                 PIM Backup Designated Router Procedure
                      draft-mankamana-pim-bdr-01

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 a mechanism to elect backup DR on a shared
   LAN. A backup DR on LAN would be useful for faster convergence.
   This draft introduces the concept of a Backup Designated Router (BDR)
   and the procedure to implement it.

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 April 25, 2019.

Copyright Notice

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 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.

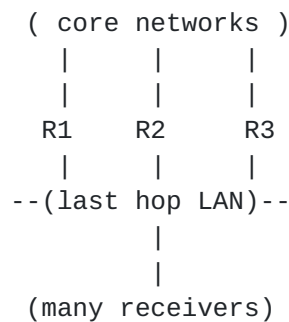   Consider the following last hop LAN in Figure 1:

```
                    ( core networks )
                     |     |     |
                     |     |     |
                    R1    R2    R3
                     |     |     |
                  --(last hop LAN)--
                           |
                           |
                    (many receivers)
```

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 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

When any of above network event occurs, PIM DR re-election would be
triggered. When a new DR is elected in shared LAN, new DR would be
responsible to build a multicast tree towards source / RP. There are
some cases, where traffic is crucial and the operator wants to have
minimum traffic loss with DR failure. To address this requirement,
this draft introduces a backup DR election procedure which would
minimize traffic loss during PIM DR failure.

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] .

BDR - PIM Backup DR

With respect to PIM, this document follows the terminology that has
been defined in [RFC4601] .

3. Applicability and deviation from draft PIM DR Improvement

[I-D.ietf-pim-dr-improvement] defines procedure to solve same problem which was stated in the introduction section of this draft. [I-D.ietf-pim-dr-improvement] introduces new PIM Hello options for election of backup PIM DR.

This draft provides mechanism to elect BDR without using any new PIM Hello.

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. BDR election follows the exact same procedure and the second best PIM DR on shared LAN to be chosen as BDR on interface "I"

BDR would perform each of the responsibility of PIM DR except it would not forward traffic on shared LAN.

4.2. Existing PIM DR failure

When PIM DR fails, PIM DR re-election is triggered on shared LAN. Since BDR is second best DR in LAN, it MUST take over immediately and MUST start forwarding multicast traffic on shared LAN.

Again on a shared LAN, new BDR would be elected. and current BDR would be 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 DR election procedure from [RFC7761].

4.4. New PIM Router addition in network

When a new PIM router is added in shared LAN, It could be either one of the below defined roles.

4.4.1. New PIM router eligible to be PIM DR on shared LAN

When a new PIM router is added in a shared LAN and has the highest PIM DR priority configured, if 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, new PIM router SHOULD send
a PIM Hello with priority 0. After 2 (default value, SHOULD have way
to configure) PIM Hello interval or IGMP Query Interval (Which ever
is higher) it SHOULD start propagating its original configured DR
priority.

Even though a new PIM router propagating its priority as 0, it MUST
start building a multicast tree towards source / RP, This is So that
traffic loss could be minimized once it starts sending Hello with
configured DR priority.

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.

After a PIM Hello or IGMP Query interval, the network would get
stable with only one DR and one BDR.

4.4.2. New PIM router eligible to be PIM BDR on shared LAN

It SHOULD follow the exact same procedure defined in the previous
section.

4.4.3. New PIM router is not eligible to be PIM DR or BDR on shared LAN

First a PIM Hello MUST be sent with priority 0. Once it has gotten
Hello from other PIM neighbors, it knows that it is not eligible to
be PIM DR or BDR. It MUST send configured PIM DR priority
immediately. It MUST not wait for next hello interval.

4.5. Initial case, All new PIM router coming up in shared LAN

In this case, initially each of the PIM routers would send Hellos
with priorities of 0. If a PIM router receives all Hellos with
priorities 0, it MUST send out a Hello with a configured PIM DR
priority. Since it is initial startup case, it would take up to one
Hello interval to converge.

[4.6](). Benefit

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

2. Does not introduce any new Hello option

[5](). Compatibility

[6](). Manageability Considerations

[7](). IANA Considerations

[8](). Security Considerations

[9](). Acknowledgement

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

[10](). Normative References

[I-D.ietf-pim-dr-improvement]
          Zhang, Z., hu, f., Xu, B., and m. mishra, "PIM DR
          Improvement", [draft-ietf-pim-dr-improvement-04]() (work in
          progress), December 2017.

[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]()>.

[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]()>.

Author's Address

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

   Email: mankamis@cisco.com