

支持 QoS 的交换技术研究

李秀芹 著



中国水利水电出版社
www.waterpub.com.cn

支持 QoS 的交换技术研究

李秀芹 著



中国水利水电出版社

www.waterpub.com.cn

内 容 提 要

本书从交换结构和调度算法两方面介绍了交换技术的 QoS 保障机制和性能评价模型。主要内容包括: 绪论、支持 QoS 的交换结构及调度算法分析、基于标识支持区分 QoS 的 CICQ 调度机制研究、基于标识支持区分 QoS 的 PPS 解决方案 PSVIOQ-CICQ、基于标识支持区分 QoS 的 PPS 解决方案 PSCICQ、支持 QoS 的多级交换结构研究、基于效用函数支持 QoS 的交换结构性能评价模型、MPLS 协议在新一代网络交换路由应用中的研究。

本书可供计算机、网络工程、通信工程等相关专业的科研人员、工程技术人员和高等学校师生参考。

图书在版编目 (C I P) 数据

支持QoS的交换技术研究 / 李秀芹著. — 北京: 中国水利水电出版社, 2014. 9
ISBN 978-7-5170-2388-3

I. ①支… II. ①李… III. ①通信交换 IV.
①TN91

中国版本图书馆CIP数据核字(2014)第199677号

策划编辑: 向 辉 责任编辑: 宋俊娥 加工编辑: 孙 丹 封面设计: 李 佳

书 名	支持 QoS 的交换技术研究
作 者	李秀芹 著
出版发行	中国水利水电出版社 (北京市海淀区玉渊潭南路 1 号 D 座 100038) 网址: www.waterpub.com.cn E-mail: mchannel@263.net (万水) sales@waterpub.com.cn
经 售	电话: (010) 68367658 (发行部)、82562819 (万水) 北京科水图书销售中心 (零售) 电话: (010) 88383994、63202643、68545874 全国各地新华书店和相关出版物销售网点
排 版	北京万水电子信息有限公司
印 刷	三河市天润建兴印务有限公司
规 格	170mm×227mm 16 开本 13.75 印张 210 千字
版 次	2015 年 1 月第 1 版 2015 年 1 月第 1 次印刷
定 价	48.00 元

凡购买我社图书, 如有缺页、倒页、脱页的, 本社发行部负责调换
版权所有·侵权必究



作者简介

李秀芹，女，汉族，1967年生，河南召陵人。工学博士，华北水利水电大学教授，河南省教育厅“555”学术技术带头人。长期从事计算机教学与科研工作。主要研究领域：新一代网络交换技术、资源分配、软件工程等，先后在国内外权威期刊公开发表科研论文42篇，作为主编先后出版了《计算机网络基础教程》、《计算机网络与Internet应用》、《新编ASP.NET 2.0网站开发从入门到精通》等著作7本；主持、参与国家973计划项目、国家自然科学基金、水利部、省级重大科研项目15项；获省科技进步奖、省自然科学优秀论文等科技成果奖励17项。目前主持省科技攻关计划项目1项，参与水利部公益性专项科研项目1项。

前 言

随着信息网络多媒体业务的兴起,路由交换设备交换结构对服务质量支持的问题再次受到关注,如何实现具有 QoS 保证的交换技术也得到了学术界和产业界的广泛重视。随着不同应用的增加,将来的交换结构必须能够考虑到对更加丰富的 QoS 保证的支持。现有信息网络的原始设计思想基本上是“一种网络支撑一种主要服务”的解耦模式,在此基础上的演进与发展难以突破原始设计思想的局限,无法满足网络及服务的多样性需求。

“一体化网络”将多种网络设计成一种网络。通过三次解析映射,将标识理论贯穿了从服务层到网通层的整个网络体系架构,可以把不同业务类型通过映射关系反映到网通层。课题研究依托国家 973 计划信息技术领域子课题《一体化网络体系结构模型及交换路由理论与技术》,为了更好地应对交换技术所面临的严峻挑战,对一体化网络体系下新型的交换机理、原理与关键技术展开研究,研究内容具有重要的理论意义和实用价值。

针对路由交换设备对于交换技术的 QoS 需求,本文首次将标识的概念引入交换结构,提出了基于标识的交换技术,给出了适合一体化网络基于标识支持 QoS 的交换技术应具有的特征;基于 CICQ 交换结构设计出支持区分服务的 DS-CICQ 交换结构,在不同类业务和不同输出端口,采用基于标识支持区分 QoS 的分布式动态双轮询 ID-DDRR 调度算法,调度复杂度大大降低;在对主流 PPS 交换技术现状分析的基础上,针对并行交换中的保序和区分 QoS 问题,提出一种新型的 PSVIOQ-CICQ 解决方案;设计出一种新型的 PPS 体系结构 PSCICQ,基于 PSCICQ 体系结构,提出一种基于标识支持区分 QoS 的 PPS 调度机制,证明了 PSCICQ 体系结构能够保证业务信元的传输顺序;在汇聚模块设置少量缓存,采用双指针轮询算法 DPRR 实现区分 QoS 保障,保证了交换对高层不同业务类的有效支持;在分析研究三级 Clos 网交换结构和调度算法的基础上,提出了一种支持 QoS 的三

级 Clos 分布式交换结构, 并基于该结构提出了 DHIRRM 调度算法, 使该交换结构的设计能提供优良的 QoS 策略。针对交换结构性能评价问题, 提出基于效用函数支持 QoS 的交换结构性能评价模型; 最后, 将一体化网络的重要思想引入到 MPLS 网络, 研究了 MPLS 协议在新一代网络交换路由的应用, 提出了基于一体化 MPLS 网络架构的协调数据流抢占机制。

基于标识和带缓存交叉开关交换单元研究了交换技术的 QoS 保障问题, 研究内容属于宽带信息网络学科前沿, 研究成果避免交换结构及实现机制方面缺乏 QoS 保证的缺陷, 能够满足不同业务对 QoS 的需求, 实现网络一体化并为用户提供普适服务, 有望对新一代信息网络特别是交换技术的发展与应用起到促进作用。

本书将作者近几年在一体化网络下支持 QoS 交换技术方面研究中取得的成果进行归纳、总结。内容主要取材于作者在各种期刊和国际会议上公开发表的论文, 同时融入国内外学者在该领域取得的优秀研究成果, 力求全面、系统地概括国内外最新研究成果, 反映发展动态。但难免会存在一些不足之处, 欢迎专家及同行多提宝贵意见。

作者

2014 年 8 月

目 录

前言	
第一章 绪论	1
1.1 引言	1
1.1.1 研究背景	1
1.1.2 研究目的意义	5
1.2 一体化网络研究	7
1.2.1 一体化网络体系结构	7
1.2.2 一体化网络中的交换与路由	12
1.3 课题研究思路与研究工作	16
1.4 本章小结	18
参考文献	20
第二章 支持 QoS 的交换结构及调度算法分析	22
2.1 引言	22
2.2 典型交换结构分析	23
2.2.1 分组交换 Crossbar 型交换结构	23
2.2.2 交叉点无缓存 Crossbar 型交换结构	24
2.2.3 交叉点带缓存 Crossbar 型交换结构	27
2.3 支持 QoS 的调度算法分析	29
2.3.1 基于 OQ 结构的典型调度算法分析	30
2.3.2 基于 IQ 结构的典型调度算法分析	31
2.3.3 基于 CIOQ 结构的典型调度算法分析	32
2.3.4 基于 CICQ 结构的典型调度算法分析	34
2.4 基于标识支持 QoS 的交换技术分析	36

2.5 本章小结	38
参考文献	38
第三章 基于标识支持区分 QoS 的 CICQ 调度机制研究	44
3.1 引言	44
3.2 相关术语及定义	45
3.2.1 区分 QoS	45
3.2.2 分组信息头格式	52
3.3 DS-CICQ 交换结构	53
3.4 ID-DDRR 调度算法	55
3.4.1 带宽分配方案	55
3.4.2 两个调度阶段	57
3.4.3 关于份额函数	59
3.4.4 ID-DDRR 算法描述	59
3.5 性能评估	62
3.5.1 算法理论分析	62
3.5.2 算法公平性证明	63
3.5.3 算法仿真实验分析	64
3.6 本章小结	68
参考文献	69
第四章 基于标识支持区分 QoS 的 PPS 解决方案 PSVIOQ-CICQ	71
4.1 引言	71
4.2 并行交换现状分析	72
4.3 PSVIOQ-CICQ 交换结构	77
4.4 调度算法	79
4.4.1 相关概念和约定	79
4.4.2 负载均衡算法	80
4.4.3 DDRRP 信元重组算法	82
4.5 仿真结果与分析	85

4.5.1	仿真环境	85
4.5.2	仿真结果分析	87
4.6	本章小结	92
	参考文献	93
第五章	基于标识支持区分 QoS 的 PPS 解决方案 PSCICQ	96
5.1	引言	96
5.2	PSCICQ 交换结构	97
5.3	调度算法	98
5.3.1	相关定义	98
5.3.2	解复用器负载分担算法	99
5.3.3	中心交换平面调度机制	101
5.3.4	DPRR 汇聚模块调度机制	104
5.4	性能分析	107
5.5	仿真结果与分析	110
5.6	PSCICQ 和 PSVIOQ-CICQ 的比较	116
5.7	本章小结	120
	参考文献	121
第六章	支持 QoS 的多级交换结构研究	122
6.1	引言	122
6.2	支持 QoS 的三级 Clos 分布式交换结构	123
6.2.1	输入端口的设计	124
6.2.2	输出端口的设计	130
6.3	交换结构的特性分析	131
6.3.1	分布式队列调度	131
6.3.2	一次调度	132
6.3.3	负载均衡	133
6.4	一种基于时延优先级可预测匹配的调度机制——DHRRM	134
6.4.1	任务模型	134

6.4.2	DHIRRM 算法描述	135
6.4.3	预测匹配	136
6.5	本章小结	137
	参考文献	138
第七章	基于效用函数支持 QoS 的交换结构性能评价模型	139
7.1	引言	139
7.2	业务流的效用函数现状分析	140
7.3	业务流的时延效用函数	141
7.3.1	语音流的时延效用函数	141
7.3.2	视频流的时延效用函数	142
7.3.3	TCP 流的时延效用函数	143
7.3.4	三种业务流的时延效用函数	144
7.4	基于时延的效用函数评价模型	145
7.4.1	评价等级定义	145
7.4.2	评价模型构建	146
7.4.3	评价模型分析	146
7.4.4	评价模型的扩展	147
7.4.5	仿真实验分析	151
7.5	双指标效用函数评价模型	153
7.5.1	双指标模型的构建	153
7.5.2	双指标评价模型分析	154
7.6	本章小结	156
	参考文献	157
第八章	MPLS 协议在新一代网络交换路由应用中的研究	158
8.1	引言	158
8.2	基于协调数据流抢占机制总体架构	160
8.2.1	工作流程	160
8.2.2	添加边缘路由表	161

8.2.3 MPLS 通知报文格式	163
8.3 发送吞吐量信息封包的设计	164
8.3.1 标志变量	165
8.3.2 回传吞吐量信息封包	165
8.3.3 抢占通知封包	166
8.3.4 启动抢占机制条件	167
8.4 TPN 传输等级定义	168
8.5 TPN 机制工作原理	170
8.5.1 TPN 工作流程	170
8.5.2 TPN 带宽保留算法	172
8.6 LSR 工作原理	174
8.6.1 Ingress LSR 工作模块	175
8.6.2 Core LSR 工作模块	177
8.6.3 Egress LSR 工作模块	178
8.7 仿真实验与结果分析	179
8.7.1 仿真软件	179
8.7.2 具有 TPN 功能的 NS 扩展	181
8.7.3 仿真与分析	183
8.8 本章小结	193
参考文献	194
附录 A 并行交换系统仿真平台 PPS-CICQ-SIM 的设计	196
A.1 引言	196
A.2 并行交换系统仿真平台 PPS-CICQ-SIM 建模	198
A.2.1 系统中的主要类及其相互关系	199
A.2.2 系统对象模型	204
A.2.3 系统动态模型	207
参考文献	209

第一章 绪论

本章介绍了课题的研究背景和面临的挑战，讨论了课题研究的目的和意义，分析了一体化网络中的交换与路由，最后给出了所做的主要工作。

1.1 引言

1.1.1 研究背景

随着科学技术的发展，信息已成为当今推动社会向前发展的巨大动力。信息网络在各国经济与社会发展中起着重要作用。随着网络的传输技术、交换技术和应用技术的蓬勃发展，互联网正成为全球范围的信息资源、存储资源以及计算资源共享的信息基础平台，并成为影响世界经济发展模式、推动生产力进步的巨大动力。路由交换设备则是这一全球信息平台的基础设施，因此它一直是网络发展和演进中的重中之重，路由交换设备的技术水平在很大程度上决定了网络的发展水平。而路由交换设备中所采用的交换技术直接影响着吞吐率、数据包时延以及交换容量等关键性能指标，因此交换技术一直是学术界研究的重点和热点问题，很多重要学术成果已经在实际路由交换设备中得到了应用。新一代网络正朝着传输高速化、接入宽带化和应用多样化方向发展，对交换技术提出了新的挑战。具体来说，现有的交换技术面临的挑战主要表现在以下两个方面：

首先，网络规模的扩大和链路传输速率的提高，要求交换设备能够支持高速率大容量交换。根据中国互联网络信息中心（CNNIC）发布的《第 27 次中国互联网络发展状况统计报告》^[1]，中国总网民人数和宽带网民数与比例增长趋势图如图 1-1 所示。截至 2010 年 12 月 30 日，中国网民规模达到 4.57 亿人，网民规模较 2009 年年底增长 7330 万人，年增长率为 19.1%，宽带网民规模达到 4.5 亿人，

较 2009 年增长 1.03 亿人，年增长率为 30%。



图 1-1 中国网民规模与普及率

无论是网民的增加还是网络规模的扩大，都导致网络所承载的信息量不断提高，这就要求交换设备能够提供大容量交换能力。在网络规模扩大的过程中，连接交换节点和交换节点之间的链路速率也在不断提高。WDM 和 DWDM 等光纤传输技术近年来取得了突破性进展；在一根光纤中复用大量的传输通道，而每个通道可以工作于 OC-48 (2.5Gb/s)、OC-192 (10Gb/s) 甚至 OC-768 (40Gb/s)；目前 DWDM 最多能够复用 128 种颜色^[2]，在单个通道工作于 OC-768 (40Gb/s) 的情况下，一根光纤最高能够支持 5Tb/s 的传输速率。虽然硅晶体技术已经取得了快速的进步，但光传输数据率和电子交换设备所能处理的数据率之间的差距仍然在不断扩大，而且在可以预期的将来，传统 CMOS 技术的物理极限将会出现。在这种情况下，交换节点就成为制约网络数据传输的一个瓶颈，造成链路传输带宽的浪费。纯光交换可以很好地解决这个问题，但是光存储等文件技术问题在短期内还无法解决，这就导致在很长一段时间内，网络的基础数据交换功能依然需要依赖于已有的光电混合交换设备。因此，基于已有的交换技术，在高速光纤链路环境下支持高速率数据交换，仍然是目前亟待研究和解决的问题。

并行分组交换^[3]是一种提高交换速率和交换容量的有效手段。许多典型的交换系统在设计时都使用了这一方法，并行分布式处理机制对提升路由交换设备的

交换容量十分有益,克服了电子器件工艺水平的瓶颈限制,弥补了单平面交换结构的先天不足。但是即便采用并行分布式处理机制构建大容量交换结构,单平面交换结构的研制也依然十分关键。单平面交换结构的性能很大程度上决定了并行分组交换结构的性能,而且过多交换平面的并行处理机制不仅浪费硬件资源,同时导致交换结构的控制管理和流量分配机制过于复杂,甚至无法实现。因此研究具有良好可扩展性的交换技术是研制能够支持更高的端口速率和提供更大交换容量的路由交换设备的客观需要。

其次,网络应用日益丰富多样化,要求交换设备由原先的“尽力而为”的服务方式转变为能够提供服务质量保障的交换。随着三网合一进程的演进以及 Everything over IP 思想的提出,互联网需要并正在承载各种不同的业务,这就要求原先的为单纯数据传输提供“尽力而为”服务的 IP 网络演变为可以为综合业务(文本、语音、图像、流媒体等)提供具有 QoS 保障服务的高性能网络。伴随实时多媒体应用和技术的飞速发展,网络上多媒体信息和实时任务的数量与日俱增。根据 CNNIC 发布的第 27 次中国互联网络发展状况统计报告,表 1-1 给出了我国网民各类主要网络应用的使用情况。截至 2010 年 12 月,网络应用使用率排名前三甲的仍然是搜索引擎(81.9%)、网络音乐(79.2%)和网络新闻(77.2%),搜索引擎使用率首次超过了网络音乐,成为我国网民规模最庞大的应用。但从发展速度上看,商务类应用用户规模继续领涨。网络购物用户年增长 48.6%,是用户增长最快的应用,网上支付和网上银行全年增长也分别达到了 45.8%和 48.2%。此外,微博和团购的用户数已初具规模。在网络应用日益丰富和多样化的过程中,人们逐渐意识到传统的网络交换性能指标并不能保障各种应用获得最佳的用户体验。在这种情况下,就需要为不同种类的应用提供不同种类的性能,即服务质量保证。例如语音聊天应用需要较低的时延性能,而文件传输类应用则需要较低的丢包率,甚至不允许丢包。交换结构和调度算法作为交换设备的两个主要部分,直接影响了网络的吞吐率、时延等性能。在网络发生拥塞和资源争用的情况下,交换结构和调度算法的设计对语音和视频业务的质量起到决定性的作用。因此,研究具有 QoS 保障能力的交换技术是实现具有 QoS 保障的高性能网络的必然要求,也成为了新一代网络发展的必然要求。

表 1-1 各类网络应用使用状况及用户增长

类型	应用	2010 年 使用率	2009 年 使用率	用户 增长率	使用 率排名	增长 率排名
信息获取	搜索引擎	81.9% ↑	73.3%	33.1%	1	5
网络娱乐	网络音乐	79.2% ↓	83.5%	12.9%	2	16
信息获取	网络新闻	77.2% ↓	80.1%	14.7%	3	14
交流沟通	即时通信	77.1% ↑	70.9%	29.5%	4	7
网络娱乐	网络游戏	66.5% ↓	68.9%	15.0%	5	13
交流沟通	博客应用	64.4% ↑	57.7%	33.0%	6	6
网络娱乐	网络视频	62.1% ↓	62.6%	18.1%	7	12
交流沟通	电子邮件	54.6% ↓	56.8%	14.6%	8	15
交流沟通	社交网站	51.4% ↑	45.8%	33.7%	9	4
网络娱乐	网络文学	42.6% ↑	42.3%	19.8%	10	10
商务交易	网络购物	35.1% ↑	28.1%	48.6%	11	1
交流沟通	论坛/BBS	32.4% ↑	30.5%	26.6%	12	8
商务交易	网上银行	30.5% ↑	24.5%	48.2%	13	2
商务交易	网上支付	30.0% ↑	24.5%	45.9%	14	3
商务交易	网络炒股	15.5% ↑	14.8%	24.8%	15	9
交流沟通	微博客	13.8%	—	—	16	—
商务交易	旅行预订	7.9% →	7.9%	19.5%	17	11
商务交易	团购	4.10%	—	—	18	—

因此,随着我国信息化进程的不断发 展,互联网所承载的业务无论是规模还是类别都在迅速增长,发生了深刻的变化,特别是多媒体、商务交易等业务在互联网上的应用快速膨胀,使得互联网在数据传输方面面临两大挑战性问题,要求交换设备:提供大的交换容量和对多种网络业务提供良好的 QoS 保障。而随着并行交换及多级交换等技术的发展,吞吐量最大化的重要性要低于对 QoS 等的保证。尤其是站在服务提供商 ISP 的角度,吞吐量的最大化并不意味着收益的最大化,而在用户的角度,吞吐量的最大化也不能保证自己关键性业务的 QoS,目前国内外的交换技术均朝着提高吞吐量的同时有效地保证 QoS 的方向发展^[4]。

1.1.2 研究目的意义

“十一五”期间和国家中长期发展规划都将新一代信息网络关键技术与服务作为优先发展领域。

现有信息网络的原始设计思想基本上是一种网络支撑一种主要服务,在此基础上的演进与发展难以突破原始设计思想的局限,无法满足网络及服务的多样性需求。例如电信网当初是面向语音业务传输设计的,它能够提供对称话务质量,但是其以电路交换为基础的通信机制决定了其网络效率低下,同时,电信网带宽受限,导致其难以适应宽带流媒体业务等的需要。互联网当初是面向数据业务传输设计的,遵从 TCP/IP 的四层体系结构(包括子网层、网络层、传输层、应用层),采用面向无连接的分组交换技术传输数据,并提供“尽力而为”的服务。

现有信息网络由于原创模式的局限,存在着诸多难以解决的问题,而当前关于新一代信息网络的研究还没有形成完整的体系,缺乏基础理论的原创性创新,因此迫切需要突破原有网络的局限,设计全新的网络体系结构,创建出兼有各家之长又能适应长远应用需求的一体化网络,解决现有信息网络在服务扩展、可信性(安全性、可靠性、可控性、可管性)以及移动性等方面存在的问题。

国内外为解决这些问题,很多机构正在研究下一代网络^[5-9],如 NewArch 计划及美国国家科学基金委员会(NSF)的 GENI、FIND 等项目都投入了大量精力开展新一代信息网络基础理论研究。NewArch 计划提出了一种抽象的网络体系结构模型 FARA^[10]仍然沿用了现有互联网技术,仅在应用层进行了功能性验证;GENI 和 FIND 项目还处于起步阶段,虽然基本确定了研究方向^[5]并取得了少量初步成果^[11],但还缺乏实质性的进展,至今尚未形成一个较为清晰和成熟的思路。文献[12-15]则分别从业务拓展、通信方式和质量控制等几个方向对新一代信息网络提出一些初步的设计目标。“一体化可信网络与普适服务体系基础研究”是由北京交通大学主持的国家 973 重大科研项目^[16-18]。该科研项目提出了一体化可信网络与普适服务的概念,力图在一种网络上支持多种服务,并解决可信、移动、传感网络接入等问题。

一体化网络的基本思想是将网络划分为接入层和核心层,然后在此基础上提

出分离映射机制，原理就是接入层使用接入标识，核心层使用交换路由标识。这样就实现了接入标识用来表示终端的身份信息，不再携带地址信息，交换路由标识仅仅用来在核心网中的路由数据包使用，不包含终端的其他信息。这种设计方案很好地解决了网络可信性、移动性等方面的问题。

973 课题中提出了广义交换路由理论、标识分离映射机制，并以此为核心创建了一体化网络模型与理论，以解决多种网络一体化问题。从业务层面上讲，可以承载各种不同类型的业务，即普适服务，包括互联网业务、电信网业务以及未来可能出现的新型信息业务。

通过三次解析映射（从服务标识到连接标识的解析映射，从连接标识到接入标识的解析映射、从接入标识到交换路由标识的解析映射），将标识理论贯穿从服务层到网通层的整个网络体系架构；而 MPLS 只是一个 2.5 层交换技术，MPLS 标识只能用于 2.5 层数据包交换。由此可见，我们可以把不同业务类型通过映射关系反映到网通层，从而为我们研究基于一体化网络支持多业务类 QoS 的交换技术提供了有力的依据。

根据 973 项目总任务书的要求，子课题 2 主要研究支持普适服务的一体化网络体系结构模型及交换路由理论与技术，主要包括以下两个方面：

- 支持普适服务的一体化网络体系结构机理与模型；
- 一体化网络中的交换与路由理论（主要包括骨干路由器的核心结构、如何支持普适服务 QoS 保证、动态路由协议）。

当前，Internet 是由一些数目相对较少的高速骨干网络连接很多小网络组成的。骨干网络的链路速率基本上是以 30% 的增长速率递增。这表明，传输线路已经不是解决网络拥塞的瓶颈。所以，设计高性能的核心路由器是提高 Internet 整体性能的关键所在^[19-21]。这就对交换结构和调度算法的配合及路由处理能力提出了更高的要求。随着不同应用的增加，将来的交换结构也必须能够考虑到对更加丰富的 QoS 保证的支持。因此，受 973 子课题“一体化网络体系结构模型及交换路由理论与技术”的支持，本文将对这些问题展开研究。

“一体化网络体系结构模型及交换路由理论与技术”（No.2007CB307102）课题的主要研究内容之一就是在一体化信息网络体系结构下，研究新型交换机理、