

Agenda

- Problem Statement
- Solution Overview and Applicability
- Protocol Change
- Example
- Discussion

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

Applicability

- Last hop only
 - First hop router is determined by incoming Join not DR state
 - The complexity of supporting this at FHR outweighs the benefits of distributing load for sending registers
- SM/SSM/DM only
- Bidir TBD (probably not)

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
 - 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

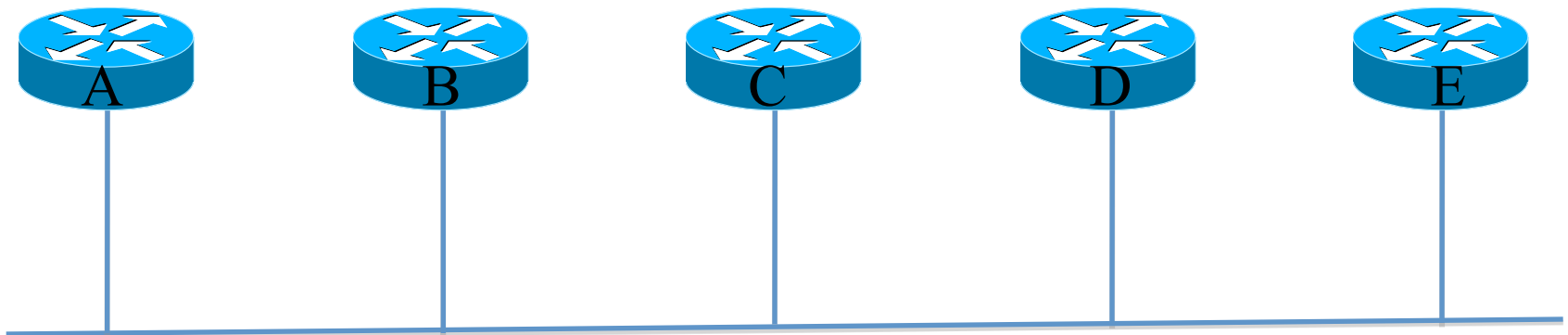
- Creating forwarding states
 - Upon receiving IGMP reports, a candidate GDR runs a hash algorithm to determine if it is the GDR for the RP of the group, the group and/or source
 - If it is, it becomes the forwarder on the LAN
 - Forwarding states are recalculated if the list of the candidate GDR changes 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 Assert
 - GDR state is preferred before using IP address as a tie-breaker
 - See discussion for non-ECMP case

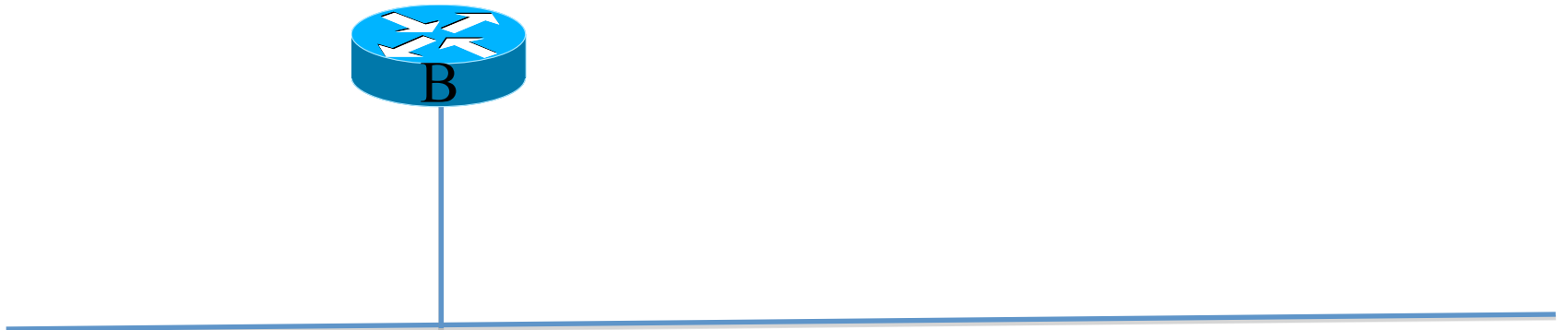
Example For SM (239.0.0.0/8)

- A, B, C, D have the same DR priority
- A would become the DR (per RFC4601)
- RP mask is 0.0.0.0
- Group mask is 0.255.0.0 (IPv4)
- Source mask is 0.0.0.0



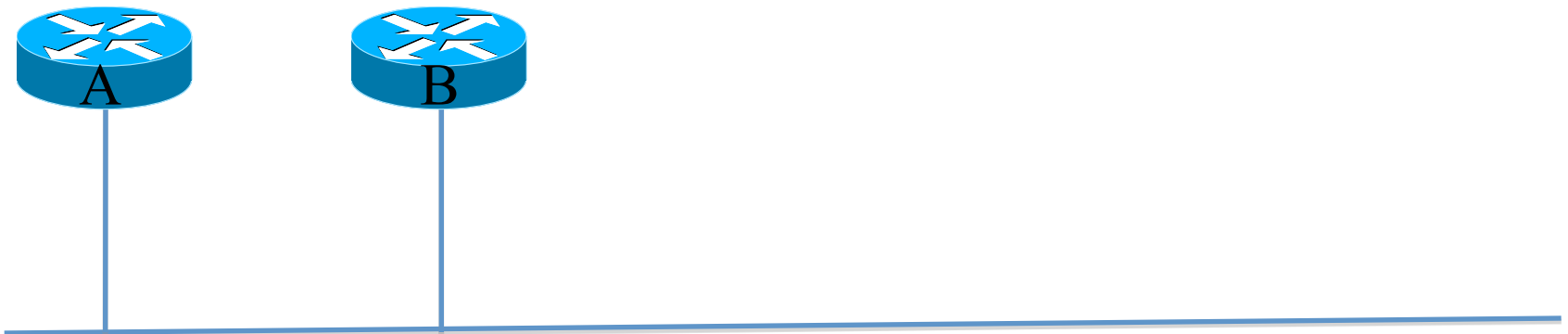
Example For SM (239.0.0.0/8)

1. B comes up first, announces mask (0, 0.255.0.0, 0)
2. When B becomes the DR, it also includes its own address in the LB TLV



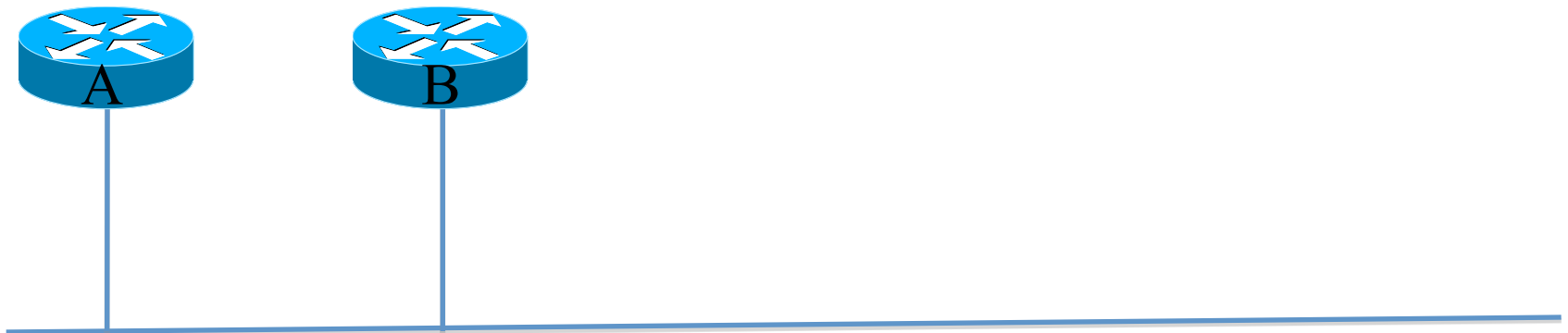
Example For SM (239.0.0.0/8)

3. A comes up, announces mask (0, 0.255.0.0, 0) first
4. A wins DR election, announces both A and B in the LB TLV.
5. B loses DR election, announces mask only
 - Some groups for which B isn't GDR anymore may experience traffic disruption



Example For SM (239.0.0.0/8)

6. Since there are only two routers, A will be the GDR for groups with hash value 0
 - E.g., 239.0/16, 239.2/16, ...
7. B will be the GDR for groups with hash value 1
 - E.g., 239.1/16, 239.3/16 ...



Example For SM (239.0.0.0/8)

8. E comes up, announces mask (0, 0.255.0.0, 0)
9. E loses DR election due to DR priority
 - Continues to announce mask (0, 0.255.0.0, 0)
 - A will not include E in the LB TLV

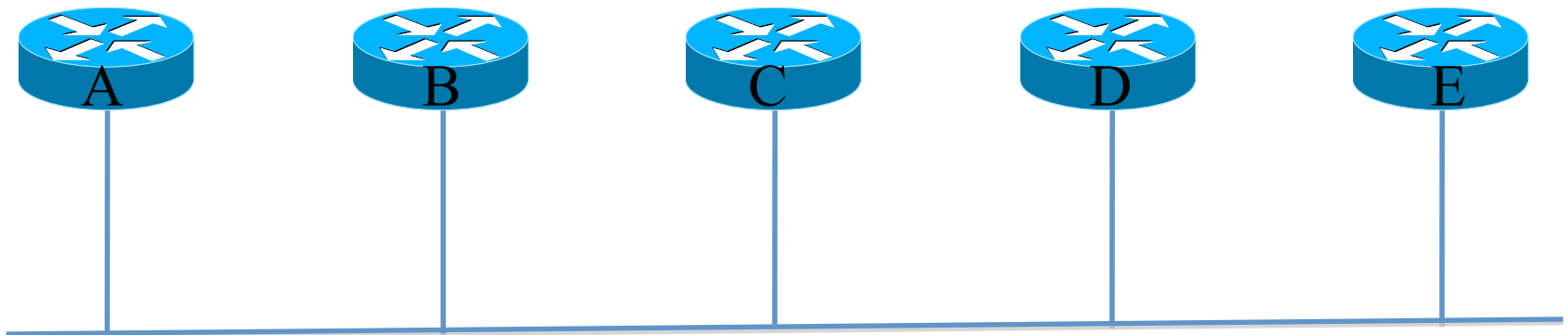


Example For SM (239.0.0.0/8)

10. C and D come up, announce mask (0, 0.255.0.0, 0)

11. A now announces A, B, C, D in LB TLV

- A will be GDR for 239.{0, 4, 8...}.0.0/16
- B will be GDR for 239.{1, 5, 9...}.0.0/16
- C will be GDR for 239.{2, 6, 10...}.0.0/16
- D will be GDR for 239.{3, 7, 11...}.0.0/16

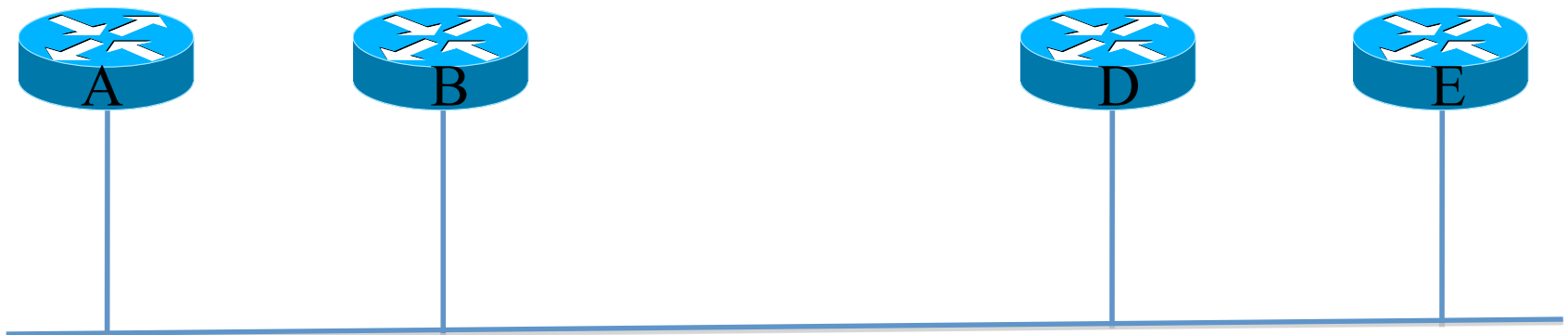


Example For SM (239.0.0.0/8)

12. C is shutdown, or its DR priority decreases

13. A now announces A, B, D in LB TLV

- A will be GDR for 239.{0, 3, 6...}.0.0/16
- B will be GDR for 239.{1, 4, 7...}.0.0/16
- D will be GDR for 239.{2, 5, 8...}.0.0/16



Discussion

- Welcome comment/collaboration, in particular,
 - Hash masks priority between RP mask and Group mask
 - Hash algorithm
 - GDR Assert: should metric be considered first after GDR state?
- Working Group document?
- Thank You!