第34卷 第20期 计 算 机 工 程 2008年10月
Vol.34 No.20 Computer Engineering October 2008
·网络与通信· 文章编号：1000—3428(2008)20—0112—03 文献标识码：A 中图分类号：TP393

组播路由协议HBH的分析与研究

张新常[1,2]，李晓东[1,3]，王 峰[1,3]，阎保平[1]
(1. 中国科学院计算机网络信息中心，北京 100080；2. 中国科学院研究生院，北京 100049；
3. 中国互联网络信息中心，北京 100080)

摘 要：组播路由协议HBH具有良好的可扩展性且内存需求低，但其对路由路径变化的适应能力及本地组播效率有待改进。该文分析了HBH协议产生上述问题的原因，提出新的组播转发树构建方式，通过模拟试验对其进行验证。结果表明，改进的HBH协议对路由路径变化具有良好的适应能力，并在本地具备较好的组播效果。

关键词：IP组播；组播转发树；HBH协议

Analysis and Study of Multicast Routing Protocol HBH

ZHANG Xin-chang[1,2], LI Xiao-dong[1,3], WANG Feng[1,3], YAN Bao-ping[1]
(1. Computer Network Information Center, Chinese Academy of Sciences, Beijing 100080;
2. Graduate University of Chinese Academy of Sciences, Beijing 100049; 3. China Internet Network Information Center, Beijing 100080)

【Abstract】Hop By Hop(HBH) multicast routing protocol has many invaluable features such as low memory requirement and high scalability, but it can not adjust its forwarding tree well when some routing paths are changed, and local multicast efficiency in HBH is low. This paper analyzes HBH protocol, explains the cause of the above shortages, and brings forward some ways and mechanisms to improve HBH protocol. Experimental result proves that the modified HBH overcomes the drawbacks above.

【Key words】 IP multicast; multicast forwarding tree; Hop By Hop(HBH) protocol

1 概述

IP组播曾被认为是实现组播最有效的途径，但由于其会使路由表表项数量迅速增长[1]，因此没有得到广泛应用。

近年来，作为IP组播的替代方案之一，应用层组播得到了迅速发展，如Narada, Scattercast, Yoid, ALMI, HMTP, NICE。由于应用层组播在端主机之上实现，不需要对网络层做任何修改，因此容易部署。但应用层组播与网络层无关，在网络层上没有任何数据包的转发分支，势必造成数据包的重复传输，从而降低了组播效率。

IP组播的另一个替代方案是BP-based组播，如REUNITE[2], HBH[3], BMP[4]，其中，BP是指组播转发树中的分支节点(向多个节点转发数据包)。在BP-based组播中，数据包的目标地址不采用D类组播地址(或IPv6中的组播地址)，而是采用普通的单播地址，因此，容易在现有Internet上部署。BP-based组播的非BP不需要保存组播路由表MFT，从而可节约路由表资源。HBH(Hop By Hop)是一种典型的BP-based组播，它基于REUNITE的核心理念，并对其进行了优化。本文将对HBH的优缺点进行分析，并作进一步完善。

2 HBH组播路由协议

2.1 HBH概述

在HBH协议中，群组由<*S*, *G*>标识，其中，*S*表示发送方单播地址；*G*表示D类IP组播地址。

HBH中的组播路由器有2个路由表：MCT和MFT。与REUNITE不同的是，在某路由器的MCT或MFT中，路由表表项存储的是下一个分支节点或群组接收者的单播地址，如图1所示，其中，r_1加删除线表示由fusion消息去掉了该表项；不加标记的路由表为MFT。

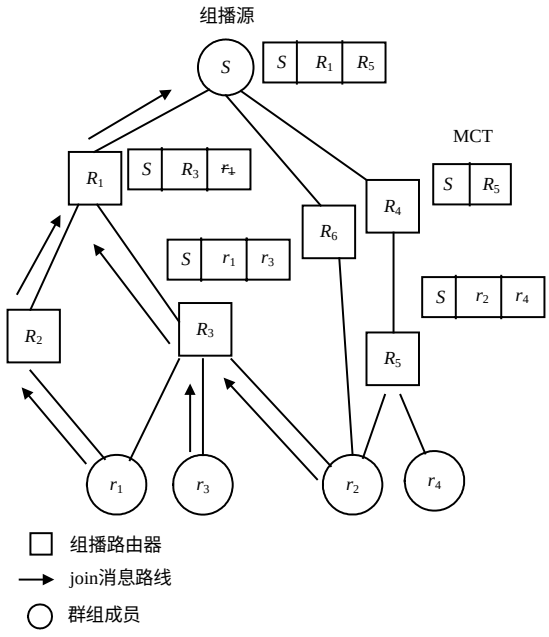


图1 HBH组播转发树

基金项目：中国下一代互联网示范工程基金资助项目(CNGI-04-16-2S)
作者简介：张新常(1975－)，男，博士研究生，主研方向：IPv6，网络协议，组播技术；李晓东、王 峰，博士；阎保平，研究员、博士生导师
收稿日期：2007-12-06 **E-mail：**zhangxinchang@cnnic.cn

HBH协议中的组播路由表表项对应 2 个定时器t_1和t_2。t_1期满后将不再转发数据包，仅转发消息；t_2期满后将不再转发数据包和消息。

HBH 利用 3 种消息完成组播转发树的构建和维护，分别为 join 消息、tree 消息和 fusion 消息。群组接收者或分支路由器周期性地以单播形式向源 S 发送 join 消息，在首次发送(加入群组)时，该消息将到达群组源，在其他发送过程中，该消息可能被该接收成员加入的分支路由器截获(经过该路由器时)。

数据发送方周期性地组播 tree 消息，每条 tree 消息指明了一个接收者，其作用是添加或更新路由表表项。当收到 join 消息则立即发送 tree 消息，以便接收者迅速加入群组。

fusion 消息用于合并相关的路由表表项，使在相关共享链路上只传输一次相同的数据包，具体过程可参见文献[3]。

2.2 HBH 的主要优缺点

HBH 是目前 BP-based 组播中较为完善的组播路由协议，其主要优点如下：

(1)在组播流通过不支持组播的单播网络时，HBH 不采用隧道技术，而是直接利用单播方式。该机制从本质上与单播网络融合，极大地提高了传输效率，扩大了组播的应用范围。

(2)在使用群组时不需要分配一个全球唯一的群组地址，提高了组播的可延伸性。

(3)数据发送方能对群组进行有效控制，不存在未授权用户向群组发送数据包的现象。

(4)在网络拓扑及路由路径稳定的情况下，建立的组播转发树是 SPT 树(REUNITE 不能构建 SPT)，因为 HBH 的组播转发树是由数据发送方向接收者发送的 tree 消息完成的。

(5)源和路由器的MCT和MFT表项是下一个组播转发树的分支路由器或接收者(从拥有该路由表的节点到接收者无分支路由器)。这一优势在文献[3]中没有过多地提及，但可靠组播可以利用此性能提高报文修复性能[5]。

HBH 尚有一些不完善之处，主要表现在以下方面：

(1)存在接收者无法正常接收数据包现象。

以图 1 为例，当R_4-R_5通路中断后，r_2将无法正常接收数据包。因为r_2发送的join消息此时是非首次join消息，所以其被R_3拦截，无法重建组播转发树。尽管R_5得不到及时更新而退出转发树，新的转发树也不会建立。

(2)对路由最短路径的变化适应能力不强。主要原因是在构建好一棵组播转发树后，HBH 利用组播 tree 消息来更新和维护转发树，源在收到 join 消息后有时可以重建转发树，但是非首次 join 消息拦截机制限制了这一能力。

以图 1 为例，假定由于拥塞控制等原因改变了从S到r_2的最短路径，即从S-R_4-R_5-r_2改变为S-R_6-r_2。r_2会向源S定期发送join消息，但被R_3拦截而到达不了源S，因此，无法根据新的路径信息进行相应的更新，从而无法建立新的SPT树。

(3)由于路由表表项是接收者的单播地址，在组播路由器所在的本地网络中有大量接收者的情况下，数据包要在该本地网络上出现 n 次(接收者数量)，从而影响了组播效率。

3 HBH 组播转发树构建改进

针对 HBH 协议的不足之处，本文进行了如下的改进。

3.1 相关消息

由于 HBH 组播转发树的构建是通过一系列消息来完成的，因此为克服上述缺点，取消了定期组播的 tree 消息，并引入 4 个新消息：join_r(restruction), expire, newpath 和 tree_u(unicast)。

join_r 消息与 join 消息基本一致，所不同的是发送时机和处理过程。在 join_r 消息转发过程中，没有拦截机制，其能自由到达数据发送方。另外，数据发送方在收到 join_r 消息后将发送 tree_u 消息，而不是 tree 消息。发送时机有 2 种情况：(1)相关网络中断，接收者检测到较长时间没有收到任何数据包时立即发送；(2)针对可能发生的最短路径改变，接收者或分支路由器定期发送(具体见下文)。一旦接收者发送一个 join_r 消息，则其在相应的重构周期内不再发送 join 消息。此后，该接收者将周期性地发送 join_r 消息，以进行相应的更新。

tree_u 消息与 tree 消息基本相同，但前者是向发送者单播，这点与数据发送方收到 join 消息发送单播的 tree 消息是一样的。路由器对 tree_u 消息进行相应的处理(与处理 tree 消息一样，包括更新所经过的路由器路由表项)，但是其所建立的表项并不立即转发数据包，只转发消息。收到对应 expire 消息后，启动数据包的转发。

expire消息格式为expire(S, R)，其中，S为组播群组标识；R为发送者路由表中t_1到期的表项地址。当某路由器表项的t_1定时器期满时，向数据发送方发送expire消息，由后者向群组组播。当由tree_u消息新建的R表项收到此消息后，启动数据包的转发。

newpath消息格式为newpath(S, R)，其中，S为组播群组标识；R为节点地址。当含有新建R表项的路由器收到expire (S, R)时，向数据发送方发送newpath消息(对应图 2 中的t_2时刻)，由后者向群组组播。收到此消息时，所有到期的R表项从路由表中退出。

修改后的协议中会涉及一些周期，如图 2 所示，其中，t_1对应HBH中表项的t_1定时器周期概念。

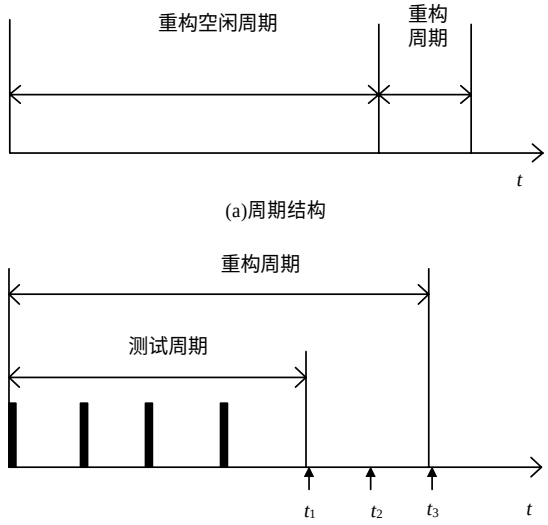


图 2 周期关系

定义 1 重构期限是指新建表项消亡的时间期限，其值大于HBH中表项t_2定时器的周期。在重构过程中，如果原路径得到更新，则新建的表项应该消亡，重构期限指明了其消亡时刻。

定义 2 重构周期是一次 join_r 消息所引起的重构过程所用时间，是从某节点发送 join_r(S, R)消息起，到所有到期的 R 表项消亡或重构期限到期的这段时间。

定义 3 重构空闲周期是 2 次重构过程的时间间隔。

定义 4 测试周期是从某节点发送join_r(S, R)消息起，到某R表项的定时器t_1到期的这段时间。这段周期可防止路径因临时改变而引起组播转发树重构。

3.2 转发树重构建

本文主要以图 1 为例说明重构过程，并假定路径按照上文所述发生了改变。

在某时刻r_2发送join_r(S, r_2)消息启动重构周期，其每隔一定时间发送一次join_r消息。这些消息均到达了数据发送方S，并由后者向r_2单播tree(S, r_2)。由于此时从S到r_2的最短路径是S-R_6-r_2，因此在S-R_4-R_5-r_2路线上的r_2表项得不到更新，而在S-R_6-r_2路径上将构建完新的路径，这条路径现在仅启动转发消息功能。当R_5的r_2表项的t_1到期后，发送expire(S, r_2)消息，在S和R_6(新路径在它们的路由表中新建了表项)收到expire(S, r_2)后，启动新路径转发数据包功能，并向S发送newpath(S, r_2)，由后者以组播的形式向组播通告新路径启动，并做相应的删除和修改工作。当含有相应expire(S, r_2)消息中非新建r_2表项的路由器R_5收到newpath(S, r_2)后，立即停止向该表项对应的地址转发数据包。本文使用fusion消息处理一些合并工作，具体可参见文献[3]。

在上述过程中，只要在测试周期内原有的r_2得到更新，新的路径就在重构期限内得不到expire消息，从而在重构期限到期时将新建路径删除。

3.3 本地组播扩展

如上文所述，在HBH中群组由<S, G>标识，G是D类组播地址。如图 1 所示，在末端网络的组播路由器上(如R_3和R_5)，要对每个加入的组播接收者建立表项。在本地有较多接收者的情况下，其效率显然不高，因为在本地的数据包可以通过硬件级组播来完成，即一个数据包在本地网络仅传输一次。另外，可以利用IGMP对接收者进行有效管理，同时取消接收者发送的join消息，从而提高本地的管理效率和扩展性。

4 模拟试验

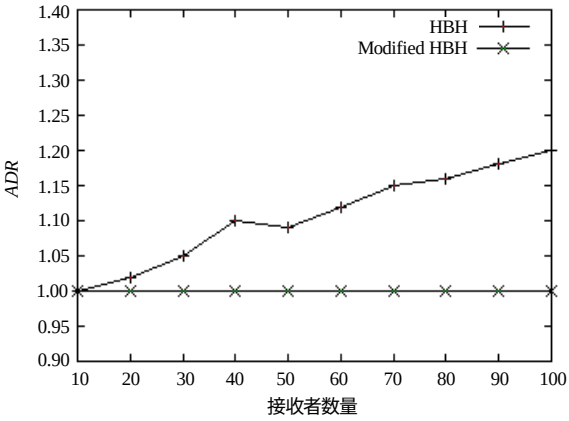
模拟试验用 NS-2.31 完成，所用拓扑用 GT-ITM 生成。拓扑由 5 000 个路由器节点组成，拓扑模型为 trans-stub。主机节点与随机的路由器节点相连接，且每个路由器节点只连接一个主机节点，即忽略本地组播的具体细节。

为了评价 HBH 组播转发树的质量，引入 ADR(Average Delay Ratio)指标：

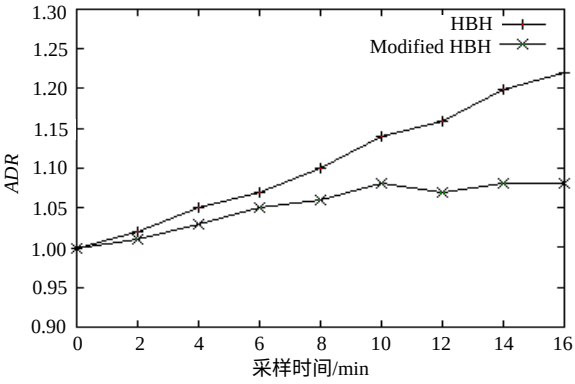
$$ADR=\frac{1}{n}\sum_{i=0}^{n}\left(\frac{\text{从发送方到}r_i\text{的实际距离}}{\text{从发送方到}r_i\text{的最短距离}}\right)$$

其中，n为接收者数量；r_i为第i个接收者。

在网络拓扑中 2 点间最短路径不变的情况下，HBH 构建的转发树是 SPT。为了测试转发树对路径变化的适应能力，试验定期(120 s)中断拓扑中的某些链路(事件发生时一条链路的中断概率为 0.03)。在上述变化环境下，HBH 修改前后的性能比较如图 3 所示，其中，图 3(a)的采样时间为拓扑开始变化后 200 s。结果验证了修改后的协议具有较好的路径变化适应能力。



(a) HBH 修改前后在不同群组下的性能比较



(b) HBH 修改前后适应能力比较

图 3 HBH 修改前后性能比较

5 结束语

由于 IP 组播存在一些缺陷，并且缺乏基础网络的支持，因此无法在 Internet 上得到广泛应用。近年来，为了克服 IP 组播的缺陷和促进组播的实际应用，BP-based 组播作为一种替代方案被提出。本文分析了这类协议中性能较好的 HBH 协议，并进行了改进，使其在相关路径发生变化时能作相应的调整；还对其进行了本地组播扩展，以提高本地组播效率。

参考文献

[1] Diot C, Levine B, Lyles B. Deployment Issues for the IP Multicast Service and Architecture[J]. IEEE Network, 2000, 14(1): 78-88.

[2] Stoica I, Eugene N T, Zhang Hui. REUNITE: A Recursive Unicast Approach to Multicast[C]//Proc. of IEEE INFOCOM'00. San Francisco, USA: IEEE Computer Society Press, 2000.

[3] Costa L K, Fdida S, Duarte O M B. Hop by Hop Multicast Routing Protocol[C]//Proc. of ACM SIGCOMM'01. San Diego, USA: ACM Press, 2001.

[4] Barzoki S. Bag-Mohammadi B, Yazdani M. BMP: An Efficient and Scalable Multicast Protocol[C]//Proc. of Conference on Electrical and Computer Engineering. Ontario, Canada: [s. n.], 2004.

[5] 张新常, 杜学东. 一种可靠组播的报文修复机制[J]. 计算机工程与设计, 2006, 27(16): 3058-3061.