LDP IGP Synchronization for broadcast networks

draft-lu-ldp-igp-sync-bcast-00

by

Wenhu Lu & Sriganesh Kini

74th IETF - San Francisco, CA, USA

(March 22-27, 2009)

LDP IGP Sync - RFC5443

- Motivation to have LDP IGP Sync
  - Applications (e.g. VPN) require end-to-end LSP to be operational
  - If LDP is used, then the LDP-LSP must be operational even during IGP path changes
  - When IGP becomes operational but until LDP is not, traffic will be blackholed

- Idea behind RFC5443
  - Avoid using link for transit traffic until LDP becomes operational
  - Advertise a 'high cost' for the link in the link-state IGP's LSA
 - After LDP becomes operational, advertise the link's real cost

RFC 5443 procedures applied to a broadcast network

- Link state IGP representation of broadcast network
  - Pseudo-node LSA (originated by DR) advertises graph edge to each router on broadcast network that the DR is adjacent to
  - The LSA of each router advertises a graph edge to the pseudo-node
- RFC 5443 procedures applied to link-state IGP broadcast network representation
  - When a router's IGP becomes operational on a broadcast network, **all** routers on that network start advertising a high cost (to the pseudo-node) in their LSA
  - When LDP becomes operational on the broadcast network IGP cost is reverted back to the real cost

Problem applying RFC 5443 procedures to broadcast networks

- Quoting RFC 5443 sec 3 `"policy decision has to be made whether the unavailability of LDP service to one peer should result in the traffic being diverted away from all the peers on the link"`
  - If policy is not to divert traffic away from all the peers on the link
    - The 'high cost' advertisement mechanism of RFC 5443 cannot be applied
    - When the IGP nexthop points to the link that came up, traffic will be blackholed
  - If policy is to divert traffic away from all peers
    - Routes (and traffic) will be churned throughout the network
    - Traffic may be diverted to the link that just came up ! This will result in traffic being blackholed

Redback
NETWORKS
AN ERICSSON COMPANY

Topology to illustrate problem
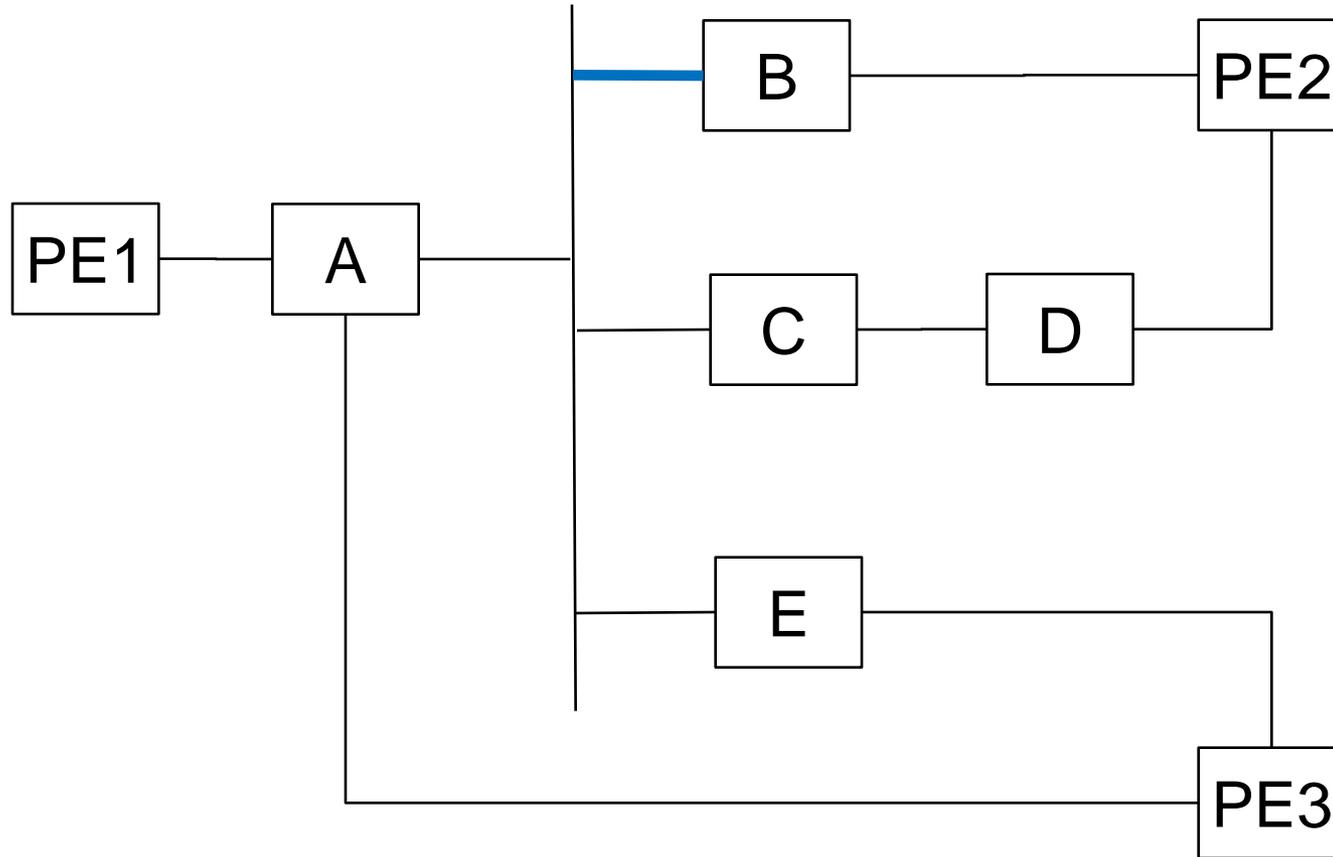
A-PE3 link cost is 10
All other links have cost of 1

Redback NETWORKS
AN ERICSSON COMPANY

Explanation of the problem for the topology

- Before B's interface to the broadcast network comes up
  - PE1 to PE2 traffic takes bi-directional path PE1-A-C-D-PE2
  - PE1 to PE3 traffic takes bi-directional path PE1-A-E-PE3

- On applying the mechanism of RFC 5443 when B's interface to the broadcast network comes up
  - A, B, C, E advertise high cost to the pseudo-node LSA
  - A's nexthop (and traffic) to PE3 changes from E to PE3. This is unnecessary route and traffic churn
  - A's nexthop (and traffic) to PE2 changes from C to B ! This results in PE1 to PE2 traffic to be now blackholed

Proposed solution

- Remove link from LSDB unless absolutely necessary

- When is the link absolutely necessary ?
  - When the link is a 'cut-edge' of the topology it is needed to bring up the LDP session

- How to detect that link is a 'cut-edge' ?
  - Check if directly connected broadcast network is reachable via an alternate path
  - Straightforward to do this during SPF computation

- If link is a 'cut-edge' advertise link with the real cost

- If link is not a 'cut-edge' the LSA of the router does not advertise a link to the pseudo-node until LDP is operational to all neighbors

Characteristics of proposed solution

- Only the router that is establishing its first adjacency on the broadcast network is involved in these procedures

- Avoids churn of routes and traffic

- No protocol changes

- No backward compatibility issues. Solution can co-exist with RFC5443

- Solution can be extended easily for point-to-point links but is not necessary

Explanation of the solution for the topology

- In the topology illustrated earlier, router B detects that the link is not a cut-edge and does not advertise a link to the pseudo-node until LDP is operational with all neighbors

- PE1 to PE3 traffic continues to go along PE1-A-E-PE3. There is no churn of routes and traffic unlike RFC 5443

- PE1 to PE2 traffic continues to go along PE1-A-C-D-PE3 until B's LDP becomes operational to all neighbors. After that traffic shifts to PE1-A-B-PE3. Traffic is not backholed