

Internet 区分服务体系结构及其实现机制*

吕 佳

(重庆师范大学 数学与计算机科学学院, 重庆 400047)

摘 要 近年来,Internet 研究人员及其运营商都致力于研究和开发具有服务质量保障能力的新型网络技术,以满足越来越多的高质量服务的需求。IETF 提出的区分服务体系结构成为人们讨论和研究的热点,它根据用户的需求将业务分为多种类型,网络节点根据数据包的 DSCP 值选择相应的每跳行为对数据包进行处理。本文深入分析了区分服务的体系结构,并在此基础上提出了一种区分服务的实现方式。

关键词 体系结构;区分服务;每跳行为;确保服务

中图分类号:TP393

文献标识码:A

文章编号:1672-6693(2006)02-0030-04

An Architecture for DiffServ in the Internet and its Implementation

LV Jia

(College of Mathematics and Computer Science, Chongqing Normal University, Chongqing 400047, China)

Abstract In recent years, some newly-emerging network technologies which have the ability of quality of service(QoS) are researched and developed for satisfying the requirement of more and more high QoS by Internet researchers and managers. Some more and more attentions shift to the architecture of Differentiated Services(DiffServ) proposed by IETF. DiffServ classifies all the packets according to users' requirement, these packets get different Per-Hop Behavior to guarantee their QoS in each node through their path. In this paper, the architecture of DiffServ is deeply analyzed and an implementation based on DiffServ is presented.

Key words architecture; differentiated services; Per-Hop behavior; assured service

随着多媒体技术的飞速发展,Internet 已逐步由单一的数据传送网过渡到数据、语音、图像等多媒体信息的综合传输网,同时用户对视频传输的要求也日益提高。基于传统的 TCP/IP 技术,对所有的网络流均提供单一的服务,即尽力而为(Best-effort)服务,没有明确的时间和可靠性传送保障,因此尽力而为服务无法满足不同的应用在本质上对服务质量(QoS)的要求。为此,Internet 任务工程组 IETF 提出了一些保障服务质量的协议和技术,如综合服务体系模型/资源预留协议^[1](IntServ/RSVP),区分服务体系模型^[2](DiffServ, Differentiated Service),多协议标签交换(MPLS)等。在综合服务体系模型中,通过资源预留协议来满足其 QoS^[3],但由于需要增加复杂的信令机制,收集和处理的每一个业务流的信息,

大大增加了网络的额外开销,可扩展性较差,严重妨碍了其在大型网络,特别是重负载网络中的应用。区分服务体系模型则通过对具有不同 QoS 要求的应用分类,并根据不同的类采用不同的优先级进行处理,即在网络节点中对不同种类业务采用不同的每跳行为^[4](PHB, Per-Hop Behavior),以此来实现业务所需的服务质量。

本文从分析区分服务体系结构^[2,5]出发,介绍了其各部分功能,并给出了一种 DiffServ 模型的实现框架。

1 区分服务体系结构

区分服务基于 IP 流分类聚合,区别对待不同等级的聚合流(Stream Aggregate),根据包头的 DSCP

* 收稿日期 2005-04-25

资助项目:重庆市教委科学技术研究项目(No. KJ050802);重庆师范大学科研资助项目(No. 05xly003)

作者简介:吕佳(1978-),女,四川达州人,讲师,硕士研究生,研究方向为人工智能、计算机网络。

(DS Codepoint)选择提供特定质量的调度转发服务 ,其特性是每跳行为。根据文献[2] ,区分服务的体系结构是由在网络节点上实现的若干功能模块组成的 ,包括 PHB、包分类器(Classifier)包标识和重标识、流量调节器(Traffic Conditioner) ,其中流量调节器又包括计量器(Meter)、标记器(Marker)、整形器(Shaper)和丢包器(Dropper)等。下面详细介绍区分服务体系结构的各个组成部分^[6]。

1.1 DS 区域和 DS 区

DS 区域(DS Domain)是由一些相连的 DS 节点(DS Node)构成的集合 ,其中 DS 节点可看作实现了区分服务功能的网络设备(如路由器)的别称。一般 DS 区域由属于同一网络管理机构的网络构成 ,DS 区域有明确定义的边界 ,边界由边界节点构成。通过边界节点将 DS 区域和非 DS 区域互相连通起来。同时边界节点又可根据业务的流向分为流入节点和流出节点。流状态信息的保存与流监控机制的实现等都在边界节点进行 ,内部节点与状态无关 ,只进行简单的调度转发来实现一组或几组 PHB。DS 内外区域通过服务等级约定(SLA ,Service Level Agreement)与流量调约定(TCA ,Traffic Conditioning Agreement)来提供跨区域服务。

连续的 DS 区域构成 DS 区(DS Region) ,区内支持跨越若干区域的区分服务。区内的各区域可能支持不同的 PHB 组 ,并且各自区域的 DSCP 到 PHB 的映射函数也可能不同。

1.2 区分服务标记域(DS Field)及 DSCP

IP 包头的区分服务标记域定义为原 IPv4 包头的 ToS(Type of Service)字节和 IPv6 的流类型字节的前 6 位 ,并重新定义了各位的含义。以此标识业务分组类型 ,供数据包经过 DS 节点是选择特定的 PHB。DS 域定义如图 1。其中 ,DSCP 表示区分服务的标识域 ,共 6 位 ,可以支持 64 种不同的 PHB ; CU(Currently Unused)表示目前尚未使用的 2 位。

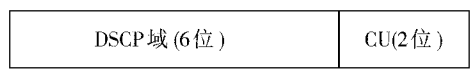


图 1 DS 字节结构

DS 域仍保留了 ToS 值 ,这样在局部不支持 DS 业务而支持 ToS 方式的网络设备中仍可以通过 ToS 方式保证服务质量 ,因此区分服务具有向下兼容性。

1.3 包分类器

包分类器按照 TCA 中的特殊规定 ,根据 IP 包头的某些域将包分成不同的聚集流。目前已定义了 BA(Behavior Aggregate)分类器和 MF(Multi-Field)

分类器两种包分类器。前者根据 IP 包头的 DSCP 将包分类 ;后者则根据包头的多个域 ,如 IP 源地址、IP 目的地址、DSCP、源端口地址、目的端口地址、协议 ID 等 ,对 IP 包分类。

1.4 流量调节器

流量调节器在逻辑上又分为计量器、标记器、整形器和丢包器。如图 2 所示。

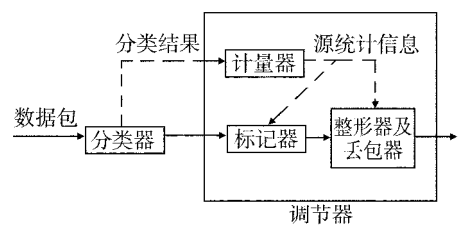


图 2 流量调节器

计量器根据 TCA 中的流规格计量器的某些实时属性 ,如速率等 ,并将统计信息传给标记器、整形器和丢包器。标记器在包头的 DS 标记域中标记适当的 DSCP ,即将分组划入某个行为聚集。整形器、丢包器通过延迟、丢弃等手段强制使入流(或出流)符合 TCA 的流规范。

1.5 每跳行为

PHB 是整个区分服务网络的核心 ,本质上描述的就是单个节点为特定流聚集分配资源的方式 ;区分服务体系的整体资源分配策略也就通过这一个个单个节点资源分配实现的。多个 PHB 由于彼此关系密切而必须同时定义 ,则在实现时就构成一个 PHB 组。若干 PHB 组可以同时定义 ,则称其属于同一 PHB 组簇。目前已标准化的 PHB 组有缺省型 BE(Best-Effort)、加速型 EF^[7](Expedited Forwarding)、确保型 AF^[8](Assured Forwarding)以及兼容 IP 优先级的类选择 CS(Class Selector)4 种。

2 一种区分服务模型的实现方式

为了便于理解区分服务体系结构 ,结合一种实现方式来具体说明。在一个 DS 域内实现 AF ,提供确保服务^[8](AS ,Assured Services)。AF 用于需要可靠传送而不是固定时延范围的应用。即在网络拥塞的情况下仍能保证用户拥有一定量的预约带宽。AS 的服务原则要求 (1) 无论是否拥塞 ,保证用户占有预约的最低限量的带宽 (2) 当网络负载较轻而有空闲资源时 ,用户可使用更多带宽。这样 ,用户最终实际得到的带宽分为预定最小保证值和与其它 AS 流或 BE 流竞争剩余资源而获得的额外带宽两部分。

实现调节器的算法都比较成熟,如令牌桶(Token Bucket)、漏桶(Leaky Bucket)等。还有一种通用调节器^[9],包括了调节器的各逻辑模块,且各模块及模块间的各种属性关系都有参数可以设置,通用性强,通过合理设置参数可以实现 AS 服务。本文中采用最简单的带优先级的队列和通用调节器技术来实现。

网络向用户提供 AS 时,可按以下步骤进行。

(1)首跳路由器对进入网络的分组进行分类过滤,设置与 AS 对应的 DSCP 值,预约带宽以内的流量记为 IN(in profile),超出预约带宽的流量记为 OUT(out of profile);采用通用调节器对进入该网络的分组进行必要的整形,以使其符合流规格。采用该方案的优点在于主机的功能不需要作任何改动,有利于区分服务网络的早期开发。

(2)内部路由器实现简单的存储转发功能,即在内部路由器上维持两个不同级别的队列。将已置 AS 标记的分组送入高优先级的缓冲队列,其它没有设置 AS 标记的分组进入低优先级的缓冲队列;按简单的优先级调度算法,首先将高优先级队列中的分组全部发送出去。同时还要完成按优先级丢弃的工作。网络拥塞时包头标记决定分组的丢弃概率,OUT 的丢弃概率大于 IN,从而一定程度上保证 IN 流;内部路由器在调度转发时要保证源头相同的流不乱序,不管其中分组是 IN 或 OUT。

(3)边界路由器对跨 DS 区域的 AS 服务分组进行必要的监控。若边界路由器为出口路由器,则根据该 DS 区域与下游 DS 区域之间签订的 TCA,采用通用调节器对分组进行必要的整形(也可选择丢弃)。若边界路由器为入口路由器,则只需对不符合 TCA 规范的分组进行简单的丢弃即可。对不符合 TCA 的 AS 分组进行丢弃操作的目的是为了防止恶意用户过度使用 AS 服务。

AF PHB 组簇包含 N 个相互独立的 AF PHB 组,每组中有 M 个 PHB 分属 M 个相对丢弃优先级,目前的定义是 $N=4, M=3$ 。根据资源预留规格,各 DS 节点为每个 AF PHB 组预留一定量资源,以保证 AF 组对应的流在任何时候能获得预约最小带宽。同时,DS 节点还应保证,在同一 AF 组内,低丢弃优先级流聚集的丢失率应小于高丢弃优先级流。无论属于哪个优先级,同一 AF 组内不能改变流内分组的顺序。

3 具体实例

下面给出一个具体实例来说明区分服务模型实

现 AS 的过程。如图 3 所示。

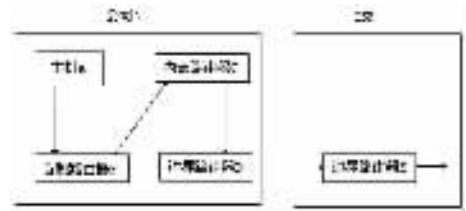


图 3 区分服务模型的一个具体实例

公司 A 网络内的某主机 A 需要通过公司 A 的 ISP 到达某个 IP 地址(主机 B)。若要对主机 A 上发出的分组提供 AS,则按如下步骤操作。(Flow A-B 表示主机 A 到主机 B 的分组)。

(1)首跳路由器 B。首先对来自主机 A 的分组 Flow A-B 作标记,预约带宽以内的流量标为 IN,超出预约带宽的流量标为 OUT;用通用调节器对 Flow A-B 进行整形;将已置 AS 标记的 Flow A-B 送入高优先级的缓冲队列,其它的分组送入低优先级的缓冲队列;按简单的优先级调度算法,首先将高优先级队列中的分组全部发送出去。

(2)内部路由器 C。将到达的分组按照是否设置了 AS 标记送入不同队列,再按简单的优先级调度算法,首先转发高优先级队列中的 Flow A-B 分组,且要保证源头相同的分组按顺序转发。

(3)边界路由器 D。按 AS 是否设置相应标记将分组分别送入两个优先级队列;按公司 A 与 ISP 之间达成的 TCA,配置通用调节器来整形使分组符合流规格;按简单的优先级调度算法首先转发高优先级队列中的分组。

(4)边界路由器 E。按公司 A 与 ISP 达成的 TCA,配置通用调节器;对 Flow A-B 进行监控,符合 TCA 的分组进入高优先级队列,不符合的丢弃;其它分组进入低优先级队列;按简单的优先级调度算法,首先转发高优先级队列中的分组。

最终 Flow A-B 到达目的主机 B。

4 结论

区分服务体系是一种基于业务分类及相关质量保证策略的体系。由于其较好的扩展性和实现简单已成为支持 IP QoS 的首选方式和研究的热点,具有如下优点。

(1)将 Internet 的各种业务继承到少数几种业务类型和 PHB 方式中,通过包头的 DSCP 选择分组的转发方式,简化了网络内部节点的服务机制。

(2)简化了网络内部节点的服务对象,采用聚

集传输控制,服务对象是流聚集(Stream Aggregate)而非单流,单流信息只在网络边界保存和处理。

(3)DS区域内采用自定的PHB方式,DS区域外采用TCA,使整个网络具有良好的兼容性和可扩展性。

(4)兼容其它的QoS技术。如可以与IntServ/RSVP或MPLS等技术结合使用。兼容IPv4的ToS。

但是,由于区分服务是一种新的体系结构,还有许多问题^[10]需要解决:如何确保AS中TCP和UDP流的公平性;如何有效结合IntServ和区分服务机构以实现端到端QoS保障;如何在区分服务体系结构中解决组播问题和安全计费问题等。

总之,区分服务作为一种新的网络技术,还需要更多的研究和实践。

参考文献:

- [1] BRADEN R, CLARK D, SHENKER S. Integrated Services in the Internet Architecture: an Overview[J/OL]. <http://www.rfc.net/rfc1633.html>. 1994-06/1999-07-15.
- [2] BLAKE S, BLACK D, CARLSON M, et al. An Architecture for Differentiated Services[J/OL]. <http://www.rfc.net/rfc2475.html>. 1998-12/1999-07-15.
- [3] 袁平,李明. 基于IP网络的QoS体系结构[J]. 重庆师

范学院学报(自然科学版),2003,20(2):9-13.

- [4] CLARK D, FANG W. Explicit Allocation of Best-effort Packet Delivery Service[J]. IEEE/ACM Transaction on Networking, 1998, 6(4):362-373.
- [5] NICHOLS K, BLAKE S. Definition of the Differentiated Services Field(DS Field) in the IPv4 and IPv6 Headers[J/OL]. <http://www.rfc.net/rfc2474.html>.
- [6] 吕佳. 基于IP区分服务的实时视频流传输[J]. 重庆师范大学学报(自然科学版),2005,22(1):23-26.
- [7] JACOBSON V, NICHOLS K, PODURI K. An Expedited Forwarding PHB[J/OL]. <http://www.rfc.net/rfc2598.html>. 1999-06.
- [8] HEINANEN J, BAKER F, WEISS W. Assured Forwarding PHB Group[J/OL]. <http://www.rfc.net/rfc2597.html>. 1999-06.
- [9] LIN L, LO J, OU F. A Generic Traffic Conditioner[J/OL]. <http://www.ietf.org/internet-drafts/draft-lin-diffserv-gtc-ol.txt>. 1999-08.
- [10] 林闯,单志广,盛立杰,等. Internet区分服务及其几个热点问题的研究[J]. 计算机学报,2000(4):419-433.

(责任编辑 黄颖)