Exposure of Telefonica network topology through ALTO for integration with Telefonica CDN

Luis M. Contreras (*), Patricia Diez | Telefónica GCTIO – Transport Group
Francisco Cano, Anais Escribano | Telefónica CCDO – Video Group

MOPS WG meeting @ IETF 114, Philadelphia, July 2022

(*) luismiguel.contrerasmurillo@telefonica.com / contreras.ietf@gmail
Background

• Telefónica is an international telecom operator with presence in 15 countries in Europe and Latin America (+ a Tier-1 international carrier)
  • Multiservice Networks: fixed, mobile, enterprise, etc

• In house development of CDN solution (Telefonica CDN → TCDN)
  • Distribution of content associated to the Telefonica video services (Movistar+ & Movistar Play) and 3rd party video content distribution
  • Worldwide deployed (Spain, Brazil, Argentina, Chile, Colombia, Peru, Ecuador, Mexico, Uruguay and USA. Ongoing deployment also in Germany)
  • Delivery of video using HTTP Adaptive Streaming (HAS) protocols for both internal and external customers to Telefónica

• State-of-the-art request routing logic (RRL) considers multiple information in order to maximize both QoE and video delivery efficiency
  • RRL inputs: streamer health status and load level, cache hit ratio maximization, content popularity, ..., and network topology (PIDs and cost matrix)
  • By now, network topology is generated and provisioned manually ← Initial driver for this activity
Rationale for making TCDN to be transport network aware

• One of the main objective of TCDN is to provide an efficient delivery of contents within the network

• Content delivery is based nowadays on a (semi-)static view of the network, decoupled from the real situation along time

• In order to make a complete and efficient usage of the network, TCDN would benefit from a real time knowledge of the status and characteristics of the network
  • For instance, allowing delivery decisions in TCDN to quickly adapt to network status variation (e.g., topology changes, congestion, etc.)
How the network looks like?

High level view of a network in a mid-size country

IP Transport Hierarchy Levels (HL)

1. Interconnection level
   - Interconnection Routers ~10 nodes

2. Backbone level
   - P Routers ~20 nodes

3. Province level
   - PE Platforms ~200 nodes

4. Local level
   - IP Edge (BPE, BRAS) ~1K nodes
   - OLT
   - DSLAMS

5. Access/ Capillarity level
   - Cell Site Gateway ~10K nodes
   - OLT
   - DSLAMS

Different levels of distribution of TCDN delivery points

HL1. Located at the top level of the IP network. Interface between IP network and Internet/ISP providers.

HL2. Located after the HL1 routers and aggregate traffic to/from the HL3 routers.

HL3. Provide connectivity to the platforms.

HL4. This is the IP Edge of the network with the Business PE and BRAS functionalities. HL4s aggregate the traffic of the fixed and mobile customers.

HL5. They are in charge of aggregating the mobile traffic and some fixed customers.

End users connected to HL4 / HL5
Example – excerpt representing some of the existing HL1/HL2/HL3
How TCDN leverages on ALTO information?

• **Pid_file** and **cost_map** generated with BGP and BGP-LS updates showing information from the network
  - PIDs associated to customer’s IP prefixes represent consumers of video streaming
  - PIDs associated to the connection of CDN streamers represent the potential sources of TCDN traffic.

• By checking network and cost map info from both kinds of PIDs it’s easy to match CDN streamers with customers.

• In order to select the more convenient streamer in each case, the RRL can be complemented with the view of the lowest cost between PIDs of CDN streamers and PIDs of customers.
  - For example, for a given PID of customers, e.g. pid0:0a0a0a01, the more convenient streamer can be determined from the lowest cost of pid0:0a0a0a05 and pid0:0a0a0a06 (assuming the rest of considerations in RRL is similar).
Network map & Cost map

Network Map

```json
{
    "pid0:0a0a0a01": [
    "1.1.1.0/24"
    ],
    "pid0:0a0a0a02": [
    "2.2.2.0/24"
    ],
    "pid0:0a0a0a03": [
    "3.3.3.0/24"
    ],
    "pid0:0a0a0a05": [
    "11.11.1.0/30"
    ],
    "pid0:0a0a0a06": [
    "22.22.2.0/30"
    ]
}
```

Cost Map

```json
{
    "pid0:0a0a0a01": {
        "pid0:0a0a0a01": 0,
        "pid0:0a0a0a02": 2,
        "pid0:0a0a0a03": 2,
        "pid0:0a0a0a05": 2,
        "pid0:0a0a0a06": 1
    },
    "pid0:0a0a0a02": {
        "pid0:0a0a0a01": 2,
        "pid0:0a0a0a02": 0,
        "pid0:0a0a0a03": 2,
        "pid0:0a0a0a05": 2,
        "pid0:0a0a0a06": 1
    },
    "pid0:0a0a0a03": {
        "pid0:0a0a0a01": 2,
        "pid0:0a0a0a02": 2,
        "pid0:0a0a0a03": 0,
        "pid0:0a0a0a05": 2,
        "pid0:0a0a0a06": 1
    },
    "pid0:0a0a0a05": {
        "pid0:0a0a0a01": 2,
        "pid0:0a0a0a02": 2,
        "pid0:0a0a0a03": 2,
        "pid0:0a0a0a05": 0,
        "pid0:0a0a0a06": 3
    },
    "pid0:0a0a0a06": {
        "pid0:0a0a0a01": 1,
        "pid0:0a0a0a02": 1,
        "pid0:0a0a0a03": 1,
        "pid0:0a0a0a05": 3,
        "pid0:0a0a0a06": 0
    }
}
```

A PID can represent a set of prefixes assigned either to clients or to CDN delivery points.
Process followed

Initial tests in lab environment with simple topology. Simplistic network configuration (e.g., OSPF) for understanding viability of the approach and get experience.

Pre-production Lab

Integration on a pre-production environment with realistic network topologies, and network configuration as in the production network, with the purpose of assess the solution and solve problems found.

Production Network

Deployment of ALTO as an element of the production network and full integration with TCDN
"Engineering" path followed

<table>
<thead>
<tr>
<th>Technology lab tests</th>
<th>Pre-production network tests</th>
<th>Integration in production network</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Initial tests with ALTO module of ODL</td>
<td>• Migration to exaBGP</td>
<td>• Adaptation to production processes and rules</td>
</tr>
<tr>
<td>• Integration with ODL BGP (originally LLDP)</td>
<td>• Fixing of issues in exaBGP (3 tickets raised and solved) mainly related to BGP-LS[*]</td>
<td>• Hardening of all the environment to prevent security issues (HW, SW, …)</td>
</tr>
<tr>
<td>• Monovendor router scenario</td>
<td>• Multivendor router scenario</td>
<td>• Limited activation of BGP-LS by now</td>
</tr>
<tr>
<td>• Virtualized routers</td>
<td>• Physical routers</td>
<td>• Coexistence with many other services in the network</td>
</tr>
<tr>
<td>• Virtualized ALTO</td>
<td>• Dedicated ALTO server</td>
<td>• Complete deployment expected for Q3’22</td>
</tr>
</tbody>
</table>
| • Simple IP network based on OSPF as IGP | • Complex MPLS network combining OSPF and IS-IS | [^]https://github.com/Exa-Networks/exabgp/issues/1071
| • Single AS | • Multiple private ASs | [^]https://github.com/Exa-Networks/exabgp/pull/1075
| • Simple metrics (= hopcount) | • More sophisticated metrics in IGP | [^]https://github.com/Exa-Networks/exabgp/issues/1077
| • Some of the routers acting as RR | • Dedicated RR, separated for BGP and BGP-LS |
Next steps

• Integration of ALTO in Telefonica Spain’s network, for later planned integration in other TCDN countries (Brazil, Germany, Argentina, ...)

• Full development of the TCDN logic for automatic consumption of ALTO generated topology

• Characterization of the ALTO performance on a real environment (scalability, processing, etc)

• Analysis / implementation of the other ALTO capabilities / use cases (network performance / congestion exposure, access network info, ...)

Conclusion

• ALTO enables an operational integration between TCDN and the Transport Network allowing the automatic and timely acquisition of network topology (e.g., network outages), eliminating manual generation of network topology information and its periodic update. This allow both to increase the service QoE and the video delivery efficiency.

• Further work considers to enrich existing information, e.g., adding performance metrics, leveraging on ALTO extensions

Feedback

• Feedback from MOPS is very welcomed
  • It can also help to gather further directions from MOPS in ALTO WG (alto@ietf.org mailing list is a perfect venue for discussion)

Acknowledgements

• We are extremely thankful with Kai Gao (Sichuan University) and Jensen Zhang (Tongji University) for their continuous support in the initial steps of this work
Backup
How to make it? Integration of ALTO and Transport Network to assist TCDN decisions

- ALTO abstract topology (Network Maps) can be automatically generated from the physical Network topology, allowing also the application of policies.
- Information about IP prefixes and TE can be obtained. Prefixes can be distributed by BGP while TE information can be obtained from the IGP protocol (OSPF / IS-IS).
- BGP and BGP-LS can advertise all that information to the ALTO Server

RFC7752, “North-Bound Distribution of Link-State and Traffic Engineering (TE) Information Using BGP” – e.g. ALTO server using topology and TE information
Network map creation

Pid name: \textit{Pid} + \textit{ASN (default value = 0)} + IP next-hop in hexadecimal:

\begin{align*}
\text{“pid”} + 0 + \text{hex(192.168.255.12)} \rightarrow \text{pid0: c0a8ff0c}
\end{align*}

BGP advertisement with address family ipv4 unicast. The next-hop value (192.168.255.12) is used for the PID identifier and the NLRI value (80.58.102.64/26) is one of the prefixes of such PID.

\begin{verbatim}
{ "exabgp": "5.0.0", "time": 1651759327.1712353, "host": "localhost.localdomain", "pid": 72259, "ppid": 1, "counter": 135, "type": "update", "neighbor": { "address": { "local": "80.58.171.201", "peer": "192.168.252.200" }, "asn": { "local": 64531, "peer": 64531 } }, "direction": "receive", "message": { "update": { "attribute": { "origin": "igp", "med": 0, "local-preference": 100, "originator-id": "192.168.255.12", "cluster-list": [ "3.3.2.1", "192.168.251.172" ] }, "announce": { "ipv4 unicast": { "192.168.255.12": [ { "nlri": "80.58.102.64/26" } ] } } } }
\end{verbatim}

Leveraging on \textit{UPDATE messages from BGP RR}
Cost map creation

Leveraging on UPDATE messages from BGP-LS RR

bgpls-link messages

```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-Is bgp-Is": { "192.168.252.178": { "ls-nlri-type": "bgpls-link", "l3-routing-topology": 2, "protocol-id": 2, "local-node-descriptors": { "autonomous-system": 3352, "bgp-Is-identifier": "178" }, "router-id": "d500b8070000" } }, "remote-node-descriptors": { "autonomous-system": 3352, "bgp-Is-identifier": "178" }, "interface-addresses": [ "192.168.204.197" ], "neighbor-addresses": [ "192.168.204.197" ], "multi-topology-ids": [ ], "link-identifiers": [ ] } } }
}
```

bgpls-node messages

```json
{
"exabgp": "5.0.0", "time": 1652431307.7986917, "host": "localhost.localdomain", "pid": 81254,
"ppid": 1, "counter": 774, "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-Is": { "node-flags": { "O": 1, "T": 0, "E": 0, "B": 1, "R": 0, "V": 0, "RSV": 0 }, "area-id": "1346612317385085419520", "local-te-router-ids": [ "192.168.181.3" ] }, "announce": { "bgp-Is bgp-Is": { "192.168.252.178": { "ls-nlri-type": "bgpls-node", "l3-routing-topology": 2, "protocol-id": 2, "node-descriptors": { "autonomous-system": 3352, "bgp-Is-identifier": "178" }, "router-id": "192168181009" } }, "nexthop": "192.168.252.178" } ] } }}
}
```
How to use ALTO information?
Information retrieval

BGP Update messages decoded in JSON format, then translated to CSV (internal format used by TCDN)
CDN-ALTO logic

Read decoded BGP message from STDOUT

BGP Announce

No

BGP Withdraw

Yes

Remove from topology

Extractor BGP attributes (igp-metric)

Yes

Load nodes and links to topology

is BGP-LS

No

is BGP

Yes

Load to PIDS

RR BGP-LS

RR BGP

STDOUT

BGP Update messages decoded in JSON format, then translated to CSV (internal format used by TCDN)

exaBGP

CDN-ALTO

ALTO
Technology lab topology

Pre-production topology (excerpt)