

PIM-BIDIR RP Resiliency

Jeffrey Zhang, Kurt Windisch, Jaroslaw Adam Gralak

Juniper Networks

88th IETF, Vancouver

PIM-BIDIR Pros & Cons

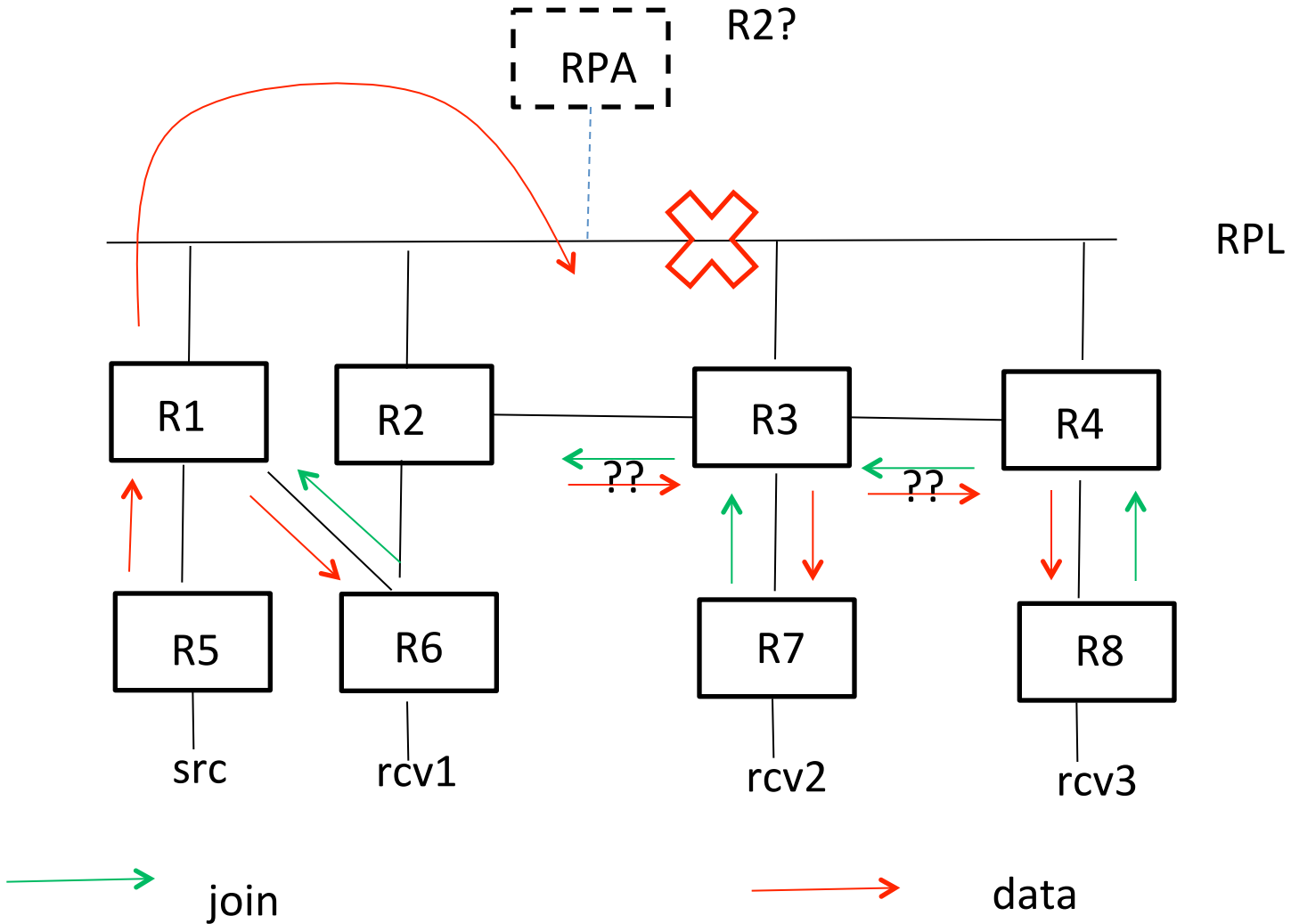
- Simple, powerful and scalable
 - 100% control-driven, no data events
 - RPT only, no source states, no RPT->SPT switch
 - No registration/encapsulation needed
 - native forwarding via RPL
 - Some RP resiliency built in
- Single Point of Failure (RPL) – somewhat

Existing RP Resiliency

- Traffic converges toward the RPA and get exchanged at the routers on the path closest to the sender and receiver
- RPA can be a just a reachable address
 - Not a real router
 - no single point of failure at router level
- As long as RPL is not partitioned itself, or partitioned from some parts of the network, things will work fine

Partitioned RPL

What if R3 & R4 realize that RPL is partitioned so it stops treating it as RPL, and sends join towards R2?



Deal with RPL Partitioning

- Detect partitioning and elect one partition as active partition to act as RPL
- Routers on the active partition advertise host route for RPA
 - R1~R4 all do so before/without the partition
 - R3/R4 stops advertising that after the partition
 - Ensure optimal path towards the RPA (via host route)
- Routers on the in-active partition subnet stops treating it as RPL
 - DF election, among other things

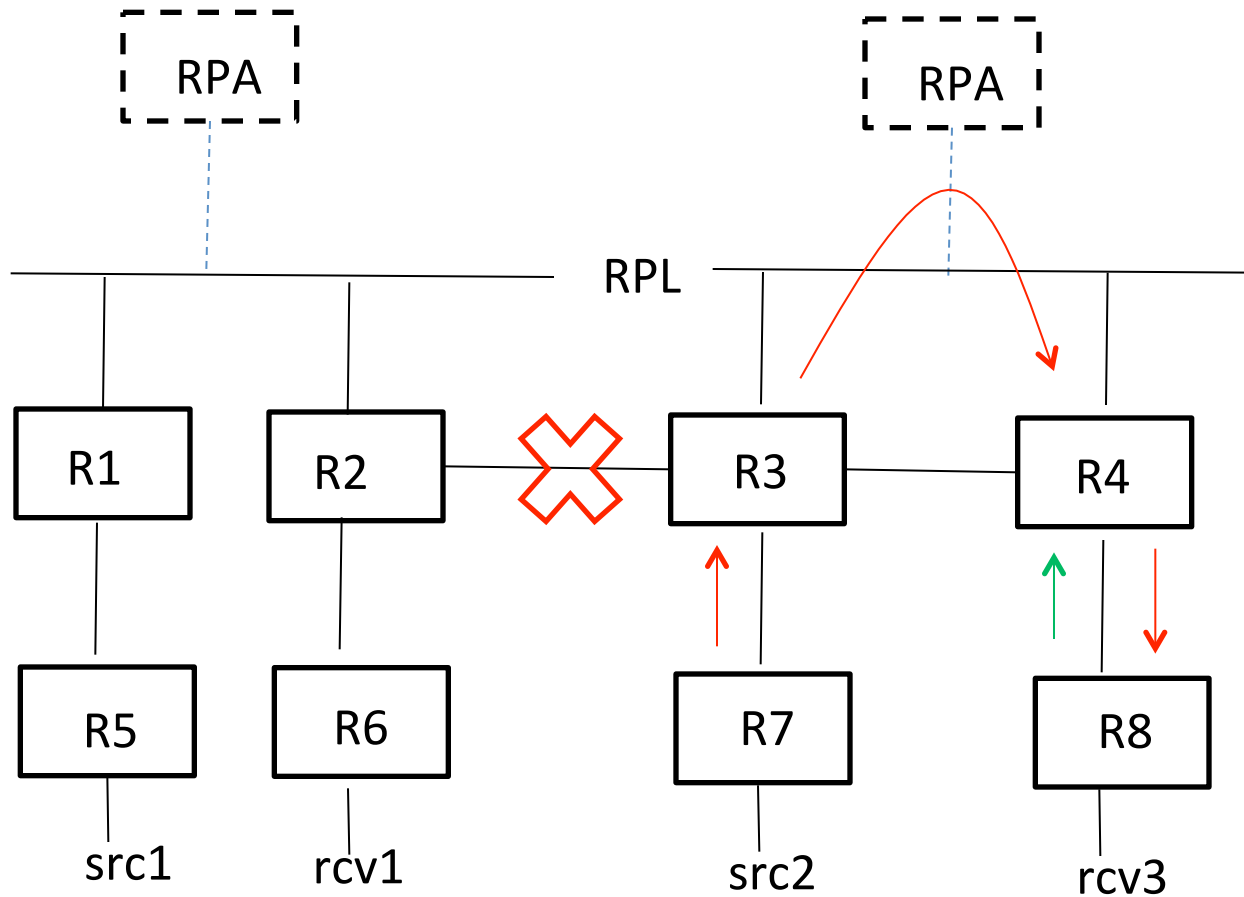
Detection & Election via IGP

- One or more DRs and Net LSAs for the RPL subnet
 - Potentially one or more Stub links as well
 - From routers that are not adjacent to the DR
 - One for each partition
- Routers on the RPL subnet pick the lowest Advertising Router to represent the active partition
 - of all reachable Net LSA for the RPL subnet, or,
 - of all reachable Router LSA with the stub link for the RPL subnet
 - When there is no Net LSA for the subnet
- Routers on the RPL subnet check if they are neighbors on the subnet with the picked router
 - If yes, they are on the active partition

Detection & Election via Host Route

- Of all the neighbors on the RPL subnet, the one with the lowest interface address advertises a host route for ITSELF
 - The one with second lowest address, or even everyone may also do so
 - Ensures smooth transition when the current lowest host route is withdrawn
- Routers on the RPL subnet pick the lowest from all host routes within the subnet
 - The router who advertised that lowest host route represents the active partition

Entire RPL unreachable?



Rcv3 can't even receive from src2 because nobody has route to RPA

→ join

→ data

Anycast RPL

- Put several Anycast RPA and links in different parts of the network
 - As if a regular RPL partitioned
 - When the entire network partitions, each partition can still operate independently as long as that partition has a reachable RPL
 - Before the partition, only one active RPL will be chosen, using the same procedures previously described
- As long as there are no hosts (vs. routers) on those links there should be no routing problem
 - The prefix is not really used for forwarding traffic because there are always host routes via which the routers and RPA are reached

IGP vs. Host Route method

- Additional routes with Host Route method
 - Works with any protocol, including RIP/BGP
- No additional signaling with IGP method
 - Works only with Link State Routing
 - Additional signaling required for inter-area
 - Applicable only in case of intentional partitioning across areas - Anycast RPL
 - Better control on election result
 - e.g. preferring the partition in backbone area

Summary

- No signaling changes
- Local PIM/routing interaction on RPL routers
 - PIM: treat a link as RPL only if all the following are met
 - The RPF route to RPA is a direct route (existing criteria)
 - The DR that this router is adjacent to has the highest address among all reachable Net LSAs for the link (additional criteria)
 - Routing: advertise host route for RPA per PIM' s instruction
 - if and only if the link is treated as RPL
- Still need standardization for consistent behavior on all RPL routers
- In completely partitioned network, both partitions of the RPL are active and traffic flows completely within each partition are not affected

Plan

- Seek review & feedback
- Polish the solution
 - IGP extension to support Anycast RPL across areas
 - Better control on partition election?
 - E.g. Avoid electing a disadvantaged partition
- Request WG review & adoption
 - After the polishing