

Constrained Route Distribution with Multiple Address Families

Saikat Ray

Arjun Sreekantiah

Keyur Patel

Prologue

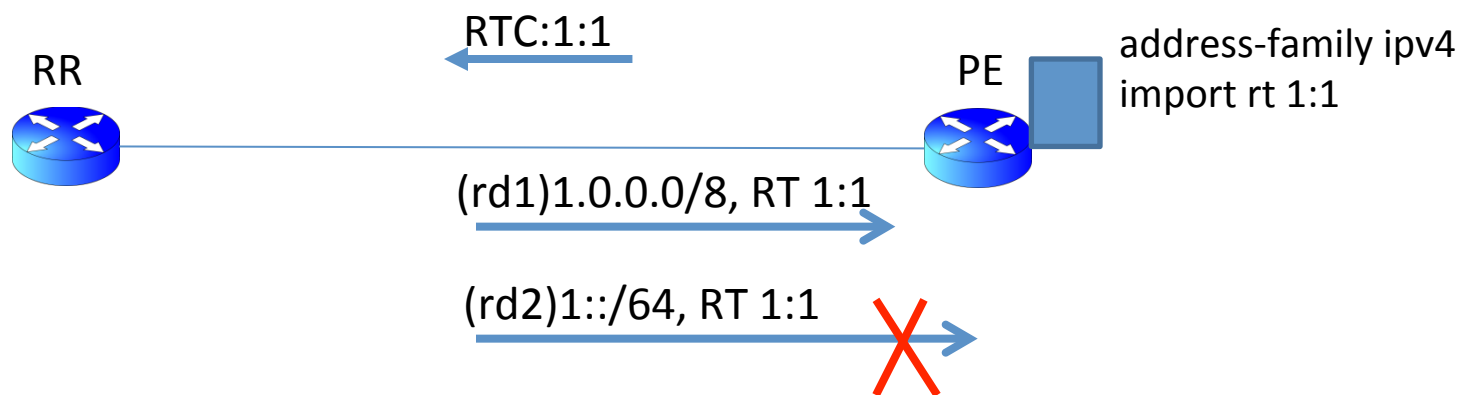
- draft-ray-idr-route-constrain-scope-00 presents two problems
- Problem 2 and suggested solution are presented here
 - We will make it a into a different draft
- Problem 1
 - Seeking input from WG about if the problem is useful to work on

Route target Constrain

- RR sends all VPN routes to a PE for a negotiated address-family (say, 1/128)
- PE only keeps the routes whose RT is imported by at least one VRF
 - RR didn't have to send the other routes
- Optimization
 - Let PE tell RR which RTs the PE is “interested” in
 - RR sends only the matching routes to PE
- Address-family 1/132 is used for exchanging RTs of “interest”
 - NLRI encodes the RT of interest

RTC and Multiple VPN Address-family

- An RTC NLRI signals sender's interest in routes with the given RT from all (VPN) address-families
 - If the PE does not need the route, it drops them
 - E.g., no VRF with IPv6 address-family importing 1:1 (but there are VRFs with IPv6 address-family so that PE and RR negotiates VPNv4 Unicast address-family)



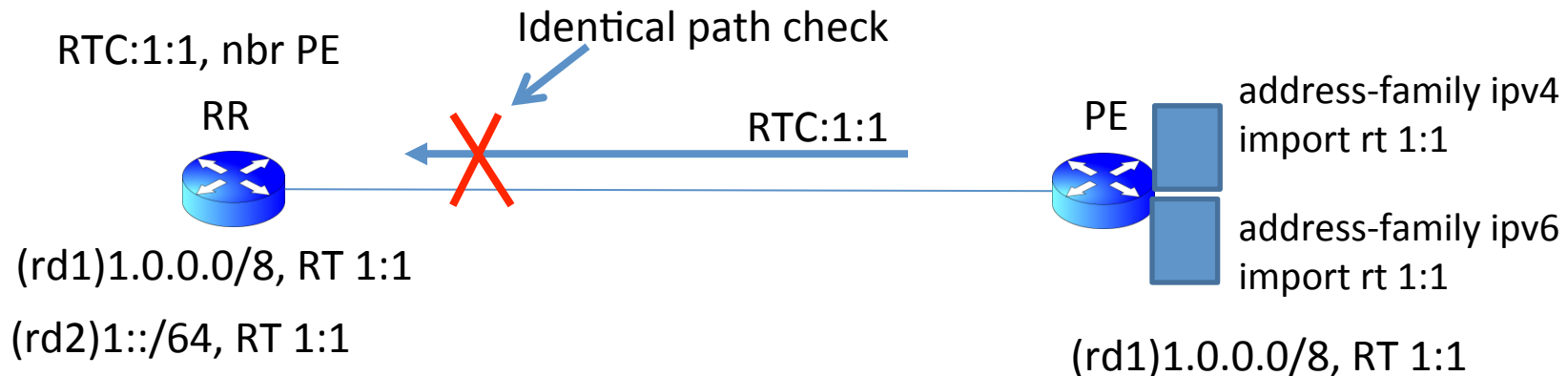
Problem: Incremental VRF addition



- **Steady-state**

- PE has a VRF with IPv4 that imports RT:1:1, but no VRF with IPv6 that imports RT:1:1
- RR has v4 and v6 routes with RT:1:1, and RTC:1:1 from PE
- PE has v4 route with RT:1:1 from RR, but no v6 route with RT:1:1

Problem: Incremental VRF addition

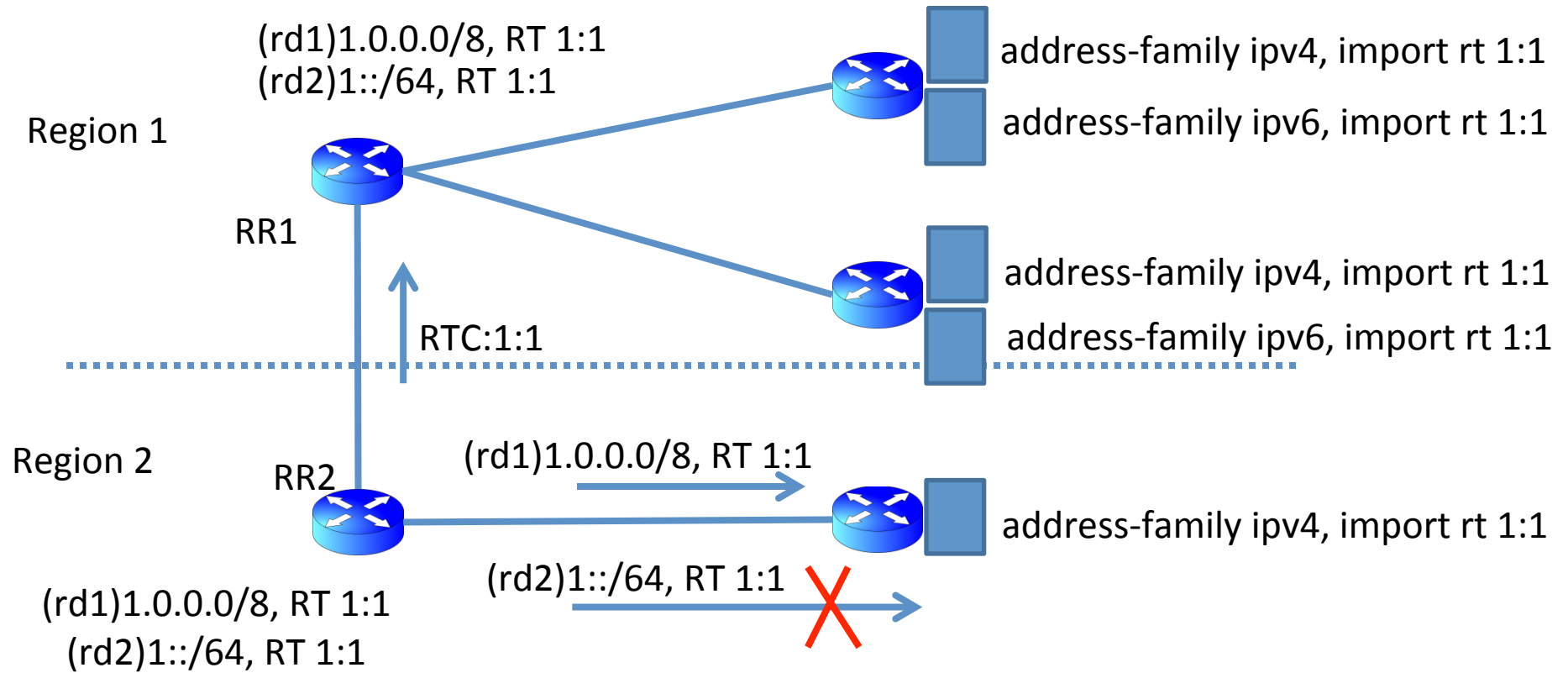


- VRF addition
 - A VRF with IPv6 that imports RT:1:1 is added to PE
 - PE sends RTC:1:1 to RR
 - RR drops it due to identical path check
- Work around
 - PE needs to send a route-refresh for 2/128 to RR
 - RR will send all VPNv6 routes to PE whose RTs match some RTC NLRI

Proposed Rules

- Rules
 - A BGP speaker that receives an identical RTC path from a neighbor must treat it as equivalent to a route-refresh request for the given RT for all (VPN) address-families.
 - If a new (VPN) address-family is negotiated between two BGP speakers without a session reset (e.g., using dynamic capability, or using multi-session feature), then existing RTC NLRIs' scope must be extended to the new address-family
- No changes to the protocol on the wire
- Why do we want to standardize this?
 - Prudent BGP implementations use identical path check; this would mandate a change in the behavior for RTC
 - The PE needs to know whether the RR supports this or not
 - A CLI knob on the RR
 - Indicate support in the Capability
 - Maybe we can use a reserved bit to indicate this for 1/132 MP capability?

Unnecessary route retention



- Region 1 RR has clients that require both IPv4 and IPv6 routes with RT:1:1
- Region 2 RR does not have any client that need IPv6 route with RT:1:1
- Region 2 RR still retains IPv6 routes with RT:1:1 and advertises them to its clients (who sent RTC:1:1)

Unnecessary route retention

- This is a problem if
 - Different address-families in different VPNs use the same RT
 - Not the usual operational practice
 - Not all sites have the same set of address-families
 - Regional differences - E.g., IPv6 in only one region
 - Transitions – e.g., IPv6 turned on “temporarily” on some sites
- Previous proposal to solve this add safi in NLRI
 - Not backward compatible
- Current proposal is to use extcomm to encode afi/safi scope in RTC path
 - Backward compatible, incrementally deployable, can be rolled back after transition period
- Question for the WG – is this an interesting enough problem for the WG?