

# Transitive BGP Graceful Restart (draft-zhang-idr-transitive-gr-00)

Haifeng Zhang, Alvaro Retana

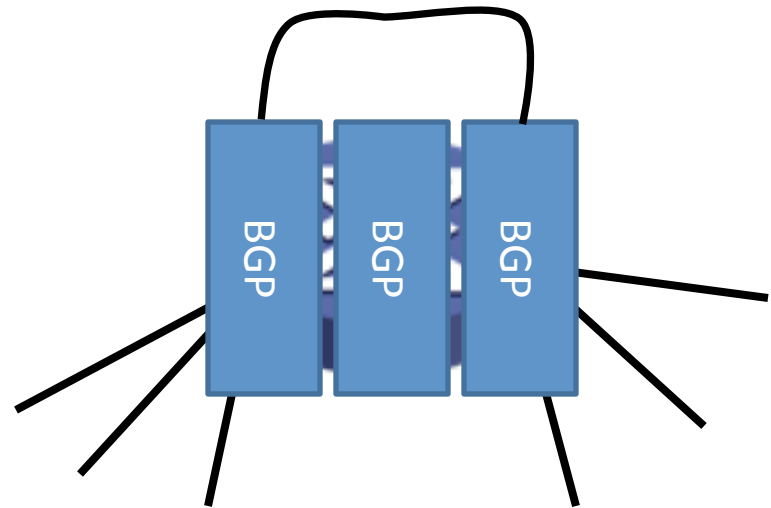
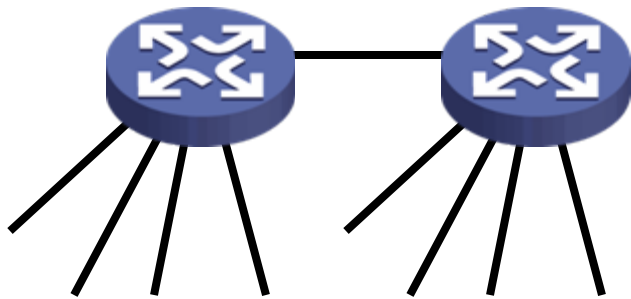
zhanghf@h3c.com,  
alvaro.retana@hp.com

# Background

- Forwarding in a Network is Transitive
  - routing protocols independently calculate paths based on consistent information distributed throughout the network
  - adjacencies confirm the transitive relationship between neighbors
  - graceful restart verifies transitivity by signaling that the forwarding state has been preserved through a restart...even if routing is not active

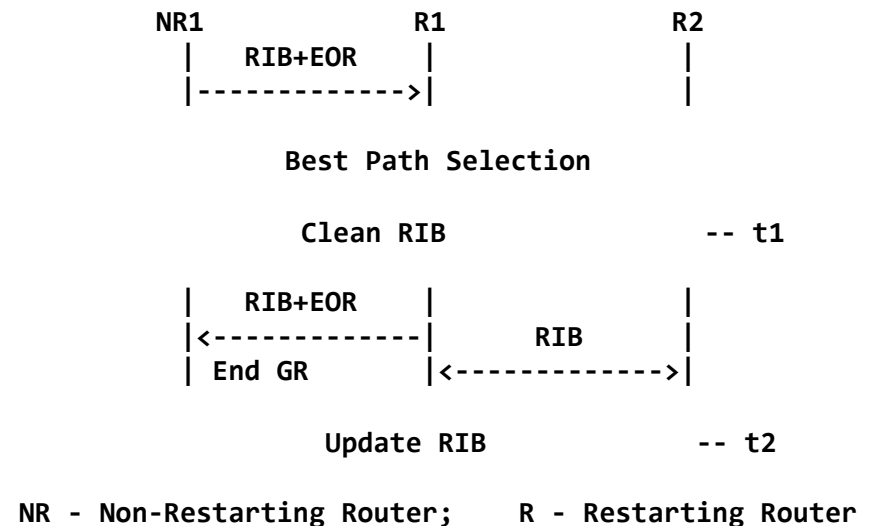
# What problem are we trying to solve?

- Two restarting routers next to each other.



# Transitive BGP Graceful Restart

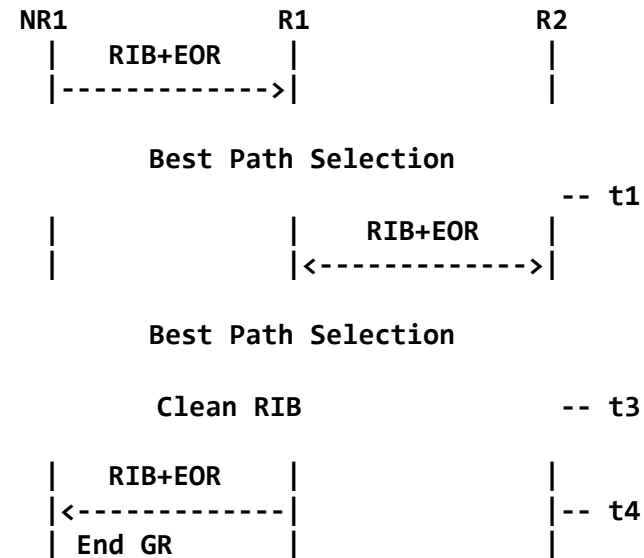
- Objective:
  - Reduce the impact of multiple restarting routers on the data plane.
- Problem:
  - BGP Graceful Restart (rfc4724) defines a mechanism that clears the forwarding table for any restarting peer, even if the forwarding state has been maintained by both routers. The effect is route instability, traffic loss, etc.



*"...the forwarding state of the speaker MUST be updated and any previously marked stale information MUST be removed."* [rfc4724]

# Transitive BGP Graceful Restart

- Objective:
  - Reduce the impact of multiple restarting routers on the data plane.
- Solution:
  - Extend the operation of BGP Graceful Restart to allow a router to fully synchronize with other restarting peers before cleaning the forwarding table.



NR - Non-Restarting Router; R - Restarting Router

*"...the BGP speaker MAY advertise the Adj-RIB-Out to the remaining peers ...and MAY wait for the corresponding End-of-RIB marker from the restarting ones." [Proposed Addition.]*

# Not used Routing Information

- During the recovery period of multiple restarting routers, a BGP speaker may advertise routing information that is not being used at the time.
  - the forwarding state of the speakers remains unchanged (from that at the restart)
- There are 3 types of routes being advertised by the restarting router (all through its non-restarting peers):
  1. Routes that correspond to routes in the RIB. Represent the majority!
  2. Routes not previously in the RIB. (Installation will be delayed until t3.)
  3. Routes that in the RIB point to the other restarting router.

*“...minimize the effect of routing flaps, it is noted that when a BGP Graceful Restart-capable router restarts...there is a potential for transient routing loops or blackholes in the network if routing information changes before the involved routers complete routing updates and convergence.” [rfc4724]*

# Next Steps

- Comments/Questions/Feedback
- Adopt as a WG Item
  - Updates rfc 4724
  - Standards Track