

主动层次多目通信路由模型*

何丹 陈道蓄 谢立

(南京大学计算机软件新技术国家重点实验室 南京 210093)

(南京大学计算机科学与技术系 南京 210093)

E-mail: hdan@dislab.nju.edu.cn/cdx@nju.edu.cn

摘要 许多应用需要 IP 多目通信. 在 Internet 大规模应用 IP Multicast 时, 有效的路由是关键. 这样的多目路由协议必须是有效的、可伸缩的和增量可配置的, 但是传统的 Internet 路由对性能是不敏感的, 不能平衡负载和处理拥塞. 现有的大多数多目通信路由协议不仅负责数据转发, 还负责路由树的构造, 这给路由器带来了极大的复杂性, 而且协议的配置是手动的、费时费钱的工作. 该文提出一个主动层次式 Multicast 路由的体系结构, 采用主动网络技术将多目通信路由协议的数据转发和控制机制分开, 根据链路的状态信息用主动报文控制路由树的构造, 可以在 Internet 上自动增量配置大规模多目通信路由. 文章阐述了 AHMR(active hierarchical multicart routing) 的协议体系结构和控制协议, 并证明了协议的正确性和有效性.

关键词 多目通信, 路由协议, 主动网络, 主动路由, 协议.

中图法分类号 TP393

IP Multicast^[1]是一种高效的多目通信协议, 它利用 UDP 实现多目通信方式. 协议用 D 类地址作为抽象标识多目通信的组. 每个有兴趣接收消息的信宿通过 JOIN/LEAVE 加入/离开一个组. 由于 D 类地址是无结构的, 仅为一个标识, 这给路由协议寻址, 特别是在 Internet 上组的成员稀疏时的寻址带来了困难. 已有的路由协议如 DVMRP^[2], CBT^[3,4], PIM^[5,6]等已被提出来用于解决大规模 IP Multicast 应用的问题.

要在 Internet 上大规模应用 IP 多目通信, 有效的路由是关键. 设计这种多目通信路由协议面临着 3 个方面的问题: (1) 有效性. 路由协议应该有效地维护组成员之间的关系, 并且使得 Multicast 路由树路径尽可能走优化路径, 最小化建立路由的开销. 有效性要求路由协议间的控制报文数量尽可能少^[7]; (2) 伸缩性. 路由协议应该最小化协议控制报文的数量和网络状态, 不应该随着组成员数量的增大而线性增加; (3) 增量可配置. 即可以根据应用的需要增加 Multicast 路由器和路由算法, 而不必要求同时改变所有路由器和端点的协议. 但现有的 Internet 路由对性能是不敏感的, 路由策略的调整和协议的配置是手动的. Internet 是个黑匣子, 我们不知在什么地方丢包拥塞, 因此不能平衡负载. 当前的 Multicast 路由协议分两类: 基于距离向量路由和链路状态路由. 基于距离向量路由协议如 DVMRP(distance vector multicast routing protocol), 采用遍历/剪枝算法, 对稀疏组被证明是低效的; CBT(core based tree)路由协议要求选一个核(core), 核选择得好坏对性能的影响很大, 因此常常不是一种优化路由方案, 而且对核的选择是 NP 难的问题; PIM(protocol independent multicast)分为稠密(dense mode, 简称 DM)和稀疏(sparse mode, 简称 SM)两种工作方式. 但还有一些问题没有解决, 如两种工作方式之间的互操作、汇聚点(rendezvous point)的选择、共享树和最短路径树间的切换等. 基于链路状态路由协议如文献[8,9], 只是近似优化路由协议. 这些路由协议都不能同时满足上述 3 个方面的要求, 主要是因为已有的多目路由协议都参与数据转发和路由控制, 增加了路由器的复杂性; 而且所有协议都采用分布式算法, 缺乏全局路由知

* 本文研究得到国家 863 高科技项目基金(No. 863-306-ZT02-03-01)资助. 作者何丹, 1967 年生, 博士生, 主要研究领域为计算机网络, 分布/并行计算. 陈道蓄, 1952 年生, 教授, 博士生导师, 主要研究领域为计算机网络, 高性能分布/并行计算. 谢立, 1942 年生, 教授, 博士生导师, 主要研究领域为分布/并行计算.

本文通讯联系人: 何丹, 南京 210093, 南京大学计算机软件新技术国家重点实验室

本文 1999-02-09 收到原稿, 1999-06-10 收到修改稿

识,需要路由器间状态同步等,因此不可能达到最优化路由的目的;应用不能改变路由的策略,路由器不可编程.

我们提出的主动层次多目路由模型 AHMR(active hierarchical multicast routing)可以同时满足上述多目路由协议设计的 3 个方面的要求.首先,要使路由协议有效,最主要的是将协议包含的数据转发和路由控制操作分开,使得路由器只能作为数据转发的引擎,而路由计算由专门的容错网关进行.为了计算路由树,还要求知道网络的拓扑结构、组成员的位置以及采用性能敏感的链路状态路由算法^[8,9].为了增加可伸缩性,多目路由必须采取层次式结构. AHMR 模型的基本思想是将路由协议的数据转发操作与路由计算功能分开,用主动网络^[16-12]技术实现的主动路由协议代理称为 ARD(active routing daemon),用 SNMP(simple network manage protocol)协议获取网络的拓扑知识和链路状态,并修改路由表.每个域选择一个节点作为主动网关(active gateways,简称 AG),主动网关与 ARD 间用主动报文通信,用于域内(intra-domain)路由的计算.主动管理网关(active managing gateways,简称 AMG)负责组地址分配和域间(inter-domain)路由计算.由于主动网络报文不仅传输数据而且传输可执行的代码,因此采用主动网络技术实现 Multicast 路由协议可以动态配置路由算法,自动配置协议,减少路由器间交换报文的数量,增加增量可配置能力^[10].

1 AHMR 的协议体系结构

1.1 模型的假定

在进一步讨论 AHMR 模型之前,先对网络环境作些假定:

- (1) 每个主机和路由器都运行 IGMP 成员报告协议;
- (2) 每个自治系统都有一个 SNMP 网管,可以获得自治域内的静态拓扑结构;

(3) 自治系统中的路由器,我们假定能够运行 Java 虚拟机 JVM,因此可以配置主动节点^[10,11]运行基于 Java 的主动路由代理 ARD,这是我们获取动态网络拓扑和链路状态、修改路由表的基础.虽然传统的路由器都被看成是硬件和很少的配置软件,但未来的路由器如 Active Bridge^[12],则具有即时(on-the-fly)配置、开放的信令等功能,可以很容易地修改并转发路由表(forwarding table).很多厂商已将 SUN 公司的 Jini 集成到设备中,许多研究机构和公司也采用 Java 实现网管*.

以上的假定在当今的 Internet 世界是可以成立的.因为模型的思想是将路由协议的数据转发操作与路由计算功能分开,用 ARD 存取路由表,以构造多目通信路由树.

1.2 基本概念

主动节点(active node,简称 AN).参与 Multicast 路由树构造的主动网络节点^[10,11].在该模型的原型系统实现中,我们采用了 MIT 基于 Java 的主动网络协议开发工具 ANTS^{[11]*},按照 ANTS 的主动节点概念,即是能执行 ANTS 主动报文 Capsule 的网络节点.Capsule 可以传送数据,也可以传送程序代码.

主动路由代理(ARD).实现主动路由协议代理模块,目的是获得动态网络拓扑信息、报告链路状态、修改路由器的转发(forwarding)路由表、参与 Multicast 路由树的构造.

主动网关(AG).负责自治系统内的路由树计算的主动节点.其作用是参与 Multicast 组的创建和撤销操作;通过 SNMP 协议获取自治系统的网络拓扑信息,并为每个 Multicast 组选择一条正确的多目路由树.

主动管理网关(AMG).负责组管理主动路由节点.有 3 个方面的作用:(1)负责 Multicast 组地址的分配.目前,组地址的分配问题是一个研究热点但不是本文的重点.国际上尚未有一个标准化的 IETF(Internet engineering task force)文档论述组地址的分配.已有的组地址分配方案如会话目录(session directory)等存在着伸缩性方面的问题.本模型采用全局集中式分配方案,翻译一个组名如 News.Group 成 D 类地址,其功能如同 DNS.这种集中式控制的优点是协议简单;(2) AMG 参与多个自治域间的 Multicast 路由树的构造;(3) AMG 优化 Mul-

* 如 SUN 公司的 Jini, <http://www.javasoft.com>. Compu-Shack Software 公司的 Active SNMP, <http://www.cscare.com/activesnmp> 等.

* * ANTS 的源程序在 <http://www.sds.lcs.mit.edu/activeware>

icast 路由树. 每个 AMG 可以管理不同的主动网关(AG),不同的 AMG 间通过同步保持组地址的一致性.

1.3 体系结构和功能

1.3.1 AHMR 模型体系结构

AHMR 模型的体系结构如图 1 所示,表示两个自治域 AS1,AS2 的 Multicast 路由树的构造.其中自治域 AS1 和 AS2 分别选出一个路由器作为主动网关 AG1 和 AG2. 每个路由器 $R_i (i=1,2,\dots,n)$ 上运行 ARD 代理程序. 每个主动管理网管 AMG 可以管理 10~100 个 AG.

其中 G_1, G_2 表示 Multicast 组
 $R_i (i=1,2,3,4)$ 表示路由器
 $H_i (i=1,2,3,4,5,6,7)$ 表示主机
AS1, AS2 表示自治域
AG1, AG2 表示主动网关
AMG1, AMG2 表示主动管理网管
MG 表示自治域内的网管
BR1, BR2 为边界路由器
虚线表示虚连接

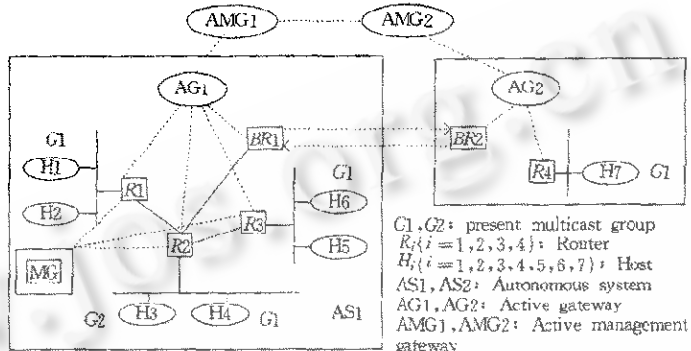


Fig. 1
图 1

G_1, G_2 : present multicast group
 $R_i (i=1,2,3,4)$: Router
 $H_i (i=1,2,3,4,5,6,7)$: Host
AS1, AS2: Autonomous system
AG1, AG2: Active gateway
AMG1, AMG2: Active management gateway
MG: Network management gateway
BR1, BR2: Border router
Dash line: virtual link

1.3.2 ARD 的功能

ARD 的功能如同 SNMP 代理. 每个路由节点都运行 ARD, 它实现 4 个功能: (1) ARD 在系统初启时, 首先从网管获得 ipRoutingTable, 并向子网中的每个主机发 ICMP 报文确认动态拓扑信息; (2) 它实现本地路由表的存取原语 GetRoute 和 UpdateRoute; 当 AG 的主动报文移动到 ARD 时, 调用这些原语修改路由表; (3) 存取本节点 MIB 链路状态信息, 并周期性地向 AG 报告; (4) 应用可以定制的路由转发策略, 每个应用可以对路由器编程. 这些机制和功能是该模型主动性的表现之一. 协议栈的结构如图 2 所示, ARD 的详细设计将另文给出.

1.3.3 AG 的功能

主动网关 AG 的主要作用是负责路由计算, 从而减轻传统路由器参与路由控制的任务. AG 用 SNMP 协议从本自治域的网管 MG 处获得全局的拓扑信息. 根据这些拓扑信息和 ARD 报告的动态路由信息和链路状态, AG 可以计算每一个要加入特定组成员的最优路径, 并用主动报文向路由器并发修改路由表. 这种集中路由计算, 简化了协议, 并且具有路由计算所需要的全局知识. 在图 1 中, 每个成员与组的关系有 3 种不同的情况, 以 H1, H7 为例: (1) 由 H1 创建一个新组; (2) H1 要加入的组在 AS1 域内; (3) H7 要加入的组不在域内.

(1) 由 H1 创建新组

当一个新的成员创建一个组时, 通过设置 JOIN($G_1, H_1, flag$) 的标志位 $flag=1$, 表明该成员是组 G_1 的第 1 个成员. 主机的 IGMP 向路由器 R1 报告有一个新的成员加入组 G_1 . 路由器向主动网关 AG 发出 JOIN($G_1, H_1, 1$) 的请求, 主动网关简单地把该请求转向主动管理网管 AMG1, AMG1 检查组分配记录. 若 AMG1 发现组 G_1 已存在, 则向主机报告组冲突. 若无 G_1 组存在, 此时 AMG1 可能尚未与其他 AMG 同步, 通过延迟到下一个同步点, 则为其分配一个 D 类地址. 一般 G_1 用组名来表示, 如 News. Group. AMG1 创建组后, 告诉 AG1 组创建成功, 将其登记在组地址分配表 GROUPALLOCATED 中, 并向其他 AMG 登记.

(2) H1 要加入的组在 AS1 域内

若一个主机要加入一个存在的组 G_1 , 如 IIG. 路由器 R3 收到 H6 的 IGMP 响应消息, 发给 AS1. 由于已分配的组 G_1 在 AS1 内, AS1 可以为 H6 寻找本域中到 Multicast 路由树的最优路径路由器, 如 R2. AG 并发向路由器 R2 和 R3 发送包含 UpdateRoute 路由 (S, G) 请求, 路由表的修改由 ARD 完成.



Fig. 2
图 2

(3) H7 要加入的组不在域内。

若主机 H7 要加入的组不在该自治系统 AS2 内,当路由器 R4 收到 H7 的 IGMP 请求报文后,向 AG2 发送 JOIN(G1,H2,0)请求。AG2 发现组 G1 不在自治系统 AS1 内,将请求送给 AMG2。AMG2 计算到 G1 的最近的自治域是 AS1,用消息(IP(AG1),G1)应答 AG2。由于不同自治系统间要通过 IP 隧道连接,AS2 选择一个边界路由器的 IP(BR2)和本自治域的标识 AS2 告诉 AS1,AS1 也选择一个边界路由器 BR1 的 IP 告诉 AS2。当两个自治域中的 AG 都收到对方的边界路由器 IP 地址后,由 AG1,AG2 分别向各自选出的边界路由器 BR1,BR2 发出建立 IP 隧道请求,并将主动报文定向 R2,R4,修改路由表。

1.3.4 AMG 的功能

如第 1.3.3 节所述,主动管理网关的一个功能是分配组地址和帮助建立域间(Inter-Domain)路由。AMG 的另一个功能是优化 Multicast 路由树。

当一个主机要加入非本地自治系统的 Multicast 组时,IP 隧道必须建立。因为 IP 隧道是虚链路,可以跨越几个自治系统,而中间的自治系统不对隧道报文解释。而且 AHMR 路由协议是将数据转发和路由树的构造分开,因此,隧道不必建立在自治域间的主动网关间,可以在 AS 中选择若干个路由器作为边界路由器,根据链路的延迟、拥塞等链路状态信息优化选择适当的边界路由器建立 IP 隧道。但一定时间后,这样建立的 IP 隧道会使得路由树不是最优的。

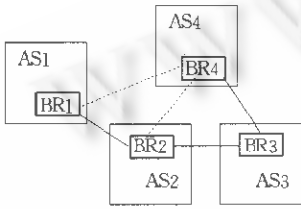


Fig. 3
图3

另一种情况是如图 1 所示,考虑到当 H7 加入组 G1 时,AS1 和 AS2 间建立了 IP 隧道。由于 AMG 选择的总是以最近的一个自治域作为跨域 IP 隧道的产生点,因此在一定时间后,路由树也不是优化的了。如图 3 所示,若信源在 AS4,则 AS1 的接收者距离信源最远。此时,采用“断枝”路由替换优化算法(make-then-break)^[13],其基本思想是先建立一棵优化路由树,然后将原树删除。该方法中优化路由树的缺点是在某一时刻路由树是有环的,部分接收者将收到两份报文。对有环的情况,在 Multicast 的 TTL 域减为 0 时,可以丢弃;双份报文的情况,可由应用程序过滤。通过定期计算优化路径,重新配置 Multicast 路由树,可以达到优化目的,如图 3 中虚线所示。

1.3.5 路由表管理

对于单目路由表,每个表项有一个三元组(dst, interface, age),当一个报文到达路由器时,通过查找路由表可以转发到 interface 上去,这对 Multicast 路由也应该成立^[2]。由于每个组可以有多个 interface 分支,所以路由表成为一个变长的记录。为了减少实现的复杂性,节省特别是共享路由树的路由表空间,可以采用文献[13]所述的称为 io-set 的方法,即每个组 G 的路由表只用项(G, {interface 表}, age)来表示。这种方案可以在主干网(backbone)上使用,以提高 Multicast 路由的伸缩性。对于最短路径树的情况,每个 Interface 上的路由表与 DVMRP^[2]相同,即用(S,G)表示。IP 隧道是虚连接,因此每个虚接口上都要设置(S,G)的路由表项。

为了增强路由器的强壮性,路由表可以组织成软状态(soft-state)。每个路由表设置一个超时值。当超时起作用后,刷新路由表项。从而可以解决因网络分段和主机失效引起的路由错误。使用软状态保存路由表给路由器带来了额外的开销,因为每个路由器的路由表要定时更新,刷新主机的 JOIN 请求。为了减轻主机刷新 JOIN 的负担,路由器可以向主机发 IGMP 请求消息,检查主机的存活(keepalive)情况,由主动网关 AG 定期向路由器查询子网成员的存活情况,刷新路由表。

1.3.6 可靠性技术

大家看到,AHMR 模型中每个自治域只有一个主动网关。主动网关保存自治域中 Multicast 组的全部路由状态,所以 AG 必须很可靠,否则 AG 出错以后,所有的路由状态都将丢失。为了解决这个问题,可以采用主副备份容错技术。

另一方面,我们没有特别提到在模型中路由器间的消息交换的可靠性机制,这是因为主动网络具有请求装入(on-demand loading)的报文可靠性机制^[11]。主动网关、主动管理网关以及 ARD 间的报文采用这种请求装入的主动报文,因此保证了协议报文的可靠传输。而主动网关 AG 与网关(MG)间的报文交换必须设置重传定时

器,保证其间信息的可靠性。

为了防止因网络分片和主机失效而引起路由错误,AHMR模型中的每个路由器上运行的ARD定期向每个儿子节点发ICMP报文,记录子节点的拓扑变化;主动网关AG定期向ARD发送获取GetRoutingTable报文,记录拓扑的改变。一旦发现拓扑改变,则重新计算Multicast路由,建立最优路径;若发现路径不可达,则向信源报告网络分片;若IP隧道破裂,则由边界路由器失效的自治域网关重新建立IP隧道。

2 路由间的报文交换

路由协议由一组报文和行为构成。为了进一步了解协议的内在关系,需要知道路由器之间的报文交换。以下网元间的报文是主动报文,内嵌程序,可以在相应的主动节点运行。

2.1 主机与AG间报文交换

主机与主动网关AG间交换两种报文,即JOIN和LEAVE。

JOIN报文是一个三元组,JOIN(组地址,用户标识,创建标志)。组地址是一个如同域名一样的字符串,若标志位为1表示创建一个组,此时创建只需指定一个组名,AMG会根据组地址分配表分配一个空闲组地址,标志位为0表示要加入一个已存在的组;用户标识决定于组的语义,对一个封闭的组,用户标识为用户名和访问控制表,安全性的审计过程可以由AG进行。当JOIN成功后,AG会向主机发出JOIN-ACK(状态,组地址)消息。状态域的值为{成功|失败},并包含失败的原因如组已存在、安全冲突等。

LEAVE报文只需一个变量即组地址G,LEAVE(G),主机只需要清除本机IGMP路由表项即可。路由器会周期性地询问主机,更新AG的JOIN状态。

2.2 主动网关AG与AMG间的报文交换

主动网关AG与AMG间的报文有JOIN,LEAVE,T.CREATE(隧道创建),T.DELETE(隧道删除)4种。

JOIN报文是一个四元组JOIN(组地址,创建标志,自治系统标识,主动网关IP)。组地址和创建标志的语义与第3.1节的JOIN相同;自治系统标识唯一标明该自治系统。主动网关的IP地址在产生IP隧道时需要用到。AMG应答AG的消息与第2.1节的JOIN-ACK相同。

LEAVE报文只含组地址,仅当自治系统AG内无此组时,AG会向AMG自动报告本域无组G了。

当主机加入非本地自治域时,AMG会向AG发送一个T.CREATE报文,该报文是个二元组T.CREATE({对方AGIP},{对方AS标识})。当AG收到T.CREATE报文后,向对方自治系统的主动网关AG交互,选择本域内的一个边界路由器与对方AS的边界路由器建立IP隧道。

为了确保主动网元与主动管理网元间消息的可靠传送,每个AG发出的每个消息都由主动网络ANTS的请求装入的可靠性机制保证,确保在有限的时间内完成AG与AMG间的消息交换。若当重传次数大于一个阈值N时($N=3$),可以让AG自动选择向另一个AMG发送消息,AMG的选择数亦可设置一个阈值。若候选的AMG数超过一个阈值,则由AG报告网络分片。这样可增加协议的健壮性。

2.3 主动管理网关AMG间的消息交换

主动管理网关的主要作用是用管理组地址和优化Multicast路由树。其地位是所有AMG是平等的,其内容是一致的,因此要维护主动管理网关间的同步。同步主要通过定期复制AMG中保存的表格,包括已分配的地址表GROUPALLOCATED=(DomainID,GroupIP)、域间IP隧道逻辑关系表TUNNELIP=(DomainID1,DomainID2)来完成。为了减少同步数据量,可以定时复制部分数据,其中每个表中新的软状态表项优先复制。

如果整个Internet配置的AMG太多,会使AMG间同步时间很长,增加了创建新组的延迟。解决的办法是将全部D类地址分区,像IPv4的地址一样。在一个区内,有一些AMG管理一段连续的组地址,减少组创建的延迟。

3 协议的正确性

我们的模型有两点假定:

- (1) 主动管理网关 AMG 间在有界时间内定期交换消息;
- (2) 主动网关 AG 和主动管理网关 AMG 的通信是有界的.

基于这两点,我们要证明定理 1 和推理 1.

定理 1. AHMR 协议是无死锁(safe)和无活锁(live)的.

证明:分两种情况.

- (1) AHMR 协议是无死锁的.

由主动网络实现的 AHMR 协议的每个消息都是与定时器相关的,在网络节点失效的情况下不会引起死锁.若 AG 与 AMG 间的消息交换失败,则当定时器超时次数超过一个阈值($N=3$)时,AG 会选择与另一个 AMG 通信;若与 M 个 AMG 通信都失败,则报告网络分片.若两个 AG 间通信 N 次失败后,AG 报告网络分片.每个主机、路由器和边界路由器失效后,都会引起 AG 重新计算路由,因此不会引起死锁.在组创建和成员加入一个组时,AG 与 AMG 间的通信以及 AG 与 AG 间的通信都可在有限的时间内完成,因此,也不会引起死锁.

- (2) AHMR 协议是无活锁的.

任何一个组的路由树都由 AG 集中计算出来.若一个组在一个自治系统内,则构成的树是无环的,因此不会活锁.若路由树横跨几个自治域,域间是用 IP 隧道连接的,中间路由器不会对 IP 隧道报文解释,虽然 AMG 在一定时间后优化域间路由,在优化其间可能会使路由树形成环,但是,由于 IP 报文中的 TTL 域值限制了报文在环上的传播跳数(hop),当 TTL 值是 0 时,报文被路由器抛弃,而且优化后的路由树是无环的,因此不会引起活锁. □

推理 1. AHMR 能在有界时间内构造正确的 Multicast 路由树.

由定理可知,协议的无死锁特性可以保证在有界时间内构造出路由树,而协议的无活锁特性保证了构造的路由树是正确的.

4 相关工作比较

我们分析协议控制报文的开销.

DVMRP^[2]采用遍历/剪枝算法,若有 N 个路由器,为每个信源建立最短距离路径需要 $O(N)$ 份消息;若有 S 个信源,则建立最短距离路径树的开销是 $O(SN)$.若 $N-K$ 个路由不在 Multicast 路由树上,则要发送 $O(N-K)$ 个剪枝报文,因此总的报文数是 $O(N)+O(SN)+O(N-K)=O(SN)$.

CBT^[3-4]采用核心共享树的方法,一个组成员 $S1$ 向核心发一个 JOIN 报文,每个有 Multicast 能力的路由器确定到核心的下一跳(next hop),并设置虚接口.当路由树上的一个路由器看到此消息时,则设置 $S1$ 到路由树的一条路径,并不再传播此消息,因此,建立路由树的开销除建立最短路径接口的开销 $O(N)$ 外,还有每个 JOIN 的开销.如果核心选择得好,可以保证 JOIN 的开销是常量^[4],因此,若有 S 个信源,总的开销是 $O(N)+O(S)$.

PIM^[5-6]协议的开销取决于其工作方式.对于稠密路由,采用类似于 DVMRP 的遍历/剪枝,其开销是 $O(SN)$.对于稀疏路由,其开销与 CBT 相同,为 $O(N)+O(S)$.

AHMR 协议,每个 JOIN 请求送到主动网关 AG 如同 CBT 送到核心,但是,AG 需要发送修改路由表的报文,因此,其开销比 CBT 多一个常量因子,但还是 $O(S)$ 数量级.总开销为设置最短路径接口的开销 $O(N)$ 加 JOIN 开销,为 $O(N)+O(S)$.

所以,AHMR 的有效性跟 CBT,PIM-SM 的有效性相当,其缺点是创建组时的延迟比其他路由方案长,但我们相信这一点不是严重的缺点,因为在 Internet 上,一旦组建立后会有一定的稳定性.

5 结论

AHMR 是一个有价值的 Multicast 路由模型,它简化了路由控制协议,采用主动网络使得路由构造具有更大的灵活性,可以减少控制报文的开销;采用集中控制计算路由,具有全局拓扑知识,比分布式路由算法更有效;用主动路由代理修改路由表,可将路由器的控制算法和数据转发分开,简化路由器的设计;主动管理网关不仅具

有地址分配的功能,而且增加了域间路由的伸缩性,可以对每个路由器编程增加定制路由和策略的目的;基于主动网络 ANTS 环境,可以动态地自动配置路由协议。本项目的研究正在进行之中,我们已用 ANTS 工具在 IBM/RS6000 和 Sun Ultra 上部分实现了 AHMR 原型。进一步的工作是完善原型的实现,研究基于这个模型的资源保留^[4]、基于 QoS 的路由算法,验证模型的可扩展性,改进模型的性能,减少组创建延迟等。

参考文献

- 1 Deering S E. RFC-1112: Host extension for IP multicasting. 1989
- 2 Deering S E. Multicast routing in Internetworks and extended LANs. *ACM Computer Communications Review*, 1998, 19(4):55~54
- 3 Ballardie A. Core based tree (CBT Version 2) multicast routing-protocol specification. RFC2189, 1997
- 4 Ballardie A, Francis P, Crowcroft J. Core based tree (CBT): an architecture for scalable inter-domain multicast routing. *ACM Communication Review*. 1993, 23(4):85~95
- 5 Deering S E, Estrin D *et al.* The PIM architecture for wide-area multicast routing. *IEEE/ACM Transactions on Networking*, 1996, 4(2):153~162
- 6 Estrin D, Farinacci D, Helmy A *et al.* Protocol independent multicast-spare mode (PIM-SM): protocol specification. RFC-2117, 1997
- 7 Rajagopalan B. Reliability and scaling issues in multicast communication. *ACM Computer Communications Review*, 1992, 22(4):188~198
- 8 Moy J. Multicast extensions to OSPF. RFC1584, Network Working Group, 1994
- 9 Garcia-Luna-Aceves J J, Spohn M. Scalable link-state Internet routing. In: Geoffrey Xie ed. *Proceedings of the 6th IEEE International Conference on Network Protocols (ICNP'98)*. New York: ACM Press, 1998
- 10 Tennenhouse D L, Smith J M, Sincoskie W D *et al.* A survey of active network research. *IEEE Communications Magazine*, 1997, 35(1):80~85
- 11 Wetherall D J, Gurtug J, Tennenhouse D L *et al.* ANTS: a toolkit for building and dynamically deploying network protocols. In: Campbell A T ed. *Proceedings of the 1st IEEE Conference on Open Architecture and Network Programming*. New York: ACM Press. 1998
- 12 Alexander D S. Active bridging. In: Biersack E ed. *Proceedings of ACM Special Interest Group on Data Communication (SIGCOMM'97) on Applications, Technologies, Architectures, and Protocols for Computer Communication*. New York: ACM Press, 1997
- 13 Aggarwal S, Paul S. A flexible protocol architecture for multi-party conferencing. *Communications of the ACM*. 1993, 35(12):37~53
- 14 Braden R, Zhang L, Berson S *et al.* Resource Reservation Protocol (RSVP) — Version 1: Functional Specification. RFC2205, Network Working Group, 1997

A Scheme of Active Hierarchical Multicast Routing

HE Dan CHEN Dao-xu XIE Li

(State Key Laboratory for Novel Software Technology Nanjing University Nanjing 210093)

(Department of Computer Science and Technology Nanjing University Nanjing 210093)

Abstract Many applications need IP multicast. In order to support large-scale multicast applications in Internet, effective multicast routing is crucial. These multicast routing protocols must possess efficacy, scalability and incremental deployment capability. But the existing multicast routing protocols are performance insensitive and cannot deal with traffic load balancing and congestion-oriented routing. Many multicast routing protocols not only take charge of data forwarding but also participate in routing control algorithms, which results in more complexity for routers. Furthermore, large-scale deployment of these multicast routing protocols in Internet is a manual, costly process. In this paper, the authors present a novel Active Hierarchical Multicast Routing scheme (AHMR) based on active network technologies by separating data forwarding and routing control protocol. Based on link-state protocol, it uses active packet to set up multicast routing as an efficient, automatically upgrading routing scheme for large-scale multicast applications in Internet. The architecture, control protocol, the proof of its correctness and its advantages over other schemes are demonstrated.

Key words Multicast, routing protocol, active network, active routing, protocol.