

Internet 服务质量管理协议及协议体系结构^{*}

QoS Protocols and Architectures in the Internet

27-29

陈庶樵 鄢江兴

TP393.4

(国家数字交换系统工程技术研究中心 郑州 450002)

Abstract Future Internet will become an integrated services network that will be able to provide high quality of service level to various services. This paper describes the needs of QoS, QoS protocols and QoS protocol architectures. It also gives a representative model of end-to-end QoS.

Keywords Internet, QoS, RSVP, MPLS, DiffServ, SBM

一、引言

随着 Internet 近年来的飞速增长,各种各样的新业务不断出现,包括电话、电视、广播网在内的各种网络融合入 Internet 已经提上日程。网络的数据流量随着用户及应用的增多以 Moore 定律的速度增长,问题在于是否单纯增加网络带宽就可以满足用户越来越高的要求?答案是否定的。Internet 的数据流量不仅仅是简单的增长了,关键在于多种新业务的特性不同于传统的 Internet 业务。比如 VOD 业务,不仅要求有较高的带宽资源,更对传输时延和时延抖动有很高的要求。要解决这些问题,单靠传统的 Internet 协议是不够的,只是在原有的网络中加入服务质量(QoS)管理协议,才能从根本上解决网络用户及应用的增长与网络有限带宽资源之间的矛盾。

二、什么是 QoS

QoS 是网元(如主机、路由器或网关)在一定程度上能够满足流量及业务需求的能力。QoS 并不能增加网络的带宽资源,这一点是肯定的。QoS 只是根据业务的需求和网络管理配置管理带宽,因此, QoS 是将带宽资源分配给特定的数据流,而且 QoS 并不保证那些尽力型(best-effort)业务的服务质量。

从本质上讲,存在两种形式的 QoS:一是资源预留:网络资源以业务的 QoS 要求为根据进行分配,并且服从于带宽管理策略。RSVP 协议就提供资源预留机制。二是优先级划分:根据带宽管理策略,网络将各种数据流加以划分,赋予不同的优先级别。对那些 QoS 要求较高的数据流分配高优先级,反之则分配低优先级。DiffServ 协议提供这种优先级划分机制。应该指出

的是,这两种 QoS 并不是互斥的,与此相反,它们是完全兼容的。只有充分考虑、合理设计这两种 QoS,利用它们的互操作性,才能在 Internet 这一异种网络互联的网络上为用户提供满意的服务质量。

按照 QoS 对应用数据流的控制方式,又可以将 QoS 分为以下两类:

- 单一流 QoS:一个流被定义为一个独立的、单向的两个应用之间的数据流,它具有唯一的五元组(传输协议、源地址、源端口号、目的地址、目的端口号)。单一流 QoS 指的就是对每一个这样的单一数据流分别进行控制。

- 聚合流 QoS:聚合流指两个或两个以上的流组成的数据流的集合。通常每一个聚合里的数据流具有一些相同的特征,如一样的优先级、相同的认证信息或相似的五元组等等。聚合流 QoS 指的是对聚合里的数据流进行统一的 QoS 控制。

三、QoS Protocols

应用、网络拓扑和网络策略决定着哪一种 QoS 最适合应用于具体的单一流或聚合流。为了适应不同种类的 QoS,产生了多种 QoS 控制协议。

3.1 资源预留协议:RSVP(ReSerVation Protocol)

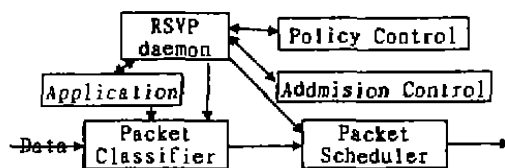


图1 RSVP节点主机软件结构

^{*} 本课题得到 863 重大课题资助(863-300-01-03-99)。陈庶樵 博士研究生,主要从事 IP 网络体系结构分析研究工作。鄢江兴 教授,博导,国家 863 高技术研究发展计划通信技术主题专家组成员,主要从事信息系统和计算机应用研究工作。

RSVP 用于建立和控制网络预留资源,以在 IP 网络上提供尽可能与电路仿真相近的效果为最终目的。对于主机及网元(路由器、交换机)来说,RSVP 是最复杂的 QoS 控制协议。RSVP 的简略工作过程如下:

- 数据流的源主机为将要发送的数据流作出一个规范的描述 Tspec,包括传输数据流所需要占用带宽的上限和下限、时间延迟和延迟抖动。主机中的 RSVP 信令模块则向目的主机发送 PATH 消息,其中包括 Tspec 信息。在源主机到目的主机的下行线路上的每一个支持 RSVP 的路由器在收到 PATH 消息时都在内部建立起链路状态标识 Path-state。

- 为建立起资源预留,目的端主机在上行线路上发送 RESV 消息,包括预留服务的种类及数据流描述符。

- 当上行线路上的路由器收到 RESV 消息时,路由器中的接纳控制进程负责消息的认证及资源预留控制,若预留失败,则发送错误信息至目的端主机,反之则向上行线路的下一跳路由器发送 RESV 消息。

- 当上行线路上的最后一个路由器资源预留成功时,则向目的端主机发送确认信息。

- 结束 RSVP 控制的基本工作原理与建立 RSVP 控制的基本工作原理类似。

RSVP 协议的特点可以总结如下:

- RSVP 不是传输协议,而是网络控制协议。它并不负责传输数据流,而是与 TCP 或者 UDP 协议并行工作。

- 为适应多个广播用户之间的差异,资源预留类型及预留资源多少的决定是基于目的端主机的。

- 资源的预留需要目的端主机周期性地刷新。

- 若线路上有不支持 RSVP 的路由器,则这些路由器对 RSVP 控制的数据流的传输方式是传统的尽力传送。

RSVP 节点主机软件结构如图 1 所示。

3.2 多协议标签交换: MPLS (Multi-Protocol Label Switching)

MPLS 是以优先级划分为基本工作原理的 QoS 控制协议,与本文所讨论的其它 QoS 控制协议不同的是, MPLS 只是工作在路由器或交换机中,在主机上则不存在 MPLS 控制模块,并且 MPLS 是独立于网络协议的,这就是说, MPLS 可以工作在 IP、IPX、ATM 或者 Frame Relay 等等各种网络中。

通常将支持 MPLS 协议的路由器称为标签交换路由器 LSR (Label Switching Router)。LSR 的简要工作原理如下:

- 在 MPLS 网络中,由第一跳的 LSR 根据目的地址及网络管理策略来决定数据流的适当标签 FEC (Forwarding Equivalence Class),并将标签和数据包封装成一个 MPLS 数据包。

- 在下一跳 LSR 中,以上一跳的标签作为索引,查出本地的转发表中的下一跳的地址及新的标签,重新封装数据包并发送至下一跳。

在众多的 LSR 组成的 MPLS 网络中,发布和管理标签是一件既复杂又重要的事,因为所有的 LSR 要在各种标签上达成协议。标签发布协议 LDP (Label Distribution Protocol) 就是为达到这一目的而建立的。虽然 LDP 是一个非常复杂的协议,但它对大多数网络管理者来讲是透明的。网络管理者只需知道 MPLS 分发标签的策略,而不需要去管理具体哪一个标签该如何分配。

3.3 分类服务协议: DiffServ (Differentiated Services)

分类服务协议提出了一种简单的、较粗糙的方法来对不同的业务进行分类。对每一跳节点行为 PHB (Per Hop Behaviors) 来说,分类服务协议定义了以下两个服务水平:

- 快速转发 EF (Expedited Forwarding): EF 是最高服务等级,在分类服务中,服务等级被称为 Diff-Serv value。分类服务协议对具有 EF 服务等级的业务数据流给以最好的 QoS 保证,包括尽可能地减少时延和时延抖动。超过本地管理策略所限定的 EF 最大流量的具有 EF 等级的业务数据包将被丢弃。

- 确保转发 AF (Assured Forwarding): AF 是较低的服务等级。AF 又可以分为四个子类,在每个子类中又可以分三个优先权。超过本地管理策略所限定的 AF 最大流量的业务数据流不会得到更多的带宽资源,其数据包将被降低服务等级,在必要时会被丢弃。

分类服务协议简单、灵活。当分类服务协议利用 RSVP 参数来定义业务服务等级时,可以建立起一条具有相当好的 QoS 聚合流。这一点将在 QoS 体系结构中进行进一步的说明。

3.4 子网带宽管理协议: SBM (Subnet Bandwidth Management)

与 ATM 网络不同,局域网 LAN 在设计之初并没有考虑在数据链路层加入 QoS 控制协议。LAN 采用共享的传输媒体,使它提供给业务的服务质量类似于 IP 网络的尽力服务,影响了 LAN 提供实时、多媒体业务的能力。为弥补这方面的不足,IEEE 制定了 SBM 协议,包括 802.1p、802.1Q 和 802.1D 标准。

SBM 的基本思路是在每个 LAN 网段上,都指定一个 SBM 控制交换机 DSBM (Designated SBM),该网段上的所有数据包都经过这台交换机。DSBM 与网络应用程序建立请求模型 RM (Requestor Module),在带宽分配 BA (Bandwidth Allocator) 与 RM 以及 RM 与 RM 之间建立信令机制,初始化网络预留资源,并和 BA 协调管理网络资源。SBM 服务等级划分如表 1 所示。

