

Video Traffic Analysis  
and QoS Management

# 视频流量分析 与QoS管理

黄天云 著



电子科技大学出版社

# 视频流量分析与 QoS 管理

## Video Traffic Analysis and QoS Management

黄天云 著

电子科技大学出版社

## 图书在版编目 (CIP) 数据

视频流量分析与 QoS 管理 / 黄天云著. —成都: 电子科技大学出版社, 2013.1

ISBN 978-7-5647-1422-2

I. ①视… II. ①黄… III. ①数据通信—视频信号  
IV. ①TN919

中国版本图书馆 CIP 数据核字 (2013) 第 001822 号

## 内容简介

如何对编码视频数据流进行有效传输并保证其服务质量 QoS, 是近年来宽带业务研究的热点。本书提供了对视频流量进行模拟和性能分析的一般化方法, 以及保证 IP 上视频流传输服务质量的关键技术。本书研究了常见的视频流量模型、典型的视频流量管制和整形算法, 以及速率-失真理论在流视频传输中的应用。在此基础上分析了 MPEG-2 的分层编码、MPEG-4 的基于对象体系结构, 以及 MPEG-4 精细粒度可扩展编码 FGS 的优化传输方案和调度策略。

本书可作为计算机、通信、电子、应用数学等专业的研究生教材, 也可作为相关领域的大学生和科技人员的参考书。

## 视频流量分析与 QoS 管理

黄天云 著

出版: 电子科技大学出版社 (成都市一环路东一段 159 号电子信息产业大厦 邮编: 610051)

策划编辑: 张 鹏

责任编辑: 张 鹏 李燕苓

主 页: [www.uestcp.com.cn](http://www.uestcp.com.cn)

电子邮箱: [uestcp@uestcp.com.cn](mailto:uestcp@uestcp.com.cn)

发 行: 新华书店经销

印 刷: 河北永清县晔盛亚胶印有限公司

成品尺寸: 170mm×240mm 印张 14.5

字数 292 千字

版 次: 2013 年 3 月第一版

印 次: 2013 年 3 月第一次印刷

书 号: ISBN 978-7-5647-1422-2

定 价: 45.00 元



■ 版权所有 侵权必究 ■

◆ 本社发行部电话: 028-83202463; 本社邮购电话: 028-83201495。

◆ 本书如有缺页、破损、装订错误, 请寄回印刷厂调换。

# 前 言

流媒体技术的发展得益于 3 个方面的因素：网络技术、视频压缩编码标准、应用的实际需求。近 10 年来，随着应用的实际需求和终端业务的不断发展，引入新的媒体服务，如视频（流媒体）传输，这使得网络流量管理和服务质量 QoS（Quality of Service）变得越来越重要。为视频传输提供应用所需的 QoS 是一项具有挑战性的工作。

然而，传统的 Internet 并不能有效地提供并满足用户的需求，因为 Internet 是基于 IP 的尽力而为（Best Effort）服务，它无法提供多媒体通信所需的服务质量参数如带宽、时延和时延变化、差错和丢失率等。当前，众多的研究者致力于从多个方面来解决该问题，如：增加网络的传输带宽，研究新的网络传输和资源预留协议，构造新的 Internet 通信模型和体系结构，以及在有限的网络资源情况下发展可扩展的自适应媒体编码和表征机制。针对 IP 网络已经提出了一些业务模型，如综合业务模型 InteServ、区分业务模型 DiffServ、多协议标识交换 MPLS 等。在 ATM 网络中，更将业务质量 QoS 作为首要考虑的问题力图加以解决。

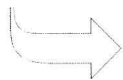
正因为意识到视频传输对终端用户和实际应用的重要性，它已经成为当今国际学术界的一个研究热点。在 IEEE 和国内的众多期刊上常见到这方面的研究成果发表。

在这个过程中，视频压缩编码标准和网络技术的进展起了至关重要的作用。视频编码的目的是以更低的码率压缩、存储和传输视频数据。从 1992 的 MPEG-1 和 1994 年的 MPEG-2 开始，视频压缩编码技术已经有了长足的进步，MPEG-2 的分级可扩展、MPEG-4 的面向对象编码和精细粒度可扩展 FGS，以及 H.264/AVC 的高级编码技术等，已经成为国际标准并得到广泛应用。我国也已建立相应的视频编码标准草案 AVS。同时，网络技术已经从传统的 IP 和 ATM 发展到 P2P 对等网络，并进一步扩展到无线环境，如无线 IP 和移动 Ad Hoc 网络。特别是移动 Ad Hoc 网络在最近几年内得到广泛的关注和研究。各种移动终端设备如 GPRS、CDMA 的推广和普及，也使得适合这些网络承载的流媒体传输如多媒体通信和视频会话业务的需求与日俱增。

因此，本书的出版目的是总结近年来 IP 视频传输应用的主要研究成果，讨论视频传输所需的关键技术和当前研究状况，这些研究反映了当前视频传输应用研究的主要进展及存在的问题。

本书第一章首先给出 QoS 参数定义，对 IP QoS 的几种主要体系结构进行对比。

第二章讨论视频数据模型和编码表征，对图像和视频内容的表示、基于内容的索引和检索、MPEG 系列视频数据的编码表征和位流语法描述等进行概述。



第三章介绍网络分析中用到的主要业务流量模型，如 Poisson 过程、Markov 调制过程、TES 过程等；同时，对近年的热点研究——自相似网络流量模型也进行详细的分析，介绍多种不同的自相似流量模型，讨论其参数估计问题；对随机网络流量的模型选择问题也进行分析。

第四章讨论视频流量模型和性能分析，给出基于 Markov 链和场景迁移的 VBR 视频流量的完整建模过程；同时对聚合视频流量的  $\alpha$  平稳自相似模型以及基于 Markov 到达过程 MAP 的分层视频流量模型进行分析。

第五章分析 IP 和 ATM 网络常见的流量管制和流量整形算法；根据令牌桶和漏桶算法，实施一种在 IntServ 模型的可保证服务 GS 下，基于源端的 VBR 视频流令牌桶整形方案，以保证视频流传输的延迟和延迟变化在可控制的 QoS 参数范围内；同时讨论常见的开环控制和闭环流量控制算法，建立 VBR 视频流基于预测控制的流量控制算法及自适应的流量传输方案。

第六章讨论速率-失真 (Rate Distortion) 理论在流视频传输中的应用。首先分析基于发送端的 MPEG 流视频传输的 R-D 模型，这也是最早提出的 R-D 流传输方案；并推广到基于接收端流传输和自适应回放。利用逼近理论，讨论可扩展编码如 MPEG-2 分级可扩展、MPEG-4 FGS 和 PFGS 的 R-D 优化模型；同时给出两种低复杂度的 R-D 优化视频流传输方案。最后讨论基于 R-D 和有效带宽模型的 MPEG 视频的优化包调度算法，对 SRD、LRD、FARIMA 和多分形流量模型下算法的不同性能指标进行对比分析。

第七章分析 MPEG-2 的分层编码和 MPEG-4 的基于对象体系结构；讨论视频对象 VO 的优先级分配算法和传输数据单元的调度序。分析服务数据单元 SDUs 向 DiffServ PHB 的映射机制，以及视频对象 VO 在 DiffServ 上的传输体系结构；同时通过对 ATM 适配层 AAL5 的修改，建立了带优先级的 ALL-5-AAL-5p，以保证基于对象分层编码视频在 ATM 网络的有效传输。

考虑到 MPEG-4 的精细粒度可扩展编码 FGS 在 IP 流视频传输研究中的广泛应用，本书第八章详细分析 MPEG-4 FGS 流视频在 IP 的优化传输方案和调度算法。首先给出一般化情形下的 MPEG-4 FGS 包分配和调度方案；其次基于 PSNR (MSE) 建立 MPEG-4 FGS 流视频的相关客观质量和质量变化指标定义，提出视频分段场景基础上的质量最优化问题；通过状态迁移图和动态规划解决 FGS 流视频的最优传输策略问题。同时将 MPEG-4 FGS 的位率分配方案推广到更细粒度，如 GoV 或单帧的精细位率控制，对比不同分配方式的优缺点。最后，建立基于错误消除 EC 的 FGS 流视频 R-D 矩阵，分析基于 Markov 决策过程 MDP 的 FGS 流视频的优化传输模型及其求解。

本书每一章的参考文献在文中相应位置标注并在章末给出。并且，从第三章开始包含文献评述，不仅对该章内容的引用处进行注释，也对未阐述的内容和当前研究重点进行概括和推广。这也便于读者参考其他文献和进行扩展研究。本书没有专

门讨论 P2P 对等网络和无线环境如无线 IP、移动 Ad Hoc 网络的流视频传输研究，这方面的内容作者将另行著书阐述。

本书是作者在博士学习期间以及后续研究工作的基础上，参考国内外相关研究成果（包括技术报告）概括和总结而成。本书的出版得到国家民委科研项目基金（05XN09）和西南民族大学博士创新基金（234138）的资助。感谢电子科技大学计算机学院 803 教研室全体同仁以及西南民族大学计算机科学与技术学院的领导对作者的研究工作提供的大力支持和帮助。

由于作者水平有限，加之时间仓促，对文献的理解可能不够透彻，书中的不足之处在所难免，敬请广大专家和读者批评指正。

作 者

# 目 录

|                                     |    |
|-------------------------------------|----|
| 第一章 QoS 参数定义和网络控制 .....             | 1  |
| 1.1 QoS 参数定义和网络映射 .....             | 1  |
| 1.1.1 QoS 参数定义的参考模型 .....           | 1  |
| 1.1.2 网络特性向 QoS 参数的映射 .....         | 2  |
| 1.2 IP 和 ATM 网络的 QoS 参数定义 .....     | 2  |
| 1.2.1 ATM 网络的 QoS 参数定义 .....        | 2  |
| 1.2.2 IP 网络的 QoS 参数定义 .....         | 3  |
| 1.2.3 IP 和 ATM 网络的 QoS 参数规范对比 ..... | 4  |
| 1.3 IP QoS 的几种主要体系结构对比 .....        | 5  |
| 1.3.1 综合业务模型 InteServ .....         | 5  |
| 1.3.2 区分业务模型 DiffServ .....         | 8  |
| 1.3.3 多协议标识交换 MPLS .....            | 9  |
| 1.3.4 小结 .....                      | 11 |
| 参考文献 .....                          | 11 |
| 第二章 视频数据模型与编码表征 .....               | 13 |
| 2.1 图像和视频内容的表示 .....                | 13 |
| 2.1.1 图像的分析与分类 .....                | 13 |
| 2.1.2 视频数据的表示与索引 .....              | 14 |
| 2.1.3 运动分析与视频分段 .....               | 16 |
| 2.1.4 视频数据的组成模型 .....               | 17 |
| 2.2 基于内容的索引和检索 .....                | 18 |
| 2.2.1 基于内容的检索与相似匹配 .....            | 18 |
| 2.2.2 查询公式化和表现 .....                | 19 |
| 2.2.3 查询优化 .....                    | 20 |
| 2.2.4 查询结果及检索的重构 .....              | 20 |
| 2.3 视频数据的编码表征 .....                 | 21 |
| 2.3.1 MPEG 系列标准概述 .....             | 21 |
| 2.3.2 MPEG 视频数据的编码表征 .....          | 22 |
| 2.3.3 编码视频数据的位流语法描述 .....           | 35 |



|                                                     |    |
|-----------------------------------------------------|----|
| 2.3.4 小结                                            | 37 |
| 参考文献                                                | 38 |
| 第三章 网络业务流量分析模型                                      | 41 |
| 3.1 传统网络业务流量模型                                      | 41 |
| 3.1.1 引言                                            | 41 |
| 3.1.2 网络业务流量模型                                      | 41 |
| 3.2 自相似网络流量模型                                       | 49 |
| 3.2.1 随机游动和布朗运动 (Random Walk & Brownian Motion)     | 49 |
| 3.2.2 二阶自相似过程 (Second-Order Self-Similar Processes) | 51 |
| 3.2.3 自相似现象的随机模型                                    | 53 |
| 3.2.4 随机过程模型的选择——自相似或非自相似                           | 55 |
| 3.3 文献评述                                            | 56 |
| 参考文献                                                | 56 |
| 第四章 视频流量模型与性能分析                                     | 59 |
| 4.1 MPEG VBR 视频流——统计特性及其模型                          | 59 |
| 4.1.1 引言                                            | 59 |
| 4.1.2 主要的工作                                         | 60 |
| 4.1.3 视频分段                                          | 60 |
| 4.1.4 场景聚类                                          | 61 |
| 4.1.5 场景切换——四状态 Markov 调制链                          | 63 |
| 4.1.6 场景建模                                          | 63 |
| 4.1.7 小结                                            | 66 |
| 4.2 聚合 VBR 视频流的 Alpha 平稳自相似随机模型分析                   | 66 |
| 4.2.1 引言                                            | 66 |
| 4.2.2 定义和参数估计                                       | 67 |
| 4.2.3 Alpha 自相似随机变量的产生                              | 71 |
| 4.2.4 聚合 VBR 视频流模型分析                                | 74 |
| 4.2.5 试验仿真和定量比较                                     | 75 |
| 4.2.6 小结                                            | 77 |
| 附录 4.1 多聚合视频流收敛性分析                                  | 77 |
| 附录 4.2 $\alpha$ -平稳自相似输入排队系统的分析                     | 79 |
| 4.3 基于 Markov 到达过程 MAP 的分层视频流量模型                    | 80 |
| 4.3.1 Markov 到达过程 MAP                               | 80 |
| 4.3.2 分层视频的 MAP 流量模型                                | 81 |



|                                        |     |
|----------------------------------------|-----|
| 4.3.3 试验对比和分析                          | 83  |
| 4.3.4 小结                               | 86  |
| 4.4 文献评述                               | 87  |
| 参考文献                                   | 88  |
| 第五章 IP 和 ATM 网络的视频流量控制算法               | 93  |
| 5.1 流量管制和流量整形算法                        | 93  |
| 5.1.1 令牌桶和漏桶流量控制算法                     | 93  |
| 5.1.2 业务量参数控制——管制和整形                   | 98  |
| 5.1.3 MPEG VBR 视频流流量整形                 | 103 |
| 5.1.4 小结                               | 112 |
| 5.2 MPEG 视频传输的流量控制算法                   | 113 |
| 5.2.1 开环流量控制算法                         | 113 |
| 5.2.2 闭环流量控制算法                         | 117 |
| 5.2.3 基于预测控制的 VBR 流视频传输控制              | 120 |
| 5.2.4 小结                               | 127 |
| 5.3 文献评述                               | 127 |
| 参考文献                                   | 128 |
| 第六章 R-D 优化模型与 MPEG 流视频传输               | 132 |
| 6.1 基于发送端 (Sender-Driven) 的 R-D 优化模型   | 132 |
| 6.1.1 数据单元敏感度调整和 R-D 优化                | 133 |
| 6.1.2 数据单元优化传输策略                       | 134 |
| 6.1.3 几种不同方案的 R-D 性能比较                 | 135 |
| 6.2 基于接收端 (Receiver-Driven) 的 R-D 优化模型 | 136 |
| 6.2.1 固定回放速率的 R-D 优化                   | 136 |
| 6.2.2 自适应回放 AMP 和优化流传输                 | 137 |
| 6.2.3 试验和对比分析                          | 138 |
| 6.2.4 小结                               | 139 |
| 6.3 分级可扩展编码视频的 R-D 优化模型                | 140 |
| 6.3.1 R-D 曲线及其逼近模型                     | 140 |
| 6.3.2 可扩展编码视频的 R-D 模型                  | 141 |
| 6.3.3 与传统 R-D 模型和 Mallat 模型的试验对比       | 143 |
| 6.4 两种低复杂度的 R-D 优化流传输方案                | 144 |
| 6.4.1 信道特征描述                           | 145 |
| 6.4.2 两种 LC RaDio 优化流方案                | 145 |



|       |                                             |     |
|-------|---------------------------------------------|-----|
| 6.4.3 | 算法复杂度分析                                     | 146 |
| 6.4.4 | 试验结果分析                                      | 147 |
| 6.5   | MPEG 流传输的 R-D 优化和有效带宽调度                     | 148 |
| 6.5.1 | 视频序列 GoP 的 R-D 分析                           | 148 |
| 6.5.2 | 基于有效带宽的优化包调度算法                              | 150 |
| 6.5.3 | 试验分析和性能评价                                   | 155 |
| 6.5.4 | 小结                                          | 156 |
| 6.6   | 文献评述                                        | 157 |
|       | 参考文献                                        | 157 |
| 第七章   | 基于对象分级编码视频及多优先级传输                           | 160 |
| 7.1   | 分层可扩展编码与 MPEG-4 对象描述                        | 161 |
| 7.2   | MPEG-4 视频对象 VO 的优先级分配算法                     | 163 |
| 7.3   | ATM 的视频对象 VO 多优先级实时传输                       | 166 |
| 7.3.1 | 对 ATM 适配层 AAL-5 的修改——带优先级的 AAL-5 (AAL-5p)   | 168 |
| 7.3.2 | 多优先级缓冲访问控制算法 MP-BAC (Buffer Access Control) | 169 |
| 7.3.3 | 实验仿真                                        | 170 |
| 7.3.4 | 小结                                          | 174 |
| 7.4   | IP DiffServ 的视频对象 VO 多优先级实时传输               | 174 |
| 7.4.1 | 优先级向 DiffServ PHB 的映射                       | 175 |
| 7.4.2 | 实验仿真                                        | 176 |
| 7.4.3 | 小结                                          | 177 |
| 7.5   | 文献评述                                        | 178 |
|       | 参考文献                                        | 178 |
| 第八章   | MPEG-4 FGS 视频传输质量定义及调度算法                    | 180 |
| 8.1   | 一般化情形下的包分配与调度方案                             | 180 |
| 8.1.1 | MPEG-4 FGS 的优化包分配和调度算法                      | 181 |
| 8.1.2 | 用户定义缓冲时间与视频表示质量的平衡                          | 185 |
| 8.1.3 | 小结                                          | 186 |
| 8.2   | 一类 FGS 流视频的质量定义与基于场景的优化调度算法                 | 187 |
| 8.2.1 | 引言                                          | 187 |
| 8.2.2 | 定义及定理                                       | 188 |
| 8.2.3 | 问题求解及算法                                     | 191 |
| 8.2.4 | 试验分析                                        | 194 |

|                                                  |     |
|--------------------------------------------------|-----|
| 8.2.5 小结 .....                                   | 198 |
| 8.3 推广到更细粒度如 GoV 和单帧的精细位率控制 .....                | 198 |
| 8.3.1 TFRC 流量和 GoV 模式下的位率分配和调整方案 .....           | 199 |
| 8.3.2 几种方案的试验仿真和对比分析 .....                       | 201 |
| 8.3.3 小结 .....                                   | 203 |
| 8.4 MPEG-4 FGS 的 R-D 矩阵与 MDP 最优决策模型 .....        | 204 |
| 8.4.1 假设及问题描述 .....                              | 205 |
| 8.4.2 失真度 R-D 矩阵 .....                           | 205 |
| 8.4.3 单约束条件 MDP $\{X_n, A_n\}$ 优化模型及其求解 .....    | 207 |
| 8.4.4 推广到多约束条件 MDP $\{X_{n-1}, X_n, A_n\}$ ..... | 209 |
| 8.4.5 部分可观察 POMDP $\{A_{n-1}, A_n\}$ 及其求解 .....  | 210 |
| 8.4.6 小结 .....                                   | 210 |
| 8.5 文献评述 .....                                   | 211 |
| 参考文献 .....                                       | 211 |
| 附录 缩写词汇表 .....                                   | 214 |

# 第一章 QoS 参数定义和网络控制

近年来,随着终端业务的不断发展,引入了多种新的媒体服务,这使得网络流量管理和服务质量 QoS (Quality of Service) 变得越来越重要。因为对包括文本、图像、音频、视频在内的综合媒体服务,要提供这些业务所需的 QoS 是一项具有挑战性的工作。IP 和 ATM 都从不同角度对流量管理和服务质量进行研究,力图提供相应的解决方案以迎接新的用户业务的挑战。本章首先讨论 QoS 参数的一般定义和参考模型,在此基础上分析 IP 和 ATM 网络是如何提供不同业务的 QoS 保证的。目的是为以后章节研究 MPEG 视频流在 IP 和 ATM 网络的传输作一个铺垫。相关的参考文献和文档,读者可以从 IETF 的 RFC 或 ATM 的正式文档得到。

## 1.1 QoS 参数定义和网络映射

### 1.1.1 QoS 参数定义的参考模型

标准从终端用户的角度提出了基于端到端的 QoS 参数定义 (End to End QoS)。这是从用户应用的交互性而言的。它可能包含多个网络,每个网络可能包含多个节点。而每个网络因交换、复用、传输等都会引入时延、丢失、差错,这些均会影响到端到端的 QoS。网络节点的输入/输出链路的拥塞、缓冲争用也会导致丢失和时延等,如图 1.1 所示。

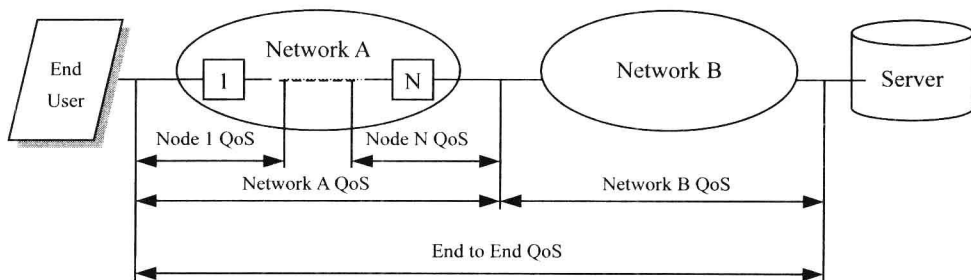


图 1.1 QoS 参数定义参考模型

因此,不论是网络还是节点,都可能需要实施一定程度的流量整形方案来最小化网络延迟变化或数据丢失。对终端用户而言,他们所能看到的只是端到端的 QoS,但也可能需要在入口级上进行流量管制,以保证输入的流量与预先协商的 QoS 参数

相一致。

### 1.1.2 网络特性向 QoS 参数的映射

表 1.1 列出 IP 和 ATM 网络中相应的网络特性对 QoS 参数的影响<sup>[1]</sup>。

可以看出，受 IP 路由器和 ATM 交换机影响最大的 QoS 参数可归结为：因缓冲溢出而导致的时延（Delay）、时延变化（Delay Variation）、丢失率（Loss）。而传输的数据包或信元的位差错和突发差错（Bit and burst errors）可能会导致随机差错（Random Errors）或信元误插（Misinsertion），从而引起差错倍增。数据包或信元穿越的网络节点越多，越可能造成 QoS 质量的下降（Degeneration）。所有的网络特性都会或多或少地影响到最终的端到端 QoS，即终端表示质量 QoP（Quality of Presentation）。

表 1.1 网络特性向 QoS 参数的映射

| Characteristic                       | Delay | Delay Variation | Loss | Random Errors | Misinsertion |
|--------------------------------------|-------|-----------------|------|---------------|--------------|
| Propagation Delay                    | ✓     |                 |      |               |              |
| Switch/router queuing architecture   | ✓     | ✓               | ✓    |               |              |
| Switch/router link rate              | ✓     | ✓               |      |               |              |
| Packet Size                          | ✓     | ✓               | ✓    |               |              |
| Switch/router buffer capacity        | ✓     | ✓               | ✓    |               |              |
| Switch/router resource allocation    | ✓     | ✓               | ✓    |               |              |
| Variations in traffic load           | ✓     | ✓               | ✓    |               | ✓            |
| Switch/router and link failures      |       |                 | ✓    |               |              |
| Bit and burst errors                 |       |                 | ✓    | ✓             | ✓            |
| Number of switches/routers traversed | ✓     | ✓               | ✓    | ✓             | ✓            |

## 1.2 IP 和 ATM 网络的 QoS 参数定义

### 1.2.1 ATM 网络的 QoS 参数定义

ATM 论坛 UNI 4.0 流量管理规范定义如何在用户/网络接口 UNI 上协商特定的 QoS 性能参数。表 1.2 是 ATM 标准定义在 ATM 层（ATM Layer）的部分 QoS 参数。表 1.2 最后一列表明用户是否可以与网络就相应的 QoS 参数进行协商。不能协商的部分在网络交换节点中隐式（Implicit）地实施。读者可以参考 ITU-T 建议 I.356<sup>[2]</sup>、I.610<sup>[3]</sup>以获得 QoS 参数更详细的定义。

表 1.2 ATM 网络的 QoS 参数定义

| QoS Acronym      | QoS Parameter Name              | Negotiated? |
|------------------|---------------------------------|-------------|
| Peak-to-peak CDV | Cell Delay Variation            | Yes         |
| Max CTD          | Maximum CTD                     | Yes         |
| CLR              | Cell Lose Ratio                 | Yes         |
| CER              | Cell Error Ratio                | No          |
| SECBR            | Severely Erred Cell Block Ratio | No          |
| CMR              | Cell Misinsertion Rate          | No          |

ATM 允许应用以两种不同的方式来协商 QoS 参数。首先, PVC 或 SVC 应用可能需要某种服务类型(如: CBR, rt-VBR, nrt-VBR, UBR 或 ABR 等)。另一种则在 UNI 4.0 中定义, 它允许应用通过信令(Signaling)传送来显式(Explicit)地定义表 1.1 中规定的某个或几个 QoS 参数。表 1.3 是 I.356 对三种不同的 QoS 等级的定义。

QoS 等级 1 是为满足具有严格要求的 CBR 业务需求而定义的; QoS 等级 2 的应用不对信元丢失优先级 CLP 进行区分; QoS 等级 3 的应用则对 CLP=0(高优先级)的信元提供可保证的服务 GS (Guaranteed Service), 对 CLP=1(低优先级)的信元则不作任何保证, 除非网络还有可用资源以传输优先级为 1 的信元。这特别适合于 MPEG-2 中的二级 VBR 编码技术。后面将根据此思想建立 MPEG 视频的面向对象的分层视频编码及多优先级传输方案。

表 1.3 I.356 的 QoS 分级及性能指标

| QoS Parameter | Notes              | QoS Class 1        | QoS Class 2        | QoS Class 3        |
|---------------|--------------------|--------------------|--------------------|--------------------|
| CTD           | Mean value         | 400 ms             | Unspecified        | Unspecified        |
| CDV           | $10^{-8}$ quantile | 3ms                | Unspecified        | Unspecified        |
| CLR(0+1)      | Applies to CLP=0+1 | $3 \times 10^{-7}$ | $10^{-5}$          | Unspecified        |
| CLR(0)        | Applies to CLP=0   | N/A                | Unspecified        | $10^{-5}$          |
| CER           | Upper Bound        | $4 \times 10^{-6}$ | $4 \times 10^{-6}$ | $4 \times 10^{-6}$ |
| CMR           | Upper Bound        | Once per day       | Once per day       | Once per day       |
| SECBR         | Upper Bound        | $10^{-4}$          | $10^{-4}$          | $10^{-4}$          |

### 1.2.2 IP 网络的 QoS 参数定义

IETF RFC 2216 定义了 IP 网络的 QoS 参数规范。表 1.4 列出在 IETF 的 InteServ<sup>[4]</sup>模型下的两种不同的服务: 可控负载 CL (Controlled Load) 和可保证服务 GS (Guaranteed Service) 下对应的 QoS 参数。这也在 RFC 2221、RFC2222 中进行总结。

表 1.4 IP 的 InteServ 模型下 QoS 参数规范和定义

| QoS Parameter           | Controlled Load                                                              | Guaranteed QoS                                         |
|-------------------------|------------------------------------------------------------------------------|--------------------------------------------------------|
| Maximum Delay Variation | Not specified                                                                | Automatically measured<br>Using Adspec                 |
| Packet loss rate        | Little or no congestion loss,<br>Should be error-limited                     | Sufficient buffer required<br>for zero congestion loss |
| Minimum delay           | High percentage of packets<br>Do not exceed minimum delay                    | Not specified                                          |
| Average delay           | Little or no queuing delay over<br>Time frames greater than round-trip delay | Not specified                                          |

与 ATM 的预分配方式相比, IP 的 QoS 参数定义没有参考配置。在 IP 中, 应用无法动态地改变已经与网络协商好的 QoS 参数值, 除非重新进行协商。当然, 也可以实施相应的流量控制算法, 如闭环控制策略来进行反馈, 以获得可接受的 QoS 性能。可保证服务 GS 定义了端到端的最大允许时延、时延变化和缓冲需求, 因此所有网络节点都可以保证数据传输过程中零排队丢失率。实时应用可以利用 GS 的最大时延变化以保证可进行连续回放 (Playback), 避免造成接收缓冲的上溢和下溢。RFC 2005~2216 已经定义了 RSVP<sup>[5]</sup>, 作为 IP 网络中资源预留和保证应用所需 QoS 的协议。

### 1.2.3 IP 和 ATM 网络的 QoS 参数规范对比

ATM 是基于源端初始化连接 (硬状态——Hard State), 因此可为应用提供所需的 QoS 和带宽保证, 无法接入的应用在准入阶段就已被拒绝。IP 通过 RSVP 进行资源预留, RSVP 是基于接收端进行初始化的 (软状态——Soft State), 网络中的交换节点或路由器不能都保证预留连接所需的所有资源, 因而无法可靠保证用户应用的连接需求。相比而言, ATM 在提供音频/视频或对网络性能敏感的业务上更显优势。

RSVP 要求接收端周期性地刷新资源预留请求, 而且资源预留和路由是独立进行的, 不能保证预留资源都恰好被利用; ATM 则是在连接建立阶段同步进行的, 因此可保证更有效地利用预留资源。

面向连接的 ATM 网络不能很好地支持已有的绝大多数 IP 应用, 然而可以通过在主干网 (Backbone) 上实施 ATM 以获得更好的连接管理和资源预留。两者都存在一定的可扩展性问题, 因为资源预留是针对每条连接进行的, 必须要采取某种形式的聚合以获得可扩展性, 如表 1.3 中的 QoS 分级或 ATM 中的 VP (虚通路) 聚合。

ATM 从 1995 年起就致力于为不同类型的业务提供 QoS 保证, IP 的 RSVP 则直

到 1999 年才正式成文。ATM 标准曾经提及采取相应的 QoS 协商和资源预留机制,如基于 ATM 承载的 IP 数据传输 (IP over ATM) 等。然而,随着近年来媒体业务传输的研究更多地集中到 IP 上,对 ATM 的 QoS 研究已经逐渐淡化。毕竟,实际应用已证明是成功并且日渐完善的 IP 技术,比采用虽然有更严格的 QoS 规范,但却更复杂、更难于实施的 ATM 网络,有更大的优势。

### 1.3 IP QoS 的几种主要体系结构对比

鉴于本书主要集中在对 IP 流媒体传输的研究,因此以下对 IETF 提出的几种主要的 QoS 体系结构进行对比分析;并且后面将讨论 InteServ 的可保证服务 GS 下的 VBR 视频流传输和 DiffServ 下的 MPEG-4 视频对象 VO 的优先级传输等。

可以从不同角度理解 IP QoS。ISO、ITU、IETF 分别给出不同的定义,其内涵不尽相同。由于 IP 技术最早应用于互联网,所以 IETF 的关于 IP QoS 的定义(即 IP QoS 是指 IP 层的 QoS)在 IP 技术研究领域得到了广泛的认可,其性能参数包括丢包率、吞吐量、传输时延、时延变化等。与 IETF 关于 IP QoS 的定义相比,ITU-T 的 IP QoS 不单指网络层 QoS,还包括应用层 QoS,它从网络运营的角度来定义 IP QoS,认为 IP QoS 存在于各个网络运营实体之间。而 ISO 的定义则更广泛,基于 7 层协议模型,ISO 认为各个网络层次之间、对等层次之间都存在服务和被服务关系,因此就存在服务质量 QoS 的概念。所以,QoS 不只局限于网络层和应用层,它存在于网络的各个协议层次之间 (ISO/IEC JTC1-SC21)。更进一步的研究则可以从用户、ISP 和网络营运商的角度进行分析。目前,从网络分析的角度,主要采用 IETF 定义的量化性能参数。为此,IETF 建议了 3 种不同的服务模型:综合业务模型 InteServ<sup>[4-7]</sup>、区分业务模型 DiffServ<sup>[8-10]</sup>、多协议标识交换 MPLS<sup>[11-14]</sup>。

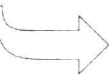
#### 1.3.1 综合业务模型 InteServ

InteServ<sup>[4]</sup>强调为保证会话的 QoS,所有的资源都必须预先保留。简化的 InteServ 通信模型如下:应用向系统提交 QoS 请求;系统决定是否有足够的资源以满足应用的请求。如果满足,则接受并为其分配相应资源;否则,拒绝请求或建议一个较低的能够提供的 QoS 服务(此时,如果接受,则该应用在较低的 QoS 级别上执行)。可见,为保证 QoS,系统必须提供以下几方面的机制:QoS 规范、准入控制、QoS 协商和再协商、资源分配和调度、流量管制。

##### 1.3.1.1 QoS 规范

QoS 规范规定应用与系统在建立连接之前就相关参数如带宽、时延、时延变化等进行协商的方式。对突发流量,可以定义均值速率、峰值速率、突发持续时间、突发容量等。规范要求会在话建立之前业务量参数必须在所有的子系统中协商一致,





并且这些参数可以在不同的系统层次上进行映射。因此, QoS 规范通常也被称为业务量描述 (Traffic Characterization)。业务量描述定义网络可以发送的流量大小, 它决定了是否支持某种业务的网络流量, 网络可以据此对流量进行管制。好的业务量规范能够描述不同类型的流量, 并能有效实施准入控制和流量管制。业务量规范包括: 多参数描述和业务量整形 (Traffic Shaping)。

### 1. 多参数描述 (规范)

研究表明, 通过选择适当的参数, 如: 峰值速率、均值速率、位速率方差等, 可以有效地对源流量进行描述。然而, 很难用有限参数对输出排队的多个流量进行描述或性能分析, 特别是多个具有较大突发性的 VBR 业务流量。随机过程, 如 Poisson 过程或 Markov 链曾被用于描述声音、视频或 LAN 的业务量统计特性, 近年来亦提出了很多新的流量模型。其优点在于: 如果每个交换都可以被描述成独立的、无记忆的排队系统, 则输出流量也可以同样地描述, 系统性能就可以很容易地被精确度量。然而, 因为只使用有限个状态来描述一个源业务流量, 对包含多种类型的聚合流量, 其准确性会受到很大影响 (聚合流可能表现出与单流完全不同的流量特征)。

### 2. 流量整形方案 (Shaping)

流量整形是对源流量进行规范和约束, 以便对流量进行描述和管制。当业务量模式过于复杂而难以直接进行描述或源流量不适合承载网络直接传输时, 往往要进行流量整形。如: 突发性较大的 VBR 视频流, 网络的交换缓冲往往没有这么大的空间来处理其业务流量。为解决该问题, 流量整形被用来在流量被送到网络之前, 先被整形成某特定的模式。一个好的流量整形方案可以对多种流量进行整形, 并且其输出能很好地被描述和管制。为满足应用的 QoS, 交换/路由也必须对流量进行重新整形 (Reshape)。典型的流量整形方案有: 漏桶、令牌桶等。目前已提出很多 VBR 视频流整形方案。

#### 1.3.1.2 准入控制、QoS 协商和再协商

当具有某特定 QoS 参数的连接建立时, QoS 参数必须在所有的子系统被协商; 只有当所有的子系统都同意并保证特定的 QoS 参数时, 端到端的 QoS 才能够被满足。QoS 协商包括: (1) 参数映射——QoS 参数被从一个层次 (或子系统) 映射到另一个层次 (或子系统); (2) 资源预留——每个层次 (或子系统) 必须决定是否支持请求的参数; 如果可以, 特定的资源为该会话保留, 只有当所有的子系统都接受请求的参数时, 该会话才建立。如果存在某层次或子系统不能满足请求的参数, 则该会话被拒绝。一个较复杂的系统还可以向用户建议它可以支持的参数, 而由用户决定是否建立相应会话。

而对动态变化的会话环境, 有可能需要更改 QoS 参数。因此, 需要提供 QoS 再协商机制以满足在多媒体通信中用户请求的动态改变。在网络拥塞的时候, 用户对资源增加的请求或许并不能满足。这时对某些特定应用如视频会议、实时监控等,