

Network Working Group
Internet-Draft
Intended status: Informational
Expires: August 26, 2013

W. Zhou
Cisco Systems
February 22, 2013

VRRP PIM Interoperability
draft-zhou-pim-vrrp-01.txt

Abstract

This document introduces VRRP Aware PIM, a redundancy mechanism for the Protocol Independent Multicast (PIM) to interoperate with Virtual Router Redundancy Protocol (VRRP). It allows PIM to track VRRP state and to preserve multicast traffic upon failover in a redundant network with virtual routing groups enabled.

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 August 26, 2013.

Copyright Notice

Copyright (c) 2013 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 (<http://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 the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction	3
2. Tracking and Failover	4
3. PIM Assert Metric Auto-Adjustment	5
4. DF Election for BiDir Group	6
5. Tracking Multiple VRRP Groups on an Interface	7
6. Support of HSRP	8
7. Security Considerations	9
8. Acknowledgments	10
9. Informative References	11
Author's Address	12

1. Introduction

Virtual Router Redundancy Protocol (VRRP) [RFC5798] is a redundancy protocol for establishing a fault-tolerant default gateway. The protocol establishes a framework between network devices in order to achieve default gateway failover if the primary gateway becomes inaccessible .

PIM has no inherent redundancy capabilities and its operation is completely independent of VRRP group states. As a result, IP multicast traffic is forwarded not necessarily by the same device as is elected by VRRP. The VRRP Aware PIM feature provides consistent IP multicast forwarding in a redundant network with virtual routing groups enabled.

In a multi-access segment (such as LAN), PIM designated router (DR) election is unaware of the redundancy configuration, and the elected DR and VRRP master router (MR) may not be the same router. In order to ensure that the PIM DR is always able to forward PIM Join/Prune message towards RP or FHR, the VRRP MR becomes the PIM DR (if there is only one VRRP group). PIM is responsible for adjusting DR priority based on the group state. When a failover occurs, multicast states are created on the new MR elected by the VRRP group and the MR assumes responsibility for the routing and forwarding of all the traffic addressed to the VRRP virtual IP address. This ensures the PIM DR runs on the same gateway as the VRRP MR and maintains mroute states. It enables multicast traffic to be forwarded through the VRRP MR, allowing PIM to leverage VRRP redundancy, avoid potential duplicate traffic, and enable failover, depending on the VRRP states in the device.

2. Tracking and Failover

With VRRP Aware PIM enabled, PIM listens to the state change notifications from VRRP and automatically adjusts the priority of the PIM DR based on the VRRP state, and ensures VRRP MR (if there is only one VRRP group) becomes the DR of the LAN. If there are multiple VRRP groups, the DR is determined by user-configured priority.

PIM triggers communication between upstream and downstream devices upon failover in order to create mroute states on the new MR. Depending on the requirements, there are various implementation options:

- o PIM sends additional PIM Hello message using the VRRP virtual IP addresses as the source address for each active VRRP group when a device becomes VRRP Active. The PIM Hello will carry a new GenID in order to trigger other routers to respond to the failover. When a downstream device receives this PIM Hello, it will add the virtual address to its PIM neighbor list. The new GenID carried in the PIM Hello will trigger downstream routers to resend PIM Join messages towards the virtual address. Upstream routers will process PIM Join/Prunes (J/P) based on VRRP group state.
- o An alternative solution is to have all passive routers maintain mroute states and record the GenID of current MR. When a passive router becomes MR upon switchover, it uses the existing mroute states and the recorded MR GenID in its Hello message. This solution avoids resending PIM J/P upon switchover and eliminates the requirement of additional PIM Hello with virtual IP address.

If the J/P destination matches the VRRP group virtual address and if the destination device is in VRRP active state, the new MR processes the PIM Join because it is now the acting PIM DR. This allows all PIM Join/Prunes to reach the VRRP group virtual address and minimizes changes and configurations at the downstream routers side.

3. PIM Assert Metric Auto-Adjustment

It is possible that, after VRRP active switched from A to B; A is still forwarding multicast traffic which will result in duplicate traffic and PIM Assert mechanism will kick in. PIM Assert with redundancy is enabled.

- o If only one VRRP group, passive routers will send a large penalty metric preference (PIM_ASSERT_INFINITY - 1) and make MR the Assert winner.
- o If there are multiples VRRP groups configured on an interface, Assert metric preference will be (PIM_ASSERT_INFINITY - 1) if and only if all VRRP groups are in passive.
- o If there is at least one VRRP group is in Active, then original Assert metric preference will be used. That is, winner will be selected between routers using their real Assert metric preference with at least one active VRRP Group, just like no VRRP is involved.

4. DF Election for BiDir Group

Change to DF offer/winner metric is handled similarly to PIM Assert handling with VRRP.

- o If only one VRRP group, passive routers will send a large penalty metric preference in Offer (`PIM_BIDIR_INFINITY_PREF- 1`) and make MR the DF winner.
- o If there are multiples VRRP groups configured on an interface, Offer metric preference will be (`PIM_BIDIR_INFINITY_PREF- 1`) if and only if all VRRP groups are in passive.
- o If there is at least one VRRP group is in Active, then original Offer metric preference to RP will be used. That is, winner will be selected between routers using their real Offer metric with at least one active VRRP Group, just like no VRRP is involved.

5. Tracking Multiple VRRP Groups on an Interface

User can configure PIM to track more than one VRRP groups on an interface. This allows other applications to exploit the PIM/VRRP interoperability to achieve various goals (e.g., load balancing). Since each VRRP groups configured on an interface could be in different states at any moment, the DR priority is adjusted. PIM Assert metric and PIM Bidir DF metric if and only if all VRRP groups configured on an interface are in passive (non-Active) states to ensure that interfaces with all-passive VRRP groups will not win in DR, Assert and DF election. In other words, DR, Assert, DF winner will be elected among the interfaces with at least one Active VRRP group.

6. Support of HSRP

Although there are differences between VRRP and Hot Standby Router Protocol (HSRP) [RFC2281] including number of backup (standby) routers, virtual IP address and timer intervals, the proposed scheme can also enable HSRP aware PIM with similar switchover and tracking mechanism described in this draft.

7. Security Considerations

The proposed tracking mechanism has no negative impact on security.

8. Acknowledgments

I would like to give a special thank you and appreciation to Stig Venaas for his ideas and comments in this draft.

9. Informative References

- [RFC2281] Li, T., Cole, B., Morton, P., and D. Li, "Cisco Hot Standby Router Protocol (HSRP)", RFC 2281, March 1998.
- [RFC5798] Nadas, S., "Virtual Router Redundancy Protocol (VRRP) Version 3 for IPv4 and IPv6", RFC 5798, March 2010.

Author's Address

Wei Zhou
cisco Systems
Tasman Drive
San Jose, CA 95134
USA

Email: weizho2@cisco.com

