Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: January 3, 2019

Mankamana. Mishra
Stig. Venaas
Cisco Systems
Mahesh. Sivakumar
juniper networks
Zheng(Sandy). Zhang
ZTE Corporation
Mikael. Abrahamsson
July 2, 2018

PIM Designated Router graceful shutdown
draft-mankamana-pim-graceful-dr-shutdown-00

Abstract

On a multi-access network, one of the PIM routers is elected as a Designated Router (DR). On the last hop LAN, the PIM DR is responsible for tracking local multicast listeners and forwarding traffic to these listeners if the group is operating in PIM-SM. In case of a network maintenance, where we want to bring down the current DR, there is currently no way to gracefully handover the PIM DR role to a new DR on the shared LAN. In this document, we propose a modification to the PIM-SM protocol that allows PIM DR to gracefully shutdown or go down for maintenance. We also provide a procedure for PIM DR to gracefully handover its role to a new PIM DR in the network.

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 https://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 January 3, 2019.

Table of Contents

1. Introduction

On a multi-access LAN such as an Ethernet, one of the PIM routers is elected as a DR. The PIM DR represents the LAN segment/broadcast domain in the PIM topology tree and has two roles to play in the PIM-SM protocol. For sources connected to the segment, the PIM DR is responsible for registering one or more active sources with the Rendezvous Point (RP) if the group is operating in PIM-SM. In addition, on the last hop LAN, the PIM DR is responsible for tracking local multicast listeners and forwarding data traffic to these listeners if the group is operating in PIM-SM.

Consider the following last hop LAN in Figure 1:

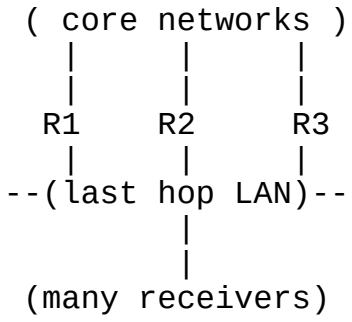

Figure 1: Last Hop LAN

Assume R1 is elected as the Designated Router. According to [RFC4601], R1 will be responsible for forwarding traffic to that LAN on behalf of any local members. In addition to keeping track of IGMP and MLD membership reports, R1 is also responsible for initiating the creation of source and/or shared trees towards the sources or the RPs.

If R1 needs to go on planned maintenance, the current approach is to lower the DR priority which would make sure that another PIM router on the LAN gets elected as the new DR and starts forwarding multicast traffic.

With this approach, R1 gives away DR role as soon as new priority is configured and a new PIM DR (lets assume R3) starts building a multicast tree and starts forwarding multicast traffic on the LAN. However, this could cause traffic disruption for the duration it takes for R3 to build the upstream multicast tree.

This draft defines a mechanism in the PIM protocol to handover DR role gracefully and as a result minimize traffic disruption.

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 [RFC2119] .

With respect to PIM, this document follows the terminology that has been defined in [RFC4601] and [RFC7761] . Many places this draft would refer to PIM RFC [RFC4601] but it MUST be considered [RFC7761] as well.

3. Protocol Specification

In this draft, we define a new hello option to enable the graceful handover of a DR during planned maintenance.In Section 3.1, we describe the proposed mechanism. In Section 3.2, we evaluate the impact of the mechanism on the network under different conditions. Section 4 describes the proposed hello option.

3.1. Proposed Mechanism

1. In Figure-1, assume that R1 is current PIM DR that needs to go on planned maintenance. R1 MUST sends out a PIM Hello with option described in Section 4. The DR Priority MUST be set to 0. R1 MUST also set its assert metric to (PIM_ASSERT_INFINITY - 1)

2. The PIM assert metric modification would make sure that R1 does not become an assert winner

3. Sending DR priority as 0 would make sure to have default transition in case new DR does not support the new specification

4. The current PIM DR (R1 here) MUST not stop forwarding traffic to intended receivers unless it starts getting duplicate flows from newly elected PIM DR.

5. A failsafe timer SHOULD be used to stop forwarding multicast traffic towards receiver. It SHOULD be set to at least two PIM Hello intervals. But it SHOULD also be a configurable value.

3.2. Impact on the network

This section covers impact of PIM hello with Section 4 option

3.2.1. Every PIM router supports the new specification on the shared LAN

1. In Figure-1, if each of the PIM routers on shared LAN supported this specification, new DR election would be done as per [RFC4601]

2. The newly elected DR MUST start building the multicast tree towards the source/RP. It MUST start fail safe timer (default value 2 PIMHello interval) and MUST not generate a data driven assert. Once the timer expires, it can move back to the default assert mechanism. The reason to avoid an assert is to allow the old PIM DR on LAN to forward multicast traffic until such time the new DR is completely ready to forward multicast traffic.

3. It MUST forward multicast flow to receivers as soon as it gets the multicast flow from the source/RP

3.2.2. Hybrid shared LAN, some of PIM router does not support specification

There are two cases to consider,

1. If the new DR supports this specification, it would follow Section 3.1

2. If the new DR does not support this specification, there is no need for any special handling as the new DR would take over as it does today. It would assert as soon as it gets elected as DR and the old DR would become the assert loser as it had already adjusted its assert metric to PIM_ASSERT_INFINITY - 1

4. PIM Hello option

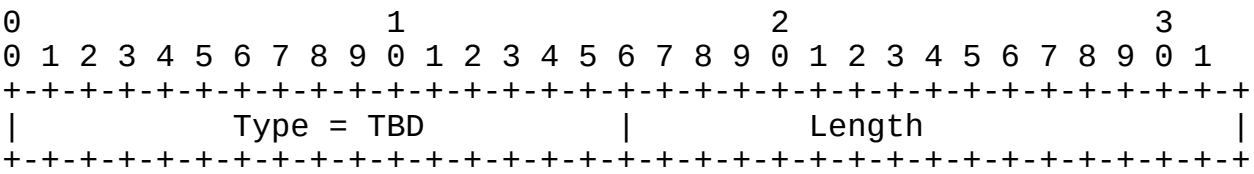

Figure 2: Graceful DR handoff Hello Option

where

Type : DR Graceful handoff

Length: 2

5. IANA Considerations

A new PIM Hello option is TBD..

6. Security Considerations

Security of the new PIM Hello Options is only guaranteed by the security of PIM Hello message, so the security considerations for PIM Hello messages as described in PIM-SM [RFC4601] apply here.

7. Acknowledgement

8. Contributors

In addition to the authors listed on the front page, the following co-authors have also contributed to original idea.

Krishna Muddenahally Ananthamurthy

Cisco Systems

Sameer Gulrajani

Cisco systems

Rishabh Parekh

Cisco systems

9. References

9.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <https://www.rfc-editor.org/info/rfc2119>.

[RFC4601] Fenner, B., Handley, M., Holbrook, H., and I. Kouvelas, "Protocol Independent Multicast - Sparse Mode (PIM-SM): Protocol Specification (Revised)", RFC 4601, DOI 10.17487/RFC4601, August 2006, <https://www.rfc-editor.org/info/rfc4601>.

[RFC6395] Gulrajani, S. and S. Venaas, "An Interface Identifier (ID) Hello Option for PIM", RFC 6395, DOI 10.17487/RFC6395, October 2011, <https://www.rfc-editor.org/info/rfc6395>.

[RFC7761] Fenner, B., Handley, M., Holbrook, H., Kouvelas, I., Parekh, R., Zhang, Z., and L. Zheng, "Protocol Independent Multicast - Sparse Mode (PIM-SM): Protocol Specification (Revised)", STD 83, RFC 7761, DOI 10.17487/RFC7761, March 2016, <https://www.rfc-editor.org/info/rfc7761>.

9.2. Informative References

[HELLO-OPT] IANA, "PIM Hello Options", IANA PIM-HELLO-OPTIONS, March 2007.

Authors' Addresses

Mankamana Mishra
Cisco Systems
821 Alder Drive,
MILPITAS, CALIFORNIA 95035
UNITED STATES

Email: mankamis@cisco.com

Stig Venaas
Cisco Systems
821 Alder Drive,
MILPITAS, CALIFORNIA 95035
UNITED STATES

Email: svenaas@cisco.com

Mahesh Sivakumar
juniper networks
1133 Innovation Way
Sunnyvale, CALIFORNIA 94089
UNITED STATES

Email: sivakumar.mahesh@gmail.com

Zheng(Sandy) Zhang
ZTE Corporation
No. 50 Software Ave, Yuhuatai Distinct
Nanjing
China

Email: zhang.zheng@zte.com.cn

Mikael Abrahamsson

Email: swmike@swm.pp.se