

BGP Enhanced Route Refresh Capability

draft-keyur-bgp-enhanced-route-refresh-01

Keyur Patel, Enke Chen, Balaji Venkatachalapathy

IETF 79, November 2010, Beijing, China

Motivation

- Current Route Refresh mechanism defined in RFC2918 allows dynamic exchange of BGP Adj-Rib-Out table
- Current Route Refresh mechanism does NOT provide any kind of table demarcation
 - Route Refresh Start of RIB
 - Route Refresh End of RIB
- Table demarcation is useful in
 - performing consistency checks for missing withdraws
 - Measurements related to complete BGP Adj-Rib-Out, etc

Enhanced Route Refresh Capability

- New capability used to exchange newly defined Route Refresh message subtypes
- Route Refresh subtypes are used to signal
 - Route Refresh Start of Rib
 - Route Refresh End of Rib
- Every Route Refresh response begins with Route Refresh Start of Rib message, complete announcement of Adj-Rib-Out table and ends with Route Refresh End of Rib message
 - Allows sending BGP speaker to deterministically signal the begin and the end of Adj-Rib-Out table
 - Works well when announcing Adj-Rib-Out in presence of ORFs

Enhanced Route Refresh Capability (Cont'd)

- BGP speaker receiving the complete Adj-Rib-Out table can use the Route Refresh Start of Rib as an indication of a new table transfer that is about to get started
 - Uses this indication to mark all the routes in its Adj-Rib-In as stale
- BGP speaker receiving the complete Adj-Rib-Out table can use Route Refresh End of Rib as an indication of completion of the table transfer
 - Uses this indication to safely purge all the stale routes that were not updated as part of the table transfer
- Both the Route Refresh Start of Rib and Route Refresh End of Rib can be use towards measuring the table announcement times, etc

Extreme Network Churn and Route Refresh End of RIB

- Typically Route Refresh End of Rib message is generated once the table announcement is complete
- Generation of Route Refresh End of Rib message can be delayed for prolong period of time in case of an extreme network churn where a BGP speaker is not able to announce entire table
 - Results in delaying of purging of stale routes (possibly hours!)
- Possible Solutions
 - Delay the generation of Route Refresh End of Rib message
 - Implementation specific timer to generate Route Refresh End of Rib message
 - Throttling of inbound update message processing till the peer gets to generate the End of Rib message

Questions?