以前的 IP over ATM 和 IP over SDH/SONET 减小了网络功能冗余,节省了协议开销,避免了 ATM/SONET 设备花费;同时它结合光层和 IP 层的路由和恢复功能,提供了动态的波长路由及类似于 SDH 的恢复性能。但相比高速发展的传输领域,由于电子器件数据处理速率的固有限制,光通信网络需在电域进行的控制、交换等电处理能力已严重滞后,渐渐成为光网络的瓶颈。光信号仅被用于信息的传输,光纤仅当作“导线”,并未发挥出巨大带宽等优越性。为解决电子瓶颈限制,降低交换成本,研究人员开始在交换系统中引入光子技术,实现光交换。

本文即在光交换的研究基础上探索新的 WDM 环网交换控制技术。文中各章内容安排如下:

第一章介绍新型光交换技术和已有的几种 WDM 环网控制方案；

第二章论述本文提出的新型光突发环网控制方案的结构和工作原理；

第三章描述实现这种新型光突发环网控制方案的仿真模型，该模型是在 OPNET 网络仿真软件平台下实现的；

第四章根据仿真结果对这种新型光突发环网控制方案的一些重要性能进行了分析。

## 第一章 光交换技术和 WDM 环形光网络

### 1.1 基于 WDM 的新型光交换技术

基于 WDM 的新型光交换技术是指由源结点发出的信号在中间节点不经过任何光/电/光转换, 直接从输入端交换到不同的输出端, 在宿结点才将信号转换到电域处理。这样端到端的全光数据传送技术有以下特点:

- (1) 提高结点吞吐量。光交换不受检测器、调制器等光电器件响应速度的限制, 可以大大提高交换单元的吞吐量。
- (2) 降低交换成本。光信号在通过交换单元时, 不需要经过光/电/光转换, 可以省掉昂贵的光电接口器件。
- (3) 透明性。光交换对比特率、信号调制方式和通信协议透明, 具有良好的升级能力。

WDM 全光交换技术包括光线路交换 (OCS, Optical Circuit Switching), 光分组交换 (OPS, Optical Packet Switching) 和光突发交换 (OBS, Optical Burst Switching) 等, 它们具有不同的交换控制方式和颗粒度, 对网络交换结点的控制处理能力、光域交换能力、器件响应速度和光域存储能力也有各自不同的要求。

#### 1.1.1 光线路交换技术 (OCS)

OCS 采用类似于电话交换系统中的信令方式, 通过双向预约资源的信令机制建立起边缘结点之间的透明光通道。透明是指交换结点将输入波长上的信号不加改变地转接到指定的输出波长上, 与这些波长上的传输速率、调制方式、编码格式和承载业务等无关。OCS 中的交换结点采用光交叉连接设备 (OXC, Optical Cross-Connect), 通过光纤链路相互连接, 中间结点不需要使用光环存。每条链路中光纤的数目、每根光纤上波长信道的数目都会影响到网络上可以建立的波长通道连接的总数。

OCS 网络交换过程分为三个阶段: 光路建立、光路保持、光路释放。OCS 通过双向信令建立连接, 如果连接持续时间较短, 开销就会增加。建立连接所需的时延在大规模网络中就成为一个问题。

波长转换器 (WC, Wavelength Converter) 对 OCS 网络的性能具有重要的影响。在光网络中, 不具备波长转换能力的 OXC 称为波长选择交叉连接设备

(WSXC, Wavelength-Selective Cross-Connect)。经过 WSXC 结点的波长通道, 输入波长和输出波长必须相同, 否则连接无法建立, 这一限制条件称为“波长一致性限制”。具备波长转换能力的 OXC 称为波长转换交叉连接设备(WIXC, Wavelength-Interchanging Cross-Connect)<sup>[5]</sup>。通过 WIXC 建立起来的波长通道可在不同的光纤链路上使用不同的波长, 称之为虚波长通道。虚波长通道放松了对波长一致性的要求, 但是要想完全不受波长一致性限制的束缚, 网络中必须全部采用具有全范围波长转换能力的 WIXC, 这无论从技术能力或经济成本上考虑, 目前都不现实。然而使用有限范围波长转换器, 也存在器件经济成本的问题。WC 的使用提高了波长利用效率, 但以波长为单位动态建立连接制约了带宽颗粒度, 所以 OCS 很难直接支持颗粒度小动态性很强的 IP 业务。

### 1.1.2 光分组交换技术 (OPS)

OPS 是将电分组交换的思想应用在光域。数据净荷以光分组的形式在波长上传送, 控制所需的地址、优先级、净荷类型等分组头信息与净荷一起传送, 在交换结点不对全部分组进行光/电/光转换的情况下实现分组头的提取、改写和插入, 净荷在光域进行缓存和转发。

由于信道速率极高, 这种操作并非易事。目前一种较有希望的方案是采用副载波调制(SCM, Sub-Carrier Modulation)技术<sup>[6]</sup>, 将分组头信息以低频副载波的形式调制到分组净荷部分, 不同的波长上采用不同的副载波频率。在交换机的接收侧通过滤波器分离副载波, 进行信息提取、处理和重写。之后在输出端口通过光波调制技术将更新后的分组头加载到输出分组上。由于头信息的处理需要一定的时间, OPS 的输入端口会采用一定长度的光纤延迟线(FDL, Fiber Delay Line)来抵消分组头和净荷之间的时延差异。

OPS 中的关键技术还包括冲突解决、突发分组接收、快速时钟恢复、交换结构和光缓存技术, 全光分组传输和交换中基于光分组的功率均衡、色散、串扰、非线性等。一些技术的实现目前存在以下难点:

(1) 光分组交换通常工作于同步状态, 提取和插入分组头, 需要时钟同步, 难度较大。

(2) 光分组交换本质上是一种存储转发机制, 现阶段缺乏可以实现随机存取的光缓存设备。在研究中采用光纤延迟线(FDL)充当光缓存, 起抵消分组头和净荷之间时延差作用时, 需要配合光开关等器件使用; 起减小接收冲突的作用时, 只能将相应的光分组数据按序延时一定的时间, 而且通过增加延迟线数目来满足大容量缓存需求不实际。

(3) 光分组交换中净荷的转发是通过对光交换矩阵的配置实现的, 而配置光交

换矩阵的延时有较严格的要求。

综上, OPS 虽然直接对应电域分组交换, 对 IP 分组业务的支持最为理想, 但目前人们在光域的处理手段还不够成熟, 造成这些关键问题没有理想的解决方案, OPS 的实用化受到影响。

### 1.1.3 光突发交换技术 (OBS)

OCS 以波长作为网络资源的基本调度单位, 因颗粒度太大而缺乏灵活性, 尤其不适应高突发性和未来数字视频业务; OPS 又因其所需器件难以实现的限制未能走出实验室; OBS 作为 OCS 和 OPS 的折中, 充分利用了电子技术的处理优势和光子技术的速度优势, 逐渐成为光通信领域交换技术的研究热点之一。

OBS 采用控制信道和数据信道分离, 单向预约带宽资源的方法, 实现以数据突发为基本单元的信息转发。网络中传输的基本数据单元是将一些具有相同特性的分组按某种汇聚算法聚合而成的数据突发 (DB, Data Burst)。DB 的长度可以是固定的, 也可是变化的。突发粒度的交换比波长粒度更灵活, 且更适应网络中数据的强突发性。每个数据突发配有一个控制头, 称突发控制分组 (BCP, Burst Control Packet), BCP 中含有该数据突发的相关信息, 用于在其经过的网络结点预留带宽。在空间域上, BCP 使用专用的控制波长信道, 经光/电/光转换在电域处理, 而 DB 不经光/电/光转换, 在数据波长信道上全光传输; 在时间域上, BCP 比 DB 提前一段时间发送, 这段时间被称作偏置时间 (OT, Offset Time), 保证途经结点有足够的时间处理控制信息, 为后续到达的 DB 预约资源, 这样可以避免使用光缓存设备。

OBS 的关键技术如边缘结点的组装算法, 交换控制协议, 资源调度机制, 竞争解决等, 将在下一节中介绍。

## 1.2 光突发交换技术

图 1.1 给出了一般的 OBS 网络结构。OBS 网络由边缘结点 (EN, Edge Node) 和核心结点 (CN, Core Node) 组成。边缘结点负责实现用户业务的接入、汇聚、DB 的组装/拆分和 BCP 的生成等功能, 核心结点主要负责 BCP 中控制信息的提取、识别和重新插入、DB 的路由与转发等功能。承载用户业务的 IP 分组在边缘结点按照目的地址、业务优先级等信息进行汇聚并组装成较大的数据突发 DB, 同时边缘结点生成一个突发控制分组 BCP 并先于 DB 发送, 该 BCP 携带相应 DB 的到达时刻、持续时间、目的地址、业务级别等控制信息 (其作用类似于 OPS 的分组头部), 核心交换结点根据 BCP 确定转发路由、设置内部交换通路。



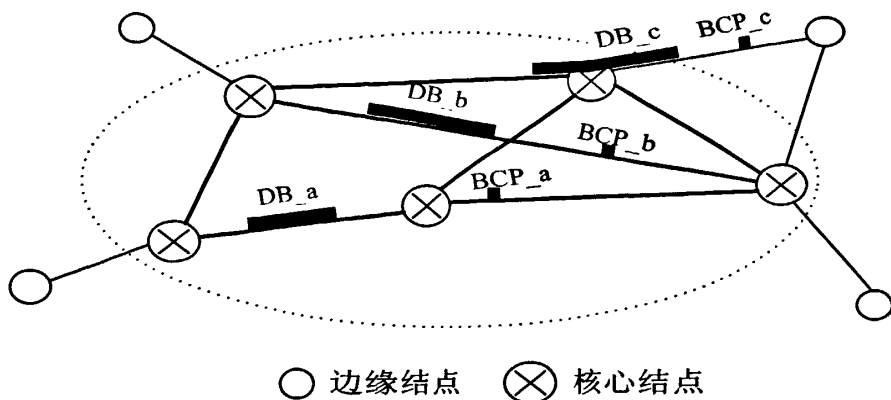


图 1.1 OBS 网络结构

### 1.2.1 组装算法

边缘结点的组装算法对 OBS 网络性能具有重要影响。突发的组装一般是按照数据分组的目的地、优先级信息将数据置入对应的组装缓存器中，再按一定的组装算法将分组组装成 DB。组装算法可分为基于固定门限和基于可变门限两种。

固定组装时间 (FAT, Fixed Assemble Time) 算法、固定组装长度 (FAL, Fixed Assemble Length) 算法和同时考虑时间和长度的基于最大组装时间最小组装长度 (FMTML, Fixed Max-Time Min-Length) 算法都属于固定门限组装算法。FAT 算法是当最大组装时延到一定时间后就组装 DB，并产生一个 BCP。这种方式比较简单，但当网络负载较大时，DB 可能过长，影响整个网络的性能。FAL 算法是当组装缓存器中的分组长度达到某一个大小时就组装 DB。这种方式也比较简单，但当网络负载过小时，组装时延不可接受，尤其对实时业务。FMTML 算法是一种比较常用的方式，当分组等待组装的最大时延达到允许的最大值或缓存器中的分组长度达到最小组装长度时就生成 DB，但它缺乏门限的灵活性。

可变门限算法包括随机变化门限组装算法、根据业务到达改变组装门限的自适应组装算法<sup>[7]</sup>和基于预测的组装算法<sup>[8]</sup>。随机变化门限组装算法在组装时控制长度或时间随机变化，以获得预期的突发间隔或长度分布特性。自适应门限组装算法结合业务强度来调整长度和时间门限，一定程度上降低了突发时延，但在 BCP 发送出去之后，由于 BCP 中记录的突发长度信息无法改变，DB 不能再继续装入新到的数据，从而造成数据组装时延增加。预测组装算法针对这一问题，利用到达业务流的实时统计特性来预测出数据突发可能达到的长度，从而降低不必要的突发时延。突发长度预测主要利用线性预测滤波 (LPF, Linear Predictive Filter) 技术来实现。

### 1.2.2 OBS 网络中的信令协议

如何使 DB 与 BCP 高效协调的工作, OBS 网络中的信令协议解决这个问题。按照控制分组是否明确通知交换结点 DB 的发送开始时间和结束时间, 以及结点何时给 DB 分配波长信道, 可对信令协议分类, 但文献中存在着不同的分类方法。本文按通信路径的建立和拆除方式将信令协议分四类:

(1) 显示建立、显示拆除。当建立消息 SETUP 到达时, 立刻进行交叉矩阵的配置及输出波长的预约。该配置直到收到 RELEASE 消息时释放。

(2) 显示建立、估算拆除。SETUP 消息到达时立刻建立, DB 的结束根据 SETUP 消息的到达时刻和 SETUP 消息中携带 DB 的持续时间信息来估计。

(3) 估算建立、显示拆除。SETUP 消息中包含 DB 的偏置时间信息, 用于建立时间的估计, 而拆除是当 RELEASE 消息到达时释放资源。

(4) 估算建立、估算拆除。利用 SETUP 消息中 DB 的偏置时间和持续时间信息估算通信路径的建立和释放时刻。

JET (Just-Enough-Time) 协议是当前被广泛应用的一种估算建立、估算拆除信令协议。源结点在发送 DB 之前, 先在控制信道上向目的结点发送一个 BCP, BCP 中包含 DB 的长度信息和偏置时间信息, 在中间各结点进行光/电/光转换, 并每经过一个结点减去一倍的结点处理时间; 各结点自动完成波长通路的选择、带宽资源的预留释放和交叉连接的建立拆除; DB 等待一个初始偏置时间后才发出, 之后完全在光域中传输, 适当的偏置时间可抵消链路的建立和处理时延, 因此每个中间结点都无需光缓存。

### 1.2.3 OBS 网络资源调度机制

OBS 网络交换结点中数据信道的调度是设计网络时需要考虑的关键问题。调度算法的目标是最大限度地利用带宽, 减少冲突的发生。这里考虑的调度机制都是基于结点具有波长转换能力的, 因为若无波长转换也就无所谓调度, 当前波长可用数据就可发送, 不可用就要丢弃。

数据信道调度机制一般可分为不带时隙填充和带时隙填充两类。最迟可用未占用信道 (LAUC, Latest Available Unused Channel) 算法属于不带时隙填充的调度算法, 其基本思想是选择信道上最后一个突发的结束时刻与新到 DB 的开始时刻间隔/时隙最小的信道。LAUC 算法的主要优势是实现简单, 但因只记录一个值, 即未调度时间, 使得突发间隔不能被充分利用。可插空最迟可用未占用信道 (LAUC-VF, Latest Available Unused Channel with Void Filling) 算法属于带时隙

填充的调度算法，与 LAUC 不同，LAUC-VF 算法中信道上两个 DB 占用的时间段之间的间隔可以填充新到的 DB。根据新到 DB 的头或尾与旧 DB 的间隔大小的比较，可将 LAUC-VF 的实现分为最小开始时隙（Min-SV, Min-Starting Void）填充算法，最小结束时隙（Min-EV, Min-Ending Void）填充算法和最适合（Best Fit）填充算法，它们的实现复杂度越来越高，但能更有效充分的利用突发间隔。

### 1.2.4 冲突的解决

当来自不同输入端口的分组要在某个时刻从同一个输出端口输出的时候，分组输出发生了冲突或称端口竞争。冲突引起分组丢失，因此冲突的解决也是 OBS 网络中的关键技术之一。

目前 OBS 网络中的冲突解决可从以下几个方面入手：

(1) 波长域。通过波长转换的手段，可将冲突的数据在不同波长上发送。这一技术对 WC 器件有很强的依赖。

(2) 时间域。采用光缓存（如 FDL），将 DB 延时，直到不会产生冲突为止。但现在的光缓存还不能像电域中的缓存一样随机存取，FDL 只能提供固定有限的时延，数据按先进先出的顺序离开 FDL。

(3) 空间域。将 DB 发送到结点另一空闲端口上，并沿不同的路由到达目的端，这种利用网络空间维度解决冲突的方法称偏射路由技术。

(4) 部分丢弃方式。在发生冲突时，并不丢掉整个 DB，而是只丢弃发生冲突的部分。这有赖于 OBS 发送的是多个 IP 分组组成的突发，部分丢弃后还有大量 IP 分组可到达目的地。这种方法并不能完全解决冲突，但可减少冲突带来的数据损失。

除以上关键技术外，OBS 网络中还有 QoS 支持、路由选择等重要问题，这里就不再介绍。

## 1.3 WDM 环形光网络的控制问题

城域网（MAN, Metropolitan Area Network）承载的业务量随着高速数据、IPTV 和视频通信等宽带实时应用的推广正迅猛增长。由于巨大的带宽资源和网络结构的确定性，光纤环网成为最广泛应用的城域网。这些环网的周长一般在几十到几百公里，环上通常有几个到几十个称为分插复用器（ADM, Add-Drop Multiplexer）的结点。随着 WDM 技术的成熟，当前许多基于单纤或者双纤的城域环网将升级为 WDM 环网。WDM 技术带来了传输速率的大幅度提高，传统环网中基于光/电

/光变换的 ADM 结点的处理速度难于继续支持如此高的传输速率，并且环网中每个结点处理的业务多数是转发给下游结点的过路业务，而非发给本地结点的下路业务。这种情况下，解决 WDM 环网结点处理瓶颈的主要思路就是让过路业务以直通（Cut Through）方式在光域直接传递到下游结点，无须变换到电域进行处理。

为了在光域实现业务信息从源点到目的点的直通传送，通常采用的方法是在 WDM 环网中单列出一个或多个波长来传输控制信息，称为控制（波长）信道。各结点将这些控制信息变换到电域处理，从而实现网络控制面的功能。光纤中其余的波长信道用于业务传送，相应称为数据（波长）信道。这种控制信道和数据信道分离、控制信道对传送业务进行控制的特点，和 OBS 网络有很多相似之处，因此，将 OBS 技术应用于环网的研究受到了很大重视。

OBS 在 WDM 环网中直接应用，可得到一个分布式随机接入的环形光网。每个结点需要发送数据突发时先观察是否有空闲数据波长，如果有，就发送突发控制分组 BCP，在一段偏置时间之后发出数据突发 DB。BCP 在控制信道中传输和处理，通知环上其它结点要到达的 DB 是过路业务还是下路业务，使中间结点可以提前设置交换矩阵。显然，这是一种随机接入方式，存在本地上路业务和过路业务之间的冲突问题。但是由于环网拓扑的特殊性，如果控制协议设计合理，一般 OBS 网络中难以解决的突发冲突完全可以在不采用波长转换器或者光缓存的情况下予以避免。

综合 WDM 网络的特点，利用 OBS 技术实现环网控制方案需要解决以下几个问题：

- (1) 如何确定光域数据信道的接入颗粒度；
- (2) 如何利用控制波长信道实现对数据波长信道的划分和标识；
- (3) 控制信道中的 BCP 与相应的数据信道中的 DB 的时间关系问题；
- (4) 高优先级上路业务与低优先级过路业务之间的资源竞争问题。

数据信道的接入颗粒度，是指媒质接入控制方案允许结点占用数据波长信道的最小单位，以及该占用单位是否有多种选择。较大的颗粒度易于实现，但是可能造成带宽浪费；较小的颗粒度有利于带宽重用，使得资源利用率高，但是控制开销大。不进行固定划分的随机占用方式快速灵活，但会相应引起冲突问题。如果进行波长信道的再次划分，则数据信道和控制信道之间必须引入某种对应机制，解决数据（子）信道识别问题。由于控制信息经过每个结点都需上升到电域进行处理而数据信息直接在光域穿过结点，因而跨越结点时控制信息经历的时延大于数据信息，控制方案必须保证携带控制信息的 BCP 在 DB 之前到达目的结点，才能实现正确控制。当用户数据存在不同的优先级时，高优先级应该比低优先级享有更高的发送权限，对于本地业务，这种服务区分通过发送队列调度即可实现，但是在上路业务和过路业务之间难以实现这种区分。

因为在普通的 OBS 环网中, 结点根据本地数据突发组装情况来决定发送时刻和信道占用时间, 占用颗粒度的最小值由组装出的最小突发长度决定, 分布式控制的环网中, 这种随机占用不可避免的引起冲突问题。BCP 在控制信道中通过偏置时间 OT 和长度指示来标识数据信道中传送的 DB, 因而源结点在发送 BCP 时, 必须根据目的结点地址的距离远近合理选择 OT 的初值, 以保证突发能正确抵达目的地。由于 OT 与目的地址相关, 而且突发的长度不定, 结点很难占用过路低优先级业务的带宽, 无法很好的支持全网服务等级区分。

## 1.4 几种已有的 WDM 城域环形光网络技术

### 1.4.1 采用 TTFR 的 OBS 环网方案<sup>[9]</sup>

在可调谐发送机固定接收机 (TTFR, Tunable Transmitter and Fixed-Tuned Receiver) 方案的环网结点中, 采用一个可调谐发射机向任意数据波长发送 DB, 一个定频接收机从分配给结点的固定波长上接收发送给该结点的数据, 这是一种接收波长固定、发送波长随意的资源分配方式。控制信道独立采用一对固定频率收发机接收和发送 BCP。突发和控制信息的时间对应采用 JET 协议。为解决多个结点向同一个目的结点发送 DB 时可能产生的冲突, 该方案采取了类似 CSMA/CD 的控制机制。当一个结点需要发送 DB 给某个结点时, 它首先通过检查记录的 BCP 信息来确认该目的结点的固定接收波长上在本结点要发 DB 时没有突发在传输。确认之后立即发出 BCP, 经过一段偏置时间后发出对应的 DB。由于发送的随机性, 该 DB 发送期间上游结点可能有突发到达, 从而引起冲突。可以采取上游优先和最早到达优先两种策略来解决冲突。上游优先策略是指发生冲突时, 本地结点立即停止当前 DB 发送, 优先保证上游的 DB 通过。该策略可以保证源结点一旦成功发送数据即可成功送达, 但是中间结点发送的突发会被中断并形成突发碎片。最早到达优先策略则是在冲突发生时, 优先保证本地结点完成发送, 上游发来的突发被本地结点终止。这种策略可以避免环上产生突发碎片以及相应的带宽浪费, 但是会造成上游结点发出的数据突发丢失 (这同样属于资源浪费)。

总体来讲, 采用 TTFR 的 OBS 环网方案实现简单, 控制容易, 但是因为接收采用专用固定波长, 限制了网络的可扩展性并降低了资源利用率。同时由于无法避免发生冲突, 存在已经占用信道的 DB 由于冲突而被丢弃的问题, 导致带宽浪费。最后, 由于 DB 发送偏置时间与目的结点的位置有关且突发长度不定, 源结点发送操作比较复杂, 中间结点难以实现对过路低优先级突发的资源抢占。

### 1.4.2 基于 ODD 协议的 OBS 环网<sup>[10]</sup>

Lisong Xu, 和 Harry G. Perros 等人针对 OBS 环网控制信道和数据信道的速度不匹配问题, 提出了一种仅考虑目的结点时延的 ODD (Only Destination Delay) 协议。在这种环网结构中, 环上共有  $N$  个结点, 链路上的波长数目为  $N+1$  个, 其中一个用作控制信道, 其余  $N$  个分配给每一个结点用于 DB 发送, 称为结点的固有波长。除了控制信道收发机, 每个源结点用一个固定频率发射机来发送 DB, 用一个可调谐的接收机从各波长上接收 DB。因此该方案不存在发送冲突; 而由于结点在同一时间只能从一个波长上接收数据, 存在接收冲突。

该方案将控制波长信道的时间分成周期性出现的控制帧, 每个控制帧分为  $N$  个时隙, 结点使用帧中的对应时隙来发送目的地址、偏置时间和突发持续时间等信息。选择固定长度的控制帧和与结点对应的控制时隙, 有利于采用硬件实现高速处理和控制。控制帧在每个结点都需要进行电处理, 因此存在控制信道和数据信道速率不匹配问题。为了避免每个结点在发送数据时都要根据目的结点的位置重新计算偏置时间 OT, 引入 ODD 协议。该协议假设控制信息在每个结点的处理时延相同; 而目的结点的处理时延可能比中间结点稍长一些。在每个中间过路结点上使用一段固定长度的 FDL 弥补控制信道与数据信道之间的时延差异, 而偏置时间设置为目的节点的处理时间。ODD 协议实现了控制信道和数据信道的速率一致, 使结点发送数据时, 无须根据源宿结点之间的跳数来调整 OT, 只需要保证 OT 足够目的结点正确接收 DB 即可, 简化了源结点的发送操作。

为解决接收冲突, 发送结点必须设法利用控制信道将 DB 的到达时刻和持续时间通知目的结点。每个结点都需要接收并处理这些时隙中的控制信息, 以判断是否有数据发给自己。文中提出的接收冲突解决方案可以分为基于重传机制、基于应答机制和基于令牌 (Token) 机制三种。基于重传的机制本质上并没有解决冲突, 仅在冲突发生后让竞争失败的一方重新发送; 基于应答机制的方案虽然可以完全避免冲突, 但是引入了更大的时延, 降低了环网资源的利用率; 基于令牌机制中的令牌表示对应结点是否允许接收数据, 可以有效的减少冲突 (采用每结点单令牌方案时可以完全避免冲突), 时延性能也优于基于应答的机制, 但是该方案的资源利用率比基于应答机制的方案更低。

基于 ODD 的环网方案优点是采用了一种较为简单的方法解决了控制信道和数据信道的速率不匹配问题, 使控制信息可以和数据之间保持“同步”; 采用半固定的方式将冲突限制于接收结点; 定长的控制帧和时隙有利于硬件实现。但是这种方案有如下不足: 由于每结点分配一个固有发送波长, 对信道资源的复用不充分, 资源利用率低, 扩展性差; 结点必须配备 FDL, 不仅增加成本, 而且会使信

号的质量有所下降。

### 1.4.3 LightRing 协议

文献[11]提出了一种称为 LightRing 的 OBS 环网协议。在 LightRing 环网方案中, 环结点被称为“光突发分插复用器”(OB-ADM, Optical Burst-Add-Drop Multiplexer)。环上仍然分 1 个控制波长多个数据波长。每个数据波长都采用一个固定的接收机进行接收, 一个固定的发射机进行发送, 这种配置可以允许结点在同一时刻发送多个突发。文献中提到可以采用可调谐收发机来减少所需的收发机的数目。在数据信道上, 结点使用光开关控制每一个波长的通断, 以此决定到达结点的数据业务是否在光域直通。

N 个 OB-ADM 通过光纤依次连接, 构成单向环网。其接入控制采用令牌(Token)方式实现。每一个波长对应一个令牌, 用以控制结点对该波长的使用。令牌中使用一个矢量记录环上相应波长在每一段链路上的忙闲状态; 令牌还记载有每一个目的结点的可用接收缓存容量以避免缓存溢出; 此外有一个流量计数器(Traffic Counter)来实时统计整个环上等待发送的高优先级业务量, 供结点在进行发送调度时决定是否可以发送低优先级业务。这种格式导致令牌大小跟环上结点总数相关(文中给出一个有 16 个结点的环中令牌的大小约 160bit)。令牌在环的周长上保持严格的顺序并且尽量均匀分布, 以保证各个 OB-ADM 能够及时获取令牌。令牌只在结点停留很短的时间, 刚好够结点完成所需信息的接收和处理。

当一个 OB-ADM 结点收到一个令牌时, 它首先检查发送队列中是否已经有完整的突发。若有, 它会查看本地到目的结点之间的各个链路是否空闲, 以及目的结点是否有足够的接收缓存。若本地到目的结点之间存在可用通路, 而且目的结点可以接收, 则将突发按“Tell-and-Go”的方式发出。Tell 操作需要修改令牌中的信息域, 通知下游结点这次发送已经执行, 让中间结点适时设置光开关, 保证目的结点正确接收。Tell 操作之后, 立即发出令牌。

通过上述方式, 在令牌的控制下, 数据突发在光域通过中间结点, 于目的结点终止。由于令牌中给出的是链路可用信息, 因此该方案允许波长的空间重用, 即同一时间在同一波长可以在环的不同链路上传送多个 DB。但是链路可用信息只能在令牌第一次返回发送结点时才会更新, 造成这种方式下的链路占用时间只能是令牌往返环路一周时间(称为令牌环回时延 Round-Trip Time)的整数倍。如果一个突发的发送时间比令牌环回时延的整数倍小, 带宽会被浪费。此外, 这个问题还受到令牌长度的影响: 随着结点数的增加令牌加长, 在结点上的处理时间增加, 会导致令牌环回时延增大, 带宽浪费加剧。

#### 1.4.4 MTIT 的令牌控制方案<sup>[12]</sup>

MTIT (Multitoken Interarrival Time) 多令牌到达间隔时间方案主要解决令牌在环上的位置分布。若环网单向光纤中有  $W$  个波长作为数据信道, 则同时有  $W$  个令牌在控制信道上循环。由于业务的抖动性, 这  $W$  个令牌可能成组到达, 即  $W$  个令牌相继到达同一结点, 允许结点在不同波长上同时发送数据, 致使该结点的下游结点无法获得空闲令牌而发送数据。这样会带来接入时延增大、接入的不公平性。而采用 MTIT 机制, 控制令牌在结点逗留时间为  $TTIT - TIAT$ 。TTIT (Target Token Interarrival Time) 为系统初始配置时由所有结点协商出的一个能满足最严 QoS 要求的目标令牌到达间隔时间; TIAT (Token Interarrival Time) 指相邻令牌到达间隔时间。当相邻令牌太近时, 即 TIAT 值较小,  $TTIT - TIAT$  控制后到令牌在结点逗留时间长些, 使两令牌间隔拉大; 当相邻令牌到达间隔较大时,  $TTIT - TIAT$  为负或接近零值, 则结点收到后到令牌立即发出, 缩短相邻令牌间隔。MTIT 机制对相关位置令牌进行自适应调节, 维持它们的分布, 最终使令牌均匀运行于环上。

文献中将 MTIT 机制应用在一个有  $W$  个数据信道和 1 个控制信道, 每个信道都有一对固定收发器, 所有中间结点都接收一份分组拷贝的环形广播系统中, 分组只有回到源结点才被销毁。虽然 MTIT 机制使令牌均匀分布, 保证了接入的公平性, 但由于广播系统本身的缺陷, 使得环网的带宽资源利用低, 网络的扩展性差。

#### 1.4.5 光突发传送 (OBT) 网

美国斯坦福大学的学者提出了一种结合令牌和确认机制的光突发传送 (OBT, Optical Burst Transport) 环网方案<sup>[13]</sup>。该方案中, 环上仍划分为一个控制波长和多个数据波长。采用可调谐激光发射器发送数据, 声光调谐滤波器 (AOTF, Acoustic-Optical Tunable Filter) 控制接收机从任意波长上接收数据。在控制波长上采用的控制信号包括令牌 (Token), 控制头 (CH, Control Header), 返回确认信号 (ACK, Acknowledgement) 和否定确认信号 (NACK, Negative Acknowledgment)。每个令牌代表一个数据波长的接入权限。令牌到达结点, 结点可以持有令牌并在有数据发送时利用相应波长信道发送数据分组, 当发送时间结束, 结点释放令牌。数据分组在结点中按照目的地址以虚拟输出队列 (VOQ, Virtual Output Queue) 方式管理, 通过汇聚减少开销。发送数据之前, 源结点先发送 CH 来提前通知目的结点将接收 AOTF 调谐到相应波长, CH 与数据之间间



隔一个与目的地址相关的偏置时间，确保 AOTF 的设置数据在数据到达之前完成。接收结点收到 CH 后，如果能在相应偏置时间内成功设置 AOTF，就返回 ACK 给源结点，否则返回 NACK。源结点并不需要收到 ACK 后才能发送数据，只是根据收到的返回信息来决定是从 VOQ 中删除已发送的数据还是保留这些数据等待重传。这种确认方式对于成功发送的数据并不引入额外的时延。发送时由流量调度器（Traffic Scheduler）来检测每个 VOQ 的发送窗口以控制发送时间。

结点通过 VOQ 锁定机制锁定当前正在被服务的发送队列，避免发送过程中有新令牌到达产生的发送冲突。

重传机制是 OBT 用来解决接收冲突问题的关键。由于存在多个令牌，到达目的结点的数据可能会由于 AOTF 不可用而无法被接收，此时目的结点向源结点返回 NACK。收到 NACK 的数据将被重新发送。当结点具有的共享资源（AOTF）较多时冲突会减少，重传率相应降低。

采用令牌，OBT 可以较好的解决发送冲突；确认和重传机制则可以减小接收冲突引起的丢失率。环上的服务区分主要依靠 VOQ 调度算法在结点内部实现。OBT 的主要不足是不能支持波长信道的空间重用，控制较为复杂，而且存在由于丢失和重传造成的带宽浪费。

## 1.5 本章小结

本章首先介绍了基于 WDM 的新型光交换技术，其中详细介绍了光突发交换（OBS）的组装、信令协议、资源调度和冲突解决等关键技术；其次总结了应用 OBS 技术的 WDM 城域网中存在的控制问题，同时介绍了已有的几种环网控制方案，分析了它们的优缺点。

针对现有问题和已有方案的不足，本文下一章将提出一种新的 OBS 城域光纤环网的接入控制方案。

## 第二章 新型虚同步光突发环网

针对环网现有问题，并参考上一章中的几种方案，我们提出一种新的 OBS 环网控制方案。该方案将数据信道划分，通过一种虚拟同步机制实现划分后的数据信道接入，支持空间重用。利用一种弹性偏置时间机制完成控制信道对数据信道的标识和时间对应关系。该方案在基本实现按需分配带宽的同时完全避免了数据波长信道上的发送和接收冲突，将数据的丢失限制在结点缓存部分。此外，这种机制可方便的支持多播业务，并实现全网的优先级区分服务<sup>[13]</sup>。我们称其为虚同步光突发环网（VS-OBR，Virtual Synchronized-Optical Burst Ring）。

### 2.1 VS-OBR 的结构

虚同步光突发环网 VS-OBR 由虚同步光突发分插复用器（VS-OBADM，Virtual Synchronized-Optical Burst Add-Drop Multiplexer）结点和光纤链路构成。每个光纤链路具有  $W+1$  个波长，其中  $W$  个波长作为数据信道，1 个波长作为控制信道（必

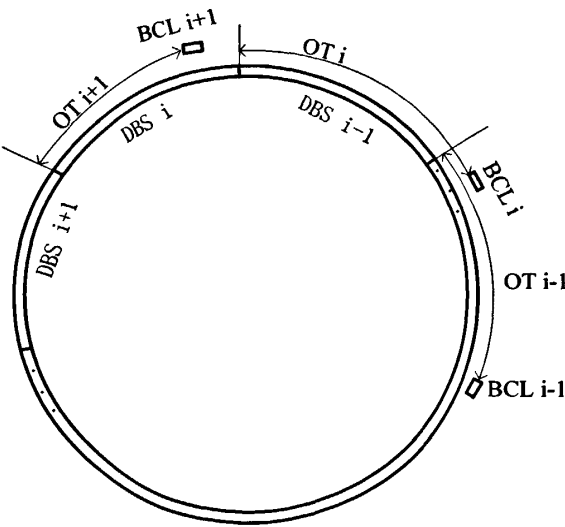


图 2.1 VS-OBR 结构

要时也可采用多个波长作控制信道）。

控制波长上传送突发控制分组，在 VS-OBR 中突发控制分组采用定长、简短的形式，称为突发控制标签(BCL，Burst Control Label)。该控制标签除了具有指示数据突发到达时刻的功能外，还兼具令牌功能，供沿途各个结点获取相应的环网带宽使用权限。该使用权限的单位是等长的时隙。为此，每个数据波长从时间

上划分为若干定长的数据突发时隙 (DBS, Data Burst Slot), DBS 首尾相连, 在整个环网数据波长上循环传播。除了接收 (下路) 或者发送 (上路) 数据外, 结点通过光开关的控制直接在光域转发 DBS。数据波长上传送的 DBS 与控制信道上的 BCL 一一对应, 而且其对应关系在系统工作期间一直维持, 这样结点通过获取 BCL 即可“使用”对应的 DBS, 无须在发送时创建自己的 BCL 和 DBS。VS-OBADM 的结构如图 2.1 所示, 其中 BCL 指明相应 DBS 的偏置时间 (OT) 和所在波长编号等。OT 的维持与更新非常关键, 因为 BCL 与 DBS 之间的时间差必须能够保证 BCL 先于 DBS 到达中间结点和目的结点, 并且留有充裕时间供结点进行相关设置。

每个结点收到 BCL 之后经过 OT 时间, 相应 DBS 即可到达同一结点。各 VS-OBADM 结点接收到 BCL 的时候, 根据 BCL 所承载的信息了解 DBS 的占用情况, 决定在 DBS 期间承载数据突发 (DB, Data Burst) 并发送、接收 DBS 所承载的 DB、或在光域转发 DBS 所承载的 DB。各结点通过修改 BCL 中的相关域通知下游结点对应 DBS 是否已被占用。下游其他结点收到占用标记为 1 的 BCL 时, 不会再使用该 DBS 来发送数据 (过路的低优先级被本地高优先级抢占除外)。因此, VS-OBADM 可以避免一般 OBS 环网中光突发之间的冲突, 提高网络性能和资源利用率。

该方案虽然划分了时隙, 但是 DBS 时隙的定界和使用都是在控制信道中 BCL 的控制下实现的, 而且 DBS 的位置通过 BCL 中的 OT 值确定, 在一定范围内可变, 有别于传统时隙交换中直接维持严格同步的工作方式, 因而称之为虚拟同步光突发环网。

## 2.2 VS-OBADM 结点的结构

VS-OBADM 的功能结构如图 2.2 所示。每个结点上配备有  $W$  个固定波长接收机和  $T$  个可调谐发射机 ( $1 \leq T \leq W$ ), 分别用来接收和发送数据; 控制波长使用一对专用收发机进行 BCL 的接收和发送。数据波长使用可调谐发射机可以使结点使用任意一个波长上的 DBS, 同时节省器件费用。结点配备的可调谐发射机个数决定该结点可以同时占用 DBS 进行数据发送的最大能力。接收单元为每个波长配备一个固定接收机, 避免了目的结点的接收冲突。

光路信号到达结点后, 先经过分路器变为两路。一路经过波长解复用分离出控制波长和数据波长, 控制波长上到达的 BCL 由控制波长接收机接收, 然后送至控制处理模块, 处理完毕则由控制信道固定波长发射机发送, 和上路业务的数据突发复用在一起进行传输。根据 BCL 的处理结果, 若数据波长上有到达本地的下路业务, 则 DBS 中的 DB 会被突发接收机接收, 经缓存、解封装后将所获得的数

据分组传给用户。分路后的另一路光信号解复用后经过光开关矩阵再进入复用器。光开关矩阵由半导体光放大器 (SOA, Semiconductor Optical Amplifier) 光开关组成, 受交换控制模块控制, 用于转发各数据波长上的 DBS。来自用户的本地上路业务进入上路业务发送处理模块之后, 根据目的地址和优先级等分别进入不同的输出队列进行突发组装, 组装可以采用基于长度门限和时间门限的组装算法, 组装后与 BCL 复用。通过光开关矩阵和上路来的两路复用后的信号经光合路器合并后继续传输到下游结点。图中调度器是 VS-OBADM 结点的核心, 它同时控制上路模块、下路模块以及 BCL 处理/交换控制模块, 分别完成数据突发的组装、发送、接收以及 BCL 的处理等操作。

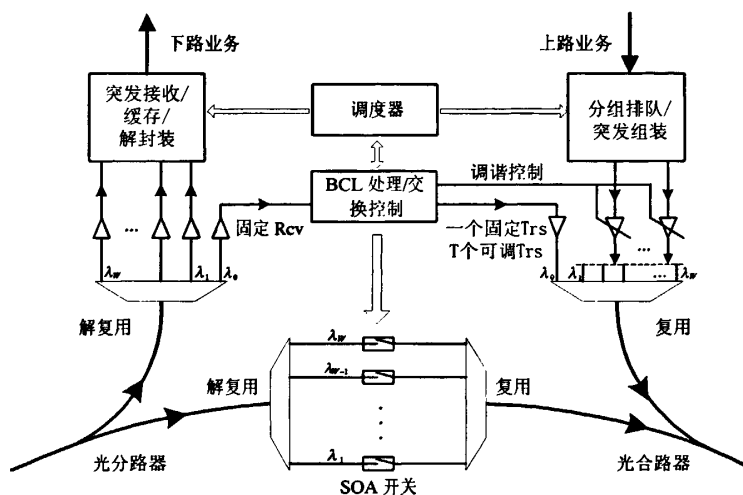


图 2.2 VS-OBADM 功能结构

环网中, 每个结点都可充当 3 种角色: 源结点、中间结点和目的结点。做源结点时, 结点将接入网送来的 IP 分组放入相应的发送队列, 缓存、等待发送机会并组装, 一旦该结点获得空闲 BCL 就可以在适当时刻将数据突发发送到相应 DBS 中。做中间结点时, 需要完成 BCL 的控制处理并在相应 DBS 到达之前完成光交换矩阵的配置工作, 在该 DBS 到达时刻将其转移到下一结点。做目的结点时, 结点接收来自源结点的数据突发, 并将突发拆分为 IP 分组发送给用户。

### 2.3 BCL 和 DBS 的格式

BCL 在控制波长上采用异步方式传送, 其传输速率与数据波长传输速率可以不同。在 VS-OBADM 结点中 BCL 的处理速度对环网的整体性能有较大的影响, 应该选用尽量简短的格式。图 2.3 给出了一种 BCL 格式, 其中各个字段的含义和长度说明如下:

- ◆ **Preamble:** 前导序列，供接收时快速获得比特同步并准确定界，4 字节；
- ◆ **DBS ID:** DBS 标识，在同一波长中的 DBS 具有不同的标识，可为连续整数，长度为 15 比特，与 I 比特共占用两字节；
- ◆ **I:** DBS 占用指示比特，置为 1 表示占用，置为 0 表示空闲，1 比特；
- ◆ **Lamda:** 所在波长编号，用于后续结点设置光开关和接收机，1 字节；
- ◆ **H-CRC:** BCL 头校验，用于头部检错和纠错，1 字节；
- ◆ **DBL:** Data Burst Length，用以指示突发数据占用长度，4 字节；

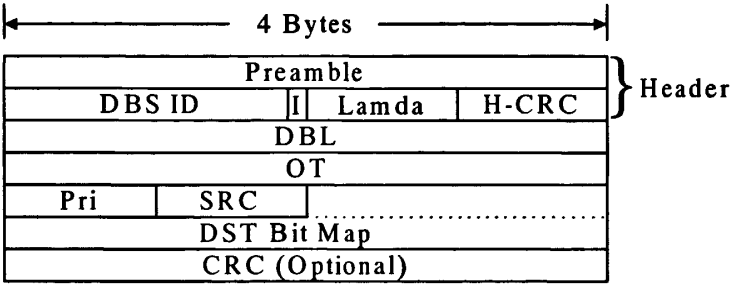


图 2.3 BCL 格式

- ◆ **OT:** Offset Time，对应 DBS 的时间偏移量，也称偏置时间，4 字节；
- ◆ **Pri:** Data Priority，数据优先级标识，用于优先级控制，1 字节；
- ◆ **SRC:** Source Address，源结点地址，1 字节；
- ◆ **DST Bit Map:** 目的结点地址（Destination Address）比特码表，其中的比特位对应各个结点，置为 1 表示相应结点为目的结点。采用比特码表可方便实现环网多播和广播协议。长度为 6 字节（可根据环网规模调整）；
- ◆ **CRC:** BCL 内容校验，用于 BCL 本身的差错控制，4 字节，可选。

算上最后的 CRC 校验，采用上述格式的 BCL 总长度为 28 字节。由于 BCL 长度固定，因此无须定义长度域。



图 2.4 DBS 承载的 DB 格式

在 VS-OBR 中，DBS 并不完全等同于一般 OBS 中的 DB，这里 DBS 仅仅表示数据波长信道中一段可以利用的时间片，类似于一个容器，真正承载的数据是这段时隙中传送的 DB。因此 DBS 格式实际是指其间传送的 DB 的格式。DBS 中 DB 的格式只要能够使接收结点快速获得比特同步、方便实现突发中 IP 分组的定界即可。为了支持数据波长的高速传输，参考光复合突发交换（OCBS，Optical Composite Burst Switching）网络中突发的格式<sup>[14]</sup>，图 2.4 给出了一种采用 SDL（Simple Data Link）帧的 DB 格式。其中“SYNC”字段用于突发接收时快速获得比特同步及初步确定帧起始；净荷长度指示后续用户分组的长度。利用净荷长度和 CRC 校验，当“SYNC”字段受到破坏，或者由于信号质量原因只接收到部分 DBS 时，也能恢复出剩余的完整数据分组。当然也可以采用其它方式进行 IP

分组封装。为了消除 OT 的定时不确定性, DBS 之间需要设置一定长度的保护区间。

## 2.4 弹性偏置时间控制机制

随着 DBS 及其对应的 BCL 在环上运行, 每经过一个结点, OT 都会减小一个 BCL 的处理时间, 如果不采取控制措施, OT 会持续减小直至造成下游结点没有足够时间提前判断并设置光开关。为了保持 DBS 与 BCL 之间的对应关系 (如图 2.1 所示), VS-OBR 采用弹性偏置时间控制 (ROTC, Resilient Offset Time Control) 机制对 BCL 和 DBS 之间的 OT 值进行调整, 保持两者之间的对应关系, 并使环结点始终能够根据 BCL 信息准确定位 DBS。

系统在初始化时确定 DBS 和 BCL 的数目, 系统运行期间保持不变; 每一个 DBS 都与控制信道中的一个 BCL 相对应, 初始化时设置两者之间的偏置时间 OT 初值。随着 BCL 经过的结点数增加, OT 会不断减小。按照 ROTC 机制, 在 OT 减小到某个特定门限时对 OT 进行调整, 将旧值自动增加 DBS 绕环传播一周的时间 RPT (Round Propagation Time), 作为更新后的 OT 值。OT 门限的设定与各个结点对 BCL 的处理速度相关, 设每个结点处理 BCL 引起的延迟时间为  $T_{in}$ , 目的结点正确接收、处理 BCL 和进行相应配置所需的时间为  $T_{id}$ , 环上结点总数为  $N$ , 则偏置时间门限  $T_{thr}$  可设定为:

$$T_{thr} = (N - 2)T_{in} + T_{id} \quad \text{式(2-1)}$$

偏置时间门限并不一定是 OT 的最小可能值, 结点仅在需要将收到 BCL 后实时计算出的 OT 值与门限对比以确定是否需要进行 OT 调整。

ROTC 机制对偏置时间调整的具体操作步骤描述如下:

- (1) 在系统初始化时设定 OT 初值和门限;
- (2) BCL 每经过一个结点, 需要从 OT 中减去它在该结点的处理时间;
- (3) 如果 BCL 状态为闲, 则经过每个结点时都需要检查 OT 值是否小于门限, 若小于则 OT 自动加上 RPT, 将 BCL 的 OT 域更新;
- (4) 若 BCL 状态为忙, 则由目的结点检查 OT 值是否小于门限, 若小于则 OT 自动加上 RPT, 作为新的 OT 值。

RPT 是光信号绕环传播一周的时间, 因此, OT 增加一个 RPT 后, BCL 与 DBS 之间在环上的位置对应关系并没有发生变化。这是弹性偏移时间控制 (ROTC) 机制的关键。每个波长上的 ROTC 操作是相互独立的, 可以在调整时使用实测的 RPT, 也可以使用外部测量或计算出的 RPT。采用弹性偏移时间控制 (ROTC) 机制后, BCL 与 DBS 一一对应的关系得以保持。从发送结点获取 BCL 并决定发

送数据，到数据抵达目的结点的时延为 OT 与传播时延之和。采用 ROTC，该时延最大约为两倍的 RPT。当环网的周长为数百公里时，该值在毫秒量级。在缓存等待时间不长的情况下，可满足绝大多数业务的时延要求。

ROTC 偏置时间调整虽然保持了 DBS 和 BCL 的对应关系，但是在 BCL 进行 OT 值调整后，第一个获得该 BCL 的结点必须等到 DBS 转过一周后才能使用该 DBS 传输数据，相当于 DBS 被空置了一周，造成部分带宽的浪费。下面我们来分析 ROTC 偏置时间调整引入的容量损失估计。

仍规定  $T_{in}$  为每个结点对 BCL 的延迟时间， $T_{id}$  为目的结点正确接收、处理 BCL 和进行相应配置所需的时间，环上结点总数为  $N$ 。由于 ROTC 并不随时对偏置时间 OT 进行检测，因此对于任意一个 DBS，其执行 OT 调整操作前一时刻的偏置时间值是一个小于偏置时间门限  $T_{thr}$  的随机变量，用  $T_{min}$  表示。偏置时间调整就是在原 OT 值上增加  $T_{RPT}$ ，即 DBS 环回传播时间，则调整后的新偏置时间  $T_{new}$  为：

$$T_{new} = T_{RPT} + T_{min} \quad \text{式(2-2)}$$

新偏置时间决定了该 DBS 多久之后会再次被调整一次，设下次进行调整时的 OT 最小值为  $T'_{min}$ ，显然  $T'_{min}$  与  $T_{min}$  是相同分布的随机变量。那么两次调整之间 OT 的差值的期望为：

$$T_e = E[T_{RPT} + T_{min}] - E[T'_{min}] = T_{RPT} \quad \text{式(2-3)}$$

$T_e$  实际上是该 BCL 在两次调整之间，在每一个结点上消耗掉的停留时间的总和。由于 VS-OBRR 的特性，目的结点和中间结点对 BCL 处理所要求的时间非常接近，若假设两者都等于  $T_n$ ，那么  $T_e$  可供 BCL 在环上循环运行的周数  $C$  为：

$$C = T_e / NT_n \quad \text{式(2-4)}$$

该 BCL 对应的 DBS 可以在这  $C$  周中被正常占用，之后 BCL 中 OT 立即会进行一次调整，DBS 空转一周，无法被占用。这部分无法被占用的容量同环带宽容量的比率  $w_b$  应为：

$$w_b = \frac{1}{C+1} = \frac{NT_n}{T_e + NT_n} = \frac{NT_n}{T_{RPT} + NT_n} \quad \text{式(2-5)}$$

式(2-5)说明最大资源利用率由结点处理速度和环回时延（RPT）决定。由于光纤波长链路的传输速率很高，前面给出的 BCL 长度最长为 28 字节，在 10Gb/s 的链路速率下 BCL 的时间长度只有 22.4ns，链路速率为 1Gb/s 时为 224ns。利用高速硬件处理电路，BCL 在结点上的停留时间可以控制在百纳秒到微秒量级。假设  $T_n$  为 1 微秒，结点数  $N=16$ ，环网周长为 160km，链路速率 1Gb/s，按照上式，此时损失的容量比率为 0.0196，几乎可以忽略不计。

ROTC 的另一个任务是确保 DBS 之间的虚拟同步关系不会随着网络的运行而逐渐混乱。不同于严格的同步传输网络, VS-OBR 中 DBS 之间的同步关系是由 BCL 中的 OT 操作来维持的。每一个结点收到 BCL 之后, 对于 DBS 的定位根据 OT 值使用本地时钟计时器来确定, 而各个结点时钟独立运行, 计时器对同样的 OT 值产生的计时间隔会有偏差, 造成定位不准确。因此, 经过一段时间之后, 需要对 DBS 的“位置”进行调整, 使得 DBS 之间的位置偏差不超过 DBS 之间保留的保护区间, 确保发出的数据不会因为定位偏差而损失。由于 VS-OBR 采用控制信道来产生和控制 DBS 所在时隙, 因此位置调整实际上只需要改变 BCL 中偏置时间的数值即可实现。下面是时钟偏差对 DBS 定位精度影响的分析。

对于任意一个 DBS, 当其不承载数据时, 其定位操作 (主要是沿途各结点光开关的切换) 几乎没有任何实质影响; 每次发送数据时, 发送结点总是根据本地时钟对 OT 计时, 定位 DBS 后发出自己的数据突发, 沿途结点和目的结点对 DBS 的定位误差仅和这些结点与发送结点之间的时钟精度误差有关。在发送到接收的有限时间里, 这种计时误差很小, 而且不会影响到下一次发送和接收。

假设环上有  $N$  个结点, 环上各个结点时钟都具有相同的标称值, 两两结点之间的时钟误差率的最大值为  $\delta$ 。 $\delta$  的含义是, 当这  $N$  个时钟中的任意两个独立测量一段标准时间  $T$  时, 它们产生的最大计时误差不会超过  $\pm T\delta$ 。同时, 每个 BCL 在结点进行处理所需的最大停留时间记为  $T_n$ 。为了便于分析, 这里首先假设对于 BCL 的接收时刻的计时是准确的, 并且无论中间结点还是源宿结点, BCL 在其中将停留相同的时间  $T_n$ 。将某一发送结点记为  $n_0$ , 然后将其下游结点依次标记为  $n_1, n_2, \dots, n_{N-1}$ 。对于一个 BCL 及其对应的 DBS, 由于  $n_0$  采用本地计算的时间发送数据, 从  $n_0$  发出的 BCL 具有“准确”的偏置时间为  $T_o$ , 即 DBS 中的数据位置是准确的。

当 BCL 到达  $n_1$  结点时, 该结点将根据其中记载的  $T_o$  值进行计时, 并在根据本地时钟计量出的时间  $T_o'$  时刻切换光开关, 从而产生偏差。由于两个时钟相互独立,  $T_o'$  与  $T_o$  之间的误差最大为  $\pm T_o\delta$ 。 $n_1$  处理完 BCL 并发出之后, 后续结点  $n_2$  从中读出的偏置时间应该为  $T_o - T_n$ 。这样, 结点  $n_2$  经过与结点  $n_1$  相同的操作, 对 DBS 的定位产生的偏差最值为  $\pm(T_o - T_n)\delta$ 。注意, 虽然  $n_1$  可能已经改变了 DBS 的定位, 但是在假设 BCL 在各个结点停留相同时间的情况下, BCL 中的偏置时间值与上一跳结点同  $n_0$  结点之间的误差无关。按以上方法类推, 可以将所有结点上可能产生的偏差最大 (正或负) 可能值写出如下:



$$\begin{aligned}
n_1 & \pm T_o \delta \\
n_2 & \pm (T_o - T_n) \delta \\
n_3 & \pm (T_o - 2T_n) \delta \\
& \vdots \\
n_{N-2} & \pm [T_o - (N-3)T_n] \delta \\
n_{N-1} & \pm [T_o - (N-2)T_n] \delta
\end{aligned} \tag{2-6}$$

显然经过的结点数越多, 并且 BCL 中记录的初始偏置时间值越大时, 上述偏差越大。为获得最大可能偏差, 假设在每个结点, 时钟都偏向同一方向 (比如都产生正偏差)。当目的结点是距离  $n_0$  最远的结点  $n_{N-1}$  时, 在目的结点收取 DBS 中的数据, 产生的最大定位误差应该为:

$$e_{\max} = \left[ (N-1)T_o + \frac{(N-1)(N-2)}{2} T_n \right] \delta \tag{2-7}$$

对于一个周长为 200km 的环网 (RPT 为 1ms), 若时钟频率的精度为  $10^{-6}$ , 环上结点总数为 21, 并且取  $T_o$  的最大值为  $T_{RPT}$ , BCL 在结点的逗留时间为  $1 \mu s$ , 发送与接收之间产生的最大定位误差约为 20ns, 相对于 DBS 的长度非常微小。

上面考虑的是单个 DBS 从发送到接收的最大误差, 该误差在每次发送和接收时都重新计算, 这次接收的误差不会影响下一次接收, 因此不存在误差积累问题。但是, 环中每个波长上都有多个 DBS 在同时运行, 而且每个 DBS 都独立计算自己的位置。出于时钟误差的随机性, DBS 之间的相互位置关系可能由于上述误差的累积而导致严重错位 (尽管出现的概率很小, 但是对环网性能有较大影响)。为此, ROTC 利用环网每一结点都能了解全部 DBS 位置信息的特点, 在某一结点上根据 BCL 中的 OT 值来利用自己的时钟计算 DBS 之间的相互位置关系, 当发现某一 DBS 与其紧邻 DBS 之间的错位超过某一门限时, 主动校正其位置。为此, 负责 DBS 位置校正的结点需要记录每个 DBS 最近一次校正之后的“准确”位置, 当该 DBS 对应的 BCL 再次回到校正结点时, 根据最近一次校正后的 DBS 所处时刻  $T_{last}$ 、BCL 收取时刻  $T_a$  和 BCL 中记录的 OT 值  $T_o$ , 分别计算出该 DBS 当前应该处于的“正确”位置  $T_r$  和当前实际出现的位置  $T_c$ , 比较两者的关系。当两者之间的误差超过某个安全门限  $T_s$ , 例如保护区间的 1/2 时, 将  $T_o$  改为校正结点认为正确的值; 没有超过安全门限的 BCL, 不进行调整。

设本地结点掌握一个正确的环回传播时延  $T_{RPT}$ , 当经过最近一次调整后, 同一 BCL 第  $k$  次返回校正结点时, 有:

$$\begin{aligned} T_r &= T_{last} + kT_{RPT} \\ T_c &= T_a + T_o' \end{aligned} \quad \text{式(2-8)}$$

这时校正结点计算出的位置偏差为：

$$T_{err} = |T_r - T_c| = |T_{last} + kT_{RPT} - T_a - T_o'| \quad \text{式(2-9)}$$

上面两式中  $T_{last}$ ， $T_a$  都是“时刻”，即相对于时间原点的时间值，而  $T_o'$  和  $T_{RPT}$  则是“时段”，即两个时间点之间的时间段值。这里我们要估计经过一周之后，返回的 BCL 对应的 DBS 相对于“标准”位置的偏差大小。当沿途各个结点都和校正结点时钟完全严格同步时，返回的 BCL 中记录的时间应为  $T_o' = T_o - (N-1)T_n$ ，其中  $T_o$  是该 BCL 上次从校正结点发出时具有的偏置时间。而根据记录计算出的 DBS 到达时刻应该和根据 BCL 到达时刻加上正确的偏置时间计算出的完全相同，即

$$T_{last} + kT_{RPT} - T_a - T_o' = T_{last} + kT_{RPT} - T_a - [T_o - (N-1)T_n] = 0 \quad \text{式(2-10)}$$

但是在各个结点之间最大时钟偏差率为  $\delta$  的情况下，返回的偏置时间记录存在误差：

$$T_o' = T_o - (N-1)T_n \pm (N-1)T_n\delta \quad \text{式(2-11)}$$

根据(2-9)，(2-10)和(2-11)，有：

$$T_{err} = |T_o - (N-1)T_n - T_o'| = |(N-1)T_n\delta| \quad \text{式(2-12)}$$

可见，在校正结点实际上产生的绝对位置偏差非常小，考虑到偏差的随机性，存在大量正负偏差相抵的情况，运行中 DBS 位置校正所需的操作频度很低。

## 2.5 结点工作过程

环网中每个 VS-OBADM 结点接收来自接入网的 IP 业务。根据业务的目的结点、服务类别等信息，IP 分组进入相应的上路待发队列。结点在控制波长上收发 BCL，根据 BCL 判断其对应的 DBS 何时到达本结点、是否已经被占用、以及占用数据的优先级等，然后决定本地业务的收发操作以及过路业务的光开关配置。结合 ROTC，每个结点中进行的主要操作步骤可描述如下：

(1) 如果到达的 BCL 未被占用而且其 OT 大于门限，结点检查本地上路数据队列中是否有足够的待发数据，若没有，修正 OT 后直接发出 BCL；若有，则记载其发送时刻（即该 BCL 对应的 DBS 到达时刻），同时检查发送持续时间内是否有可用的调谐发送器，若没有可用的调谐发送器则放弃本次发送，发出 BCL；有可用的调谐发送器则修改 BCL 的相关域值（I、DST、SRC 等）之后发出 BCL。

根据本地时钟计数值,等待 OT 时间之后,发送数据到对应的 DBS 中。在等待 OT 时刻的过程中,如果突发长度未达到最大组装长度,仍可以继续组装。

(2) 如果到达的 BCL 未被占用并且其 OT 值小于门限,则根据 ROTC 控制机制增加 RPT,之后再检查本地队列是否有可发数据,其余同步步骤(1)。

(3) 如果到达的 BCL 已被占用,则判断其对应的数据突发的目的地址是否是本结点。若不是,只需修正 OT,转发 BCL 到下游紧邻结点,并根据记载的到达时刻提前配置好光开关,使得 DBS 从本地直接通过;若是,则根据 BCL 中的信息预计对应的 DBS 到达时刻并准备接收数据;检查 BCL 中 OT 值是否小于最小门限,若小于则加上 RPT;将 BCL 中的 DBS 占用标志域置为空闲。此时如果本地发送队列中有待发数据,则仍可利用收到的 DBS 按照前述步骤发送数据;否则直接将空闲 BCL 发往下游结点。

(4) 若结点负责 DBS 的位置调整,则该结点收到 BCL 后,根据所记录的最近一次调整后该 BCL 中的偏置时间值,计算出 DBS 当前应该所处的位置(实际就是记录的 OT 值加上 RPT),并将其与当前根据 BCL 接收时刻和其中的 OT 计算得到的实际位置相比较。若两者偏差超过某个特定门限,则记录该偏差,在下次收到同一 BCL 时根据偏差微调 OT 值。

## 2.6 多播的实现

多播(multicast)是网络中一个源结点向多个目的结点发送相同数据的通信过程,组播即是一种多播。当目的结点扩大到除源结点之外所有其它结点时,多播即成为广播(broadcast)。环网拓扑上实现多播比较简单,但是在基于令牌、基于返回确认或者突发长度随机变化、OT 与源宿距离相关的 OBS 环网中,实现多播需要付出额外的代价。在 VS-OBR 中,结点发送数据时无须考虑 OT 与目的地址的关系,利用 BCL 中的目的地址比特码表,只需要将多播目的结点的对应地址比特设为 1,就可以很容易的利用环网特性实现多播。结点接收多播 DBS 和单播 DBS 的过程并无本质的不同。在接收到 BCL 之后,如果发现本结点对应的地址比特为 1,即将对应的 DB 作为下路业务接收;如果此时地址比特码表中仍有非 0 值,即意味着后续结点中仍有本 DBS 的接收者,那么在本地下路接收的同时,结点控制转发的光开关为“通”状态,让 DBS 继续通过。从图 2.2 不难看出,VS-OBADM 的结构完全能支持这种本地接收和转发同时进行的操作。当最后一个目的结点完成数据接收后,目的地址表中全部比特均为 0,多播 DBS 会被自动终止。

## 2.7 优先级区分服务

VS-OBR 的一个特点是支持全网范围内的优先级区分。之所以强调全网范围，是因为在 1.4 节介绍的其它 OBS 环网方案中，优先级调度仅仅局限于本地结点的上路业务队列。VS-OBR 采用了固定长度的 DBS，而且通过 ROTC 实现了 OT 与源宿结点之间的距离无关，可以比较方便的实现本地高优先级业务对过路的低优先级业务的资源抢占。简而言之，如果高优先级队列在预定的时间门限之内一直没有得到发送机会，则可启动抢占机制。此时，如果收到对应 DBS 承载低优先级业务的 BCL，调度控制单元就可以抢占该 BCL（及其对应的 DBS）发送高优先级业务。

被抢占的业务可以直接被丢弃，也可以在本地进行暂存，跟本地低优先级队列一起调度，等待再次发送的机会。这样，在本地高优先级业务、低优先级业务以及本地暂存的被抢占业务之间可以设计多种不同的调度控制方法，可以使用长度门限、计时器超时处理等算法。方便的支持可抢占的优先级区分是 VS-OBR 的一个显著优点，在这个意义上，不仅可以在不同结点之间实现业务优先级区分，而且可以在业务不均衡时，通过抢占机制来避免环形拓扑所固有的不公平性。

## 2.8 本章小结

本章详细介绍了一种新的 OBS 环网控制方案——虚同步光突发环网（VS-OBR）。2.1 节至 2.5 节描述了 VS-OBR 的基本结构和工作机制，它使用虚同步光突发分插复用器（VS-OBADM）做环上结点，完成 BCL 和 DBS 的控制处理机制；弹性偏置时间控制（ROTC）解决了控制信道和数据信道的速率不匹配问题，并形成了 BCL 与 DBS 一一对应的关系。其中 2.3 节还给出了 BCL 和 DBS 的理论格式。在 2.4 节中专门对 ROTC 的引入而引起的容量损失和时钟偏差问题进行了严密的理论分析。此外，2.6 节和 2.7 节论述了 VS-OBR 可支持多播和区分服务的优良性能。

下一章我们将在本章的理论基础上设计 OPNET 平台下的仿真方案。

## 第三章 VS-OBRR 的仿真建模

仿真是目前国内外使用较多的一种网络规划和设计技术，它以计算机为主要工具，借助软件技术，在网络建设的前期对网络性能和状态进行了解和规划。通过仿真可以及时发现设计问题，提出改进和优化方案，从而极大地提高效率，节约投资，缩短开发时间，为通信网络的发展提供强有力的支持。随着网络新技术的不断出现和数据网络的日趋复杂，网络仿真技术已成为研究、规划、设计网络不可缺少的工具。本章基于目前业界广泛应用的仿真工具 OPNET Modeler，介绍为考察虚同步光突发环网（VS-OBRR）性能而设计的仿真模型。

### 3.1 OPNET 仿真平台简介

OPNET Modeler 是 OPNET 公司开发的一套集开发和应用为一体的通信系统模拟软件，它采用面向对象的建模方法和图形化的编辑方式，支持在网络各个层次的设备、链路和协议的精确建模，并提供了丰富的外界开发接口。本文使用的是版本号为 10.5 的 OPNET Modeler。

在仿真驱动机制上，OPNET Modeler 采用了离散事件驱动（Discrete Event Driven）的模拟机制。所谓“事件”是指网络状态的变化，也就是说，只有网络状态发生变化时，模拟机才工作，否则不执行任何模拟计算。因此，与时间驱动相比，离散事件驱动的模拟机计算效率得到很大提高<sup>[5]</sup>。

在建模机制上，OPNET Modeler 以包为基本单位模拟实际物理网络中数据的流动，可以生成、编辑任何标准的或自定义的包格式，并且利用调试功能，还可以在模拟过程中察看任何特定包的包头（Header）、净荷（Payload）等内容。从而使得仿真过程清晰明了，易于控制。

在建模方式上，OPNET Modeler 采用层次化建模结构，分网络（Network）、结点（Node）和进程（Process）三个层次建模。其中最底层为进程模型，它以有限状态机（FSM, Finite State Machine）来描述协议，用于模拟单个对象的状态。状态（states）表示模块的行为；状态转移（transitions）表示状态改变以响应一个事件。其次为结点模型，由相应的协议模型构成，反映设备特性，通过它可将进程模型中的对象互联成设备。最上层为网络模型，用于将设备互联组成网络。

OPNET Modeler 具有丰富的统计量收集和分析功能，可以直接收集常用的网络层性能统计参数，并有多种统计参数的采集和处理方法，还可以通过底层网络模型编程，收集特殊的网络参数。OPNET 还具有丰富的图表显示和编辑功能、模

拟错误提示和告警功能，能够方便地编制和输出仿真报告。

### 3.2 仿真模型整体设计

图 3.1 是虚同步光突发环网在 OPNET 平台下的网络仿真模拟示例，它主要由初始化结点（Initial\_Node）、源结点（SRC）和环上结点（VS\_OBADM 结点）组成。源结点与 VS\_OBADM 结点之间通过接入链路相连，环上 VS\_OBADM 结点之间通过控制波长链路、数据波长链路相连。环网规模（如结点数、链路长度等）是可变的。

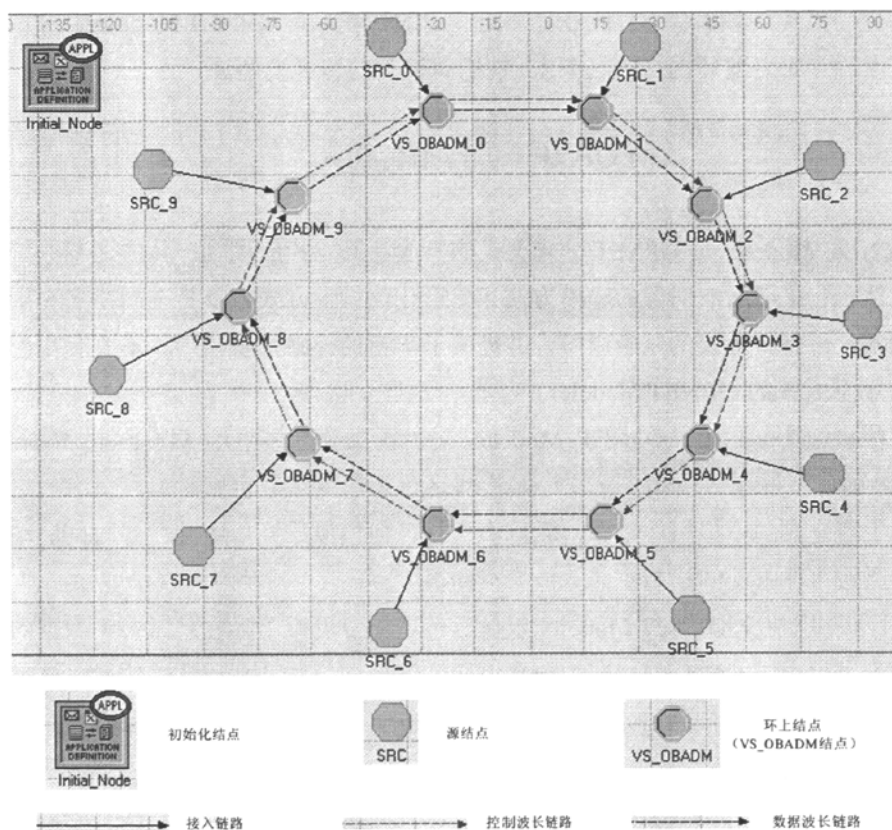


图 3.1 VS-OBDR 网络仿真模型

初始化结点主要负责网络拓扑的自动发现。源结点的功能是发出服从一定分布的 IP 数据分组，其网络原型是有源和目的功能的接入网，源结点只是仿真中对接入网的简化。环上的 VS\_OBADM 结点是进行核心控制和操作的单元，VS-OBDR 的上路、下路、交换、配置等功能均由这种 VS\_OBADM 结点完成。

网络中有三种链路：接入链路（access\_link）、控制波长链路（control\_link）和数据波长链路（data\_link）。接入链路上传输的是源结点发出的 IP 分组，这些

IP 分组经由接入链路进入到环网并由环上结点转发到其目的地, 仿真中不考虑分组接入时延, 所以设定接入链路上时延为零。环上结点之间的控制波长链路是用来传输控制分组的, 即数据突发的控制头, 在 VS-OBDR 中是 BCL; 数据波长链路是用来传输数据突发的, 突发是由源结点发来的 IP 分组在环上结点组装后生成的, 这里数据波长链路上传播的是 DBS。控制波长链路和数据波长链路可采用相同的传输速率, 也可采用有所区分的高速和低速链路, 仿真中均设定为常用的传输速率 2.5Gb/s。

### 3.3 三种结点模型

#### 3.3.1 初始化结点模型

初始化结点不是 VS-OBDR 必须的一部分, 设计这个结点是为了方便获得以下功能:

- (1) 自动扫描网络拓扑, 获取结点类型、各类型结点数目、链路类型等;
- (2) 获取用户配置参数, 如数据波长链路上的波长数目等;
- (3) 初始化和全网相关的变量和统计量。

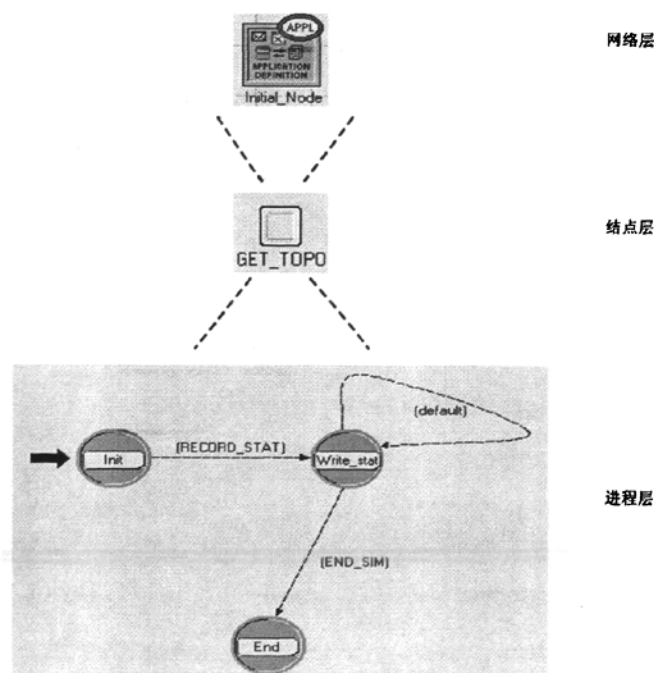


图 3.2 初始化结点模型

初始化结点功能较单一, 因此结构也相对简单。图 3.2 展示了它的三层模型。GET\_TOPO 模块是初始化结点的核心, 它内部的状态转移情况是: 系统初始化时,

仿真核心调度状态机 Init 的入口（状态机上部）代码，自动扫描全网拓扑结构，获得相关的属性；仿真即将结束时，满足状态转移条件 RECORD\_STAT，转移到 Write\_stat 状态机的入口处，记录统计量；仿真结束时间到达，转向 End 状态，结束仿真。

### 3.3.2 源结点模型

作为城域骨干网的 VS-OBR，其数据业务由各种接入网提供。但这里的研究集中在环上的控制处理机制，我们需要的仅是有一定特征的输出业务，对接入网本身并不关心。因此，为了简化仿真，设计了源结点来模拟各种接入网的输出业务。

#### 一 泊松源

泊松源产生的 IP 分组服从泊松到达，其数据发生有一定的突发性，常被用来模拟网络输入业务。我们就采用泊松源作为几种源之一来模拟接入网的数据业务。

##### ◆ 泊松源的数学描述<sup>[16]</sup>

设一个随机过程为  $\{A(t) | t \geq 0\}$ ， $A(t)$  的取值为非负整数，到达速率为  $\lambda$ 。如果该过程满足下列条件，则称为 Poisson 过程。

(1)  $A(t)$  是一计数过程，它表示在  $[0, t]$  区间内到达的用户总数。任给两个时刻  $s$  和  $t$ ，且  $s < t$ ，则  $A(t) - A(s)$  即为  $[s, t]$  之间到达的用户数；

(2) 在两个不同时间区间（区间不重叠）内到达的用户数是相互独立的；

(3) 任何长度为  $\tau$  的区间内，到达的用户数服从参数为  $\lambda\tau$  的 Poisson 分布，即

$$P(A(t + \tau) - A(t) = n) = e^{-\lambda\tau} \frac{(\lambda\tau)^n}{n!} \quad n=0,1,2,\dots \quad \text{式(3-1)}$$

其基本特征为：

(1) 到达时间间隔  $\tau_n = t_{n+1} - t_n$  相互独立，且服从指数分布，其概率密度函数为：

$$P(\tau_n) = \lambda e^{-\lambda\tau_n} \quad \text{式(3-2)}$$

分布函数为：

$$P(\tau_n < s) = 1 - e^{-\lambda s} \quad s \geq 0 \quad \text{式(3-3)}$$

(2) 对于一个任意的小区间  $\delta \geq 0$ ，将 Poisson 分布用泰勒 (Taylor) 级数展开，即利用

$$e^{-\lambda\delta} = 1 - \lambda\delta + \frac{(\lambda\delta)^2}{2} - \dots \quad \text{式(3-4)}$$

可得：



$$P\{A(t+\delta) - A(t) = 0\} = 1 - \lambda\delta + o(\delta) \quad \text{式(3-5)}$$

$$P\{A(t+\delta) - A(t) = 1\} = \lambda\delta + o(\delta) \quad \text{式(3-6)}$$

$$P\{A(t+\delta) - A(t) \geq 2\} = o(\delta) \quad \text{式(3-7)}$$

式中,  $o(\delta)$  表示  $\delta$  的高阶无穷小, 即  $\lim_{\delta \rightarrow 0} \frac{o(\delta)}{\delta} = 0$

(3) 多个相互独立的 Poisson 过程之和  $A = A_1 + A_2 + \dots + A_k$  仍是一个 Poisson 过程, 其到达率为  $\lambda = \lambda_1 + \lambda_2 + \dots + \lambda_k$ , 式中  $\lambda_k$  是 Poisson 过程  $A_k$  的到达率。

#### • 泊松源的仿真建模

在 OPNET Modeler 中泊松源可利用模型库 (Model LIB) 中的 simple\_source 进程模型, 建成后的模型如图 3.3。

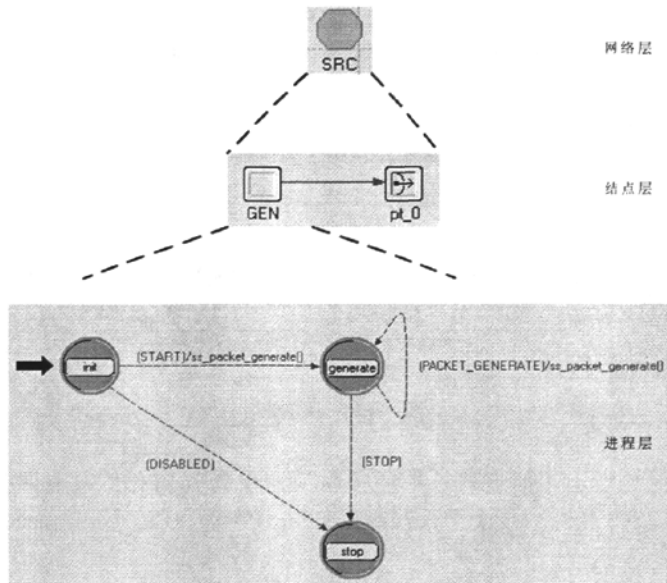


图 3.3 泊松源仿真模型

结点层模型由一个数据发生模块和一个点到点发射器组成。数据发生模块 GEN 在进程层提供属性接口参数, 可设置产生分组的格式 (Packet Format)、长度 (Packet Length)、到达间隔时间 (Packet Interarrival Time)、开始产生时间 (Start Time) 等。

将分组到达间隔时间选择为指数分布 (exponential), 则分组到达为泊松流。分组长度可在接口参数中选则服从常数分布 (constant)、均匀分布 (uniform)、指数分布 (exponential)、正态分布 (normal) 等; 也可修改函数 ss\_packet\_generate() 中的产生分组长度代码, 使分组长度取一些离散值, 这些离散值服从某种特定概率分布。把分组到达间隔时间设为常量, 分组长度也设为常量, 就是最简单的常

量源。

## 二 自相似源

传统的网络业务一般以泊松模型描述,然而近年来,通过对局域网 LAN 和广域网 WAN 等业务量的大量研究表明,有一些网络业务(如以太网、WWW、7号信令、TCP/FTP/TELNET 等<sup>[17]</sup>)在较大尺度范围内呈现出统计自相似性,突出表现为:业务没有明确的长度,在不同的时间尺度下表现出相同的突发特性,业务长程相关(long-range dependence),不能被平滑掉。对于这些业务,我们将设计具有自相似特性的源结点模型来模拟。

### ◆ 自相似源的数学描述和生成方式<sup>[17][18]</sup>

考察一个协方差平稳(又称为广义平稳)统计过程  $X = \{x_t : t = 0, 1, 2, 3, \dots\}$ ,也就是说,该过程的数学期望  $\mu = E[x_t]$  为常数,方差  $\sigma^2 = E[(x_t - \mu)^2]$  为有限值,自相关函数  $r(k) = E[(x_t - \mu)(x_{t+k} - \mu)] / \sigma^2, k = 1, 2, 3, \dots$ , 仅与  $k$  相关。其中  $x_t$  可理解为第  $t$  个单位时间里到达的网络业务实体(Packet)数目。假定  $X$  的自相关函数有如下形式:

$$r(k) \sim \alpha k^{-\beta}, 0 < \beta < 1, \text{当 } k \rightarrow \infty \quad \text{式(3-8)}$$

其中  $\alpha$  为一个正常数,对于每一个  $m = 1, 2, 3, \dots$ , 令  $X^{(m)} = \{X_t^{(m)} : t = 0, 1, 2, 3, \dots\}$  表示一个新的序列,其中  $x_k^{(m)} = 1/m(x_{km-m+1} + \dots + x_{km}), k = 1, 2, 3, \dots$  代表程度为  $m$  的聚集过程。如果  $X^{(m)}$  也定义了一个广义平稳随机过程,  $r^{(m)}(k)$  表示  $X^{(m)}$  的自相关函数,如果对于所有的  $m$ , 均有

$$r^{(m)}(k) = r(k) \sim k^{-\beta}, 0 < \beta < 1, \text{当 } k \rightarrow \infty \quad \text{式(3-9)}$$

我们称  $X$  为精确二阶自相似过程,并称  $H = 1 - \beta/2$  为其自相似参数,也就是说如果聚集过程  $X^{(m)}$  与过程  $X$  是不可区分的(最少是二阶统计特性不可区分),我们就称  $X$  是精确自相似过程。如果

$$r^{(m)}(k) \rightarrow r(k) \sim k^{-\beta}, \text{当 } m \rightarrow \infty, k = 0, 1, 2, 3, \dots \quad \text{式(3-10)}$$

则称  $X$  为渐进二阶自相似过程。

目前用于模拟自相似业务较为成熟的方法大致分为两类:一类是基于自相似过程物理意义的,典型代表为流叠加法<sup>[19]</sup>;另一类是基于统计模型的,主要方法有分形布朗运动(FBM, Fractional Brown Motion),分形高斯噪声(FGN, Fractional Gaussian Noise),逆小波变换法等。使用物理特性建立的模型在精度和灵活性上虽没有在统计特性下建立的模型好,但其物理意义明确,易于理解且建立简单,因此得到广泛应用。本文采用流叠加法生成自相似业务。

流叠加法将自相似过程看成是无数用户数据源叠加的结果,利用 ON/OFF 业

务源模拟用户数据源产生。原理如下：ON/OFF 模型只有 ON 和 OFF 两个状态，当业务处于 ON 状态时以恒定的速率输出数据，而处于 OFF 状态时不产生任何数据。假设网络上传输的多业务源模型为上述若干个单个 ON/OFF 业务源的叠加，设有  $m$  个同类的业务源，单个业务源 ON 状态持续时间为  $\tau^{(i)}$ ，OFF 状态持续时间为  $\theta^{(i)}$ 。随机变量  $\tau^{(i)}$ 、 $\theta^{(i)}$  独立同分布，其分布函数为

$$p(\tau > t) : t^{-\alpha} \quad t \rightarrow \infty, 1 < \alpha < 2 \quad \text{式(3-11)}$$

即  $\tau^{(i)}$ 、 $\theta^{(i)}$  均服从 Pareto 分布，具有有限均值  $\mu = E(X) < \infty$  和无限方差。注意到式 (3-11) 与传统 Poisson 过程的负指数分布  $p(\tau > t) : e^{-\lambda t} \quad t \rightarrow \infty$  有很大不同，或者说 (3-11) 中的分布函数比负指数分布有更重的拖尾 (heavy tail)。可以证明<sup>[19]</sup> 当叠加的业务源个数趋于无穷时，总的业务是渐进自相似的。

#### ◆ 自相似源的仿真建模

仿真中采用流叠加法生成自相似业务时，使用了 40 个 ON/OFF 源生成模块汇聚业务，其网络层和结点层模型如图 3.4。

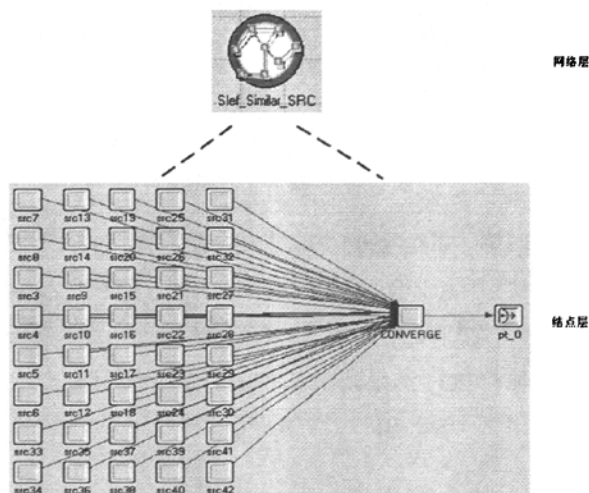


图 3.4 自相似源仿真模型

结点层左边的 40 个 ON/OFF 源完全相同，他们的功能是在 ON 周期产生 IP 分组；所有的 ON/OFF 源将生成的 IP 分组发给 CONVERGE 模块，由此模块汇聚 IP 业务，通过点到点发送器发往环上。

ON/OFF 源的内部结构如图 3.5。init 状态进行模型的初始化工作，如获取属性值，初始化变量、参数等；on 状态表示 ON 期；off 状态表示 OFF 期。初始化工作完成后，模型以等概率进入 ON 期或 OFF 期，ON 期和 OFF 期相互独立，且都服从 Pareto 分布，为方便仿真将 Pareto 分布的两个参数  $k$  和  $\alpha$  设置为属性接口并进行了提升，使用时可以根据具体情况进行设定以改变 ON/OFF 期的分布规律。当处于 ON 期时，模型将以恒定速率发送数据，发送速率通过式 (3-12) 计算得

到:

$$v = \rho \cdot R \cdot C \cdot (L_{on} + L_{off}) / (N_{source} \cdot L_{on}) \quad \text{式(3-12)}$$

其中:

$$L_{on} = \alpha_{on} \cdot k_{on} / (\alpha_{on} - 1) \quad \text{式(3-13)}$$

$$L_{off} = \alpha_{off} \cdot k_{off} / (\alpha_{off} - 1) \quad \text{式(3-14)}$$

$(\alpha_{on}, k_{on}), (\alpha_{off}, k_{off})$  分别是 ON 周期和 OFF 周期的  $\alpha$  和  $k$  参数,  $L_{on}, L_{off}$  为均值,  $v$  是数据发送速率,  $\rho$  是单波长上的业务强度,  $R$  是单个信道的数据传输速率,  $C$  是每个端口包含的信道数。

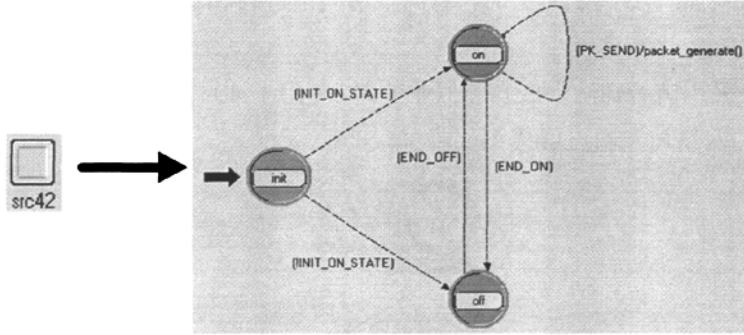


图 3.5 ON/OFF 数据源仿真模型

根据大量测量实际网络中 IP 分组长度的统计值, 我们设定产生的 IP 分组长度遵循以下统计规律<sup>[20][21]</sup> (近似值): 44 字节占 56%, 52 字节占 4.5%, 576 字节占 16.5%, 1 500 字节占 23%。

对自相似源进行单独测试, 结果表明其生成的业务流具有自相似性, 符合自相似突发源的特性, 可用于 VS-OBDR 网络仿真的研究。

### 3.3.3 VS-OBADM 结点模型

VS-OBADM 结点是整个仿真模型的核心部分, 它聚集了普通 OBS 网络中边缘结点和核心结点的所有功能, 完成环网中重要的 BCL 控制交换处理, DBS 上路、下路、过路操作及从源结点接收数据。图 3.6 展示的是 VS-OBADM 结点在网络层和结点层的模型。

#### 一 结点工作流程

点对点接收器 IP\_RCV 接收源结点发来的 IP 分组业务, 并把这些分组传输给 ASSEMBLE 模块, 分组在 ASSEMBLE 模块中插入以目的地和优先级标识的缓存队列。

从上游结点发来的 BCL 经 BCL\_RCV 接收器接收后交付 BCL\_PRO 模块进行处理。若该 BCL 空闲, 则 BCL\_PRO 模块询问 ASSEMBLE 模块是否有足够的数据要发送并查询偏置时间后有无空闲可用的发送机 (这里的发送机指可调谐发送

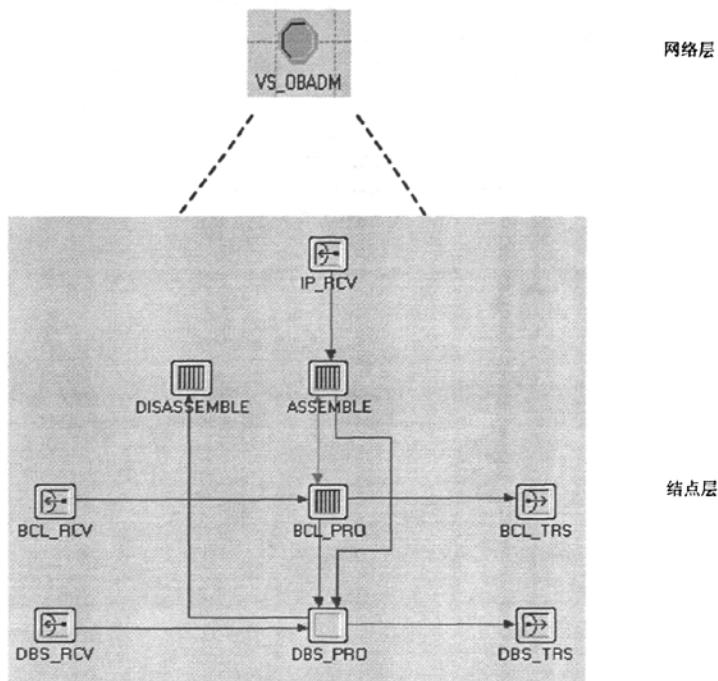


图 3.6 VS-OBADM 结点模型

机, 是结点中的逻辑部分, 不是图中的发送器), 如果两个条件有一个不满足, 则将空闲 BCL 经点对点发送器 BCL\_TRS 发送给下游结点; 如果条件都满足, 则改写该 BCL 中的忙闲标识, 源、目的地址等相关域, 发送给下游结点。若该 BCL 正忙, 则检查目的地址是否是本结点, 如果是则配置接收机以便偏置时间后接收与其对应的 DBS (下路操作), 并清空该 BCL, 再次查询数据, 经 BCL\_TRS 发走该 BCL; 如果不是, 则配置光开关, 在偏置时间后让与其对应的 DBS 穿过该结点, 并经 BCL\_TRS 发走该 BCL。

在空闲 BCL 到达时, ASSEMBLE 模块有足够的发送数据, 则在 ASSEMBLE 模块设定一个自中断, 该自中断于偏置时间后组装相应的数据分组为一个 DB, 由 ASSEMBLE 模块发给 DBS\_PRO 模块, 再由 DBS\_TRS 点对点发送器将该 DB 装载在 DBS 中发送给下游结点 (上路操作)。过路和下路的 DBS 均是由 DBS\_RCV 点对点接收器接收, 过路的再由 DBS\_TRS 发送给下游结点, 下路的则由 DBS\_PRO 发送给 DISASSEMBLE 模块, 在 DISASSEMBLE 模块拆分 DB 并销毁 IP 分组。在网络原型中, 下路的 DBS 本应发到用户终端, 而这里为了将研究集中在处理 BCL 上, 于是在环上结点就销毁了拆分后的 IP 分组。

每个空闲 BCL 到来或是目的地为本结点的 BCL 到来时, BCL\_PRO 都要先

检查偏置时间 OT 是否大于一个门限值, 若小则加上一倍的 RPT。

详细的结点工作步骤在 2.5 节中已介绍。为简化仿真, 减少运行时间, 可先将调谐发送机的数目设为等于波长数目。总结以上文字, 可得 VS-OBADM 结点工作流程图 3.7。

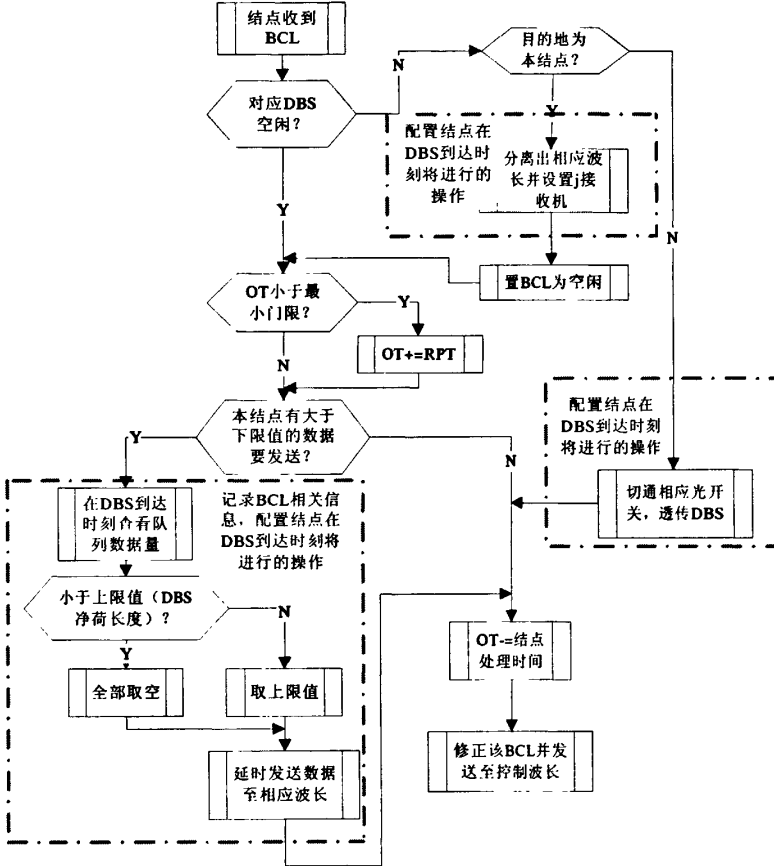


图 3.7 VS-OBADM 结点工作流程

## 二 ASSEMBLE 模块

ASSEMBLE 模块的功能是将 IP 分组插入缓存队列, 并在适合时刻进行组装。其进程层模型如图 3.8。下面介绍其中几个重要的状态机。

- ◆ **Insert\_subq 状态:** 将源发来的 IP 分组按目的地址插入不同的缓存队列, 若相应的缓存队列已满, 则丢弃该分组, 将它记录在丢失统计量中。如果数据分组有优先级的区分, 则还要按照优先级的不同插入目的地址和优先级共同标识的缓存队列。
- ◆ **RR 状态:** 当 BCL\_PRO 模块发来询问有无数据要发送的消息时, 进入该状态机入口处 (状态机上半部), 轮询 (Round Robin) 所有缓存队列, 遇到第一个满足的队列, 即给 BCL\_PRO 返回肯定消息, 并在偏置时间后设置自中断, 转向 Assemble\_and\_send 状态; 若所有队列都没足够数据

发送, 则给 BCL\_PRO 返回否定消息。入口代码执行完毕, 从出口处返回中间 Judge 状态, 等待后续事件发生。如果有优先级区分, 可先轮询高优先级缓存队列, 只有当所有高优先级队列没有足够数据发送时, 才能轮询低优先级队列; 如果允许本地高优先级业务抢占过路低优先级业务, 则查询高优先级队列时可不按照轮询算法, 而采用计时器, 超时则认为可发。

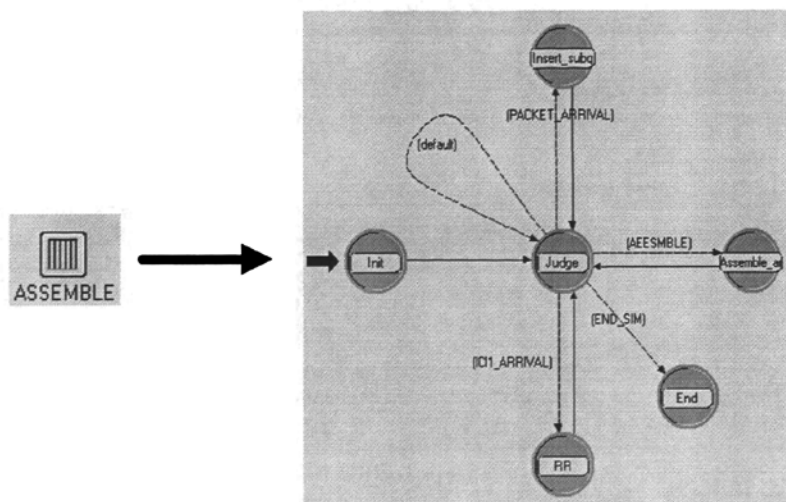


图 3.8 ASSEMBLE 模块进程层模型

- Assemble\_and\_send 状态: 当有足够数据可组装成 DB 时, RR 状态机会设置一个中断, 在收到相应 BCL 时刻的偏置时间后, 进入本状态机的入口处。按照 1.2 节中介绍的固定组装长度算法将所记录的缓存队列中的 IP 分组组装成 DB, 并将它发往 DBS\_PRO 模块。对于允许抢占的机制可用同时考虑时间和长度的组装算法。

### 三 BCL\_PRO 模块

BCL\_PRO 模块负责处理到达本结点的 BCL, 它对 BCL 及其对应的 DBS 进行忙闲判断, 并针对忙闲状态分别进行处理, 同时它还接收来自 ASSEMBLE 模块的应答消息。BCL\_PRO 的进程层模型如图 3.9。其中,

- Dispose\_BCL 状态: 当本结点收到上游结点发来的 BCL 时, 进入该状态。状态入口处代码首先检查 OT 值是否不足门限 (不足则加一倍的 RPT)、判断 BCL 是否空闲, 若空闲则查询是否有可用调谐发送机, 有才向 ASSEMBLE 模块发出询问是否有组装数据的消息, 无则在 OT 值中减去 BCL 处理时间并立即发送至下游结点; 若 BCL 忙, 就检查其对应的 DBS 是否以本结点为目的地, 是则配置接收机以便接收 DBS 并同上述过程查询本结点是否可发送据, 不是则配置光开关以便 DBS 穿过。

- ◆ **Dispose\_ICI2 状态:** 当 BCL\_PRO 模块收到来自 ASSEMBLE 模块的应答消息时, 进入该状态。若应答消息肯定有数据发送, 则将消息中携带的目的地址、优先级、长度等信息写入 BCL, 修改 OT 值, 预约调谐发送机, 将 BCL 发给下游结点; 若应答消息否定有数据要发, 则修改 OT 值, 将 BCL 发往下游结点。

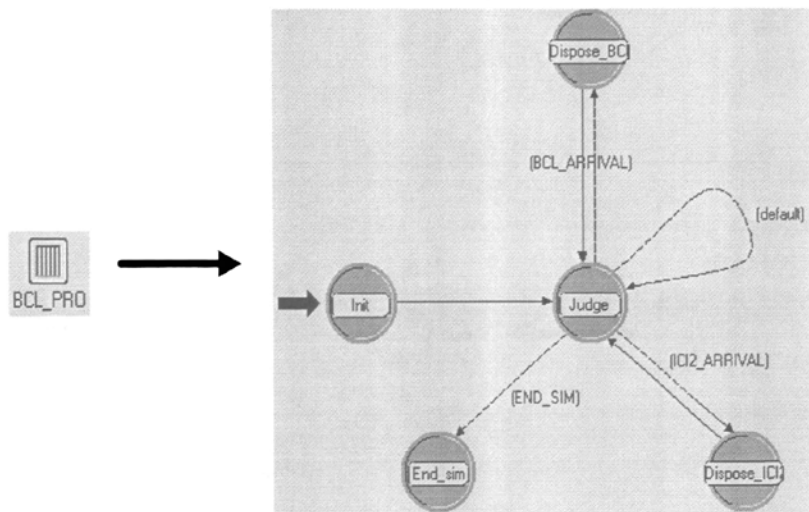


图 3.9 BCL\_PRO 模块进程层模型

负责 DBS 的收发和拆分及拆分后 IP 分组销毁的 DBS\_PRO 模块和 DISASSEMBLE 模块功能单一, 其进程层结构就不再详细介绍。

#### 四 OPNET Modeler 中的接口控制信息 (ICI, Interface Control Information) [22]

图 3.6 中 ASSEMBLE 模块和 BCL\_PRO 模块之间的双向箭头表示的是模块之间消息的传递。它不同于单向箭头表示的分组传递, 分组传递的目的是携带数据净荷, 而有时模块之间只需将很少的一部分控制信息互相告知, 并不需要传递大量的数据。OPNET Modeler 库中提供了适用于这种目的的灵活的接口控制信息, 可根据用户特定需求设置 ICI 格式, 并在模块之间传递。

ICI 有自己的编辑器, 即 ICI Editor。通过 ICI Editor, 用户可创建所需的 ICI 格式, 其中包括 ICI 属性, 各属性数据类型等。仿真核心调用 `op_ici_create()` 函数创建一个特定格式的 ICI, 通过 `op_ici_attr_set()`, `op_ici_attr_set_dbl()`, `op_ici_attr_set_ptr()` 等设置其属性值, 再通过 `op_ici_install()` 函数绑定在当前进程的某种中断上, 一旦这种中断发生, ICI 就随中断到达另一模块的某一进程。中断可以是 `op_pk_send()` 发送数据分组, `op_intrpt_force_remote()` 远程强制中断调用等。接收进程调用 `op_intrpt_ici()` 和 `op_ici_attr_get()` 函数获取该 ICI 及其属性值。这样模块间的消息传递完成。当不再需要某一 ICI 时, 用 `op_ici_destroy()` 函数销



毁之，并释放内存。

图 3.6 中 BCL\_PRO 模块将格式为 ICI1 的消息发给 ASSEMBLE 模块，询问其缓存队列中是否有足够的数据要发送。ASSEMBLE 模块收到 ICI1 后经查询，给 BCL\_PRO 模块一个应答消息 ICI2。ICI2 中携带了是否有数据可发、数据长度、目的地址等信息。

ICI 的使用是不同模块内进程之间传递消息的一种灵活方式。

### 3.4 分组格式

仿真模型中共有三种数据分组：IP 分组、突发控制标签 BCL、数据突发时隙 DBS。设定它们的格式分别如下：

#### (1) IP 分组

packet_count (8 bits)	packet_priority (8 bits)	packet_bit_len (16 bits)
packet_src (16 bits)		packet_dest (16 bits)
packet_start_time (32 bits)		
packet_payload		

图 3.10 IP 分组格式

根据 TCP/IP 协议设计 packet\_priority、packet\_bit\_len、packet\_src、packet\_dest、packet\_payload 各域，分别表示 IP 分组优先级，比特长度，源、目的地址，分组净荷。根据仿真需要，设计了 packet\_count 唯一标识 IP 分组，packet\_start\_time 以便获得分组时延统计量。

#### (2) BCL

Lamda_num (8 bits)	DBS_id (16 bits)	
offset_time (32 bits)		
priority (8 bits)	DBS_src (8 bits)	DBS_dest_bit_map (32 bits)
burst_bit_length (32 bits)		

图 3.11 BCL 仿真格式

根据 2.3 节中 BCL 的理论格式，设计了仿真中的 BCL 格式（图 3.11），其中 Lamda\_num 表示 BCL 对应 DBS 的所用波长序号，DBS\_id 记录对应 DBS 的编号并标志其忙闲状态，offset\_time、priority、DBS\_src、DBS\_dest\_bit\_map、burst\_bit\_length 则分别表示 BCL 与其对应 DBS 的偏置时间、优先级、源地址、目的地址比特表、数据突发长度。理论格式中的 Preamble，H-CRC 等域因仿真上不需要，没有设计。

(3) DBS 格式

DBS_id (16 bits)			burst_packet_count (16 bits)			
field_4 (0 bits)	field_5 (0 bits)	field_6 (0 bits)	field_7 (0 bits)	field_8 (0 bits)	field_9 (0 bits)	field_11 (0 bits)
field_23 (0 bits)	field_22 (0 bits)	field_21 (0 bits)	field_20 (0 bits)	field_19 (0 bits)	field_18 (0 bits)	field_17 (0 bits)
field_43 (0 bits)	field_42 (0 bits)	field_41 (0 bits)	field_40 (0 bits)	field_39 (0 bits)	field_38 (0 bits)	field_37 (0 bits)
field_33 (0 bits)	field_32 (0 bits)	field_31 (0 bits)	field_30 (0 bits)	field_29 (0 bits)	field_28 (0 bits)	field_27 (0 bits)
field_83 (0 bits)	field_82 (0 bits)	field_81 (0 bits)	field_80 (0 bits)	field_79 (0 bits)	field_78 (0 bits)	field_77 (0 bits)

图 3.12 DBS 仿真格式

仿真中不考虑保护带等因素，设计的 DBS 格式如图 3.12。其中 DBS\_id 唯一标识一个 DBS，burst\_packet\_count 域记录该 DB 携带了多少个 IP 分组，剩余的 field 域均承载了 IP 分组。

3.5 本章小结

本章给出了 OPNET Modeler 10.5 网络仿真平台下 VS-OBR 方案的详细建模过程。仿真模型由三种结点、三种链路、三种分组构成一个整体，其中 VS-OBADM 是环上的核心结点，负责重要的 BCL 处理、DBS 交换配置及 IP 分组组装等任务；IP 分组、控制分组 BCL、数据突发时隙 DBS 分别在接入链路、控制波长链路和数据波长链路上流动。

使用本章所述模型进行仿真，可得大量仿真数据，下一章根据这些数据结果对 VS-OBR 的一些重要性能进行分析。

## 第四章 VS-OBRR 性能分析

使用第三章描述的模型进行仿真,本章对 VS-OBRR 的一些关键性能做了分析。

### 4.1 最大吞吐量

VS-OBRR 方案中,我们假定接入链路和数据波长链路都不会发生故障,数据丢失不会发生在链路上;接收机数目等于数据波长数目,环上结点不会有接收冲突,而方案本身又避免了发送冲突,则数据丢失的唯一原因只可能是源结点的上路业务数据缓存器溢出。

当环网输入的业务量超过其吞吐能力时,环网将工作在饱和状态,这时的输出业务量即为 VS-OBRR 的最大吞吐量。设环上  $N$  个结点,每个结点发送数据的概率相等,同时发往其他结点的业务是均衡的,即通过环上一个结点发往其他  $N-1$  个目的结点的分组概率相等;每个结点都有一定的缓存容量容纳上路和下路业务,饱和状态下,缓存器应该是充满的,每段链路上都不会存在空闲的 DBS;这时,每个结点只能利用本地下路业务占用的 DBS 卸载完数据后空闲下来,发送自己的上路业务。不考虑由于 ROTC 机制对 BCL 中偏置时间的调整带来的容量损失。

用符号  $\ominus$  表示“模  $N$  减”运算,即若在环上任选一结点  $N_i$ ,则其紧邻上游结点为  $N_{i\ominus 1}$ ,次邻上游结点为  $N_{i\ominus 2}$ ,以此类推;链路用两个相邻结点编号对来表示。观察通过链路  $(i\ominus 1, i)$  的任一 DBS,如果该 DBS 发往目的结点  $N_i$  的概率为  $p_{dest\_i}$ ,那么根据前面的分析,结点  $N_i$  可以获得 DBS 进行发送的业务量  $A_i = p_{dest\_i}$  (设数据链路总容量为单位 1)。在业务均衡假设下各个结点的发送量相同,因而  $A_i = A$ ,  $p_{dest\_i} = p_{dest}$ ,于是环网在饱和情况下的总吞吐量为:

$$A_o = \sum_{i=0}^{N-1} A_i = \sum_{i=0}^{N-1} p_{dest\_i} = Np_{dest} \quad \text{式(4-1)}$$

对于结点  $N_i$ ,经过链路  $(i\ominus 1, i)$  上到达  $N_i$  的 DBS 可能来自环上各个上游结点,因此:

$$p_{dest\_i} = \sum_{j=i' \atop k=1}^{N-1} p_{dest\_i|src\_j} p_{src\_j} \quad \text{式(4-2)}$$

其中  $p_{dest\_i|src\_j}$  表示 DBS 来自结点  $N_j$  同时目的地为结点  $N_i$  的条件概率,  $p_{src\_j}$  表示 DBS 来自结点  $N_j$  的概率。在业务均衡情况下,每个结点发出的业务将有不同的份额会经过链路  $(i\ominus 1, i)$ ,例如,从  $N_{i+1}$  发出的到达其它  $N-1$  个结点的业务

全部要经过  $(i \ominus 1, i)$ , 从  $N_{i \ominus 2}$  发出的只有  $N_{i \ominus 1}$  之外的其余  $N-2$  个目的地址的业务会经过  $(i \ominus 1, i)$ , 以此类推, 对于  $N_{i \ominus N-1}$ , 即单向环上距离  $N_i$  最远的下游紧邻结点, 其发出的业务中只有目的地为  $N_i$  的会经过链路  $(i \ominus 1, i)$ 。因此对于结点  $j = i \ominus k$  有,

$$p_{src\_j} = \frac{N-k}{(N-1)+(N-2)+\cdots+1} = \frac{2(N-k)}{N(N-1)}, \quad j = i \ominus k \quad \text{式(4-3)}$$

$$p_{dest\_i|src\_j} = \frac{1}{N-k}, \quad j = i \ominus k \quad \text{式(4-4)}$$

将(4-3)和(4-4)代入(4-2)中, 可得:

$$p_{dest\_i} = \sum_{j=i \ominus k, k=1}^{N-1} \frac{1}{N-k} \frac{2(N-k)}{N(N-1)} = \frac{2}{N} \quad \text{式(4-5)}$$

于是由(4-5)可得 VS-OBRR 环网在饱和业务输入条件下的极限吞吐量为:

$$A_o = Np_{dest} = 2 \quad \text{式(4-6)}$$

即负载饱和且业务量均衡的情况下, VS-OBRR 的最大吞吐量为环路数据容量的两倍。这正是由于 VS-OBRR 控制方案允许同一时间同一波长在环上不同位置传送多个 DBS, 即空间重用带来的增益。

对此, 我们在 OPNET Modeler 下, 使用不同业务类型的源、针对不同的环网规模进行了仿真。其仿真条件为 6 个结点的环网, 设置每段链路的长度为 30km, 传播时延 0.00015s, 每个 DBS 可承载 375 000bit 的业务, 不考虑保护带, 将组装上下门限设置为 DBS 总长的 50% 和 100%, 每个结点处理 BCL 需用时 0.1  $\mu$ s。

我们使用以下三种不同的源: 包长为常量 3 000bit 到达间隔为常量的常量源, 包长为指数分布 (均值为 3 000bit) 到达服从泊松分布的泊松源, 包长按概率取

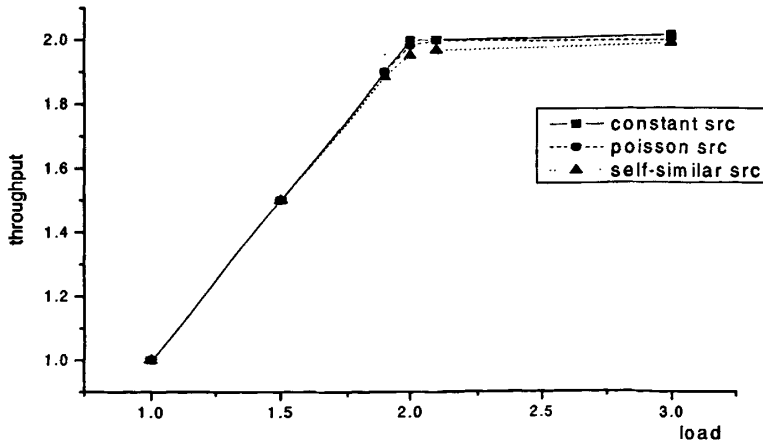


图 4.1 不同业务源对 VS-OBRR 最大吞吐量的影响

40byte、52byte、576byte、1500byte（均值为 3 718.24bit）的自相似源<sup>[20][21]</sup>进行源对环网最大吞吐量影响的测试。

图 4.1 的仿真结果表明当输入业务量小于 1.9 时,三种业务源的丢失率几乎为 0,即环网接收多少业务就能成功传输多少;当输入业务量接近 2 时,开始有所丢失;当输入业务超过 2 时,丢失增加,环网的吞吐量基本维持在 2 附近。常量源分组定长、固定到达的特性使得它的丢失率最小,泊松源分组长度变化、到达离散,丢失分组多于常量源,而自相似源的分组长度离散、到达具有很强的突发性,丢失较前两种源最多。对这三种源,环网的最大吞吐量均保持在 2 左右,这是 VS-OBDR 方案可空间重用和无冲突的控制机制带来的良好性能,同时也说明 VS-OBDR 对于不同的业务源有良好的适应性。

为了进一步分析 VS-OBDR 的吞吐量,我们在 6、9、12 个结点的环网上分别加载自相似源,测试不同规模下环网的最大吞吐量。这里每段链路的长度仍设为 30km。

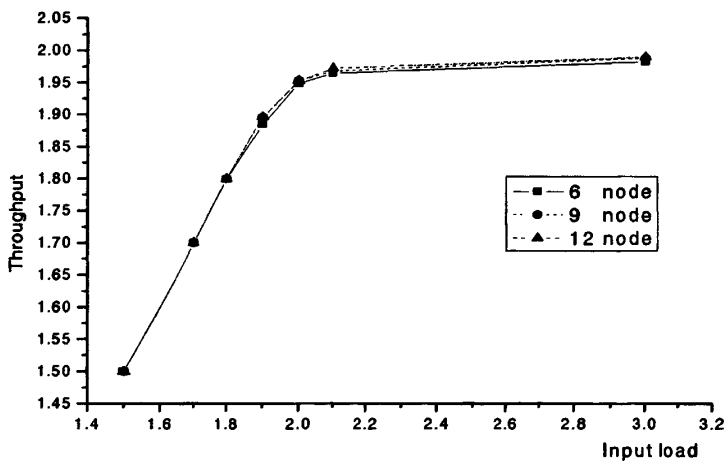


图 4.2 不同结点数对 VS-OBDR 最大吞吐量的影响

图 4.2 是不同结点数下环网吞吐量的仿真结果,由图中可看出环网的最大吞吐量仍然保持在 2 附近;而随着结点数的增加,在相同总业务负载下,每个结点的输入业务量减小,同时缓存队列个数增多,相当于缓存容量增加,使得丢失下降,吞吐量有所上升。

## 4.2 缓存利用率

在实际网络中,缓存作为一种消耗资源,在满足要求的条件下应减少其使用量。我们在 VS-OBADM 结点中设置不同的缓存队列大小,研究缓存的最佳利用。

仿真是在 6 个结点的 VS-OBDR 上进行的,接入自相似业务源,其他参数设置

与 4.1 节中第一个仿真相同，全环输入业务量为 1.9 倍的链路总容量（这里链路总容量设为单位 1），结点上每个缓存队列分别置 10 倍、15 倍、20 倍、25 倍、30 倍的 DBS 总长。表 4.1 是仿真结果。

表 4.1 缓存与丢失率和时延的比较

缓存容量 (Mbit)	平均比特丢失率	平均端到端时延(s)	平均缓存利用率
3.75 (10×DBS)	0.007809	0.001505	36.29107%
5.625 (15×DBS)	0.007085	0.001636	26.85255%
7.5 (20×DBS)	0.005904	0.001821	23.26523%
9.375 (25×DBS)	0.005154	0.001988	20.84782%
11.25 (30×DBS)	0.004600	0.002142	19.16676%

表 4.1 中数据显示，随着缓存容量的增加，丢失率依次减小；缓存无溢出时，分组在缓存中滞留等待组装的时间变长，从环上源结点到目的结点的端到端时延增加；虽然缓存容量增大，但使用的缓存增长并不多，缓存利用率减小。在对时延不很敏感的业务中，可选用 10 倍 DBS 总长的缓存，这样时延较小、缓存利用率高，而丢失率也不大。

4.3 丢失率和时延性能

网络中数据丢失率和端到端时延是两个重要的性能指标，这里我们采用 6 个结点、自相似源的仿真模型对 VS-OBR 的丢失率和平均端到端时延性能进行测试和分析。

仿真条件为：环长 180km，每条链路长 30km，链路时延 0.00015s，一个 DBS 可承载 375000bit，组装上下限为 DBS 总长度的 50%和 100%，结点中每个缓存队列设置为 10 倍的 DBS 长度，每个结点 BCL 处理时间为 0.1  $\mu$ s。

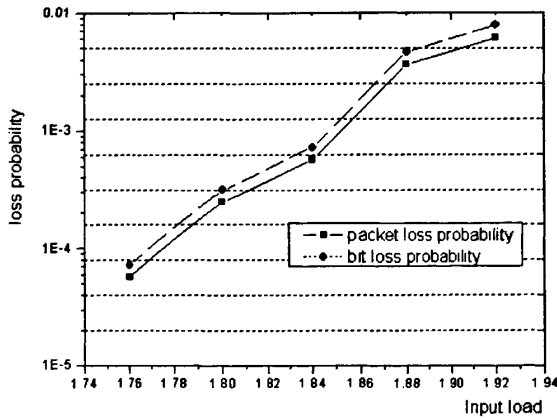


图 4.3 自相似源 6 个结点环的丢失率

定义：分组丢失率=丢失分组数/进入环总分组数；比特丢失率=丢失比特数/进入环总比特数。端到端时延指分组进入环网后从环上源结点到目的结点的时延。

图 4.3 是输入业务一丢失率对比曲线，可看出小于最大吞吐量时输入业务增加，分组丢失率和比特丢失率增加；而同一业务负载下分组丢失率要小于相应的比特丢失率，根据分组长度分布容易得出长分组丢失概率要大于短分组丢失概率。

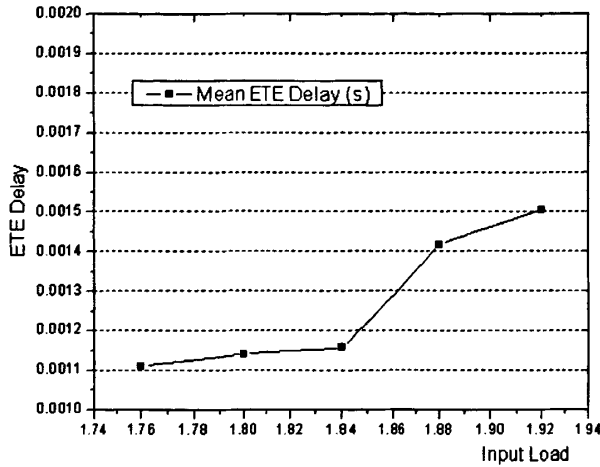


图 4.4 自相似源 6 个结点环的平均端到端时延

图 4.4 展示的是平均端到端时延，输入业务量较小时，时延很小且平稳；输入业务接近最大吞吐量时，平均时延增大，但数量级是几个毫秒，这对于城域网是可以接受的时延。

#### 4.4 区分服务的性能分析

2.7 节中提到 VS-OBRR 可支持全网范围的优先级区分服务，并可在此基础上实现本地高优先级业务抢占过路低优先级业务。相对于 VS-OBRR 的优先级区分服务，现有环网区分服务控制机制复杂且不易在硬件上实现，其原因在 1.3 节中已详细阐述。

针对 VS-OBRR 的这种优良性能，我们首先设计了有两个优先级区分的服务模型。这一点主要体现在 VS-OBADM 结点中 ASSEMBLE 模块的查询和组装算法的改动上。在有优先级区分的情况下，查询时先轮询高优先级，再轮询低优先级；组装时将高低优先级的业务分开组装。

图 4.5 是 6 个结点的环网在自相似业务下有两个优先级区分的丢失率和时延性能曲线。在全环总业务量一定的情况下，图 4.5a 显示高优先级所占比例增大，分组和比特丢失率也随着增大，而低优先级的分组和比特丢失率略有减小。这是因为当高优先级业务溢出缓存丢失后，就不再能组装成突发，它省下的空闲 BCL

被低优先级业务使用, 导致低优先级的丢失有所下降。图 4.5b 显示的是同种业务量下的高低优先级时延曲线, 随着高优先级业务量的增加, 它到达长度组装门限的速度加快, 于是较频繁的被组装发送, 致使高优先级时延减小; 相反, 低优先级业务在总业务量中不断减小, 缓存中长时间达不到数据组装下限, 不能及时被

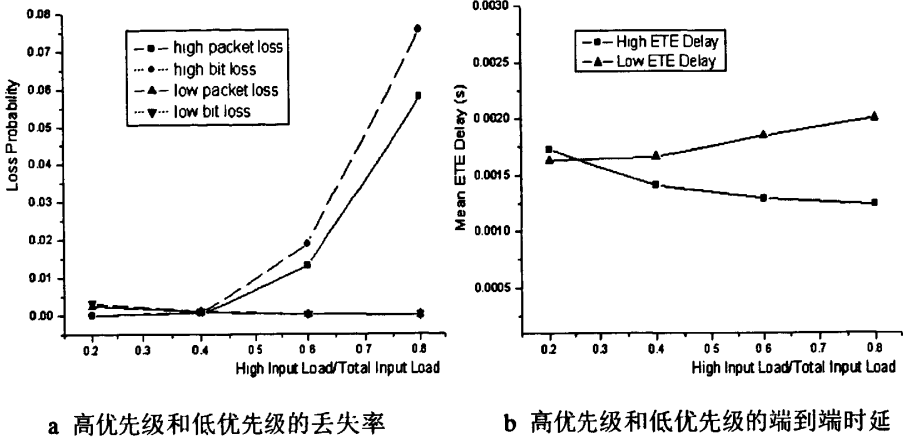


图 4.5 区分服务的丢失率和时延

组装, 导致低优先级业务时延增加。

从丢失率曲线可看出, 当高优先级业务比例越来越重或突发性增强时, 高优先级的丢失率会增大, 这时就需要进行高优先级业务对低优先级业务的抢占。其改进为: 首先, 在环上 VS-OBADM 结点的 ASSEMBLE 模块中, 查询高优先级缓存队列中等待组装时间最长的, 若该等待时间超过组装时间门限则允许它抢占过路的低优先级业务; 若没有超时缓存队列, 则找到第一个达到长度组装门限的队列准备组装; 若以上两项都没有, 则进行低优先级队列的轮询。其次, 组装时可将高低优先级业务分开组装成突发, 也可在高优先级组装不满时搭载低优先级业务。另外, 低优先级业务的处理策略可有直接被丢弃、缓存于本地结点的一个单独队列中或是和本地低优先级业务混合缓存。

仿真中应是在区分了优先级的基础上加入抢占机制。但因时间仓促, 我们只进行了高优先级查询、超时组装和抢占时机的条件研究, 没能在仿真中实现。

## 4.5 本章小结

本章依据仿真得到的结果, 对 VS-OBDR 的最大吞吐量、缓存利用、丢失率、端到端时延和区分服务等性能进行了数据和理论分析。仿真结果证明 VS-OBDR 因空间重用最大吞吐量可达环链路总容量的 2 倍; 其仿真最佳缓存为设定的 DBS 总长的 10 倍; 丢失率和端到端时延随业务的增大而增加; 仿真中很好的实现了优先级区分服务, 在此基础上的抢占机制只进行了理论研究。



## 结束语

本文针对现有 WDM 光纤环网中的控制问题,提出了一种新的 WDM 环网控制方案,称之为虚同步光突发环网 VS-OBRR。该方案将数据信道划分时隙,使用弹性偏移时间控制机制实现控制信道中突发控制标签和数据信道中数据突发时隙的持续映射和有效控制。VS-OBRR 可提供光纤链路的空间重用、信道的灵活接入、支持可抢占的优先级,不需要在结点配置光纤延迟线、也不需要全网的严格同步。

本文的主要工作包括论述 VS-OBRR 的结构、BCL 和 DBS 的格式、结点的工作机制和可支持的性能,在 OPNET Modeler 10.5 下建立结点模型、构建仿真网络,并根据仿真结果对 VS-OBRR 的最大吞吐量、缓存利用、丢失率、时延等关键性能进行数据和理论分析。

在研究支持可抢占的优先级区分服务时,由于时间所限,我们只给出了策略分析,并没有在仿真平台上实现,从而得到仿真结果,这是本文的一大遗憾,也是后续工作的内容之一。此外,为了将 VS-OBRR 尽快应用于实际,有待进一步研究的问题还有:

(1) BCL 的再生。当链路或结点发生故障致使 BCL 丢失后,要在正确位置再生丢失的 BCL。

(2) 环的保护与恢复。可利用双环对当前使用的链路进行保护。

(3) 环的公平性研究。对 VS-OBRR 抢占的公平性进行分析,针对业务特性提出公平策略。

(4) BCL 核心控制单元的硬件实现。用 FPGA 等硬件实现核心的处理单元,构建一个由实际器件组成的实验平台。

## 致谢

一位师姐曾跟我说，学术研究就像走在多叉的路上，一不小心就会走进没有出路的分叉口。每当我钻进这样的死胡同时，感谢我的导师文爱军副教授总将我拉回正轨。文老师这种方向性导引如茫茫大海上的明灯，指引我的学术研究在正确的航向上航行。同时，文老师将求新、严谨而又踏实的工作精神通过自己的行为传递给我们，使我从内心深处开始追求创新、追求不苟、追求不懈。值此论文完成之时，特别感谢我的导师，感谢老师一直以来对我的指导和关怀！

感谢姚明昨博士为我开启了 OBS 研究的绚烂大门，使我第一次深刻认识了 OBS 研究中的重大问题，第一次体会将新的思想用仿真实现的快乐。极为可贵的是，无论什么时候找他讨论问题，他总是满心热情并毫无保留的指导我帮助我，使我从中受益良多。

感谢年轻的留校老师易湘博士、岳鹏博士，将自己工作和生活上的许多经验与我分享。

感谢 OBS 项目组的所有成员，从最初对我耐心的指导到后来激烈的讨论，都为我带来宝贵的团队合作经验。他们是刘建平博士，已毕业的黄伯虎、靖杰师兄，有着深厚同级之谊的毋丹芳和正在读研的张志卿、何阳师弟。

从 2004 年夏进入实验室，到 2007 年春离开，我先后与五届师兄师姐师弟师妹们一起学习和生活，感谢大家用真诚和宽容营造了如家般温暖的实验室，感谢一起走过那些风雨兼程的日子。

真挚的爱和感激致给亲爱的父母和姐姐，感谢你们倾家所有支持我读书至今，并在生活上给予无微不至的关爱。你们的支持和疼爱是我一生不竭的动力。

最后感谢各位评审和答辩老师，耐心的将论文看至最后一页，谢谢你们辛勤的劳动和宝贵的批评意见。

## 参考文献

- [1] A. I. Siahlo, J. Seoane, A.T. Clausen, et al.. 320 Gb/s Single-Polarization OTDM Transmission over 80km Standard Transmission Fiber. Optical Fiber Commun. Conf. 2005. Mar 2005, Technical Digest vol. 6
- [2] Christian Rasmussen, Tina Fjelde, Jon Bennike, et al.. DWDM 40G transmission over trans-pacific distance (10 000 km) using CSRZ-DPSK, enhanced FEC, and All-Raman-Amplified, 100-km UltraWave Fiber Spans. Journal Of Lightwave Technology. January 2004, Vol. 22, No. 1. pp.203-207
- [3] G. Charlet, E. Corbel, J. Lazaro, et al.. WDM transmission at 6 Tbit/s capacity over transatlantic distance, using 42.7 Gb/s differential phase-shift keying without pulse carver. Proc. Optical Fiber Conference. PD36,2004
- [4] J. X. Cai, D. G. Foursa, C. R. Davidson, et al.. A DWDM demonstration of 3.73 Tb/s over 11 000 km using 373 RZ-DPSK channels at 10 Gb/s. OFC03. PD22, 2003
- [5] Byrav Ramamurthy, Biswanath Mukherjee. Wavelength Conversion in WDM Networking. IEEE J. Select. Areas Commun.. Sep. 1998, Vol. 16, No. 7. pp.1061-1073
- [6] S. Yao, B. Mukherjee, and S. Dixit. Advances in photonic packet switching: An overview. IEEE Commun. Mag.. Feb. 2000, vol. 38. pp. 84-94
- [7] Dong-Hee Park and Myungsik Yoo. Comparative study on burst assembly schemes for optical burst-switched networks. Proc. SPIE Int. Soc. Opt. Eng. 4910, 266 (2002)
- [8] Jingxuan Liu, Ansari N., Ott T.J.. FRR for latency reduction and QoS provisioning in OBS networks. IEEE Journal On Selected Areas In Commun.. September 2003, Vol. 21, No. 7. pp. 1210-1219
- [9] Yutaka Arakawa, Naoaki Yamanaka, and Iwao Sasase. Performance of Optical Burst Switched WDM Ring Network With TTFR System. in Proc. 1st IFIP Optical Networks & Technologies Conf. (OpNeTec), Pisa, Italy. Oct. 2004. pp. 95-102
- [10] Lisong Xu, Harry G. Perros, George N. Rouskas. A simulation study of optical burst switching and access protocols for WDM ring networks. Computer Networks 41 (2003). pp.143–160

- [11] Fumagalli, A., Krishnamoorthy, P.. A low-latency and bandwidth-efficient distributed optical burst switching architecture for metro ring. Commun., 2003. ICC '03, Proceedings of IEEE International Conf.. May 2003, Volume 2, 11-15. pp. 1340-1344
- [12] James Cai, Andrea Fumagalli, and Imrich Chlamtac. The Multitoken Interarrival Time (MTIT) Access Protocol for Supporting Variable Size Packets Over WDM Ring Network. IEEE Journal on Selected Areas in Communications. Oct. 2000, Vol. 18, No. 10. pp. 2094-2104
- [13] 姚明昨. 光突发交换网络核心结点关键技术研究. 西安电子科技大学. 博士学位论文. 2006
- [14] Marcel Neuts, Zvi Rosberg, Hai Le Vu, Jolyon White, and Moshe Zukerman. Performance Analysis of Optical Composite Burst Switching. IEEE Commun. Letters. August 2002, Vol. 6, No. 8. pp. 346-348
- [15] 陈敏. OPNET 网络仿真. 清华大学出版社. 2004
- [16] 李建东. 信息网络理论基础. 西安电子科技大学出版社. 2001
- [17] William Stallings. “高速网络与互联网——性能与服务质量”. 电子工业出版社. 2003. pp. 174-188
- [18] 胡严, 刘星成, 张光昭. “快速 RMD 法产生的自相似业务流的性质研究”. 中山大学学报 (自然科学版). May 2001, Vol. 40, No. 13
- [19] 吴援明, 宁正容, 梁恩志. “网络自相似业务模型进展”. 通信学报. March 2004. Vol. 25, No. 3
- [20] "Packet Length Distributions". [http://www.caida.org/analysis/AIX/plen\\_hist](http://www.caida.org/analysis/AIX/plen_hist)
- [21] "WAN packet size distribution". <http://www.nlanr.net/NA/Learn/packetsizes.html>
- [22] OPNET Modeler 10.5 Help Document

## 在读期间研究成果

在硕士研究生期间取得的成果如下：

### 一 科研情况

1. IP/MPLS over WDM 网络中动态业务疏导算法的研究
2. 虚同步光突发环网的研究，属华为高校基础研究基金资助项目“光突发交换技术”

### 二 发表论文

1. 宁波, 文爱军. IP/GMPLS over WDM 网络中动态业务疏导算法的比较. 西安电子科技大学通信工程学院 2006 研究生学术年会
2. 姚明旻, 文爱军, 刘增基, 宁波, 毋丹芳. VS-OBR: 一种无冲突虚同步光突发环网. 投稿

### 三 专利申请

- 一种虚同步的光突发交换城域网 （专利审查中）