Steering Traffic with the ALTO Protocol to Optimize Network-Aware Application Performance

OPSAWG Meeting 7/29/2022

on behalf of IETF ALTO WG (Jordi Ros Giralt, Y. Richard Yang Jacob, Jensen, Kai, Alex, Mahdi, Roland, Sabine ...)

Outline

• What is the ALTO?
  - ALTO protocol overview
  - Basic ALTO Network Abstraction
  - Basic ALTO Transport Framework
  - ALTO Usage Examples

• ALTO WG Documents Status Update
• ALTO Implementation and Application Integration
• Summary and Takeaway
What is the ALTO?

• ALTO is an effort of the ALTO Working Group in the Transport Area of Internet Engineering Task Force (IETF)

• **ALTO high-level goal:** provide a standard for applications and networks to work together to optimize traffic patterns to improve both network and application performance
  • Initial for data distribution by peer-to-peer applications (e.g., BitTorrent)
  • Evolved to include content distribution networks (CDN)

ALTO includes two components: (1) abstractions of network states/services, and (2) transport of network information
General ALTO Protocol Overview

ALTO client

Abstractions, services
Transport

ALTO protocol

ALTO server

East-west (server-to-server)

Backend/infrastructure

Network substrate
(NMS, routing protocols)

Beyond scope of the ALTO
Basic ALTO Network Abstractions

A network consists of:

• A set of entities, where each entity can be an
  • Endpoint
  • Aggregations of endpoints (PID)
  • Abstract network element
• Entities have properties that can be inherited
  • Entities can have capabilities

• A set of paths, where each path is from a src entity to a dst entity
• A path has path properties:
  - Cost metrics (originally cost was called distance)
  - Complex properties: multicost, calendar, path vector
• Path Vector: A set of src-dst pairs can form a co-flow, with shared abstract network elements
ALTO Network Abstractions and Services

Legend:
PV: Path Vector
ECS: Endpoint Cost Service
EPS: Endpoint Property Service
Basic ALTO Transport Framework

- Bootstrap server provided by server discovery
  - Server discovery (RFC7286), xdom discovery (RFC8686)
- Network information divided into (network) information resources
- List of available information resources provided by Information Resource Directory (IRD)
- Each individual information resource is provided as a RESTful service
- Information resources distributed using either pull (RFC 7285) or incremental-push (RFC8895)

GET /maps HTTP/1.1 Host: custom.alto.example.com Accept: application/alto directory+json,application/alto-error+json

Transport Protocol: HTTP 1.1
Encoding: JSON Plaint Text for Debug
Message Format: Independent of HTTP protocol
Example: Endpoint Cost and HTTP Transport

Many performance metrics: one-way delay (delay-ow), round-trip delay (delay-rt), delay variation (delay-var), loss rate (lossrate), residual bw (bw-residual), available bw (bw-available), TCP throughput (tput), hop count.
Example: Entity Properties and SSE Incremental Push Update

- More details see RFC8895
Example: Path Vector Supporting Co-Flow

POST /endpointcost/pv HTTP/1.1
Host: alto.example.com
Accept: multipart/related;
    type=application/alto-endpointcost+json,
    application/alto-error+json
Content-Length: 362
Content-Type: application/alto-endpointcostparams+json

```
{  
    "cost-type": {  
        "cost-mode": "array",
        "cost-metric": "ane-path"
    },
    "endpoints": {  
        "srcs": [  
            "ipv4:192.0.2.34",
            "ipv6:2001:db8::3:1"
        ],
        "dsts": [  
            "ipv4:192.0.2.2",
            "ipv4:192.0.2.50",
            "ipv6:2001:db8::4:1"
        ]
    },
    "ane-property-names": [  
        "max-reservable-bandwidth",
        "persistent-entity-id"
    ]
}
```

ALTO Deployment Update

• Current implementations/deployments:

• Wiki list of implementations: https://trac.ietf.org/trac/alto/wiki/Impl
  • Examples:
    • Telefonica: https://dl.ifip.org/db/conf/im/im2021mini/212012.pdf

• Forthcoming new deployments (work in progress):
  • Pacific Research Platform
  • CERN Rucio
  • UCSD 5G
  • NY City Cosmos 5G
  • ESnet
Example ALTO Deployment Setting: Flow Director

Steering Hyper-Giants’ Traffic at Scale

Figure 9: Flow Director: High-level system architecture.

Figure 10: Flow Director processing pipeline.


CoNEXT 2019 Best Paper Award; IETF/IRTF 2020 Applied Networking Research Prize

07/28/2022

OPSAWG IETF 114 Hybrid meeting
Example ALTO Deployment Setting: Telefonica CDN deployment roadmap

Technology lab tests
- Initial tests with ALTO module of ODL
- Integration with ODL BGP (originally LLDP)
- Monovendor router scenario
- Virtualized routers
- Virtualized ALTO
- Simple IP network based on OSPF as IGP
- Single AS
- Simple metrics (= hopcount)
- Some of the routers acting as RR

Pre-production network tests
- Migration to exaBGP
- Fixing of issues in exaBGP (3 tickets raised and solved) mainly related to BGP-LS[
- Multivendor router scenario
- Physical routers
- Dedicated ALTO server
- Complex MPLS network combining OSPF and IS-IS
- Multiple private ASs
- More sophisticated metrics in IGP
- Dedicated RR, separated for BGP and BGP-LS

Integration in production network
- Adaptation to production processes and rules
- Hardening of all the environment to prevent security issues (HW, SW, …)
- Limited activation of BGP-LS by now
- Coexistence with many other services in the network
- Complete deployment expected for Q3’22

[∗] https://github.com/Exa-Networks/exabgp/issues/1071
     https://github.com/Exa-Networks/exabgp/pull/1075
     https://github.com/Exa-Networks/exabgp/issues/1077

IETF 114 / ALTO WG : ALTO Code Bases and Deployments
ALTO Hackathon Project Update (IETF 113,114)

• Implementation of an ALTO Client in Python (RFC 7285)

• Integration with CERN Rucio replica download
  • Submitted pull request to Rucio Project: –
    https://github.com/rucio/rucio/pull/5364
  • 3 Demos [https://github.com/openalto/ietf-hackathon/issues/8]
    • [D1] Single-flow replica node selection using ALTO BW Cost Map
    • [D2] ALTO Estimator: Multi-flow BW prediction
    • [D3] ALTO Scheduler: SLA-constrained multi-flow node selection

• Southbound ALTO integration with SDN:
  • Mininet/Pox, OpenDaylight
  • Scrum dashboard: https://github.com/orgs/openalto/projects/1/views/1
Summary and Take away

• ALTO is originally designed for P2P application, has evolved to support CDN application
  • ALTO further development can be driven by real application such as CAN (Compute aware Networking)
    • ALTO provide centralized approach to support service placement
    • Both network metric and compute metric can be exposed to the client
• ALTO enable network and application integration and provide a good Network visibility
  • Intelligence: AI, ML
  • Control: Service Invocation, Peer Selection, Path Selection, Resource Schedule Control, Congestion Control, etc

• Further work plan (https://github.com/ietf-wg-alto/wg-materials):
  • Telefonica TCDN deployment is coming in the 2\textsuperscript{nd} half of this year
  • Integration w/ Rucio (automatic workflow)
    - Manual workflow (client download)
    - Southbound: Provide automatic, network-aware, dynamic distance []
    - Northbound: Automatic replication QoS aware scheduling