Existing ECMP RPF Overview

- There are two ways to choose an RPF path when ECMP is present
  - Select the path whose gateway is the PIM neighbour with the largest IP address
  - Use a hash algorithm
- ECMP RPF selection is downstream driven
  - Limited by routing/hash algorithm, no other factors considered

Existing ECMP RPF Issues

- Load-balancing is based on IP addresses instead of “loads”
- Same flow might be sent onto two links
  - Waste of bandwidth
  - Especially if an implementation chooses to stick to its RPF selection after link/node failure
- “Assert” only chooses an RPF neighbour within a LAN, but not between ECMP paths

PIM ECMP Assert

- PIM ECMP Assert is proposed to improve control of RPF path selection.
  - Initiated by upstream routers (similar to Assert)
  - Used to choose a path
    - based on administrative choice
    - from ECMP path
  - Allow downstream routers to use information such as available bandwidth to choose an RPF neighbour

PIM ECMP Assert

- Design Consideration
  - Minimize control traffic in steady state
  - Minimize unnecessary traffic disruption
  - Allow for future enhancement to include more criteria for choosing a path
- We are OPEN to a different name

PIM ECMP Assert

- Key features
  - Triggered by PIM Joins
  - Sent in a different subnet (used to choose a path, instead of an RPF neighbour)
  - New PIM Hello Options

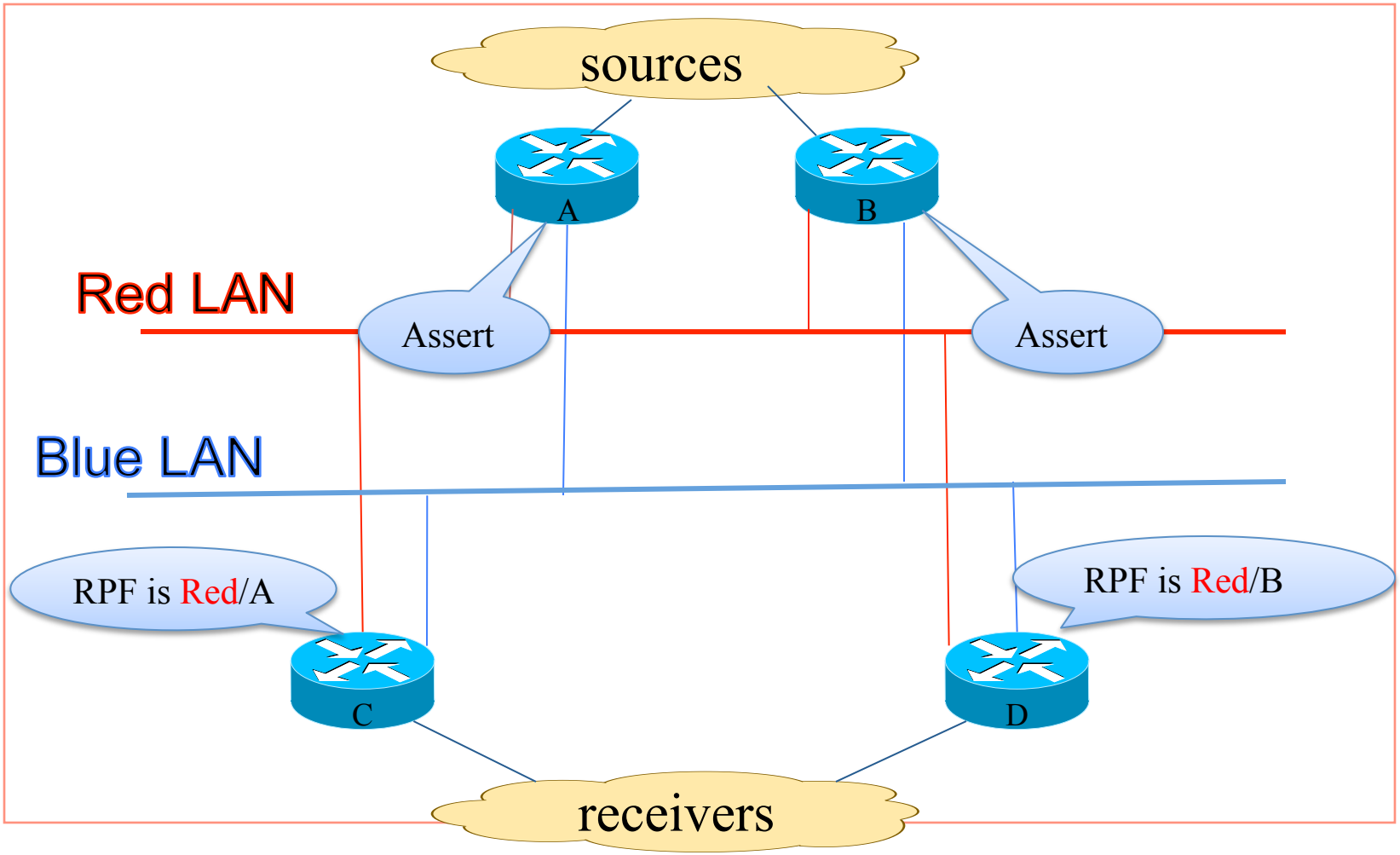
Comparing to PIM Assert

- Trigger
  - Assert is data driven
  - ECMP Assert is triggered by Join
- Application
  - Using Assert to choose an RPF neighbor within a subnet
  - Using ECMP Assert to choose a path from ECMP

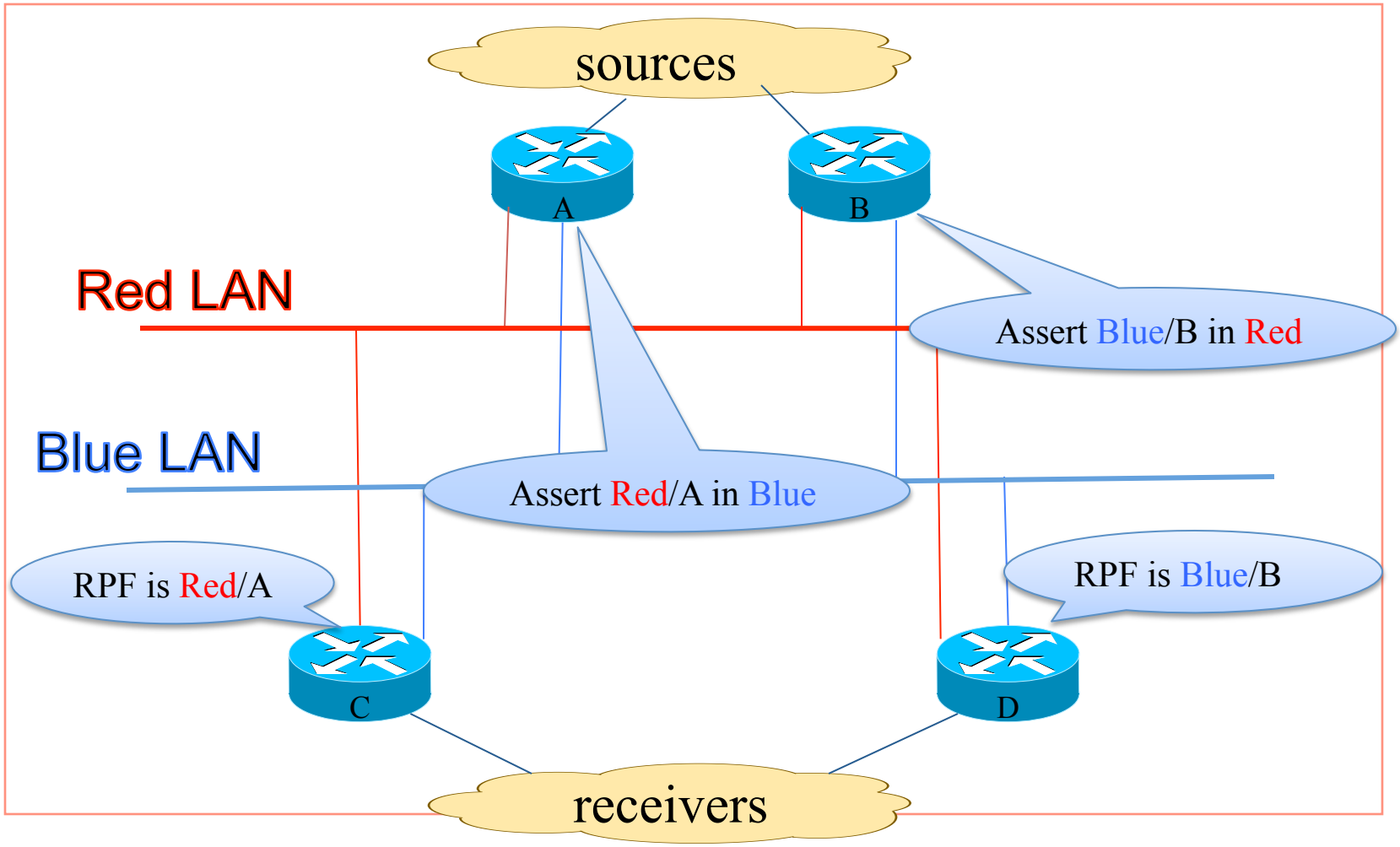
Comparing to PIM Assert

- Impact
  - Assert modifies "routing" decision by comparing routing metrics sent by upstream routers
  - ECMP Assert preserves routing decision (ECMP)
  - ECMP Assert compares non-routing metric (such as uptime/timestamp, bandwidth etc…)

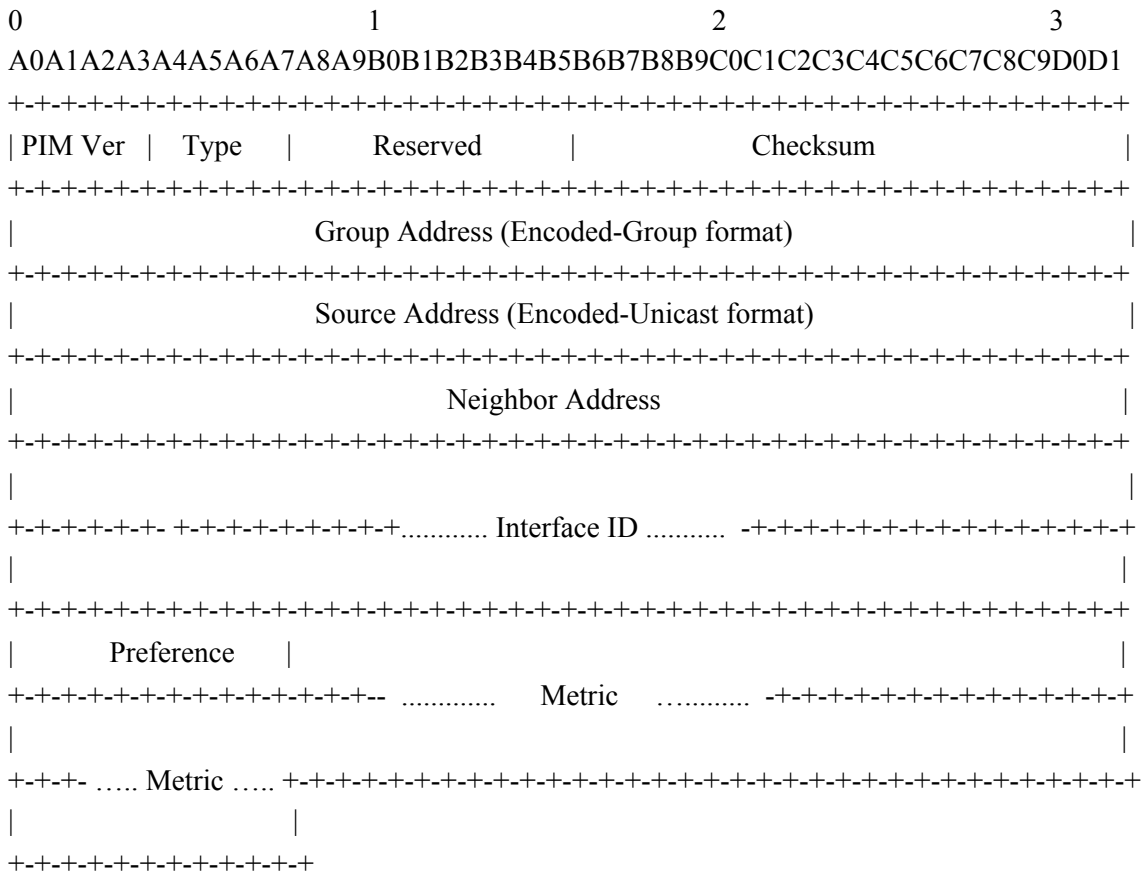
Example (PIM Assert)



Example (PIM ECMP Assert)



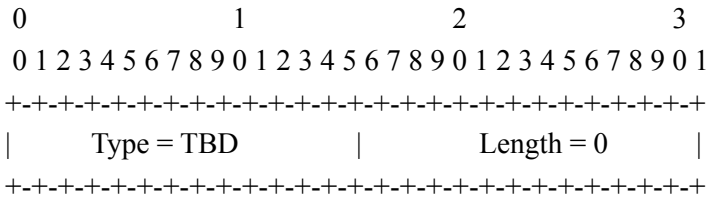
Packet Format: ECMP Assert



Packet Format: Hello Option

- PIM Hello Options

ECMP Assert Hello Option



Update From -00

- Added new authors
- Clarified operation on transient cases
- Clarified use of PIM Interface-ID

For The Working Group

- The draft addresses a weakness in PIM RPF selection
- There is practical application that requires a solution like this
- We welcome comments/suggestion from the working group
- We'd like to request the working group to adopt this I-D