Experience and Practical Issues of ALTO Deployment using BGP-LS

draft-zhang-alto-bgp-ls

J. Jensen Zhang
Kai Gao
Luis M. Contreras
Anais Escribano
Patricia Cano
Francisco Cano

Outline

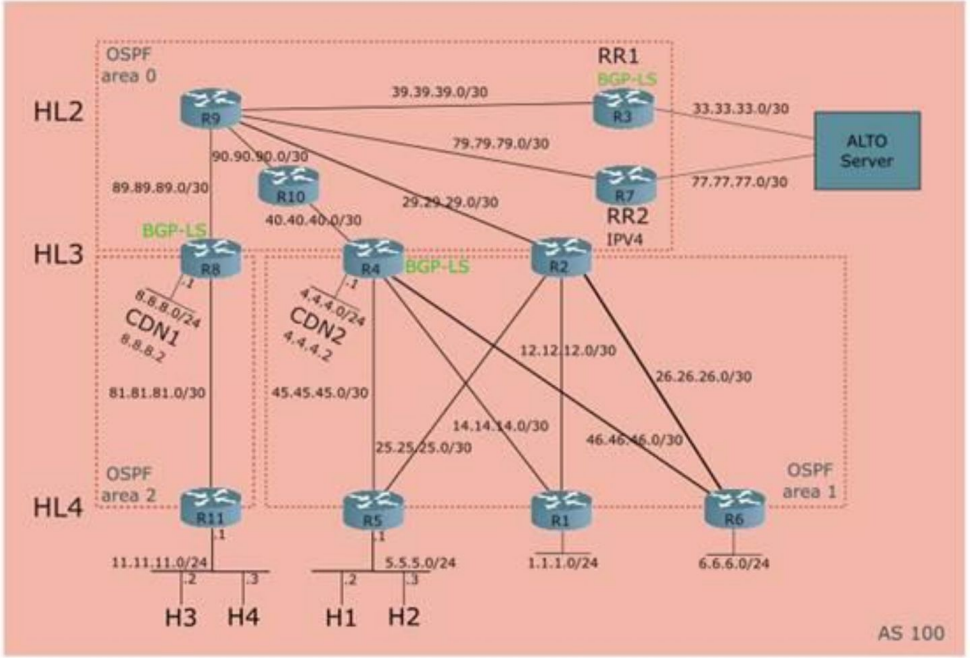
- Progress update
  - Based on OpenDaylight BGPCEP plugin
  - IFIP IM (IFIP International Symposium on Integrated Network Management)
  - To be deployed in Telefonica (reported by Luis in IETF 114/115)
- Experiences of building ALTO maps from BGP and BGP-LS using OpenDaylight
  - Deployment requirement (how BGP and BGP-LS are configured to connect to ALTO)
  - Abstract model for BGP and BGP-LS
  - Map computation
  - Issues and limitations
- Potential inputs to the ALTO WG

Progress update

- Implemented in OpenDaylight based on BGPCEP plugin
  - https://github.com/openalto/sextant
- Early deployed in Telefonica technology lab tests
- Reported technical details in IFIP IM 2021
- Migrated to Exabgp in Telefonica pre-production network tests (reported by Luis in IETF 114/115)
  - https://datatracker.ietf.org/meeting/114/materials/slides-114-mops-exposure-of-telefonica-network-topology-through-alto-for-integration-with-telefonica-cdn-01
  - https://datatracker.ietf.org/meeting/115/materials/slides-115-mops-exposure-of-telefonica-network-topology-through-alto-for-integration-with-telefonica-cdn-00

Deployment Requirements in a Single AS

- All the AS Border Routers (ASBRs) must logically connect to at least one Route Reflector (RR)
- In each IGP area, BGP-LS must be configured on at least one node
  - To minimize number of nodes configuring BGP-LS, typically, BGP-LS is only configured on Area Border Routers (ABRs), for example, R8/R4
- All the nodes configuring BGP-LS must logically connect to at least one RR
- The ALTO server must have a data source listener connecting to RRs

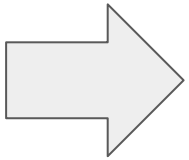


BGP Route Update Abstraction

- BGP Update Message [RFC4271]
  - Path Attribute
    - Originator
    - Next hop
    - AS path
    - Local preference
    - …
  - Network Layer Reachability Information (NLRI)
    - IPv4/IPv6 prefix
    - …

Building Network Map in a Single AS

- BGP Update Message [RFC4271]
  - Path Attribute
    - Originator
    - Next hop
    - AS path
    - Local preference
    - …
  - Network Layer Reachability Information (NLRI)
    - IPv4/IPv6 prefix
    - …



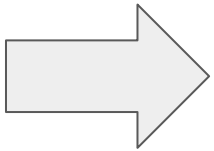
- Network Map
  - PID: ASN + Next hop
  - Prefixes: NLRI

BGP-LS NLRI Abstraction

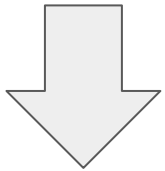
- Link-state NLRI [RFC7752]
  - Node descriptor
    - ASN
    - BGP-LS Id
    - Area Id
    - Router Id
  - Link descriptor
    - Local node descriptor
    - Remote node descriptor
    - Attributes (e.g., IGP metric)
  - Prefix descriptor
    - Advertising node descriptor
    - IPv4/IPv6 prefix

Building Cost Map in a Single AS

- Link-state NLRI [RFC7752]
  - Node descriptor
    - ASN
    - BGP-LS Id
    - Area Id
    - Router Id
  - Link descriptor
    - Local node descriptor
    - Remote node descriptor
    - Attributes (e.g., IGP metric)
  - Prefix descriptor
    - Advertising node descriptor
    - IPv4/IPv6 prefix

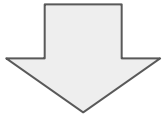


IGP Topology with Link Metric



run link-state routing algorithm (e.g., OSPF)

End-to-end Routes (i.e., Path Vector)



compute PID-to-PID costs upon network map

Cost Map

Building ALTO Maps for Multiple ASes

Known issues:

- In a downstream AS, the ingress point for an end-to-end flow may be unknown
  - Potential solutions:
    - Traffic monitoring
    - Compute routes (path vectors) for all the potential ingress points
    - Wait for additional information from the upstream AS (require multi-domain extension)

Data Encoding

- The data source listener can choose different formats to store information retrieved from RRs for BGP and BGP-LS; It depends on the implementation, e.g.,
  - In OpenDaylight, it is YANG data model
  - In Exabgp, it is JSON


```
{ "exabgp": "5.0.0", "time": 1651676896.0184126, "host" :
"localhost.localdomain", "pid" : 70559, "ppid" : 1, "counter": 733, "type":
"update", "neighbor": { "address": { "local": "80.58.171.201", "peer":
"192.168.255.89" }, "asn": { "local": 64531, "peer": 64531 } , "direction":
"receive", "message": { "update": { "attribute": { "origin": "igp",
"local-preference": 100, "originator-id": "192.168.252.178", "cluster-list":
[ "12.4.1.1" ], "bgp-ls": { "generic-lsid-258": ["0x00000C0D0000008B"],
"igp-metric": 1000 } }, "announce": { "bgp-ls bgp-ls": { "192.168.252.178": [
{ "ls-nlri-type": "bgpls-link", "l3-routing-topology": 2, "protocol-id": 2,
"local-node-descriptors": [ { "autonomous-system": 3352 }, {
"bgp-ls-identifier": "178" }, { "router-id": "d500b8070000" } ],
"remote-node-descriptors": [ { "autonomous-system": 3352 }, {
"bgp-ls-identifier": "178" }, { "routerid": "c0a8ff210000" } ],
"interface-addresses": [ "192.168.204.198" ], "neighbor-addresses": [
"192.168.204.197" ], "multi-topology-ids": [ ], "link-identifiers": [ ] } ] }
} } } } }
```


```
:(linkstate-routes-case)
   +--ro linkstate-routes
      +--ro linkstate-route* [route-key path-id]
         +--ro route-key                        string
         +--ro path-id                          path-id
         +--ro protocol-id                      protocol-id
         +--ro identifier                       identifier
         +--ro (object-type)?
         |  +--:(node-case)
         |  |  +--ro node-descriptors
         |  |     +--ro as-number?         inet:as-number
         |  |     +--ro area-id?           area-identifier
         |  |     +--ro domain-id?         domain-identifier
         |  |     +--ro (c-router-identifier)?
         |  +--:(link-case)
         |  |  +--ro local-node-descriptors
         |  |  +--ro remote-node-descriptors
         |  |  +--ro link-descriptors
         |  +--:(prefix-case)
         |  |  +--ro advertising-node-descriptors
         |  |  |  +--ro as-number?         inet:as-number
         |  |  |  +--ro area-id?           area-identifier
         |  |  |  +--ro domain-id?         domain-identifier
         |  |  |  +--ro (c-router-identifier)?
         |  |  ...
         |  |     +--ro multi-topology-id?            topology-identifier
         |  |     +--ro ospf-route-type?              ospf-route-type
         |  |     +--ro ip-reachability-information?  inet:ip-prefix
         ...
         +--ro attributes
              ...
```

Other Potential BGP/BGP-LS Extensions

- BGP-LS inter-AS topology extension
  - https://datatracker.ietf.org/doc/html/draft-ietf-idr-bgpls-inter-as-topology-ext-13
- BGP-LS TE path extension
  - https://www.ietf.org/archive/id/draft-ietf-idr-bgp-ls-te-path-00.txt
- BGP-LS segment routing extension
  - https://datatracker.ietf.org/doc/html/draft-ietf-idr-bgp-ls-sr-policy-00
- BGP-LS flexible algorithm extension
  - https://datatracker.ietf.org/doc/rfc9351/

Backup

Related Materials

https://datatracker.ietf.org/meeting/114/materials/slides-114-mops-exposure-of-telefonica-network-topology-through-alto-for-integration-with-telefonica-cdn-01

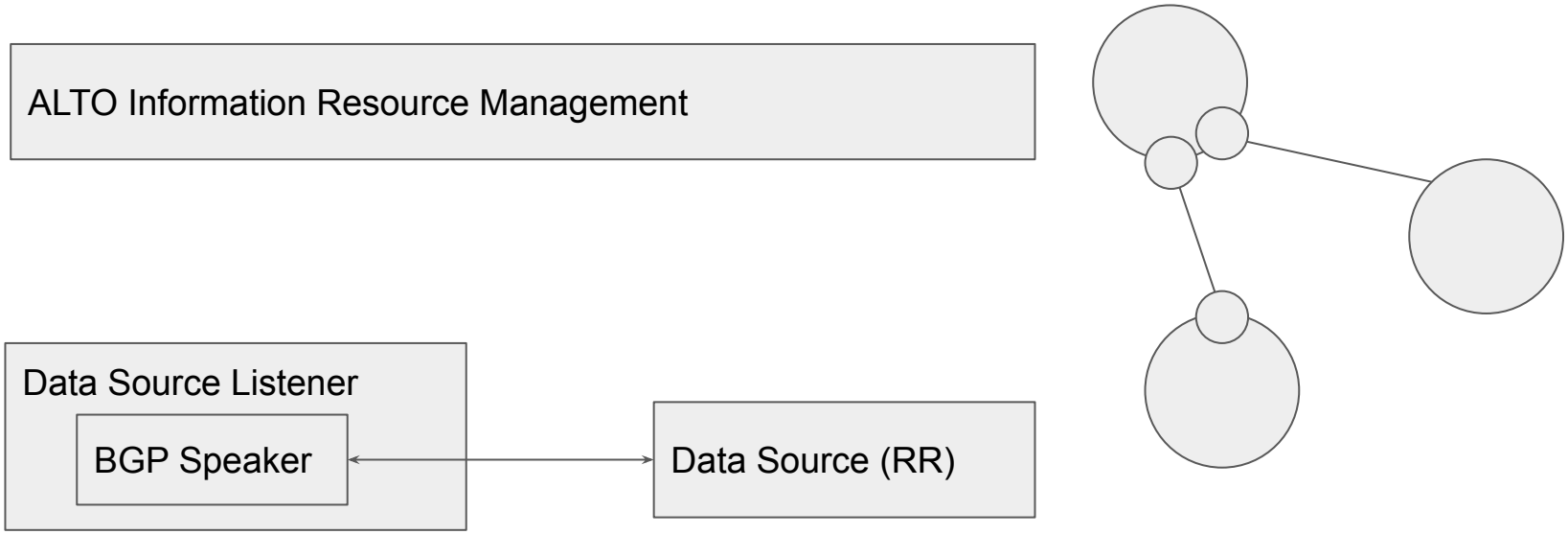
https://datatracker.ietf.org/meeting/115/materials/slides-115-mops-exposure-of-telefonica-network-topology-through-alto-for-integration-with-telefonica-cdn-00

https://content.cisco.com/chapter.sjs?uri=/searchable/chapter/content/en/us/td/docs/ios-xml/ios/iproute_bgp/configuration/xe-16/irg-xe-16-book/bgp-ls.html.xml

Network Information -> ALTO Maps

- Network Map
  - Topology information
    - Inter-domain topology and intra-domain topology
    - CIDRs distribution
- Cost Map
  - Routing information
    - end-to-end routes
  - Performance metrics information
    - routing cost
- BGP-LS is a potential approach to collect all the information above

Setup


ALTO Information Resource Management
Data Source Listener
BGP Speaker
Data Source (RR)

Settings:

1 AS - 1..n IGP domains

1 AS - all **border routers with BGP-LS enabled** (logically) connected to **ALTO**

BGP-LS Benefits and Limitations

- BGP-LS [RFC7752] allows a BGP speaker to advertise link state database or traffic engineering database of its connected IGP areas
- Benefits
  - Could be a unified interface to advertise IGP topology, routing information and additional performance metrics
  - Reuse existing BGP sessions (no extra connection is required)
- Limitations
  - Only one-hop advertisement: Cannot be propagated to remote routing servers

Map Computation

1. Map prefixes to border routers
2. Get topology of each area using BGP-LS
3. Compute the topology of the AS by combining the topology of each area, with information that how areas are connected
4. Compute end-to-end routes using the corresponding link-state routing algorithm
5. Compute the distance between any two border routers (i.e., PIDs)

BGP

prefix -> next-hop (router-id?)

BGP-LS

node (asn, BGP-LS id, area, router-id)?

link (node1, node2, metric(s))

prefix (node, route type, prefix)

ALTO Deployment Consideration

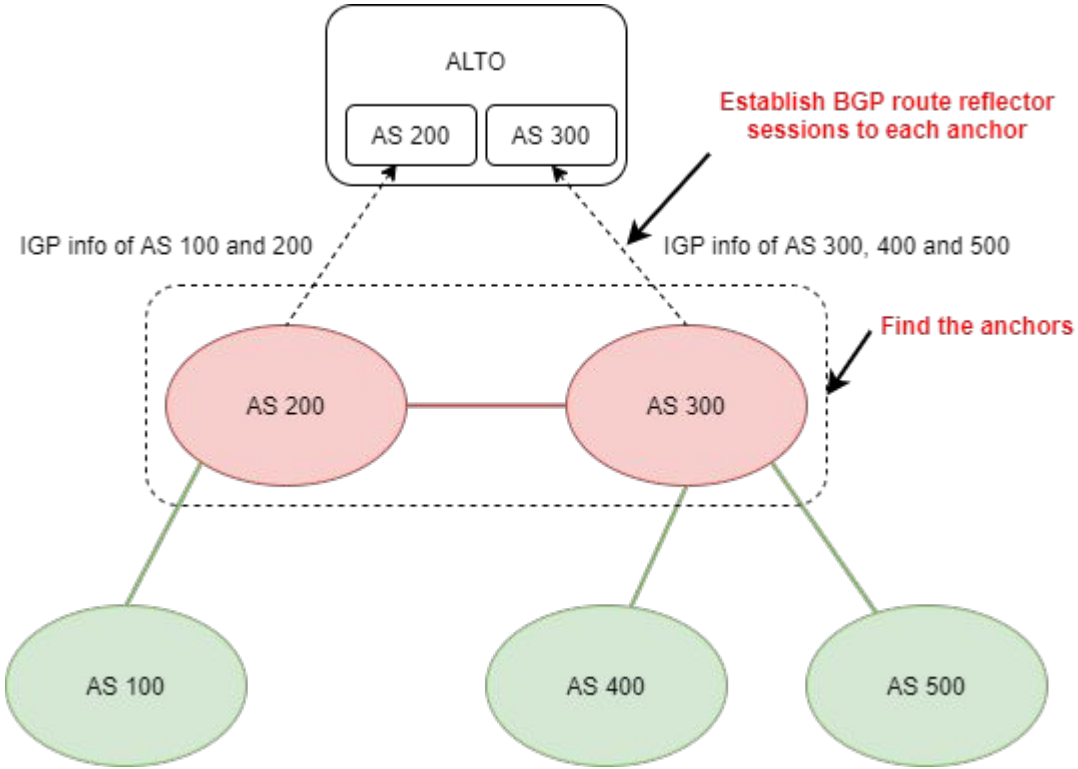
Talking to every BGP routers is inefficient

Efficiency requirements on deploying ALTO using BGP-LS:

- **Req 4**: The ALTO server SHOULD only communicate with necessary BGP speakers.
- **Req 5**: The ALTO server SHOULD only enable BGP-LS advertisement on necessary BGP sessions between BGP speakers.

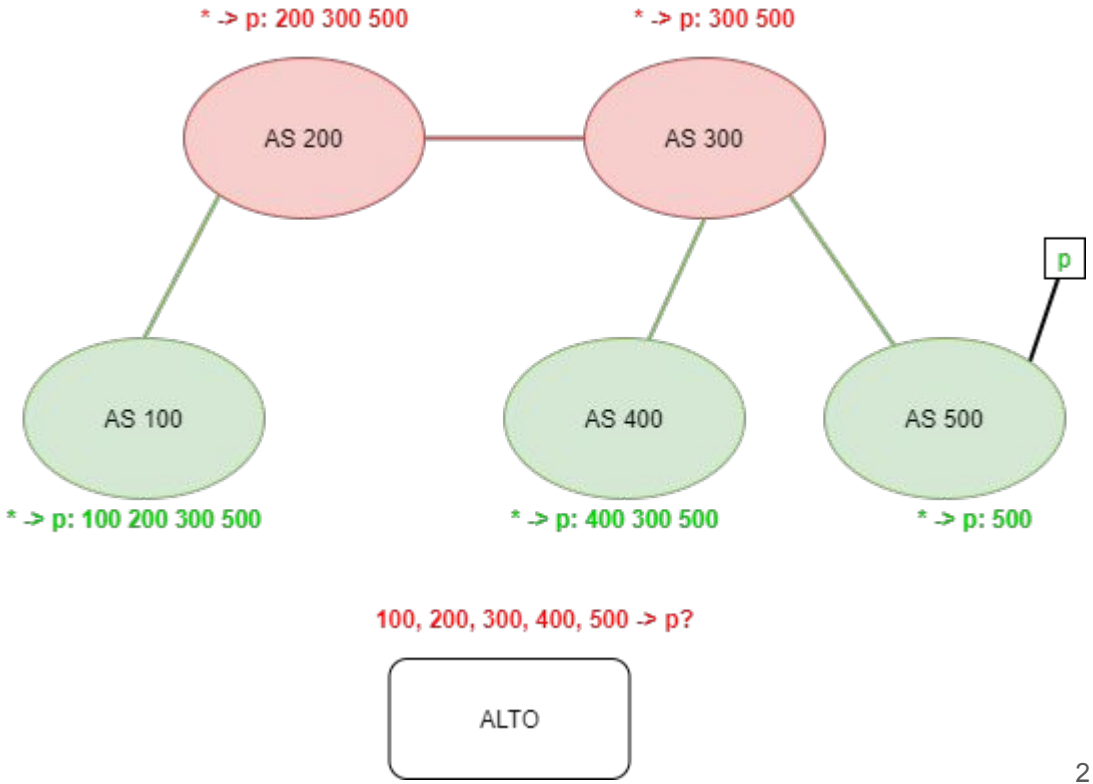
IGP Topology Collection: Full-coverage Problem

- Select a minimal set of ASes as anchors
- Each AS establishes a BGP-LS session with at least one anchor
- Each anchor mirrors its received BGP-LS advertisements with the ALTO server



Routing Information Collection

- The route from a downstream AS can be inferred by the route from an upstream AS
- Select a minimal set of ASes as anchors
- Each AS is a provider or remote provider of at least one anchor
- Each anchor mirrors its BGP RIB with the ALTO server



Summary

- BGP-LS can be used to provide necessary information
- How to collect necessary information by establishing minimal BGP sessions should be considered