North-Bound Distribution of Link-State and TE Information using BGP

Hannes Gredler (hannes@juniper.net)
Jan Medved (jmedved@cisco.com)
Stefano Previdi (sprevidi@cisco.com)
Adrian Farrel (adrian@olddog.co.uk)
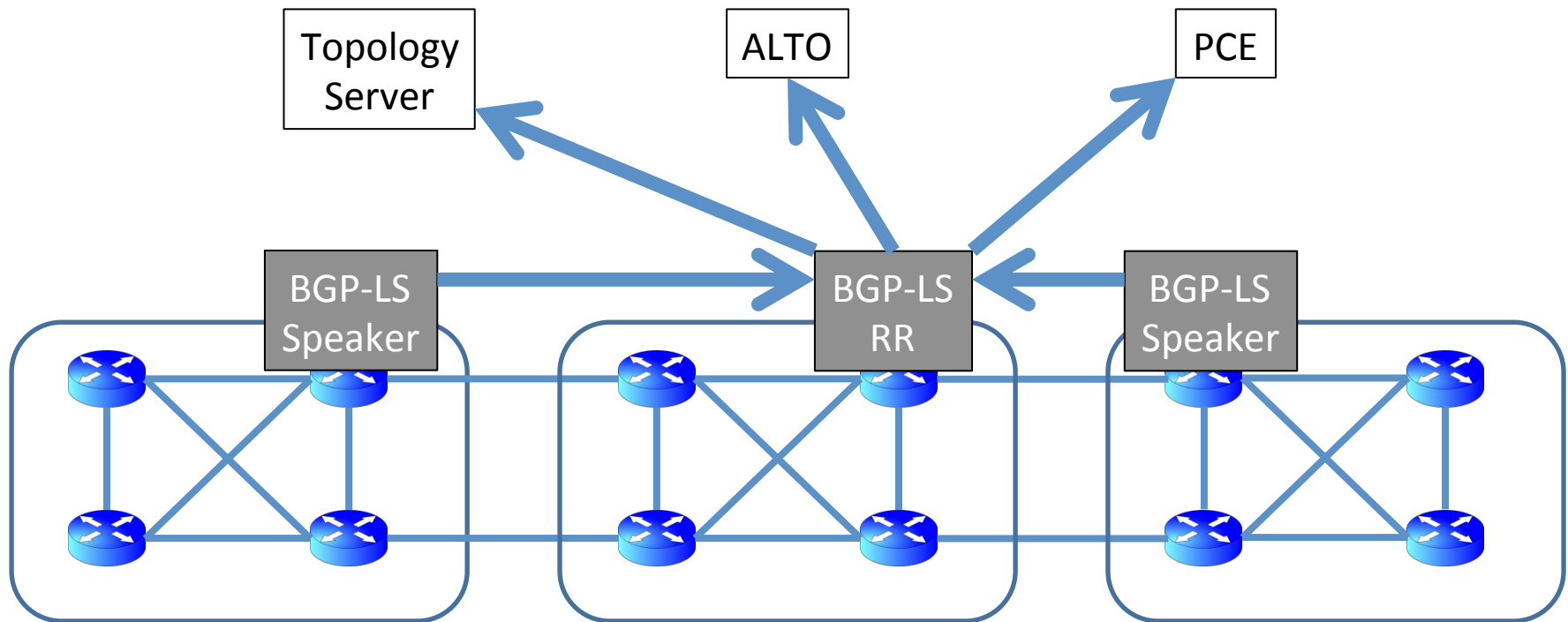Saikat Ray (sairay@cisco.com)

IETF 87   August 2013, Berlin

BGP-LS Overview

- BGP-LS is an address-family (afi=16388, safi=71) defined to carry IGP link-state database via BGP
  - Supports both IS-IS and OSPF(v2/v3)
  - Delivers topology information to outside agents
    - Topology servers, orchestration elements, ALTO servers
  - Allows a topology server to construct the full topology (even across ASes)
  - BGP allows policy-based control to aggregation, information-hiding, abstraction, etc.
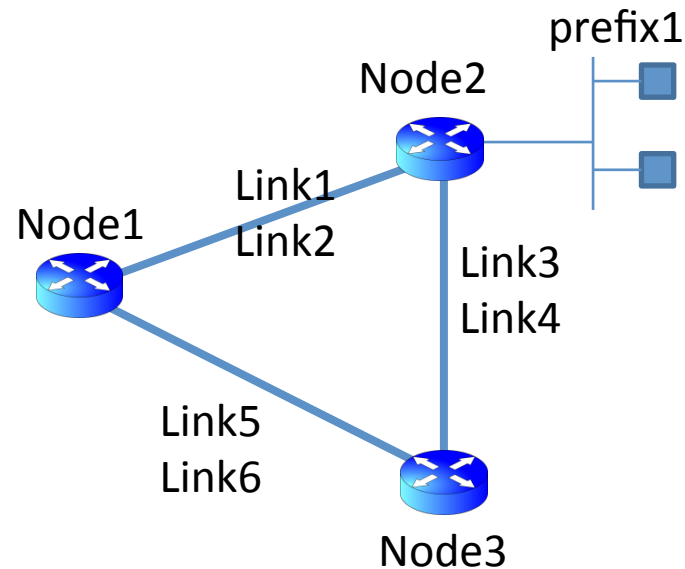 - Out of scope: Leak LS information back to routing

BGP-LS Overview

- Deployment model
  - IGP redistribution into BGP-LS
  - Advertisement of BGP-LS NLRIs to RR.
  - RR sends information to external agents

Topology Server

ALTO

PCE

BGP-LS Speaker

BGP-LS RR

BGP-LS Speaker

BGP-LS Overview
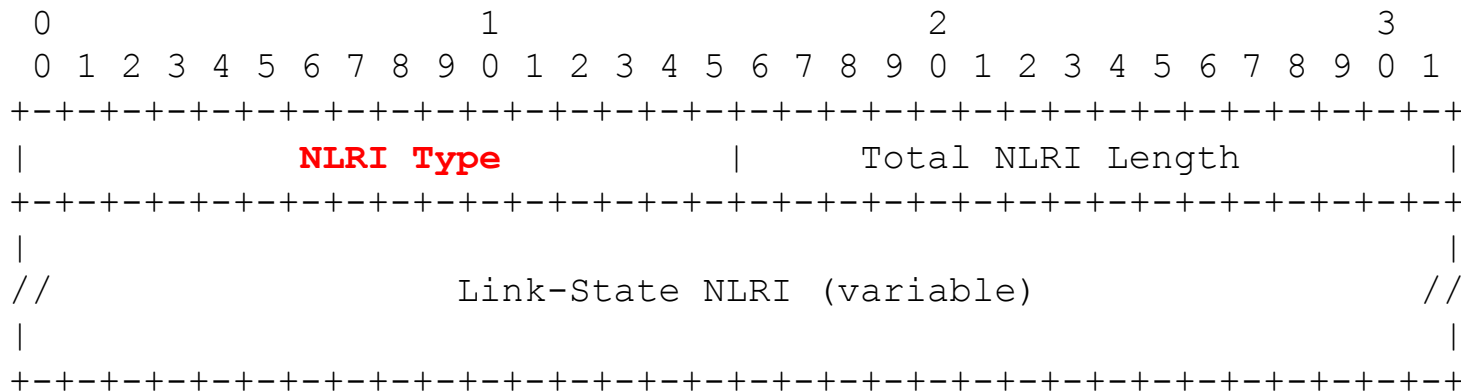
- A common topology abstraction model: An IGP network is modeled as three classes of objects
  - Nodes, Links (pair of nodes), prefixes

prefix1

Node2

Link1
Link2

Node1

Link3
Link4

Link5
Link6

Node3

BGP-LS Objects:
- 3 nodes
- 6 links
- 1 prefix

BGP-LS Overview

- BGP-LS NLRI
  - NLRI Type defines the object class (node/link/prefix)
  - NLRI body is a set of TLV
  - NLRI contains the data that identifies an object
    - NLRI is the key for the object
      - Minimal data needed to remove ambiguity

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|            NLRI Type           |     Total NLRI Length         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
//               Link-State NLRI (variable)                    //
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

BGP-LS Overview

- BGP-LS attribute
  - Optional non-transitive
  - Encode properties of the object
  - Data consists of TLVs
    - TLVs are specific to the object class
      - Node attribute TLVs (MT-ID, Flag bits, Node-name, etc.)
      - Link attribute TLVs (local/remote ipv4/ipv6 router-id, admin-group, link BW, SRLG, etc.)
      - Prefix attribute TLVs (IGP flags, (Exteneded) route tags, etc.)

Changes from -02

- One 64 bit *identifier*
  - Identifies the IGP instance
    - Needs to be "globally" unique
    - No semantics imposed.

- NLRI types are
  - Node = 1, Link = 2, IPv4 Prefix = 3, IPv6 Prefix = 4

- Node descriptor TLV uses a uniform "IGP Router-ID"

- OSPF route-type in prefix descriptor. Only one prefix in a prefix NLRI.

Changes from -02

- Node name and link name TLVs.

- OSPF area goes in the NLRI (key). New TLV for IS-IS area (goes in node attribute)

- Specifying auxiliary IPv4/IPv6 local/remote router-id in link attribute (for IS-IS links) is MUST
  - Helps with TE

- All TLV code-points from one space

- Some section reorganizations and clean-up

Status

- 3(.5) implementations
  - Cisco/Juniper
- Inter-op planned around September/October
- Still need BGP-LS Path Attribute code-point from IANA
- Comments?