

PIM-SM 协议的建模与改进^{*}

李丹¹⁺, 吴建平¹, 崔勇¹, 徐恪¹, 刘莹²

¹(清华大学 计算机科学与技术系, 北京 100084)

²(清华大学 信息网络工程研究中心, 北京 100084)

Modeling and Improvement of PIM-SM Protocol

LI Dan¹⁺, WU Jian-Ping¹, CUI Yong¹, XU Ke¹, LIU Ying²

¹(Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China)

²(Network Research Center, Tsinghua University, Beijing 100084, China)

+ Corresponding author: Phn: +86-10-62785822, E-mail: lidan@csnet1.cs.tsinghua.edu.cn, http://netlab.cs.tsinghua.edu.cn/~lidan

Li D, Wu JP, Cui Y, Xu K, Liu Y. Modeling and improvement of PIM-SM protocol. Journal of Software, 2006,17(2):285-294. <http://www.jos.org.cn/1000-9825/17/285.htm>

Abstract: PIM-SM (protocol-independent multicast-dense mode) is currently the preferred intra-domain multicast routing protocol. One problem that impedes its widely use is its high overhead of control messages. In order to improve and optimize the PIM-SM, its performance model should be established and the nicety performance analysis should be made above all. In this paper, the Stochastic Petri Net (SPN) model of the whole PIM-SM protocol is established, and the analysis and the experiments are made on the router processing load caused and the network bandwidth consumed by each type of the protocol messages, based on the model and router realization. It is discovered that register message and Join/Prune message cause most router processing load, while Join/Prune message and Bootstrap message consume most network bandwidth. According to the conclusion of performance analysis, an improvement is made on PIM-SM, which is achieving better performance compared with the former protocol.

Key words: PIM-SM (protocol-independent multicast-dense mode); SPN (stochastic Petri net); performance analysis; multicast

摘要: PIM-SM(protocol-independent multicast-dense mode)协议是目前 Internet 首选的域内组播路由协议.影响其广泛应用的一个主要问题是该协议的控制报文负载比较大.为了对协议进行改进和优化,首先需要建立性能模型并进行准确的性能分析.利用随机 Petri 网(stochastic Petri net,简称 SPN)模型对整个 PIM-SM 复杂的协议行为进行了建模,并在其 SPN 模型的基础上,结合路由器的实现,对协议中每种消息消耗的路由器处理负载和占用的网络带宽进行了分析和实验,发现 Register 消息和 Join/Prune 消息消耗的路由器处理负载比较多,而 Join/Prune 消息和 Bootstrap 消息占用的网络带宽比较大.根据性能分析的结论对 PIM-SM 协议进行了改进.与原来的协议相比,改进后的协议性能明显提高.

^{*} Supported by the National Grand Fundamental Research 973 Program of China under Grant No.2003CB314801 (国家重点基础研究发展规划(973)); the National Natural Science Foundation of China under Grant Nos.60303006, 90104002 (国家自然科学基金)

Received 2004-09-07; Accepted 2005-07-28

关键词: PIM-SM(protocol-independent multicast-dense mode);SPN(stochastic Petri net);性能分析;组播
中图法分类号: TP393 文献标识码: A

组播是一种介于单播和广播之间的通信模式,它允许一台或多台主机作为组播源一次性地发送单一数据包到多台主机^[1],有效地节省了网络带宽,减轻了网络负载和服务器负载.随着 Internet 的发展,网络会议、远程医疗、订阅/分发系统、电子白板、网络游戏等应用都要求 Internet 能够支持组播.在下一代 Internet 核心协议 IPv6 中,组播已经被确定为必须实现的技术^[2].

为了传送组播分组到所有的组播组成员,路由器必须构造并维护一棵连接所有组播成员的树,即组播分发树.因此,组播包括两部分功能:组播分组转发和组播分发树的维护.组播路由协议的主要任务就是维护组播分发树.由于组播数据在传送过程中不但要被路由,还要在适当的地方被复制,这也使得组播路由协议的实现比单播路由协议要复杂.

组播路由协议可以分为域内协议和域间协议两种.常见的域内组播路由协议又可以根据组播树的结构特征分为两类:基于源树的协议和基于共享树的协议.前者的代表有距离向量组播路由协议(distance vector multicast routing protocol,简称 DVMRP)^[3]、开放最短路径优先的组播扩展(multicast extensions open shortest path first protocol,简称 MOSPF)^[4]、协议独立组播密集模式(protocol-independent multicast-dense mode,简称 PIM-DM)^[5]等;后者的代表有基于核的组播树(core-based tree,简称 CBT)^[6]、协议独立组播稀疏模式(protocol-independent multicast-sparse mode,简称 PIM-SM)^[7]等.在这些组播路由协议中,PIM-SM 由于其与单播路由协议的无关性,并且采用成员显式加入的方式,支持组播树从共享树到源树的切换,从而成为域内组播路由协议的首选.

PIM-SM 协议实现的复杂性和其控制报文的负载,一直是影响其广泛使用的主要问题,也是研究者关注的焦点问题.大量的协议控制报文带来了至少两个问题,一个是路由器处理负载的增加,另一个是占用较多的网络带宽.其解决途径有两个,一是路由器对协议报文处理的优化,二是协议机制本身的改进,但这两种方法都是以协议的性能分析为基础的.

通常较直观的性能分析方法是在实际网络上进行协议性能测试.但是,对大规模网络环境的测试很难操作,而且也难以进行定量分析.而随机 Petri 网(stochastic Petri net,简称 SPN)^[8]模型则是一种很优秀的性能评价方法.本文采用 SPN 模型对整个 PIM-SM 协议进行了建模,并在其 SPN 模型的基础上,结合路由器的实现,分析了 PIM-SM 协议的性能,最后,在分析结论的基础上对协议进行了改进,并且对改进后的协议和原来协议进行了性能比较.

本文第 1 节介绍相关工作.第 2 节建立 PIM-SM 协议的 SPN 模型.第 3 节在 SPN 模型和路由器实现的基础上,针对每种协议消息,对 PIM-SM 协议进行性能分析.第 4 节在性能分析结论的基础上对 PIM-SM 协议进行改进,并与原来协议作性能比较.第 5 节总结全文.

1 相关工作

自从 PIM-SM 协议的 RFC 标准^[7]于 1998 年 6 月诞生以来,PIM-SM 已经被业界视为首选的域内组播路由协议,并在各大厂商的路由器上得到了广泛的支持.

但是,关于 PIM-SM 协议的研究大多局限在 PIM-SM 协议的安全扩展、QoS 扩展等方面.在安全扩展方面,Hardjono 提出了一种通过路由器上 PIM-SM 协议的协助来实现组播接收者的访问控制的方法^[9];而在 QoS 扩展方面,Biswas 和 Izmailov 提出了一种基于 PIM-SM 协议的支持 QoS 扩展的组播框架^[10].

随着 PIM-SM 协议的大规模部署,对协议本身的性能分析和优化越来越重要. Billhartz 把 PIM-SM 和 CBT 这两种基于共享树的组播路由协议进行了性能比较,发现 PIM-SM 协议的控制负载远远大于 CBT 协议^[11]. Ying-Dar Lin 建议根据网络拓扑的变化,周期性地选举 RP,实验表明,可以减少维护组播树的开销^[12]. Holt 提出根据网络状况,动态、自适应地让组播树在共享树和源树之间切换,并通过实验证明,这样会大量减少控制

负载^[13].

但是,上述研究都没有对整个 PIM-SM 协议进行建模.SPN 模型是一种很优秀的性能评价方法,已经在许多复杂系统的建模和性能评价中得到了广泛的应用.陆正福等研究者利用扩展 Petri 网模型,对 PIM-SM 协议的一些协议细节进行了建模^[14],但却没有建立整个 PIM-SM 协议的 SPN 模型,而且没有在模型基础上进一步分析协议性能.本文针对这几方面研究的不足,首次采用 SPN 模型对整个 PIM-SM 协议进行建模,并在其 SPN 模型的基础上结合路由器的实现,分析了 PIM-SM 协议的性能,然后基于分析结果对协议作了改进.

2 PIM-SM 协议的 SPN 模型

如果把 PIM-SM 协议的功能都集中到一个抽象的路由器上(因为并不是每个路由器都会运行所有功能),那么,这台路由器在启动 PIM-SM 协议后将执行如下操作: 周期性地发送 Hello 消息; 周期性地发送 Bootstrap 消息; 周期性地发送 Candidate-RP-Advertisement 消息; 周期性地发送 Join/Prune 消息; 如果是 DR,则在收到组播数据的初期发送 Register 消息到 RP,并在此后周期性地探测发送 Register 消息; 从错误入接口收到组播数据包时立即发送 Assert 消息; 对收到的各种 PIM-SM 消息进行处理.

由此得到 PIM-SM 的 SPN 模型如图 1 所示.其中,各个位置的含义见表 1,各个变迁的含义见表 2.在图 1 中,T2,T3,T5,T6 为确定时间变迁,T0,T1,T4,T7 为随机时间变迁,T8~T14 为变迁子网.

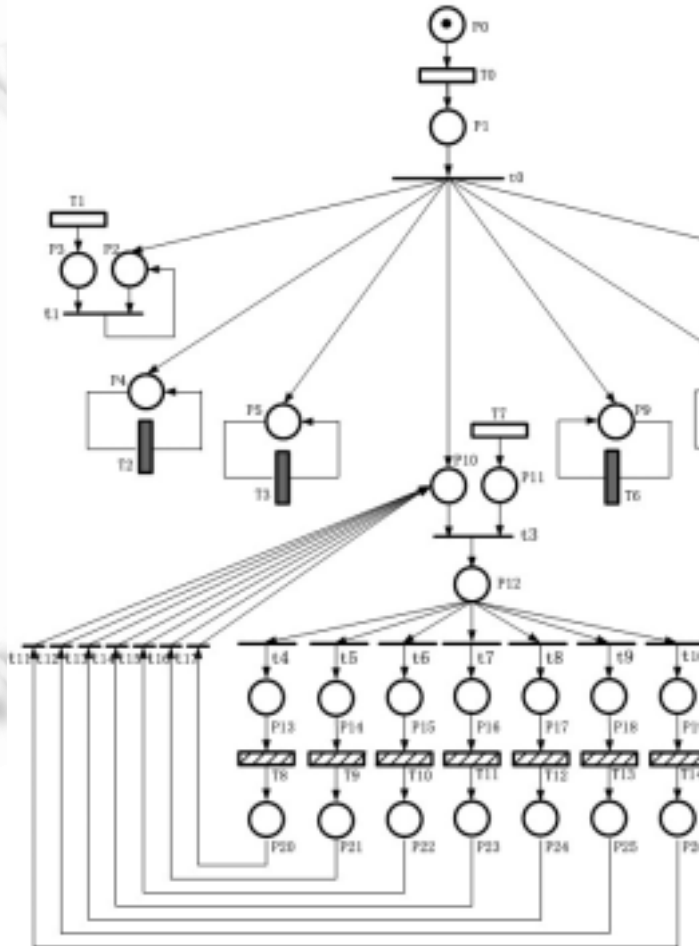


Fig.1 SPN model of PIM-SM protocol

图 1 PIM-SM 协议的 SPN 模型

Table 1 Significations of places in Fig.1

表 1 图 1 中各个位置的含义

Place	Signification	Signification of the token
P0	Ready to start PIM-SM	PIM-SM protocol running token
P1	PIM-SM running normally	PIM-SM protocol running token
P2	Ready to send register message	Register message sending token
P3	Register message triggered	Multicast packet
P4	Ready to send Hello message	Hello message sending token
P5	Ready to send Bootstrap message	Bootstrap message sending token
P6	Ready to send Assert message	Assert message sending token
P7	Assert message triggered	Multicast packet
P8	Ready to send candidate- RP-Advertisement message	Candidate-RP-Advertisement message sending token
P9	Ready to send Join/Prune message	Join/Prune message sending token
P10	Ready to receive PIM-SM message	PIM-SM message receiving token
P11	PIM-SM message arrived	PIM-SM message
P12	PIM-SM message accepted	PIM-SM message
P13	Hello message accepted	Hello message
P14	Bootstrap message accepted	Bootstrap message
P15	Candidate-RP-Advertisement message accepted	Candidate-RP-Advertisement message
P16	Join/Prune message accepted	Join/Prune message
P17	Register message accepted	Register message
P18	Register-Stop message accepted	Register-Stop message
P19	Assert message accepted	Assert message
P20	Hello message process finished	Hello message
P21	Bootstrap message process finished	Bootstrap message
P22	Candidate-RP-Advertisement message process finished	Candidate-RP-Advertisement message
P23	Join/Prune message process finished	Join/Prune message
P24	Register message process finished	Register message
P25	Register-Stop message process finished	Register-Stop message
P26	Assert message process finished	Assert message

Table 2 Significations of transitions in Fig.1

表 2 图 1 中各个变迁的含义

Transition	Signification
T0	PIM-SM initializing
T1	Non-RP router receiving a multicast packet for which there is no corresponding routing entry, or valid time for Register-Stop expiring during a multicast session
t1	Sending Register message
T2	Sending Hello message
T3	Sending Bootstrap message
T4	Receiving a multicast packet from a wrong incoming interface
t2	Sending Assert message
T5	Sending Candidate-RP-Advertisement message
T6	Sending Join/Prune message
T7	Arriving a PIM-SM message
T8	Processing received Hello message
T9	Processing received Bootstrap message
T10	Processing received Candidate-RP-Advertisement message
T11	Processing received Join/Prune message
T12	Processing received Register message
T13	Processing received Register-Stop message
T14	Processing received Assert message

瞬时变迁 $ti(i=4,5,6,7,8,9,10)$ 对应的可实施谓词 yi 分别为

- y4: 收到Hello消息
- y5: 收到 Bootstrap消息
- y6: 收到 Cand - RP - Adv消息
- y7: 收到Join/Prune消息
- y8: 收到 Register消息
- y9: 收到Register - Stop消息
- y10: 收到Assert消息

3 PIM-SM 协议的性能分析

我们从 PIM-SM 协议的各种协议消息的角度来分析协议性能,找出主要矛盾,为协议的改进提供依据.在 PIM-SM 协议的 7 种消息中,Assert 消息实际上是对协议构造组播树的一种修正,在经过修正后将保持较长时间的正常状况.并且,Assert 消息只在多路接入网络中出现,虽然在今后的网络发展中多路接入网络将会越来越多,但我们认为,Assert 消息仍然不是 PIM-SM 协议各种协议消息中的主要矛盾.因此,在本文的讨论中将其忽略.下面以一个具体的组播会话为实例,从 PIM-SM 协议另外 6 种消息的角度来分析协议性能.性能分析分两个方面:消耗的路由器处理负载和占用的网络带宽.

3.1 组播会话的网络拓扑及参数配置

文献[11]中采用了一种比较有代表性的网络拓扑,图 2 在其基础上进行了扩充.它表示一个由 23 台路由器和 34 条链路构成的自治域.域内的每台路由器都运行 PIM-SM 协议,组播会话就运行在这个自治域上.

根据上面的讨论,PIM-SM 协议运行后,路由器会周期性地发送 Hello 消息、Bootstrap 消息、Candidate-RP-Advertisement 消息和 Join/Prune 消息.按照协议 RFC 规定,它们的发送周期分别为 30s,60s,60s,60s^[7].因此,路由器也会周期性地收到这几种消息.除此之外,网络振荡会引起 Hello 消息的额外发送,组播组成员的动态变化会引起 Join/Prune 消息的额外发送.在组播会话的初期,组播源的 DR 会把组播数据封装后,通过 Register 消息发送到 RP,并在其后周期性地探测发送.如果 RP 已经通过源树收到组播包,再收到封装相同数据的 Register 消息后,就会立即发送 Register-Stop 消息.

我们在国家 863 项目“高性能 IPv6 路由器协议栈软件”的支持下,设计和实现了组播系统,包括 PIM-SM 协议.我们以此为平台进行测量和估算,对组播会话作参数假设,见表 3.

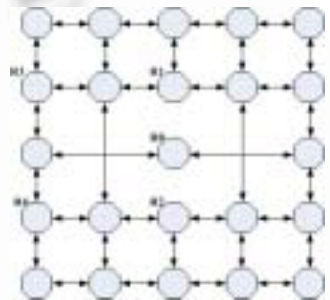


Fig.2 Network topology
图 2 网络拓扑

Table 3 Parameters of the multicast session
表 3 组播会话参数

Parameter name	Value
Number of candidate BSRs	1 (R0)
Number of candidate RPs for the multicast group	2 (R1,R2)
Number of multicast sources	2 (in LANs where R3, R4 stand)
Persistent time of the multicast session	2 hours
Average data sending rate of multicast sources	2 M bit/s
Average packet size	1 000 bytes
Average time interval for Register message unicast from DR to RP	30 milliseconds
Average time interval for RP to join Source-based shortest path tree and receive data from the tree	100 milliseconds
Average Time interval for Register-Stop message unicast from RP to DR	30 milliseconds
Average time interval for a router to send additional Hello messages because of network fluctuation	5 minutes
Average time interval for a router to send additional Join/Prune messages because of dynamic change of multicast group members	1 minute

3.2 消耗的路由器处理负载

路由器处理 PIM-SM 协议的负载可以分为两部分:发送协议消息和接收协议消息.与消息的接收处理相比,发送消息所消耗的负载是很小的,而且大部分消息的发送操作实际上是在对消息的接收处理中完成的,因此只考虑消息的接收处理所消耗的路由器负载.

对于路由器收到每种消息的数量,参考网络中收到这种消息最多的路由器.结合 PIM-SM 的协议机制、如图 2 所示的网络拓扑和表 3 中的参数,可以计算出在一个组播会话周期内,路由器收到 Hello 消息 1 056 条、Bootstrap 消息 120 条、Candidate-RP-Advertisement 消息 240 条、Join/Prune 消息 720 条、Register 消息 3 680 条、Register-Stop 消息 900 条.据此可以计算出,在图 1 的 SPN 模型中,t4~t10 的变迁实施概率分别为 0.157,0.018,0.036,0.107,0.548,0.134,0.

我们根据自己实现的 PIM-SM 协议来考察路由器处理每种消息的耗时权重.首先,定义有效代码和有效代码长度的概念.

定义 1(有效代码). 程序中表项搜索、内存拷贝等耗用处理器时间较多的代码称为有效代码.

定义 2(有效代码长度). 把有效代码中的循环语句展开(边界值参考上述组播会话实例),条件选择语句以其执行概率最大的分支代替,子函数以子函数的有效代码代替(递归进行,直到完全代替为止),所得到的代码总行数称为有效代码长度.

显然,路由器对每种消息的处理速度,近似地与处理这种消息的有效代码长度成反比.通过计算,处理各种消息的有效代码长度分别为:Hello 消息 174、Bootstrap 消息 511、Candidate-RP-Advertisement 消息 109、Join/Prune 消息 925、Register 消息 260、Register-Stop 消息 89.假设路由器处理 PIM-SM 协议的速度为单位速度 1,容易得到:图 1 中 T8~T13 的变迁速率分别为 0.174,0.059,0.278,0.033,0.116,0.340.

由于 PIM-SM 协议消息到达的随机模型并不会影响 $t_4 \sim t_{10}$ 的大小,因此对我们的讨论也不会有影响.我们不妨设 T7 的随机模型为默认随机模型,且变迁速率为 0.01,那么,P13~P18 中的平均标记数就表示 6 种消息消耗路由器处理负载的相对大小.运行 SPNP 软件进行模拟,所得到的结果如图 3 所示.

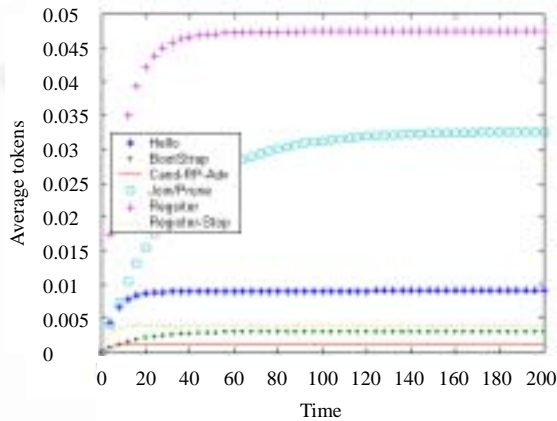


Fig.3 Router processing load caused by 6 types of PIM-SM messages

图 3 6 种 PIM-SM 消息消耗的路由器处理负载

在这 6 种消息中,处理 Register 消息所消耗的路由器处理负载是最大的,其次是 Join/Prune 消息.从协议本身来分析,由于组播源的 DR 需要在组播会话初期及其后周期性地把组播包封装在 Register 消息中发送到 RP,直到收到 RP 的 Register-Stop 消息为止,所以 RP 将会收到大量的 Register 消息.RP 的解封操作会消耗较多的路由器资源,在组播源发送速率较大时问题尤其突出.所以,得到如图 3 所示的结论是合理的.而 Join/Prune 消息包含了所有的组播组成员信息,是构造和维护组播树的关键,在大规模组播组成员动态变化时,消息会更加频繁.所以,Join/Prune 消息也会消耗较多的路由器处理负载.

3.3 占用的网络带宽

对于 PIM-SM 协议报文占用的网络带宽,计算组播会话期间出现在域内所有链路上的各种消息的流量.根据协议机制、如图 2 所示的拓扑结构和表 3 中的参数可以计算出:Hello 消息 17 952 条、Bootstrap 消息 4 080 条、Candidate-RP-Advertisement 消息 960 条、Join/Prune 消息 11 040 条、Register 消息 9 200 条、Register-Stop 消息 4 500 条.

再根据协议的报文格式计算上述情景中每种消息的最小载荷长度(不包括分组头).消息内的地址长度都以 IPv6 地址长度来计算.可以计算得到:Hello 消息 12 字节、Bootstrap 消息 100 字节、Candidate-RP-Advertisement 消息 48 字节、Join/Prune 消息 92 字节、Register 消息 8 字节(由于 Register 消息中封装了组播数据,因此计算协议报文长度时不包含数据长度)、Register-Stop 消息 44 字节^[7].

网络中各种消息的流量如图 4 所示.

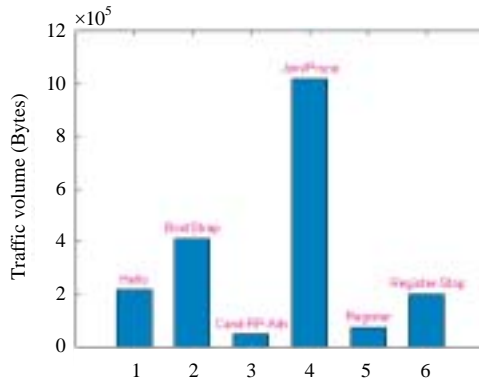


Fig.4 Network bandwidth consumed by 6 types of PIM-SM messages

图 4 6 种 PIM-SM 消息占用的网络带宽

Join/Prune 消息所占用的网络带宽在所有消息中是最多的,其次是 Bootstrap 消息.这是因为 Join/Prune 消息包含所有的组播组成员信息、报文数量和报文长度都比较大,而 Bootstrap 消息向网络中洪泛组播组和候选 RP 的对应消息也要占用较大的网络带宽.

4 PIM-SM 协议的改进

从上面的分析结果可以看出,在 PIM-SM 协议中,Register 消息和 Join/Prune 消息所消耗的路由器处理负载比较大,而 Join/Prune 消息和 Bootstrap 消息占用较多的网络带宽.因此,协议改进主要集中在这几种消息上.减轻这些协议消息的控制负载的方法有两个:一是优化路由器对协议报文的处理;二是通过改进协议机制,减少协议报文在网络上的流量.

4.1 协议改进和优化处理

对于 Register 消息,在路由器的处理上进行优化.在通常的路由器设计中,像 Register 消息这样的协议报文,需要在路由器的数据层面复制后,再上交到控制层面来处理.但 Register 消息的特殊性在于,它封装了组播数据,而且以发送组播数据的速率到达,大量的 Register 消息携带组播数据上交到路由器控制模块,会消耗很多路由器资源.所以,在优化设计中把 Register 消息放到数据层面即组播转发模块来处理,避免数据复制后再上交到控制层面,减轻了路由器的处理负担.通过优化设计和实现,处理 Register 消息的有效代码长度,折合成原来控制模块的有效代码长度约为 150.

按照 PIM-SM 协议的 RFC 规定,每当组播组成员发生动态变化时,都要立即发送额外的 Join/Prune 消息.对此进行如下改进:当组播组成员动态变化需要发送额外的 Join/Prune 消息时,并不立即发送,而是检查当前时间与下一次周期性发送 Join/Prune 消息时刻的时间间隔,如果这个间隔大于一定的阈值(规定为发送周期的 1/10),则立即发送;否则等到下一次周期性发送的时候一起发送.经过这样的改进,在上述组播会话周期内,网络中收到 Join/Prune 消息最多的路由器,收到的 Join/Prune 消息将减少为 684 条,而网络中所有链路上的 Join/Prune 消息总数也减少为 10 344 条.

对于 Bootstrap 消息机制,协议规定:候选 RP 路由器向自引导路由器申请成为某个组播组的 RP,然后,自引导路由器把每个组播组和它的所有候选 RP 的映射信息,通过 Bootstrap 消息周期性地网络中逐跳洪泛,收到 Bootstrap 消息的路由器,通过统一的哈希函数计算出该组播组的 RP^[7].对此机制进行如下改进:候选 RP 路由器向自引导路由器发送请求后,由自引导路由器通过哈希函数或者其他规则,计算出该组播组的唯一 RP,然后,将组播组和它对应的 RP 的信息向网络中洪泛.这样,原来网络中的每台路由器都需要从候选 RP 中经过计算选举出唯一的 RP,改进后只需要自引导路由器一台路由器来进行此项工作,而且 Bootstrap 消息的报文长度也减

小了.由于收到 Bootstrap 消息的路由器避免了选举 RP 的操作,在协议改进后的实现中,处理 Bootstrap 消息的有效代码长度减少为 400.而在上述组播会话情景中,Bootstrap 消息的报文长度也减少为 76 字节.显然,经过改进,处理 Bootstrap 消息的路由器负载减轻了,而且 Bootstrap 消息所占用的网络带宽也大为减少.

下面,再次从路由器处理负载和网络带宽两个方面来分析经过改进后的协议性能.

4.2 消耗的路由器处理负载

经过上述改进后,再次分析各种 PIM-SM 消息消耗的路由器处理负载.同样,在如图 2 所示的网络拓扑中,参照表 3 中的参数,可以计算出在该组播会话周期内,收到各种类型消息最多的路由器上收到的消息条数分别为:Hello 消息 1 056 条、Bootstrap 消息 120 条、Candidate-RP- Advertisement 消息 240 条、Join/Prune 消息 684 条、Register 消息 3 680 条、Register-Stop 消息 900 条.因此,图 1 中 t4~t10 的变迁实施概率分别为 0.158,0.019,0.037,0.102,0.549,0.135,0.

路由器处理每种消息的有效代码长度也发生了变化,分别是:Hello 消息 174、Bootstrap 消息 400、Candidate-RP-Advertisement 消息 109、Join/Prune 消息 925、Register 消息 150、Register-Stop 消息 89.与原来假设的路由器处理速度对应,T8~T13 的变迁速率分别为 0.174,0.075,0.278,0.033,0.201,0.340.

仍然假设 T7 的随机模型为默认随机模型,但由于协议报文总数的减少,与原来的协议报文外部到达速率对应,T7 的变迁速率变为 0.009 946.P13~P18 中的平均标记数仍然表示 6 种消息消耗路由器处理负载的相对大小.运行 spnp 软件进行模拟,所得到的结果如图 5 所示.

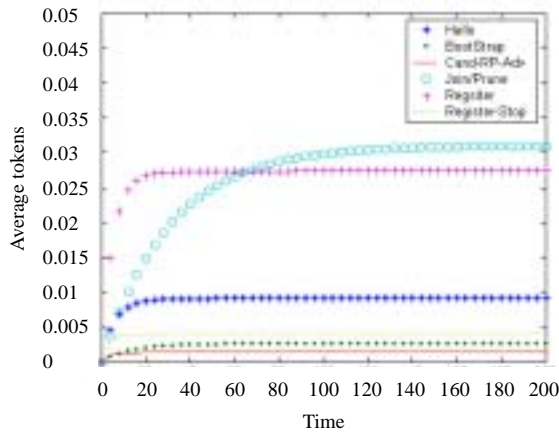


Fig.5 Router processing load caused by 6 types of PIM-SM messages after improvement

图 5 协议改进后 6 种 PIM-SM 消息消耗的路由器处理负载

对比图 3 和图 5 可以看到,经过协议改进和优化处理,处理 Register 消息消耗的路由器负载大为减少,Join/Prune 消息和 Bootstrap 消息所消耗的路由器处理负载也得到了一定程度上的降低.因此,协议的改进和优化大幅度减轻了路由器的处理负载.

4.3 占用的网络带宽

再看看协议改进后的各种消息占用的网络带宽.参照表 3 中的参数,如图 2 所示的网络中在该组播会话周期内所有链路上的各种消息总数分别为 Hello 消息 17 952 条、Bootstrap 消息 4 080 条、Candidate-RP- Advertisement 消息 960 条、Join/Prune 消息 10 344 条、Register 消息 9 200 条、Register-Stop 消息 4 500 条.

经过协议改进,各种消息的最小载荷长度分别为 Hello 消息 12 字节、Bootstrap 消息 76 字节、Candidate-RP- Advertisement 消息 48 字节、Join/Prune 消息 92 字节、Register 消息 8 字节、Register-Stop 消息 44 字节.

因此,网络中各种消息的流量如图 6 所示.

对比图 4 和图 6 可以看出,Bootstrap 消息占用的网络带宽减少了原来的 24%,Join/Prune 消息占用的网络带

宽也减少了 6.3%.因此,协议的改进大幅度地减少了协议控制报文占用的网络带宽.

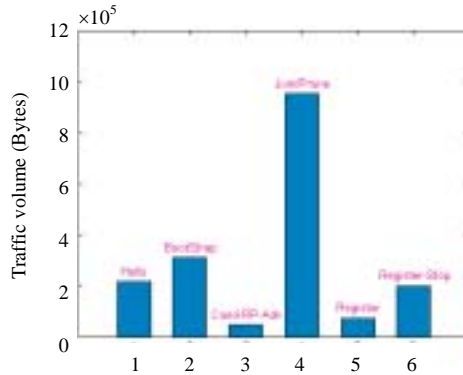


Fig.6 Network bandwidth consumed by 6 types of PIM-SM messages after improvement

图 6 协议改进后 6 种 PIM-SM 消息占用的网络带宽

4.4 付出的代价

根据上面的讨论,协议的改进和优化既减轻了处理协议报文消耗的路由器负载,又减少了协议报文占用的网络带宽,但也付出了相应的代价:(1) 对 Register 消息的优化处理,部分地破坏了 PIM-SM 消息处理程序的模块性;(2) Join/Prune 消息的实时性略有降低;(3) 原来每台路由器上收到 Bootstrap 消息后选举 RP 的操作实际上转移到 BSR 一台路由器在周期性发送 Bootstrap 消息之前来完成.

但是,考虑到 Register 消息本身的特殊性(封装了大量组播数据,并且以组播数据的发送速率到达),Join/Prune 消息过于频繁,以及被配置为 BSR 的路由器本身应该具有的高性能,这些代价都是值得的.经过改进和优化以后,协议性能得到了明显的提高.

5 结 论

PIM-SM 协议是目前首选的域内组播路由协议,但其控制报文负载比较大,因此对协议的建模和性能分析是必要的.本文针对目前这方面研究的不足,首次建立了整个 PIM-SM 协议的 SPN 模型.在该模型的基础上,结合在路由器上对 PIM-SM 协议的实现和测量,从 PIM-SM 协议各种消息的角度对协议性能进行了分析和实验,发现 Register 消息和 Join/Prune 消息消耗较多的路由器处理负载,而 Join/Prune 消息和 Bootstrap 消息占用较多的网络带宽.最后,针对这几种消息对协议进行了改进和优化处理.通过分析和实验表明,这种改进既减轻了路由器的处理负载,也节省了网络带宽,明显提高了协议性能.

References:

- [1] Deering S, Cheriton D. Multicast routing in datagram Internetworks and extended LANs. ACM Trans. on Computer Systems, 1990, 8(2):85-110.
- [2] Bradner S, Mankin A. The recommendation for the IP next generation protocol. RFC1752, 1995.
- [3] Waitzman D, Deering S. Distance vector multicast routing protocol. RFC1075, 1988.
- [4] Moy J. Multicast extensions to OSPF. RFC1584, 1994.
- [5] Adams A, Nicholas J, Siadak W. Protocol independent multicast-dense mode (PIM-DM): Protocol specification (revised). RFC3973, 2005.
- [6] Ballardie A. Core based trees (CBT) multicast routing architecture. RFC2201, 1997.
- [7] Estrin D, Farinacci D, Helmy A. Protocol independent multicast-sparse mode (PIM-SM): Protocol specification. RFC2362, 1998.
- [8] Lin C. Performance Evaluation in Computer Networks and Computer System. Beijing: Tsinghua University Press, 2001 (in Chinese).

- [9] Hardjono T. Router-Assistance for receiver access control in PIM-SM. 2000. <http://csdl2.computer.org/persagen/DLabsToc.jsp?resourcePath=/dl/proceedings/&toc=comp/proceedings/iscc/2000/0722/00/0722toc.xml&DOI=10.1109/ISCC.2000.860719>
- [10] Biswas S, Izmailov R. A QoS-aware routing framework for PIM-SM based IP-Multicast. 1999. <http://www3.ietf.org/proceedings/99jul/slides/pim-framework-99jul/>
- [11] Billhartz T, Cain J, Goudreau E. Performance and resource cost comparisons for the CBT and PIM multicast routing protocols. IEEE Journal on Selected Areas in Communications, 1997,15(3):304-315.
- [12] Lin YD, Hsu NB, Pan CJ. Extension of RP relocation to PIM-SM multicast routing. 2001. <http://speed.cis.nctu.edu.tw/~ydlin/>
- [13] Holt J, Peng WX. Improving the PIM routing protocol with adaptive switching mechanism between its two parse sub-modes. 1998. <http://136.199.54.185/~ley/db/conf/icccn/ icccn1998.html>
- [14] Lu ZF, Wang M, Wang GD. Formal analysis of multicast routing protocol PIM-SM based on extended Petri net. Journal of Yunnan University (Natural Science Edition), 2004,26(2):127-131 (in Chinese with English abstract).

附中文参考文献:

- [8] 林闯. 随机 Petri 网和系统性能评价. 北京:清华大学出版社,2001.
- [14] 陆正福,王敏,王国栋. 多播路由协议 PIM-SM 的扩展 Petri 网形式分析. 云南大学学报(自然科学版),2004,26(2):127-131.



李丹(1981 -),男,四川阆中人,博士生,主要研究领域为计算机网络体系结构,组播,P2P,Overlay 网络.



徐格(1974 -),男,博士,副教授,主要研究领域为计算机网络体系结构,路由器体系结构,路由算法与协议,组播与服务质量控制.



吴建平(1953 -),男,教授,博士生导师,CCF 高级会员,主要研究领域为计算机网络体系结构,协议工程学,互联网络.



刘莹(1973 -),女,博士,副研究员,CCF 高级会员,主要研究领域为计算机网络体系结构,组播路由和算法.



崔勇(1976 -),男,博士,助理研究员,主要研究领域为计算机网络体系结构,服务质量控制,路由算法和性能评价.