

MPing: A Ping Utility for IP Multicast

[draft-sarac-mping-00.txt](#)

Status of this Memo

This document is an Internet-Draft and is subject to all provisions of [Section 10 of RFC2026](#). Internet Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet Drafts.

Internet Drafts are draft documents valid for a maximum of six months. Internet Drafts may be updated, replaced, or made obsolete by other documents at any time. It is not appropriate to use Internet Drafts as reference material or to cite them other than as a "working draft" or "work in progress."

The list of current Internet Drafts can be accessed at <http://www.ietf.org/ietf/1id-abstracts.txt>

The list of Internet Draft Shadow Directories can be accessed at <http://www.ietf.org/shadow.html>.

Copyright Notice

Copyright (C) The Internet Society (2003). All Rights Reserved.

Abstract

This document describes a mechanism for discovering multicast reachability between end systems within/between multicast enabled networks. It uses request/response messages to verify multicast reachability between the local site and a remote site. With this utility, multicast users can test if they can successfully join a multicast group of a remote source and receiver its data.

1. Introduction

Verifying reachability between multicast enabled networks is an important management task. Multicast data propagates on forwarding paths toward the receivers. As the receivers join and leave multicast groups, these paths are dynamically created and maintained by the routers. In addition, the join mechanism does not use any feedback. As a result, problems due to lack of multicast reachability become difficult to detect.

Current approaches to verify multicast reachability consist of a number of monitoring efforts that depend on voluntary participation from multicast users in collecting the reachability information among a number of multicast enabled sites. These efforts include sdr-monitor of UC Santa Barbara, Multicast Beacon of NLANR, and Multicast Tester of MulticastTech.com. The collected information in these systems give a general reachability picture for the overall multicast infrastructure. Even though these systems are useful for the multicast community in understanding the reachability status of the multicast infrastructure, they do not really help a network administrator to detect and debug a particular reachability problem between his/her network and a remote one. As a result, mechanisms are needed to help network administrators to verify the reachability status between their networks and the remote ones.

In unicast, ICMP Echo Request/Reply based PING tool provides a convenient approach to discover unicast reachability between two systems in the network. For multicast, this mechanism has not been considered very useful. In multicast, PING requests are sent to multicast group address and these requests trigger group receivers to send PING responses to the PINGing host via unicast. This operation informs an end system about the fact that there are a number of receivers that received the PING request on the group address. Contrary to the utility of unicast PING, this mechanism does not help us to verify multicast reachability between two given end systems in the network. In addition, the mechanism is vulnerable to feedback implosion problem. Because of these reasons, this approach is not put into practical use by network operators.

In this draft, we propose a multicast ping utility that can be used to discover multicast reachability between a local site and a remote site. We call this utility MPING. MPING can be used to verify multicast reachability between the local site (as potential receiver) and a remote site (as potential multicast source). This utility can be very helpful for network operators to debug potential multicast reachability problems between their networks and the remote ones.

Sarac, Almeroth

[Page 2]

2. MPING Protocol Operation

The main idea in MPING is to use a well known multicast group address (MPING.MCAST.NET) to exchange an MPING request and response messages between the local host and a remote host. More specifically, MPING involves the below message exchanges between the local host R and a remote host S:

- 1) MPING request: This can be implemented as a source specific join request to (S,MPING.MCAST.NET) multicast group address. Using the source filtering support provided by IGMPv3, R sends a Membership Report that includes a join request to above group address. The Designated Router (DR) at R's site creates a PIM-Join message for (S,MPING.MCAST.NET) and forwards it toward S.

When the join request reaches the DR at S's site, this router needs to inform S about the MPING request. This can be achieved by introducing a new message type to IGMP protocol. When a DR receives a join request specifically for the MPING.MCAST.NET group of a local host, it will use this new message to inform the host about the incoming MPING request.

- 2) MPING response: On receiving the MPING request, the host S creates and sends a MPING response back to the MPINGing host R. The MPING response is a special control message that is sent to (S,MPING.MCAST.NET) multicast group address by S. This control message propagates on the established multicast forwarding path toward R.
- 3) LEAVE message: On receiving the MPING response or on timeout, the MPINGing host R sends a IGMP Membership Report to leave the (S,MPING.MCAST.NET) multicast group. Consequently, the DR at R's site sends a PIM-Leave message toward the S's site to resolve the multicast forwarding path in the network.

The above mechanism helps an end host R to verify multicast reachability between itself as a potential receiver and a remote host S as a potential multicast sender. The MPING response returned by S may include some useful information as follows:

- 1) S may include the list of multicast group addresses that it is currently sourcing data for.
- 2) MPING request can be answered by the DR router at S's site and the response can still contain the above information. This can remove the dependency on the host machines to implement the required MPING functionality in their operating systems. This approach assumes that the router based deployment of MPING is

easier than end host based deployment.

3. Multiple Simultaneous MPING Requests

One issue with the above mechanism is the interaction of multiple simultaneous MPING requests going toward an end host S. Due to the possibility of multiple simultaneous requests, a MPING request may not reach the source site but instead may graft on the corresponding multicast tree in the middle of the network. By the time the MPING request reaches a router on the multicast tree, the MPING response may have already been sent by the queried host and this response may have already crossed this grafting on-tree router. In this situation, the PIM-Join message corresponding the second MPING request may not return a MPING response. As a result, the second host may mistakenly conclude the lack of reachability toward the source.

As an example, consider the example scenario in Figure 1. In this scenario, an end host K sends a MPING request toward a remote end host S. This request creates the multicast forwarding path between K and S in the network which includes a router A. On receiving the request, the end host S sends a MPING response back to K on the (S,MPING.MCAST.NET) multicast group. Assume that just after this MPING response passes the router A toward its destination K, the router A receives a second MPING request that has originated from end host L. Since the router A is already on the multicast tree, instead of forwarding this new MPING request (which is nothing but a PIM-Join message for the (S,MPING.MCAST.NET) multicast group) it will graft this new branch on the tree. However, since the MPING response has already passed the router A, this second MPING request will not result in a MPING response. And this situation may be mistakenly interpreted by the end host L as lack of reachability between L and S.

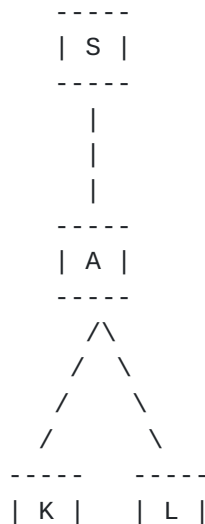


Figure 1. A sample scenario with multiple MPING requests.

In the above situation, before the end host L decides on the lack of reachability, it may need to run several MPING requests toward the end host S. In addition to this, we require the end host S to send multiple MPING responses with some periodicity (every 5 seconds) up to some defined time (up to 10 responses). This will increase the chances for L to receive a response to its MPING request. In the above example scenario, when the end host L sends its MPING request, it will wait for over 5 seconds to receive a response from the end host S. By the time the MPING request of L arrives the router A, if S has not sent out 10 responses yet, the end host L will receive a response for its MPING request. On the other hand, if the end host S already sent out 10 responses, we expect the end host K to receive a MPING response and then send a Leave message to resolve the previously created multicast path between its site toward the S's site. The Leave message will cause the multicast states to be removed in the routers between K and A. At this point, when the end host L waits for 5 seconds without receiving any MPING responses, it will also send a Leave message and this message will resolve the multicast forwarding path between itself and S. Then, the end host L will initiate a second MPING request which will most likely go all the way toward S's site and will result in a MPING response. Note that even if this second MPING request gets grafted on the multicast tree before it reaches S's site (due to potentially another MPING request from another MPINGing end host), because of the repeated MPING responses coming from the end host S, it will most likely receive a MPING response from S.

4. Required Modifications

MPING requires a few modifications to routers and end hosts to be able to realize its function. These modifications are discussed below.

4.1 Required Modifications to End Systems

End systems need to support a new IGMP message type for MPING. On receiving this IGMP message, they will create and send a number of responses to their MPING.MCAST.NET group address. Optionally these responses may include information about the multicast groups that this host is sourcing data currently.

4.2 Required Modifications to Routers

Similar to end systems, edge routers need to be modified to recognize MPING requests and act on them. That is, on receiving a PIM-Join for

(S,MPING.MCAST.NET) for a host S in its directly

Sarac,Almeroth

[Page 5]

attached subnet, a router will create an IGMP message and forward it to the end system to inform it about the MPING request. Optionally, routers may also participate in MPING operation by sending MPING responses on behalf of their local end hosts as discussed in [Section 3](#).

5. Limitations

The MPING mechanism presented in this draft is useful for verifying multicast reachability among remote end hosts. The maximum benefit of this mechanism can be realized with wide scale deployment in the network. In other words, the lack of MPING deployment at a multicast site may cause a MPINGing host to draw wrong conclusions about the reachability status to this site. Compared to end host deployment, we expect that MPING deployment in DR routers is relatively easier. In addition, the service provided by this utility is most useful for network administrators for debugging multicast services in their network. This fact should give strong incentives to network administrators for deploying the service in their routers.

6. Intellectual Property

The IETF takes no position regarding the validity or scope of any intellectual property or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; neither does it represent that it has made any effort to identify any such rights. Information on the IETF's procedures with respect to rights in standards-track and standards-related documentation can be found in [BCP 11](#) [3]. Copies of claims of rights made available for publication and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementors or users of this specification can be obtained from the IETF Secretariat.

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights which may cover technology that may be required to practice this standard. Please address the information to the IETF Executive Director.

7. Full Copyright Statement

Copyright (C) The Internet Society (2003). All Rights Reserved.

This document and translations of it may be copied and furnished to others, and derivative works that comment on or otherwise explain it or assist in its implementation may be prepared, copied, published and distributed, in whole or in part, without restriction of any kind, provided that the above copyright notice and this paragraph are included on all such copies and derivative works. However, this document itself may not be modified in any way, such as by removing the copyright notice or references to the Internet Society or other Internet organizations, except as needed for the purpose of developing Internet standards in which case the procedures for copyrights defined in the Internet Standards process must be followed, or as required to translate it into languages other than English.

The limited permissions granted above are perpetual and will not be revoked by the Internet Society or its successors or assigns. This document and the information contained herein is provided on an "AS IS" basis and THE INTERNET SOCIETY AND THE INTERNET ENGINEERING TASK FORCE DISCLAIMS ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

8. Author's Address

Kamil Sarac <ksarac@utdallas.edu>
Department of Computer Science
University of Texas at Dallas
Richardson, TX 75083, USA

Kevin C. Almeroth <almeroth@cs.ucsb.edu>
Department of Computer Science
University of California Santa Barbara
Santa Barbara, CA 93106, USA

