ALTO Topology Service: Use Cases, Requirements and Framework

Greg Bernstein, Young Lee, Y. Richard Yang

IETF 88

Nov. 4, 2013
Overview

- **Goal:** provide *simple, abstracted topologies* to applications and network tools for better network and application optimization.

- Such topologies allow a spectrum of use cases, spanning “simple” visualization, scheduling, diagnosis, to “complex” application-network programming.
Use Case I: Topology Advertisement

- Many networks already provide public views (advertisement), in their specific formats, of their topology infrastructures

- Goal: standard format for consistent information disclosure
Comcast was founded in 1968 as a single-system cable operator and now is the nation's largest with over $35 billion in revenue in 2003. We are one of the nation's leading providers of information, communications and entertainment products and services with over 16 million Internet customers, 8 million digital voice customers, 23 million video customers and hundreds of thousands of business customers.

With over 100,000 employees, we currently serve 18 of the top Metropolitan Statistical Areas (MSAs) in the U.S. and provide service to customers in 39 states and the District of Columbia.

Our high-speed, high-capacity broadband and Ethernet services operate across our private, diverse enhanced fiber network. With over 147,000 national route miles of fiber, our network is the largest facilities-based last mile alternative to the phone company.

With the first and largest fully 40G backbone, and the deployment of the first 100G router interface, Comcast's advanced network delivers reliable and scalable services for businesses of any size.

**COMCAST'S EXTENSIVE NATIONWIDE FIBER OPTIC NETWORK**

**THE LARGEST FACILITIES-BASED, LAST MILE ALTERNATIVE TO THE PHONE COMPANY IN THE UNITED STATES.**
Network Maps

View the network! The CenturyLink network covers the entire continental United States and has one of the largest fiber footprints in the U.S., capable of supporting 40Gbps data transmission rates. Internationally, CenturyLink provides coverage to over 200 destinations, including landline and mobile terminations and currently handles almost four billion international calls per year.
BT Global Services: Network Maps

The map below shows the current core topology of the BT IP Network within Europe, and its connectivity to the USA. BT also offers IP services globally through strategic supply arrangements. These supplier networks can be viewed by clicking on the regional tabs below. In addition to the backbone links shown, the network is connected to all the major Internet Exchanges (see Peering Information) and allows customer access via a wide number of cities in each country, increasing our network depth in-country. To view these Access PoP locations, click on the countries shown in the map.
Use Case I: Example Workflow

ISP 1
ALTO Server

ISP 2
ALTO Server

ISP 3
ALTO Server

ALTO Client

Business Analysis

ALTO Client
Use Case II: Large Flow Scheduling
Large-Flow Scheduling

1. obtain transfer tasks \{(src, dst, data), \ldots\}

2. Obtain path for each src-dst
   At least 3 possibilities: (1) trace route; (2) CostMap with a “path” metric; (3) Topology Service provides computation hint (e.g., shortest path of the topology)

3. while (tasks not done)
   query available bw
   schedule data transfer
Large-Flow Scheduling

- **Hyatt Regency Vancouver, Burrard Street, Va**
- **Vancouver International Airport, 3211 Grant M**

**Suggested routes**
- BC-99 S and Grant 13.8 km, 22 mins
- McConachie Way W 14.3 km, 26 mins

**Driving directions to Vancouver International Airport**

**Hyatt Regency Vancouver**
655 Burrard St
Vancouver, BC V6C 2R7, Canada

1. Head southwest on Burrard St toward W Georgia St/BC-1A S
Use Case III: Two-Level Traffic Engineering

- Traffic engineering algorithms based on optimization typically cannot scale to fine-grained, large-scale topologies. They often work well only on abstract, small-scale (e.g., aggregated PoP-level) topologies.

- Multiple level optimization is a commonly used system design (e.g., load balancing design, OS device driver) structure
Use Case III: Two-Level Traffic Engineering
Use Case III Example: R3 TE

Compute \((r,p)\) to minimize MLU (Max Link Utilization) for original demand \(d + \) rerouted traffic \(x \in X_F\)
- \(r\) carry \(d\), \(p\) carry \(x\)

\[
\min_{(r,p)} \text{MLU}
\]

[C1] \(r\) is a routing, \(p\) is a routing;

[C2] \(\forall x \in X_F, \forall e:\)

\[
\left[ \sum_{a,b \in V} d_{ab} r_{ab}(e) + \sum_{l \in E} x_l p_l(e) \right] / c_e \leq \text{MLU}
\]

- [C2] contains infinite number of constraints, we apply LP duality to convert to polynomial number of constraints
- \(r\) may be given (e.g., as an outcome of OSPF)

R3: Resilient Routing Reconfiguration, Y. Wang (Yale), H. Wang (Google), Ajay Mahimkar (UT Austin), Richard Alimi (Yale), Yin Zhang (UT Austin), Lili Qiu (UT Austin), Y. Richard Yang (Yale). In SIGCOMM 2010.
Use Case III Example: R3 TE
Use Case III: Example Workflow

ALTO Server

ISP

ALTO Client

fine-grained real topology

LSPs, link weights in abstract topo

Top Level TE

Low level TE

coarse abstract topology
Use Case IV: SDN

Use Case IV: SDN

Current programming model

Policy

Programming

**p.TcpDst t=22?**

- **yes**
  - drop
  - match: {TcpDst=22}
  - action: drop

- **no**
  - send along lowest-weight path from p.EthSrc to p.EthDst
  - action: outPort(1)
  - action: outPort(1)
  - action: outPort(1)
Use Case IV: SDN

Current programming model

**Policy**

- **p.TcpDs t=22?**
  - yes: drop
  - no: send along lowest-weight path from p.EthSrc to p.EthDst

**Programming**

- **match:{TcpDst=22}**
  - action: drop
  - priority: HIGH

- **tcpDst!=22**
    - action: outPort(1)
    - priority: LOW
    - action: outPort(1)
    - action: outPort(1)
Use Case IV: SDN (Maple)

Route f(Packet p, Env e) {
  if (p.tcpDstIs(22))
    return null();
  else {
    Location sloc = e.location(p.ethSrc());
    Location dloc = e.location(p.ethDst());
    Path path = myPathAlg (e.links(),
                             sloc, dloc);
    return unicast(sloc, dloc, path);
  }
}

Use Case IV: Example Workflow

ISP

ALTO Server

flow table entries

fine-grained real topology

Maple run time

Routes in abstract topology

User f

ALTO Client

course abstract topology
Use Case V: e2e Joint App/Net Routing

• Instead of a design that Network Path Computation Element understands every single possible application path selection consideration, a higher level orchestrator computes both network routing, and app-level load balancing, considering network resource availability/constraints, compute/storage resource availability/constraints, and other policies (e.g., geo restrictions).
Use Case V: Example

Source: <draft-bernstein-alto-large-bandwidth-cases-01>
Use Case V: Example

Chosen paths considered both net and compute/storage

Source: <draft-bernstein-alto-large-bandwidth-cases-01>
Use Case V: Example Workflow

ALTO Server

- real topology
- simplified abstract topology
- compute/storage/policy

Network Service Interface

- network control guide

Global Orchestration

- compute/storage guide
Basic Requirements:
Extend the Base ALTO “Single Switch” Abstraction

Access port/pop connecting endhosts (PID)

Inter-PID “distances” define a Cost Map
Basic Requirements:
Extend the Base ALTO “Single Switch” Abstraction

Access port/pop connecting endhosts (PID)

Inter-PID “distances” define a Cost Map

Req1: Add nodes

Req2: Add links
Req 3: Annotate Properties on Links and Nodes

- Properties depend on application desire
- Example properties
  - Typical link performance metrics such as BW, cost, latency, jitter, ...
  - Typical node properties such as geo location
Req 4: Information Abstraction

- Simplification of application programming
  - Present as simple a representation of a network to an application as possible, but not simpler

- Information hiding for network privacy
  - Allow network layer to hide as much information as desired
Req 5: Policy Preserving Routing

- Real network topology T
- Application is given abstract topology Ta derived from T
- Application computes route Ra from Ta
- Network converts Ra expressed in Ta to real route R in real topology T

- An example of non-preserving representation:
  - T = Ta
  - Application uses arbitrary routing, but real network can use only shortest path (OSPF)

- An example of non-preserving realization
  - Application chooses path avoiding region X, but realization uses a path that visits region X
Summary of Key Requirements

- Add links
- Add nodes
- Allow link/node properties
- Information abstraction
- Policy preservation
Related Work

• GMPLS Routing
  – Not intended for application use. Does not provide network information abstraction/hiding

• I2RS

• OGF NML
  – General framework for multi-layer network modeling in XML/RDF based on ITU-T G.800. Very detailed. Could be useful as a starting point. Does not include a JSON representation
ALTO Topology Framework

- **Sources of raw topology information**
  - From management systems, to proprietary interfaces to routing systems, to i2rs... (how to transport to ALTO server is out of scope)

- **Abstract topology representation**

- **Service/client specific topology transformation**
  - [I-D.lee-alto-app-net-info-exchange] and [I-D.yang-alto-topology]

- **Network Service Interface compatibility**
Next Step

• Specify a complete ALTO Topology Services Resource

• Propose WG re-chartering to include Topology Service and go beyond P2P
Backup Slides
Framework

- How the service may interface with service interfaces

```
Topology Client

ALTO Topology Server

Network Topology
```

Application optimizer

Reservation Interface (NaaS)

Network provisioning system

Out of scope
Requirements Wish List

• Topology Representation
  – Paths or Graphs
  – Bandwidth constraint information
  – Metrics: cost, latency, jitter, SRLG, ...
  – Location information

• Incremental Updates...

• Compatibility
  – With “reservation interfaces”
  – Route computation system, i.e., for process of transforming a “loose route” provided by application to detailed route implemented in the network.