Authors: J. Zhang   K. Gao
         Tongji University Sichuan University
         LM. Contreras   A. Escribano   P. Cano   F. Cano
         Telefonica   Alten   UST Global   Telefonica

**Considerations of Deploying ALTO using BGP - Link State (BGP-LS)
Advertisement**

Abstract

   This document discusses the requirements and deployment
   considerations of providing Application-Layer Traffic Optimization
   (ALTO) information in the inter-domain scenario using Border Gateway
   Protocol - Link State (BGP-LS) extension.

Requirements Language

   The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT",
   "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this
   document are to be interpreted as described in [RFC2119].

Status of This Memo

Copyright Notice

Table of Contents

1. **Introduction**

   The major component of the Application-Layer Traffic Optimization
   (ALTO) [RFC7285] deployment is the network information collection.
   [RFC7971] discussed multiple options to collect the network
   information from the inter-domain networks.

   To collection the related network information for ALTO, the
   following high-level questions should be considered:

     *Can the ALTO service realistically discover that information?

*Is the distribution of that information allowed by the operators
   of that service?

  *Is it information that a client cannot find easily some other
   way?

The Border Gateway Protocol - Link State (BGP-LS) extension
[RFC7752] is one of the popular options and has been deployed in
many Autonomous Systems (ASes) in recent years [TODO: Need some
reference].

BGP-LS enables ALTO server to provide underlay inter-domain topology
information using the link-state information in IGP domains.

To leverage BGP-LS to generate ALTO information effectively, some
requirements for deployment should be considered.

This document discusses these requirements and the corresponding
deployment considerations.

Additionally, this document describes some inter-domain scenarios to
test the deployment.

2. New Problem Statement and Working Items

This document was initially written to summarize ALTO deployment
consideration using BGP-LS. However, authors are finding new
interesting practical problems when pushing the deployment to large-
scale ISP networks.

Authors identify two problems:

Problem 1: how to efficiently obtain fine-grained global ALTO
information from multiple networks?

Problem 2: how to efficiently reconstruct and disseminate the ALTO
information upon dynamics?

Considering the scale of the ISP carrier networks and the frequency
of network dynamicity, the previous design cannot survive. We have
to target a systemtic design to support (1) distributed information
collection, and (2) calculation and incremental information
recomputation.

Thus, authors propose a hierarchical architecture to deploy ALTO
servers. To make the deployment in different small networks
compatible with each other, the interfaces to allow interoperations
between different ALTO servers are required. Although we can design
and implement those interfaces outside the scope of ALTO, we believe

making ALTO provide this capability by itself can be a more coherent approach and also have more potential benefits.

We are still heavily working on the initial specification. No more details are added in the current document. But we have already done a paper submission to talk about the initial design. For people interested in this work, please feel free to contact us.

3. Background

3.1. ALTO Inter-domain Deployment Problem

[RFC7971] discusses considerations of ALTO deployment in different network scenarios. The inter-domain network is the most common scenario to deploy ALTO.

In practice, the following approaches are used to collect information from the network:

  *Interior Gateway Protocols (IGPs, e.g., OSPF, IS-IS): intra-domain topology, link weights

  *Border Gateway Protocol (BGP): inter-domain topology, prefixes, AS numbers, AS distances, or other BGP metrics

  *Network Management Protocols (NMPs, e.g., SNMP, Netconf): latency, utilization, bandwidth

3.2. BGP-LS Background and Benefits for ALTO

BGP-LS [RFC7752] is designed to allow a BGP speaker to advertise the link state database (LSDB) or traffic engineering database (TED) of its connected IGP area.

BGP-LS defines a new address family, link state, in the BGPv4 framework [RFC4271].

Using BGP-LS, the ALTO server can communicate to only BGP speakers to collect all those information.

3.3. ALTO Deployment Problem using BGP-LS

A simple deployment solution is to connect the ALTO server as a BGP reflector client of every BGP speakers in the network. However, this solution is expensive and redundant. And because of the BGP updates, the ALTO server could receive a lot of inconsistent redundant informaiton. To avoid the redundancy and inconsistency of the collected information, a deployment solution should be minimal.

To understand what is a minimal solution to deploy ALTO using BGP-LS, the following questions are raised:

  *Is it necessary to connect the ALTO server to every AS within a BGP session?

  *Does the session between the ALTO server and each AS have to enable BGP-LS?

  *If using BGP-LS, can the number of necessary BGP sessions be reduced?

The following example shows a minimal deployment in a simple example topology.

Consider the following AS-level topology as an example. Assuming all the BGP sessions between ASes have enabled BGP-LS, the BGP speaker on AS B can received the IGP topologies from all the three ASes. Thus, to make sure the ALTO server collect all the inter-domain and intra-domain topology information, the minimal deployment could be to set up the the ALTO server as a BGP reflector of the BGP speaker on AS B.

```
+--------+   +--------+   +---------+
| AS A   |---| AS B   |---| AS C    |
+--------+   +--------+   +---------+
     | BGP     / BGP-LS
     |        /
     |       /
+-------------+
| ALTO Server |
+-------------+
```
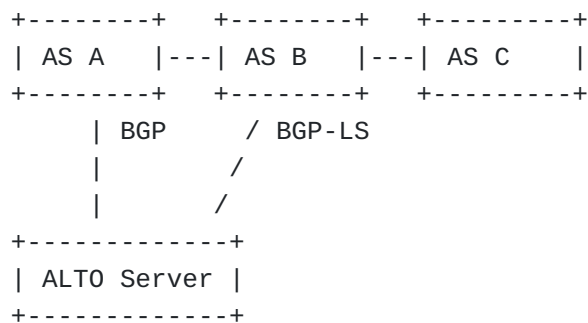
Figure 1: Example AS-level Topology

However, it is not enough for collecting the routing information. As the BGP is a destination-based routing protocol, AS B could not receive the routing information between endpoints from AS A and AS C. To get the missing routing information, the ALTO server should also connect read the BGP RIB of AS A or AS C at least.

As the result, the minimal solution is to establish a BGP session to AS B with BGP-LS and another BGP session to AS A (or AS C) without BGP-LS.

The following part of this document will discuss how to achieve the minimal ALTO deployment using BPG-LS in detail. Specifically, two questions are required to be answered:

   *Which BGP speakers are required to be connected to the ALTO server?

   *Which BGP sessions are required to enable BGP-LS?

4. Requirements for Deploying ALTO in the Inter-domain Scenario using BGP-LS

4.1. Basic Requirements

The following basic requirements are required by ALTO inter-domain deployment in any case.

Req 1: The ALTO server MUST be able to collect topology information from multiple IGP areas.

Req 2: The ALTO server MUST be able to collect routing information for any pairs of endpoints.

Req 3: The ALTO server MUST be able to collect performance metrics across routes.

4.2. BGP-LS specific Requirements

The following additional requirements are required by ALTO deployment when 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 in necessary BGP sessions between BGP speakers.

5. ALTO Deployment Considerations using BGP-LS

This section discusses some deployment considerations about how to address the basic requirements (Req 1-3) when satisfying the BGP-LS specific requirements (Req 4-5).

5.1. Provisioning of Topology Information

As BGP-LS advertisement cannot be propagated to remote the remote ASes, each BGP speaker can only discover directly peered IGP topologies using BGP-LS.

To satisfy Req 4, the ALTO server should only communicate to transit
networks or IXPs using BGP-LS. As the IGP topology of a stub network
can always be discovered by its peered transit networks or IXPs, so
it is not necessary to communicate with the stub network.

Specifically, the ALTO server should find a minimal BGP speaker set
whose peered networks can cover all IGP domains.

5.2. Provisioning of Routing Information

As BGP is a destination-based routing protocol, a stub network can
receive all the inter-domain routing information from all the
reachable destinations via BGP.

Thus, to satisfy Req 4, the ALTO server should only communicate to
stub networks using BGP, as the inter-domain routing information
from the transit networks is not necessary.

Assuming the ALTO server has already collected the complete topology
information using BGP-LS, the ALTO server will have the LSDB of
every IGP domain.

To satisfy Req 5, all the BGP sessions connected to the stub
networks do not have to enable BGP-LS.

5.3. Provisioning of Performance Metric Information

TBD.

6. Configuration Interfaces of Map Calculation

6.1. Configuration Interface of Network Map Calculation

```
rw network-map-config* [resource-id]
+--rw resource-id      alto-types:resource-id
+--rw description?     string
+--rw (params)
|  +--:(bgp)
|     +--rw bgp-params
|        +--rw bgp-rib* [rib-id]
|           +--rw rib-id        rib:rib-id
|           +--rw topology-id? topology:topology-id
|           +--rw bgp-ls?       boolean
+--rw (algorithm)
   +--:(first-hop-cluster)
      +--rw first-hop-cluster-algorithm
         +--rw inspect-igp      boolean
```

To generate a network map, one or more BGP RIBs that could provide
the topology information MUST specified. Each BGP RIB MAY include a

pre-computed topology from the RIB, and an option indicating if the
BPG-LS is enabled.

The inspect-igp option in the first-hop-cluster-algorithm field
indicates if the ALTO server exposes information about the IGP
topologies. If it is true, the ALTO server will inspect all the IGP
topolgies from the BGP RIBs that enalbe BGP-LS (whose bgp-ls option
is true).

6.2. Configuration Interface of Cost Map Calculation

```
rw cost-map-config* [resource-id]
+--rw resource-id            alto-types:resource-id
+--rw description?           string
+--rw dependent-network-map  alto-types:resource-id
+--rw (general-params)
|  +--:(bgp)
|     +--rw bgp-params
|        +--rw alternative-bgp-rib* [rib-id]
|           +--rw rib-id            rib:rib-id
|           +--rw topology-id?      topology:topology-id
|           +--rw bgp-ls?           boolean
+--rw cost-type* [cost-mode,cost-metric]
   +rw cost-mode                 alto-types:cost-mode
   +rw cost-metric               alto-types:cost-metric
   +rw (params)?
```

To generate a cost map, besides the dependent network map, one or
more alternative BGP RIBs could be specified to provide necessary
routing information to the ALTO server.

6.3. Configuration Examples

Example Network
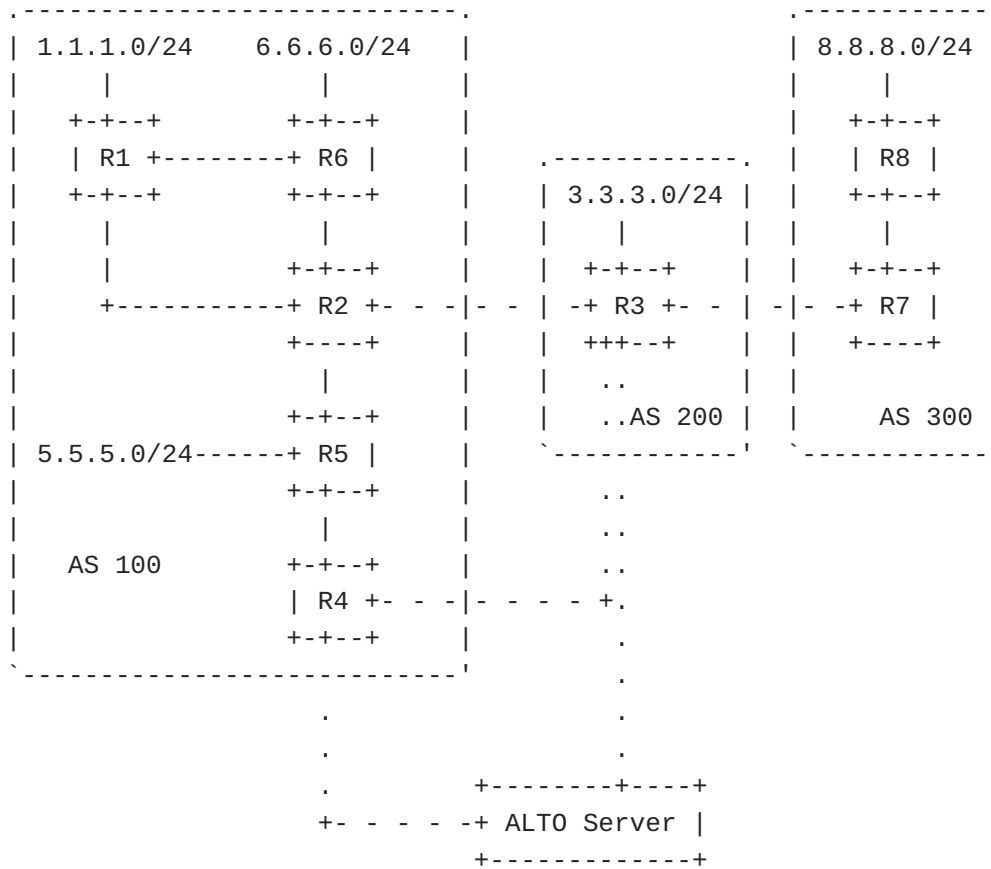
```
 .-----------------------------.                        .------------.
 | 1.1.1.0/24    6.6.6.0/24    |                        | 8.8.8.0/24 |
 |      |             |        |                        |     |      |
 |   +-+--+        +-+--+      |                        |   +-+--+   |
 |   | R1 +--------+ R6 |      |      .------------.     |   | R8 |   |
 |   +-+--+        +-+--+      |      | 3.3.3.0/24 |     |   +-+--+   |
 |     |             |         |      |      |     |     |     |      |
 |     |           +-+--+      |      |   +-+--+   |     |   +-+--+   |
 |     +-----------+ R2 +- - -|- - |  -+ R3 +- - | -|- -+ R7 |   |
 |                 +----+      |      |   +++--+   |     |   +----+   |
 |                   |         |      |     ..     |     |            |
 |                 +-+--+      |      |     ..AS 200|    |    AS 300  |
 | 5.5.5.0/24------+ R5 |      |      `------------'    `------------'
 |                 +-+--+      |            ..
 |                   |         |            ..
 |    AS 100       +-+--+      |            ..
 |                 | R4 +- - -|- - - - +.
 |                 +-+--+      |            .
 `-----------------------------'           .
                   .                       .
                   .                       .
                   .           +--------+----+
                 +- - - - -+ ALTO Server |
                             +-------------+

 R2 - R3: BGP-LS
 R4 - R3: BGP-LS
 R7 - R3: BGP-LS
 R3 - ALTO: BGP-LS
 R4 - ALTO: BGP
```

Figure 2

Config a network map:

```
POST /restconf/config/alto-maps/network-map-config/bgp-networkmap
HOST: alto-config.example.com
Content-Type: application/json
Content-Length: TBD

{
  "network-map-config": {
    "resource-id": "bgp-networkmap",
    "bgp-params": {
      "bgp-rib": [
        {
          "rib-id": "as200-r3",
          "bgp-ls": true
        }
      ]
    },
    "first-hop-cluster-algorithm": {
      "inspect-igp": true
    }
  }
}
```

Test to fetch the network map:

```
GET /alto/networkmap/example
HOST: alto.example.com

HTTP/1.1 200 OK
Content-Length: TBD
Content-Type: application/alto-networkmap+json

{
  "meta": {
    "vtag": {
      "resource-id": "bgp-networkmap",
      "tag": "da65eca2eb7a10ce8b059740b0b2e3f8eb1d4785"
    }
  },
  {
    "network-map": {
      "PID100.R1": {
        "ipv4": [ "1.1.1.0/24" ]
      },
      "PID100.R5": {
        "ipv4": [ "5.5.5.0/24" ]
      },
      "PID100.R6": {
        "ipv4": [ "6.6.6.0/24" ]
      },
      "PID200.R3": {
        "ipv4": [ "3.3.3.0/24" ]
      },
      "PID300.R8": {
        "ipv4": [ "8.8.8.0/24" ]
      }
    }
  }
}
```

   Config a cost map:

```
POST /restconf/config/alto-maps/cost-map-config/bgp-costmap
HOST: alto-config.example.com
Content-Type: application/json
Content-Length: TBD

{
  "cost-map-config": {
    "resource-id": "bgp-costmap",
    "dependent-network-map": "bgp-networkmap",
    "bgp-params": {
      "alternative-bgp-rib": [
        {
          "rib-id": "as100-r4",
          "bgp-ls": false
        }
      ]
    },
    "cost-type": [
      {
        "cost-mode": "numerical",
        "cost-metric": "hopcount"
      }
    ]
  }
}
```

   Test to fetch the cost map:

```
GET /alto/costmap/bgp-costmap
HOST: alto.example.com

HTTP/1.1 200 OK
Content-Length: TBD
Content-Type: application/alto-costmap+json

{
  "meta": {
    "vtag": {
      "resource-id": "bgp-costmap",
      "tag": "c0ce023b8678a7b9ec00324673b98e54656d1f6d"
    },
    "dependent-vtags": [
      {
        "resource-id": "bgp-networkmap",
        "tag": "da65eca2eb7a10ce8b059740b0b2e3f8eb1d4785"
      }
    ],
    "cost-type": {
      "cost-mode": "numerical",
      "cost-metric": "hopcount"
    }
  },
  "cost-map": {
    "PID100.R1": {
      "PID100.R1": 1, "PID100.R5": 3, "PID100.R6": 2,
      "PID200.R3": 3, "PID300.R8": 5
    },
    "PID100.R5": {
      "PID100.R1": 3, "PID100.R5": 1, "PID100.R6": 3,
      "PID200.R3": 3, "PID300.R8": 5
    },
    "PID100.R6": {
      "PID100.R1": 2, "PID100.R5": 3, "PID100.R6": 1,
      "PID200.R3": 3, "PID300.R8": 5
    },
    "PID200.R3": {
      "PID100.R1": 3, "PID100.R5": 3, "PID100.R6": 3,
      "PID200.R3": 1, "PID300.R8": 3
    },
    "PID300.R8": {
      "PID100.R1": 5, "PID100.R5": 5, "PID100.R6": 5,
      "PID200.R3": 3, "PID300.R8": 1
    }
  }
}
```

7. Test Scenarios

7.1. Test Environment Setup

7.2. Test Approach

7.3. Test Results

8. References

8.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate
           Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/
           RFC2119, March 1997, <https://www.rfc-editor.org/info/
           rfc2119>.

[RFC4271] Rekhter, Y., Ed., Li, T., Ed., and S. Hares, Ed., "A
           Border Gateway Protocol 4 (BGP-4)", RFC 4271, DOI
           10.17487/RFC4271, January 2006, <https://www.rfc-
           editor.org/info/rfc4271>.

[RFC7285] Alimi, R., Ed., Penno, R., Ed., Yang, Y., Ed., Kiesel,
           S., Previdi, S., Roome, W., Shalunov, S., and R. Woundy,
           "Application-Layer Traffic Optimization (ALTO) Protocol",
           RFC 7285, DOI 10.17487/RFC7285, September 2014, <https://
           www.rfc-editor.org/info/rfc7285>.

[RFC7752] Gredler, H., Ed., Medved, J., Previdi, S., Farrel, A.,
           and S. Ray, "North-Bound Distribution of Link-State and
           Traffic Engineering (TE) Information Using BGP", RFC
           7752, DOI 10.17487/RFC7752, March 2016, <https://www.rfc-
           editor.org/info/rfc7752>.

[RFC7971]
           Stiemerling, M., Kiesel, S., Scharf, M., Seidel, H., and
           S. Previdi, "Application-Layer Traffic Optimization
           (ALTO) Deployment Considerations", RFC 7971, DOI
           10.17487/RFC7971, October 2016, <https://www.rfc-
           editor.org/info/rfc7971>.

[RFC8571] Ginsberg, L., Ed., Previdi, S., Wu, Q., Tantsura, J., and
           C. Filsfils, "BGP - Link State (BGP-LS) Advertisement of
           IGP Traffic Engineering Performance Metric Extensions",
           RFC 8571, DOI 10.17487/RFC8571, March 2019, <https://
           www.rfc-editor.org/info/rfc8571>.

8.2. Informative References

[I-D.ietf-alto-cost-calendar]

Randriamasy, S., Yang, Y., WU, Q., Lingli, D., and N. Schwan, "Application-Layer Traffic Optimization (ALTO) Cost Calendar", Work in Progress, Internet-Draft, draft-ietf-alto-cost-calendar-21, 17 March 2020, <http://www.ietf.org/internet-drafts/draft-ietf-alto-cost-calendar-21.txt>.

[I-D.ietf-alto-incr-update-sse]
          Roome, W. and Y. Yang, "ALTO Incremental Updates Using Server-Sent Events (SSE)", Work in Progress, Internet-Draft, draft-ietf-alto-incr-update-sse-22, 20 March 2020, <http://www.ietf.org/internet-drafts/draft-ietf-alto-incr-update-sse-22.txt>.

[I-D.ietf-alto-path-vector]
          Gao, K., Randriamasy, S., Yang, Y., and J. Zhang, "ALTO Extension: Path Vector", Work in Progress, Internet-Draft, draft-ietf-alto-path-vector-10, 9 March 2020, <http://www.ietf.org/internet-drafts/draft-ietf-alto-path-vector-10.txt>.

[I-D.ietf-alto-unified-props-new]
          Roome, W., Randriamasy, S., Yang, Y., Zhang, J., and K. Gao, "Unified Properties for the ALTO Protocol", Work in Progress, Internet-Draft, draft-ietf-alto-unified-props-new-11, 9 March 2020, <http://www.ietf.org/internet-drafts/draft-ietf-alto-unified-props-new-11.txt>.

[I-D.wang-idr-bgpls-inter-as-topology-ext]
          Wang, A. and H. Chen, "BGP-LS Extension for Inter-AS Topology Retrieval", Work in Progress, Internet-Draft, draft-wang-idr-bgpls-inter-as-topology-ext-02, 13 August 2018, <http://www.ietf.org/internet-drafts/draft-wang-idr-bgpls-inter-as-topology-ext-02.txt>.

[RFC8189] Randriamasy, S., Roome, W., and N. Schwan, "Multi-Cost Application-Layer Traffic Optimization (ALTO)", RFC 8189, DOI 10.17487/RFC8189, October 2017, <https://www.rfc-editor.org/info/rfc8189>.

Authors' Addresses

Jingxuan Jensen Zhang
Tongji University
4800 Cao'An Hwy
Shanghai
201804
China

Email: jingxuan.zhang@tongji.edu.cn

Kai Gao
Sichuan University
No.24 South Section 1, Yihuan Road
Chengdu
610000
China

Email: kaigao@scu.edu.cn

Luis M. Contreras
Telefonica
Ronda de la Comunicacion, s/n
28050 Madrid
Spain

Email: luismiguel.contrerasmurillo@telefonica.com

Anais Escribano
Alten
Carrer de Josep Pla, 2
08019 Barcelona
Spain

Email: anais.escribano@alten.es

Patricia Cano
UST Global
Ramirez de Arellano 29
28043 Madrid
Spain

Email: Patricia.Diez@ust-global.com

Francisco Cano
Telefonica
Avenida del Conocimiento, 12
18016 Granada
Spain

Email: franciscojose.canohila@telefonica.com