

PTN 网络中基于 Diff-Serv TE 的 QoS 研究

陈 胜¹, 李兴明¹, 陈 捷²

(1. 电子科技大学 宽带光纤传输与通信网技术教育部重点实验室, 四川 成都 611731;

2. 中兴通讯股份有限公司, 广东 深圳 518057)

摘 要:在电信业务 IP 化趋势推动下, 传送网承载的业务从以 TDM 为主向以 IP 为主转变。分组传送网 (PTN) 正是 IP/MPLS、以太网和传送网 3 种技术相结合的适合分组传送的下一代传送网技术。在当前的 T-MPLS、PBT 二种主流的 PTN 技术中, T-MPLS 在标准化过程中抢得先机, 较为成熟。IP 网络对服务质量 (QoS) 的支持一直是网络技术研究的热点, 有几种协议支持 IP 网络的 QoS。文中描述了综合服务结构模型 (IntServ/RSVP)、区分服务结构模型 (Diff-Serv) 和 MPLS TE 的原理, 分析了 Diff-Serv TE 的技术原理并提出了基于 T-MPLS Diff-Serv TE 实现多种业务服务质量保证的方案。

关键词:分组传送网; 综合服务; 区分服务; 流量工程; 传送多协议标签交换; 服务质量

中图分类号: TP393.0

文献标识码: A

文章编号: 1673-629X(2011)09-0128-05

QoS Research Based on Diff-Serv TE in PTN Network

CHEN Sheng¹, LI Xing-ming¹, CHEN Jie²

(1. Education Ministry Key Lab. of Broad-band Optic Fiber Transmission and Communication Systems,

Univ. of Electronics Sci. and Tech. of China, Chengdu 611731, China;

2. ZTE Corporation, Shenzhen 518057, China)

Abstract: Driven by the IP-based trend of telecommunication traffic, traffic carried by transport network changes from TDM-based to IP-based. Combining three technologies including IP / MPLS, Ethernet and transport network, packet transport network (PTN) is next-generation transport network technology which is proper to packet transport. In the current three PTN technologies including T-MPLS, PBB-TE, PVT, T-MPLS seize the initiative in the standardization process and is more mature. IP network supports Quality of Service (QoS) has been a hot research of network technology. There are several protocols to support IP network QoS. Describe the Integrated Services model (IntServ / RSVP), Differentiated Services model (Diff-Serv) and MPLS TE, and then analyze the technology principle of Diff-Serv TE and propose an implementation of technology to support data transmission QoS based on T-MPLS Diff-Serv TE.

Key words: PTN; IntServ; Diff-Serv; TE; T-MPLS; QoS

0 引 言

随着 Internet 规模的不断扩大, 网络数据流量日益膨胀, Internet 上的应用层出不穷, 这使得网络中的业务不仅包括传统的 IP 分组数据业务, 也包括 VoIP、IPTV 以及多媒体业务等。IP 网络是基于无连接的数据传送, 在网络中传输的协议数据单元是长度可变、结构统一的分组数据。这种无连接的传输方式正适合于数据通信, 不仅提高了网络的带宽利用率, 也增强了通信的可靠性, 但也这种通信方式也有如下的问题:

(1) 无连接的方式无法保证通信时延, 只是尽力

而为转发;

(2) 随着网络业务流量的变化和传输路由的不同, 分组数据的传输时延将发生变化, 即产生时延抖动;

(3) 如果分组数据的时延超过限制, 数据会发生丢失;

(4) 实时业务以及多种不同级别的业务分别要求满足相应质量的服务。

因此 IP 网络在新的网络业务应用背景下, 提供不同服务质量这一关键问题已经成为 IP 网络向全业务网方向发展的瓶颈。如何实现 Internet 对多种业务 QoS 的支持一直都是通信网络研究的热点, 必然成为未来互联网技术发展的重要方向之一。

业务的 IP 化是网络发展的一个必然趋势^[1,2], PTN 是一种能够很好处理 IP 和以太网等分组信号的新型传送网, 而在当前的 T-MPLS、PBT 二种 PTN 技

收稿日期: 2011-02-21; 修回日期: 2011-06-19

基金项目: 2009 中兴通讯基金资助项目 (编号略)

作者简介: 陈 胜 (1985-), 男, 四川广安人, 硕士研究生, 研究方向为下一代网络、分组传送网; 李兴明, 教授, 博士, 研究方向为网络管理、光纤通信、光交换与自动光网络。

术中, T-MPLS 在标准化过程中抢得先机, 较为成熟^[3-5]。PTN 网络的 QoS 研究已经取得了一些成果^[6-8]。

文中介绍了 IntServ 和 Diff-Serv 两种服务质量的实现技术, 阐述了 MPLS 流量工程的原理, 着重分析 Diff-Serv TE 的实现技术, 提出了基于 T-MPLS Diff-Serv TE 支持网络业务 QoS 保障的技术方案。

1 IntServ 和 Diff-Serv 的技术原理

1.1 综合服务(Integrated Services and RSVP)模型

1.1.1 IntServ 的工作原理

综合服务模型分为两个部分: (1) 网络流量调度和控制, 其中包括接纳控制、包调度器、包分类器等; (2) 建立路径所需要的信令信息, 通过资源预留协议(RSVP)来实现。

1.1.2 IntServ 存在的问题

IntServ 模型从理论上能够确保 IP 网络提供高质量的实时性服务, 然而在实际的网络部署时却面临不少问题。

(1) 可扩展性较差, 在大规模网络中无法适用。

Internet 是一个无连接的网络, 网络不保留数据流状态, 但资源预留协议引入了数据流状态的概念, 要求网络中的每个路由器为每一个网络数据流都建立并维持一个“软状态”, 在大规模的网络中这样的代价太大, 且几乎是不可能实现这样的要求。

(2) 策略控制问题。

策略控制问题即确定用户预留资源的权限问题。

(3) 资源预留和路由协议之间的矛盾。

这是 IntServ 模型目前最难解决的问题之一。从路由协议的角度来看, 某条路径可能是好的路由, 因为这条路径完全满足路由协议的路由策略, 但从 RSVP 的角度来看, 如果该路径上的路由器没有足够的资源可供预留, 就不能为业务数据流建立起这条路径, 因此整个进程只能停留在这里, 直到上层超时终止这个应用进程, 再重新寻找另外的路径。

(4) 对预留操作, IntServ 缺乏相应的身份验证和授权机制, 使网络可能受到攻击, 存在安全问题。

区分服务模型正好能解决这些问题。

1.2 区分服务(Differentiated Services, DS)模型

1.2.1 Diff-Serv 工作原理

Diff-Serv 体系在具体实现中由以下两种主要功能部件组成^[9,10], 如表 1 所示。

Diff-Serv 为具有相同属性的业务数据流集合提供服务, 不同分类集合的数据分组用相异的 Diff-Serv (DS) 字段值来标识。在 IPv4 协议中, DS 字节是通过定义头部 TOS 字段来实现的, 而在 IPv6 协议中, 通过

表 1 Diff-Serv 功能部件

功能部件	位置	功能	措施
网络边界流量调节器	位于 Diff-Serv 域中边界路由器的输入接口	分组分类和标记 流量整形和控制	控制输入流量, 并根据通信配置文件设置 DSCP 字段
单中继段行为	Diff-Serv 域所有路由器	资源分配策略 分组丢弃策略	根据 DSCP 定义的服务特征应用 PHB

定义业务类(Traffic Class)字节, 每个具有标记的数据分组在网络中每一个转发节点得到特定的单中继段行为(PHB, Per-Hop Behavior)或转发处理。Diff-Serv 在实施中将策略控制功能与核心网络剥离, 将 QoS 管理和控制功能转移到网络边界, 网络传送平面不再关心完成控制功能, 大量控制功能通过边界的服务平面完成, 网络中的核心路由器只负责数据分组的存储转发, 即依据 DS 字段值对不同等级的数据分组进行不同级别的转发处理, 通过数据流整合以及逐跳行为方式降低对整个网络信令的依赖程度。其思想与 IP 网络将寻路功能和分组转发功能相分离的设计思路十分契合。基于这样的原因, 区分服务模型具有良好的可扩展性, 非常适合骨干网络的应用。

1.2.2 Diff-Serv 实现技术

(1) DSCP 码点分配。

IPv4 协议报头中定义了一个 byte 的 TOS 域字段, 如图 1 所示。其中包含一个 3bit 的优先级子字段和一个 4bit 的服务类型(TOS)子字段。TOS 子字段的中 4 个 bit 依次代表最小时延、最大吞吐量、最高可靠性和最低费用四个特性, 4 位中只能置其中 1 位。如果所有 4 位均为 0, 那么就意味着是一般服务。

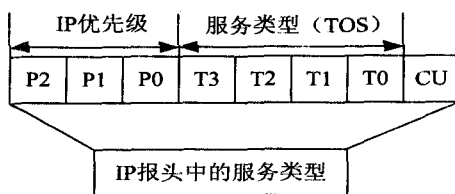


图 1 TOS 域字段分配

由于 TOS 域定义的这些服务选择非常有限, IETF Diff-Serv 工作组重新定义了 TOS 字节(称为 DS 字段)来标记分组的区分服务码点(DSCP), 并定义了一套基本的数据包转发方式(称作 Per-Hop Behaviors-PHB)。通过在数据分组 DS 字段中的不同标记, 以及对依据 DS 字段值进行数据分组的处理, 可以映射成不同的服务等级。DS 字段中的 6 个比特已经被定义, 2 个比特保留为将来之用, 如图 2 所示。

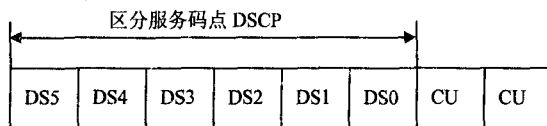


图 2 DS 字段分配

(2) 每一跳行为(PHB)。

Diff-Serv 模型中,网络路由器根据 DSCP 字段来为数据分组执行特定的 PHB,PHB 定义了转发路由器对具有相同 DSCP 值的数据分组执行的转发处理。Diff-Serv 标准定义了 4 种 PHB,两个最常用的 PHB 是 EF PHB 和 AF PHB。

①BE PHB(默认 PHB)。

②CS PHB(类别选择 PHB)。

③EF PHB(加速转发 PHB)。

EF(Expedited Forwarding)PHB 通常用来实现 Diff-Serv 域中的低丢失率、低延迟、低抖动,确保带宽、端到端服务,也叫奖赏服务(Praised Service,PS)。可采用 PQ(Priority Queuing)、CBWFQ(Class-Based Weighted Fair Queuing)、WRR(Weighted Round Robin)、DRR(Deficit Round Robin)等分组调度技术实现 EF 功能。

④AF PHB(可靠转发 PHB)。

AF(Assured Forwarding)PHB 对 Diff-Serv 域内的分组提供不同等级的转发保证,适合大多数基于 TCP 的应用程序。

(3) DSCP 码点与 PHB 的映射。

DSCP 码点与 PHB 的映射如表 2 所示。

表 2 DSCP 编码点与 PHB 的映射

PHB 类别	编码点			
BE PHB	000000 或无法识别的 DSCP 值			
CS PHB	xxx000			
EF PHB	101110			
AF PHB		级别 1	级别 2	级别 3
	低丢包优先级	001010	010010	011010
	中丢包优先级	001100	010100	011100
	高低包优先级	001110	010110	011110

Diff-Serv 最核心的优势在于它很大程度上降低了对信令的依赖,节点路由器只需根据特定的分组标记选择不同的 PHB。可以预见,Diff-Serv 将成为未来广域网中居统治地位的 QoS 技术。

尽管如此,Diff-Serv 也存在一定的局限。Diff-Serv 通过将业务定义为有限的类,可以很好地解决扩展性的问题,但由于各种业务的流量模型和业务模型的不同,各种业务叠加在一起后,其流量模型和业务模型将会是非常复杂的,这使得在现实网络中无法去真正部署 Diff-Serv,必须采用 MPLS TE 等相关技术才能实现 QoS 要求。

2 T-MPLS Diff-Serv TE 实现技术

2.1 T-MPLS TE 的基本原理

流量工程(TE)的目标就是保障网络既能高效率且可靠运行,又能最大限度地均衡网络的业务流量,提高网络资源的利用率。

· 面向应用的性能指标。

这是一种与每种特定应用服务流的流量特性相关的指标,面向应用的性能指标包括了增强业务 QoS 性能的各个方面。

· 面向网络的性能指标。

这是一个与网络资源相关的指标,面向网络的性能指标包括了优化资源利用的各个方面。有效的网络资源管理是实现面向资源优化目标所使用的手段。

理想的流量工程解决方案是根据业务需要分配网络资源、具有将通信流量映射到特殊路径和专用资源上以实现负载均衡的方法^[11,12]。

T-MPLS 流量工程的包括以下几个方面的内容:LSP 路径规划,均衡网络流量,选择 LSP 路径优先级,备份 LSP 路径,网络故障恢复。其中 LSP 路径规划是利用改进版的最短路径优先(SPF)算法(也叫基于限制的 SPF(CSPF)),计算符合特定限制条件的路径。

2.2 T-MPLS Diff-Serv TE 的解决方案

2.2.1 为何采用 T-MPLS Diff-Serv TE

实际网络中,存在一些应用问题无法通过单独使用 Diff-Serv 或 MPLS TE 来解决。比如如下场景:

(1)限制链路上特定级别流量的比例;

(2)限制链路上特定级别流量的比例;

(3)有保证带宽服务。

2.2.2 T-MPLS Diff-Serv TE 具体实现

2.2.2.1 级别类型

Diff-Serv TE 基本的要求是为不同级别的业务流量预留单独的带宽资源。这就要求在设定时间内跟踪统计全网中每个路由器上每一级别业务流量的可用带宽。

为了实现这样的要求,RFC 3564 提出了级别类型(CT)的概念^[13],特定的一组带宽约束条件管理跨越链路的一组业务流量中继段。级别类型在如下情况中使用:分配带宽、寻找约束条件下的路由和许可权的控制。

由于 PHB 是同时由队列和丢弃优先级进行定义的,因此,CT 可以传递多个 Diff-Serv 服务等级的流量,前提是它们全都映射到相同的调度队列中(如 AF2x PHB)。

IETF 建议最多支持 8 个级别类型,从 CT7 到 CT0。在采用 MPLS TE 来保证带宽的前提下,由特定 CT 进行带宽分配的 LSP 叫做 Diff-Serv TE LSP。在当前的实现模式下,一个 CT 的业务流量只能在一个 Diff-Serv TE LSP 中传送。传送相同 CT 业务流量的 LSP,使用的抢占机制可以相同或不同。按照惯例,尽力而为流量映射到 CT0,由于所有预留 Diff-Serv TE LSP 都被看作是尽力而为流量,因此都被映射到 CT0。

2.2.2.2 路径计算

Diff-Serv TE 为 8 种类型的 CT 分配了可用带宽, 以此作为 LSP 路由的路径约束条件。因此在寻找 LSP 路径中需要改进 CSPF, 必须在路径计算中, 结合给定优先级 CT 的带宽约束条件加以考虑。要找到符合约束条件的路径, 必须了解所有链路中优先级别基于 8 种 CT 类型的可用带宽。

这表明链路状态 IGP 必须将每条链路上基于 CT 的各个优先级的可用带宽信息在网络中进行广播。由于有 8 个优先级别 8 个 CT 类型的组合, 因此, 链路状态协议必须承载总共 64 个值。在理想情况下, 所有这 64 个值都将被存储和通过信令进行传递。然而, IETF 决定从可能的 64 个值中提取 8 个值进行广播。

鉴于此, TE 级别被定义成 CT 和优先级组合。Diff-Serv TE 最多可支持 8 个 TE 级别, 从高到底为 TE7 到 TE0。TE 级别可人工配置, 根据实际需要, 从 64 种优先级和 CT 的组合值中进行选择。在最极端的情况下, 存在一个带 8 个优先级的 CT, 极象现在的 TE 实施。在另一个极端情况下, 存在优先级相同的 8 个不同 CT。

链路状态 IGP 广播每个 TE 级别的可用带宽。[DSTE-PROTO] 要求必须使用现有的非预留带宽 TLV 来做广播, 非预留带宽 TLV 此前曾用于传播 TE 的非预留带宽。因此, 通过 IGP 获得的 CSPF 信息只与构成 TE 级别的 CT 和优先级组合相关。因此, 为了使 CSPF 进行有意义的计算, 为 LSP 所选择的 CT 和优先级必须与一个已配置的 TE 级别相对应。

总的来说, 基于 CT 带宽限制条件来计算路径, 需要增强 CSPF 以对每个 CT 的预留要求进行处理, 同时需增强 IGP 以便在不同的优先级承载每个 CT 的可用带宽。

2.2.2.3 带宽限制模式

可用带宽的计算实质就是设置不同 CT 级别带宽分配的比例。带宽限制 (BC) 即表示某个 CT (或一组 CT) 所占用的链路带宽比例。带宽限制模式来表示了 BC 与 CT 之间的关系。

(1) 最大分配模式 (MAM, Maximum Allocation Model)。

MAM 是一种最简单和直接的带宽限制模式, 是在 BC 和 CT 之间一对一的映射。直观来看, 不同的 CT 占用相应百分比例的链路带宽, 如图 3 所示。

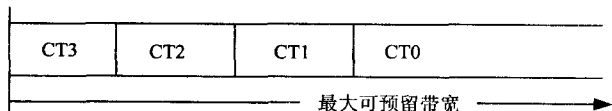


图3 以 MAM 模式将带宽分配给 CT

(为简单起见, 只显示了 4 个 CT)

MAM 的缺点: 由于 CT 间不能共享没有使用的带宽, 因此造成了带宽浪费, 这部分带宽不会被使用来传递其他 CT。由于 CT 和 BC 间一对一映射, 每个 CT 占用各自的带宽, 彼此之间的带宽不可共享, 使得带宽资源浪费, 带宽利用率比较低。

MAM 的优势: 简单, 很容易实现, 不同 CT 之间没有影响; LSP 不会发生抢占, 承载不同 CT 业务流量的 LSP 间无需优先级的配置。

(2) 俄罗斯套娃模式 (RDM, Russian Dolls Model)。

RDM 模式中, 不同的 CT 间可以共享带宽, 从而提高了链路的带宽利用率。俄罗斯套娃模式中, CT0 的业务是尽力而为转发, 而 CT7 是服务质量要求最高的业务。不同 CT 之间共享的带宽是动态变化的。等级越高的 CT 可能抢占等级低的 CT 的 BC。在链路带宽的分配中, 除 BC7 占用链路带宽的比例只能预留给 CT7 的业务流量, 不会发生共享外, 其余的 BC 从 BC6 到 BC0 都可能在某些 CT 之间共享。共享的机制就好比俄罗斯套娃一样, 一层套一层: BC6 由 CT6 和 CT7 的业务流量共享; BC5 由 CT5、CT6 和 CT7 的业务流量共享, 以此类推, BC0 由 8 个 CT 的业务流量间共享, 也就是链路中的全部可用带宽, 在所有 CT 间共享。图 4 形象说明了俄罗斯套娃模式的原理。

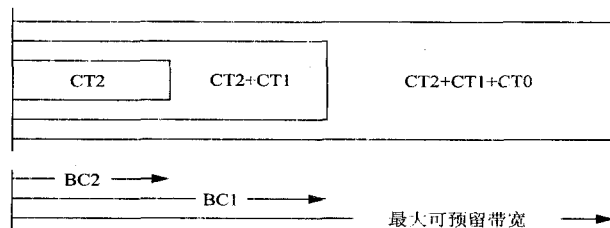


图4 俄罗斯套娃带宽分配模式

(为简单起见, 只显示 3 个 CT)

相对于 MAM, RDM 的优势是通过 CT 间的带宽共享, 不仅提高了网络的带宽利用率, 同时可以为网络中实时性要求高的业务提供一种简单高效的轻载方式。这是因为实时性要求高的业务具有等级高的 CT 值, 通过共享低 CT 值的业务带宽, 既能满足实时性业务的带宽需求, 又不会在很大程度上影响全网的性能。

RDM 的不足之处在于它不能像 MAM 那样隔离不同的 CT, 实现也相对困难一些。而且不同 CT 间可能发生抢占, 这就必须采用相应的机制来保证每个 CT 获得指定的带宽份额。

表 3 总结了 MAM 和 RDM 之间的区别。

综上, CT 占用链路的可用带宽由带宽约束模式来确定。BC 模式的实现包括 MAM 和 RDM 两种形式。这两种模式在链路带宽分配方面采用不同的机制, 需要根据实际网络的情况和业务的服务质量要求采用相

应的模式。

2.2.3 部署 Diff-Serv TE 解决方案

综合前面的内容,部署 T-MPLS Diff-Serv TE 解决方案可采用如下步骤:

1) 根据网络类型和业务的 QoS 要求,采用合适的带宽约束(BC)模式,然后确定网络中每条链路上各个 BC 占用链路带宽的比例。

2) 根据第一步的分析结果,在网络中每条链路上配置缓冲区和带宽分配比例。

3) PTN 网络中一般的业务等级不超过 8 种,从 64 种优先级和 CT 的组合值中进行选择,选取某些优先级和 CT。

4) 选择内部网关协议(IGP, Interior Gateway Protocol),通过 IGP 交换链路带宽的分配信息,实现路由(LSP)。

5) 配置 LSP 信息,包括路径需要预留的带宽、优先等级和 CT 类型。

6) 在需要的情况下,还需对预留的带宽和速率限制进行配置。

7) 如果是 E-LSP 规划,由 EXP 位决定 Diff-Serv 处理方法,如果是 L-LSP 规划则由标签决定。如果基于 EXP 位决定 Diff-Serv 处理方法,则在整个 Diff-Serv 域中配置一致的 EXP-PHB 映射,并确保流量被正确进行标记。

表 3 MAM 与 RDM 的区别

	MAM	RDM
BC 映射	BC 和 CT 一一对应的映射,简单直观,容易实现,便于管理	一个 BC 可能容纳多个 CT,管理上不是很方便
带宽保证	不同的 CT 间相互隔离,每个 CT 有保证的带宽份额,不会发生抢占	CT 间不能相互隔离,采用共享(抢占)机制保证每个 CT 获得相应的带宽份额
链路带宽利用率	带宽利用率较低	通过共享,带宽利用率较高
抢占和重路由	无须抢占和重路由	需要抢占和重路由,影响传输效率,引起流量跌落
使用范围	适用于不允许抢占的网络	不建议在不允许抢占的网络中使用

现在简单了解一下流量工程网络如何实现到 Diff-Serv TE 的迁移。第一步,PTN 网络的运营商应该根据网络需求选择采用哪些 CT 类型和优先级之间的组合。如果是无 CT 要求的 LSP,则用 CT0 进行映射。然后,在迁移时将现有流量工程网络中的 TE LSP 所映射的 CT 要求与优先级进行组合。接着将 TE 级别与上一步中选定的(CT,优先级)组合进行映射。TE 级别的信息覆盖原流量工程网络中非预留带宽 TLV 的信息。网络迁移的过程通常分成几个步骤。这就造成不同步的问题,新旧节点之间广播的非预留带宽 TLV 信

息的语义不同。旧节点将用非预留带宽 TLV 的字段 i 填充(CT,i)的可用带宽。而新节点将为 TE 填充可用带宽。将(CT,i)映射到 TE,就能实现新旧节点间可用资源图的一致性。这样的处理就可以顺利实现不支持 Diff-Serv TE 扩展的节点和支持 Diff-Serv TE 扩展的节点之间的互操作。

3 结束语

差分业务(Diff-Serv)可通过多级服务实现可扩展的网络设计。MPLS 流量工程(TE)可实现资源预留、容错和传输资源优化。T-MPLS Diff-Serv TE 结合了 Diff-Serv 和 MPLS TE 二者的优势,能够为分组数据业务提供多种等级的服务质量(QoS)保证,同时提高网络资源的利用率。由 T-MPLS Diff-Serv TE 提供的 QoS 使网络运营商能够提供话音等需要严格性能保证的业务,并将 IP 和 ATM/FR 网络整合到一个通用网络核心中。

参考文献:

- [1] 龚倩,徐荣,李允博,等. 分组传送网[M]. 北京:人民邮电出版社,2009.
- [2] 黄晓庆,唐剑锋,徐荣. PTN-IP 化分组传送[M]. 北京:人民邮电出版社,2009.
- [3] 张成良. PTN 技术与组网应用[J]. 电信科学,2008(8):1-6.
- [4] 李健,顾晓仪,张杰,等. T-MPLS 分组传送技术[J]. 电信科学,2007(1):85-88.
- [5] 吴江,赵慧玲. 下一代的 IP 骨干网络技术——多协议标记交换[M]. 北京:人民邮电出版社,2001.
- [6] 林娜,齐红满. 基于 MPLS 的下一代网络 QoS 模型[J]. 计算机工程,2008(12):107-109.
- [7] 周敏,龙昭华,王俊峰. VoIP 在基于 MPLS 的集成模型中的 QoS 实现技术[J]. 中国数据通信,2004(2):30-37.
- [8] 秦君,林晓明. 基于 IP/WDM 网络的 QoS 技术研究[J]. 计算机工程与设计,2004,25(5):744-745.
- [9] Nichols K, Blake S, Baker F, et al. Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers[S]. RFC 2474, 1998.
- [10] Blake S, Black D L, Carlson M A, et al. An Architecture for Differentiated Services[S]. RFC 2475, 1998.
- [11] 李蓬,黄河. 基于 NS2 的 MPLS 流量工程仿真研究[J]. 计算机技术与发展,2008,18(9):53-56.
- [12] 许经彩,王新华,薛健,等. 一种新的 MPLS 流量工程最小干涉算法[J]. 计算机技术与发展,2009,19(10):77-80.
- [13] Le Faucheur F, Lai Wai Sum. Requirements for Support of Differentiated Services-aware MPLS Traffic Engineering[S]. RFC 3564, 2003.