Abstract

Graceful Insertion and Removal (GIR) of routers is often adopted by many network administrators as an alternative to ISSU. This document discusses various scenarios, requirements and possible solutions to make PIM gracefully shut down and isolate the multicast router with minimal network disruption when a router goes through maintenance procedures.

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 December 30, 2018.

Copyright Notice

Copyright (c) 2018 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 Simplified BSD License text as described in Section 4.e of
1. Introduction

When a network administrator wishes to perform maintenance activity on a router, a system maintenance mode command need to be configured. This isolates the router from the network by gracefully shutting down various protocols running on the router. Multicast protocols will also require graceful migration in order to achieve minimal traffic disruption when a maintenance activity is performed on a router. This document proposes a possible solution to perform GIR with PIM routers gracefully.

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 [RFC7761].

3. Applicability

The proposed change described in this specification applies to PIM routers only.

4. Graceful RPF change

Multicast routing protocol uses Unicast reachability information to find unique Reverse Path Forwarding Neighbor (RPF). Any change in unicast routing triggers multicast RPF changes. Multicast flows need
to change the RPF in a graceful manner to have minimal or no disruption in traffic flow. To achieve graceful RPF change, PIM should not change RPF immediately following unicast routing change. PIM should join the new path and wait for the traffic to arrive on the new path before pruning the old path. Until the packets arrive on the new path, the packets are accepted and forwarded on the old path. Since we have not changed the RPF to new one, we would see RPF failures. The RPF failures on the new path will indicate that the flow is available on the new path, upon which the RPF for the flows will be changed from old to new. Using this method, we will be able to achieve non-stop forwarding of multicast traffic thereby minimizing traffic disruption. The graceful RPF change, however, is not advisable in a normal RPF change scenario. This is because old path could be down due to link failures and the RPF change may take more time which could increase convergence time. Multicast flows can do a graceful RPF change in a GIR scenario since the flow will be available via the old path.

5. GIR removal procedure for a PIM router

A multicast router undergoing a graceful insertion/removal must indicate the same to all the routers in PIM domain. This will ensure that all routers will gracefully change RPF for the multicast flows within the GIR window. This information needs to be propagated before the unicast metrics are altered by the GIR router. To achieve this, a PIM Flooding Mechanism message (PFM) [RFC8364] TLV is originated from the router undergoing GIR. The GIR details will be carried in the PFM TLV. This message is flooded periodically in the PIM domain and the RECOMMENDED interval to send this message is 60 secs. The propagation of this message will ensure that all the routers (in the PIM domain) knows the router undergoing GIR and can gracefully migrate flows from old path to new path when the unicast infinite metrics are advertised from the router undergoing GIR in the GIR window.

Procedure for PIM routers (GIR mode)-

1. The router undergoing GIR (GIR router) will send a PFM message with a new TLV option (GIR TLV) to all its PIM neighbors indicating that the router wishes to go to maintenance mode. The router could send more than one PFM message so that the loss of the PFM messages are minimized. The value fields in the TLV will be populated with the following hold-time values -

   1. graceful-rpf-start - This value indicates the seconds until the PIM router will start doing graceful RPF change
2. graceful-rpf-stop - This value indicates the seconds after which the PIM router will stop doing graceful RPF change. The time period between the graceful-rpf-start and graceful-rpf-stop indicates the duration during which the routers in the network will do graceful RPF changes for multicast flow.

2. Upon receipt of the PFM message with GIR TLV option from GIR router, PIM neighbors will compute the RPF towards the originator ip address in the incoming PFM message. If the RPF matches with the interface where this message is received, the router will perform the following, else pim will just drop the message.

1. Start a timer for graceful-rpf-start, if not already started. Once the graceful-rpf-start timer expires, the routers will be in graceful RPF change mode until the hold-time period stop. During this time period, router will do Graceful RPF change (as described in section above) as soon as it receives unicast metric change. The unicast infinite metric change from the router undergoing GIR has to be sequenced between the advertised graceful-rpf-start and graceful-rpf-stop.

2. Forward PFM message with the GIR TLV to its immediate PIM neighbors. This is to propagate PFM message with the GIR TLV to all the routers in the PIM domain.

The above method could be further optimized if PFM messages could carry the source prefixes of the multicast states present on the GIR router. The routers receiving the PFM GIR message can examine the source prefixes and move to Graceful RPF change mode only if the routers have the multicast state for the source prefixes. With this, not all routers needs to do Graceful RPF change. This will ensure Graceful RPF change occurs only on the routers that are impacted by the router undergoing GIR.

6. GIR insertion procedure for a PIM router

The same method as described above will be used to gracefully insert a router with no traffic disruption after system maintenance mode. When a router is inserted into network, it will send the PFM GIR message with the graceful-rpf-start timer value set to 0, and graceful-rpf-stop timer set to seconds until Graceful RPF stop. During this window, the inserted router will start bringing up the unicast protocols. Every router in the PIM domain will examine the PFM message and do a Graceful RPF change for the window specified in the message. Once the GIR router is inserted and fully operational, it should send at least one message with both timers (graceful-rpf-start and graceful-rpf-stop) set to 0.
7. Compatibility

The router undergoing GIR must set the Transitive bit to 1 in the PFM message so that when a router that doesn’t support GIR, receives a PFM GIR TLV, it will forward the message to the PIM neighbors.

8. PIM GIR TLV

```
+---------------+---------------+---------------+---------------+
| 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 |
+---------------+---------------+---------------+---------------+
|   +---------------+---------------+---------------+---------------+
|   |       Type = TBD       |       Length = 8       |
|   +---------------+---------------+---------------+---------------+
|   |                   Graceful RPF Start |                   |
|   +---------------+---------------+---------------+---------------+
|   |                   Graceful RPF Stop |                   |
|   +---------------+---------------+---------------+---------------+
```

Figure 1: PIM GIR TLV

where

Type : TBD

Length : The length of the value in octets

Graceful RPF Start : Timer value in seconds until PIM router start doing graceful RPF change

Graceful RPF Stop : Timer value in seconds after which the PIM router will stop doing graceful RPF change

9. IANA Considerations

A new PIM PFM option is TBD for GIR.

10. Security Considerations

Security of the new PFM TLV is only guaranteed by the security of PFM message, so the security considerations for PFM message as described in [RFC8364] apply here.

11. Normative References


Authors’ Addresses

Ramakrishnan Chokkanathapuram
Cisco Systems, Inc.
Tasman Drive
San Jose CA 95134
United States of America

Email: ramaksun@cisco.com

Raunak Banthia
Cisco Systems, Inc.
Tasman Drive
San Jose CA 95134
United States of America

Email: rbanthia@cisco.com