

Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: 4 January 2026

M. Mishra
A. Budhiraja
J. Neema
Cisco Systems
I. Romdhani
Edinburgh Napier University
G. Mishra
Verizon Communications Inc. (VZ)
3 July 2025

PIM Backup Designated Router Procedure
draft-ietf-pim-bdr-03

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 this document, we propose mechanisms to minimize traffic losses during DR re-election. This document introduces the concept of a Backup Designated Router on a shared LAN and a procedure to perform Delayed DR election. A backup DR on LAN would be useful for faster convergence, whereas a Delayed DR election would be useful for minimizing traffic losses when new routers join a LAN.

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 4 January 2026.

Copyright Notice

Copyright (c) 2025 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<https://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Revised BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Revised BSD License.

Table of Contents

1. Introduction	2
2. Terminology	4
3. Applicability and deviation from draft PIM DR Improvement . .	4
4. Protocol Specification	4
4.1. PIM Backup DR (BDR) election procedure	4
4.2. Existing PIM DR failure	5
4.3. Existing PIM BDR failure	5
4.4. New PIM Router addition in network	5
4.4.1. Delayed DR Election Algorithm	6
4.5. Benefit	7
5. PIM Hello Message Format	8
6. Backward Compatibility	8
7. Manageability Considerations	8
8. IANA Considerations	8
9. Security Considerations	9
10. Acknowledgement	9
11. Normative References	9
Authors' Addresses	9

1. Introduction

On a multi-access LAN such as an Ethernet, one of the PIM routers is elected as a DR. The PIM DR has two roles in the PIM-SM protocol. On the first hop network, the PIM DR is responsible for registering an active source with the Rendezvous Point (RP) if the group is operating in PIM-SM. On the last hop LAN, the PIM DR is responsible for tracking local multicast listeners and forwarding to these listeners if the group is operating in PIM-SM.

Consider the following last hop LAN in Figure 1:

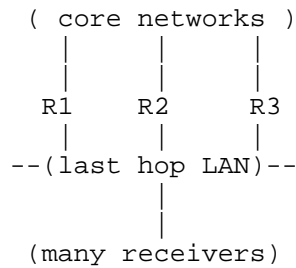


Figure 1: Last Hop LAN

Assume R1 is elected as the Designated Router. According to [RFC7761], R1 will be responsible for forwarding traffic to that LAN on behalf of any local member. 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 senders or the RPs.

There are multiple reasons for why network could potentially trigger DR re-election. Some of the reasons are

1. R1 going down
2. Access interface towards shared LAN going down
3. Config changed with lower DR priority
4. A new router (R4) joins the LAN with higher DR priority

When any of above network event occurs, PIM DR re-election would be triggered. Unfortunately, the DR election procedure under section 4.3.2 of RFC 7761 can lead to unnecessary traffic loss in certain scenarios. Two different scenarios are described based on figure 1 to illustrate potential issues with DR re-election and the approach this draft suggests to solving them - Initial state: All the three routers are in the LAN network and R1 is the elected DR.

1. If R1 router fails, multicast packets are discarded until either R2 or R3 is elected as DR, and it assumes the multicast flows on the LAN. To address this requirement, the draft introduces a backup DR election procedure which would minimize traffic loss during PIM DR failure.
2. If a new router R4 with higher DR priority than R1 joins the LAN network, then it will be elected as the new DR even before it is ready to pull multicast traffic. R1 will stop forwarding

multicast traffic and the flow will not recover until R4 assumes them. To address this scenario, the draft introduces a procedure for Delayed DR election.

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] .

Backup Designated Router(BDR): A PIM router which is second best to the PIM DR router. A single BDR is elected per interface like the DR.

Designated Router Other (DROther): A router which is neither a DR nor a BDR

With respect to PIM, this document follows the terminology that has been defined in [RFC7761] .

3. Applicability and deviation from draft PIM DR Improvement

[I-D.ietf-pim-dr-improvement] defines a procedure to partly solve same problem which was stated in the introduction section of this draft.

[I-D.ietf-pim-dr-improvement] introduces two new PIM Hello options for election of backup PIM DR. This draft provides mechanism to elect BDR without using any new PIM Hello options.

[I-D.ietf-pim-dr-improvement] makes the elected DR sticky in nature. This draft honours the DR priority of each router (old or new) and performs the DR and BDR election in a non-sticky manner, while still minimizing the traffic loss and allowing faster convergence.

4. Protocol Specification

4.1. PIM Backup DR (BDR) election procedure

[RFC7761] defines procedure for PIM DR election. PIM DR is elected on interface "I" among all PIM routers for which "I" has received PIM Hello. The current design tells that DR election between 2 PIM routers on an interface "I" will happen based on DR priority and PIM neighbor's IP address. The PIM router with highest priority for interface "I" will be chosen as DR. If priorities are same for both PIM routers on interface "I" then PIM neighbor having highest IP address on interface "I" wins the election. BDR election follows the exact same procedure and the router with second best DR priority or

second highest IP address on interface "I" will be elected as PIM BDR.

BDR would perform each of the responsibility of PIM DR except it would not forward traffic on shared LAN or send a register message to avoid duplication.

4.2. Existing PIM DR failure

When PIM DR fails, PIM BDR MUST immediately take over all the responsibilities of PIM DR and start forwarding multicast traffic on shared LAN. A PIM DR re-election is triggered on shared LAN, which will select a new PIM BDR, and the current BDR will become the new DR.

4.3. Existing PIM BDR failure

When an existing PIM BDR fails, the shared LAN MUST have BDR re-election using the BDR election procedure specified in section 4.1

4.4. New PIM Router addition in network

When a new PIM router is added in a shared LAN and has the highest PIM DR priority configured, if this a new router starts propagating its configured DR priority right away, the existing PIM DR would give up its role. Then there would be potential traffic loss till the new DR learns about membership states and builds a multicast tree to the source or RP.

To avoid any such traffic loss situation, the new router SHOULD participate in the DR re-election after a logical delay. To achieve this, the draft introduces the logic of DR priority 0 and a new PIM Hello option of Delayed DR Priority. A PIM router advertising a value of 0 in its DR Priority Option would mean that it is not participating in the DR election (even though it may be capable of becoming DR on the LAN). Other Routers on the LAN, who receive Hello message with DR priority 0 SHOULD not include this router in the DR election procedure stated in RFC 7761 yet.

Following steps SHOULD be followed when a router starts or interface comes up :

- a) When a router first starts or its PIM interface is enabled, it SHOULD send a Hello message with DR priority 0 and start a timer set to Hello_interval [RFC 7761]. At this point, the router considers itself in a DROther state and is explicitly telling its peers that it is not participating in DR election yet. During this time, the router SHOULD not build any multicast trees.

- b) When the router receives Hello messages from other routers on the same shared-media LAN, the router checks the value of DR priority. The advertised DR-priority can be either a Zero or Non-Zero value. During the first Hello interval, if the router receives any Hello with a non-zero DR-priority, the router assumes that it is joining an already stable LAN network with an elected DR and it starts the process of entering in a Delayed DR Election Phase. However, during the first Hello interval, if the router receives all Hello messages with DR priority Zero, the router assumes it to be a bring up phase.
- c) At the expiry of the first Hello_Interval, the router enters one of two phases - Bring-Up phase or Delayed DR election phase. Based on the phase, the router follows the respective DR-election process.
- d) In the Bring-up phase, the router follows the DR-election procedure mentioned in RFC 7761, i.e. the Router starts advertising its configured DR priority (or default DR priority) in the Hello message and the router with higher DR-priority or IP address wins the DR-election.
- e) In the Delayed DR election phase, the router first executes the algorithm defined in section 4.4.1. After that, the router takes one of the roles in LAN: DR, BDR, or DROther

4.4.1. Delayed DR Election Algorithm

A router starts this algorithm when it enters Delayed-DR election phase as specified in section 4.4 (b). In this phase, the router determines whether it is eligible to become the new DR on the LAN, by comparing the received DR priorities with self DR priority. So, the PIM router will be in one of the below defined roles.

- a) New PIM router eligible to be PIM DR on shared LAN
 - 1. The router MUST send Hello message with DR priority 0, and also send its real DR priority in a new PIM Hello Option - Delayed_DR_Priority.
 - 2. The router SHOULD start a Delayed_DR_Election_Timer. This timer SHOULD have way to be user configured, with default value of (2 x PIM_Hello_Interval) or (2 x IGMP Query Interval) (whichever is higher).

3. While the timer is running, and even though the router is propagating DR priority 0, it MUST start building the multicast tree towards the source/RP. The router SHOULD also program itself to drop any multicast traffic received during the Delayed_DR_Election_Timer.
4. During the course of Delayed_DR_Election_Timer, if the router receives any Hello message with DR priority 0 and the Delayed_DR_Priority higher than self, then it SHOULD - a) Restart its Delayed_DR_Election_Timer AND b) Prune the multicast tree and accept that there is another router in the network which will win the DR election after the Delay_timer_expiry. Note: The restarting of Delayed_DR_Election_Timer is required so that the current router doesn't expire its timer before the soon-to-be-DR-winner Router.
5. After the timer expires, the router SHOULD advertise its real priority in the DR priority option and start the DR election as per RFC 7761. The Delayed_DR_Priority option will become a no-op in the presence of a non-zero DR_priority option.
6. If the router wins the election, it SHOULD start forwarding traffic; otherwise, prune the multicast tree.

With this approach, for a brief amount of time, there would be multiple copies of flows present in the multicast core, but a user SHOULD be able to configure whether to send hello with 0 priority or a configured priority. Depending on the application tolerance (Traffic loss Vs Extra traffic in core) the operator can choose option whichever is suitable for network.

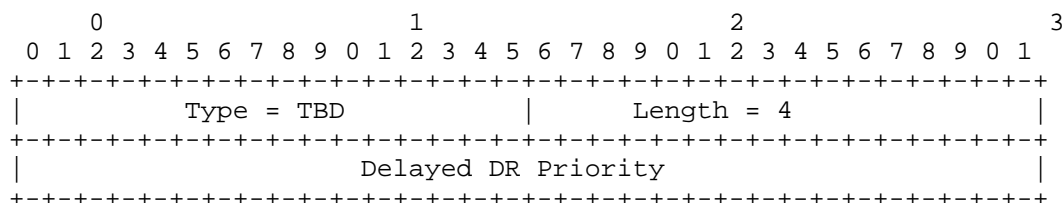
- b) New PIM router eligible to be PIM BDR on shared LAN In this role, the router SHOULD follow the exact same procedure defined in section 4.1.
- c) New PIM router is not eligible to be PIM DR or BDR on shared LAN In this role, the router SHOULD follow the same procedure for DR election as mentioned in RFC 7761. The router assumes the role of DROther. As soon as the router knows that its not eligible to be PIM DR/BDR on the LAN, it SHOULD start sending its original DR priority without waiting for the next hello interval.

4.5. Benefit

1. Easy to implement as it uses an existing PIM procedure to elect DR.

2. Honors the configured DR priority of the router for selecting the DR winner.

5. PIM Hello Message Format



OptionType : The value is TBD.

OptionLength : 4 bytes.

Delayed DR Priority : is a 32-bit unsigned number and SHOULD be considered in the Delayed DR election as described in Section 4.4. This Delayed DR priority option SHOULD be considered valid for any DR election procedure, only if Type 19, DR priority option has a value of 0. However, if DR-priority (Type 19) has a non-zero value, then the Delayed DR priority option value MUST be ignored.

6. Backward Compatibility

TBD

7. Manageability Considerations

TBD

8. IANA Considerations

IANA is requested to allocate a new code point from the "PIM-Hello Options" registry.

Type	Name	Reference

TBD	Delayed DR Priority Option	[this document]

9. Security Considerations

TBD

10. Acknowledgement

The author would like to thank Stig Venaas, Tharak Abraham, Anish Kachinthaya, Anvitha Kachinthaya and Aravind for helping with original idea.

11. 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>>.
- [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>>.
- [I-D.ietf-pim-dr-improvement] Zhang, Z., hu, F., Xu, B., and M. P. Mishra, "Protocol Independent Multicast - Sparse Mode (PIM-SM) Designated Router (DR) Improvement", Work in Progress, Internet-Draft, draft-ietf-pim-dr-improvement-14, 4 December 2022, <<https://datatracker.ietf.org/doc/html/draft-ietf-pim-dr-improvement-14>>.

Authors' Addresses

Mankamana Mishra
Cisco Systems
821 Alder Drive,
MILPITAS, CALIFORNIA 95035
United States
Email: mankamis@cisco.com

Anuj Budhiraja
Cisco Systems
821 Alder Drive,
MILPITAS, CALIFORNIA 95035
United States
Email: abudhira@cisco.com

Joya Neema
Cisco Systems
Bangalore
India
Email: jneema@cisco.com

Dr Imed Romdhani
Edinburgh Napier University
United Kingdom
Email: I.Romdhani@napier.ac.uk

Gyan S. Mishra
Verizon Communications Inc. (VZ)
13101 Columbia Pike FDC1 Rm 304-D
Silver Spring MD 20904
United States
Email: gyan.s.mishra@verizon.com