

代理型应用层组播体系结构研究

曹继军 苏金树

(国防科技大学计算机学院 长沙 410073)

摘要 组通信技术是未来互联网应用的关键支撑技术。目前,应用层组播是组通信技术研究中的热点,而代理型应用层组播体系结构已成为组播体系结构研究的重要方向之一。首先概述和比较了 IP 组播、主机型应用层组播和代理型应用层组播 3 种组播体系结构,接着重点研究和分析了代理型应用层组播体系结构的研究现状。以此为基础,进一步总结了代理型应用层组播体系结构设计中需要考虑的基本问题,并针对现有体系结构的不足,探讨了设计新的代理型应用层组播体系结构中值得关注的问题。最后对未来工作进行了展望。

关键词 IP 组播,应用层组播,主机型,代理型,组播体系结构

中图分类号 TP393 **文献标识码** A

Research on Proxy-based Application Layer Multicast Architecture

CAO Ji-jun SU Jin-shu

(School of Computer, National University of Defense Technology, Changsha 410073, China)

Abstract Group communication is a key technique for the future Internet applications. Nowadays, Application Layer Multicast (ALM) is an active research area in group communication technology, and Proxy-based ALM architecture is becoming one of the most important research directions in multicast architectures. Firstly, the three kinds of multicast architectures, i. e., IP multicast, Host-based ALM and Proxy-based ALM were surveyed and compared. Secondly, the six kinds of start-of-art Proxy-based ALM architectures, which are Overcast, RMX, AMcast/ALMI, Scattercast, OMNI and TOMA, were presented and analyzed. Thirdly, three basic design considerations for Proxy-based ALM architecture were summarized and some issues for the future Proxy-based ALM architecture, which are worth attention, were discussed. In the end, the prospect and the future work were discussed.

Keywords IP multicast, Application layer multicast, Host-based, Proxy-based, Multicast architecture

1 引言

相对于单播而言,组播由于借助中间节点进行分布处理,从而实现了分布式并发传输模式,因此成为一种高效的数据传输机制。IP 组播是最早提出的互联网组播方案,但是由于技术和市场等诸多方面原因,目前并未得到广泛部署。为了避开 IP 组播部署遇到的难题,相关研究提出了应用层组播。最早提出的应用层组播直接建立在网络层单播基础之上,组播功能完全由参与组播的用户主机完成,因此称为主机型应用层组播(Host-based ALM)。为了提高组播传输的稳定性和规模可扩展性,相关研究者进一步提出了基于代理的应用层组播(简称为代理型应用层组播,即 Proxy-based ALM)。在代理型应用层组播中,策略性地部署在互联网中的代理服务器成为构建组播覆盖网络的基础设施,用户主机通过访问代理服务器而获得组播服务。目前,代理型应用层组播体系结构已逐渐成为组播体系结构研究的热点之一,它也被认为是未来互联网组播的一种长期解决方案^[1]。

现有研究对主机型应用层组播进行了不少综述工作,但

还未出现综述代理型应用层组播体系结构发展的研究。为此,本文对近年来代理型应用层组播体系结构的最新研究成果进行了综述。论文第 2 节概述组播研究发展,重点研究和分析 3 种组播体系结构的基本思想、研究现状和优缺点等。第 3 节调研 6 种主要的代理型应用层组播体系结构,并对它们进行了比较与分析。第 4 节主要探讨如何设计代理型应用层组播体系结构的问题。最后给出全文总结和 research 前景展望。

2 组播发展技术概述

2.1 IP 组播

1988 年,Deering 博士首次提出 IP 组播,即将组播功能增加到网络层^[2]。IETF RFC 1112 描述和定义了 IP 组播的业务提供方式和形式,成为 IP 组播的业务模型标准。主机组(Host Group)是 IP 组播概念的核心,它由多个有“共同兴趣”的主机组成并由 IP 组播地址标识,以组地址为目的地址的组播报文被以“尽力转发”方式传输给主机组中的各个成员主机。组播路由和转发控制功能由路由器承担。

收稿日期:2008-07-15 本文受 973 国家重点基础研究发展规划(2009CB320503),国家自然科学基金(906004006)资助。

曹继军(1979-),男,博士生,主要研究方向为高性能路由器技术和互联网组播技术等,E-mail: caojijun@nudt.edu.cn;苏金树(1962-),男,博士,教授,博士生导师,CCF 高级会员,主要研究方向为计算机网络和信息安全等。

历经近 30 年的发展,IP 组播已经形成较为完善的组播协议体系。组播路由协议是该协议体系的核心,研究者已经提出多种组播路由协议,例如 DVMRP, MOSPF, PIM-SM, PIM-DS, CBT, OCBT 和 BGMP 等。互联网上的 IP 组播部署尝试始于 1992 年建立第一个组播试验网——MBone。当前,IP 组播使用的协议体系为 MBGP/MSDP/PIM-SM,而其长期解决方案将可能以 MASC 和 MGMP 为基础。虽然目前的大多数商用路由器都实现了 PIM-SM 和 IGMP 协议,因而具备支持 IP 组播业务的能力,但就互联网 IP 组播应用现状而言,IP 组播远未取得预期的成功。一方面,在互联网中,只有很少的路由器启动了 IP 组播功能,至今还未开展全互联网范围的组播业务;另一方面,基于 IP 组播的上层应用也屈指可数。CAIDA 利用 MANTRA^[3]在美国骨干网上的检测数据显示,IP 组播数据流量呈现逐年下降趋势。国内目前也只有 CERNET 开展了 IP 组播业务。总之,IP 组播当前并未得到普遍的商业应用,而且相对于同时产生的 WWW 等技术,IP 组播的发展显得非常缓慢。

IP 组播发展困境引起了人们对制约其部署的因素进行深入思考。文献[4]从服务模式和体系结构等方面总结了阻碍 IP 组播业务发展的主要因素:体系结构缺乏可扩展性、模型难以支持有效的管理和控制机制、破坏了路由器迁移模式、缺乏清晰的商业费用模式、缺乏完善的安全机制等。Kevin Almeroth 在 2007 年召开的组播专题会议 P2PM 2007 中阐述了自己的观点^[5],他认为经济学因素是阻碍 IP 组播部署的最重要因之一。文献[6]将影响 IP 组播部署的因素概括为两方面,即可行性(Feasibility)和收益性(Desirability)。

2.2 主机型应用层组播

面对互联网 IP 组播发展困境,研究者开始反思 IP 组播体系结构自身的问题,提出主机型应用层组播。在主机型应用层组播中,成员主机之间建立以 IP 网络为基础的覆盖网络,组播路由控制和数据转发均由成员主机完成。主机型应用层组播思想的提出和发展增进了人们对端到端思想(End-to-end Arguments)^[7]的认识。

自从 Francis^[8]和 Zhang^[9]在 1997 年和 1998 年各自独立地提出主机型应用层组播思想以来,国内外多家研究机构开展了主机型应用层组播研究项目。研究者提出的主机型应用层组播协议包括:ESM(Narada)^[9],ALMI^[10],Yoid^[8],NICE^[11],Bayeux^[12],BTP^[13],CAN-Multicast^[14],Scribe^[15],Bullet^[16],SplitStream^[17],HostCast^[18],oStream^[19]等。目前成功应用的典型系统或中间件包括:ESM,RelayCast,Emma,AnySee,TVNet,PPLive 等。最早的应用层组播系统 ESM 曾成功应用于 SIGCOMM 2002/2003,OSP 2003 和 NOSSDAV 2004 等著名国际会议。总之,主机型应用层组播从理论研究到实践都取得了成功。

相对于 IP 组播而言,主机型应用层组播的优势主要体现在两方面:(1)易部署性。主机型应用层组播以 IP 单播服务为基础,主机型应用层组播应用系统的部署不需要路由器和主机操作系统等提供额外的支持,因此较 IP 组播易于部署。(2)组播高层需求的易实现性。由于主机型应用层组播中主机间的数据传输能够根据应用需求选择使用传输协议,因而便于解决 IP 组播面临的可靠组播传输和组播拥塞控制等难题。

当然,主机型应用层组播也存在局限性和不足。首先,由

于主机对网络层了解有限,当前的大部分主机型应用层组播协议通过探测获得某些网络性能参数(如带宽和延迟),这增加组播控制通信开销。其次,主机型应用层组播对网络资源的利用效率也比 IP 组播低许多。再者,主机型应用层组播中单个主机的动态性(如频繁加入或离开)可能会对组播会话的稳定性和可靠性造成严重影响。

2.3 代理型应用层组播

以主机型应用层组播思想为基础,为了提高组播传输稳定性和规模可扩展性,研究者提出代理型应用层组播。通过在互联网中部署专门的代理服务器以构建骨干覆盖网,组播树在代理之间产生,用户主机通过与代理节点之间的单播或本地组播(Native Multicast)通信而获得组播服务。

目前,对代理型应用层组播的研究主要集中于提出可行的体系结构并给出该体系结构下的组播覆盖网拓扑构造方法。典型的代理型应用层组播体系结构有 Overcast^[20],RMX^[21],AMcast^[22],Scattercast^[23],OMNI^[24]和 TOMA^[25]等。这 6 种代理型应用层组播体系结构是本文第 2 节研究的重点。

从某种意义上而言,代理型应用层组播是对 IP 组播和主机型应用层组播的折衷,因此它结合了这两种组播方法的优点。其优点主要体现在:(1)骨干覆盖网络和本地簇有限的节点数目使成员关系和组播树的管理更有效,同时降低了控制消息交换的范围。(2)在主机型应用层组播中,每个组播覆盖网络只支持一个组或一种应用,而在代理型应用层组播中,骨干覆盖网可以同时为多个组或多种应用提供服务。探测和维护骨干覆盖网络所产生的控制开销不会受到同时支持的组数目的影响。因此,当存在大量组时,代理型应用层组播在控制开销方面的性能将优于主机型应用层组播。(3)主机型应用层组播中提供组播服务的都是普通主机,它们随时可能会因为人为原因或机器的不稳定性而突然离开组播组,这将对组播数据传输质量造成很大影响,因此组播系统的稳定性较差。代理型应用层组播从一定程度上改善了组播系统的稳定性。(4)代理型应用层组播将组播路由维护和应用开发分离,组播应用开发者只需要调用相关编程接口就可以使用代理型应用层组播功能,代理型应用层组播服务提供商负责代理间的组播路由维护。因而,与普通的主机型应用层组播相比,它大大缩短了应用开发周期。代理型应用层组播的缺点主要是需要部署组播代理服务器。目前,国内外对代理型应用层组播体系结构的研究处于起步阶段。

2.4 小结

IP 组播、主机型应用层组播和代理型应用层组播是目前 3 种主流的互联网组播解决方案,图 1 直观比较了这 3 种组播体系结构的特征。

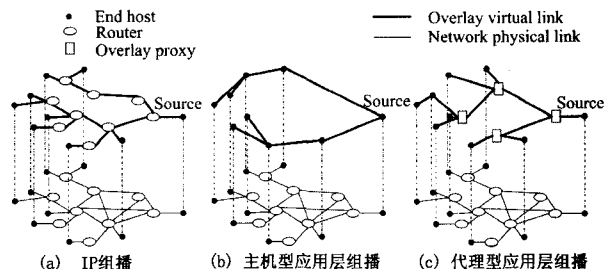


图 1 3 种组播体系结构特征的比较

表 1 定性比较了 3 种组播体系结构在易部署性、组播效率、控制开销、组播稳定性和组播应用开发的难易程度等方面的表现。在该表中, ALM-H 表示主机型应用层组播, 而 ALM-P 表示代理型应用层组播。显然, 代理型应用层组播是其余两种组播方式的折衷。

表 1 组播体系结构比较

Metrics	IP	ALM-H	ALM-P
Ease of Deployment	low	high	medium
Multicast Efficiency	high	low	medium
Control Overhead	low	high	medium
Multicast Stability	high	low	high
Ease of Application Development	high	high	low

由于代理服务器比普通主机具有更高的处理能力、稳定性和安全性, 而一个代理服务器同时可以服务多个客户主机。因此, 代理型应用层组播更适合于构建通用的组播服务平台, 并能提供更为强大、灵活和可扩展的组播传输服务。

3 代理型应用层组播体系结构研究现状

本节主要调研了 6 种典型的代理型应用层组播体系结构, 并对其进行了分析与比较。这 6 种体系结构提出的时间为 2001—2005 年, 下面将按照提出的先后顺序对其进行介绍。

3.1 Overcast 体系结构

Overcast^[20] 提供高扩展性、可靠单源组播服务, 因此特别适合于视频点播或直播等应用。Overcast 应用层组播系统包括数据源、内部 Overcast 节点(具有高存储容量的服务器)和标准的 HTTP 客户(用户主机)。Overcast 使用简单的组播树构建协议将内部节点组织成以数据源为根节点的组播树。每个组播组提供的数据被复制并存放于各个 Overcast 节点, 每个 Overcast 节点可以为多个组提供服务。

组播树的根节点不但是数据源, 而且也是组的管理中心。当用户节点加入组时, 它们首先向根节点发送加入请求消息, 然后被重定向到适当的 Overcast 节点。组播树的构建和维护采用一种自组织算法, 该算法通过最大化从根节点到 Overcast 节点带宽启发式对网络带宽进行优化。初始时, 新加入节点选择根节点作为父节点。经过多次动态调整后, 该节点将会处于组播树的固定位置。每一次调整过程都需要探测该节点到父节点和到父节点的其余子节点的带宽。子节点选择其父节点的最优子节点作为新的父节点。每个节点维护一个祖先节点信息列表, 这样, 当其某个祖先节点失效时, 它可以重新加入组播树。该体系结构设计了 UP/DOWN 协议, 用来维护节点的状态信息并管理节点的加入和离开行为。网络中的每个节点也维护着其所有子孙节点的信息列表。

3.2 RMX 体系结构

RMX(Reliable Multicast proXies)^[21] 体系结构如图 2 所示, 它使用 TCP 协议实现组播数据传输, 因而具有拥塞控制能力。该体系结构将参与组播的用户节点分为多个组播数据组 DG(Multicast Data Groups), 每个 DG 内的用户节点性能相近, 而不同 DG 内的节点性能差异较大。这种分而治之的策略通过将组播用户节点分类为多个异构的集群, 从而将大范围网络内的异构性问题简化为各个同构 DG 内的速率适应问题和跨 DG 间的点到点速率调整问题。每个 DG 包括一个代理服务器节点 RMX。用户通过订阅到本地 DG 而加入

RMX 会话。RMX 之间通过 TCP 单播构成组播树。RMX 体系结构将应用级的智能和语义整合进了数据转发服务。

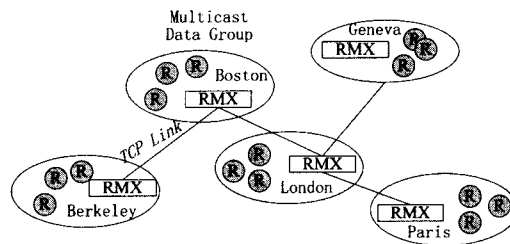


图 2 RMX 体系结构

RMX 体系结构适应于对实时性和可靠性要求较高的组播应用。RMX 使用应用级命名机制对数据对象进行命名, 这将能够标识 RMX 覆盖网中传输的数据对象并记录丢失的数据分组。为了解决异构节点间的可靠数据传输问题, 作者提出了在 RMX 体系结构中用语义可靠性概念代替传统的数据可靠性概念。语义可靠性允许 RMX 使用应用级知识动态转化数据内容以适应速率变化。语义可靠性概念和应用级封装 ALF(Application Level Framing)^[26] 机制的引入使得 RMX 能够适应于应用的需求并在数据转发前进行数据转化。

3.3 AMcast/ALMI 体系结构

AMcast/ALMI(Advanced Multicast Infrastructure)^[22] 体系结构如图 3 所示。在该体系结构中, 由多个 AMcast 服务器组成的服务器集群分布于各个网络域内。终端用户通过 IP 单播连接到本地 AMcast 服务器而接收或发送组播数据。多个 AMcast 服务器之间构成组播树并向终端用户提供组播服务, 新加入的 AMcast 服务器通过访问骨干网络并加入到组播树中。在 AMcast 服务器中, 本地网络接口和骨干网访问接口是分离的。

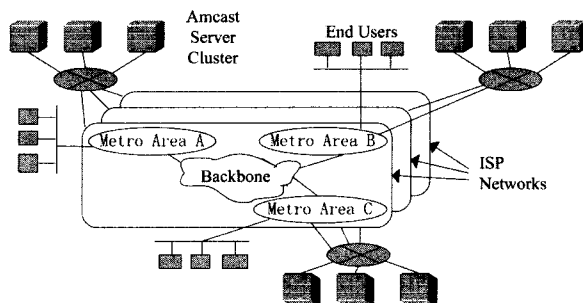


图 3 AMcast/ALMI 体系结构

AMcast 服务器有 3 种差错控制和流控机制相结合的方式: Hop-by-Hop 的 TCP 连接、层次型数据编码和基于速率的流控。在该体系结构中, 数据路径分成 4 个层次: 组播会话层、数据复制和转发层、传输层、传统 IP 单播和转发功能层。交换控制模块 SC(Switch Control)在 AMcast 服务器和输出接口之间建立互连, 它具有单播和组播路由控制功能。SC 还周期性地从各个服务器收集负载信息, 从而将新的会话请求分配给负载最轻的服务器。通过与服务器通信, SC 为每个组播会话建立转发表。

关于如何构造 AMcast 服务器间组播树的问题, S. Shi 等人在文献[22]中提出了两种路由算法——MDDBST 和 BDRBST, 它们分别以网络直径和负载均衡度作为优化目标。S. Shi 等人在文献[27, 28]中进一步研究了该代理型组播服务模型下的端到端延迟和接口带宽利用率的优化问题。

3.4 Scattercast 体系结构

Scattercast^[23] 体系结构如图 4 所示,组播服务由 Scattercast 代理 SCX(ScatterCast proXies)间协作提供。在该体系结构中,不同的组播组用类似于 URL 的唯一名称进行标识。每个成员都通过网络距离最近的 SCX 加入到组播组。每个 SCX 可以同时为多个用户节点提供组播传输服务。成员节点可以采用组播或单播方式与 SCX 通信。SCX 之间的单播数据传输可以根据组播组的传输要求而采用不同的传输协议(UDP 和 TCP 等)。Scattercast 还可以根据一定的语义对组成员分类,并对组播传输内容作自适应调整,从而有效解决组播会话的异构性问题,这类似于 RMX 体系结构使用的速率调整机制。

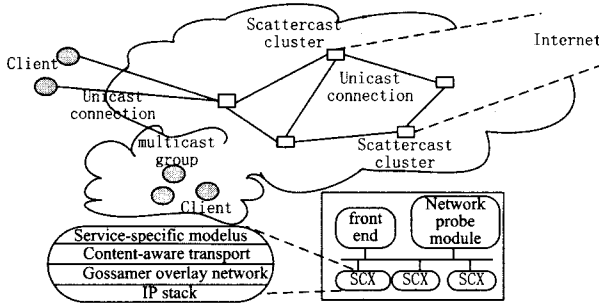


图 4 Scattercast 体系结构

Scattercast 组播由两个层次上进行通信构成,即组播成员和 SCX 的通信以及 SCX 之间的通信。当节点加入某个组播会话时,首先定位到网络距离最近的 SCX 节点,然后向该 SCX 节点发送请求加入消息。成员节点和 SCX 之间通过周期性交换 Keep-alive 消息保持连接。将离开组播会话的成员向 SCX 节点发送离开请求消息。SCX 节点可以根据 Keep-alive 消息的接收情况判断成员节点是否失效。当一个 SCX 节点的所有成员都离开组播会话时,该 SCX 节点将向其它 SCX 节点通告离开消息并离开该组。SCX 节点之间首先通过 Gossamer 协议构造 Mesh 网,然后在 Mesh 网络基础上运行分布式的距离向量路由协议而构建组播树。当 SCX 节点加入某个组时,它使用节点发现协议来发现其它 SCX 节点并从中选择部分节点作为自己的邻居节点。

3.5 OMNI 体系结构

图 5 所示,OMNI(Overlay Multicast Network Infrastructure)^[24]是一种具有两层结构的代理型应用层组播体系结构,它适应于大规模流媒体应用。OMNI 主要由提供数据分发服务的多个 MSN(Multicast Service Nodes)和参与应用层组播的多个用户主机构成。用户主机通过从单个 MSN 订阅组播服务而接收组播数据。MSN 之间通过运行分布式协议而组织成组播数据传输的骨干覆盖网络。从 MSN 到其各个客户节点的数据传输路径可以采用 IP 组播、应用层组播或顺序直接单播等方式。

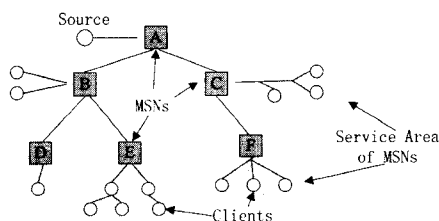


图 5 OMNI 体系结构

针对流媒体应用对组播传输的实时性提出的要求,OMNI 覆盖网拓扑构造算法以建立节点度受限的最小平均延迟或最小最大延迟生成树为目标。构造 OMNI 覆盖网的过程采用全分布式的方法,因此它更适应于网络状况和用户节点高度动态变化的网络环境。每个 MSN 维护着组播树中所有邻居和祖先节点的状态信息。初始化时,每个 MSN 向根 MSN 发送加入请求信息(包含到根 MSN 的延迟和度限制信息)。根 MSN 收集其它各个 MSN 的加入请求消息后,使用简单的集中式算法产生最初的组播树。组播树中的每个节点周期性地执行本地节点位置转换操作,并且基于模拟退火算法概率性地执行非本地节点位置转换操作。新 MSN 节点加入时,首先向根 MSN 节点发送加入请求,每个请求都对 3 种可能的加入选择,即直接加入到根节点、分割边加入或重新加入到下一级。MSN 节点离开时,如果是叶节点则直接离开组播树,否则,该 MSN 节点的某个子节点将被提升到它的位置。

3.6 TOMA 体系结构

TOMA(Two-tier Overlay Multicast Architecture)^[25]也是一种具有两层结构的代理型应用层组播体系结构。在该体系结构中,分布在网络中的各个服务代理构成各自的 MSON(Multicast Service Overlay Network)服务域。MSON 的构建依赖于 MSON 提供者、网络服务提供者和组播组创建者之间的商业关系。因此 TOMA 体系结构设计着重解决 MSON 提供者面临的问题,例如可扩展的 MSON 管理和高效的成簇算法等。TOMA 结构使用 OLAMP(OverLay Aggregated Multicast Protocol)协议管理各个 MSON,该协议采用聚合组播(Aggregated Multicast)^[29]方法使多个组播组共享一棵组播树。用户节点通过连接到成员代理而订阅 MSON 提供的组播传输服务。各个成员代理将本地用户组织成应用层组播树。TOMA 体系结构如图 6 所示。

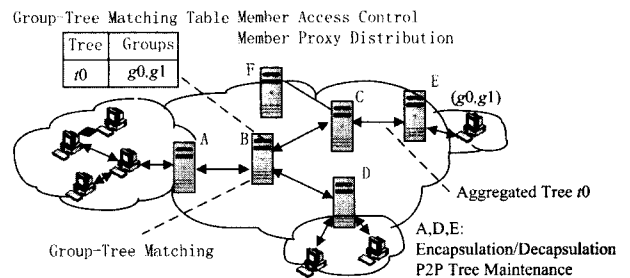


图 6 TOMA 体系结构

整个聚合树结构由分布在各个代理所维护的组播路由表构成。首先,终端用户在获得成员代理地址后,向其发送加入请求消息。然后,成员代理在本地建立一棵应用层组播树,同时将加入请求消息转发给管理此组播组的主机代理。接收到加入请求的主机代理更新相应的匹配表并执行组播树匹配算法,从而将请求加入的组映射到合适的聚合树。当成员代理发现无终端用户相连时,它向相应的主机代理发送离开消息,这将触发从组播树到聚合树(Group-to-Tree)的匹配过程。如果无其它成员代理属于该组,则主机代理将删除表中相应项。组播树为空时,将进一步导致删除组播树的操作。主机代理可以随机选择,也可利用从组播组到主机(Group-to-Host)的哈希函数获得。当且仅当簇首次加入或完全离开一个组,从而导致成员代理向相应主机代理发送加入或离开请求时,聚合树结构才会受到影响。而单个终端用户加入和离开组播组

时并不影响聚合树结构。终端用户加入时总是选择低延迟轻负载的成员代理。各个成员代理在簇内采用基于共享树的方法构造应用层组播树。

3.7 小结

表 2 对上述 6 种代理型应用层组播体系结构进行了总结与比较,其比较内容包括所支持的应用种类(Applications)、主要设计需求(Requirements)、协议优化目标(Optimization)、覆盖网管理方式(Overlay Management)和主要技术特点

(Characteristics)等 5 个方面。从表中可见,大部分的代理型应用层组播体系结构都支持基于流媒体技术的应用。而在设计需求方面,可扩展性、可靠性和实时性是大部分代理型应用层组播体系结构需要优先考虑的问题。各种体系结构下的组播路由协议的优化目标通常包括延迟和带宽两个方面。在覆盖网管理方面,除了 AMcast/ALMI 采用集中式控制方式外,其它的体系结构都采用分布式的控制方式。而这 6 种体系结构在具体实现技术上又各具特色。

表 2 代理型应用层组播体系结构比较

Architectures	Applications	Requirements	Optimization	Overlay Management	Characteristics
Overcast ^[20]	On-demand/live data delivery	①Scalability ②Reliability	Bandwidth	Distributed	Building Bandwidth efficient distribution trees
RMX ^[21]	①Shared whiteboard ②Multi-player games ③Software distribution	①Reliability ②Real time	Latency	Distributed	Using knowledge of application semantics
AMcast/ALMI ^[22]	Many kinds	①Real time ②Broadband access	①Cost ②Latency ③Balance	Centralized	ALMI is a middleware package
Scattercast ^[23]	Many kinds	①Efficiency ②Flexibility+ ③Ease of deployment	Customizable	Distributed	Customizable overlay architecture
OMNI ^[24]	Multimedia streaming applications	Real time	Latency	Distributed	Using localized transformations to adapt with changing distribution of MSNs*
TOMA ^[25]	Many kinds	①Scalability ②Efficiency	①Control overhead ②Multicast State	Distributed	Using Aggregated Multicast approach to improve the scalability

+ “flexibility” means to provide a range of broadcasting services rather than single standardized service.

* MSNs; the Multicast Service Nodes.

4 代理型应用层组播体系结构设计

通过对上述 6 种代理型应用层组播体系结构的分析,可以看出,借助代理服务器在互联网上实现能够为多种组播用户提供传输服务的基础设施是完全可行的。同时,还可以看出,设计代理型应用层组播体系结构需要考虑的基本问题包括:(1)代理服务器间组播树的构建和动态维护。代理节点间的组播树是构建于物理网络之上的覆盖网络,它可以是“有源树”也可以是“共享树”。构建与维护与物理拓扑相匹配的、自组织的、动态优化的组播树是代理型应用层组播路由协议设计的主要目标。(2)组成员和组播会话的管理。组播用户如何定位到最优(网络距离最近或带宽资源最丰富等)的代理节点,以及代理节点如何为用户节点提供组播传输服务,即采用直接相连的方式组播(简称为直接组播)还是使用户节点间进行应用层组播。同时,如何标识、产生和维护组播会话是组成员和组播会话管理方面需要重点考虑的问题之一。(3)物理链路状态探测。链路状态探测机制是有效构建和维护组播树的基础。组播系统需要周期性地检测代理节点间的物理链路状态,根据链路状态构建和动态优化组播覆盖网络和组播树,以及管理组成员的动态行为。链路状态探测机制直接影响着应用层组播系统性能。

目前,代理型应用层组播体系结构是较新的研究领域,因此还面临着许多的问题和挑战。我们认为,在设计新的代理型应用层组播体系结构时,如下两个方面的问题值得关注:(1)组播基础设施的应用感知性。覆盖网组播基础设施为多种组播通信应用提供数据传输服务,应用层组播应用需求的多样性对覆盖网组播基础设施的可复用性提出了挑战。在上述 6 种体系结构中,尽管相关文献称 AMcast, Scattercast 和 TOMA 能够支持多种类型的组播应用,但并没有具体设计出

适应于各种应用需求的机制或方法。如何使代理型应用层组播基础设施具备感知应用需求的能力,并智能地根据不同类型的应用需求构建优化的覆盖网组播树,从而增强覆盖网络基础设施的可复用性,这是一个值得研究的问题。(2)用户获得组播传输服务方式的自适应性。在类似 RMX 的体系结构中,成员节点通过直接组播的方式获得组播传输服务。当用户数量较大时,这种方式容易产生代理服务器瓶颈问题。而在类似 TOMA 的体系结构中,代理节点完全充当汇聚点 RP (Rendezvous Point)的角色,它将其服务域内的成员节点组织一棵应用层组播树。该方式可能会浪费代理节点的带宽资源。因此,如何使代理节点能够根据组播域中成员节点的数量自适应地调整用户节点获得组播传输的服务方式,从而保证既能充分利用代理节点的性能,又能提高代理节点的服务容量,这也是一个值得研究的问题。

结束语 代理型应用层组播体系结构是 IP 组播和主机型应用层组播两种体系结构相互借鉴的结果,它既在一定程度上解决了 IP 组播难于部署的问题,又解决了主机型应用层组播性能低下的问题。该组播体系结构通过部署高性能代理服务器而构建组播服务平台,从而能够为用户提供优质的组播传输服务。因而,代理型应用层组播体系结构将是互联网组播体系结构发展的一个新方向。积极研究代理型应用层组播体系结构,广泛开展相关实验,推动代理服务器等基础设施的部署,对于构建支持组播功能的未来互联网具有重要的意义。

参考文献

- [1] Lao L, Cui J H, Gerla M, et al. A Comparative Study of Multicast Protocol: Top, Bottom, or In the Middle[C]//IEEE INFOCOM. Barcelona, Spain, 2006

- [2] Deering S. Multicast Routing in Internetworks and extended LANs[J]. *ACM Transaction on Computer Systems*, 1990, 8(2): 85-110
- [3] Mantra-Monitoring C. Multicast on a Global Scale[OL]. <http://www.caida.org/tools/measurement/Mantra/>, 2008
- [4] Diot C, Levine B, Lyles J, et al. Deployment issues for the IP multicast services and architecture[J]. *IEEE Network*, 2000, 14(1): 78-88
- [5] Almeroth K. Multicast Help Wanted: From Where and How Much[C]// *IEEE Consumer Communications and Networking Conference (CCNC)*. Vegas, Nevada, 2007
- [6] Ratnasamy S, Ermolinskiy A, Shenker S. Revisiting IP Multicast[C]// *ACM SIGCOMM*. Pisa, Italy, 2006
- [7] Saltzer J, Reed D, Clark D. End-to-end arguments in system design[J]. *ACM Transactions on Computer System*, 1984, 2(4): 195-206
- [8] Francis P, Pryadkin Y, Radoslavov P, G, et al. Yoid: Extending the Internet Multicast Architecture[OL]. <http://www.aciri.org/yoid>, 2008
- [9] Chu Y H, Rao S G, Zhang H. A case for end system multicast [J]. *IEEE Journal on Selected Areas in Communication*, 2002, 20(8): 1456-1471
- [10] Pendarakis D, Shi S, Verma D, et al. ALMI: An application level multicast infrastructure[C]// the 3rd Usenix Symposium on Internet Technologies and System. Boston, USA, 2001
- [11] Banerjee S, Bhattacharjee B, Kommareddy C. Scalable application layer multicast[C]// *ACM SIGCOMM*. Pittsburgh, USA, 2002
- [12] Zhuang S Q, Zhao B Y, Joseph A D, et al. Bayeux: An architecture for scalable and fault-tolerant wide-area data dissemination [C]// the 11th Intl. Workshop on Network and Operating System Support for Digital Audio and Video (NOSSDAV). New York, USA, 2001
- [13] Helder D A, Jamin S. End-host multicast communication using switch-tree protocols [C]// the 2nd IEEE/ACM International Symposium on Cluster Computing and the Grid. Berlin, Germany, 2002
- [14] Ratnasamy S, Francis P, Handley M, et al. A scalable content-addressable network[C]// *ACM SIGCOMM*. New York, USA, 2001
- [15] Castro M, Druschel P, Kermarrec A M, et al. SCRIBE: A large-scale and decentralized application-level multicast infrastructure [J]. *IEEE Journal on Selected Areas in Communications*, 2002, 20(8): 1489-1499
- [16] Kostic D, Rodriguez A, Albrecht J, et al. Bullet: High bandwidth data dissemination using an overlay mesh[C]// the 20th ACM Symposium on Operating System Principles(SOSP). New York, USA, 2003
- [17] Castro M, Druschel P, Kermarrec A M, et al. Splitstream: High-bandwidth multicast in cooperative environments [C]// 20th ACM Symposium on Operating System Principles(SOSP). New York, USA, 2003
- [18] Li Z, Mohapatra P. HostCast: A new overlay multicasting protocol [C]// *IEEE International Conference on Communications (ICC)*. Alaska, USA, 2003
- [19] Cui Y, Li B, Nahrstedt K. oStream: Asynchronous streaming multicast in application-layer overlay network[J]. *IEEE Journal on Selected Areas in Communications*, 2004, 22(1): 91-106
- [20] Jannotti J, Gifford D K, Johnson K L, et al. Overcast: Reliable multicast with an overlay network[C]// the 4th Usenix Symposium on Operating System Design and Implementation (OSDI). Denver, Colorado, 2000
- [21] Chawathe Y, McCanne S, Brewer E. RMX: Reliable multicast in heterogeneous environments[C]// *IEEE INFOCOM*. Tel-Aviv, Israel, 2000
- [22] Shi S. A Proposal for A Scalable Internet Multicast Architecture [R]. WUCS-0103, Washington University in StLouis, USA, 2001
- [23] Chawathe Y. Scattercast: An adaptable broadcast distribution framework[J]. *ACM Multimedia System*, 2003, 9(1): 104-118
- [24] Banerjee S, Kommareddy C, Bhattacharjee K K B, et al. Construction of an efficient overlay multicast infrastructure for real-time applications[C]// *IEEE INFOCOM*. San Francisco, USA, 2003
- [25] Lao L, Cui J H, Gerla M. TOMA: A Viable Solution for Large-Scale Multicast Service Support [C]// *IFTP NETWORKING*. Waterloo, Canada, 2005
- [26] Clark D, Tennenhouse D L. Architectural Considerations for a New Generation of Protocols[C]// *ACM SIGCOMM*. Philadelphia, USA, 1990
- [27] Shi S, Turner J, Waldvogel M. Dimensioning server access bandwidth and multicast routing in overlay networks[C]// the 11th Intl. Workshop on Network and Operating SysSupport for Digital Audio and Video (NOSSDAV). New York, USA, 2001
- [28] Shi S, Turner J. Routing in Overlay Multicast Networks[C]// *IEEE INFOCOM*. New York, USA, 2002
- [29] Fei A, Cui J H, Gerla M, Faloutsos M. Aggregated Multicast: an approach to reduce multicast state[C]// 6th Global Telecommunications Conference(GLOBECOM). Texas, USA, 2001

(上接第 32 页)

个用户进行通信,实现了空间复用的最大化,但也面临着新的问题:波束越多,用户间干扰越大,导致系统整体性能的降低。鉴于天线个数、波束个数与系统性能之间的不确定性,尤其是无线信道时变特性,给系统性能改善带来极大挑战。因此,我们还将非满秩波束的基础上,结合波束成型的瞬态特性,研究非满秩的天线传输问题,即考虑从多个天线中选出部分最优的天线,进行空间多用户复用,在现有确定天线基础上增加一维新的自由度,以期达到较高的性能。

参 考 文 献

- [1] Sharif M, Hassibi B. On the capacity of MIMO broadcast chan-

nels with partial side information[J]. *IEEE Trans. Inf. Theory*, 2005, 51(2): 506-522

- [2] Hassibi B, Marzetta T L. Multiple-antennas and isotropically random unitary inputs: The received signal density in closed form[J]. *IEEE Trans. Inf. Theory*, 2002, 48(6): 1473-1484
- [3] Viswanath P, Tse D N C, Laroia R. Opportunistic beamforming using dumb antennas[J]. *IEEE Trans. Inf. Theory*, 2002, 48(6): 1277-1294
- [4] Taesang Yoo, Goldsmith A. On the Optimality of Multiantenna Broadcast Scheduling Using Zero-Forcing Beamforming[J]. in *IEEE J. Sel. Areas Commun.*, 2006, 24(3): 528-541