ANRW'19 - July 22, 2019
Montreal, Quebec, Canada

Supporting Multi-domain Use Cases with ALTO

Danny A. Lachos[*] Christian E. Rothenberg[*] Qiao Xiang[‡] Y. Richard Yang[‡]
Börje Ohlman[#] Sabine Randriamasy[§] Farni Boten[&] Luis M. Contreras[¶]

[*]Unicamp [‡]Yale [#]Ericsson [§]Nokia Bell Labs [&]Sprint [¶]Telefónica

Outline

- **Introduction**
  - Base ALTO Protocol
  - Extending Base ALTO Protocol
- **Multi-domain Use Cases**
  - ALTO in Multi-domain
  - Use Cases
- **Requirements on Network Information Exposure**
- **Multi-domain Abstractions**
- **Conclusion**

Introduction

1. Base ALTO Protocol

ALTO RFC7285

- Basic function: provides network information to applications for better network resource consumption
  - While improving application performance.
- Network information is exposed as abstract maps
  - Network map, Cost Map, etc.
- Benefits of abstract maps include
  - Protection of information privacy
  - Improved scalability
- Typical use cases:
  - P2P applications
  - Datacenter Networks, CDN, etc

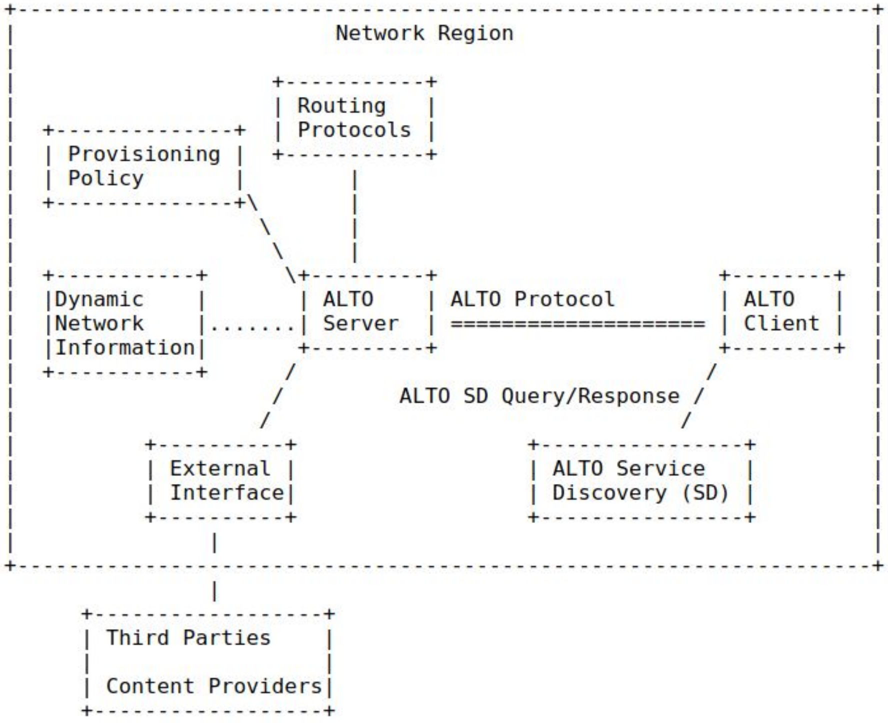
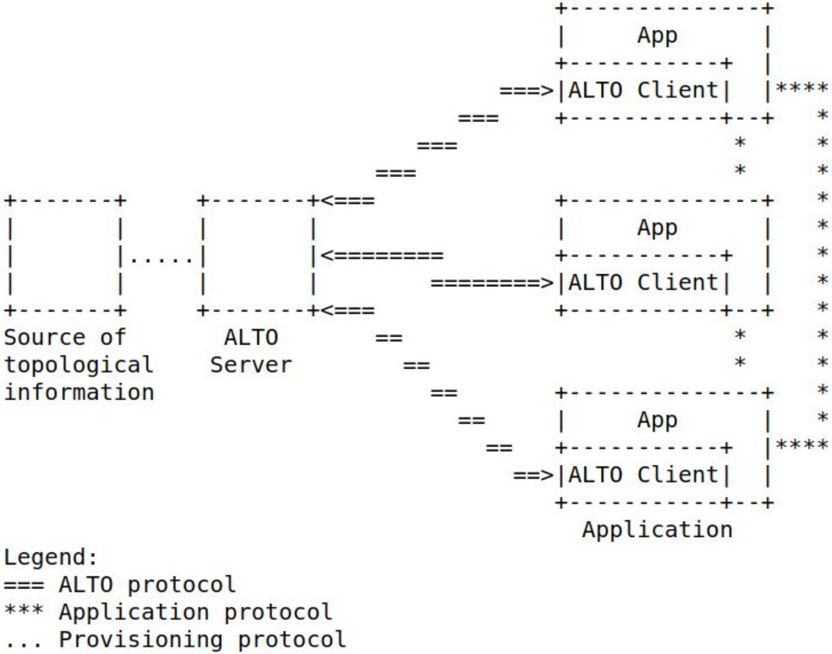


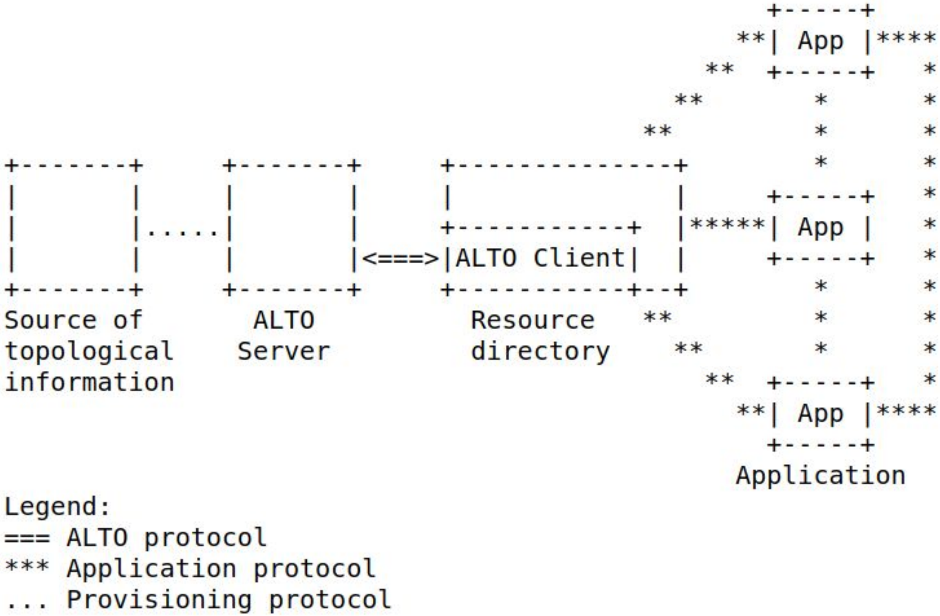
Figure 1: Basic ALTO Architecture

Source: https://tools.ietf.org/html/rfc7285

Placement of ALTO Entities



Application **without** tracker



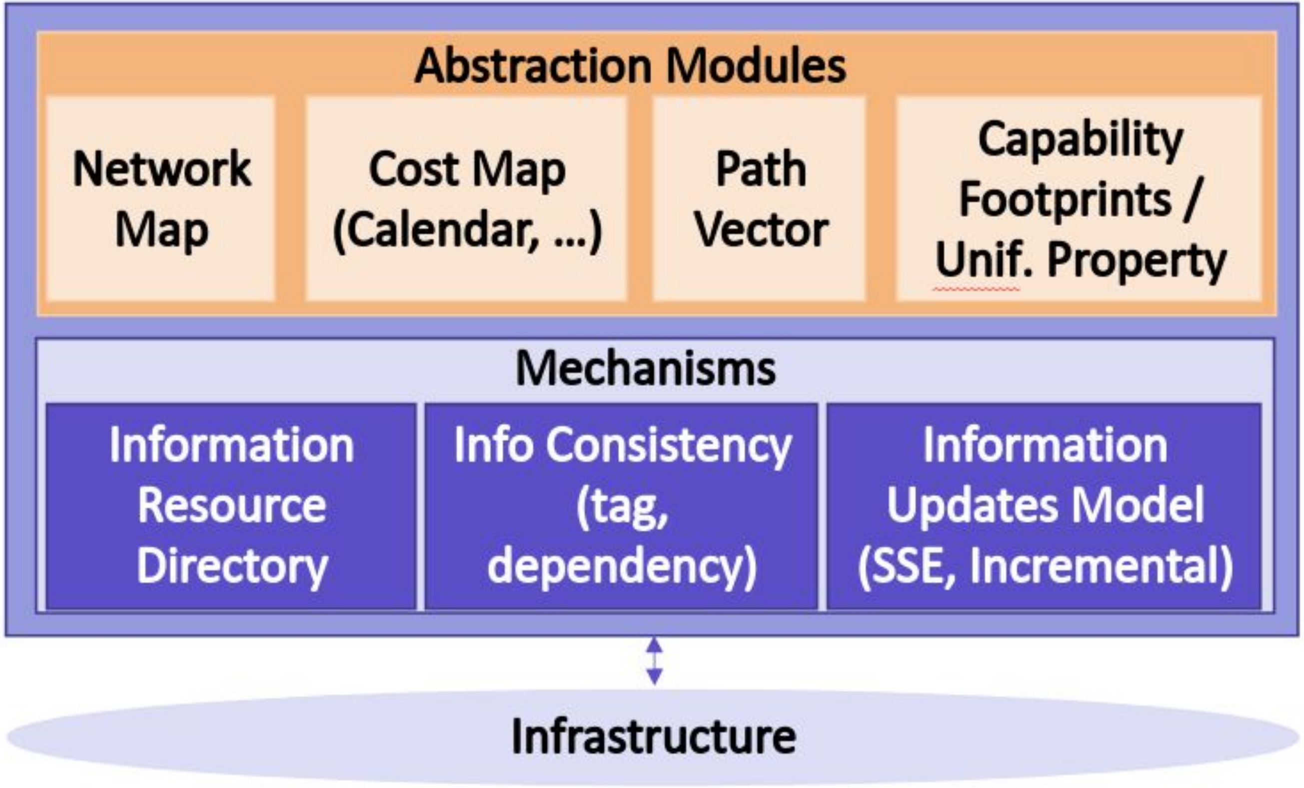
Application **with** tracker

Source: https://tools.ietf.org/html/rfc7971

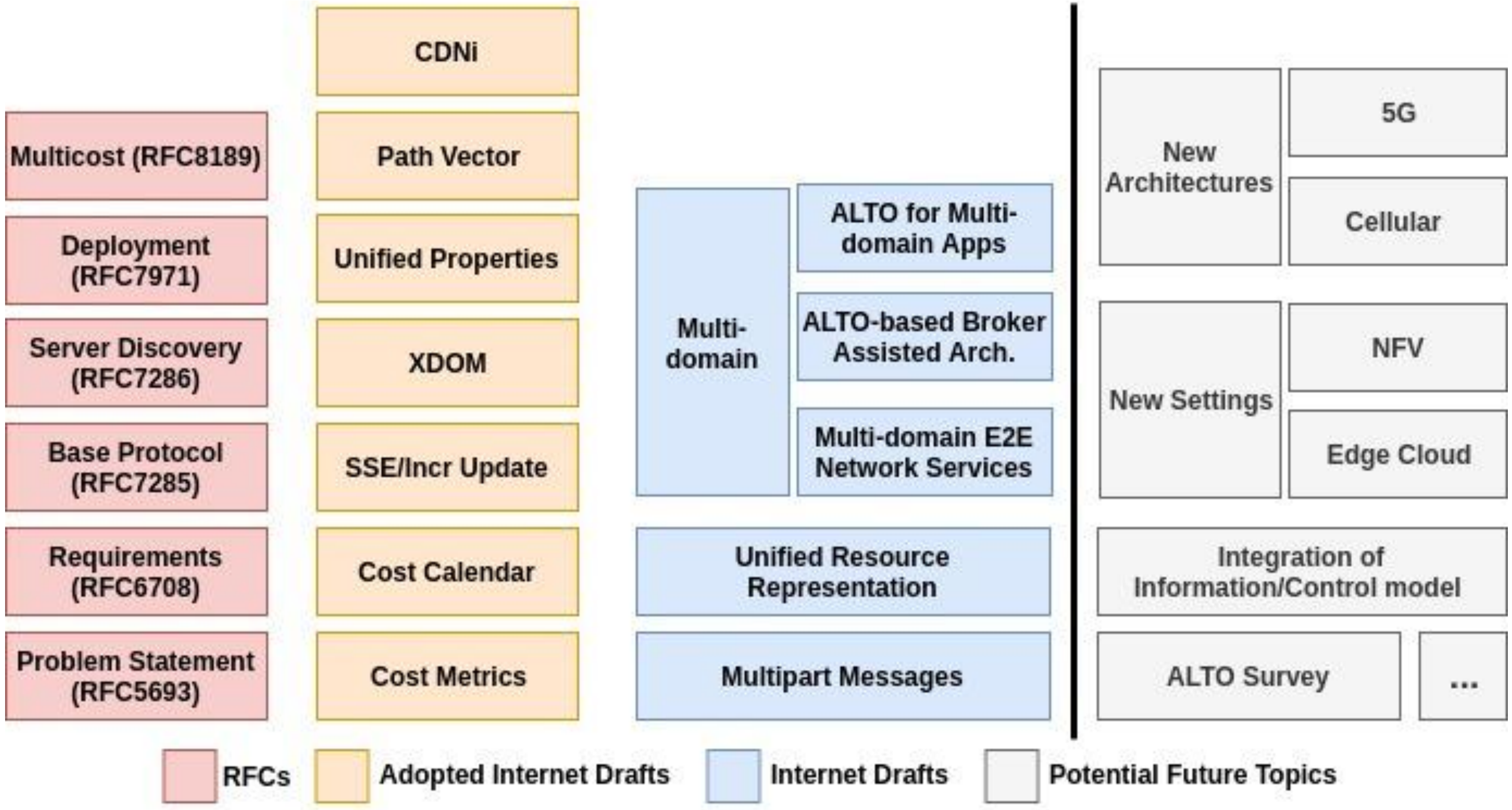
Introduction

2. Extending Base ALTO

High Level ALTO Architecture



Existing RFCs/WG Docs/Drafts



Source: https://datatracker.ietf.org/wg/alto/documents/

Multi-domain Use Cases

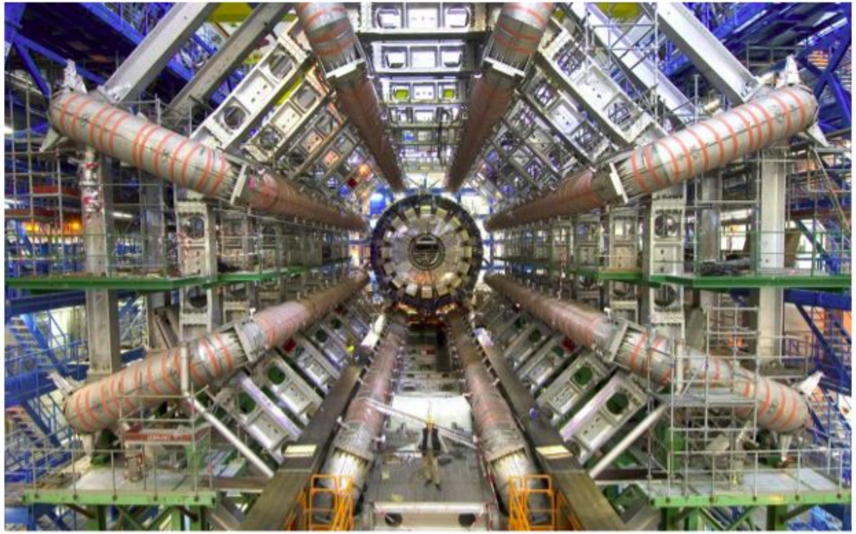
ALTO in Multi-domain

- Driven by new technologies, such as SDN, NFV, and 5G
- Driven by new use cases, such as multi-domain, collaborative data sciences, multi-domain SFC, and flexible inter-domain routing control.
- Details see individual drafts summarizing the experiences on developing multi-domain applications using ALTO
  - **Multi-domain, collaborative data sciences**
    - Draft-xiang-alto-multidomain-analytics
    - draft-xiang-alto-exascale-network-optimization
  - **Multi-domain e2e network service deployment**
    - Draft-lachosrothenberg-alto-md-e2e-ns
    - draft-lachosrothenberg-alto-brokermdo
  - **Flexible interdomain routing control**
  - **...**

Source: https://datatracker.ietf.org/meeting/104/materials/slides-104-alto-alto-for-multi-domain-applications-use-cases-and-design-requirements-01

Multi-domain, Collaborative Data Sciences

Large Hadron Collider (LHC)

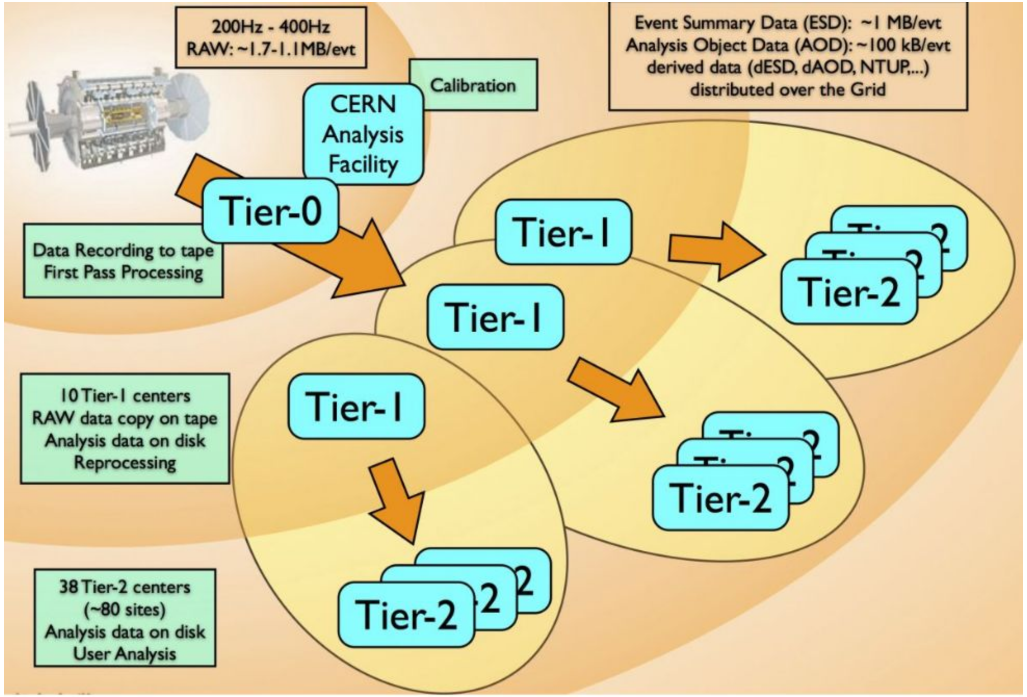


Square Kilometre Array (SKA)

- LHC and SKA push scientific discovery boundaries and rely on workflows that coordinate geographically distributed resources (e.g., compute, storage)

Figure sources: phys.org, extremetech.com

LHC Detail

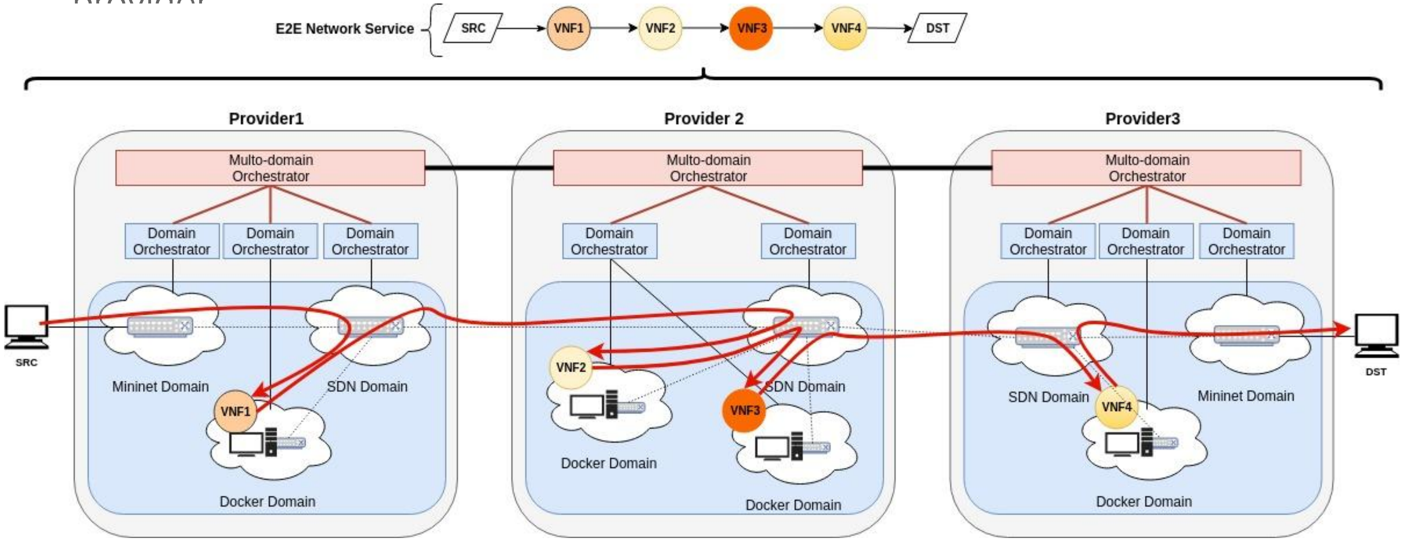


- **Multiple domains:** Tier-0 (CERN), Tier-1 (large computer centres), Tier-2 (Universities).
- **Resources:** Different domains provide heterogeneous resources (e.g., instrument, compute, storage).
- **Heterogeneous applications/jobs:**
  - Exascale dataset transfers
  - MapReduce/MPI analytics
- **REQUIREMENT:** Ability to orchestrate multiple resources across multiple domains for heterogeneous applications.

Source: https://datatracker.ietf.org/meeting/98/materials/slides-98-alto-traffic-optimization-for-exascale-science-applications-02

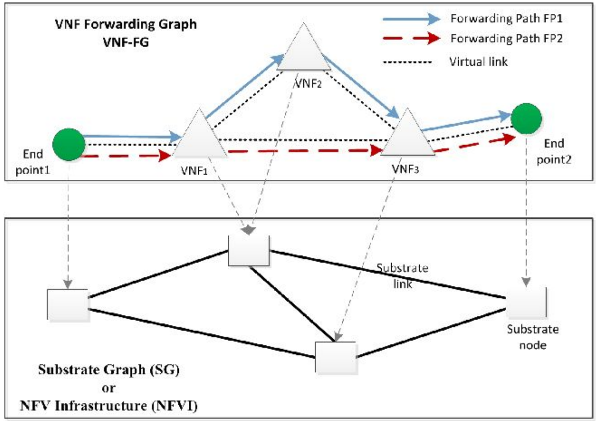
Multi-domain SFC

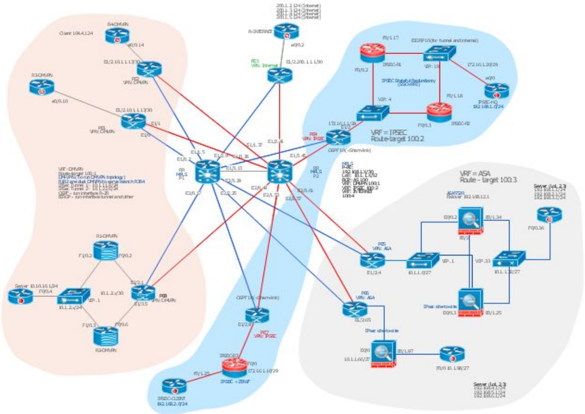
- E2E network services often require **VNFs in a specific order** [RFC7665].
  - Network services with specific requirements in terms of **resources** (e.g., cpu, memory, hard-disk) and **performance** objectives (e.g., bandwidth, latency).
  - Resources are expected to be available across **multiple domains** with different:
    - **Technology domain**: e.g., Docker domain, SDN domain, Legacy domain, etc.
    - **Administration domain**: e.g., mobile operator, cloud service provider, transport network provider

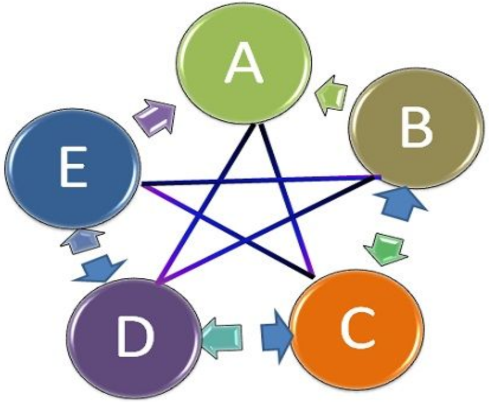


Source: https://datatracker.ietf.org/meeting/104/materials/slides-104-alto-multi-domain-e2e-network-services-00

Multi-domain SFC: Detail







Placement Decisions

Network Inventory

Publishing Information

Source: https://datatracker.ietf.org/meeting/104/materials/slides-104-alto-multi-domain-e2e-network-services-00

Multi-domain SDN

- Network providers are expanding the fine-grained capability of SDN:
  - From intradomain set-up to multi-domain setting to provide flexible interdomain routing as a valuable service.
  - Use cases: DDoS, congestion mitigation, inbound traffic control, ...
- Traditional interdomain routing protocols (e.g., BGP) are limited
  - E.g., ,single path routing, limiting client's path choices
- Flexible, multi-domain routing allows users to specify routing actions at provider networks, with
  - More flexible matching conditions (e.g. , match on TCP/IP 5-tuple).
  - More choices on routes
    - in contrast with coarse-grained protocols, provider networks can expose not only currently used routes, but also available yet unused routes
    - requires the exposure of network's routing capability.

Source: https://datatracker.ietf.org/meeting/104/materials/slides-104-alto-alto-for-multi-domain-applications-use-cases-and-design-requirements-01

Requirements on Multi-domain Network Information Exposure

Requirements for ALTO to Support New Multi-domain Cases

- **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 supports aggregation of multiple networks into a single, virtual network (“one-big-network") are a key requirement.

Multi-domain Abstraction for Application Performance Optimization

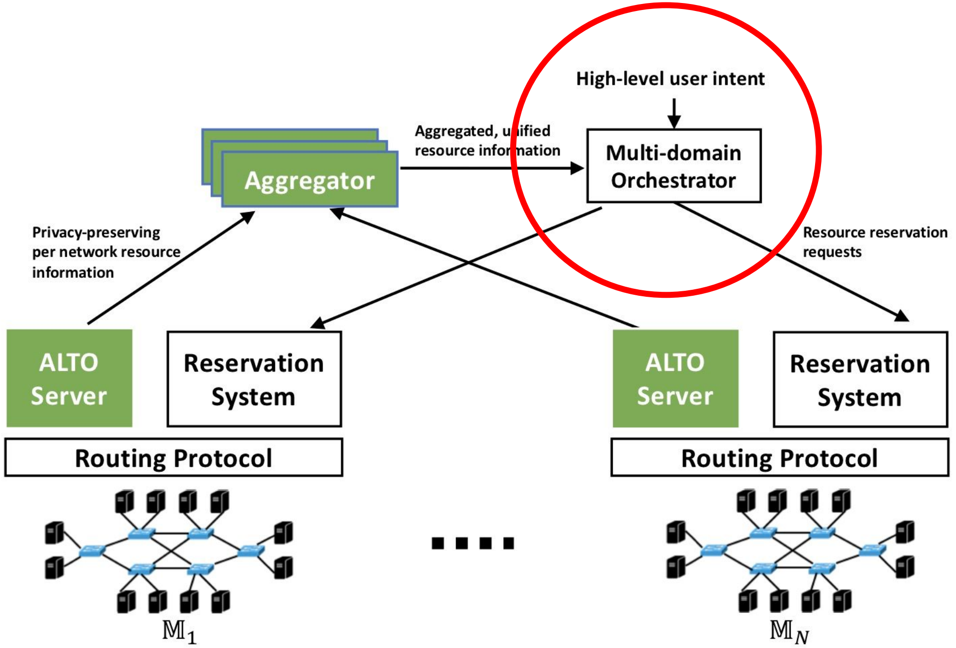
Multi-domain Optimization

- Consider an application (geo-distributed data analytics, etc.) that orchestrates large data transfers
- Typically can be modeled as an optimization problem:

optimize $f(\mathbf{x}; \mathbf{y})$

Variables representing network parameters

Variables representing application parameters



Source: Xiang, Qiao, et al. "Fine-grained, multi-domain network resource abstraction as a fundamental primitive to enable high-performance, collaborative data sciences." ACM/IEEE Supercomputing 2018.

Multi-domain Optimization

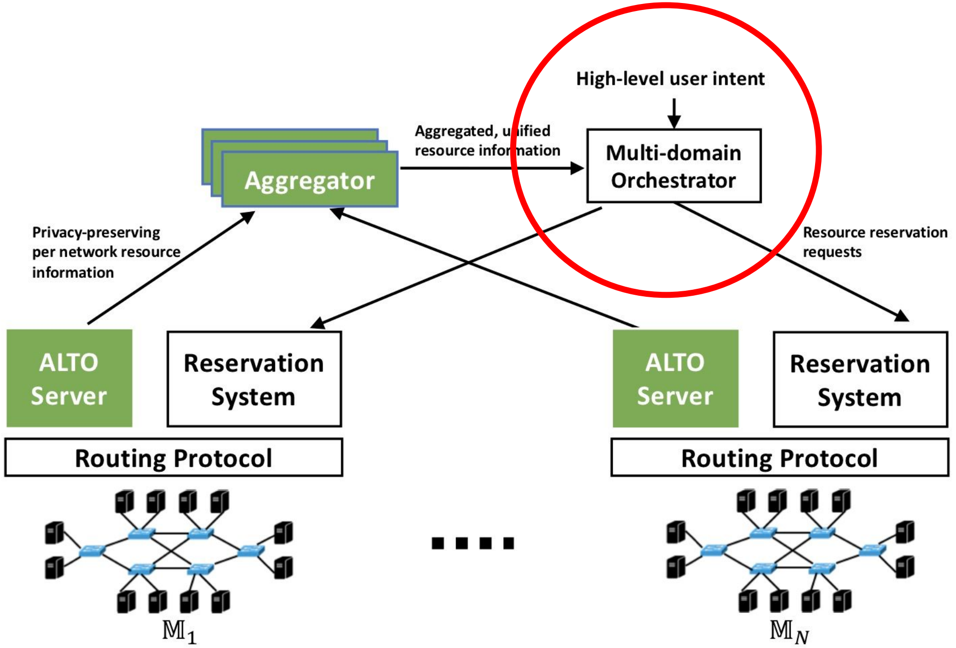
- Consider an application (geo-distributed data analytics, etc.) that orchestrates large data transfers
- Typically can be modeled as an optimization problem:

optimize $f(\mathbf{x}; \mathbf{y})$

x subject to network constraints

$\mathbf{x} \in \Omega$

Networks limit potential values of x (e.g., bw, delay, loss rate)



Source: Xiang, Qiao, et al. "Fine-grained, multi-domain network resource abstraction as a fundamental primitive to enable high-performance, collaborative data sciences." ACM/IEEE Supercomputing 2018.

Multi-domain Optimization

- Consider an application (geo-distributed data analytics, etc.) that orchestrates large data transfers
- Typically can be modeled as an optimization problem:

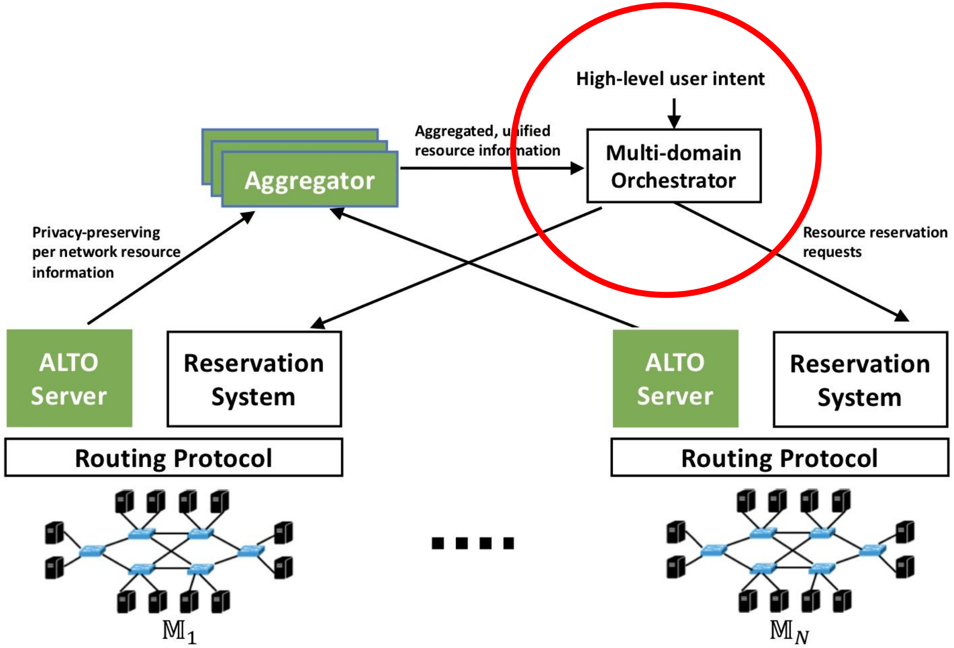
optimize $f(\mathbf{x}; \mathbf{y})$

x subject to netwo

$\mathbf{x} \in \Omega$

Multi-domain networking abstraction is aimed at providing this information

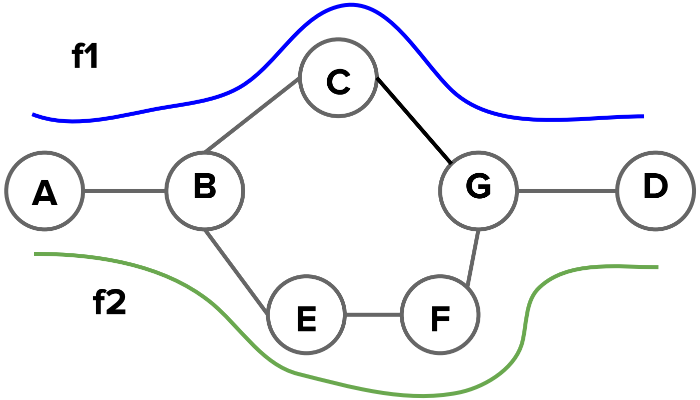
Networks limit potential values of x (e.g., bw, delay, loss rate)



Source: Xiang, Qiao, et al. "Fine-grained, multi-domain network resource abstraction as a fundamental primitive to enable high-performance, collaborative data sciences." ACM/IEEE Supercomputing 2018.

Basic Formulation

- Application interacts with networks by asking the networks to **carry traffic for a set of flows** [f1, f2, ..., fF]
- Consider services provided by the networks to flow fi as an object fi. fi has a set of **network properties**:
  - **Path (fi.path):** representing the sequence of network devices that packets of flow fi will traverse
  - **Delay (fi.delay):** representing the average delay of packets of flow fi
  - **Available bandwidth (fi.abw):** representing the bandwidth that flow fi can request
  - ...
- A network property in a multi-domain setting may involve **network properties of multiple component networks**, e.g.:
  - **fi.path** = fi.path[network1] . fi.path[network2] . ...
  - **fi.delay** = fi.delay[network1] + fi.delay[network2] . ...
  - **fi.abw** = min(fi.abw[network1] + fi.abw[network2], ...)
  - **fi.loss** = log^{-1}(log fi.loss[network1] + log fi.loss[network2] + ...)
  - ...



Example: two flows f1 and f2. f1 passes networks A, B, C, G, D, and f2 passes networks A, B, E, F, G, D.

Basic Idea

Provide the ability to discover, aggregate and expose information of multiple domain networks to provide a **single,** consistent, virtual network **view**.

Represent information using generic, **compact mathematical programming constraints**.

fi.path = fi.path[network1] . fi.path[network2] . ...

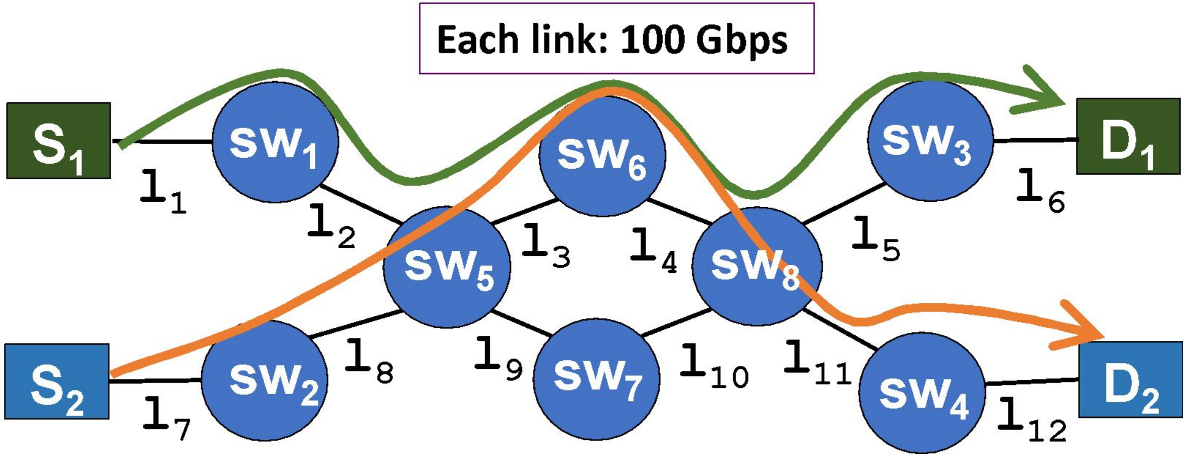
fi.delay = fi.delay[network1] + fi.delay[network2] . ...

fi.abw = min(fi.abw[network1], fi.abw[network2], ...)

fi.loss = $\log^{-1}$(log fi.loss[network1] + log fi.loss[network2] + ...)

...

Example: Mathematical Programming as Abstraction Representation

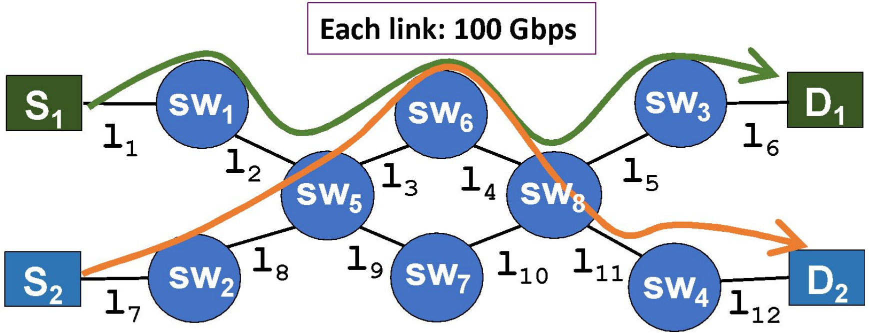


- Two circuits:
  - Flow 1 (f1) : **S1 -> D1**
  - Flow 2 (f2): **S2 -> D2**
- Share common links: **l3 and l4**
  - It is **not possible** to reserve 100Gbps (**for both circuits**)

Source: https://datatracker.ietf.org/doc/draft-ietf-alto-path-vector/

Example: Mathematical Programming as Abstraction Representation

- **GOAL:** Use mathematical programming constraints to provide a compact representation of the **available bandwidth** of flows through **a network**.



$$x_1 \leq 100 \qquad \forall l_u \in \{l_1, l_2, l_5, l_6\},$$
$$x_2 \leq 100 \qquad \forall l_u \in \{l_7, l_8, l_{11}, l_{12}\},$$
$$x_1 + x_2 \leq 100 \qquad \forall l_u \in \{l_3, l_4\},$$

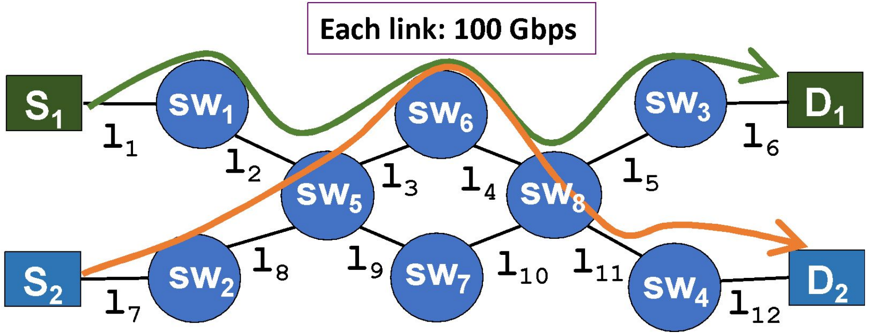
Linear Inequalities

- **x_1 (f_1.awb): flow 1's available bandwidth**
- **x_2 (f_2.awb): flow 2's available bandwidth**

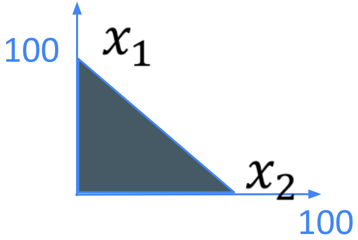
Source: https://datatracker.ietf.org/doc/draft-ietf-alto-path-vector/

Example: Mathematical Programming as Abstraction Representation

- **GOAL:** Use mathematical programming constraints to provide a compact representation of **available bandwidth** of flows through **a network**.



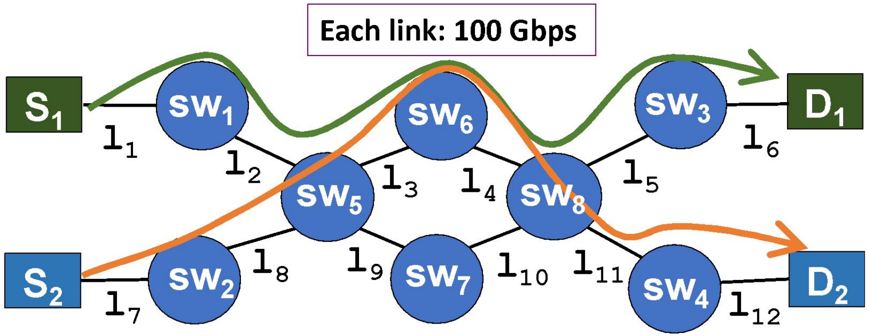
$$\begin{aligned} x_1 &\leq 100 && \forall l_u \in \{l_1, l_2, l_5, l_6\}, \\ x_2 &\leq 100 && \forall l_u \in \{l_7, l_8, l_{11}, l_{12}\}, \\ x_1 + x_2 &\leq 100 && \forall l_u \in \{l_3, l_4\}, \end{aligned}$$



- Geometrically, resource abstraction represents the **resource feasible region** of the network for providing resources to a set of flows.

Example: Mathematical Programming as Abstraction Representation

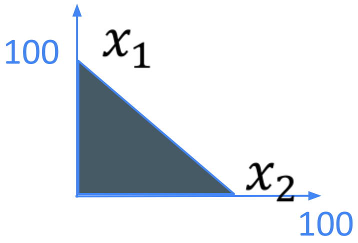
- **GOAL:** Use mathematical programming constraints to provide a compact representation of **available bandwidth** of flows through **a network**.



~~$x_1 \leq 100 \quad \forall l_u \in \{l_1, l_2, l_5, l_6\},$~~

~~$x_2 \leq 100 \quad \forall l_u \in \{l_7, l_8, l_{11}, l_{12}\},$~~

$$x_1 + x_2 \leq 100 \quad \forall l_u \in \{l_3, l_4\},$$

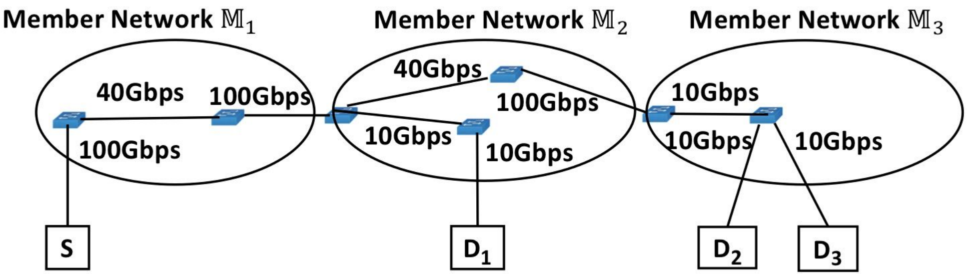


- Geometrically, resource abstraction represents the **resource feasible region** of the network for providing resources to a set of flows.
- **Redundant inequalities are removed** via a polynomial-time, optimal algorithm.

Source: Kai Gao, Qiao Xiang, Xin Wang, Yang Richard Yang, Jun Bi: An Objective-Driven On-Demand Network Abstraction for Adaptive Applications. IEEE/ACM Trans. Netw. 27(2): 805-818 (2019)

Basic Aggregation Abstraction: From a Single Network to Multiple Domains

- Problem: Domain discovery
  - E.g., flow 3 (S -> D_3) will traverse 3 networks (M_1, M_2, and M_3)
  - Why: Decompose global problem into per-domain problem
- Solution:
  - API for egress mapping at each domain: (flow, ingress) -> egress



$\mathbb{M}_1$:

$$x_1 + x_2 + x_3 \leq 100,$$
$$x_1 + x_2 + x_3 \leq 40,$$
$$x_1 + x_2 + x_3 \leq 100,$$

$\mathbb{M}_2$:

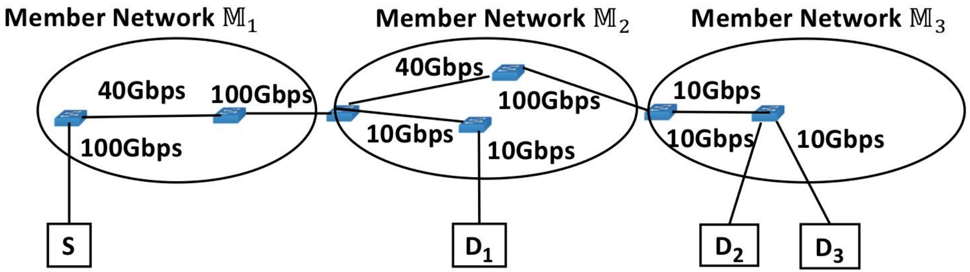
$$x_2 + x_3 \leq 40, \quad x_1 \leq 10,$$
$$x_2 + x_3 \leq 100, \quad x_1 \leq 10,$$

$\mathbb{M}_3$:

$$x_2 + x_3 \leq 10,$$
$$x_2 \leq 10,$$
$$x_3 \leq 10,$$

Multi-Domain Redundancy Optimization

- Problem: Although each domain may already conduct redundancy optimization, there can be cross-domain redundancy



$\mathbb{M}_1$:

~~$x_1 + x_2 + x_3 \leq 100,$~~

$x_1 + x_2 + x_3 \leq 40,$

~~$x_1 + x_2 + x_3 \leq 100,$~~

$\mathbb{M}_2$:

$x_2 + x_3 \leq 40,\quad x_1 \leq 10,$

~~$x_2 + x_3 \leq 100,\quad x_1 \leq 10,$~~

$\mathbb{M}_3$:

$x_2 + x_3 \leq 10,$

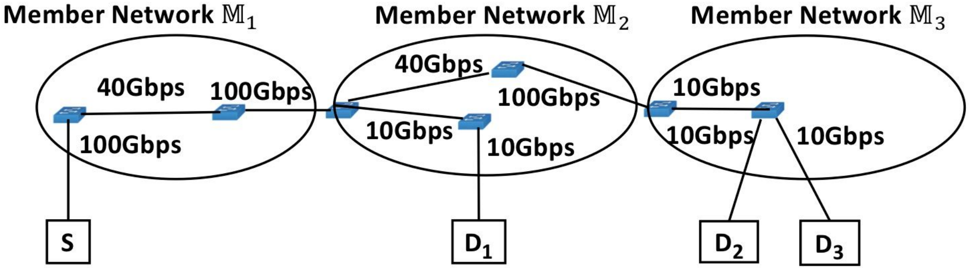
~~$x_2 \leq 10,$~~

~~$x_3 \leq 10,$~~

The constraint on flow 2 and flow 3 at M_3 (<=10) can eliminate that at M_2 (<= 40).

Multi-Domain Redundancy Optimization

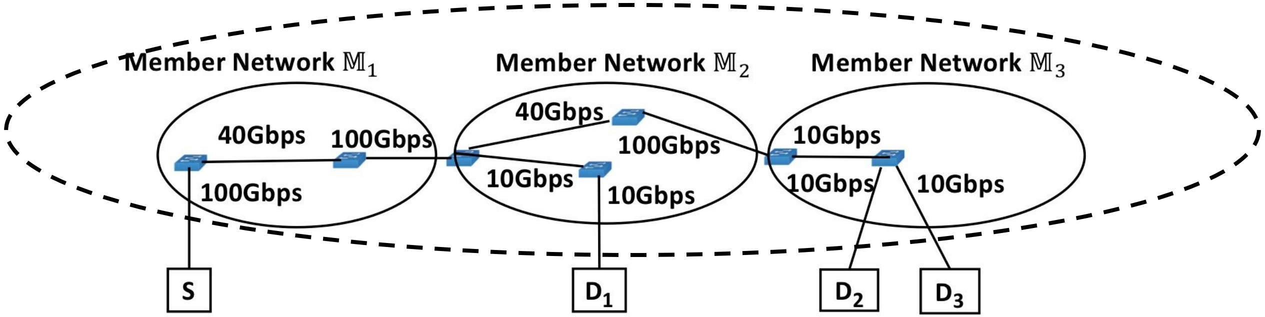
- Problem: Although each domain may already conduct redundancy optimization, there can be cross-domain redundancy



$\mathbb{M}_1$:	$\mathbb{M}_2$:	$\mathbb{M}_3$:
~~$x_1 + x_2 + x_3 \leq 100,$~~	~~$x_2 + x_3 \leq 40,$~~ $x_1 \leq 10,$	$x_2 + x_3 \leq 10,$
~~$x_1 + x_2 + x_3 \leq 40,$~~	~~$x_2 + x_3 \leq 100,$~~ ~~$x_1 \leq 10,$~~	~~$x_2 \leq 10,$~~
~~$x_1 + x_2 + x_3 \leq 100,$~~		~~$x_3 \leq 10,$~~

The constraint on flow 2 and flow 3 at M_3 (<=10) can eliminate that at M_2 (<= 40).

The Reverse View: Mathematical Constraints as Virtual Network Representation



Aggregate the abstraction in multiple networks into a **unified, single, virtual** representation:

$$x_1 \leq 10, \qquad x_2 + x_3 \leq 10,$$

Conclusion

Conclusion

- Summary:
  - Multiple important multi-domain applications can benefit substantially from ALTO
  - Corresponding new design requirements also emerge
  - Different drafts have been proposed to address some of the design requirements
- Next Steps
  - Systematic investigation of deployment concern of ALTO for multi-domain applications
    - Incentive, stability, scalability, privacy, etc.
  - Systematic design of extensions to address corresponding design requirements
- Further ALTO Information
  - IETF105 ALTO session
    - In-person: **Thursday, July 25, 2019 - 10:00-12:00 (Notre Dame room)**
    - Remote participation: https://www.meetecho.com/ietf105/alto/
  - Internal meetings
    - Wednesday weekly meetings (9:30 US ET)
    - Bridge: https://yale.zoom.us/j/8423318713

Thanks!

(more) Questions?