ALTO WG
Internet-Draft
Intended status: Informational
Expires: January 13, 2022

D. Lachos
C. Rothenberg
Unicamp
Q. Xiang
Xiamen University
Y. Yang
Yale University
B. Ohlman
Ericsson Research
S. Randriamasy
Nokia Bell Labs
F. Boten
Sprint
LM. Contreras
Telefonica
J. Zhang
Tongji University
K. Gao
Sichuan University
July 12, 2021

Supporting Multi-domain Use Cases with ALTO
draft-lachos-alto-multi-domain-use-cases-02

Abstract

The goal of this document is to summarize current standardization efforts in the IETF ALTO working group to support important multi-domain use cases and show how they can benefit from network information exposure using ALTO. Besides, key design requirements of network information exposure to support multi-domain use cases are also presented along with information about novel mechanisms and abstractions to improve the base ALTO framework in multi-domain scenarios.

Status of This Memo

This Internet-Draft will expire on January 13, 2022.

Copyright Notice

Table of Contents

1. Introduction

Many multi-domain use cases are emerging with the development of new technologies, such as software-defined networking (SDN), network function virtualization (NFV), and 5G. Examples of such use cases include multi-domain, collaborative data sciences [CMS][LCLS][LHC][SKA], multi-domain service function chaining (SFC) [NGMN-5G][SFC-MD][MD-ORCH-NFV][ETSI-ZSM], and multi-domain SDN [SFP][SDX][RFC5575]. Such use cases can benefit substantially from the exposure of network information, with which users can perform application-layer resource optimization to improve the performance.

The Application-Layer Optimization Protocol (ALTO) [RFC7285] already introduces basic mechanisms (e.g., modularity, dependency) and abstractions (e.g., map services) for applications to take optimized actions based on network information. However, exposing network information to support multi-domain use cases places additional requirements that existing solutions such as the current ALTO design do not satisfy. First, abstractions that aggregate multiple networks into a single, virtual network are required to simplify the application-layer optimization conducted by end-users. Second, such abstractions need to provide a unified representation of multiple resources (e.g., networking, computation, and storage) in multiple networks.

This document reviews the current ALTO architecture to expose network information for applications (Section 3), and it then presents several important multi-domain use cases that can benefit substantially from network information exposure (Section 4). Next, it elaborates the key design requirements of network information exposure to support these use cases (Section 5), followed by proposed extensions in the ALTO working group to satisfy the design requirements [MD-E2E-NS][UNIF-REPR] [MD-USE-CASES][BROKER-MDO][MD-ANALYTICS][MD-SFC-ALTO]. Finally, it summarizes novel mechanisms and abstractions based on recent research to improve the ALTO framework in the multi-domain set-up [MERCATOR][BOXOPT][SDI] (Section 6).

2. Changes Since Version -01

- Introduce the basic idea of the multi-domain resource abstraction obfuscation mechanism in Section 6.2.3. "Frome Single Domain to Multiple Domains".

- Fix some editorial errors in the drafts.

3. Network Information Exposure: Current ALTO Context

ALTO already provides a generic framework to expose network information for applications to improve their performance. Figure 1 presents a high-level overview of key ALTO mechanisms and abstractions.

In particular, ALTO introduces generic mechanisms such as:

- Modularity and flexibility through an explicit division of ALTO network information into (network) information resources.

- An information resource directory (IRD) providing a list of available information resources in an ALTO server.

- Information consistency (tag, dependency, multi-info resources [ALTO-MULTIPART]) to specify a dependency among different information resources.

- A generic framework with Server-Sent Events (SSEs) [ALTO-SSE] to perform stream-control, push, incremental update of information resources.

ALTO also introduces abstractions modules, such as:

- Network and cost maps to provide network location groupings and costs between them.

- A multi-cost map [RFC8189] to retrieve several cost metrics in a single query/response transaction.

- The path vector abstraction [ALTO-PATH] to provide more detailed routing information using Abstract Network Elements (ANEs).

- Capability maps (e.g., CDNI [ALTO-CDNI] and unified property Map [ALTO-PROP]).

Another generic concept introduced is filter so that information resources can be filtered (e.g., Filtered network map, filtered cost map). Besides, each individual information resource is provided as a

RESTful service with a very simple, but well-working grammar (essentially JSON grammar [RFC7159]).

Server discovery [RFC7286] and cross-domain server discovery [ALTO-XDOM] are also introduced in order to identify a topologically nearby ALTO server or ALTO servers outside of a network domain, respectively.

```
      ALTO SERVER
   +--------------------------------------------------------------------+
   |                                                                    |
   |                         ABSTRACTIONS                               |
   | +----------------------------------------------------------------+ |
   | |                                                                | |
   | |+-------------+ +-------------+ +-------------+ +-------------+ | |
   | ||             | |             | |             | | Capability  | | |
   | ||   Network   | |   Cost Map  | |    Path     | | Footprints/ | | |
   | ||     Map     | | (Calendar,  | |   Vector    | | Unified     | | |
   | ||             | |     ...)    | |             | | Properties  | | |
   | |+-------------+ +-------------+ +-------------+ +-------------+ | |
   | +----------------------------------------------------------------+ |
   |                                                                    |
   |                          MECHANISMS                                |
   | +----------------------------------------------------------------+ |
   | |                                                                | |
   | | +------------------+ +------------------+ +------------------+ | |
   | | |    Information   | |   Information    | |  Information     | | |
   | | |     Resource     | |   Consistency    | |  Update Model    | | |
   | | |     Directory    | |(Tag,Dependency)  | |(SSE,Incremental) | | |
   | | +------------------+ +------------------+ +------------------+ | |
   | +----------------------------------------------------------------+ |
   +----------------------------------^---------------------------------+
                                      |
                                      |
                                      |
       +------------------------------v------------------------------+
       |                       Infrastructure                        |
       +-------------------------------------------------------------+
```

Figure 1: High Level ALTO Architecture.

4. Multi-domain Use Cases

Multi-domain network information exposure can be beneficial in supporting multiple important use cases, including multi-domain, collaborative data sciences, multi-domain SFC, and multi-domain SDN.

4.1. Multi-domain, collaborative data sciences

Many of today's premier science experiments, such as the Large Hadron Collider (LHC) [LHC] and the Square Kilometre Array (SKA) [SKA], rely on finely-tuned workflows that coordinate geographically distributed resources (e.g., instrument, compute, storage) to enable scientific discoveries. One example is the movement of LHC data from Tier 0 (i.e., the data center at the European Organization for Nuclear Research, known as CERN) to Tier 1 (i.e., national laboratories) storage sites around the world. Another example is that the Fermilab is experimenting with moving the exascale LHC workflow to Amazon EC2 for more computation power [HEPCLOUD].

The key to supporting these distributed workflows is the ability to orchestrate multiple resources across multiple network domains to facilitate predictable workflow performance (e.g., available bandwidth, packet loss rate [ALTO-METRICS]). As such, multi-domain network information exposure is a cornerstone to enable this ability.

4.1.1. How can multi-domain resource orchestration benefit from ALTO?

One key design challenge for multi-domain resource orchestration is its resource information model. Existing design options such as resource graph and ClassAds are inadequate because they cannot simultaneously (1) allow member networks to provide accurate information on different types of resource, (2) avoid the exposure of private information of member networks such as topology, and (3) allow data analytics jobs to describe their requirements of different types of resources accurately. In contrast, ALTO is well suited for providing a generic representation that (1) allows different types of data analytics jobs to accurately describe their resource requirements and (2) allows member networks to provide accurate information on different types of resources they own and at the same time maintain their privacies.

4.1.2. Example

Consider an example of three member networks in Figure 2, where s1 and s2 are storage endpoints, and d1 and d2 are computation endpoints. Assume a data analytics job is composed of two parallel tasks T1 and T2. T1 needs dataset X as input, and T2 needs dataset Y as input.

```
                                          .-------------.
                                          | Network B   |
                   .-------------.    ingB|             |
                   | Network A   o--------|o-----d1     |
                   |            /|        '-------------'
                   |    s1\    / |
                   |       o--o  |        .-------------.
                   |    s2/    \ |        | Network C   |
                   |            \|    ingC|             |
                   |             o--------|o-----d2     |
                   '-------------'        '-------------'
```

Figure 2: Multi-domain resource orchestration.

Using the ALTO endpoint property service, an ALTO client in the resource orchestrator can discover that d1 satisfies the computing requirements of T1, and d2 satisfies the computing requirements of T2. Hence there are only two candidate endpoint pairs: (s1, d1) and (s2, d2).

Afterward, using the ALTO path vector extension, the ALTO client can retrieve the bandwidth sharing information of task T1 and T2, denoted as x1 and x2, respectively, as follows:

```
A: x1 + x2 <= 10Mbps
B:      x1 <= 3Mbps
C:      x2 <= 3Mbps
```

With such information, the resource orchestrator can make the optimal resource orchestration decision to reserve 3 Mbps bandwidth for task T1, and 3 Mbps bandwidth for task T2.

4.2. Multi-domain SFC

This use case refers to building end-to-end services by composing multiple service functions (SFs) in an abstract sequence across multiple network domains [SFC-MD]. It is identified as an important value-added service in 5G [MD-ORCH-NFV][ETSI-ZSM]. Exposing multi-domain network and resource information (e.g., link bandwidth, CPU utilization) can substantially improve the efficiency of constructing and managing such SFCs.

4.2.1. How can multi-domain SFC benefit from ALTO?

A "dialogue" between potential domains that provide multi-domain SFC could be beneficial for more efficient use of resources and increasing the SFC performance. However, constrained knowledge of the network services and underlying network topology based only on localized views from the point of view of a single domain limits the potential and scope for multi-domain SFC. ALTO (and customized ALTO extensions) can be used to offer aggregated/abstracted views on various types of information, including domain-level topology, storage resources, computation resources, networking resources, and SF capabilities. This generic representation contributes to a more simple and scalable solution for resource and service discovery in multi-domain, multi-technology environments.

4.2.2. Example

Figure 3 shows a SFC eXchange Platform (SXP), connecting three different domains (AS1, AS2, AS3). A SXP is a logical entity to make possible the negotiation between different domains, and it could be deployed, for example, in future Software-defined IXP (as a trusted third-party platform) [HH-MDSFC].

In this scenario, each domain provides different SFs: AS1 = {SF1}; AS2 = {SF2, SF3}; and AS3 = {SF3}. The SXP also includes an ALTO server component to provide abstract topology, resource, and service information for the high-level control plane in each domain [MD-SFC-ALTO].

```
                                 SFC eXchange
                                 Platform
                            +---------------------+
                            |     +--------+      |
                            |     | ALTO   |      |
                            |     | Server |      |
                            |     +--------+      |
                       +---->                    <-----+
                       |    +----------^---------+     |
                       |               |               |
                       |               |               |
                  +----v---+     +-----v--+     +------v-+
                  |Control |     |Control |     |Control |
                  |Plane   |     |Plane   |     |Plane   |
                  +--------+     +--------+     +--------+
                       |              |              |
                       |              |              |
                  +--------+     +--------+     +--------+
                  |Data    |     |Data    |     |Data    |
                  |Plane   -------Plane   -------Plane   |
                  +--------+     +--------+     +--------+
                     SF1            SF2            SF3
                                    SF3

                 [----AS1----][-----AS2-----][-----AS3-----]
```

Figure 3: Multi-domain SFC Architecture with ALTO.

The ALTO Property Map Service [DRAFT-ALTO-PM] can provide a clear global view of the resource information offered by other domains. This information allows discovering which candidate domains may be contacted to deliver the remaining requirements of a requested end-to-end service deployment. In our example, the Property Map (see Table 1) includes a property value grouped by AS. This value contains the supported SFs. Additional properties could be considered, such as resource availability (e.g., CPUs, Memory, and Storage), orchestrator entry points, etc.

	Capabilities	Entry Point	CPU	MEM	Storage	...
AS1	{SF1}	http://...	...	...	...	...
AS2	{SF2, SF3}	http://...	...	...	...	...
AS3	{SF3}	http://...	...	...	...	...

Table 1: ALTO Property Map

Once the candidate domains are discovered, it is necessary to compute multi-domain SF paths to select the SF location from those different candidate domains. The connectivity information among discovered domains can be retrieved by an ALTO Cost Map service. In our example, the Cost Map defines a path vector as an array of ASes, representing the AS-level topological distance for a given SFC request. Table 2 below shows a brief example of a service request and its multi-domain SF path response containing a list of potential domains to be traversed to deliver such service.

SFC Request	Multi-domain Service Function Path(s)
SF1->SF2->SF3	1:{AS1:SF1->AS2:SF2->AS2:SF3} 2:{AS1:SF1->AS2:SF2->AS3:SF3}

Table 2: ALTO Cost Map

4.3. Multi-domain SDN

Network providers are expanding the fine-grained capability of SDN from intradomain set-up to multi-domain settings to provide flexible interdomain routing as a valuable service [SFP][SDX][RFC5575]. Users of this service can specify routing actions at the provider network based on flexible matching conditions of flow parameters such as TCP/IP 5-tuple. This service requires provider networks to expose their available routing information to users. However, handling routing information of each network individually is too complex for users. As such, a multi-domain network exposure solution that aggregates information of multiple networks into a single abstraction can simplify the use of this service.

4.3.1. How can flexible interdomain routing benefit from ALTO?

ALTO provides provider ASes a standardized approach to expose its routing capability to client ASes. Traditional interdomain routing protocols such as BGP are not good options because they only expose the currently used routes, limiting client ASes' choices to specify flexible routes. In contrast, ALTO and its extensions provide interfaces for provider ASes to expose not only currently used routes, but also available yet unused routes, to client ASes so that they can have the flexibility to specify different routes for different data traffic.

4.3.2. Example

Consider the example in Figure 4. AS A is compromised and being used to send DDoS traffic to AS E. Without flexible interdomain routing, AS E can setup a firewall locally, but normal traffic from B to E will still be congested at C-D-E due to the existence of malicious traffic from A to E. If AS C provides flexible interdomain routing service, AS E can specify such a firewall at AS C to block DDoS traffic from A, and at the same time avoid the congestion of normal traffic from B to E.

```
        +-----+ DDoS traffic
        | AS A|_
        |     | \   +--+---+         +---+--+        +------+
        +-----+  \  |      |         |      |        |      |
                  \_| AS C +---------+ AS D |--------| AS E |
        +-----+   / |      |         |      |        |      |
        | AS B|__/  +------+         +------+        +------+
        |     |
        +-----+ Normal traffic
```

Figure 4: Flexible interdomain routing for DDoS mitigation.

5. Requirements on Multi-domain Network Information Exposure

Supporting previous use cases with multi-domain network information exposure places new, key requirements, which are not fully satisfied by existing exposure solutions. This section discusses these requirements and briefly review existing efforts in the ALTO working group aiming to satisfy them.

5.1. Design Requirements

o Unified Resource Capability Representation. Modern use cases require information on properties and capabilities of diverse in-network resources, including transport resources (e.g., available bandwidth), processing resources (e.g., SFs), and storage resources. These use cases may then conduct orchestration of multiple resources in multiple networks (e.g., RAN, transport, core in 5G). As such, a unified representation of capabilities of multiple resources is key requirement for multi-domain network information exposure to support multi-domain use cases.

- Multi-domain, easy-to-compose, end-to-end representation. Existing representations (e.g., ALTO network/cost maps, generic YANG models) tend to focus on a single domain. In multi-domain use cases, related information can be retrieved from multiple networks to compute end-to-end information. As such, abstractions that support aggregation of multiple networks into a single, virtual network ("one-big-network") are a key requirement.

- Providing interfaces for more flexible queries. With the emerging of new networking architecture (e.g., SDN and NFV) and the fine-grained resource requirement of applications (e.g., link-disjoint paths and endpoint precedence), applications need a more flexible interface to specify queries of resource information.

5.2. Existing Efforts in the ALTO Working Group

Several documents have been submitted to the ALTO working group, with the aim to satisfy one or more of the design requirements discussed above.

- [DRAFT-RSA][DRAFT-UNICORN-INFO] documents propose and apply the ALTO path vector extension to provide accurate networking resource information to support multi-domain resource orchestration.

- [BROKER-MDO] proposes to use ALTO to support resource orchestration for multi-domain SFC, and proposes a new ALTO extension to retrieve AS path of network functions across different networks.

- [DRAFT-CONTEXT] proposes to extend cost information specified in [RFC7285] by providing several possible cost values for the same cost metric where each value depends on qualitative criteria as opposed to quantitative criteria such as time.

- [UNIF-REPR] makes a proposal to use mathematical programming constraint as a generic representation of multiple resources.

- [DRAFT-FCS] proposes a flexible flow query extension service to allow applications to specify query entities based on flexible matching conditions (e.g., TCP/IP 5-tuple) instead of IP addresses only.

6. Novel Multi-domain Mechanisms and Abstractions

This section presents novel mechanisms and abstractions based on recent research to allow the ALTO framework supporting the important multi-domain settings.

6.1. Multi-domain Aggregation

Figure 5 shows a generic aggregation framework of using ALTO in multi-domain use cases [BROKER-MDO][MERCATOR][BOXOPT]. It includes a multi-domain aggregation mechanism on top of the existing single domain architecture. This new mechanism aggregates network information from ALTO servers in multiple networks to provide a single, consistent, updated, "virtual" domain abstraction. Network maps, cost maps, unified entity properties, network capabilities, and routing path abstractions (path vectors) of individual networks are consistently integrated to provide the abstraction of a single, coherent network to the users, satisfying the multi-domain aggregation requirement.

```
                                    +-------------+
                                    | Application |
                                    +------|------+
                                           |(1)
                                           |
             +---------------+ (2) +-------v-------+
          (3)|               <-----+  Multi-domain |
        xxxxx>  AGGREGATOR   |     |  Orchestrator |
        x    |               +----->               |
        x    +-----------^---^ (5) +-------|-------+
        x                x   x       -/    |(6)  \--
        x                x   x     -/      |        \--
        x                x   x   -/        |           \-
        x                x   xxxxxxxxxxxxxxxxxxxxxxxxxx  \--
        x                x   -/            |          x     \--
        x(4)             x -/              |          x        \--
   +----v---+ +------+   x/ +--------+ +---v--+      +v-------+ +-\>---+
   | ALTO   | | Dom  </-/x  | ALTO   | | Dom  |      | ALTO   | | Dom  |
   | Server | | Orch |   xxx> Server | | Orch |      | Server | | Orch |
   +--------+ +------+      +--------+ +------+      +--------+ +------+
   +-----------------+      +-----------------+      +-----------------+
   | Infrastructure  |      | Infrastructure  |      | Infrastructure  |
   +-----------------+      +-----------------+      +-----------------+

        Network 1                Network 2                Network 3
```

Figure 5: Generic framework of using ALTO in multi-domain use cases.

6.1.1. Workflow

The basic workflow of this framework is as follows:

- o A user (e.g., an application) submits a network service request to the Multi-domain Orchestrator (MdO).

- o The MdO receives the network service and queries the aggregator to discover multi-resource network information (e.g., bandwidth, delay, and loss rate).

- o The aggregator determines the member networks that the network service will traverse, and queries the ALTO server in each of these member networks to discover their resource information.

- o Upon receiving the query from the aggregator, each ALTO server computes the resource abstraction of the corresponding member network, and send the network resource information to the aggregator.

- o The aggregator collects the resource abstractions from the relevant member networks, and derives this aggregated, unified resource information to the MdO.

- o Based on the received information, the MdO determines the resource allocation for the network service, and sends a reservation request to the Domain Orchestrators (DOs).

6.2. Multi-resource Abstraction

Although the existing abstractions (network/cost map, unified property, and path vector) are already powerful, they cannot handle the multi-resource information requirement. To this end, this document presents a recent, unified resource abstraction, based on mathematical programming constraints as a generic representation of the feasible resource capability of networks [MERCATOR], which users can consume via different resource management systems.

6.2.1. Mathematical Programming Constraints: Basic Idea

Consider a network as shown in Figure 6. Assume that the bandwidth of every link is 100 Gbps. An application (e.g., a large data analytics system) wants to reserve two circuits from S1 to D1 and S2 to D2, respectively. Assume the application is only interested in the bandwidth of both circuits. The route of f1 (S1 -> D1) is S1 -> sw1 -> sw5 -> sw6 -> sw8 -> sw3 -> D1, and the route of f2 (S2 -> D2) is S2 -> sw2 -> sw5 -> sw6 -> sw8 -> sw4 -> D2. The routes for the two circuits share common links l3 and l4, making it infeasible for both circuits to each reserve a 100 Gbps bandwidth.

```
              +-----+ l2      l3 +-----+  l4      l5 +-----+
    +--+  l1  |     -----+  +-----     |-----+  +-----     |  l6  +--+
    |S1-------- SW1 |    |  |    | SW6 |     |  |    | SW3 -------|D1|
    +--+      +-----+  +-|--|+   +-----+   +-|--|+   +-----+      +--+
                       |     |             |     |
                       | SW5 |             | SW8 |
              +-----+  +-|--|+   +-----+   +-|--|+   +-----+
    +--+  l7  |     |    |  |    |     |     |  |    |     | l12  +--+
    |S2-------- SW2 -----+  +----- SW7 |-----+  +----- SW4 -------|D2|
    +--+      +-----+ l8      l9 +-----+ l10     l11 +-----+      +--+
```

Figure 6: Network Topology Example.

The use of mathematical programming constraints is to generate a set of linear inequalities to provide a compact representation for different important properties of network resources (e.g., bandwidth, delay, loss rate). In the previous example, the following set of linear inequalities are generated:

```
            f1 <= 100, for {l1, l2, l5, l6} ..... (le1)
            f2 <= 100, for {l7, l8, l11, l12} ... (le2)
       f1 + f2 <= 100, for {l3, l4} ............. (le3)
```

which accurately captures the bandwidth sharing among two circuits' routes.

6.2.2. Removing Redundant Linear Inequalities

Taking a deeper look at the set of previous linear inequalities, one can conclude that inequalities of le1 and le2 can be implicitly derived from that of le3. Thus, these inequalities are considered redundant. The problem of finding redundant linear constraints has been widely studied [PAULRAJ10COMP]. Specifically, redundant linear inequalities are removed via a polynomial-time, optimal algorithm [KARMARKAR84]. In our example, the compressed set will only contain one inequality: f1 + f2 <= 100, for {l3, l4} (le3).

6.2.3. From Single Domain to Multiple Domains

To illustrate the basic aggregation abstraction from a single network to multiple networks, consider a collaboration network composed of three member networks, as shown in Figure 7. A user wants to reserve bandwidth for three circuits, from source host S to destination hosts D1 , D2, and D3.

```
       Member                    Member                     Member
      Network 1                 Network 2                  Network 3
   +xxxxxxxxxxxxxx+  +xxxxxxxxxxxxxxxxxxxxxxxxxxxx+  +xxxxxxxxxxxxxxxxx+
   x              x  x                            x  x                 x
   x +---+        x  x +---+         +---+        x  x +---+           x
   x |   |100Gbps x  x |   | 40Gbps  |   |        x  x |   |           x
   x |   --------------|   -----------   --------------|   ---+        x
   x |   |        x  x |   |         |   |100Gbps x  x |   |  |        x
   x +|--+        x  x +|--+         +---+        x  x +---+  |10Gbps  x
   x  |           x  x  |                         x  x        |        x
   x  |40Gbps     x  x  |                         x  x      +-|-+      x
   x  |           x  x  |             +---+       x  x      |   |      x
   x +|--+        x  x  |    10Gbps   |   |       x  x  +----   -----+ x
   x |   |        x  x  +--------------   |       x  x  |   |   |    | x
   x |   |        x  x                |   |       x  x  |   +---+    | x
   x +|--+        x  x                +|--+       x  x  |      10Gbps| x
   x  |           x  x                 |          x  x  |            | x
   x  |100Gbps    x  x           10Gbps|          x  x  |10Gbps      | x
   x  |           x  x                 |          x  x  |            | x
   +xx|xxxxxxxxxxx+  +xxxxxxxxxxxxxxxxx|xxxxxxxxxx+  +xx|xxxxxxxxxxxx|x+
      |                                |                |            |
     +|-+                            +-|+              +--+         +--+
     |S |                            |D1|              |D2|         |D3|
     +--+                            +--+              +--+         +--+
```

Figure 7: A collaboration network composed of three member networks.

The resource abstraction captures the constraints from all networks using the set of linear inequalities, as depicted in Figure 8. Specifically, the variables f1, f2, f3 represent the available bandwidth that can be reserved for (S, D1), (S, D2), and (S, D3), respectively.

Network	Linear Inequalities
Member Network 1	f1 + f2 + f3 <= 100 .... (le11) f1 + f2 + f3 <= 40 .... (le12) f1 + f2 + f3 <= 100 .... (le13)
Member Network 2	f2 + f3 <= 40, f1 <= 10 ... (le21) f2 + f3 <= 100, f1 <= 10 .. (le21)
Member Network 3	f2 + f3 <= 10 .... (le31) f2 <= 10 .... (le32) f3 <= 100 .... (le33)

Figure 8: Resource abstraction for the reservation request from Figure 4.

Each linear inequality represents a constraint on the reservable bandwidths over different shared resources by the three circuits. For example, the inequality le11 indicates that all three circuits share a common resource and that the sum of their bandwidths can not exceed 100 Gbps.

After removing the redundant inequalities of each member network, the resource abstraction of each member network is:

Network	Linear Inequalities
Network Member 1	f1 + f2 + f3 <= 40 .... (le12)
Network Member 2	f2 + f3 <= 40, f1 <= 10 ... (le21)
Network Member 3	f2 + f3 <= 10 .... (le31)

Figure 9: Resource abstraction of each member network after removing the redundant inequalities.

Although each domain may already conduct redundancy optimization, there can be cross-domain redundancy. For example, the constraint le31 at member network 3 (f2 + f3 <= 10) can eliminate those at member network 1 (f1 + f2 + f3 <= 40) and member network 2 (f2 + f3 <= 40). Using a classic compression algorithm [TELGEN83], we can remove this cross domain redundancy. Therefore, the compressed multi-domain set of linear inequalities will contain:

```
                          f1 <= 10,
                     f2 + f3 <= 10
```

In implementation, to protect the privacy of different networks, an obfuscation mechanism developed in [MERCATOR] can be adopted to prevent the applications from associating the received resource abstractions with the corresponding member networks. The key idea is to have each ALTO server obfuscate its own resource abstraction as a set of linear equations through a private random matrix of its own and a couple of random matrices shared with few other ALTO servers from other networks, and then send the obfuscated resource abstraction back to the application. From the received resource abstractions, the application can retrieve the actual resource information across multiple networks, but cannot associate any linear inequality with its corresponding network.

6.3. Multi-domain Programming Information Abstraction

Multi-domain SDN requires multi-domain SDN resource and programming abstraction. A novel multi-domain network programming framework is recently proposed to achieve programmable, end-to-end interdomain route control. Specifically, in this framework, each autonomus system (AS) deploys an ALTO server to provide a programming abstraction. Collectively, these ALTO servers provide the client a single, abstract, programmable network spanning multiple individual networks. Unlike existing SDN abstractions (e.g., OpenFlow and P4), it includes a built-in layer to extract and learn the interactions of interdomain policies of individual networks (e.g., route selection preference), providing a unified abstraction.

Specifically, given a destination IP prefix p specified by a client, each autonomous system (AS) exposes its routing information base (RIB), i.e., all available routes it has to reach p, and the corresponding price for the client to use each route, by deploying an ALTO server. A client queries the available routes and the corresponding prices at different ASes. With the retrieved information, it then itereatively selects the best route based on its internal objective function and budget, and interacts with different ASes to check if the selected route is policy compliant, i.e., whether all ASes along this route will export the preceding segment to its upstream neighbor. After finding the best policy compliant route, the client then interacts with ASes of the route in a backward order along the route to setup the AS path. A blackbox based optimization algorithm has been developed to allow the client to swiftly find the best policy compliant route, and details can be find in [SDI].

7. IANA Considerations

This document includes no request to IANA.

8. Security Considerations

TBD.

9. References

9.1. Normative References

[RFC5575] Marques, P., Sheth, N., Raszuk, R., Greene, B., Mauch, J., and D. McPherson, "Dissemination of Flow Specification Rules", RFC 5575, DOI 10.17487/RFC5575, August 2009, <https://www.rfc-editor.org/info/rfc5575>.

[RFC7159] Bray, T., Ed., "The JavaScript Object Notation (JSON) Data Interchange Format", RFC 7159, DOI 10.17487/RFC7159, March 2014, <https://www.rfc-editor.org/info/rfc7159>.

[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>.

[RFC7286] Kiesel, S., Stiemerling, M., Schwan, N., Scharf, M., and H. Song, "Application-Layer Traffic Optimization (ALTO) Server Discovery", RFC 7286, DOI 10.17487/RFC7286, November 2014, <https://www.rfc-editor.org/info/rfc7286>.

[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>.

9.2. Informative References

[ALTO-CDNI]
Seedorf, J., Yang, Y., Ma, K., Peterson, J., Lin, X., and J. Zhang, "Content Delivery Network Interconnection (CDNI) Request Routing: CDNI Footprint and Capabilities Advertisement using ALTO", draft-ietf-alto-cdni-request-routing-alto-08 (work in progress), November 2019.

[ALTO-METRICS]
WU, Q., Yang, Y., Lee, Y., Dhody, D., and S. Randriamasy, "ALTO Performance Cost Metrics", draft-ietf-alto-performance-metrics-08 (work in progress), November 2019.

[ALTO-MULTIPART]
Zhang, J., "Multiple ALTO Resources Query Using Multipart Message", draft-zhang-alto-multipart-02 (work in progress), March 2019.

[ALTO-PATH]
Gao, K., Lee, Y., Randriamasy, S., Yang, Y., and J. Zhang, "ALTO Extension: Path Vector", draft-ietf-alto-path-vector-09 (work in progress), November 2019.

[ALTO-PROP]
Roome, W., Randriamasy, S., Yang, Y., Zhang, J., and K. Gao, "Unified Properties for the ALTO Protocol", draft-ietf-alto-unified-props-new-10 (work in progress), November 2019.

[ALTO-SSE]
Roome, W. and Y. Yang, "ALTO Incremental Updates Using Server-Sent Events (SSE)", draft-ietf-alto-incr-update-sse-17 (work in progress), July 2019.

[ALTO-XDOM]
Kiesel, S. and M. Stiemerling, "Application Layer Traffic Optimization (ALTO) Cross-Domain Server Discovery", draft-ietf-alto-xdom-disc-06 (work in progress), August 2019.

[BOXOPT] Xiang, Q., Zhang, J., Wang, T., Liu, J., Guok, C., Le, F., MacAuley, J., Newman, H., and R. Yang, "Optimizing in the Dark: Learning an Optimal Solution Through a Simple Interface", Conference AAAI 2019, 2019.

[BROKER-MDO]
Perez, D. and C. Rothenberg, "ALTO-based Broker-assisted Multi-domain Orchestration", draft-lachosrothenberg-alto-brokermdo-02 (work in progress), March 2019.

[CMS] The CMS Collaboration, "The CMS experiment at the CERN LHC", 2008, <https://doi.org/10.1088/1748-0221/3/08/S08004>.

[DRAFT-CONTEXT]
Randriamasy, S., "ALTO Contextual Cost Values", draft-randriamasy-alto-cost-context-02 (work in progress), July 2017.

[DRAFT-FCS]
Zhang, J., Gao, K., Wang, J., Xiang, Q., and Y. Yang, "ALTO Extension: Flow-based Cost Query", draft-gao-alto-fcs-06 (work in progress), July 2018.

[DRAFT-RSA]
Gao, K., xinwang2014@hotmail.com, x., Xiang, Q., Gu, C., Yang, Y., and G. Chen, "Compressing ALTO Path Vectors", draft-gao-alto-routing-state-abstraction-08 (work in progress), March 2018.

[DRAFT-UNICORN-INFO]
Xiang, Q., Newman, H., Bernstein, G., duhaizhou@gmail.com, d., Gao, K., Mughal, A., Balcas, J., Zhang, J., and Y. Yang, "Resource Orchestration for Multi-Domain Data Analytics", draft-xiang-alto-exascale-network-optimization-03 (work in progress), July 2017.

[ETSI-ZSM]
ETSI, "Zero Touch Network and Service Management", 2020, <https://www.etsi.org/technologies/zero-touch-network-service-management>.

[HEPCLOUD]
Holzman, B., Bauerdick, L., Bockelman, B., Dykstra, D., Fisk, I., Fuess, S., Garzoglio, G., Girone, M., Gutsche, O., and D. Hufnagel, "Hepcloud, a new paradigm for hep facilities: Cms amazon web services investigation", Journal Computing and Software for Big Science, Volume 1, Number 1, Pages 1, Publisher Springer, 2017.

[HH-MDSFC]
Li, G., Li, G., Xu, Q., Zhou, H., Feng, B., Feng, X., and Z. Yu, "Hybrid Hierarchical Multi-Domain Service Function chaining", draft-li-sfc-hhsfc-07 (work in progress), September 2019.

[KARMARKAR84]
Karmarkar, N., "A new polynomial-time algorithm for linear programming", Book Title Proceedings of the sixteenth annual ACM symposium on Theory of computing, Pages 302--311, 1984.

[LCLS] SLAC National Accelerator Laboratory, "The Linac Coherent Light Source", 2020, <https://lcls.slac.stanford.edu/>.

[LHC] CERN: European Council for Nuclear Research, "The Large Hadron Collider (LHC) Experiment", 2020, <https://home.cern/topics/large-hadron-collider>.

[MD-ANALYTICS]
Xiang, Q., Le, F., Yang, Y., Newman, H., and H. Du, "Unicorn: Resource Orchestration for Multi-Domain, Geo-Distributed Data Analytics", draft-xiang-alto-multidomain-analytics-02 (work in progress), July 2018.

[MD-E2E-NS]
Perez, D. and C. Rothenberg, "Multi-domain E2E Network Services", draft-lachosrothenberg-alto-md-e2e-ns-00 (work in progress), March 2019.

[MD-ORCH-NFV]
Katsalis, K., Nikaein, N., and A. Edmonds, "Multi-domain orchestration for NFV: Challenges and research directions", focus 189--195, 2016.

[MD-SFC-ALTO]
Perez, D., Xiang, Q., Rothenberg, C., and Y. Yang, "Multi-domain Service Function Chaining with ALTO", draft-lachos-multi-domain-sfc-alto-00 (work in progress), July 2018.

[MD-USE-CASES]
Xiang, Q., Le, F., and Y. Yang, "ALTO for Multi-Domain Applications: A Review of Use Cases and Design Requirements", draft-xiang-alto-multidomain-usecases-00 (work in progress), March 2019.

[MERCATOR]
Xiang, Q., Zhang, J., Wang, T., Liu, J., Guok, C., Le, F., MacAuley, J., Newman, H., and R. Yang, "Fine-grained, multi-domain network resource abstraction as a fundamental primitive to enable high-performance, collaborative data sciences", focus 58--70, 2018.

[NGMN-5G] Alliance, NGMN, "5G White Paper", 2015, <https://www.skatelescope.org/>.

[PAULRAJ10COMP]
Paulraj, S. and P. Sumathi, "A comparative study of redundant constraints identification methods in linear programming problems", Journal Mathematical Problems in Engineering, Volume 2010, Publisher Hindawi Limited, 2010.

[SDI] Xiang, Q., Zhang, J., Gao, K., Lim, Y., Le, F., Li, G., and R. Yang, "Toward Optimal Software-Defined Interdomain Routing", Conference 39th Annual IEEE International Conference on Computer Communications (INFOCOM'20), 2020.

[SDX] Gupta, A., Vanbever, L., Shahbaz, M., Donovan, S., Schlinker, B., Feamster, N., Rexford, J., Shenker, S., Clark, R., and E. Katz-Bassett, "Sdx: A software defined internet exchange", focus 551--562, 2015.

[SFC-MD] Sun, G., Li, Y., Liao, D., and V. Chang, "Service function chain orchestration across multiple domains: A full mesh aggregation approach", Journal IEEE Transactions on Network and Service Management, Volumen 15, Number 3, Pages 1175--1191, Publisher IEEE, 2018.

[SFP] Xiang, Q., Guok, C., Le, F., MacAuley, J., Newman, H., and R. Yang, "SFP: Toward Interdomain Routing for SDN Networks", focus 87--89, 2018.

[SKA] SKA Organisation, "The Square Kilometre Array", 2020, <https://www.skatelescope.org/>.

[TELGEN83]
Telgen, J., "Identifying redundant constraints and implicit equalities in systems of linear constraints", Journal Management Science, Volume 29, Number 10, Pages 1209--1222, Publisher INFORMS, 1983.

[UNIF-REPR]
Xiang, Q., Le, F., and Y. Yang, "ALTO Extension: Unified Resource Representation", draft-xiang-alto-unified-representation-01 (work in progress), March 2019.

Authors' Addresses

Danny Alex Lachos Perez
University of Campinas
Av. Albert Einstein 400
Campinas, Sao Paulo 13083-970
Brazil

Email: dlachosp@dca.fee.unicamp.br
URI: https://intrig.dca.fee.unicamp.br/danny-lachos/

Christian Esteve Rothenberg
University of Campinas
Av. Albert Einstein 400
Campinas, Sao Paulo 13083-970
Brazil

Email: chesteve@dca.fee.unicamp.br
URI: https://intrig.dca.fee.unicamp.br/christian/

Qiao Xiang
Xiamen University
School of Informatics
Xiamen, Fujian Province
China

Email: xiangq27@gmail.com

Y. Richard Yang
Yale University
51 Prospect St
New Haven, CT
USA

Email: yang.r.yang@gmail.com

Borje Ohlman
Ericsson Research
S-16480 Stockholm
Sweden

Email: Borje.Ohlman@ericsson.com

Sabine Randriamasy
Nokia Bell Labs
Route de Villejust
NOZAY 91460
FRANCE

Email: Sabine.Randriamasy@nokia-bell-labs.com

Farni Boten
Sprint
USA

Email: farni.weaver@sprint.com

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

Email: luismiguel.contrerasmurillo@telefonica.com
URI: http://lmcontreras.com/

Jingxuan Jensen Zhang
Tongji University
4800 Caoan Road
Shanghai 201804
China

Email: jingxuan.n.zhang@gmail.com

Kai Gao
Sichuan University
Chengdu 610000
China

Email: kaigao@scu.edu.cn