

draft-ietf-pim-drlb

Problem Statement

- PIM DR is elected based on DR priority or IP address (per RFC4601)
- In the last hop LAN, only one router, the DR, is responsible for forwarding
- Forwarding load is not distributed
- Failover takes longer time
 - All forwarding states must be rebuilt on the new DR after a failover



Solution Overview

- Elect multiple forwarders on the last hop LAN
 - Each is called a GDR (Group DR)
 - Hashing is used to determine which candidate GDR becomes the GDR
- Forwarding load is now distributed
- During a failover only a subset of the forwarding states need to be rebuilt



Protocol Change

- GDR Election
 - DR election procedures remain unchanged
 - A router announces hash masks in new Hello Option TLV to indicate its capability
 - Hash masks include RP, Group, Source
 - Modulo hashing is specified
 - More can be added if needed
 - All candidate GDRs must have the same DR priority as the DR
 - DR announces the list of candidate GDRs and the hash masks to be used on this LAN



Protocol Change

- Creating forwarding states
 - Upon receiving IGMP reports, a candidate GDR runs a hash algorithm to determine if it is the GDR for the RP of the group, the group or/and source
 - If it is, it becomes the forwarder on the LAN
 - Forwarding states are recalculated if the list of the candidate GDR changes or the hash masks change (per announcement by DR)



Protocol Change

- GDR Assert
 - Used to reduce packet loss during GDR state change
 - A GDR becoming non-GDR MAY choose not to remove the oif immediately
 - This will lead to both routers forwarding
 - GDR asserts using “normal” metric
 - Non-GDR asserts using `PIM_ASSERT_INFINITY - 1`
 - Rely on consistent GDR state to converge



Draft Status

- First presented at IETF 82, Taipei, accepted as WG draft-ietf-pim-drlb-00
- Last updated at Vancouver IETF 88, changes proposed:
 - Introduce “algorithm” field in DRLBC option to allow flexibility in algorithm selection.
 - Specify modulo in the default to achieve load balancing, after comparing with BSR.
 - If “interface ID” option presents in Hello, use “Router ID”, instead of “interface IP address”, to calculate hash.



Other changes made in -03

- Add a 32-bit mask in Modulo hash to simplify the calculation for IPv6.
- Adjust Assert metric to $(\text{PIM_ASSERT_INFINITY} - 1)$ on non-GDR
- Suggest DRLBC and DRLBGDR type as 34 (0x22) and 35 (0x23)



Next Step

- Implementation is on the way
- Ready for LC

