

IS-IS Extension For Building Distribution Trees

draft-yong-isis-ext-4-distribution-tree-01

Lucy Yong, Weiguo Hao, Donald Eastlake
Andrew Qu, Jon Hudson

November 2013 Vancouver BC

Problem Statement

- IP network doesn't yet support multicast transport
- IP network relies on PIM protocols and solutions to carry multi-destination traffic
- Pain points in PIM protocol solutions:
 - Use own “hello” protocol and hop-by-hop message
 - Additional convergence time besides IGP's
 - A lot of soft state and heavy CPU load
 - Packet may be forwarded to RP unnecessarily
 - Scalability Challenges to support overlay applications

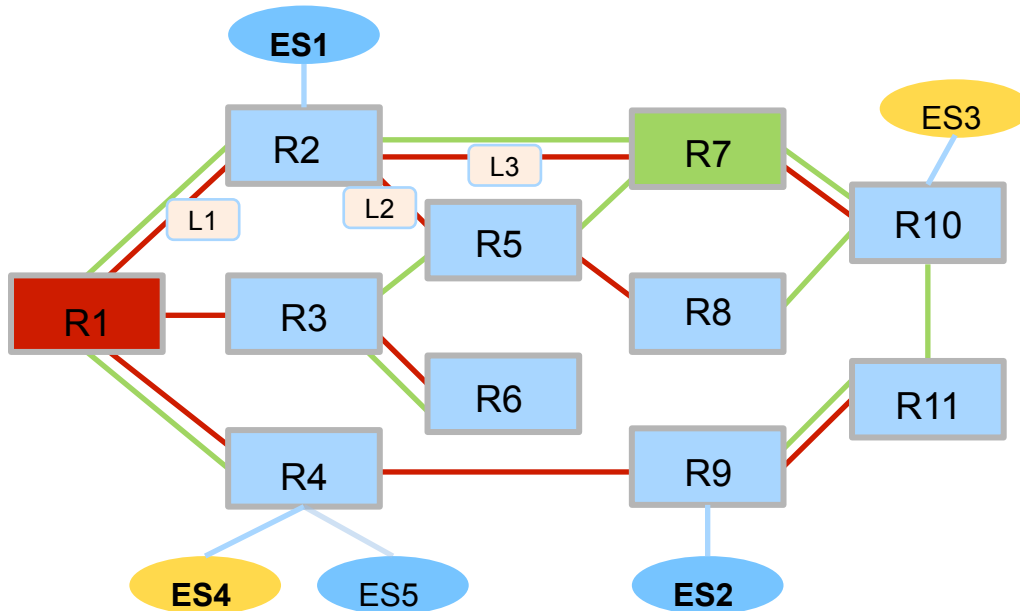
Problem Statement Cont.

- PIM solutions hardly meet the requirements for network virtualization overlays, i.e. [NVO3]
 - IP as underlying network may carry many overlay VNs that contain some BUM traffic
 - A VN maps to an IP multicast group, PIM can't scale to
 - A VN topology may be sparse and dynamic compared to underlying network topology
 - Massive PIM states burden the device
 - Slow convergence time for multicast traffic
 - Major impacts to VNs and cloud applications
- Overlay VN BUM traffic may be carried as underlying unicast traffic
 - May waste massive network capacity due to replications
 - Increase the cost on the applications

IS-IS Extension for Multicast Transport

- IS-IS has been used for unicast traffic routing
 - Can be for multicast traffic routing as well, like TRILL
- Simple IS-IS extension can achieve:
 - Build distribution rooted-trees for multicast transport
- The benefits to do this:
 - One protocol for both unicast and multicast transport
 - Use of LSDB and SPF algorithm to build a tree
 - Same convergence time for both unicast and multicast
 - Forwarding Optimization
 - Well align with network evolution direction, create programmable, self-healing, multi-service fabric
 - In the fabric, a closed IP network carry all types of traffic under one control plane protocol

ISIS based Rooted Distribution Trees



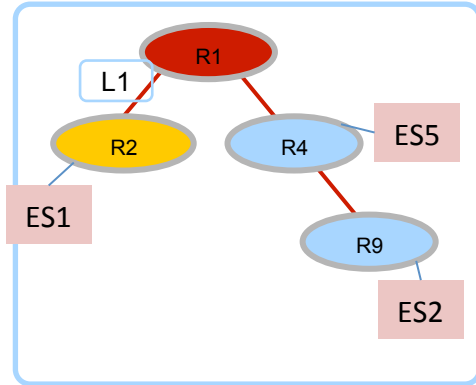
R1&R7 are the root of red and green trees

ES1、 ES2、 ES5 is the end station of Group1,
ES3、 ES4 is the end station of Group2

Group1 uses red tree and Group2 uses green

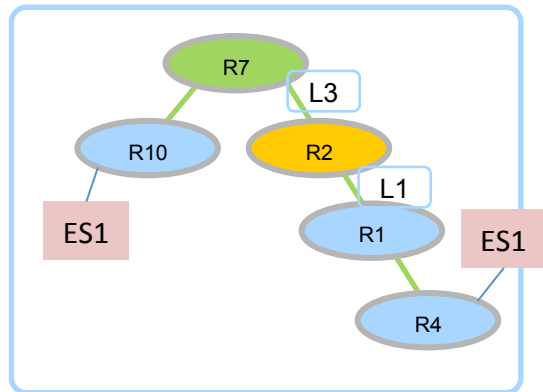
- Provision tree root and associated (*,G) at a node
- ISIS distributes tree root/ (* G) in IGP
- Nodes compute the trees with SPF/LSDB
- Supports ECMP and allow multiple rooted trees associated w/ (*, G)

Distribution Tree Pruning for (*, G)



Pruning red tree for Group1

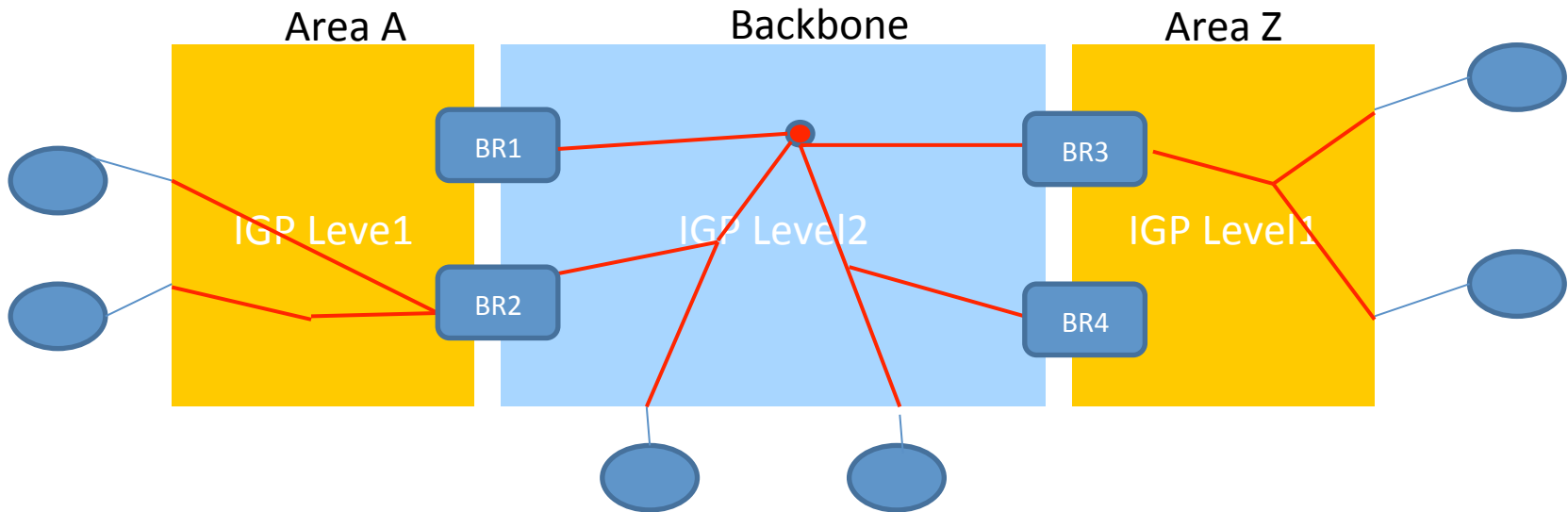
The multicast table in R2	
Multicast group	Out interface
Group1	L1 L3
Group2	L1



Pruning green tree for Group2

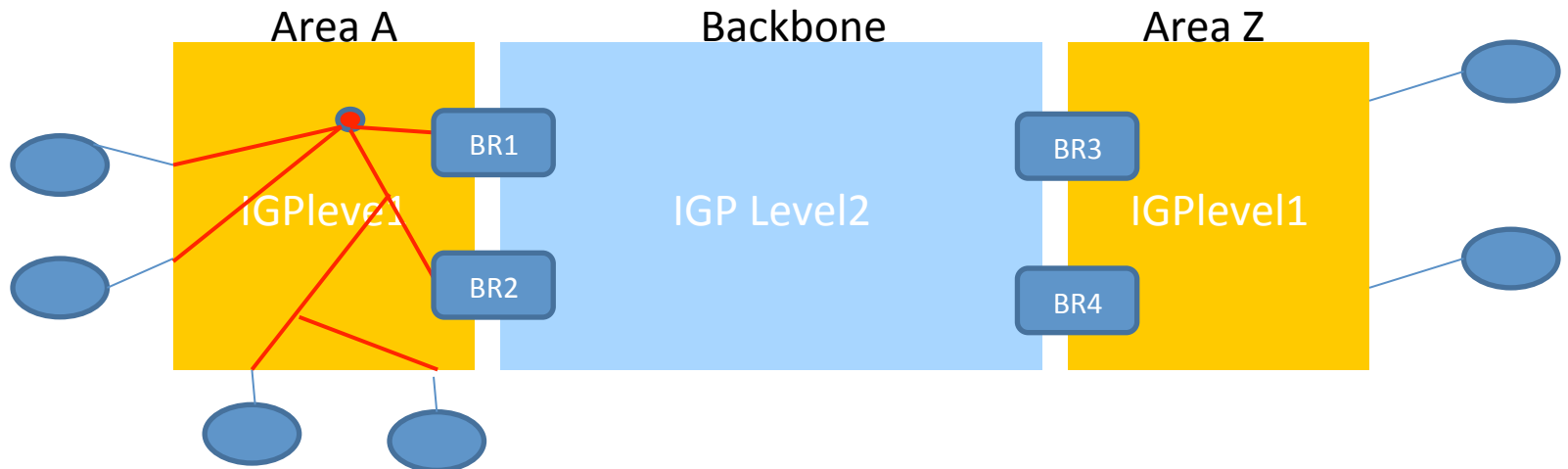
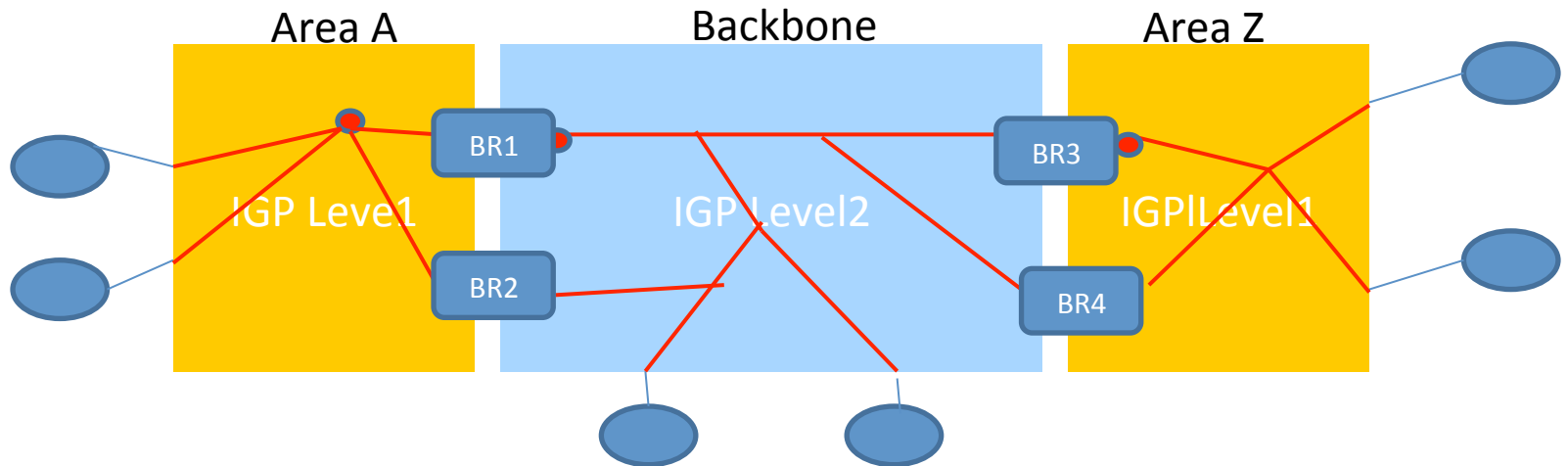
- Edge Routers use IGMP to learn host interested (*,G) or (S, *G)
- Edge routers announce their interested (*,G)
- Each router prunes the tree based on other edge router's interests on the (*, G)/(S,G)
- Multi-destination traffic for a group is forwarded over the pruned tree for the group

IGP Multi-Level Support



- The solution needs to work under multi-levels configuration
 - Multicast source(s) and listeners may be in any area, respectively
- A tree root node associated multicast groups may be in any area
 - The distributed tree spans across all other areas via BR
- Loop prevention is must
 - Single designated BR is selected for (*, G)
- Level optimization is possible
 - If an area doesn't have any interested members, the pruned tree may not have any branch in the area.

IGP Multi-Level Support



Build a distribution tree in area A only

Transport Overlay VN BUM Traffic in IGP

- Designated edge/boarder Routers (DR) maintains the mapping between the VN (*, G) to a underlying (*,G)
- DRs perform packet encaps/decap upon receiving a packet from a host or the underlying network
- Mapping of the VN (*, G) and a underlying (*,G) at DR
 - manually configured at DRs
 - automatically generated
 - dynamically informed
- The same edge/boarder router MUST be selected as the Designated Router for the VN (*, G) and the underlying (*,G) that are mapped

Next Step

- Seek some comments and feedbacks
- Should we add this work item to the WG?

Huawei has the solution demo in the Bits-n-Bytes event on Thur. 7pm. Welcome to visit and comment