

JIYU FENZU

WANGLUO DE

FUWU ZHILIANG BAOZHENG

张继军 高鹏 编著

基于分组网络的 服务质量保证



04



北京邮电大学出版社
www.buptpress.com

基于分组网络的服务质量保证

张继军 高 鹏 编著

北京邮电大学出版社
·北京·

内 容 简 介

下一代网络是基于分组技术的网络,而能否有效保证分组网络承载多类型业务时的服务质量,则是通信网络向下一代网络演进的关键。本书从多个方面对分组网络的服务质量(QoS, Quality of Service)保证问题进行了系统地论述。首先,介绍了 QoS 保证的一般性问题,包括基本概念、技术发展、标准化组织的研究工作、应用性能需求、性能参数、网络性能和 QoS 分类等,并介绍了一个通用的 QoS 体系结构框架。其次,分析了两种基于分组的网络(IP 网络和 ATM 网络)的 QoS 保证问题,并对其他一些网络技术(如以太网、RPR 等)在 QoS 保证方面的情况进行了介绍。第三,介绍了 QoS 路由和相关的流量工程方法。第四,介绍了 IP 网络 QoS 策略控制方法和移动网络的 QoS 保证问题。最后,分析了 3 个实现分组网络端到端 QoS 保证的例子。

本书系统地收集了目前国际上有关分组网络 QoS 保证的最新科研成果和研究资料,并总结了作者多年来从事通信网络研究、系统设备研究与总体规划工作的经验。

本书选材新颖、内容详尽、系统性强,可作为理工院校通信、电子、计算机等专业高年级学生的专业教材,也可以作为从事电信工作的专业技术人员和管理人员的培训教材或自学参考书。

图书在版编目(CIP)数据

基于分组网络的服务质量保证 / 张继军, 高鹏编著. —北京: 北京邮电大学出版社, 2004

ISBN 7-5635-0859-7

I. 基... II. ①张... ②高... III. 计算机通信网—通信协议—服务质量—质量管理
IV. TN915.04

中国版本图书馆 CIP 数据核字 (2004) 第 021697 号

书 名: 基于分组网络的服务质量保证

编 著: 张继军 高 鹏

责任编辑: 方 瑜

出 版 者: 北京邮电大学出版社(北京市海淀区西土城路 10 号) 邮编: 100876

发行部电话: (010) 62282185 62283578(传真)

经 销: 各地新华书店

印 刷: 北京源海印刷有限责任公司

开 本: 787 mm × 1 092 mm 1/16

印 张: 21

字 数: 494 千字

印 数: 1—3 000 册

版 次: 2004 年 5 月第 1 版 2004 年 5 月第 1 次印刷

ISBN 7-5635-0859-7/TN·322

定价: 36.00 元

如有印装质量问题, 请与北京邮电大学出版社发行部联系

序

多年来,从三网融合到下一代网络,人们对在电信网中构建统一的基于分组技术的承载网络寄予了很高的期望。然而,实际情况却不尽如人意。虽然分组网络以其固有的带宽分配灵活、带宽利用率高等特点在通信网络中扮演了日益重要的角色,但其 QoS 保证能力不足的问题始终是一个严重的瓶颈,制约了分组网络(特别是 IP 网络)在电信运营领域内发挥更大的作用。与此同时,随着电信技术的进步和电信市场竞争的日趋激烈,大量的业务类型丰富的应用开始需要由分组网络承载。不同的应用有不同的服务质量需求,这无疑进一步增加了实现分组网络 QoS 保证的难度。正因如此,众多国际标准化组织、国内外研究机构都对如何实现分组网络承载业务的 QoS 保证投入了极大的热情,并取得了可喜的阶段性的成果。

一般来说,产生 QoS 问题的原因是业务流量对网络资源的需求和网络实际能够提供资源之间的不匹配。当业务对网络资源的需求超过了网络实际供给能力时,业务的服务质量将不能得到有效保证。为了实现分组网络端到端的 QoS 保证,至少需要解决以下问题:

- 实现网络资源的最大利用;
- 有效支持在网络业务端到端范围内高度分布的智能和多种应用类型;
- 有效支持网络业务端到端范围内的多个管理域的协同工作;
- 实现在网络业务端到端范围内多种网络技术和 QoS 控制机制的并存和互操作。

时至今日,有关 QoS 保证技术问题的研究仍处于不断发展、完善之中,完全成熟尚有待时日。但是,在新技术层出不穷的今天,对这一领域内已经取得的研究和应用成果进行及时总结是非常必要的。《基于分组网络的服务质量保证》一书,系统地介绍和分析了目前国际上有关分组网络 QoS 保证的最新研究和应用成果,无疑将对人们更深入地了解通信网络的服务质量问题,积极寻求和部署更好的解决方案提供十分有益的帮助。

赵梓森

中国工程院院士
中国光谷(武汉)首席科学家

2004 年 3 月 6 日

前 言

从 20 世纪 80 年代至今,分组网络以其固有的带宽分配灵活、带宽利用率高等特点在通信网络中扮演了日益重要的角色。随着分组网络(如 Internet)的超常规发展,网络承载的业务类型逐渐呈现出多样化的发展趋势。不同的业务有不同的服务质量要求。例如,远程医疗强调业务提交的准确性,但对业务端到端的时延和分组的时延变化(如抖动)要求不高;与之相对应的是,IP 电话要求有尽可能小的业务抖动和时延性能。下一代网络是基于分组技术的网络,而能否有效保证分组网络承载多种类型应用时的业务服务质量,是通信网络向下一代网络演进的关键。

服务质量是服务性能的综合体现,这种服务性能决定了网络在多大程度上满足业务用户的要求。一般来说,产生 QoS 问题的原因是业务流量对网络资源的需求和网络实际能够提供资源之间的不匹配。当业务对网络资源的需求超过了网络实际供给能力时,业务的服务质量将不能得到有效保证。此时,我们需要根据业务类型及其对服务质量要求的不同,采用适当的 QoS 控制机制,合理分配网络资源。当前,众多国际标准化组织、研究机构都对如何实现分组网络承载业务的 QoS 保证投入了极大的热情,并已经定义、规范和应用了多种类型的 QoS 控制机制,如基于服务优先级的接入控制、拥塞反馈、流量的计量和测量、QoS 选路和流量工程、QoS 策略及策略配置、排队和调度、资源预留(如 RSVP)、服务等级管理、服务费率表征和流量标识等;同时,作为当前通信领域的研究热点,各项新的 QoS 保证机制的提出及对已有技术的改进也层出不穷。熟悉并在网络建设中适当地应用这些技术,对于提高分组网络的 QoS 保证能力是十分有益的。

本书分 11 章。第 1 章对 QoS 相关问题作了概要性的介绍,包括基本概念、技术发展,以及从事 QoS 研究工作的主要标准化组织及其活动。第 2 章自顶向下地介绍了站在端用户角度对多媒体应用的性能需求和 QoS 分类、QoS 参数,以及 ATM 网络和 IP 网络的网络性能参数、指标和 QoS 分类等问题。第 3 章介绍了目前 ITU-T 正在讨论规范的一种由多个对应于不同网络控制机制的构件模块组成的基于分组网络的 QoS 体系结构框架。第 4 章和第 5 章分别介绍了 IP 网络、ATM 网络,及其他一些网络技术的 QoS 保证问题。第 6、7、8 章围绕 QoS 路由和相关的流量工程方法进行介绍。第 9 章分析了 IP 网络的 QoS 策略控制方法。第 10 章介绍了移动网络的 QoS 保证问题。第 11 章以举例的方式介绍了 3 个实现端到端 QoS 保证的例子。

在本书的编写过程中,得到了作者的博士导师赵梓森院士的指导和帮助,范志强同志也参加了部分章节的编写,在此表示衷心的感谢。

由于作者水平有限,加之书中涉及的部分 QoS 保证技术仍处于不断发展、完善的过程中,书中某些观点和论述难免有谬误之处,恳请广大读者批评指正。

作 者

2004 年 4 月于武汉

目 录

第 1 章 QoS 概述

1.1 QoS 的基本概念	1
1.1.1 QoS 的定义	1
1.1.2 通用的 QoS 模型	2
1.1.3 描述通信服务质量的框架和 QoS 的 4 个视角	3
1.1.4 SLA、服务生命周期和 QoS	6
1.2 QoS 技术的发展	7
1.2.1 网络容量的超额供应	7
1.2.2 QoS 保证的分层实现	8
1.2.3 基于 ATM 网络的 QoS 保证机制	10
1.2.4 基于 IP 网络的 QoS 保证机制	10
1.2.5 基于策略的 QoS 控制方法	13
1.2.6 移动网络的 QoS	13
1.2.7 QoS 路由和相关的流量工程方法	14
1.2.8 端到端的 QoS 保证和 QoS 信令	15
1.3 网络安全	15
1.3.1 网络安全的基本涵义	15
1.3.2 网络安全的层次结构	16
1.4 从事 QoS 技术研究及标准化的相关组织及其活动	17
1.4.1 ITU-T	17
1.4.2 Internet 工程任务组	21
1.4.3 EURESCOM	23
1.4.4 欧洲电信标准协会	23
1.4.5 QoS 发展组	24
1.5 本书的结构	25

第 2 章 通信 QoS 分类和性能参数

2.1 端用户多媒体应用的性能需求和 QoS 分类	27
2.1.1 用户驱动的性能需求和关键性能参数	27
2.1.2 不同应用关注的性能指标	28
2.1.3 端用户 QoS 分类和性能需求	32
2.2 QoS 参数分类	33

2.2.1	基本概念	33
2.2.2	依赖参数和独立参数	35
2.2.3	独立于服务或技术的 QoS 参数	36
2.2.4	和技术或服务相关的 QoS 参数	36
2.3	ATM 网的网络性能参数、指标和 QoS 分类	38
2.3.1	ATM 网的物理层性能要求	38
2.3.2	ATM 层信元传递事件的定义	40
2.3.3	ATM 层的网络性能参数	41
2.3.4	ATM 层的网络性能指标和 QoS 分类	43
2.3.5	ATM 业务分类及其和 ATM 层 QoS 分类的对应关系	46
2.4	IP 网络性能参数、指标和 QoS 分类	49
2.4.1	IP 网络性能的通用模型	49
2.4.2	IP 网的网络性能参数	52
2.4.3	IP 网的网络性能指标和 QoS 分类	55
2.4.4	IP 网络性能指标分配	58
2.4.5	区分服务的业务分类和 IP 网络 QoS 分类的对应关系	60
2.5	ATM 网络与 IP 网络 QoS 分类、网络性能的对应关系	60
2.5.1	ATM、IP 网络 QoS 分类的对应关系中存在的“矛盾”	60
2.5.2	ATM 网络的 QoS 分类对 IP QoS 的支持	61

第 3 章 分组网络的 QoS 体系结构框架

3.1	QoS 构件模块和体系结构框架	62
3.2	控制平面机制	64
3.2.1	接纳控制	64
3.2.2	QoS 路由	66
3.2.3	资源预留	66
3.3	数据平面机制	67
3.3.1	流量分类和分组标记	67
3.3.2	流量整形和流量管制	70
3.3.3	排队和调度	72
3.3.4	拥塞控制和拥塞避免	75
3.3.5	队列(缓存)管理	78
3.4	管理平面机制	82
3.4.1	服务等级协议	82
3.4.2	流量的测量	83
3.4.3	流量恢复	84
3.5	QoS 信令	84
3.6	安全问题的考虑	86
3.6.1	数据平面的安全考虑	86

3.6.2 管理和控制平面的安全考虑	87
3.6.3 QoS 信令的安全考虑	87
第 4 章 IP 网络的 QoS 保证机制	
4.1 IP QoS 简介	88
4.1.1 传统 IP 分组传送服务与 QoS 控制技术的引入	88
4.1.2 IP QoS 模型	90
4.2 集成服务模型	91
4.2.1 背景	91
4.2.2 RSVP 协议机制	93
4.2.3 集成服务模型的特点	101
4.3 区分服务模型	102
4.3.1 背景	102
4.3.2 DiffServ 网络结构	103
4.3.3 DiffServ 服务类型定义	104
4.3.4 实现 DiffServ 的相关技术	106
4.4 集成服务模型与区分服务模型的互操作	108
4.5 引入 MPLS 对网络 QoS 保证能力的增强	111
4.5.1 概述	111
4.5.2 MPLS 流量工程对 QoS 的改进	118
4.5.3 MPLS 与区分服务模型的结合	124
第 5 章 其他类型网络的 QoS 保证机制	
5.1 ATM 网络的 QoS 保证机制	129
5.1.1 ATM 网络的流量模型和 QoS 分层控制机制	129
5.1.2 ATM 的网络资源管理	132
5.1.3 连接接纳控制	134
5.1.4 用户参数控制/网络参数控制	137
5.1.5 反馈式拥塞控制机制	138
5.2 基于以太网的 CoS 保证机制	140
5.3 子网带宽管理	141
5.4 RPR 的 QoS 保证机制	143
5.4.1 RPR 的流量分类	143
5.4.2 RPR 的带宽预留机制	145
5.4.3 RPR 的公平算法	145
第 6 章 QoS 路由和相关的流量工程方法(一)——QoS 路由方法	
6.1 概述	151
6.1.1 流量工程模型	151

6.1.2 QoS 路由和相关的流量工程方法	153
6.2 基于分组网络的呼叫选路方法	155
6.3 基于分组网络的连接选路方法	156
6.3.1 分等级固定路由	159
6.3.2 基于时间的路径选择	160
6.3.3 基于状态的路径选择	162
6.3.4 基于事件的路径选择	164
6.3.5 域间路由	164
6.3.6 小结	166
6.4 QoS 路由方法中的源路由和分布式路由	168
6.4.1 源路由	168
6.4.2 分布式路由	169
6.4.3 源路由与分布式路由的比较分析	170
6.5 QoS 路由技术的扩展	170
6.5.1 QoS 路由扩展技术的分类	170
6.5.2 减少更新负载	171
6.5.3 减少路由计算	174
6.6 QoS 路由度量参数的选择	174
6.6.1 度量值的选择原则	174
6.6.2 单混合度量参数路由算法	175
6.6.3 多度量参数路由算法	176
6.7 QoS 路由有待进一步研究的问题	178

第 7 章 QoS 路由和相关的流量工程方法(二)——资源管理

7.1 QoS 资源管理方法概述	180
7.1.1 服务类型判别	180
7.1.2 基于策略的路由表的生成	181
7.1.3 QoS 资源管理的实施步骤	182
7.2 动态带宽分配、保护和预留的原则	183
7.3 基于虚网络的带宽分配、保护和预留原则	185
7.3.1 基于 VNET 的带宽分配/预留——网状网	188
7.3.2 基于 VNET 的带宽分配/预留——稀疏网络	191
7.4 基于流的带宽分配、保护和预留原则	191
7.4.1 基于流的带宽分配/预留——网状网	192
7.4.2 基于流的带宽分配/预留——稀疏网络	193
7.5 分组级的流量控制	194
7.6 域间 QoS 资源管理	196

第 8 章 QoS 路由和相关的流量工程方法(三)——容量管理和运营需求

8.1 容量管理方法	199
------------------	-----

8.1.1 链路容量设计模型	199
8.1.2 最短路径选择模型	200
8.1.3 多时段网络设计模型	201
8.1.4 天到天负载变化的设计模型	207
8.1.5 不确定预测/保留容量设计模型	208
8.1.6 网状拓扑稀疏节点和动态传送设计模型	210
8.2 流量工程运营需求	211
8.2.1 流量管理	211
8.2.2 容量管理——预测	216
8.2.3 容量管理——日和周性能监视	219
8.2.4 容量管理——短期网络调整	220
8.2.5 在线和非在线流量工程方法的比较	221

第 9 章 IP 网的 QoS 策略控制方法

9.1 QoS 策略控制系统	222
9.1.1 概述	222
9.1.2 QoS 策略控制系统的组成	224
9.1.3 目录服务器和策略服务器	225
9.2 标准化组织的相关工作	227
9.2.1 目录使能网络	227
9.2.2 分布式管理任务组	228
9.2.3 DMI 与 SNMP 管理框架	228
9.2.4 Internet 工程任务组	229
9.3 COPS 协议	229
9.3.1 COPS 协议的特点和基本工作原理	229
9.3.2 COPS 协议报文的格式	231
9.3.3 通信、句柄的使用和同步行为	233
9.3.4 COPS 协议的通用操作	234
9.4 目录服务和轻量级目录访问协议	235
9.4.1 目录服务的组成和 LDAP 访问协议	236
9.4.2 LDAP 的特点与应用	238

第 10 章 移动分组网络的 QoS 保证技术

10.1 移动 IP 简介	239
10.1.1 移动 IPv4	239
10.1.2 移动 IP 的路由优化	243
10.1.3 移动 IPv6	243
10.2 移动 IP QoS 概述	245
10.2.1 移动环境对移动 IP 中 QoS 的影响	245

10.2.2 移动 IP QoS 解决方案的要求	245
10.2.3 移动 IP 中 QoS 解决方案概述	247
10.3 移动 IP 网络 QoS 框架结构	248
10.3.1 移动 IP 网络的 QoS 框架结构的要求	248
10.3.2 “两层、两级”QoS 框架概述	249
10.3.3 “两级”QoS 机制	251
10.3.4 端到端 QoS 保证	252
10.4 QoS 信令和协商	253
10.4.1 RSVP 对移动 IP 协议的支持	254
10.4.2 移动 RSVP	256
10.4.3 基于移动 IPv6 移动管理信令的 QoS 信令	260
10.4.4 QoS 协商过程	262
10.5 平滑切换中的缓存管理	266
10.5.1 缓存管理概述	267
10.5.2 协议扩展定义	270
10.5.3 移动节点的操作	271
10.5.4 路由器的操作	272
第 11 章 端到端 QoS 保证实现举例	
11.1 例 1:综合采用多类型 QoS 控制机制实现 IP 网络的端到端 QoS	276
11.1.1 概述	276
11.1.2 局域网 QoS 保证的实现	278
11.1.3 WAN QoS 保证的实现	282
11.1.4 LAN 和 WAN 的互联	285
11.1.5 小结	286
11.2 例 2:采用基于策略的集中控制方式实现承载业务端到端 QoS	287
11.2.1 IP 网络端到端的结构划分	287
11.2.2 基于以太网技术的 IP 网络接入部分的 QoS 保证的实现	287
11.2.3 IP 网络骨干部分的 QoS 保证的实现	291
11.3 例 3:下一代 VoIP 网络的端到端 QoS 保证	295
11.3.1 概述	295
11.3.2 下一代 VoIP 网络的体系结构和 QoS 需求	296
11.3.3 下一代 VoIP 网络的 QoS 解决方案	301
11.3.4 MSF 定义的 VoIP QoS 解决方案架构	303
附录 缩略语	310
参考文献	320

第 1 章 QoS 概述

从 20 世纪 80 年代至今,分组网络以其固有的带宽分配灵活、带宽利用率高等特点在通信网络中扮演了日益重要的角色。与此同时,随着电信运营市场竞争的加剧,网络的建设者们越来越关注于在已有的网络中开发新的、符合市场需求的应用和服务,这使网络需要承载的业务类型大幅增加。为了降低网络的建设、运营成本,人们希望采用统一的分组网络实现多类型业务的综合承载,这也是下一代网络(NGN, Next Generation Network)的发展目标之一。如何保证分组网络承载多种类型应用时的业务服务质量,已经成为人们重点关注的问题。

目前,运营网络中大量的业务类型丰富的应用开始由分组网络承载,其中,不乏对业务服务质量要求较高的实时业务。不同类型的应用对网络有不同的服务质量要求,如果网络资源是无限的,那么所有的应用都可以得到其所需的业务服务质量。然而,由于网络的资源是有限的,不可能同时满足所有业务的服务质量要求,因此需要有适当的控制机制通过对网络资源的合理分配以尽量满足各种业务的 QoS 需求,这就是所谓的 QoS 保证机制。

对对象(Object)进行准确定义是讨论任何问题的基本前提,而准确定义和描述业务服务质量是通信网实现服务质量保证的基础,因此本章将首先从介绍 QoS 的基本概念和描述方法开始。在此基础上,在 1.2 节中将简要介绍现阶段基于分组网络的常见的 QoS 保证机制,以使大家对 QoS 技术的发展历史和现状有一定了解。网络安全是人们在讨论 QoS 问题时经常忽视的一个问题,但它确实又是 QoS 保证的重要组成部分,1.3 节中将对网络安全的基本概念进行介绍。另外,QoS 概念和实现机制的标准化一直是业内专家重点讨论的问题,1.4 节将介绍目前在 QoS 方面开展工作的主要标准化组织(如 ITU-T、IETF 等),以及它们在此方面所做的工作。最后,在 1.5 节中将介绍本书的结构和内容安排,以方便读者通读本书或直接查阅感兴趣的章节。

1.1 QoS 的基本概念

1.1.1 QoS 的定义

今天,服务质量这一术语不仅在传统的电信领域里得到了广泛的使用,还逐渐扩展应用于基于分组的宽带、无线和多媒体等服务领域。与此同时,通信网络和通信系统的设计

和规划也越来越多地考虑到用户应用的端到端性能需求。然而,人们在使用 QoS 术语时,术语本身却经常没有有效地定义,使用上也往往不够严谨,甚至出现错误。

有关质量(Quality)概念的定义最初出现在 ISO8402 标准中:质量是有关实体的一组特性的总和,这些特性反映了该实体满足其被明确规定或隐含需求的能力。随后,2000 年发布的 ISO9000 取代了 ISO8402,并将质量定义为一组固有属性对需求的满足程度。

那么服务质量又是如何定义的呢?ITU-T 标准化组织在 E.800 中是这样描述的:服务质量是一种服务性能的综合体现,这种服务性能决定了网络在多大程度上满足业务用户的要求。其他大量出版物(包括许多标准、报告和技术规范等)在使用 QoS 术语时都没有再对其进行清晰地定义,或仅仅简单地引用或指明 QoS 术语定义参见 ITU-T 标准 E.800。

从根本意义上说,QoS 指网络对业务性能要求的支持能力。这里的性能要求是特定于不同业务的,即不同的业务有不同的性能要求。例如,远程医疗强调业务提交的准确性,但对业务端到端的时延和分组的时延变化(如抖动)要求不高;与之相对应的是,IP 电话要求有尽可能小的抖动和时延性能。

对于通信网中的服务质量,上述性能要求可以用性能参数来描述,例如可以包括:业务可靠性、时延、抖动、吞吐量和数据丢失率等。这些性能参数的详细描述见表 1-1。

表 1-1 QoS 性能参数描述

性能参数	描述
业务可靠性	指用户与用户或用户与通信网络间业务连接的可靠性,包括建立时间、保持时间等。
时延	又称延迟,指在两个参考点间,业务流量从发送到接收的时间间隔。
时延抖动	指同一个业务流量中的不同分组,经网络承载后传输时延的变化。
吞吐量	指一个网络中数据包的传输速率,这一参数可以用平均速率或峰值速率来表示。
数据丢失率	指某一业务在网络中传输时,可允许的最大数据丢失率。数据丢失主要是由网络拥塞引起的。

与 QoS 相对应的一个概念是 CoS(Class of Service,服务类别),它是指为特定业务流量提供的一组可用的服务特征,它的对象是一组具有相似性能要求的业务流。相对而言,QoS 具有更高的精确性,它能对业务服务过程中所表现出的各种性能进行具体、准确的描述。

实际上,通信服务质量所涉及的主要评估标准并不仅限于表 1-1 中所描述的 5 种,为了全面而准确的描述通信服务的服务质量,人们又定义了通用的 QoS 模型、描述通信服务质量的框架矩阵,并提出了评估 QoS 的 4 个视角,下面我们分别进行介绍。需要首先说明的是,这里的介绍将主要针对 QoS 的概念,而不涉及具体应用的服务性能参数及其性能指标,这些内容我们将在第 2 章中重点讨论。

1.1.2 通用的 QoS 模型

通常,我们可以将 QoS 定义划分成 3 个方面来理解:服务固有的 QoS、用户感知的 QoS 和用户评价的 QoS。这 3 个方面自上而下构成了通用的 QoS 模型,如图 1-1 所示。此模型也是 ITU(国际电信联盟)和 ETSI(欧洲电信标准协会)组织定义 QoS 时所采用的方法。

服务固有 QoS 属于和服务特性相关联的技术方面的范畴。因而,服务固有 QoS 是由

传送网络设计的优劣,以及网络接入、终止和连接的提供情况所决定的。除了其他一些必要的工作外,通过适当地选择传送协议、QoS 保证机制和相关参数的数值,就能够得到特定应用所要求的 QoS。对服务固有 QoS 的评价需要通过比较对可测量的用户所期望的性能特性来实现。用户对 QoS 的主观感知评价不会影响服务固有 QoS 的等级。

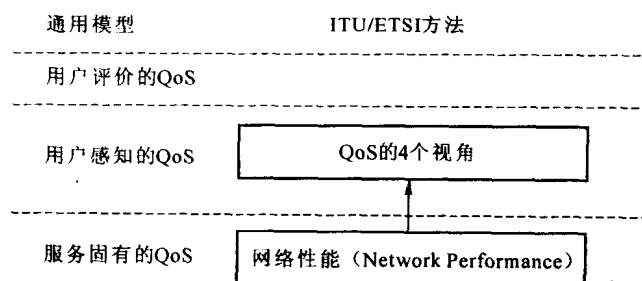


图 1-1 通用 QoS 模型

用户感知 QoS 反映了用户对使用特定服务的体验。它受到观察到的服务性能和用户期望相比较后的比较结果的影响。反过来,用户个人对 QoS 的期望经常受到用户对相似电信服务的体验和其他用户观点的影响。因而,对于不同的用户,具有相同固有 QoS 特性的服务会得到不同的评价。实际上,仅仅确保特定的服务(网络)性能参数还不足以保证能够满足用户的 QoS 要求,而这些服务的用户也不会关心服务究竟是如何提供的。服务提供者提供的 QoS 必须在反映固有 QoS 特性的同时,还反映一些对用户有意义的,以及和特定用户群体期望相关联的非技术参数。用户感知 QoS 包括 QoS 的 4 个视角,这一问题我们将在下一部分中详细讨论。

只有当用户决定是否继续使用电信服务时,服务提供者才能知道相应服务的“用户评价 QoS”。这一决定取决于被感知 QoS、服务价格,以及服务提供者受理用户抱怨和服务问题的响应情况。在这里,对客户提供服务时的服务态度是影响被感知 QoS 等级水平的一个重要因素。

QoS 定义所包含的 3 个方面(固有 QoS、用户感知 QoS 和用户评价 QoS)的用户满意程度可以分别考虑。首先,它是网络提供者的职责,并依赖于网络体系结构、规划和管理。其次,它是工程师、设计人员和网络运营者处理的一个重要的技术问题。针对提供的特定服务来相应调整固有 QoS 性能的合理使用,同时结合适当的市场分析,对于确保高水平的被感知 QoS 是非常必要的,这是服务提供者的责任。广告和市场效果也会影响被感知 QoS。被评价 QoS 主要依靠服务提供者的收费策略,以及可信赖的用户服务人员和技术支持水平。

1.1.3 描述通信服务质量的框架和 QoS 的 4 个视角

1. 通信服务质量的描述框架

在上述 QoS 定义的基础上,人们发现还需要将那些受应用影响的服务功能和这些功能执行时用来评估服务质量的不同标准关联起来。1988 年 10 月,Richters 和 Dvorak 在发表于 IEEE 通信杂志上的一篇文章中提出了一个通信服务质量的描述框架,并很好地实现

了服务功能和服务质量标准之间的关联。该框架随后被欧洲共同体电信工程师联盟 (FITCE, Federation of Telecommunications Engineers of the European Community) 采用, 并加以进一步提炼, 使其具有更好的实用性。

评估一个通信服务的质量标准可以从如图 1-2 所示的矩阵获得。实践已经证明, 这一框架对于在部署或使用一个新业务之前, 评估用户的 QoS 标准是非常有用的。FITCE 研究发现, 根据业务所需要的 QoS 粒度, 可以推导出多达 43 个评估 QoS 的标准(对于基础话音业务, 至少有 13 个重要的评估标准)。

		服务质量标准 (Service Quality Criteria)						
		速率 (Speed)	准确性 (Accuracy)	可用性 (Availability)	可靠性 (Reliability)	安全 (Security)	简易性 (Simplicity)	弹性 (Flexibility)
服务功能								
服务管理	销售和 预定活动							
	预留							
	变更							
	服务支持							
	修补							
	停止							
连接质量	连接建立							
	信息传送							
	连接释放							
计费								
网络/业务的 用户管理								

图 1-2 简化确定通信服务质量标准的矩阵

图 1-2 的矩阵可以应用于任何通信服务, 以确定其必须的服务质量标准。在确定了该业务的 QoS 标准的基础上, 我们可以进一步定义适用于评估该业务的一套性能参数。

2. QoS 的 4 个视角

图 1-2 中的“QoS 定义矩阵”提供了任何服务必须支持的通信功能质量的标准。进而, 这一定义矩阵可以从以下 4 个不同的角度进行描述:

- 用户的 QoS 需求;
- 服务提供者计划提交的 QoS(即计划/目标提交的 QoS);
- 服务提供者达到的 QoS(即服务提供者实际获得或提交的 QoS);
- 用户感知的 QoS(即用户实际测得的 QoS)。

这就是 QoS 的 4 个视角。从这 4 个方面评估都具有实际意义的框架, 才可能具有真正的使用价值, 具有足够的实用性以应用于工业领域。图 1-3 示意了这 4 个视角自顶向下的关联。

(1) 用户的 QoS 需求

用户的 QoS 需求陈述了用户对特定业务所要求的质量水平,这一要求可以使用非技术语言来表述。站在用户的角度,他们不会关心其所需的业务是如何实现的,也不会关心网络内部设计的任何细节问题,他们关心的是最终能够获得的端到端服务质量。从业务用户的角度,服务质量的描述应符合以下原则:

- 集中于用户可感知的服务效果,而不是获得该效果的网络内在的原因;
- 不依赖于网络内部设计时所做的相关假设的准确性;
- 应尽可能考虑站在用户角度的对服务质量要求的所有方面;
- 服务质量应由服务提供者向用户明确承诺,有时应采用协议或合同的形式;
- 应采用和网络类型无关的术语来描述服务质量,描述语言应能够被用户和服务提供者所理解。

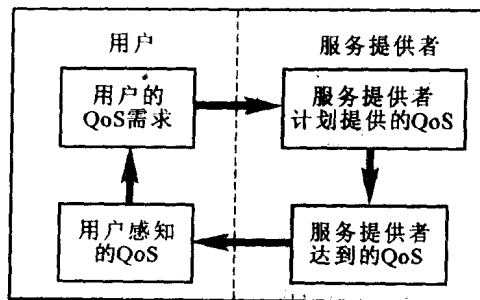


图 1-3 QoS 的 4 个视角

准确而清晰地描述用户的服务质量要求是非常重要的,它是服务提供者规划网络服务等级的基本前提。

(2) 服务提供者计划提交的 QoS

服务提供者计划提交的 QoS 相当于一份服务的质量水平声明,这一声明描述了服务提供者计划向用户提交的服务质量。在这里,服务质量水平可以通过对 QoS 参数分配具体数值来表述。每一类业务都拥有特定于该业务的一套 QoS 参数,这种对 QoS 参数分配数值的 QoS 描述形式显然是比较精确的,主要用于网络规划和服务等级协议(SLA, Service Level Agreement)。服务提供者可以从维护用户利益的角度出发,采用非技术术语描述其提交的 QoS,而同时在具体的协议或合同中采用相对精确的技术术语描述。关于服务的 QoS 分类和参数描述的详细内容,我们将在第 2 章中介绍。

服务提供者提交的 QoS 描述可以用于网络规划文档中,以准确说明相应的性能测量方法和具体指标。该描述也被用做服务等级协议的基础信息。

举例来说,一个服务提供者可以声明:为保证用户的利益,计划一年内的基础电话服务的可用性达到 99.995%,任何一次服务中断的持续时间不超过 15 min,而且一年内中断次数不超过 3 次。

(3) 服务提供者实现或实际提交的 QoS

“服务提供者实现的 QoS”是一份服务质量水平声明,这一声明描述了服务提供者实际实现并向用户提交的服务质量水平。该声明也是通过对一组 QoS 参数分配数值来表述的,这种表述方式和“用户计划提交的 QoS”的表述方式一样,因此我们可以对目标水平和实际获得的 QoS 水平进行比较。这些性能数据往往特定于某一时间周期,例如“在此前的一个月”。

举例来说,一个服务提供者可以声明:“此前的一个季度”实现的可用性为 99.95%,服务中断 5 次,其中最長一次中断达 65 min。“服务提供者实现的 QoS”可作为网络服务的调节依据,该信息的及时发布能够充分保障用户的利益。

(4) 用户感知的 QoS

“用户感知的 QoS”也是一份服务质量水平声明,声明反映了已经得到用户认同的该用户获得的服务质量水平。“用户感知的 QoS”常采用用户满意度来表示,而非采用专业技术术语。感知的 QoS 常被服务提供者用来确定服务质量的满意度。例如,某用户声称他通过网络打通一次电话的重拨次数太多,不满意,因此给 5 分制的服务满意度评分为 2(5 表示很满意)。理想情况下,用户感知的 QoS 和服务提供者实现的 QoS 间的比值为 1:1。

(5) QoS 4 个视角的相互关联

在 QoS 的 4 个视角中,“用户的 QoS 需求”可以被认为是一个逻辑出发点。用户需求可以作为服务提供者的信息输入,帮助其确定“计划提交的 QoS 水平”。服务提供者不一定需要始终向用户提交其需要的服务质量,某些因素(如质量成本、服务提供者经营战略方面的因素、市场竞争等)将影响提交的服务质量水平。“用户的 QoS 需求”也将对监测系统产生影响,设置这些系统的目的是为了实时测量实际提交的 QoS,并定期报告。显然,这 4 个 QoS 描述形式相结合,能够非常准确而全面地对服务质量进行描述,这是进一步有效实现服务质量管理的基础。

1.1.4 SLA、服务生命周期和 QoS

在电信服务领域,和 QoS 密切相关的一个概念是服务等级协议(SLA)。SLA 是服务提供者和用户通过协商,在通信服务的服务品质、优先等级和责任义务等方面达成的协议。ITU-T E.860 定义了一个在多运营商环境下的通用的 SLA 框架。通信服务是一个过程,因而它和与它相关联的 SLA 都有其自身的生命周期。SLA 和通信服务的生命周期可以分为 5 个阶段,分别是服务开发阶段、协商制定阶段、实施阶段、执行阶段和评估阶段,如图 1-4 所示。而在这 5 个阶段中,都涉及到了 QoS 的相关问题。

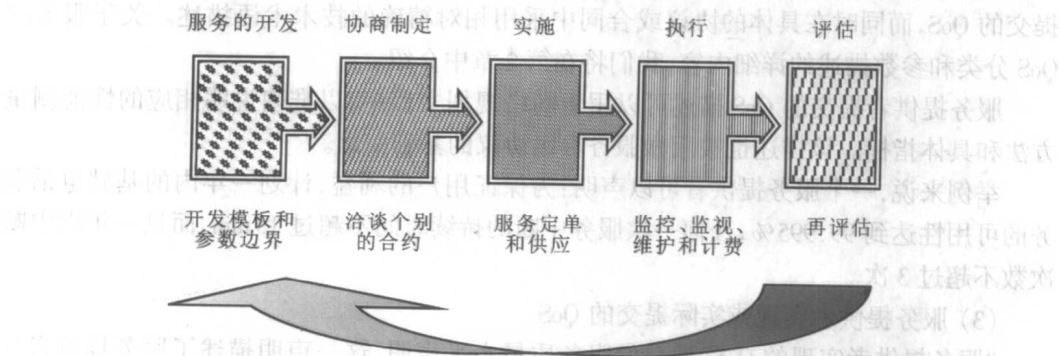


图 1-4 SLA、服务生命周期

1. 服务开发阶段

服务开发阶段是服务生命周期的起始阶段,这一阶段的主要工作包括:明确客户需