

互联网区域路由 QoS 机制研究^{*}

夏梦芹 易发胜 曾家智

(电子科技大学计算机科学与工程学院 成都610054)

摘要 本文讨论了在互联网上使用区域路由机制来满足实时应用和非实时应用不同 QoS 要求的方法。该机制综合了电路交换和分组交换方法的优点,能够解决综合服务中主干节点负载过重问题和区分服务中不能完全保证实时服务 QoS 问题,同时可以保留分组交换灵活充分利用网络资源的优点,是一种值得进行深度研究的互联网 QoS 机制。

关键词 服务质量,区域路由,资源预留,综合服务,区分服务

On Internet Partial Route QoS Mechanism Issue

XIA Meng-Qin YI Fa-Sheng ZENG Jia-Zhi

(School of Computer Science and Engineering, University of Electronic Science and Technology of China, Chengdu 610054)

Abstract This paper discusses the basic concept about the partial route mechanism, which can satisfy the QoS requests of both real-time applications and non-real-time applications. The partial route mechanism converges the merits of circuit switch mode and packet switch mode. Therefore it can solve existent problems of integrated service and differentiated service, while utilizes the network resources efficiently. As an enhanced QoS mechanism, the partial route mechanism is worthy more researching and may be suitable for the Internet.

Keywords Quality of service, Partial route, Resource reservation, Integrated service, Differentiated service

1 引言

互联网 QoS 的研究目标是有效地为用户提供端到端的服务质量控制或保证^[1]。QoS 意味着网络单元(例如应用程序,主机或路由器)能够在一定级别上确保它的业务流和服务要求得到满足。QoS 并没有创造带宽,只是根据应用程序的需求以及网络状况来管理带宽。

互联网业务的快速增长和多样化对计算机网络提出了更高的要求。对那些有带宽、延迟、抖动等特殊要求的应用来说,现有的“尽力而为”的服务显然是不够的^[2]。为了解决互联网 QoS 问题,IETF 已经提出了几种服务模型和机制,目前,有两个主流的体系结构:综合服务模型^[3](Intserv, Integrated Service)和区分服务模型^[4](Diffserv, Differentiated Service)。

在实现实时通信的策略上,Intserv 要求通信开始前建立连接,通信线路上的所有中间节点预留资源并保存链路状态,这样就能够保证满足实时 QoS 需求,但是主干节点管理操作负载过重,难以应用于大型网络^[5];Diffserv 将通信载荷进行分类,各中间节点根据载荷不同类型进行不同处理,保证了 QoS 要求高的载荷能够得到更好的服务,但是由于缺乏建立连接过程和资源预留,不能完全确保满足实时 QoS 需求^[6]。目前 Internet 网络 QoS 研究的重点是寻找既能适用于大规模网络,又能保证满足实时 QoS 要求的解决方案。

区域路由方法参考电路交换网络树型拓扑结构实时性强、能够保证服务质量的优点^[7],将互联网通信子网目前的不规则型拓扑结构规范为分级树型结合同级不规则型的拓扑结构。分组从源节点传送到目的节点的路径只有部分需要路由,从源节点到传送最上级节点和从接收最上级节点到目的节点

的路径都不需要路由,只有最上级节点之间的通路需要路由。这种方法可以保证实时 QoS 要求,也能够适用于大规模网络。

本文主要讨论在互联网上使用区域路由方法来满足 QoS 要求的相关问题,重点介绍区域路由方法的系统构架及其工作原理,最后对区域路由机制的特点进行了总结。

2 区域路由网络拓扑结构及其地址机制

2.1 网络拓扑结构

互联网通信子网目前采用点到点的不规则型拓扑结构。这种拓扑导致隶属于同一数据流的不同分组的传送途径有可能存在较大差异,难以保证满足特定的延迟和抖动等 QoS 要求;这种拓扑的另一个不足在于难以规范地进行描述和统计,从而导致难以对同一通信线路上的不同数据流进行聚合以简化化管理,主干节点的管理操作就很难避免地琐碎和复杂。

鉴于不规则网络拓扑结构的上述问题,我们考虑采用分级树型结合同级不规则型的网络拓扑结构来实现 Internet 的通信子网,其拓扑结构如图1所示。

图中同级节点(如 L1 中的 A1、B1、C1、D1 等节点)之间的虚线连接表示两个含义:一是同级节点没有全部画出,只用了4个节点来代表同级的数百或者数千个节点;二是同级各节点之间采用不规则拓扑结构,同级两个节点之间可能存在多条不同的通信线路(如 A1 可以通过直连和 C1 进行通信,也可以通过 B1 或者 D1 进行中转再和 C1 进行通信)。需要注意的是:由于图中每级只画出了4个节点,所以各节点之间是完整型的连接拓扑,但是实际上每级包含的节点数成百甚至上千,不能完全采用完整型拓扑,应该根据实际情况选择相应的不规则

^{*}国家自然科学基金项目,编号:69871005。夏梦芹 博士研究生,从事新型网络体系结构研究。易发胜 博士研究生,从事新型网络体系结构研究。曾家智 教授,博导,从事计算机网络与通信研究。

则拓扑结构。

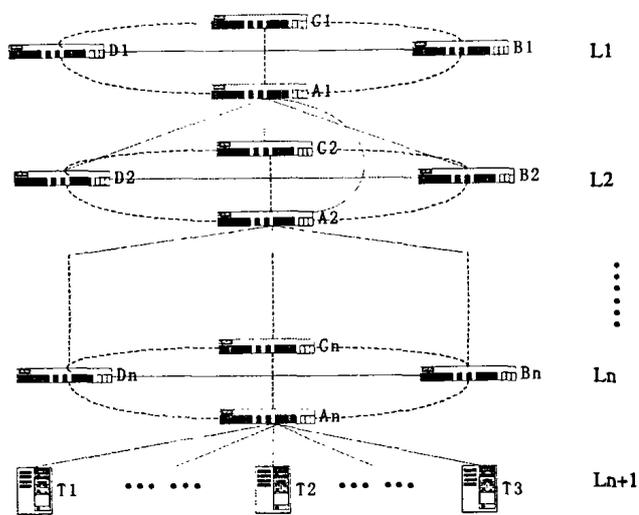


图1 区域路由网络拓扑结构

第1级节点除了和同级节点相连外,还和其所属的全部第2级节点相连,如图中的A1节点,除了和同级的B1、C1、D1等节点相连外,还和其所属的A2、B2、C2、D2等节点相连。当然,B1、C1、D1等节点也分别和其所属的第2级节点相连,不过在图中没有画出。

中间节点(除第1级节点和终端节点以外的节点)除了和同级节点相连外,还与其直接上级节点和全部直接下级节点相连。如图中的A2节点,除了和同级的B2、C2、D2等节点相连外,还和其直接上级节点A1节点相连,同时还和其全部直接下级节点A3、B3、C3、D3等节点相连。

终端节点的情况则根据实际情况的不同而有所不同:如果终端节点属于一个局域网,则部分终端节点之间存在直接互连关系,同时还要和其上级节点相连;如果终端节点采用拨号等点到点连接方式,则只和其上级节点相连,不与其它终端节点直接互连。

2.2 地址机制

在区域路由网络中,由于采用了规范的网络拓扑,使得两节点之间的通信线路是规范可知的,从而我们只需为每个网络节点分配一个网络地址即可,不需要像现在这样为连接不同网络的节点分配多个网络地址。

我们将网络地址划分为n个不同的地址段,每个地址段的值说明了该节点在相应级别在网络中所处的位置,这样一个节点的网络地址就能够准确表示该节点在网络中所处的位置,根据目标节点的网络地址就能够从网络中任意位置找到连接该节点的通信线路。

以图1中的各节点为例:

A1节点的网络地址为1.0.....0,B1节点的网络地址为2.0.....0;

A2节点的网络地址为1.1.0.....0,B2节点的网络地址为1.2.0.....0;

.....

An节点的网络地址为1.1.....1.0,Bn节点的网络地址为1.1.....2.0;

T1节点的网络地址为1.1.....1.1,T2节点的网络地址为1.1.....1.2。

一种可能的地址段的长度为8位,地址段的数量为16个。这样就将整个网络划分为16个级别,直接隶属于同一个上级

节点的节点最大数量为255个。

采用这种规范的网络地址结构还有一个好处,就是网络地址存在着无穷的可扩展性,如果将来出现网络地址不够的情况,可以将最后一个地址段的最低位全部设置为0才表示该级具体的终端物理设备,如果该位设置为1,则表示扩展的中间节点,其后再增加了一个地址段。因为该方法对前面地址段没有任何影响,所以不会影响到上级节点的操作;而该方法的重复使用也可以无穷无尽地满足对网络地址的需求。

3 区域路由 QoS 保障通信工作原理

用一个实时通信的具体例子来看看区域路由网络的QoS保障通信工作原理。

在图2中,终端节点A要和终端节点B进行实时通信,从A到其发送最上级节点D2和从接收最上级节点B2到B的通信线路唯一固定,从D2到B2的通信线路则需要在同级内通过路由选择来获得,区域路由因此得名。

在实时通信开始以前,需要建立实时连接,建立实时连接过程完成的主要工作是确定通信线路和资源预留。在本例中,从A到D2和从B2到B的各中间节点在建立实时连接过程中完成的工作是根据通信量大小在和上下级节点之间的通道上预留带宽,同时在本节点内部预留相应缓冲;D2到B2则需要根据路由协议选择合适的通路,然后在该通路上建立虚通道并预留资源。

实时连接建立以后就可以开始实时通信。从A到D2和从B2到B的各中间节点只需根据目的地址就可向下一节点进行转发,无需查找路由表项;D2和B2之间的中间节点则查找已经建立的虚通道路由表项,找到转发的下一节点,使用预留的资源完成满足QoS要求的转发操作。

实时通信结束以后,需要拆除连接释放预留的资源。可以由任一方发出拆除指令,沿传送通道传送一次后即可释放全部资源,如需保证可以确认一次。

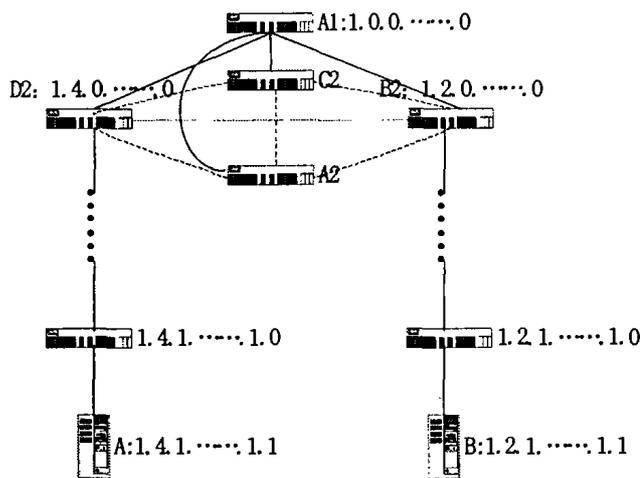


图2 区域路由网络工作原理实例图

非实时通信的情况更为简单,不需要建立连接预留资源的过程,只需要各中间节点根据目的地址和同级路由情况进行转发即可。

4 区域路由网络的特点

区域路由网络改变了目前Internet的网络拓扑和地址机制,从而能够更好地满足通信的QoS要求。其具体特点如下:

- (1)能够完全满足实时通信QoS的要求。分组从源节点

(下转第49页)

4.3.5 管理界面的设计 管理所需数据通过两个方面得到,一是充分利用 proc^[9]文件系统的相关数据,二是使用 iptables 命令进行一定内容过滤后获得相应信息,在此基础上编写 CGI^[14]程序,将所需信息通过 Web 服务传送。下面给出了 Web 管理功能部分界面,如图3与图4所示。

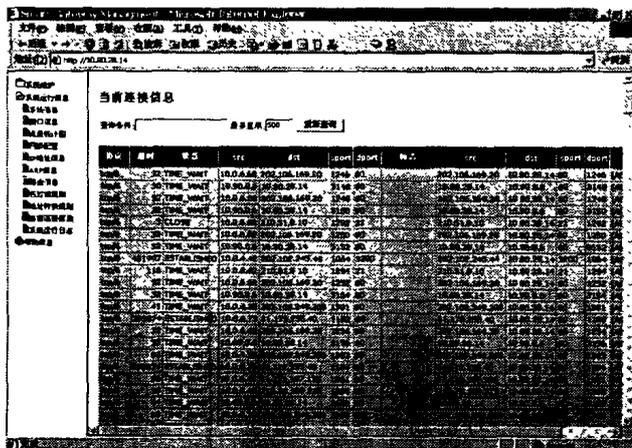


图4 当前连接信息

4.3.6 系统升级 升级前,将原来的文件进行备份,更换 Linux 文件系统的两个文件, kernel 文件和 initrd 文件后,重新启动即可,如果升级失败则可重新使用老版本,使得升级安全、方便。

5 系统性能分析

灵巧网关提供3个10/100M 自适应以太网接口,最多可以扩展到5个,适应当前网络速度不断提高,带宽不断增加的应用环境。灵巧网关的软件经过特别优化,足以应付各个接口满负荷工作。在设计时就高可靠性作为首要特性考虑,在硬件、软件层协同采用了多项技术保证系统的可靠性。经过实际运行,系统稳定,适合长期工作,最长连续工作纪录达到450

天。

由于每个 conntrack 需要占用一定内存,如果出现恶意攻击,使连接数过多时,可能导致系统无法正常工作,可以通过超时设置来防止这种情况发生。由于对 conntrack 是用 HASH 方法进行查找,如果恶意伪造数据包的报头,可能会使 HASH 出现冲突,导致系统性能下降。

结束语 灵巧网关是一个设计合理,功能齐全,性能稳定的系统,完全满足校园网的应用需求。同时由于其升级与管理非常方便,因此具有很强的实用性。根据配置的不同,灵巧网关可以作为路由器、地址转换设备、带宽分配器、包过滤器等来使用,具有高度的灵活性。还可以根据客户需求,在灵巧网关上开发新的功能。

参考文献

- 1 The netfilter framework in Linux 2.4. <http://www.gnumonks.org/papers/netfilter-lk2000/presentation.html>
- 2 <http://www.netfilter.org/documentation/HOWTO//packet-filtering-HOWTO.html>
- 3 <http://www.netfilter.org/documentation/HOWTO/de/netfilter-hacking-HOWTO.html>
- 4 <http://www.netfilter.org/documentation/HOWTO//NAT-HOWTO.html>
- 5 <http://gnumonks.org/ftp/pub/doc/conntrack+nat.html>
- 6 <http://cvs.gnumonks.org/netfilter-tools/doc/conntrack.sgml?rev=1.2>
- 7 <http://pc1.peanuts.gr.jp/~kei/Kernel-Snapshot/linux/net/ipv4/netfilter/>
- 8 毛德操,胡希明. Linux 内核源代码情景分析(上、下). 浙江大学出版社,2001
- 9 陈莉君. 深入分析 Linux 内核源代码. 人民邮电出版社,2002
- 10 Shah S. Linux 管理员指南. 机械工业出版社,2001
- 11 <http://www.burysbox.net/>
- 12 <http://www.acme.com/software/thttpd/>
- 13 <http://rpm.pbone.net/index.php3/stat/4/idpl/44124/com/dhcptd-2.1.i686.rpm.html>
- 14 Guelich S. CGI 编程-使用 Perl(第二版). 中国电力出版社,2001

(上接第39页)

到传送最上级节点和从接收最上级节点到目的节点的路径固定,最上级节点之间的通路在小范围内根据 QoS 路由建立虚通道,沿路所需资源全部进行预留,因而能够完全满足实时通信在延迟、抖动等方面的 QoS 要求。

(2)主干节点管理操作简单,适用于大规模网络。在建立实时通信连接时,可以根据所在级别实际通信统计情况设置虚通道的大小。如在最低级别上每次连接建立的虚通道可供1个实时通信应用使用,但在图2的 D2和 B2之间则可以考虑每次连接建立的虚通道可供10000个实时通信应用使用,由于虚通道的大小设置是基于实际应用统计,其实际利用率应该非常高;而且 D2下属节点和 B2下属节点间的实时通信都可以使用该通道,而 D2和 B2只需要管理1个而不是10000个虚通道,管理操作大为简化。

(3)保留分组转发机制灵活利用网络资源的特点。同级内部可以根据路由协议选择路由,能够充分利用网络资源。

(4)路由协议简单。路由局限在同级数百或者数千个节点的范围之内,可以采用最简单的路由协议, QOSPF,甚至增添 QoS 参数的 RIP 协议都可以满足要求。

(5)网络设备低成本高效率。非终端节点可以根据仅根据目的地址确定大部分分组的下一转发节点,仅当目的地址为同级其它节点的下属节点时才需要查找路由表选择转发路

径。同级节点数目有限,各级路由表表项都可以控制在很小规模以内,操作简单,转发效率大大提高,成本大幅下降。

(6)地址可无限扩展。前文已详细介绍,在此不再赘述。

小结 区域路由网络综合了电路交换和分组交换方法的优点,在保证实时应用服务质量的同时还可以灵活利用网络资源,同时能够解决实时应用下综合服务中主干节点负载过重问题和区分服务中不能完全保证实时服务 QoS 问题。其网络设备低成本高效率,特别适用于大规模网络,是一种适用于互联网的、值得进行深入研究和尝试的 QoS 实现方案。

参考文献

- 1 McDysan D. QoS & traffic management in IP & ATM networks [M]. New York: The McGraw-Hill Companies, 2000. 32~54
- 2 顾尚杰,薛质. 计算机通信网基础[M]. 北京:电子工业出版社, 2000. 27~28
- 3 RFC1633. Integrated Services in the Internet Architecture: an Overview[S]. 1994
- 4 RFC2475. An Architecture for Differentiated Services[S]. 1998
- 5 翁惠玉,刘芳,陈志英,等. Intserv/RSVP 的现状及其存在的问题 [J]. 数据通信,1999(4):4~7
- 6 李玮,卢燕飞. 区分服务 Diffserv 体系结构及其工作机制剖析[J]. 电信快报,2002(6):15~18
- 7 唐宝民. 电信网技术基础[M]. 北京:人民邮电出版社,2001. 18~20