

光突发交换技术

© 乐孜纯 朱智俊 付明磊 王强 编著



GUANGTUFA
JIAOHUA JISHU



机械工业出版社
CHINA MACHINE PRESS

光突发交换技术

乐孜纯 朱智俊 付明磊 王 强 编著



机械工业出版社

光突发交换技术是近年来受到学术界和产业界广泛关注的光交换技术。本书围绕光突发交换技术中的网络架构、节点结构、关键算法、实现技术等核心问题,介绍了光交换技术概况、网络业务量与常用网络性能指标、光突发交换网的汇聚技术、光突发交换网的信令技术、光突发交换网的调度技术、光突发交换网的冲突消解技术、光突发交换网的路由技术、光突发交换网络的生存性技术、光突发交换网络的仿真软件平台、光突发交换网络边缘节点的硬件设计、光突发交换网络核心节点的光学结构设计等内容。全书内容全面系统,专业性和针对性强。

本书可供从事光网络研究与应用的科研与技术人员,以及高校相关专业的教师、研究生和本科生学习和参考。

图书在版编目(CIP)数据

光突发交换技术/乐孜纯等编著. —北京:机械工业出版社, 2014.7

ISBN 978-7-111-46697-0

I. ①光… II. ①乐… III. ①光纤通信-通信交换
IV. ①TN929.11②TN91

中国版本图书馆CIP数据核字(2014)第092805号

机械工业出版社(北京市百万庄大街22号 邮政编码100037)

策划编辑:吉玲 责任编辑:吉玲 张丽萍 刘丽敏

版式设计:赵颖喆 责任校对:佟瑞鑫

封面设计:张静 责任印制:刘岚

北京京丰印刷厂印刷

2015年1月第1版第1次印刷

184mm×260mm·11印张·264千字

标准书号:ISBN 978-7-111-46697-0

定价:35.00元

凡购本书,如有缺页、倒页、脱页,由本社发行部调换

电话服务

网络服务

社服务中心:(010) 88361066 教材网:<http://www.cmpedu.com>

销售一部:(010) 68326294 机工官网:<http://www.cmpbook.com>

销售二部:(010) 88379649 机工官博:<http://weibo.com/cmp1952>

读者购书热线:(010) 88379203 封面无防伪标均为盗版

前 言

随着信息通信技术的蓬勃发展,人们对高速移动互联、高清视频点播、高质量视频电话等新兴业务的需求急剧增加。加之物联网技术逐步得到推广与认可,人与人、人与物、物与物的联系与互动使得当前的信息交互量与日俱增。因此,选择具有大容量、高速率和智能化特点的光交换网络作为未来电信网与互联网的基础信息承载机制已经成为业界的共识。

光突发交换技术是近年来受到学术界和产业界广泛关注的光交换技术。它与光线路交换技术和光分组交换技术并称为“光交换技术领域的三大主流技术”。尤为重要的是,光突发交换技术在网络性能、技术实现难度和现有光学器件的成熟度等方面实现了较完美的折中,有望实际应用并示范推广。本书围绕光突发交换技术中的网络架构、节点结构、关键算法、实现技术等核心问题展开,旨在帮助读者更好地了解和学习光突发交换技术的基本原理和关键技术。全书共分11章,第1章是光交换技术概述;第2章介绍网络业务量与常用网络性能指标;第3章介绍光突发交换网的汇聚技术;第4章介绍光突发交换网的信令技术;第5章介绍光突发交换网的调度技术;第6章介绍光突发交换网的冲突消解技术;第7章介绍光突发交换网的路由技术;第8章介绍光突发交换网络的生存性技术;第9章介绍光突发交换网络的仿真软件平台;第10章介绍光突发交换网络边缘节点的硬件设计;第11章介绍光突发交换网络核心节点的光学结构设计。

本书第1章、第3~5章和第11章由乐孜纯撰写;第7章、第9章、第10章由朱俊撰写;第6章和第8章由付明磊撰写;第2章由王强撰写。全书最后由乐孜纯统编定稿。

本书凝聚了笔者所在科研团队多年的科研经验和实践总结,并且自2005年起有幸得到国家自然科学基金(61172081、61007031)、浙江省科技计划重大和重点项目(2009C11051、2005C21010)、浙江省自然科学基金(LZBF010001、Y1080172、Y1110923、LQ12F05008)和浙江省科技计划公益性工业项目(2011C21011)等科研项目的资助支持。在此,对国家自然科学基金委员会、浙江省科技厅和浙江省自然科学基金委员会表示最衷心的感谢。同时,本书还包含了科研团队中张明和全必胜两位老师,以及方江平、潘迪飞、朱晓菲、郑和蒙、郭蕊、朱振华、朱冉、陈拓、陈伟、李翔、卞燕如、陈君、侯继斌硕士在他们攻读硕士学位的部分研究成果。同时,王伟文、孟轲、柴志成、张彭生、厉屹和沈立峰对文稿的文字校对和插图处理也做了很多工作,在此一并感谢。

由于作者水平有限,书中难免有错误与不足之处,敬请同行和读者们批评指正。

编 者

目 录

前言

第1章 光交换技术概述	1	3.7 OBS-Ethernet 汇聚网卡的设计方案	22
1.1 光交换技术简介	1	3.7.1 方案概述	22
1.2 光交换技术分类	2	3.7.2 以太网物理层接口	22
1.3 三种主流光交换技术	2	3.7.3 发送模块	22
1.3.1 光线路交换技术	2	3.7.4 接收模块	24
1.3.2 光分组交换技术	3	3.7.5 MII 管理模块	25
1.3.3 光突发交换技术	3	3.8 本章小结	26
1.3.4 三种光交换技术的比较	4	第4章 光突发交换网的信令技术	27
1.4 光交换技术的国内外研究进展	5	4.1 资源预留方式及分类	27
1.5 本章小结	6	4.1.1 资源预留方式	27
第2章 网络业务量与常用网络性能		4.1.2 典型资源预留协议	29
指标	7	4.2 JIT 控制协议的基本原理	30
2.1 泊松业务量模型	7	4.3 JIT 控制协议的实现过程	31
2.2 自相似业务量模型	7	4.3.1 JIT 控制协议总体设计	31
2.2.1 网络流量的自相似特性	7	4.3.2 JIT 协议控制分组格式	32
2.2.2 自相似的数学定义	8	4.3.3 JIT 协议子模块设计	35
2.2.3 网络流量的自相似模型分析	9	4.4 Just Enough Time 控制协议基本	
2.3 常用网络性能指标	12	原理	47
2.4 本章小结	12	4.5 Just Enough Time 控制协议的实现	
第3章 光突发交换网的汇聚技术	13	过程	48
3.1 OBS 边缘节点的基本结构	13	4.5.1 控制模块的总体设计	48
3.2 汇聚算法的基本原理	14	4.5.2 JET 协议控制分组格式设计	49
3.3 汇聚模块的主要设计参数	14	4.5.3 JET 协议子模块设计	49
3.3.1 汇聚时延	15	4.6 本章小结	50
3.3.2 BDP 包长	15	第5章 光突发交换网的调度技术	51
3.4 典型汇聚算法分类	15	5.1 调度模块的基本原理	51
3.4.1 固定分组长度汇聚机制	15	5.2 典型调度算法	51
3.4.2 固定汇聚时间汇聚机制	16	5.2.1 FF 算法	51
3.4.3 最大突发长度最大突发汇聚		5.2.2 LAUC 算法	52
时间汇聚机制	17	5.2.3 LAUC-VF 算法	53
3.4.4 最小-最大突发长度最大突发		5.3 新型数据信道调度算法	55
汇聚时间汇聚机制	17	5.3.1 基于 PPJET 协议的数据信道	
3.4.5 自适应突发汇聚机制	17	调度算法	55
3.5 汇聚算法影响自相似业务的理论		5.3.2 一种改进的基于 PPJET 协议的	
分析	18	数据信道调度算法	55
3.6 新型自适应汇聚策略 CBAAP	20	5.4 本章小结	57

第6章 光突发交换网的冲突消解

技术 58

6.1 光缓存技术 58

6.2 波长转换技术 59

6.2.1 波长转换技术的基本原理 59

6.2.2 OBS网络中两种波长转换方式的性能分析 60

6.3 基于有限波长转换的LAUC算法 61

6.3.1 CS集合的计算 61

6.3.2 算法执行步骤 62

6.3.3 仿真结果及分析 62

6.4 偏射路由技术 64

6.4.1 偏射路由技术的基本原理 64

6.4.2 基于AIMD控制的偏射路由技术 65

6.5 突发分片技术 68

6.6 多种冲突消解技术的联合 68

6.7 本章小结 69

第7章 光突发交换网的路由技术 70

7.1 路由协议的分类 70

7.2 典型路由协议 71

7.2.1 边界网关路由协议 71

7.2.2 内部网关路由协议 72

7.3 OBS路由协议帧格式及流程 77

7.3.1 OBS路由协议帧格式 77

7.3.2 OBS路由协议流程 78

7.4 OBS路由协议的实现 79

7.4.1 整体流程图 79

7.4.2 处理数据模块子程序流程图 79

7.5 本章小结 80

第8章 光突发交换网络的生存性

技术 81

8.1 OBS网络的故障类型 81

8.1.1 OBS网络故障的层次性 81

8.1.2 网络节点故障与网络链路故障 81

8.2 OBS网络故障检测与恢复技术的基本步骤 82

8.3 OBS网络故障检测与定位技术 83

8.4 基于BFD的OBS网络故障检测与定位技术 84

8.4.1 BFD协议简介 84

8.4.2 BFD协议的报文格式 84

8.4.3 BFD协议的状态机 85

8.4.4 BFD协议的运行模式 85

8.4.5 异步模式下的网络节点故障检测与网络链路故障检测 86

8.4.6 BFD协议在OBS网络中的报文交换过程 87

8.4.7 BFD协议在OBS网络中故障检测与定位的实施步骤 88

8.5 OBS网络故障保护与恢复技术 89

8.5.1 端到端的FRR技术 89

8.5.2 本地FRR技术 90

8.6 本章小结 90

第9章 光突发交换网络的仿真软件

平台 91

9.1 NS-2简介 91

9.2 NS-2软件构成 92

9.3 NS-2使用的仿真语言 92

9.4 几种典型的OBS仿真平台 93

9.4.1 OBS-ns Simulator 93

9.4.2 NCTUns 93

9.4.3 OBS Simulator 93

9.4.4 其他仿真工具简介 94

9.5 OBS-NP光网络仿真软件概述 94

9.6 OBS-NP软件设计 95

9.6.1 OBS网络模拟的过程 95

9.6.2 安装OBS-NP 95

9.6.3 OBS-NP软件结构 97

9.7 OBS-NP仿真实例 98

9.7.1 Pareto流叠加产生自相似流 98

9.7.2 OBS汇聚算法仿真 104

9.7.3 AWK对TRACE文件分析 115

9.8 本章小结 118

第10章 光突发交换网络边缘节点

的硬件设计 119

10.1 FPGA简介 119

10.2 基于QuartusII的FPGA开发流程 119

10.3 OBS边缘节点功能模块总图 120

10.4 三种汇聚算法的FPGA实现 121

10.4.1 FAS汇聚算法的FPGA实现 121

10.4.2 FAP汇聚算法的FPGA实现 122

10.4.3 MBMAP汇聚算法的FPGA实现 124

10.5 LAUC 数据信道调度算法的 FPGA 实现	125	10.8.6 以太网帧格式设计	141
10.5.1 LAUC 数据信道调度算法的顶层图	125	10.8.7 汇聚网卡软件部分	143
10.5.2 初始时间表为空的时序仿真分析及实际输出波形	126	10.8.8 测试结果	145
10.5.3 初始时间表为非空的时序仿真分析及实际输出波形	128	10.9 本章小结	146
10.6 资源预留协议的 FPGA 实现	130	第 11 章 光突发交换网络核心节点的 光学结构设计	147
10.6.1 JIT 协议的 FPGA 实现	130	11.1 光交叉连接器概述	147
10.6.2 JET 协议的 FPGA 实现	132	11.1.1 OXC 的主要类型	148
10.7 路由协议中帧交换过程的 FPGA 实现	134	11.1.2 两种典型的基于空间交换的 OXC 结构	148
10.7.1 Hello 帧的时序仿真与实际输出波形	134	11.1.3 两种典型的基于波长转换的 OXC 结构	149
10.7.2 链路状态分组的时序仿真与实际输出波形	135	11.1.4 多粒度光交叉连接结构 MG-OXC	150
10.7.3 链路状态数据库分组的时序仿真与实际输出波形	135	11.2 用于 OBS 核心节点的 L-WIXC	152
10.7.4 应答帧的时序仿真与实际输出波形	136	11.2.1 L-WIXC 的光学结构及其性能、成本比较分析	152
10.7.5 四类路由帧的时序总图	137	11.2.2 L-WIXC 的控制方法	154
10.8 低成本 OBS-Ethernet 汇聚网卡设计	137	11.2.3 L-WIXC 光学参数仿真平台设计	156
10.8.1 实现内容与设计目标	137	11.2.4 L-WIXC 光学参数仿真结果与分析	157
10.8.2 以太网基础知识	138	11.3 用于 OBS 核心节点的 MG-OXC	161
10.8.3 MII 接口设计	139	11.3.1 MG-OXC 的光学系统结构	161
10.8.4 以太网控制器	139	11.3.2 工作原理与控制流程	162
10.8.5 PHY 芯片 (LAN83C185)	140	11.4 本章小结	163
		附 录 常用缩略词中英文对照	164
		参考文献	166

第1章 光交换技术概述

20 世纪 60 年代,随着激光器的诞生,相继出现了许多新兴的研究领域和方向,通信中的光交换技术正是在 20 世纪 70 年代被提出的研究课题。随着光交换技术的长足发展,至今已形成种类繁多的光交换器件和不同方式的光交换机制。

目前 Internet 上存在 Tbit/s 及以上量级的语音、数据和图像信息需要传输、交换和存储,这是电子技术无法做到的。光子技术在 Tbit/s 及以上量级的传输、交换和存储上显示出电域交换无法比拟的优势,因此充分利用光子技术建立全光传输/交换网的设想应“需”而生。众多通信研究机构 and 高校致力于宽带交换技术、光交换器件和灵活高效的控制协议等技术的研究,全光试验网、示范网在全球纷纷建立,光交换技术正在快速地发展并逐步被应用和推广。

本章主要介绍光交换技术的分类、常用光交换技术的概念与比较、光交换技术的研究进展以及光突发交换技术的基本原理。

1.1 光交换技术简介

传统光纤通信网络在节点处对光纤中的信号经过光-电-光的处理后转发到下一个节点,这使得整个网络的速率主要由电域中的速率决定,造成了所谓“电子速率”瓶颈。特别当光纤速率远远超过电域速率时,这种矛盾越发明显。使得需要有一种全新的交换方式,让光信号在节点处能够直接在光域中进行传输和交换,形成一个透明的光网络。随着 IP 网络近年来呈指数型的增长,对作为其底层的光交换技术的发展要求更加迫切。

全光通信网络的关键技术主要包括光传输技术、光交换技术、光放大技术和光处理技术等几大类。其中光交换技术是全光网络系统中的一项重要支撑技术,它的发展在某种程度上也决定了光通信网络的发展。

目前光通信网络中的传输容量已经超过了 Tbit/s 量级。而电交换技术由于存在交换速率、功耗等因素,很难进行 Tbit/s 量级的信息交换。所以,光交换技术的发展是势在必行的。其中,光交换技术的优势、新型光电子器件的进步和光波技术本身在光交换中的广泛应用是推动光交换技术发展的三个主要驱动力。

1) 光交换技术具有如下技术优势:

① 光交换技术使得光信号能够高速并行地交换和传输 (WDM 技术)。

② 光信号具有极大的带宽 (大于 1Tbit/s)、抗干扰性好。

③ 光信号的比特速率和调制方式具有透明性,便于扩展新业务。

④ 光交换过程不需要光电/电光转换,避免了“电子速率瓶颈”的限制。

⑤ 光信号能够良好匹配光传输系统,信号质量好,可靠性高。

⑥ 光交换器件体积小、功耗小、便于集成、成本较低。

2) 随着各种新型的电光/声光/热光/磁光材料,半导体量子阱、波导、光纤等线性或

者非线性材料的利用,大量新型的光交换器件应运而生。这些新型的光交换器件是组成光交换系统的基本单元,并且凭借其良好的性能改善,提高了光交换系统的性能。

3) 光波技术自激光产生后被广泛地研究和发展。比如,利用光波在自由空间的反射、折射、衍射等传播特性,能够产生不同的自由空间光交换方式;利用光波的复用技术,能够产生空分、时分、波分等不同的光传输方式;利用光波的非线性效应,能够形成多种波长变换技术;而光波的高速调制、相干检测技术、滤波技术等也直接用于高速光交换网络的组件、分组或者信元的产生、报头或者信头的编码与识别等。

1.2 光交换技术分类

光交换技术是指在光域上直接将输入光信号交换到不同的输出端,而不需要任何光/电(O/E)转换。

1) 按复用方式不同,光交换技术可分为以下几种:

① 时分光交换技术:将一条复用信道划分成若干个时隙,每个基带数据光脉冲占用一个时隙,多个基带时隙复用成高速光数据流信号进行传输。时分光交换在时间轴上将复用光信号的时隙转换到另一时隙。

② 空分光交换技术:将不同信道的光信号通过改变光的传输路径来实现交换,它需要通过改变并建立两个或多个点之间的光物理通道来完成。这个通道既可以是波导,也可以是自由空间的波束。

③ 波分光交换技术:利用波分解复用器将光波分信道空间分隔开,对每个波长信道分部进行波长变换,然后再将它们复用起来,经由一条光纤传输。光信号在网络节点中不经过光/电转换,直接将信息从一个波长转换到另一个波长。

④ 码分光交换技术:根据不同的地址码来区分各路光信号,码分光交换是把一种地址码变为另一种地址码,来实现交换功能。

2) 按交换粒度划分,目前在光层上有三种光交换方式:

① 光线路交换(OCS):在光学层面,以一个波长通道上的业务流量作为最小交换单元。

② 光分组交换(OPS):以光分组包作为最小的交换单元。

③ 光突发交换(OBS):采用单向预留机制,数据包和控制包独立传输,以光突发包作为最小的交换单元。

1.3 三种主流光交换技术

1.3.1 光线路交换技术

采用OCS的光通信网络是波长路由光网络(WRON)。在WRON中,每一个连接请求都会通过端到端的光通道进行通信。并且光通道的建立方式采用了双向预留的方式,即源节点发出建立连接请求的数据分组,只有源节点在接收到来自目的节点的确认信息后才开始发送数据。每一个光通道的交换粒度是一个波长,所以WRON必须满足波长一致性条件,即同一个光纤链路上不同的光通道必须采用不同的波长。OCS交换方式的特点是适合高速率、

高带宽的业务传输，并且要求业务的生存时间相对于光通道的建立时间足够长。但是，在传输突发性较强的业务时，如 Internet 中的数据业务，OCS 由于其交换粒度粗（波长粒度）、通道建立和拆除时间长等缺点将导致较低的带宽利用率和较低的交换效率。

1.3.2 光分组交换技术

OPS 是一种典型的“存储-转发”式交换方式，并且对于连接请求采用单向预约方式。OPS 技术旨在光学层面上实现细粒度的信息交换，并且实现光网络中带宽资源的统计复用，具有较高的带宽利用率。因此，OPS 适合传输和交换类似于 IP 的突发性较强的业务。然而，在 OPS 的研究和实验过程中，需要解决的技术问题还很多。比如，光开关矩阵的交换性能、OPS 的控制协议、物理参数管理与控制、光分组的产生、同步、再生、分组头重写以及光分组之间的光功率等。这些技术问题使得 OPS 技术目前仍然处于未成熟阶段。

1.3.3 光突发交换技术

为了提高可实现性就必须尽量避免使用目前还未成熟、商用的器件，而为了提高交换性能又必须尽量减少单次通信的时延。这就需要一种新的交换方式能够像 OCS 那样提前预留资源，又能像 OPS 那样采用单向预留资源的方式，由此 OBS 应运而生。

1999 年纽约州立大学的 Chunming Qiao 和 J. S. Turnor 等人提出了光突发交换网络的概念，至今 OBS 已经发展成为光交换网络的一种主流技术。光突发交换以光突发数据包为交换单位，通过在单独的信道（一般是单独的波长）中提前发送突发控制包（BCP），以便预留相应的突发包所需的资源。后续核心节点在电域处理控制信息，突发数据包（BDP）透明地（全光）到达目的节点，途经的中间节点不需要对它作任何识别或者其他相关处理，只需要将其按预先配置的信息进行波长变换、延迟。光突发数据包是在入口边缘节点由多个具有相同特性的分组（如相同的目的节点地址或同类的服务质量要求）汇聚而成，并在出口边缘节点完成解汇聚。光突发交换网络结构如图 1-1 所示。

OBS 的基本原理可结合图 1-2，从以下三个方面进行详细说明：

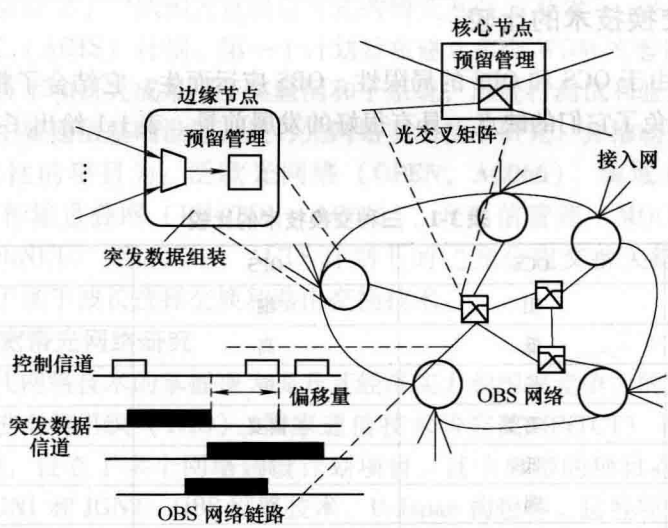


图 1-1 光突发交换网络结构

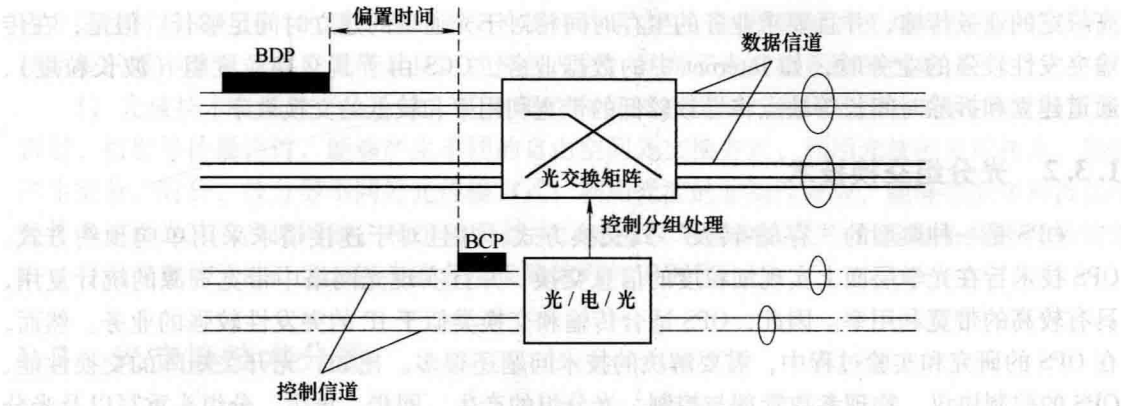


图 1-2 光突发交换原理

1) 以突发数据包 (BDP) 为基本传送单位。OBS 网络边缘节点按照一定的汇聚算法将多个输入的 IP 分组建成一个 BDP，然后再将其发送到网络中，从而增大了网络的传输和交换颗粒。这一方面保持了 OPS 网络的灵活性和高带宽利用率 (OBS 仍然采用统计复用)，另一方面又可以缓解核心节点处理速度上的瓶颈，包括消息处理速度和光开关速度。实际上，IP 分组交换通常要求光开关的速度为纳秒级，而突发数据包的交换只要求光开关的速度为微秒级。

2) BCP 和 BDP 在传送时间和信道上完全分离。在传送时间上，BCP 提前 BDP 一段时间发送，这段时间称为偏置时间 (Offset Time)。为避免 BDP 在传送过程中“超过”BCP，要求偏置时间要大于 BCP 在所经过中间节点的处理时间之和。在传送信道上，OBS 采用带外信令方式，BCP 和 BDP 利用不同的波长信道进行传送。为避免使用复杂的光逻辑器件，BCP 在核心节点进行光/电转换后在电域进行处理，为 BDP 预留网络资源。而随后到达的 BDP 以“切通”方式直接通过核心节点，不需要使用光/电转换和光存储设备。

3) 网络资源一般采用“单向预留”方式。为提高信道利用率，降低数据的端到端时延，BDP 在发送时只需等待一个偏置时间，不用等待资源预留成功的确认就可以发送。

1.3.4 三种光交换技术的比较

如上文所述，由于 OCS 和 OPS 的局限性，OBS 应运而生。它结合了粗粒度 OCS 和细粒度 OPS 的优点，避免了它们的缺点，具有很好的发展前景。表 1-1 给出了三种交换技术之间的比较。

表 1-1 三种交换技术的比较

光交换方式	OCS	OPS	OBS
粒度	粗	细	中等
带宽利用率	低	高	高
建立延迟时间	长	短	短
光缓存	不需要	需要	不需要
实现难度	低	高	中
适用性	弱	强	强
处理/同步开销	小	大	大

与 OCS 和 OPS 相比, OBS 具有如下 6 个优点:

- 1) 具有中等的交换粒度和较低的控制开销。
- 2) 支持带宽统计复用方式, 具有较高的带宽利用率。
- 3) 具有控制信道和数据信道分离传输的方式, 避免中间节点的缓存处理。
- 4) 具有单向预留机制, 有效降低连接时间和端到端时延。
- 5) 数据传输的过程是完全透明的, 中间环节不需要任何光-电-光转换。
- 6) 支持服务质量 (QoS)。

1.4 光交换技术的国内外研究进展

美国、欧洲和日本等发达国家于 20 世纪 90 年代开始确立了光网络技术及应用的战略计划和部署。而我国也通过高技术“863”计划对光网络发展相关的光器件、光系统、光传输、光交换技术等进行了研究和应用。

1. 美国的 ARPA 计划和 NGI 计划

ARPA 以政府资助、研究开发项目的形式将国家实验室、大学、研究中心、通信设备制造商和电信业务运营商联合起来, 并且制定了光网络大规模应用试验计划。ARPA 旨在验证光网络及其技术的可行性, 加快全光网络的实用化、商业化进程。该计划的第一期是实施光网络技术联合 (ONTC) 和全光网络联盟 (AON) 项目。第二期计划是建立多波长光网络 (MONET)、国家透明光网络 (NTON) 等试验网。美国国家信息基础设施建设的另一个重大计划是 NGI 计划。它包括高性能计算机和通信 (HPCC), 高带宽、高速率的全国性网络 VBNS, Internet II 以及下一代 Internet (NGI) 等建设项目。这些项目的目标是将先进技术、网络结构、控制管理, 以及网络经济等综合起来, 构造全光网络的演示、测试、试验和新技术开发平台。并且通过 NGI 计划展示全光网络的技术优势和发展前景, 为实施下一代高性能、高可靠性和经济性的全光网络打下坚实的基础。

2. 欧洲的 RACE 计划和 ACTS 计划

欧洲通信委员会设立了“欧洲先进通信方式的研究和技术开发”(RACE)计划和“先进通信技术和服 务”(ACTS)计划。第一个计划旨在建立多个 WDM 实验网, 将宽带光网络的关键器件、基础技术和研究成果用于实验网和子系统, 以进行测试和验证。第二个计划旨在利用欧洲各国的主要通信基础设施, 进行光网络实用技术研究, 并推动未来的先进通信技术研究。其中代表性的项目为: 泛欧光网络 (OPEN, AC066)、城域光网络 (METON, AC073)、泛欧电子传输重叠网 (PHOTON, AC084)、光通信管理 (MOON, AC231) 和社团光纤骨干网 (COBNET) 等。其中, ACTS 计划下的“光分组交换关键技术 (KEOPS)”项目中研究并实现了基于波长选择交换和路由交换技术。

3. 日本的高速宽带光网络研究

日本将对新一代网络技术的掌握视为提升其经济实力和 国家竞争力的重要手段。日本政府通过邮电省、先进电信组织 (TAO)、国家通信技术研究院 (NTCT) 和先进电信技术研究基金 (SCAT) 等, 设立了多个网络建设计划项目。其中典型的项目有: e-Japan 战略计划、STAR 计划、JGN1 和 JGN2、OBS 网络技术、U-Japan 构想等。这些项目逐步建立起日本的高速网络基础设施, 促进了高速宽带网络新技术的研究及国际合作研究。

4. 中国的光网络研究与发展

我国政府十分重视光网络基础设施的建设。在“八五”期间,我国以速率为 140Mbit/s 的 PDH 系统为主,建立了 20 多条省际光缆线路。在“九五”期间,我国以 622Mbit/s、2.5Gbit/s 的 SDH 系统为主,又建设了 20 多条主干光缆线路。从而将北京、上海、广州等中心城市通过光缆线路连接起来,形成了“八纵八横”的国家光纤干线网。同时,省内的长途干线网与国家干线网同步建立,光纤接入网也逐步推进。在“九五”、“十五”期间,我国继续投入大量人力、物力,通过“863”、“973”等国家级项目相继建设了“中国高速互联试验网 (NSFCNet)”、“中国高速信息示范网 (CAINONet)”、“中国网通骨干网 (CNC-Net)”、“高性能宽带信息网 (3TNet)”、“中国下一代互联网示范工程 (CNGI)”,以及“光时代计划 (O-TIME)”等重大建设项目。我国的这些建设项目旨在加快我国信息基础设施的研究和应用,形成具有自主知识产权的产品及其应用支撑环境。并且在这个过程中,培养中国光网络通信及其相关领域的技术人才,促进我国经济发展,使我国在新一代通信网核心技术方面进入世界前列。

1.5 本章小结

本章介绍了光交换技术的概念和几种典型的光交换技术,对比了三种主流光交换技术的优缺点。OBS 技术兼顾了 OCS 技术和 OPS 技术的优点,同时避免了两者的缺点,为构建下一代光网络开拓了新的思路。

第2章 网络业务量与常用网络性能指标

网络业务量模型是网络性能分析的基础,可靠的业务量模型对于网络协议的设计、网络拓扑的设计、网络性能的分析以及拥塞控制等具有重要意义。业务量模型的研究是一个分析、设计、建模、求证的循环反复的过程。为了建立正确的业务量模型,需要针对特定的网络,选择流量模型、设定模型参数、进行模型组合等。本章主要介绍网络业务量模型以及常用的网络性能指标。

2.1 泊松业务量模型

在通信网络技术发展过程中,业务量模型研究一直备受关注。20世纪70年代以来主要借鉴 PSTN 的业务量模型,用 Poisson 模型来描述数据网络的业务量模型,一般称为经典业务量模型。其数学表达如下:

定义 2-1 (泊松过程):计数过程 $\{X(t), t \geq 0\}$ 为具有参数 $\lambda > 0$ 的泊松过程,若它满足下列条件:

- 1) $X(0) = 0$ 。
- 2) $X(t)$ 是独立增量过程。
- 3) 在任一长度的 t 区间中,事件 A 发生的次数服从参数 $\lambda > 0$ 的泊松分布,即对任意 $s, t \geq 0$, 有

$$P\{X(t+s) - X(s) = n\} = e^{-\lambda t} \frac{(\lambda t)^n}{n!}, n = 0, 1, \dots \quad (2-1)$$

由于泊松过程是平稳增量过程且 $E[X(t)] = \lambda t$ 。 $\lambda = E[X(t)]/t$ 表示单位时间内事件发生的平均个数,故称 λ 为泊松过程的速率或强度。

1) 外部数据源产生流量的时间间隔为指数分布,即数据源到达为一 Poisson 过程,令 $\{D(i) | i = 1, 2, \dots, n\}$, $D(i)$ 表示数据包 i 和 $i+1$ 的间隔时间。

2) 数据源一次产生流量的长度服从指数分布,令 $\{L(i) | i = 1, 2, \dots, n\}$, $L(i)$ 表示数据包 i 的数据长度。

3) $D(i)$ 和 $L(i)$ 相互独立。

2.2 自相似业务量模型

2.2.1 网络流量的自相似特性

随着 Internet 的飞速发展,视频会议、多媒体传输等逐渐成为网络的主流业务,也使得现代网络业务呈现出越来越高的复杂性和突发性,能否揭示现代网络业务的真实性并为网络技术研究提供正确和有效的业务量模型,成为研究现代网络的关键。

20世纪70年代,由于网络业务的性能单一,人们主要基于泊松过程、指数分布等建立

模型, 这些模型是平稳的独立增量过程, 它们的共同特点是所描述的流量序列都具有短相关性。

另一种可采用的方法是具有无穷大方差的分布, 特别是采用参数 α 介于 1 和 2 之间的 Pareto 分布作为 ON/OFF 时段的模型。在这个范围内 Pareto 分布随机变量具有有限均值和无穷大的方差。可以证明: 多 Pareto 分布随机变量的叠加结果就是 Hurst 参数 $H = (3 - \alpha)/2$ 的自相似通信量。对于 $1 < \alpha < 2$, 有 $0.5 < H < 1$, 即 H 处于自相似范围之内。类似 Pareto 这样的重尾分布比较真实地反映了单个以太网信源的实际情况。直观来看, 重尾分布的较高或无限大方差表现出极大的可变化性, 即表现出在所有时间尺度上都具有可变化性。一个应用或一台工作站通常以突发的方式产生通信量, 突发之间是空闲时段。对于高方差分布而言, 则有关时间间隔的范围就可能非常宽, 既包含了许多很短的突发, 也包含了许多较长的突发和一些很长的突发。

同理万维网浏览器可以描述为一个 ON/OFF 信源并且数据刚好符合 Pareto 分布。在多个测量结果中, 发现 Pareto 分布的参数 $1.5 < \alpha < 1.6$ 。自相似行为可以这样解释, 服务器传回浏览器的报文大小, 其分布的尾部服从 Pareto 类型的分布。万维网通量反映了随机选取文件进行传输的特点。特别是如果用户是通过跟随超级链接选择要传送的文件而不考虑待传文件的大小时, 一次传送的文件大小实质上就代表了从万维网同类文件中进行的一次随机采样。因此从万维网上得到的文件大小的分布是重尾分布, 因为万维网上虽然有大量的小文件, 但也有大量的较大的以及很大的文件。其中很大文件的例子包括多媒体文件, 而这种文件正在成为万维网上的主流文件。

通过对大量的网络业务流的测量和分析发现: 网络业务在很长时间范围内都具有相关性, 即业务流具有长相关 (LRD) 特性, 而自相似 (Self-similar) 模型是最能够描述这种特性的模型之一。

2.2.2 自相似的数学定义

由于网络业务流在数学上等效为一离散时间序列, 下面先给出贝尔实验室对自相似过程的离散时间定义。

考察一个广义平稳的随机过程 $X = \{X_t : t = 0, 1, 2, \dots\}$, 其中 X_t 表示第 t 个单位时间里到达的网络业务实体的数目, 平稳时间序列的期望和方差分别为: $\mu = E[(X_t)]$, $\delta^2 = E[(X_t - \mu)]$ 。

自相关函数为: $r(k) = E[(X_t - \mu)(X_{t+k} - \mu)]/\delta^2$, 且仅与 k 有关。对每个 $m = 1, 2, 3, \dots$, 令 $X_k^{(m)} = (X_{km-m+1} + X_{km-m+2} + \dots + X_{km})/m$, $k = 1, 2, 3, \dots$, 代表长度为 m 的聚集过程, $x^{(m)} = \{X_k^{(m)}, k = 1, 2, 3, \dots\}$ 是一个根据 X 得到的 m 阶聚合序列, 记 $r^{(m)}$ 为 $X^{(m)}$ 的自相关函数。

定义 2-2 (二阶自相似): 上面的随机过程 $X(t)$, 如果对于所有的 $m = 1, 2, 3, \dots$, 都有:

$$r^{(m)}(k) = r(k) = \frac{1}{2}[(k+1)^{2H} - 2k^{2H} + (k-1)^{2H}] \rightarrow k^{-\beta}, \quad 0 < \beta < 1, \quad k \rightarrow \infty \quad (2-2)$$

其中 $H = 1 - \beta/2$ 。我们称这样的随机过程 $X(t)$ 为具有自相似参数 H 的严格自相似过程 (Exactly Second-order Self-similar), H 参数表示自相似性的强弱。对于随机过程 $X(t)$, 如果

它满足条件:

$$\lim_{m \rightarrow \infty} r^{(m)}(k) = r(k) = \frac{1}{2} [(k+1)^{2H} - 2k^{2H} + (k-1)^{2H}] \quad (2-3)$$

则称其为渐近自相似 (Asymptotically Second Order Self-similar) 的。

定义 2-3 (长程相关): 如果 X 的自相关函数 $r(k)$ 是不可求和的, 即

$$\sum_{k=-\infty}^{\infty} r(k) = \infty \quad (2-4)$$

则称 X 是长程相关的。

自相似与长程相关并不等同, 有的自相似过程不是长程相关的, 也有的长程相关不属于自相似, 不过, 对于我们感兴趣的 H 参数满足 $0.5 < H < 1$ 的自相似网络业务来讲, 它是长程相关的, 因此, 本章在不引起混淆的情况下将不加区别地使用这两个概念。

自相似程度的大小用 Hurst 参数 ($0.5 < H < 1$) 来表征, H 越大, 自相似程度越高。检测一个给定的时间序列是否具有自相似性, 只需要估计其 Hurst 参数大小即可。现在评估 Hurst 参数的方法已经有很多种, 比如方差-时间图 (V-T plot)、R/S 图、周期图、绝对矩 (Absolute Moments)、残数方差 (Variance of Residuals)、小波分析等, 其中有些可以图形化, 有些只能给出数值结果。

2.2.3 网络流量的自相似模型分析

传统的流量模型已不能很好地描述实际的网络特征, 于是人们提出用具有长相关特性的自相似模型来描述网络流量。下面就网络中广泛应用的自相似模型进行介绍。

1. ON/OFF 模型

ON/OFF 源叠加模型是对传统模型的扩展, 将自相似过程看成是无数数据源叠加的结果。网络中存在着大量的数据源, 该模型假定每个源有 ON 和 OFF 两个状态, 各个数据源相互独立且状态持续时间符合重尾分布。当数据源处于 ON 状态时, 以恒定的速率产生数据, 而处于 OFF 状态时, 不产生数据。理论上, 足够多个具有 Pareto 分布的 ON/OFF 信号源叠加, 可以产生出应用于传统以太网的通信量。

ON/OFF 模型把复杂的流量特征细化到了单个信号源, 有利于分析和解释产生自相似的原因而且其模型构造简单, 构造过程有明确的物理意义。它的缺点是假设前提太苛刻, 大多数数据源都不符合。而且由这类方法所产生序列的 Hurst 参数并不稳定。因此 ON/OFF 模型常常被用于探索网络自相似现象产生的根源, 研究模型各个参数和控制策略对网络设计的影响。图 2-1 为 ON/OFF 源叠加模型。

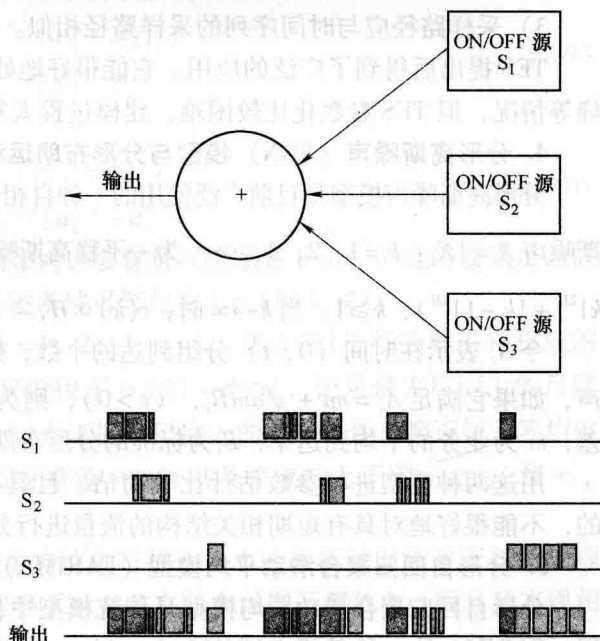


图 2-1 ON/OFF 源叠加模型

2. 确定性模型

除了用 ON/OFF 数据源叠加来生成相关业务流的方法外,有不少学者提出用基于混沌映射的确定性模型来分析网络流量的自相似特性。一维映射状态变量为

$$x(n+1) = \begin{cases} f_1(x_n), 0 \leq x_n < d \\ f_2(x_n), d \leq x_n \leq 1 \end{cases} \quad (2-5)$$

分组到达过程由与之相关联的变量 y_n 表示,即

$$y_n = \begin{cases} 0, 0 \leq x_n < d \\ 1, d \leq x_n \leq 1 \end{cases} \quad (2-6)$$

式(2-5)中第一部分为线性段, ON/OFF 时间呈几何分布, $P(L > 1) \propto \alpha^{-1}$ (L 为队长);第二部分为非线性段, $f(x) \approx \varepsilon + x + cx^\alpha$, 其中 ε 为很小的常数。ON/OFF 期间停留时间呈 Pareto 分布,即

$$P(L > l) \propto \frac{1}{l^{m-1}}, (l \rightarrow \infty), H = \frac{3m-4}{2m-2}$$

混沌映射易于进行网络排队分析,便于分析系统的性能,但是它难以给出不变密度的解析表达式,而且它所产生的业务流的自相似特性的时间尺度和阈值 d 有关。

3. TES 模型

TES 建模方法提出的最初动机是为了捕获业务流的突发性。TES 过程是一个具有任意边缘分布以及在很宽范围内的自相关函数的自相关序列。它能够产生各种各样的采样路径(包括循环和任意方向)和具有各种形式(如单调,振荡)的自相关函数。

假设有一些静态时间序列,建立一个 TES 模型需同时达到三个目标:

- 1) 模型的边缘分布与时间序列的边缘分布相匹配。
- 2) 模型的自相关函数与时间序列的自相关函数相匹配。
- 3) 采样路径应与时间序列的采样路径相似。

TES 提出后得到了广泛的应用。它能很好地处理网络业务流的非指数下降或者周期性尖峰等情况,但 TES 参数化比较困难,建模过程太复杂,建模过程中需要人工干预等。

4. 分形高斯噪声 (FGN) 模型与分形布朗运动 (FBM) 模型

分形高斯噪声模型是目前广泛使用的一种自相似模型,它是严格二阶自相似过程。分形高斯噪声 $X = \{X_k: k=1, 2, 3, \dots\}$ 为一平稳高斯噪声,其自相关函数 $r(k) = \frac{1}{2}(|k+1|^{2H} - 2|k|^{2H} + |k-1|^{2H})$, $k \geq 1$ 。当 $k \rightarrow \infty$ 时, $r(k) \propto H$, X 与 $X^{(m)}$ 具有相同的分布。

令 A_t 表示在时间 $(0, t)$ 分组到达的个数,则 A_t 是一种具有定长增量的自相似高斯噪声,如果它满足 $A_t = mt + \sqrt{am}B_t$, ($t > 0$), 则为分形布朗业务。其中 $\alpha > 0$ 为网络流量方差, m 为业务的平均到达率, B_t 为标准的分形布朗运动生成的随机变量。

用这两种模型进行参数估计比较简洁,且具有较好的数学基础。但它们是严格自相似的,不能很好地对具有短期相关结构的流量进行分析。

5. 分形自回归聚合滑动平均模型 (FARIMA)

分形自回归聚合滑动平均模型是传统模型中自回归聚合滑动平均模型 (ARIMA) 的扩展。FARIMA 是一种基于 FARIMA (p, d, q) 过程的业务生成方法。它的定义为 $\phi(B)\Delta^d X_k = \theta(B)\varepsilon_k$ 。其中 $\phi(B)$ 为自回归项 (AR), $\theta(B)$ 为滑动平均项, B 为定义的一个延