

15-18

与资源预留相结合的 QoS 路由问题研究*

QoS Routing Research Combined with Resource Reservation

冯 径 顾冠群

TP393.4

(东南大学计算机科学与工程系 南京 210096)

(东南大学教育部计算机网络和信息集成支撑技术开放实验室)

Abstract In this paper, we analysis and describe the improvement on routing technology abide by IETF integrated network framework, which provides some mechanism to guarantee network service quality. Based on the review of QoS routing problem, we present a new idea to select a path accommodating requested QoS combined with resource reservation.

Keywords Integrated service, QoS routing, Resource reservation, RSVP

1 引言

十年前,结合快速的处理器、光纤链路和分组交换机制,开发和部署综合电话网络 and 传统数据网络的能力,建立一个可提供综合服务网络的单一的基础网络结构,这种最初的概念导致了 ATM 网络的出现。ATM 在 QoS、有效资源的使用和拥塞控制方面有独到的解决方法,是对分组交换和多路复用技术的一大发展,以 IP 技术为特征的 Internet,用数据报表达了最小的网络通信服务,即通过由网络提供的最少功能来获得构造复杂功能的灵活性,具有规范的包容性与开放性。其特点是:逐跳的分布式处理、无连接、公平性、独立与基础网络和不可靠传递,对于商业应用来说,网关不能直接了解到整个发送序列的存在,它被迫孤立地处理每一个报文分组,因此,资源管理和记账必须在每个分组上独立操作。同时由于中继结点中不保存会话的状态信息,给不同类型的通信量提供不同 QoS 保证带来了困难,需要补充新的协议和控制机制。

数据传输的 QoS,至少应该包括以下几个部分的研究:应用特性分析和 QoS 级别;流分类机制;QoS 参数转换;接纳控制;资源预留;分组调度、流整形;拥塞控制、速率控制和 QoS 路由选择等。

IntServ 作为 Internet 体系结构的扩充,由两个基本元素组成:集成服务模型和参考实现框架^[1]。IntServ 模型包括两类面向实时通信量的服务:确保服务和可预测服务,它们在受控的共享链路中加以集成,即在原来 Internet 的尽力服务(B-ES)的基础上,增加确保服务(GS)和负载受控服务(C-LS)。假设网络资源

是可以显式管理的,那么资源预留和接入控制就是实现这种服务的关键。RSVP 已被 IETF 指定为 Internet 集成服务中的资源预留协议,如何保证把网络资源预留到支持应用需求的服务质量(QoS)的路径上,是合理使用网络资源并提高预留成功率的关键。本文将在剩下的部分里分别介绍 QoS 路由研究的现状并分析其与资源预留相结合的问题。

2 路由问题研究现状

2.1 路由协议的网络模型

通常,网络可以用图来建模: $G=(V,E)$ 。图中结点 V 可用来表示交换机、路由器和主机;边 E 代表通信链路,如果链路是对称的,可以用无向边表示,链路的状态可以是一个多元组:(带宽、延迟、代价,……)。

目前 Internet 路由协议如 OSPF、RIP 等,基本上是基于“最短路径”法,即路由在单个特征值(metric)下的优化。它们可以根据链路的传播时延或源、目的之间的跳数(hops)等,但这些度量标准一般不在一个算法中同时使用。为了支持更大范围的 QoS 需求,路由协议需要一个更复杂的模型,在该模型中用多个特征值(如带宽、延迟、丢失率等)来描述网络特性。人们把这种根据网络上可利用的资源和流的 QoS 需求等知识来决定一个流的路由的机制称为 QoS 路由。

因此,QoS 路由的基本问题就是在源和目的之间,如何找到一条能够同时满足多个约束条件的路径。多特征值可以在某种程度上更精确地描述网络,然而,找到一条符合多个限制条件的路径本质上是一个难度很大的问题。这个问题的多项式算法可能不存在。早在

* 本论文的研究工作受国家重大自然科学基金和 863 课题资助(项目编号分别为 9841-009 和 69896249),冯 径 博士生,研究方向为高性能网络协议处理和网络集成服务技术。顾冠群 教授,工程院院士,研究方向为高性能计算机网络体系结构和 CIMS。

1979年, M. R. Garey 和 D. S. Johnson 在“Computer and Intractability—A Guide to the Theory of NP-Completeness”一书中已经将两个约束的寻径问题“最短加权约束路径”列为 NP-完全问题, 由此推出多个特征值的相加、项乘、取极值以及这几种方式混合的寻径问题, 都属于 NP-完全问题。目前, 一般用启发式算法来求解, 有二种路由策略: 源路由、分布式路由和层次路由。

2.2 路由的基本问题

路由有两个基本任务: 一是收集网络状态信息并即时更新; 二是基于所收集的信息为新的连接找到一个可行的路径。路由问题基本上可以分为两类: 单点投递路由(Unicast Routing)和多点投递路由(Multicast Routing)。

定义 1 单点投递路由是指给定一个源结点 s , 一个目的结点 d , 一系列约束条件 C 以及可能的优化目标, 找到一条从 s 到 d 的最可行的路径, 该路径满足 C 。

定义 2 多点投递路由是指给定一个源结点 s , 一个目的结点的集合 D , 一系列约束条件 C 以及可能的优化目标, 找到一棵从 s 到 D 中所有结点的最可行的树, 该树满足 C 。

由上述定义可见, 多点投递路由的优化或约束必须要在整个树上实施, 而不是单独的路径。二者都应该包括一系列的协议(如链路状态协议、距离向量协议等)、机制(基于本地状态、基于全局状态、在何时、何地计算路由等)和寻径算法(如 Bellman-Ford, Dijkstra, DFS, BFS 等)。

2.3 路由机制的演变

在电路交换网络上, 一个简单的适应性路由方案是可替代路由。这是一个逐跳的、基于目的地址的路由策略。为每个结点计算的路由表, 对每个目的地都列出了多个输出链路, 一旦某个建立请求被接收, 它要试探所有输出链路, 直到找到一个满足要求的链路, 然后将建立请求转发到下一个结点, 并在该链路上预留资源。建立过程可能到达目的地, 也可能被中途挂起。ATM 网络中的 VC 路由问题, 也类似于电路交换网络, 其目的都是力争比传统的最短路径法获得较高的吞吐率。建立一个 VC 时, 首先沿着主要路径选择路由, 如果在该路径上无有效资源可利用, 替代的路径以到达目的地跳数的增加顺序进行试探。在沿替代路径的各链路上分配与首选路径预留一致的带宽, 首选路径带宽的预留根据假设的通信量特征而定。由于这种机制只能以一种速率工作, 其实用性受到限制。还有一种想法是将可替代路由与路由代价联系起来, 每个 VC 具有一个允许的路由代价, 当某个 VC 的路径上的代价小于等于允许的路由代价, 该 VC 被接纳。这种方案的性能还没有得到彻底的分析。

可替代路由思想可以很容易用于分组交换网络, 但要建立相应的控制机制, 避免对重负载网络的性能损害(退回和占用过多容量)。ATM 的专用网络到网

络接口协议(PNNI)可以分成两部分: 下信息协议,

一个路由协议。前者用来建立点到点和点到多点连接, 并支持源路由和可替代路由。后者是通过泛投(flooding)方式向全网广播拓扑信息来动态地、按层次地选择路由。拓扑信息是一组资源的集合(结点、链路地址等), 资源又通过特征值和属性被量化(如延迟, 可利用带宽, 抖动等), 它们组合起来支持各类通信量。该协议还有自动配置路由等级、连接接纳控制、聚集拓扑和可达性信息等功能。PNNI 路由协议不解决多点投递路由、策略路由和可替代路由的控制。目前, ATM 论坛正制定集成 PNNI(I-PNNI)规范, 想把 QoS 路由和 PNNI 的思想综合起来, 为包括 ATM 核心在内的网络提供第二层路由协议。

早期分组网络的适应性路由主要目标在于当拥塞时使端-端延迟最小, 有基于 K-最短路径表的替代路由。路由选择基于一些本地测量。原 ARPAnet 的路由机制是一个基于延迟代价度量的距离向量协议, 它导致了触发更新、泛投等技术的产生。这些技术被用于 OSPF 和 PNNI 路由协议。

目前 IP 主要使用两种路由协议: 域间路由协议 (IDRP) 和开放最短路径优先 (OSPF), 支持基于服务种类的路由。还没有一个标准的方法来使用 IP 报头中的服务种类 (ToS: Type of Service) 字段来说明 ToS 和进行路由选择, 但似乎有可能利用 IP 的 ToS 特征, 结合适当的网络工程方法, 解决一部分基于 QoS 的路由问题。

在多点投递路由方面, MOSPF 是对 OSPF 在多播上的一个扩展, 适用于延迟约束的多播路由。在于网内的组成员信息的变化由本地路由器检测, 然后广播到其他结点。每个结点附加一个全局状态, 维护路由范围内的各组成员信息。斯坦纳树 (Steiner Tree) 是一个跨越给定的结点子集上的最小代价树, 严格地说, 它不是一个 QoS 路由问题。但构造一个斯坦纳树的启发式方法对于如何构造一个其他条件的斯坦纳树是有影响的。

3 主要的 QoS 路由算法

3.1 路由度量值的选择

QoS 路由选择的目标有两个: (1) 所选择的路径必须是满足 QoS 约束的可行路径; (2) 所选择的路径必须尽量有效地使用网络资源并使网络的资源利用率最大。

为了满足 QoS 需求, 要选择不同的网络度量标准作为路由算法的操作对象。对度量值的选择, 应该满足三个标准^[3]: (1) 必须存在计算路径的有效算法; (2) 必须反映网络的基本特征和 QoS 对网络元素 (结点、链路) 的约束; (3) 度量值之间应该互为正交, 没有冗余信息。因此, 很可能出现这样的情况, 即当要求十分严格时, 可能根本找不到一条满足条件的路径; 或者, 由于一个连接过多地占用资源而使其他任务无法执行, 牺牲了网络的利用率。

通常带宽、延迟、丢失率、代价以及延迟抖动是较为有用的 QoS 参数,但对大多数应用来说,延迟比其他参数更重要,因此,在目前的一些 QoS 路由算法中,大多选择两个尺度的组合作为路由选择的约束条件,如带宽和传播延迟,带宽和抖动,带宽和代价等。这是因为第一,路由选择的约束条件越多,算法的复杂性会越大,路由选择本身的开销也就越大;第二,通过研究,我们观察到上述几个 QoS 约束之间在某些方面是相关的,如端到端的排队延迟、延迟抖动以及所需的缓冲空间取决于为某个流保留的带宽和网络上通信量的突发。

为了保证带宽和延迟这两个尺度基本上相互独立,可以把造成延迟的因素分解开来:排队延迟+传播延迟。排队延迟与网络结点的调度策略、缓冲的尺寸、端口的处理速度、链路的剩余带宽和流的通信量特征相关,我们可以从带宽上反映这些因素,因此,在延迟尺度中,只考虑传播延迟,它与两个结点间的距离成正比,瓶颈带宽是路径上所有链路中最小的剩余带宽。

瓶颈带宽和传播延迟反映了一些网络路径上的基本特征,可将其看成路径的“宽度”和“长度”。对于路径 P 上的各个链路,前者取最小(min)运算,后者取相加(add)运算。

3.2 源路由和逐跳路由

源路由主要用于网络诊断和特殊的策略路由,在源路由上,一个转发路由按要求在端系统计算并放在分组头上,路由器根据分组头中指示的路径来转发。因为路由计算是集中进行,所以非常灵活,可以使用任何一种算法或用一组算法来解决困难的问题;然而,源必须获得所有结点的路由信息,并且分组的头较大,当网络范围很大时,路由信息的获取和初始计算的开销很大。

逐跳路由是一种公共的通用路由形式,分组在每个结点逐跳转发,每个结点有路由表,表示所有的下跳,该表周期性地计算以适应路由的更新,其分组头与源路由相比较小,使用完全分布式的计算方法,只需知道相邻结点的路由信息。

这两种方法以及由它们相结合而派生的层次路由算法都可以用于 QoS 路由算法中。通常可以借助于深度优先或广度优先等算法,在(源、目的)结点对之间,找出满足 QoS 约束的可行路径集,再用优化算法选出较佳的路径,如对网络 G 、在源 S_1 和目的 S_2 之间选择一条路径 P , P 的目标函数为:带宽 $(P) \geq B$, 延迟 $(P) \leq D$, 可以采用如下算法:在所有链路中,除去带宽小于 B 的链路,在剩下的链路集合中找出延迟不大于 D 的路径(算法略)。在逐跳路由算法中,必须将目标函数分解成可以在每一跳中操作的对象,并且避免发生循环路由。

3.3 优化标准

在 QoS 路由算法中,还有两个重要的问题需要解决:一是如何在所有可行的路径中选择一个有效的路径,使得网络得到较高的吞吐率;二是在使用上述逐个

度量值遴选的过程中,虽然算法的复杂度是多项式的,但比通常的路由算法有高得多的计算复杂性,如何采用近似算法,既可减少计算开销,又不会牺牲太大的资源利用率。文[5]提出四种优化标准:

- 最小-带宽路径。选择一条路径,使使用的资源最少,即以带宽为主要路由尺度,保留最小的剩余带宽,如果有多条可行路径,保留最小剩余带宽的路径。

- 最宽-最短路径。选择一条路径,使跳数最少,即以跳数为主要路由尺度,如果有多条可行路径,选择具有最大剩余带宽的路径。

- 最短-最宽路径。选择一条路径,有最大剩余带宽,即以带宽为主要路由尺度,如果有多条可行路径,选择跳数最少的路径。

- 最短-延迟路径。选择一条路径,给定最小端到端延迟,即以带宽为主要路由尺度,如果有多条可行路径,选择跳数最少的并保留最大剩余带宽的路径。

上述最宽-最短路径性能较好,而最短-最宽路径在轻载情况下性能较好,重载情况下会导致尽力服务的通信量处于较低的吞吐率。

对于不同的特性的通信量,路由算法的性能衡量标准是不同的,如在确保服务情况下,要尽量保证网络有较低的呼叫阻塞率;而在尽力服务中服务情况下,要使网络的平均吞吐率尽量大。

4 资源预留和路由协议之间存在的问题

RSVP 现被 IETF 指定用于 Internet 上 IntServ 活动中通用的资源预留协议^[4],它在已有的 IPv4 和未来的 IPv6 上增加预留机制,依靠这些协议来交换数据和传递 RSVP 控制报文。它由接收方启动,因而解决了接收方端系统的差异问题,发送方不必了解太多接收方的情况;预留建立在流的基础上,由 IPv4 的地址字段或 IPv6 的流标识来指定;在端系统和路由器上用软状态指示预留的存在状态,路由器根据为该流建立的预留来调度数据分组的转发,便于刷新,预留的资源可以被多个发送者共享;提供三种不同的预留风格:固定过滤 FF,通配过滤 WF,共享过滤 SE,支持点一点、点一点和多点一点、多点多点的投递。

由于 RSVP 独立于路由协议,不能将 QoS 与路由紧密钩联,即链路的失败由 IP 处理,后继的数据可能会转发到无预留的链路和结点上,而原先的预留一直到有显示的撤除消息或刷新超时才释放,因而不能严格地确保 QoS。另一方面, QoS 路由是重要的网络资源管理组件之一,它支持集成服务网络的 QoS。因此,在预留建立和路由之间需要一种基本的交互,把所要传送的流的特性和预留请求沿所选择的数据分组路由设置下来。另外,当路由发生变化时,需要一种机制在新的路由上建立预留。

我们正在研究如何将具有 QoS 约束的路由机制与资源预留协议相结合,并借助于资源预留状态将所需的资源与数据传输路径绑定在一起,保证重要的实时应用在整个连接生命期,数据传输不受网上其他通

信量的影响。

RSVP 通过在其相应的消息中分别装载三种协议对象 SENDER TSPEC、ADSPEC 和 FLOWSPEC 来提供集成服务的通信量特性和 QoS 约定。

SENDER TSPEC 和 ADSPEC 由向接收方传递的 Path 消息运载。SENDER TSPEC 携带的通信量信息包含令牌桶规范(Token-Bucket)参数的设置^[3,7],包括:

r:令牌速率,以每秒 IP 数据报字节数为度量单位,范围为每秒 1—40TB;

b:令牌桶深度,以字节(byte)为度量单位,值的范围为 1—250GB;

p:通信量峰值速率,同 r;

m:最小分组,以字节(byte)为度量单位,包含应用数据和所有 IP 及 IP 以上层的协议头;

M:最大分组,以字节(byte)为度量单位,表示符合通信量规范的最大分组。

$m \leq M$,所有小于 m 的 IP 数据报均以 m 的尺寸处理;任何大于 M 的分组被认为不符合通信量的规范,因而不接受有 QoS 控制的服务。r、b、p、M 可用来计算排队延迟。

ADSPEC 是一个可选的对象,向接收方广播端到端的通信路径的特征,除了服务标志外,还包括几个重要的路径参数字段,如:

(1)可利用带宽 该字段是对路径上单个链路可利用带宽的估计,该参数每秒钟测试一次,按取小(min)原则更新。

(2)路径带宽和延时 这里的延迟等于光传播时间+分组处理时间,不包括任何队列等待延迟,它与特定的服务类型要求的队列延迟相加,给出应用的延迟。该值每一微秒测试一次,带宽取决于路径上最小的链路带宽。

(3)路径上的最大传输单元(PATH-MTU) 该值表明当前允许发送的最大分组,取中间结点所支持的最小的路径 MTU 为最终值。而 SENDER TSPEC 中的 M 仅表示发送方能够发送的最大分组尺寸。

(4)为确保服务设置的参数 用专门的确保服务参数字段填写端到端的 C_{tot} 、 D_{tot} 、 C_{sum} 和 D_{sum} 等。它们是为了计算最大端到端延迟而设,其中 C 表示与速率无关的误差,以字节为单位(如连续的分组由于尺寸的差异造成的发送时间误差);D 表示与速率相关的误差,以微秒为单位(如等待所分配的时隙的时间造成的延迟误差)。因此, C_{tot} 和 D_{tot} 分别表示端到端 C 和 D 合成值,即 C_{tot} 表示了数据流所经过路径分组模型和流模型的误差叠加, D_{tot} 表示了数据流所经过路径各个链路传播延时的叠加; C_{sum} 和 D_{sum} 分别表示局部的从上一个再整形点得到的值。

FLOWSPEC 由 Resv 消息运载,携带接收方要求网络预留的请求信息,这些信息根据从 Path 消息中的

SENDER TSPEC 和 ADSPEC 对象获得的关于通信量和路径的特征和自身接收条件而定。其通信量规范与 SENDER TSPEC 的类令牌桶参数相同。此外,在确保服务请求中,还要包括两个 RSPEC 参数:链路传输速率 R(也可以认为是带宽)和松弛值 S。R 以每秒 IP 数据报字节数为度量单位,S 的单位是微秒,它反映了所请求的延迟和能得到的延迟之间的差别。

$$D_{round_delay} = \begin{cases} \frac{(b-M)(p-R)}{R(p-r)} + \frac{(M+C_{tot})}{R} + D_{tot}, & \text{当 } p > R \geq r \\ \frac{(M+C_{tot})}{R} + D_{tot}, & \text{当 } R \geq p \geq r \end{cases}$$

由此,不难看出,可以利用 SENDER-TSPEC 和 ADSPEC 对通信量的描述生成合适的 QoS 路由约束条件,选择可行的路径,并相应地扩大结点和消息中关于路径状态的信息。若接收方请求确保服务,并给定预留的带宽和松弛度,就可以在可行路径中选择合适的路径作为数据传输路径,并将设置预留状态,否则选取一条占用资源最少的路径。

结束语 在网络发展的不同时期,路由问题都是网络传输的核心问题,它们根据不同时期网络的主要矛盾(或瓶颈),提出了不同的路由机制。早期的路由协议仅考虑单个网络特征值,以“可达”为根本目的,适应性算法也以解决拥塞为目的。而今天,面对计算机处理信息的多样化和网络系统通信服务的单一化的矛盾,希望路由协议能考虑多个网络特征值,满足上层递交任务的 QoS 约束,提出了基于 QoS 的路由等一系列概念。将 QoS 路由机制与资源预留有机地结合在一起,既可以避免路由受网络突发流量的影响而“抖动”,还可以提高预留的成功率。

参考文献

- 1 Braden R, et al. Integrated Services in the Internet Architecture: An Overview, RFC1533, July, 1994
- 2 Wang Z, Crowcroft T. Quality of Service Routing for Supporting Multimedia Application. IEEE Journal on Selected Areas in Communications, 1996, 14 (Sept.), 1288 ~ 1234
- 3 Crawley E, et al. A Framework for QoS-Based Routing in The Internet, RFC 2386, Aug. 1998
- 4 Braden R, et al. Resource Reservation Protocol(RSVP)-Version 1, Functional Specification, RFC2205, Sept. 1997
- 5 Guerin R, Blake S, Herzog S. Aggregating RSVP-based QoS Requests. Work in Progress, Internet Draft, Nov. 1997
- 6 Wroclawski J. Specification of Controlled-load Network Element Service, RFC2211, Sept. 1997
- 7 Shenker S, et al. Specification of Guaranteed Quality of Service, RFC2212, Sept. 1997