

DiffServ网络中组播的相关问题及解决方案*

The Problems and Solving Methods of Multicasting in DiffServ Network

李海霞, 宋 玲

Li Haixia, Song Ling

(广西大学计算机与电子信息学院, 广西南宁 530004)

(School of Comp., Elec. and Info., Guangxi Univ., Nanning, Guangxi, 530004, China)

摘要: 介绍 DiffServ 体系结构和组播模型, 分析在 DiffServ 网络中应用组播时产生的相关问题, 探讨基于 DiffServ 模型的组播技术解决方案。

关键词: 组播 DiffServ 模型 服务质量

中图分类号: TP393.03 文献标识码: A 文章编号: 1002-7378(2005)S0-0061-04

Abstract This paper describes Differentiated Services mechanism and multicasting technology, the related problems which will arise when IP multicast is used in Diffserv networks that are illustrated and their solutions are presented.

Key words multicasting, differentiated services model, quality of service

近年来提供满意的网络应用服务质量 (Quality of Service, 简称 QoS) 已经成为 Internet 发展的重要方向。随着 IP 网络逐步由单一的数据传送网向数据、语音、图像等多媒体信息的综合传输网演化, 大多数的这些应用都是对时延敏感的业务, 传统的传输模式已无法满足多媒体业务和各种用户对网络应用服务质量的不同要求。

尽管互连网的带宽不断增长, 但如果缺乏合理的资源调度机制, 骨干网远不能满足 QoS 的需求。目前, 区分服务 (Differentiated Services, 简称 DiffServ) 模型是 IETF 提出的一种 QoS 技术, 它通过合理利用有限的网络资源保证用户申请的服务质量。显然, 组播与区分服务是两种互补的技术, 一方面组播试图节省网络带宽, 另一方面 DiffServ 旨在有效利用网络资源为用户提供满意的 QoS。本文对在 DiffServ 网络中应用组播时产生的相关问题以及解决方案进行了探讨。

1 DiffServ 体系结构

为了满足不同应用的服务质量要求, 网络必须提供服务质量支持。最初, 人们提出利用 IntServ / RSVP 模型对有 QoS 要求的数据流进行资源预留^[1]。虽然该模型具有很好的特性, 能够为每个应用提供其要求的服务质量, 但复杂的信令协议影响了 IntServ 的可扩展性, 因此在骨干网中无法得到应用。

DiffServ 模型把应用的服务质量要求分成几大类, 可以使用某一类服务来传输数据。支持区分服务的 IP 网络中, DiffServ 服务模型包含边缘和核心路由器两类^[2]。边缘路由器根据业务所需要的服务级别在 IP 包头的 TOS 字段中做出一定的标记, 不同的标记对应不同的服务等级, 网络边缘进行的分类、标识、整形、丢弃等操作使输入/输出流符合预先指定的传输调节协议 (Traffic Conditioning Agreement, 简称 TCA)。核心路由器根据标记来实现一组或若干组 PHB, PHB 是一个 DS 节点调度转发特定聚集流这一行为的外特征描述。IETF 区分服务工作组定义了两类 PHB: EF PHB 和 AF PHB。EF PHB 用于低延迟、低丢失率的实时应用, 如 IP phone。AF PHB 用于带宽保证的应用, 这类应用没有严格的延迟和抖动要求, IETF 定义了 4 类 AF PHB, 每一类具

收稿日期: 2005-08-19

作者简介: 李海霞 (1982-), 女, 山西人, 硕士研究生, 主要从事计算机网络研究。

* 广西科学基金项目 (桂科基 0342011) 和广西教育厅科研项目 (桂教科研 [2004]20号) 联合资助。

有三个丢弃优先级

2 组播模型

传统的 IP 模型包括主机模型和组播路由协议两部分。在主机模型中,一组主机由一个组播地址决定,通过组地址来进行服务的订阅和调度转发。组播路由协议维护组播树。当前许多组播应用都对服务质量提出了要求,设计支持 QoS 保证的组播路由算法一直是人们追求的目标,也是目前研究的热点问题。尽管 IP 组播技术出现的最初目标是缓解网络瓶颈、节省带宽,而并非 QoS 观念,但它的使用给新兴应用的带宽利用产生了显著影响,如大量的传输要求 QoS 保证多媒体信息。无论从网络运营商还是用户的角度来看,组播给网络资源的合理利用和 QoS 提供了技术支持。随着 DiffServ 网络研究的开展,DiffServ 区域的组播问题也亟待解决。然而在 DiffServ 区域支持组播的研究尚没有提上日程。DiffServ 已有的框架、结构都是基于单播的,所以加入组播势必会出现一些问题。

3 DiffServ网络中组播的相关问题

3.1 可扩展性

可扩展性问题是由于组播与差分服务在体系结构的本质冲突导致^[3]。组播树的每个源、组信息被保留在每个路由器的路由表中,根据路由表的入口,组播数据包被准确无误地传输到相应链路。在 DiffServ 网络中,所有的核心路由器简单化和独立化,单流信息只在网络边缘保存和处理,内部节点进行的简单调度转发与状态无关。

3.2 被忽视的预留子树 NRS

DiffServ 的资源必须先预留再使用,在 DiffServ 网络区,当一个新的接收者加入 IP 组播组时,相应的组播路由协议,如 PIM-SM, PIM-DM 或 DVMRP,负责将该组的组播树扩展出新的子树一直延伸到新成员。IP 组播的分组复制由新子树与原树接合点的路由进程完成,它将 DS 域的内容拷贝到每个被复制的 IP 分组的报头。这样,复制的分组会获得与原来的分组完全相同的 DSCP 值,并且接受与该组播组的输入分组同样的调度转发处理及区分服务,这并没有经过接纳控制、资源预留的必要过程。由于额外的资源被绕过区域管理(如 Bandwidth Brokers,简称 BB)的新增接收节点所消耗,其他已正确预留资源的接收者的服务质量将受到负面影响,甚至遭到破坏。

问题的关键在于^[4]:负责复制的接合点无法知道输出链路(接口)上复制的分组流是否已经进行了资源预留,因为在内部节点无法得到特定流的传输规格。接合点的位置通常有两种情况:

(1)接合点正好是 DiffServ 区域的出口边界路由器(图 1),则拷贝后的新增节点的分组会与享受同级服务的合法分组争夺资源。由于新增组播流使所需的资源超过了原来预定的数量,所以出口边界路由器的管制单元将对分组进行丢弃直到传输聚集符合传输协议为止。但是由于缺乏分类功能,路由器无法识别分组是否经过预留,使丢包具有任意性,导致对特定的流失去了任何服务保证。图 2 描述了两个 EF 聚集共享 40% 的带宽引起的 NRS 问题:假定超额流量将使用 30% 的链路带宽,图 2 则显示了 EF 聚集 70% 出口链路带宽被扼杀到原始预留的 40%,标记部分表明 EF 聚集引起的包丢失,而其他服务例如 AF、BE 不被影响。

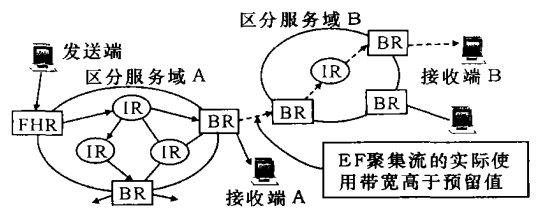


图 1 加入 B 之后引起的 NRS 问题

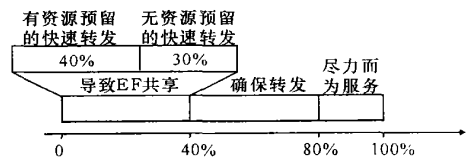


图 2 NRS 导致在边缘路由器之间共享带宽

(2)接合点在 DiffServ 区域内部(图 3),由于路由器没有计量和管制功能,所以无法识别超额的传输,导致对新增的组播流一味调度转发,使得新增加的未经预留的高级服务流能够与预留的聚集享受同样的服务,抢占了原来享受低一级服务的合法传输的资源。如图 4 描述当在 DS 域内加入 C 接收者时没有预留的 30% 的 EF 流抢占低一级的 AF 服务,AF 抢占 BE 预留资源,使 AF 产生 10% 的包丢失以及 BE 完全丢弃,显然预留给 AF、BE 的资源受到严重威胁。

这些问题的根源在于,新成员的传输绕过了接纳控制和资源预留的管理框架而直接享用了服务,或者具体地说,DSCP 的拷贝造成非法享用服务的后果。

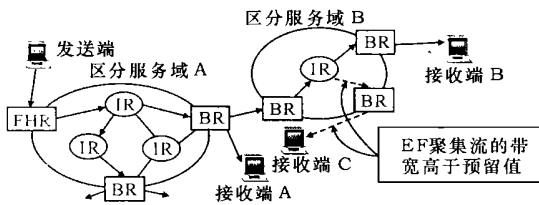


图3 加入C之后的NRS问题

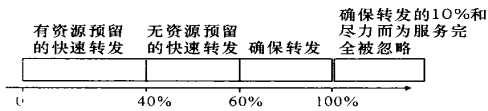


图4 在内部路由器之间共享带宽

3.3 异构组播组

同一组内的不同接收节点可希望获得不同的服务质量。对服务质量参数(比如AF服务中的预留带宽大小)要求的不同增加了问题的复杂性。异构性导致无法在边界处对报文进行标记。更严重的是,不同的服务种类可能没有可比性,无法取其较好者,使得可能在传送树的某些链路上需要同时支持这些服务。另外,因为任意一个新增接收节点都可以被嫁接到当前组播传送树的任何节点,包括内部路由器节点,那么在接合点复制分组时,必须根据接收端QoS要求设置DSCP,这与DiffServ区域内部节点不保留传输规格矛盾。

3.4 发送方任意动态改变

对于组播传输,接收组经常动态改变。组播允许组内任意成员作为发送者,而DiffServ是单向结构,这意味着如果有多个发送者同时发包的话,则其资源必须分别预留^[5]。

4 解决方案

4.1 NRS的解决方案

要解决NRS问题,就要防止任何未经资源预留就享用较高级服务的情况。当一个新的接收者要求加入组播组时,如果这两个树的接合点属于DiffServ区域,则该接合点把发向新成员的DS编码进行转换,使之对应PHB比默认PHB提供的服务级别还低,即比BE优先级还低的准尽量作好型LBE^[6]。这样可以制约新的传输流,即使对BE服务的资源也不可以抢占,维护了公平性。只有当新扩展的子树上资源预留的请求被管理实体认可之后,新成员才可以享受较高级别的服务。

上述方案需要子树生成点在调度转发复制分组时了解每个出口的传输所属的聚集,也即其DSCP解决的方法是,在每个组播的路由表条目中增加一

个包含DS域的字节,为每个输出链路增加一个DSCP值。这几乎没有额外的代价,复制时从组播路由表读取DSCP即可。新的组播组成员刚加入组播组时,其对应DSCP=LBE;当新成员的QoS要求经过准许之后,把DSCP改成对应的服务级别即可。

另外,还需要辅以一定的管理机制。对级别高于BE的服务,必须在DiffServ网络区进行接纳控制和资源预留。这也需要建立和更新区域边界节点上的传输规格,实现动态分配资源所需的信令机制和自动的接纳控制过程。总之,只有在一定的管理框架下,在组播子树上预留资源的接收者才能获得较好的区分服务,否则只能得到比BE还差的LBE服务。实现过程如图5所示。

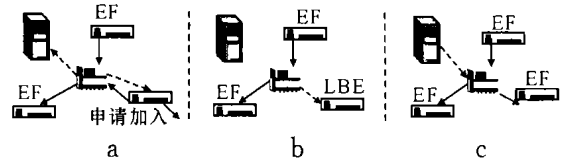


图5 NRS解决方案

a: 一个新的子树申请加入组播树; b: 未经过资源预留的新子树被分配DSCP=LBE; c: 通过BB进行资源预留允许得到更好的服务(比如:DSCP=EF)。

4.2 异构组播组的解决方案

本文针对不同的服务种类来讨论。解决方法与上述NSR问题的情况相同,只需在组播路由表中增加一个字节来存储对应接口的DSCP,并辅以管理措施。增加的这一个字节的前六位表示DSCP,第七位指示包的原始编码点是否被改变为一个特殊值(0表示被更新为更好的服务,1表示仍为LBE)。报文复制时,取出DSCP值复制到报文中,不同的接收者可以接收不同的服务级别。组播路由表项的结构为:(S,G,(OL1,DSCP1),(OL2,DSCP2),..., (OLn,DSCPn)),其中组的源地址为S,组地址为G,OLi标示一个输出接口,DSCPi为对应接口使用的服务级别。在没有预约的情况下,接收者仅允许获得LBE服务,所以在一个组播组中至少可能有两种服务存在。当两种服务类型可以比较优劣排序时,在上游链路中只需要传送较好即可。

4.3 任意发送方变动的策略

每个发送者要享用某种好的DiffServ服务质量时必须事先预留资源,而且无论组内是否有其他发送者已经预留了同种服务的资源。这就意味着为每个源组播树分别进行独立的资源预留策略。对于只需要BE服务的参加者可以在任何时间向组发送分组,无须任何其他机制。当使用上述方案时,并没有

增加 DS 体系结构的复杂性,而是在每个组播路由器的表中添加不同的 DSCP 值,没有减弱 DiffServ 的可扩展性

5 结束语

区分服务和组播共存时产生的问题主要是由于 DiffServ 体系结构的简化造成的,可以通过为单一传播和组播使用不同的 DS 编码值,即在组播路由表的每个输出链路条目中加入一项 DSCP,并辅以一定管理机制加以解决,这种方案实现简单,同时又保持 DiffServ 良好的可扩展性。总之,解决 DiffServ 网络区支持组播存在的固有缺陷,并最终实现在 DiffServ 网络区支持具有端到端 QoS 保证的组播传输,是目前研究的难点问题。

参考文献:

[1] Shenker S, et al. Specification of Guaranteed Quality of

Service. RFC 2212[M]. 1997.

- [2] Blake S, et al. An Architecture for Differentiated Services, RFC 2475[M]. 1998.
- [3] Bless R, Wehrle K. IP multicast in differentiated services (DS) networks [J]. Request for Comments, 2004 3754.
- [4] Striegel A, Manimaran G. A scalable approach to diffserv multicasting [J]. Proc of International Conference on Communications, Helsinki, 2001, 6: 2327-2331.
- [5] Striegel A, Manimaran G. A scalable protocol for member join/leave in diffserv multicast [J]. Proc of Local Computer Networks (LCN) Tampa Florida, 2001, 11: 395-404.
- [6] 林 闯,单志广,任丰原.计算机网络的服务质量 [M]. 北京:清华大学出版社,2004.

(责任编辑:黎贞崇)

(上接第 54 页)

于 2 个的数据。若待排序数据的总数为 n ,平衡二叉树树高为 i ,则有:

$$3 \times 2^{i-1} \leq n \leq 3 \times 2^i - 1$$

(2) 线程树中在同一路径上的两个线程存在逻辑依赖关系,它们在计算上不可能重叠;即便是父子关系的两个线程,它们之间也没有实质性的计算重叠,不在同一路径上的两个线程则可异步并行执行。

(3) 当待排序的数据很多时,算法将创建许多排序线程,线程创建将占用相当的时间开销。在逻辑并行度与物理并行度失配时,线程间的切换也将占用相当的时间开销。

4 结束语

多线程进行并行编程主要利用进程中线程的并行性来描述问题求解过程的并行性,用于完成子任务工作的各线程可并发执行。目前大多编程工具现均已引入线程机制,为并行计算的多线程实现提供

了途径。本文介绍快速排序异步并行算法的多线程实现,探索并行计算的多线程编程方向,对今后的进一步研究有一定的参考价值。

参考文献:

- [1] Clifford A. Shaffer. 数据结构与算法分析 (Java 版) [M]. 张 铭,刘晓丹译.北京:电子工业出版社,2002. 163-168.
- [2] 任 哲,李益民,车进辉. MFC Windows 应用程序设计 [M]. 北京:清华大学出版社,2004. 205-224.
- [3] Cameron Hughes, Tracey Hughes. C++ 面向对象多线程编程 [M]. 周良忠译.北京:人民邮电出版社,2003. 80-83.
- [4] 苏德富,钟 诚. 计算机算法设计与分析 [M]. 北京:电子工业出版社,2001. 128-130.

(责任编辑:黎贞崇)