PIM Working Group
Internet-Draft
Intended status: Standards Track
Expires: September 1, 2012

H. Liu
Huawei Technologies
T. Tsou
Huawei Technologies (USA)

February 29, 2012

PIM MTU Hello Option for PIM Message Encapsulation
draft-lts-pim-hello-mtu-00

Abstract

This memo introduces a new PIM Hello MTU Option which is carried in PIM Hello messages. The MTU option enables interface MTU information to be exchanged among PIM neighbors, and PIM messages to be encapsulated in an efficient and consistent way.

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at http://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on September 1, 2012.

Copyright Notice

Table of Contents

1. Introduction

A PIM router often needs to preserve a great many (*,G) or (S,G) states to enable traffic forwarding for large scale multicast channels. These states are usually set up and kept alive by periodical PIM messages (e.g.PIM Join) sent from its downstream neighbors. For each periodical assembling of these states into a PIM message, multiple packets will possibly be generated due to MTU limitation on the sending PIM interface.

Current implementation uses merely sending link MTU to calculate maximum PIM packet length without considering the receiving interface link MTU of the neighbor. It has some drawbacks because if the MTU of the downstream sending interface is larger than that of the upstream receiving interface, PIM protocol packets encapsulated according to the sending MTU will most possibly be discarded for exceeding the MTU limitation of the upstream receiving interface. The forwarding states cannot be properly established as a result. There are already faults being reported caused by inconsistent MTU configuration among PIM neighbors.

Even though the problem could be resolved by requiring each PIM downstream interface taking less or equal MTU value than its upstream interface, it is inflexible for operation and does not scale because the interface or link conditions across the network might be diverse in practice. As a remedy, this memo recommends exchanging link MTU information among PIM neighbors, and introduces a new PIM MTU Hello option. PIM MTU option is carried in periodical PIM Hello messages. A PIM router uses the option to inform its own receiving link MTU on an interface to its neighbor(s). The neighbor(s) will use the MTU when encapsulating and sending PIM protocol messages to this router.

PIM MTU Option can be applied to all variants of PIM protocols, i.e., PIM-SM, PIM-SSM, PIM-DM, and BIDIR-PIM, on both IPv4 and IPv6 PIM networks. Because MTU issue for unicast Register Message has already been considered in PIM-SM (4.4.1 in [1]), neighboring MTU is only

referred when encapsulating PIM messages with multicast destination.

It should be noted that PIM MTU discovery proposed here is different from multicast PMTU discovery described in RFC1981 [2]. Section 5.2 of RFC1981 requires that an implementation should maintain a single PMTU learned across the whole multicast distribution tree. This might result in using smaller packets than necessary for a lot of paths. And because the end to end paths can be very dynamic this could make the effort too complex. PMTU is used in encapsulating a 'multicast data packet' (not a 'PIM protocol packet' as here) to avoid fragmentation as the packet travels on the paths of the tree. Whereas PIM MTU option works in control plane and has a per-hop

nature - it only functions between one-hop PIM neighbors and helps PIM protocols to establish correctly the multicast forwarding states.

2. Terminology

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [3].

3. MTU Option and its Operation Rule

To record the minimum sending MTU value on an interface, a new General Purpose non-group-specific state (say Sending MTU state) is introduced in PIM protocols (for General Purpose State referring to 4.1.1 of [1] and [4], and 3.1.1 of [5]. It is 32-bit long and is unique on an interface even if what is connected is a multi-access network. The initial value of the Sending MTU state should be set to the outbound MTU, or if unavailable, set to the default MTU of the interface.

When an MTU Hello Option is received from a neighbor, the PIM router parses the MTU value in the option and decides whether or not it should accept the value and should store it in the Sending MTU field. A router should not accept too small a value to prevent extreme fragmentation deteriorating the router's performance. If the MTU value is valid from a legal neighbor, it compares the value with the MTU value currently stored in the Sending MTU field, and makes the

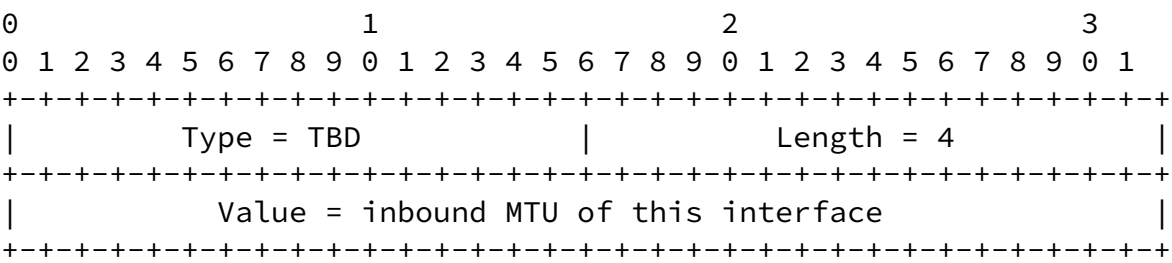
replacement if the former is less than the latter.

Unlike other PIM Hello option, MTU Option is not required being supported simultaneously by all PIM neighbors connecting to a network. An MTU-capable router only considers the MTU of a trusty neighbor from which a valid MTU option is received. An MTU-capable PIM router should use MTU option in its Hello message, and should keep the Sending MTU state to the initial value if no neighbor reports a valid MTU Option. Finally, an MTU-incapable router should ignore an MTU option on reception.

The Sending MTU state should be checked before sending a multicast PIM message, to ensure the length of the message does not exceed the MTU limit of both the sending and receiving links. It should be noted that as a convention, the length calculation starts from the beginning of an IP header.

4. Option Format

A Hello MTU Option has the following format:



Type: to be assigned by IANA if this option is accepted. The field is 16-bit long.

Length: the length of the Value field. The field is 16-bit long.

Value: inbound MTU value for this interface. The field is 32-bit long.

5. IANA Considerations

The Type field should be allocated by IANA if MTU option is accepted.

6. Security Considerations

The potential security threat for MTU option should be the denial-of-service attack of extremely fragmenting PIM messages, by advertising much smaller MTU value than necessary. A remedy is to require a PIM router to check the validity of a neighbor's MTU value before accepting it.

7. Acknowledgements

The authors would like to acknowledge Hou Yunlong and Mach Chen for their valuable comments on the work.

8. Normative References

[1] Fenner, B., Handley, M., Holbrook, H., and I. Kouvelas, "Protocol Independent Multicast - Sparse Mode (PIM-SM): Protocol Specification (Revised)", RFC 4601, August 2007.

[2] McCann, J., Deering, S., Mogul, J., and L. Vicisano, "Path MTU

Discovery for IP version 6", RFC 1981, August 1996.

[3] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", March 1997.

[4] Adams, A., Nicholas, J., and W. Siadak, "Protocol Independent Multicast - Dense Mode (PIM-DM): Protocol Specification (Revised)", RFC 3973, January 2005.

[5] Handley, M., Kouvelas, I., Speakman, T., and L. Vicisano, "Bidirectional Protocol Independent Multicast (BIDIR-PIM)", RFC 5015, October 2007.

Authors' Addresses

Liu Hui
Huawei Technologies
Building Q14, No.156, Beiqing Rd.
Beijing 100095
China

Phone: 8610-60610012
Email: helen.liu@huawei.com

Tina Tsou
Huawei Technologies (USA)
2330 Central Expressway
Santa Clara CA 95050
USA

Phone: +1 408 330 4424
Email: Tina.Tsou.Zouting@huawei.com