Multi-point BFD in VRRP use case
draft-mirsky-bfd-p2mp-vrrp-use-case

Greg Mirsky gregimirsky@gmail.com

Jeff Tantsura jefftant.ietf@gmail.com

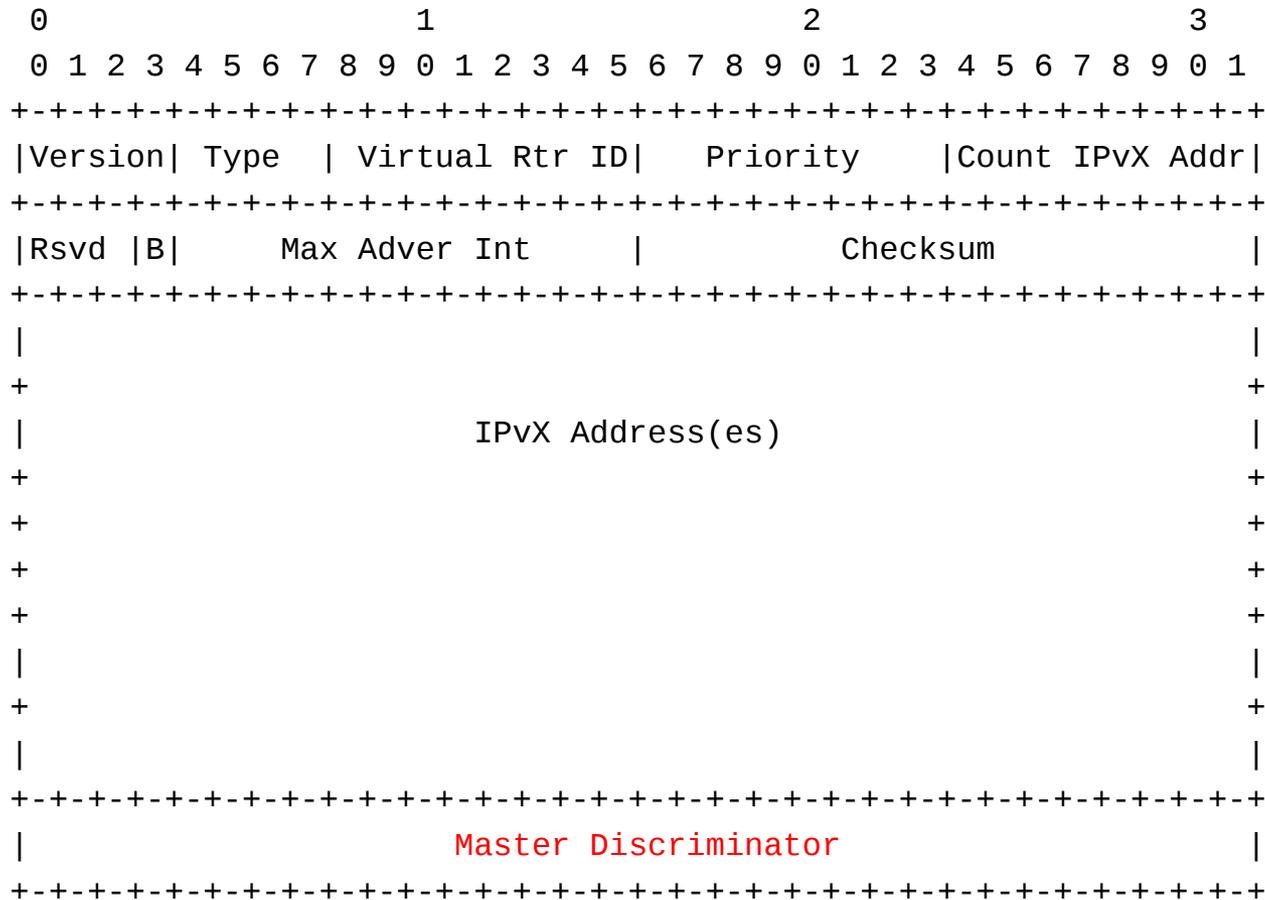IETF-100 November, 2017, Singapore

Motivation (recap)

- RFC 5798 supports 10 msec Advertisement_Interval thus enabling Master_Down_Interval to be at ~33 msec

- Backup router monitors Master

- Master is not aware of identity of Backup router(s)

- Monitoring availability of the Master is another incarnation of Continuity Check OAM

- RFC 5880 defined BFD

- Asynchronous mode can be used by a Backup to monitor availability of the Master

- Concerns with use of Async RFC 5880:

  – Master doesn't know identities of Backup routers in the given VRID on the same LAN segment

  – Master is not really interested in Backup state change. That can be mitigated by Master setting its bfd.RequiredMinRxInterval to 0

Proposed solution with update

- BFD for Multipoint Networks (draft-ietf-bfd-multipoint) extends BFD for use in multipoint and multicast networks

- Master advertises its BFD My Discriminator in VRRP control packet (extension to RFC 5798) to bootstrap p2mp BFD session

- Backup creates MultipointTail session and uses My Discriminator value from VRRP control packet and Master IP address to demultiplex BFD control packets

- Thus the p2mp BFD session per VRID has Master router as root and Backup routers as tails. The Master uses one of IP addresses associated with the VRID as source IP address for p2mp BFD control packets

- In p2mp BFD there's no P-F sequence with interval negotiation and Your Discriminator must always be zeroed

VRRP Control packet to support BFD

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|Version| Type  | Virtual Rtr ID|   Priority    |Count IPvX Addr|
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|Rsvd |B|    Max Adver Int      |           Checksum            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
+                                                               +
|                       IPvX Address(es)                        |
+                                                               +
+                                                               +
+                                                               +
+                                                               +
|                                                               |
+                                                               +
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                    Master Discriminator                       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

How is suppose to work (with update)?

- Master router
  - creates BFD session of MultipointHead type
  - allocates My Discriminator:
    - begins transmitting VRRP control packets with My Discriminator;
    - begins transmitting BFD control packets with:
      - Virtual Router MAC address, as specified in RFC 5798, as source MAC address
      - one of Virtual router IPv4/IPv6 addresses as source IP address
- a Backup router receives VRRP control packet with My Discriminator and creates new BFD session of MultipointTail type
- the tail BFD node demultiplexes BFD control packets to particular VRID based on source IP address and My Discriminator
- when a tail detects that MultipointTail is in down state, i.e. the Master router is not Up, it informs the Backup router which may elect itself as new Master router

Next steps

- Comments are welcome
- As extension to VRRP proposed, would RTG WG be more suitable to continue this work?
- Ask for WG adoption