



IP 组播网络监测系统的研究

王虎, 曹争

(东南大学计算机科学与工程学院, 南京, 210096)

摘要: 为了更合理的利用和管理安全组播服务系统, 本文设计了一套 IP 组播网络监控测量系统。首先介绍了该系统的框架结构与功能模块, 并研究了以 MCS (组播控制服务器) 为核心的分级监测方法。随后提出了组播组性能测量方案, 该方法可解决组播数据流的单向性带来的问题, 而且能在降低组播网络安全性问题威胁的同时降低测量过程中带来的网络负载。最后本文还设计了一个可以结合该系统使用的客户端, 其可完成类似于单播 ping 的功能。

关键词: 组播; 性能测量; 探测; ping

The Research of IP Multicast Network Monitoring System

Wang Hu, Cao Zheng

(School of Computer Science and Engineering, Southeast University, Nanjing, 210096)

Abstract: To use and manage the Secure Multicast Server System, an IP Multicast Network Monitoring System is designed in this paper. First of all, the frame structure and functional module of the system is built in this paper, and the method of multi-layer monitor based on the core of MCS is proposed. Then, in order to resolve the problem concerning the one-way measurement, the algorithm of multicast performance measurement is designed, and this method can decrease the threaten about multicast secure problem and decrease the load in the process of network measurement. At last, a client which can use as a part of the system is designed, and, to some extent, it can work well like the ping.

Key words: Multicast; Performance Measurement; Detect; ping

组播技术是在 1988 年通过定义 D 类地址和 IGMP (Internet 组管理协议) 正式加进 IPV4 的, 1992 年 MBONE 的建立推动了组播技术的发展。组播技术能有效地解决单点发送、多点接收的问题, 能实现网络中点到多点的高效数据传输, 从而大量节约网络带宽、降低网络负载, 因此它是大规模组通信的理想技术。虽然由于组播技术的安全性问题, 使得组播技术一直未能在 Internet 中广泛应用, 但是随着 Internet 的发展, 组播技术的应用越来越多, 组播技术也日趋成熟。同时, 作为 IPV4 的替代协议, 组播在 IPV6 中得到进一步的丰富和加强。

在组播网络的建立和发展过程中, 产生了一系列用于网络调试、管理的工具。如 Mrinfo 是用于收集 Mbone 节点和隧道当前状态信息的工具, Mtrace 是跟踪两个主机之间组播路径的工具^[1], RTPmon^[2]

可以加入到一个特定的组播地址, 接收所有老的反馈报告 (RTCP) 报文, 而 Mhealth^[3]在 RTPmon 的基础上更进一步, 将一个特定组的组播树形结构 (包括丢包信息等), 用实时、图形化的方式显示出来。总的来说, 各种工具功能比较单一, 可扩展性及应用范围都比较小。为了满足安全组播系统中的管理需求, 需要设计一个系统化的可扩展的面向大规模组通信的 IP 组播综合监测系统。该系统监测当前活跃组播组的运行状况, 包括组地址、源地址、组成员分布等组播组基本信息和带宽利用率、会话时延、时延抖动、服务等级等组播组性能参数。此外, 该系统还提供一种面向组播的调试和测量工具^[4], 以便主动测量和调试使用组播流量的网络。包括探测组播是否可使用、追踪组播转发路径和获取组播组性能参数。

基金项目: 国家科技支撑计划项目(2008BAH37B04)

作者简介: 王虎, (1986-), 男, 硕士研究生, E-mail: hwang@ninet.edu.cn; 曹争, (1958-), 男, 副教授, E-mail: zcao@njnet.edu.cn.

1 系统框架描述

本文所设计的 IP 组播网络监测系统是基于大规模可控组播试验网建立的, 其拓扑结构见图 1。



如图 1 所示，为了提供安全可靠的组播服务，实验网中使用对组播组的集中管理控制方式，引入了组播控制服务器（MCS）和组播策略服务器（MPS）。每个驻地网配置一个 MCS，驻地网与主干网相连。多个 MCS 可共用一个 MPS，MPS 主要提供组播源管理、用户管理以及服务策略管理等功能。

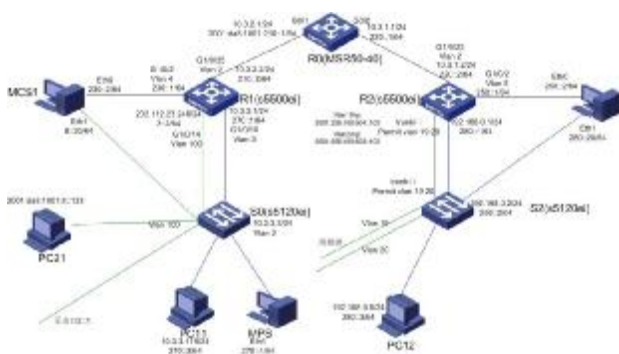


图 1 实验网拓扑图

本系统以 MCS 为核心，将使用组播流的网络分成两级：驻地网域内和驻地网域间。充分的利用 MCS 采集和统计当前活跃组播组的信息，同时分别在域内和域间测量组播组的性能参数，以便对组播组进行监控。该系统的整体框架见图 2^{5-6]}。这种框架中，驻地网域内的用户频繁的增加

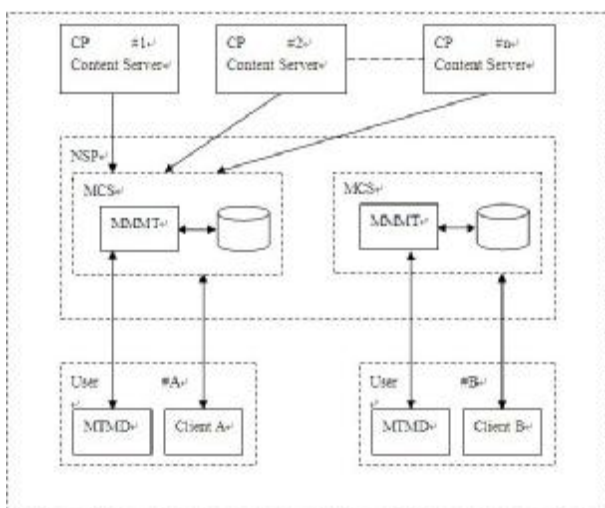


图 2 系统框架图

入退出和主动探测等操作对域间主干网部分的统计和测量的影响将大幅度减小，整个测量的复杂度也将会减小。因为安全组播服务业务的扩展是以 MCS 为单位的，本系统框架也体现了良好的可扩展性。

如图 2 所示，本系统的核心是 MCS，主要功能由 MCS 的 MMTD（组播监控测量模块）完成。

MTMD（组播组测量调试工具）可以以一个客户端的形式被组播用户使用，进行对组播组的主动探测，功能类似于单播 ping 等调试工具。

2 系统设计

2.1 系统结构

作为整个系统的核心功能模块，MMMT 完成组播组信息的统计和组播组性能的测量，并协调好域间和域内的组播组性能测量。系统总体设计如图 3。根据图 3 所示，MMMT 主要由以下四个子模块组成：统计测量模块、处理 MTMD 请求模块、与其它 MMTD 通信模块、显示模块；而客户端 MTMD 由以下两个模块组成：消息处理模块、结果显示模块。

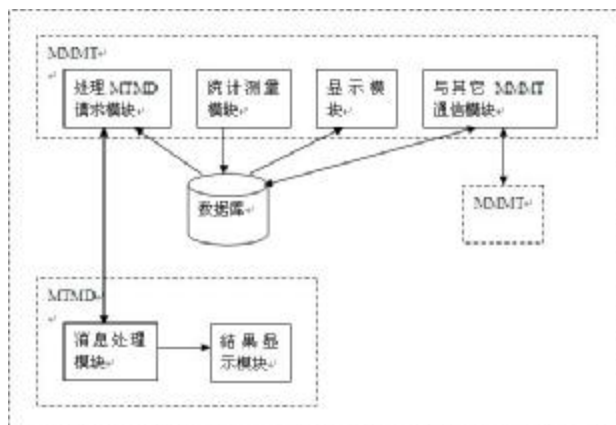


图 3 系统总体设计

2.2 功能模块

2.2.1 MMTD 子模块

1) 处理 MTMD 请求模块

该模块监听 MTMD 的请求消息，收到请求消息后解析消息中的参数和组地址，根据组地址和参数访问数据库，将从数据库中获得的信息返回给 MTMD。

2) 统计测量模块

该模块周期性的从 MCS 中获取该驻地网域中当前活跃的组播组的基本信息，必要时访问 MPS。同时测量组播组性能参数，并存储到数据库中。其中，根据组播源所在位置，该模块将采用不同的工作方式。

3) 与其它 MMTD 通信模块



该模块与其它 MMMT 交互，主要功能是协助靠近组播源的 MCS 构建该组的组播树和统计当前在线用户。

4) 显示模块

该模块提供图表式的信息表达方式，方便管理员查看。

2.2.2 MTMD 子模块

1) 消息处理模块

该模块是 MTMD 的功能模块，当使用 MTMD 时，该模块检查参数，并转发给用户所在域的 MCS，然后等到返回的结果，并对收到的结果进行处理。

2) 结果显示模块

该模块将结果显示给用户。

3 关键技术设计

3.1 组播组信息统计和性能测量算法

本系统主要用于 SSM 组播网络模型^[7-8]，组播源的地址由 MPS 统一管理，当一个新的组播源申请成功后，由 MPS 以 (S, G) 的形式通告给其管理域中的所有 MCS。MCS 也可主动向 MPS 查询组播组基本信息。MCS 主要根据数据包类型作相应的处理完成统计和测量的功能。

3.1.1 用户加入退出消息处理

- [1] 当有合法用户请求加入组播组 (S, G) 时，查询数据库：若该组已存在，将用户数加 1。否则根据 (S, G) 建立组播组项，并注明类型，用户数初始化为 1，然后查询 MPS 获取组播源所在域的 MCS 地址和服务等级等信息；
- [2] 检查组播源所在域的 MCS 地址，若是首次发现该 MCS，建立 SA 项，并初始化计数器 mc 为 1，否则计数器 mc 加 1；只要计数器 mc 大于 0，周期性的发送 mtrace 请求，并记录所经过的网络设备及其时延等值；
- [3] 当有用户退出时，用户数减 1，当用户数不大于 0 时，计数器 mc 减 1，当计数器 mc 不大于 0 时，停止 mtrace 查询，删除 SA 项；

3.1.2 组播数据处理

对组播数据的处理算法主要完成组播组性能测量的功能，这个过程中采用被动测量的方式。由于目前组播应用多偏向于音视频等多媒体服务，这些应用要求高带宽低时延，如果测量系统对每个数据包都进行统计计算，那么测量的代价将会很高，而且可能因为高负荷工作而使测量结果出现偏差或者导致测量系统崩溃。因此，我们采用组播组流的方式对组播数据进行处理^[9]，尽可能降低测量代价。

所谓组播组流是指将一段连续的组播数据根据特定的规则组合成一个个数据块，以数据块为单位进行数据统计和处理。例如，如果根据 (组地址，源地址，时间间隔) 这个规则组成组播流，就可以将所有组地址为 ff35::101:2，源地址为 2001:da8:1001:280::123 的组播数据流按时间间隔为 5 分钟分成一个个数据块。

对组播数据的具体算法又根据组播流向和组地址确定算法类型，若是从源经 MCS 流出的数据包使用算法 A，否则使用算法 B；

算法 A：

- [1] MCS 对收到的组播数据 IP 报头做一些改动或者扩展，然后再转发出去。例如：在每个数据包中确定一个标志位 (FB)，根据规则 (组地址，源地址，时间间隔) 循环的将 FB 设置为 0 或者 1，如果相邻的数据包 FB 不同，就以这两个数据包为界，判定他们属于不同的组流；
- [2] MCS 周期性的向源地址 S 发送探测报文 (使用 ICMPV6 或者 UDP)，测量 MCS 到源的 (会话) 时延、时延抖动和丢包率，并记录这些信息。

算法 B：

MCS 对收到的组播数据包根据算法 A 中设计的组流规则进行统计计算，获得本地 MCS 到该组播组靠近源的 MCS 的 (会话) 时延、时延抖动和丢包率；

3.1.3 用户主动探测消息处理

- [1] 解析探测消息参数，检查探测组地址及其它参数；
- [2] 若组地址不合法，返回给用户组不存在消息；否则
- [3] 根据组地址查询数据库，按照参数获取组运行状况返回给用户；



3.2 MTMD 的设计

MTMD 设计成一个纯软件的小工具,主要参数就是将要测试的组地址。对组地址只做地址正确性检查,确保组地址参数是一个合法的 IPV6 组播地址。还可设计一些有关探测的可选参数以便实行个性化的测量。例如,主动探测时间间隔,探测包的数量等。

假设 MCS (2001:da8:1001:230::2) 已有关于 (2001:da8:1001:280::123, ff35::101:2) 的记录。用户可使用 MTMD 对组地址 ff35::101:2 进行探测,此时,MTMD 会将 (2001:da8:1001:280::123, ff35::101:2) 封装在数据域然后单播给 MCS (2001:da8:1001:230::2),MCS 查询数据库将组播组基本信息和所测量性能参数返回给用户,用户根据收到的响应报文计算 MCS 和用户之间的性能参数,然后再结合报文中的信息得到最终的结果;同时,MCS 将 MTMD 请求以单播的形式转发给组播源,组播源将以单播的形式直接向用户发送探测响应报文,然后用户可以记录关于单播的网络性能参数。

如果用户使用的组地址是 ff35::101:5,但是 MCS 没有关于这个组播组的信息,MCS 可以查询 MPS,结果可分为两种情况:第一种情况,MPS 没有关于 ff35::101:5 的记录,则 MCS 向用户返回此组播组不存在信息;第二种情况,MPS 有这个组播组的记录,则表明 ff35::101:5 是合法组播组地址,MCS 根据从 MPS 获得的信息建立 (S, G) 组项,然后根据 3.1 所述算法执行,待获得稳定有效的测量结果后返回给用户。

4 结束语

本监测系统是一种分层结构的可扩展的组播组监测系统,采用主动与被动相结合的方式测量,分两层进行统计计算。在组播网络中实现了对组播组的监控和管理,以 NSP 的角度,可以为更有效的管理和配置组播网络提供依据;以普通用户的角度,可以使用类似于单播 ping 的功能组播组的探测,查看指定组播组的相关信息。本文所设计的系统建立在安全组播服务实验网的基础上,组播组是可控的,所以能有效的避免滥用组播地址进行主动探测带来的安全问题。

参考文献

- [1] Adams A, Bu T, Caceres R, et al. The use of end-to-end multicast measurements for characterizing internal network behavior[J]. IEEE communication Magazine, 2004, 38(2): 341. 349.
- [2] David Bacher, Andrew Swan, Lawrence A Rowe. Rtpmon: A third-party RTCP monitor[C]. Proceedings of the Fourth ACM International Conference on Multimedia. Boston, Massachusetts, United States, 1997: 437-438.
- [3] David B Makofske, Kevin C Almeroth. MHealth: A realtime multicast tree visualization and monitoring tool[C]. Proc. NOSSDAV 99, Basking Ridge New Jersey, USA, June 1999.
- [4] Internet-Draft, Multicast Ping Protocol[S].
- [5] Internet-Draft, AAA and Admission Control Framework for Multicasting[S].
- [6] Internet-Draft, Requirements for Multicast AAA coordinated between Content Provider(s) and Network Service Provider(s)[S].
- [7] RFC: 3569, An overview of Source-Specific Multicast(SSM)[S].
- [8] Stig Venaas, Tim Chown. Source Specific Multicast(SSM) with IPV6[C]. Proceedings of the The 2005 Symposium on Applications and the Internet Workshops(SAINT-W'05): 0-7695-2263-7,2005.
- [9] Internet-Draft, A method for IP multicast performance monitoring[S].