ALTO Topology Service: Use Cases, Requirements and Framework

Greg Bernstein, Young Lee, Y. Richard Yang

IETF 88

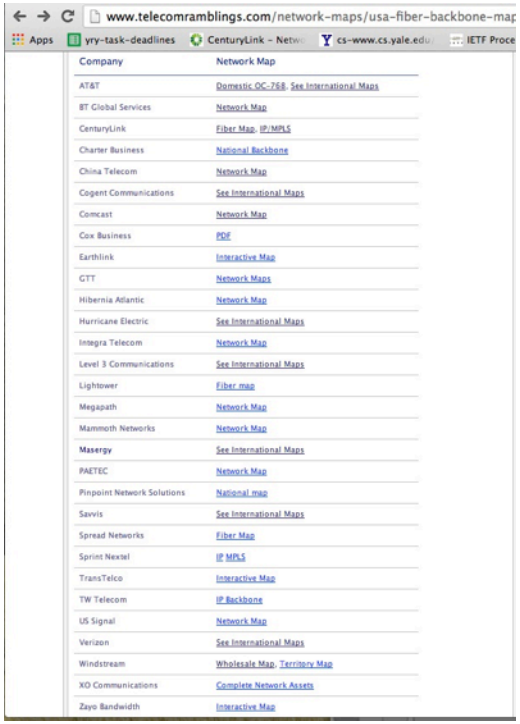
Nov. 4, 2013

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

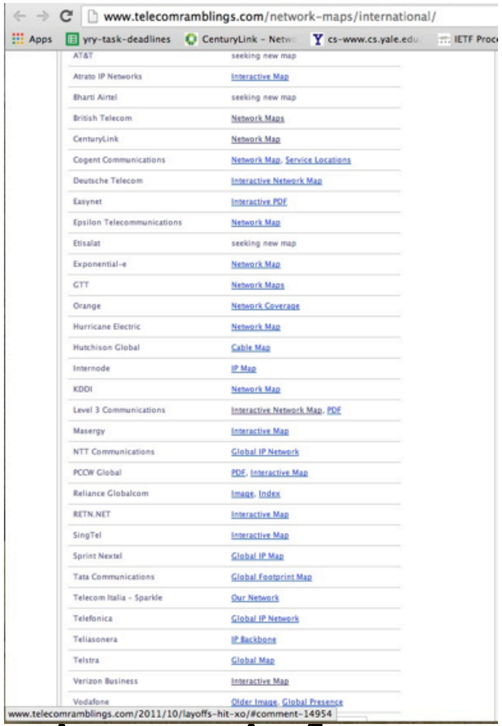
- Many networks already provide public views (advertisement), in their specific formats, of their topology infrastructures
  – http://www.telecomramblings.com/network-maps/usa-fiber-backbone-map-resources/



www.telecomramblings.com/network-maps/usa-fiber-backbone-map

Apps yry-task-deadlines CentturyLink – Netw cs-www.cs.yale.edu IETF Proce

Company	Network Map
AT&T	Domestic OC-768, See International Maps
BT Global Services	Network Map
CenturyLink	Fiber Map, IP/MPLS
Charter Business	National Backbone
China Telecom	Network Map
Cogent Communications	See International Maps
Comcast	Network Map
Cox Business	PDF
Earthlink	Interactive Map
GTT	Network Maps
Hibernia Atlantic	Network Map
Hurricane Electric	See International Maps
Integra Telecom	Network Map
Level 3 Communications	See International Maps
Lightower	Fiber map
Megapath	Network Map
Mammoth Networks	Network Map
Masergy	See International Maps
PAETEC	Network Map
Pinpoint Network Solutions	National map
Savvis	See International Maps
Spread Networks	Fiber Map
Sprint Nextel	IP MPLS
TransTelco	Interactive Map
TW Telecom	IP Backbone
US Signal	Network Map
Verizon	See International Maps
Windstream	Wholesale Map, Territory Map
XO Communications	Complete Network Assets
Zayo Bandwidth	Interactive Map



www.telecomramblings.com/network-maps/international/

Apps yry-task-deadlines CenturyLink – Netw cs-www.cs.yale.edu IETF Proc

AT&T	seeking new map
Atrato IP Networks	Interactive Map
Bharti Airtel	seeking new map
British Telecom	Network Maps
CenturyLink	Network Map
Cogent Communications	Network Map, Service Locations
Deutsche Telecom	Interactive Network Map
Easynet	Interactive PDF
Epsilon Telecommunications	Network Map
Etisalat	seeking new map
Exponential-e	Network Map
GTT	Network Maps
Orange	Network Coverage
Hurricane Electric	Network Map
Hutchison Global	Cable Map
Internode	IP Map
KDDI	Network Map
Level 3 Communications	Interactive Network Map, PDF
Masergy	Interactive Map
NTT Communications	Global IP Network
PCCW Global	PDF, Interactive Map
Reliance Globalcom	Image, Index
RETN.NET	Interactive Map
SingTel	Interactive Map
Sprint Nextel	Global IP Map
Tata Communications	Global Footprint Map
Telecom Italia - Sparkle	Our Network
Telefonica	Global IP Network
Teliasonera	IP Backbone
Telstra	Global Map
Verizon Business	Interactive Map
Vodafone	Older Image, Global Presence

www.telecomramblings.com/2011/10/layoffs-hit-xo/#comment-14954

- Goal: standard format for consistent information disclosure business.comcast.com/docs/general-docs/Network_Brochure.pdf?sfvrsn=0

Apps | yry-task-deadlines | CenturyLink - Netwo | cs-www.cs.yale.edu/ | IETF Proceedings | THE SCHOOL OF CIS

Comcast was founded in 1963 as a single-system cable operator and now is the nation's largest with over $35 billion in revenue in 2009. 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.

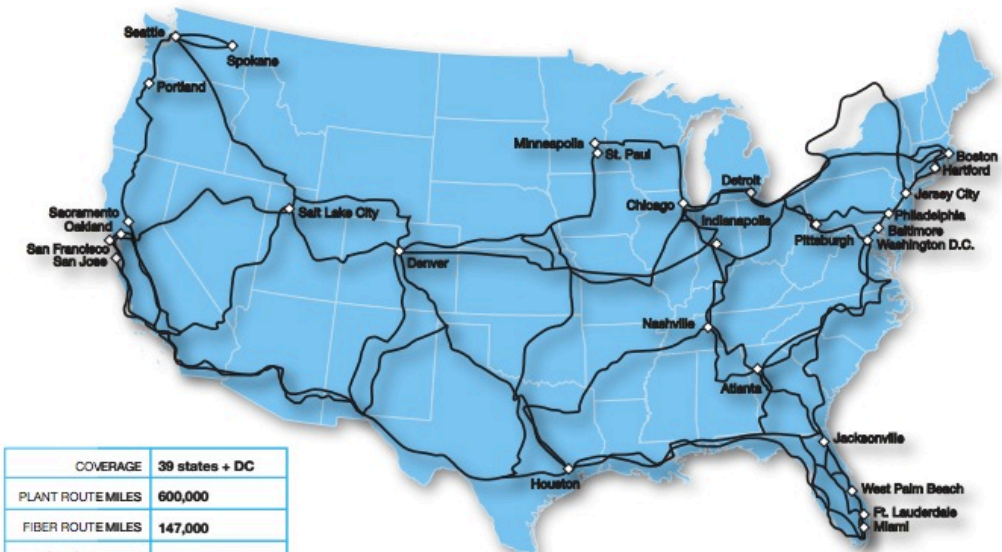
With over 100,000 employees, we currently serve 18 of the top Metropolitan Statistical Areas (MSAs) in the U.S. and provide service to customers in 39 states and the District of Columbia.

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.**

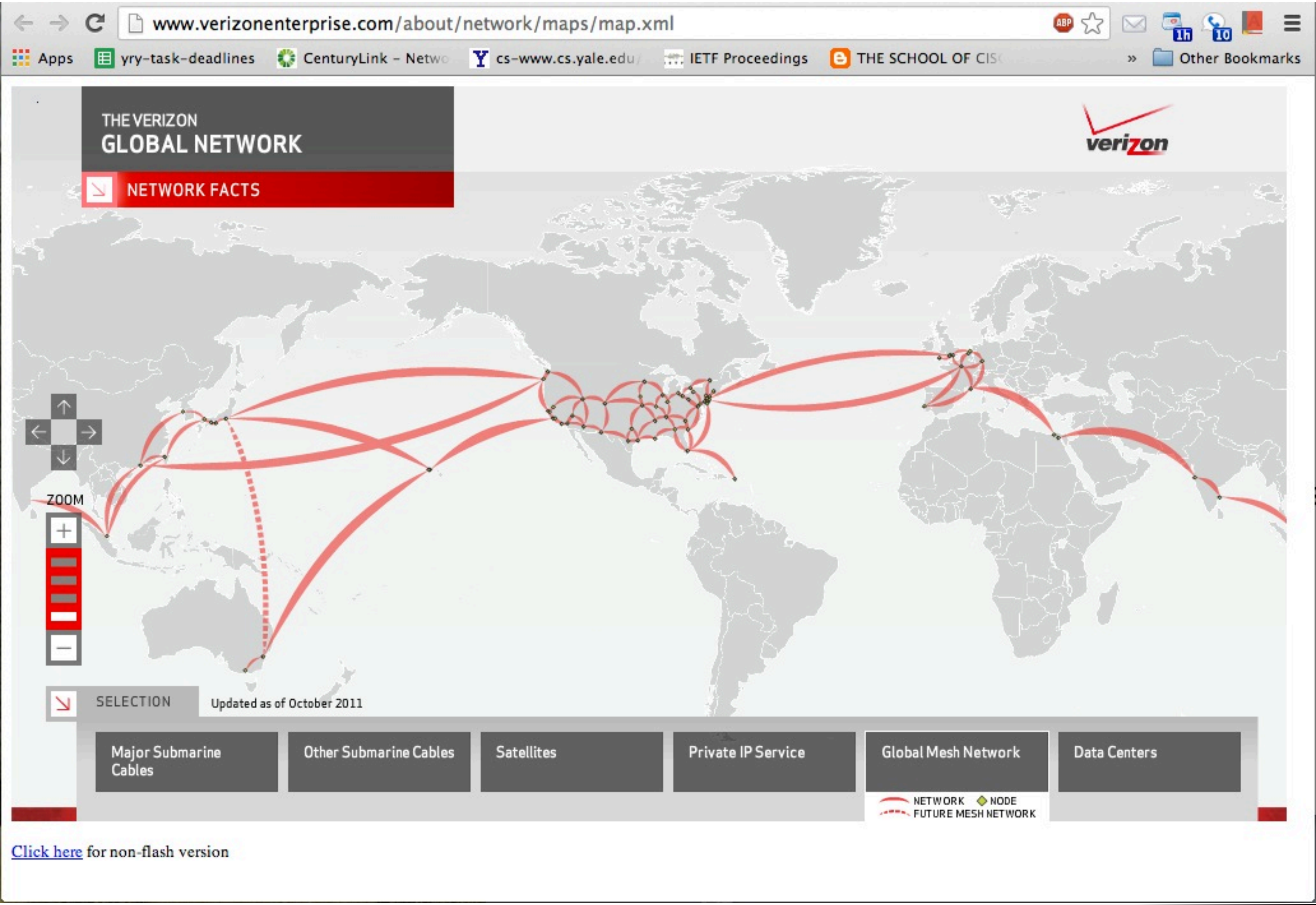
With the first and largest fully 40G backbone, and the deployment of the first 100G router interface, Comcast's advanced network delivers reliable and scalable services for businesses of any size.

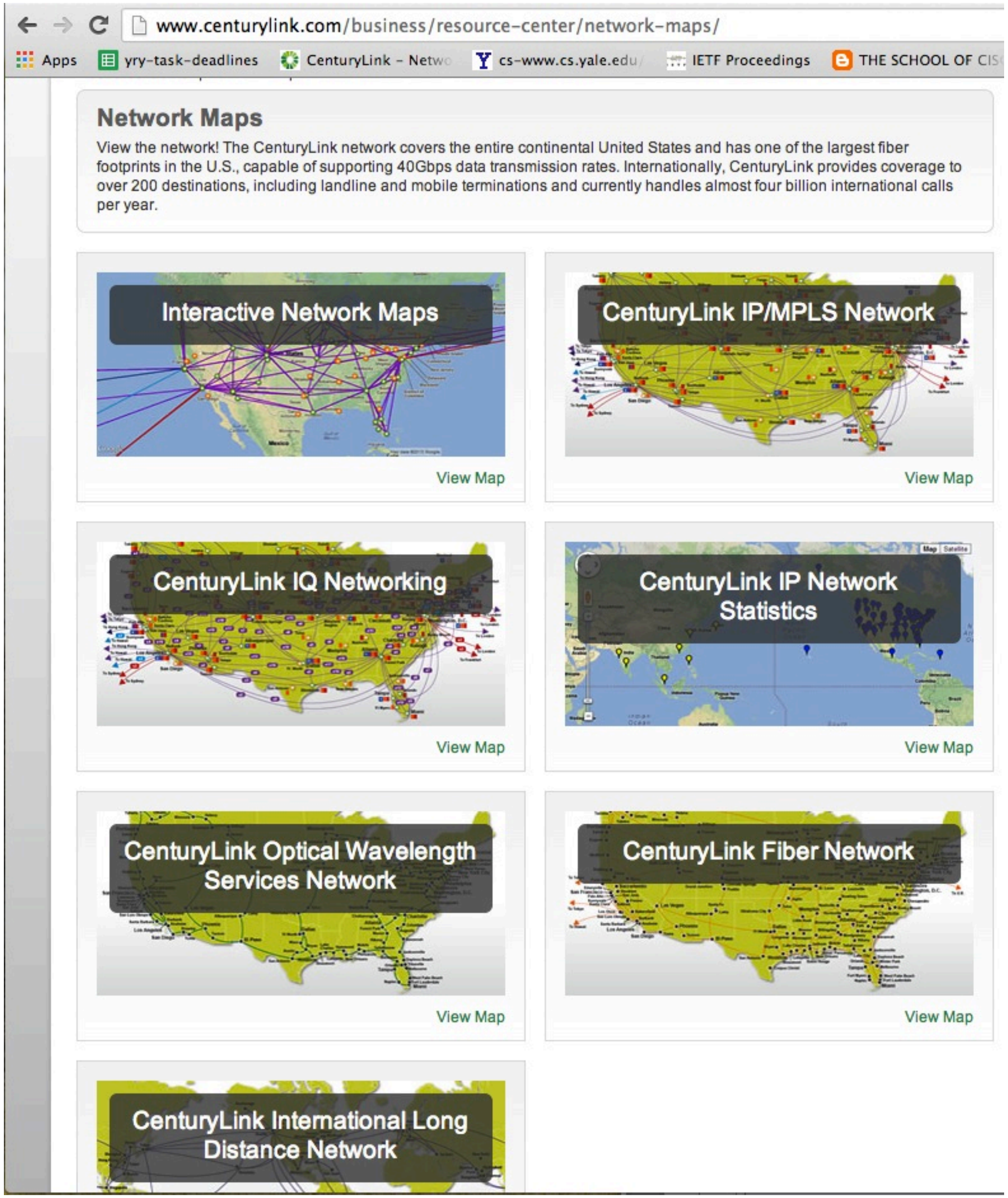
COMCAST'S EXTENSIVE NATIONWIDE FIBER OPTIC NETWORK

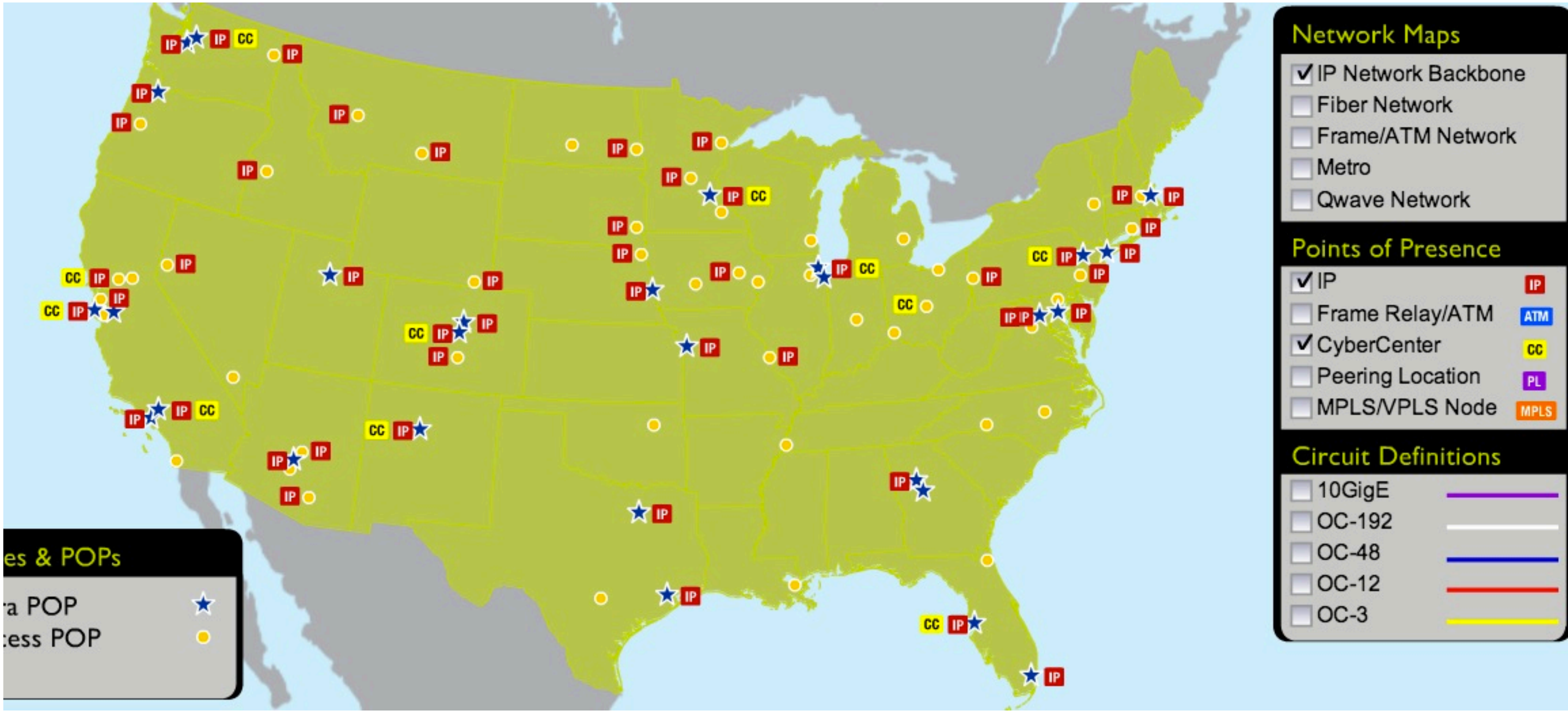
THE LARGEST FACILITIES-BASED, LAST MILE ALTERNATIVE TO THE PHONE COMPANY IN THE UNITED STATES.



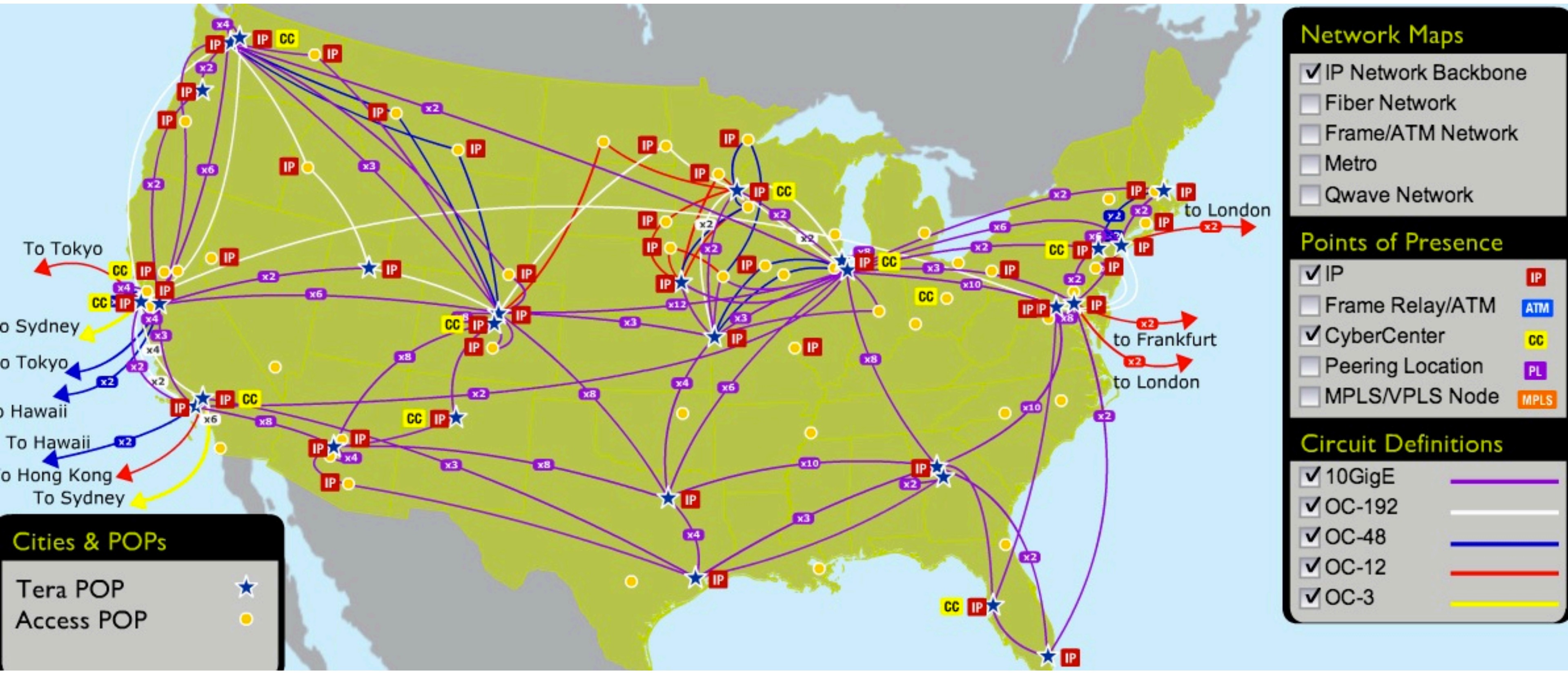
COVERAGE	39 states + DC
PLANT ROUTE MILES	600,000
FIBER ROUTE MILES	147,000
OPTICAL NODES	116,000
TRAFFIC TYPE	95% commercial or customer of a direct peer


www.verizonenterprise.com/about/network/maps/map.xml
Apps
yry-task-deadlines
CenturyLink - Netwo
cs-www.cs.yale.edu/
IETF Proceedings
THE SCHOOL OF CIS
Other Bookmarks
THE VERIZON
GLOBAL NETWORK
NETWORK FACTS
verizon
ZOOM
SELECTION
Updated as of October 2011
Major Submarine Cables
Other Submarine Cables
Satellites
Private IP Service
Global Mesh Network
Data Centers
NETWORK
NODE
FUTURE MESH NETWORK
Click here for non-flash version


www.centurylink.com/business/resource-center/network-maps/
Apps
yry-task-deadlines
CenturyLink - Netwo
cs-www.cs.yale.edu/
IETF Proceedings
THE SCHOOL OF CIS
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.
Interactive Network Maps
View Map
CenturyLink IP/MPLS Network
View Map
CenturyLink IQ Networking
View Map
CenturyLink IP Network Statistics
View Map
CenturyLink Optical Wavelength Services Network
View Map
CenturyLink Fiber Network
View Map
CenturyLink International Long Distance Network


Network Maps
IP Network Backbone
Fiber Network
Frame/ATM Network
Metro
Qwave Network
Points of Presence
IP
Frame Relay/ATM
CyberCenter
Peering Location
MPLS/VPLS Node
Circuit Definitions
10GigE
OC-192
OC-48
OC-12
OC-3
es & POPs
a POP
cess POP


Seattle
Portland
Boise
Minneapolis
Rochester
Boston
Newark
Philadelphia
New York City
Japan
Sacramento
San Francisco
Salt Lake City
Omaha
Chicago
Baltimore
To Europe
Kansas City
Cincinnati
Washington D.C.
To UK
Hawaii
Denver
Las Vegas
Los Angeles
To Japan
San Diego
Memphis
Charlotte
Atlanta
Phoenix
Dallas
Cities & POPs
Major Cities
El Paso
Austin
Houston
Tampa
Miami



Network Maps
IP Network Backbone
Fiber Network
Frame/ATM Network
Metro
Qwave Network
Points of Presence
IP
Frame Relay/ATM
CyberCenter
Peering Location
MPLS/VPLS Node
Circuit Definitions
10GigE
OC-192
OC-48
OC-12
OC-3
Cities & POPs
Tera POP
Access POP
To Tokyo
To Sydney
To Hawaii
To Hong Kong
to London
to Frankfurt

www.bt.net/info/europe.shtml

Apps | yry-task-deadlines | CenturyLink – Netw | cs-www.cs.yale.edu | IETF Proceedings

BT

Tools | Network Information | Looking Glass | IP Performance | Privacy Policy

Home | BT Global Services

Network Information

- Network Maps
- Acceptable Use Policy
- Peering Information
- IP Security
- Technical Information
  - BGP FAQ

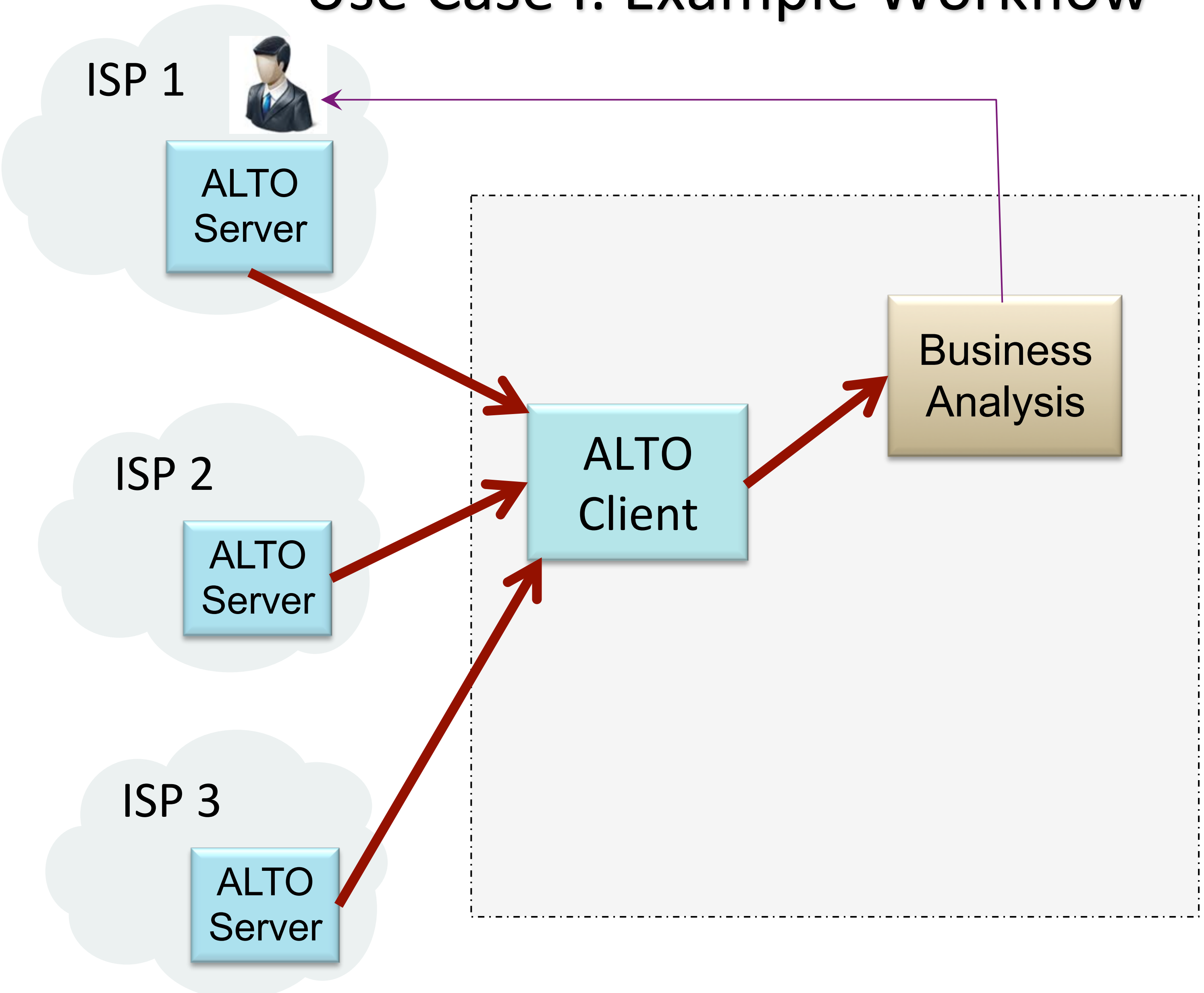
BT Global Services: Network Maps

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.

Europe | North America | Latin America | Asia Pacific

www.chinatelecomusa.com/content_images/ChinaNet_Full.jpg
Apps
yry-task-deadlines
cs-www.cs.yale.edu/
IETF Proceedings
Harbin
Changchun
Shenyang
Urumqi
Hohhot
Beijing
Dalian
Tianjin
Kashi
Yinchuan
Shijiazhuang
Taiyuan
Xining
Lanzhou
Jinan
Qingdao
International link
Zhengzhou
Xi'an
Nanjing
Suzhou
Shiquanhe
Hefei
Shanghai
Lhasa
Wuhan
Hangzhou
Chengdu
Nanchang
International link
Chongqing
Changsha
Fuzhou
Guiyang
Xiamen
Kunming
International link
Guangzhou
Nanning
Hong Kong
Haikou
International link

Use Case I: Example Workflow

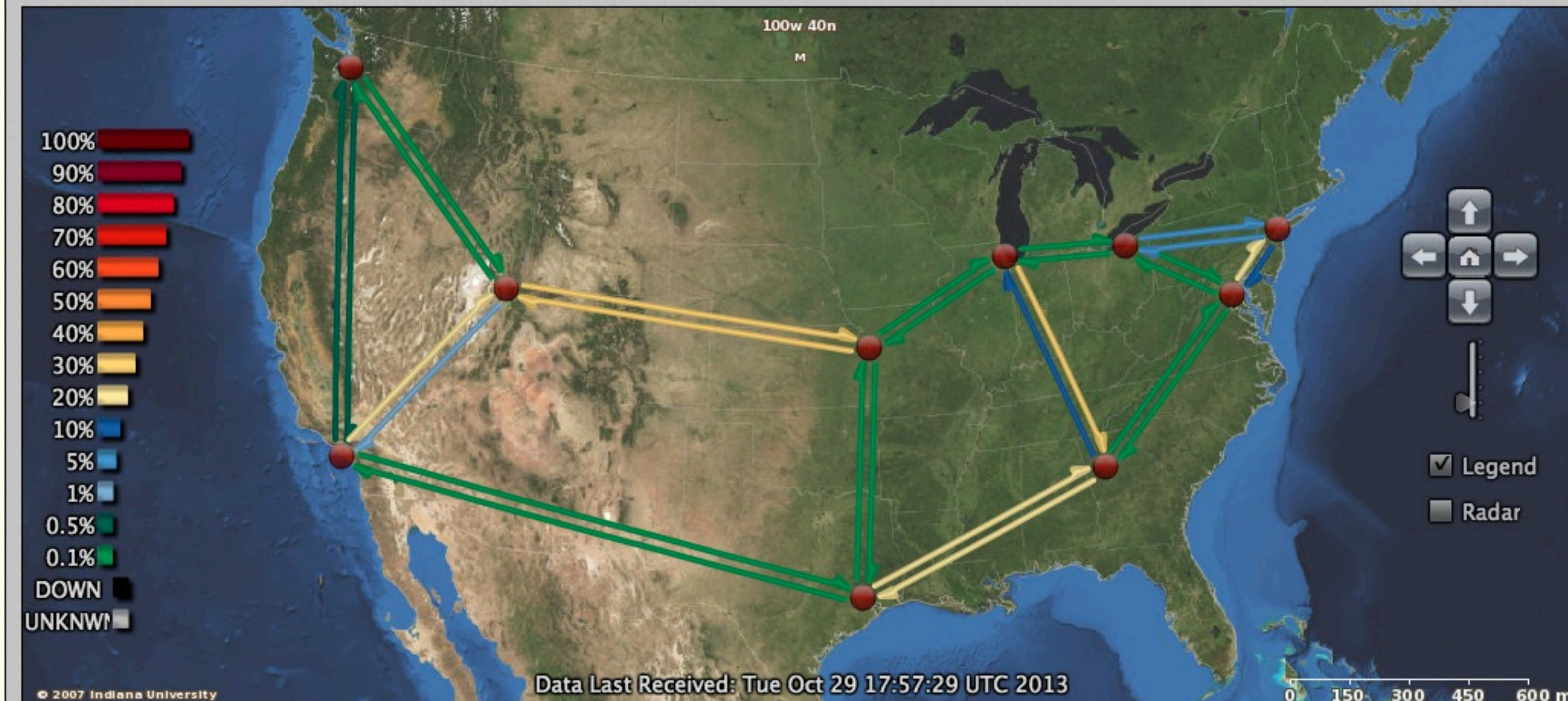


atlas.grnoc.iu.edu/atlas.cgi?map_name=Internet2%20IP%20Layer

Apps | yry-task-deadlines | CenturyLink - Netwo | cs-www.cs.yale.edu/ | IETF Proceedings | THE SCHOOL OF CIS | OpenDaylight Contr | » | Other Bookmark

Internet2 Network

GlobalNOC RealTimeAtlas

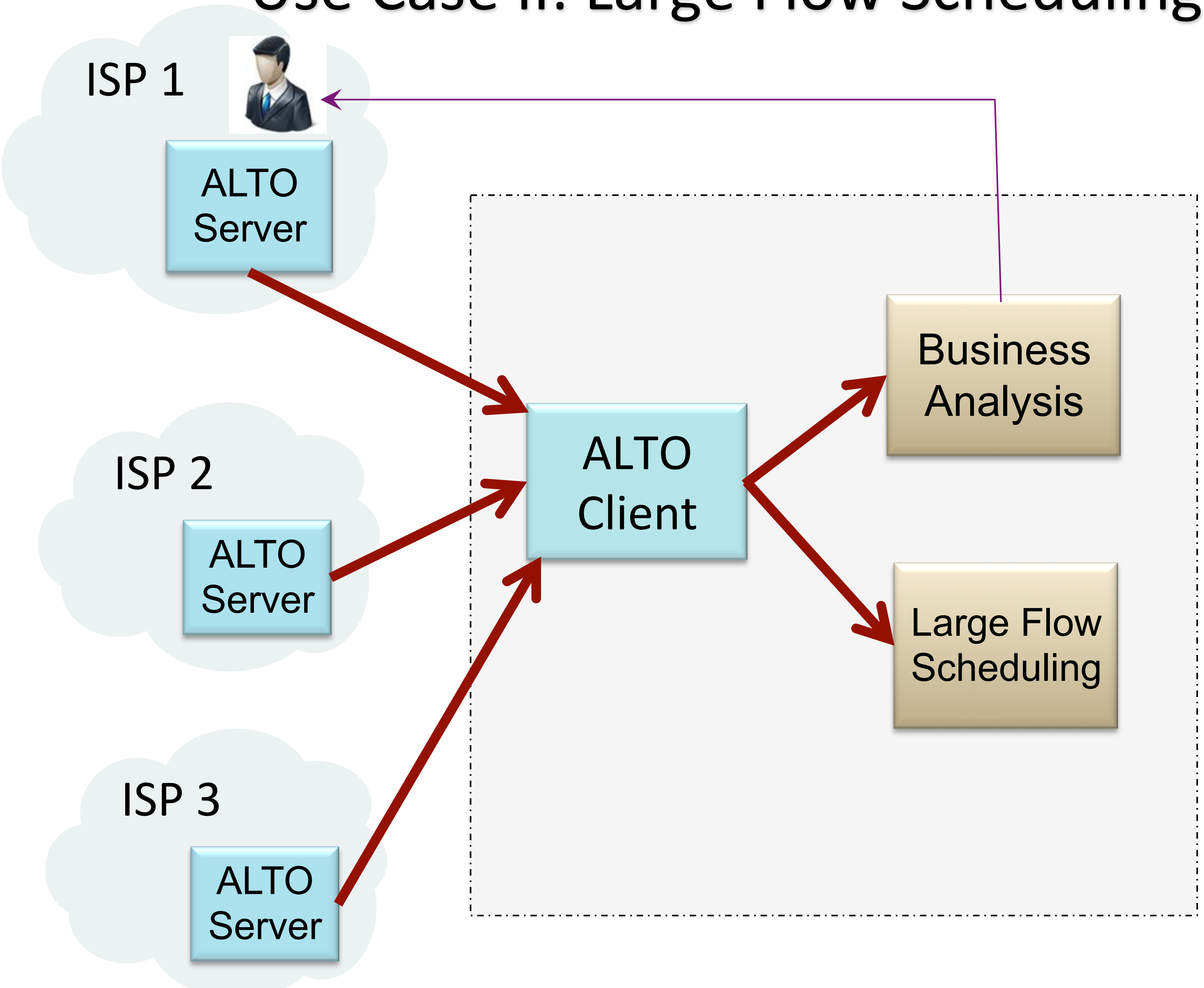


About | Statistics | Circuit Info

Statistics

Circuit Name	A -> Z	bits/sec		Packets/sec		Errors/sec	Z -> A	bits/sec		Packets/sec		Errors/sec
I2-LOSA-SEAT-LAG-06154	seat -> losa	64.78	Mbps	16.45	Kpps	0	losa -> seat	75.88	Mbps	9647	pps	0
I2-LOSA-SALT-LAG-08943	salt -> losa	133.54	Mbps	15.33	Kpps	0	losa -> salt	2.17	Gbps	209.89	Kpps	0
I2-NEWY32AOA-WASH-LAG-0902	wash -> new	2.19	Gbps	244.71	Kpps	0	newy32aoa	1.57	Gbps	273.1	Kpps	0
I2-ATLA-HOUS-LAG-08969	hous -> atla	3.75	Gbps	443.49	Kpps	0	atla -> hous	3.49	Gbps	514.98	Kpps	0
I2-HOUS-KANS-LAG-08962	kans -> hou	7331	bps	3	pps	0	hous -> kan	2012	bps	2	pps	0

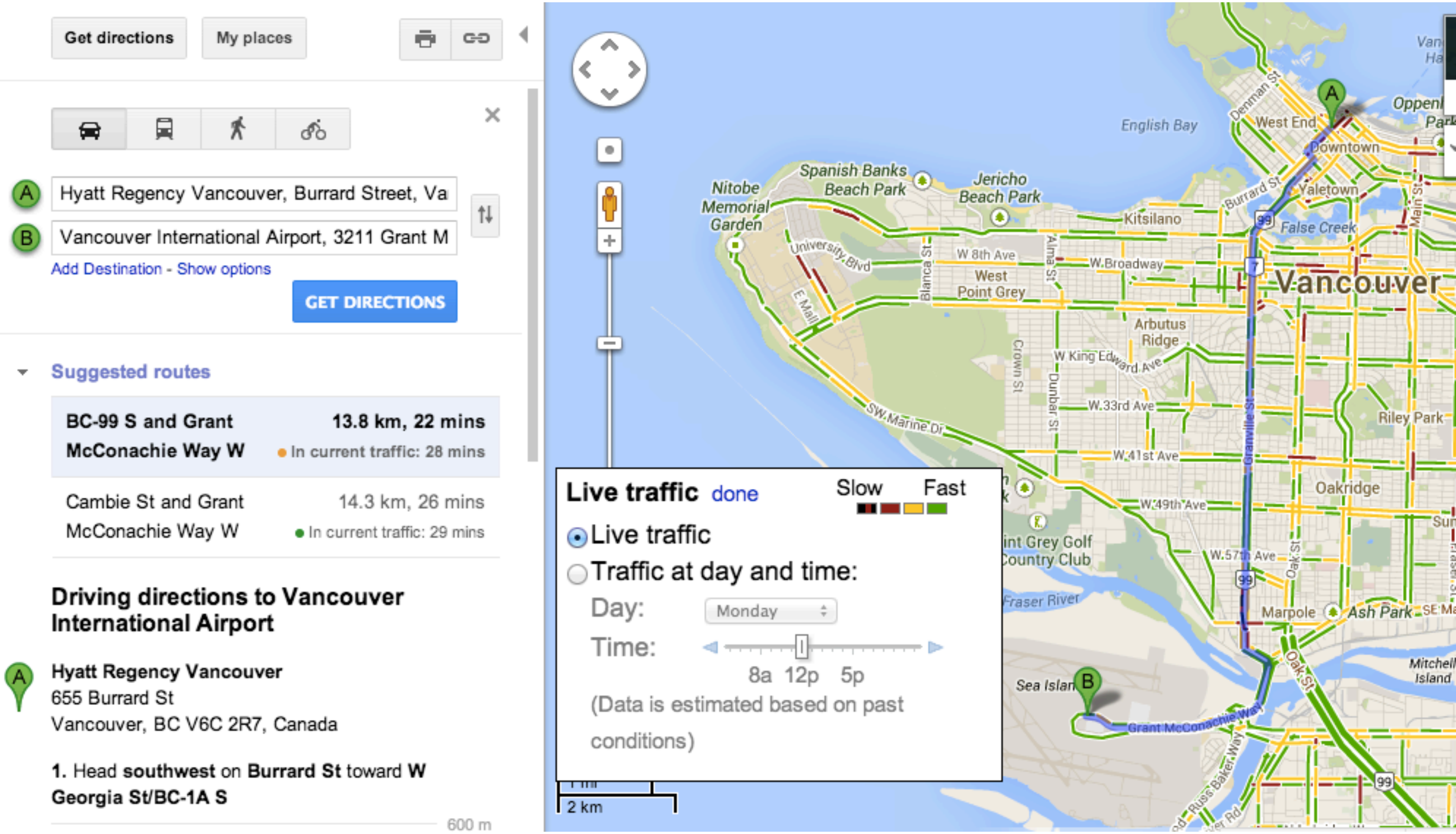
Use Case II: Large Flow Scheduling



Large-Flow Scheduling

```
1. obtain transfer tasks {(src, dst, data), …}

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

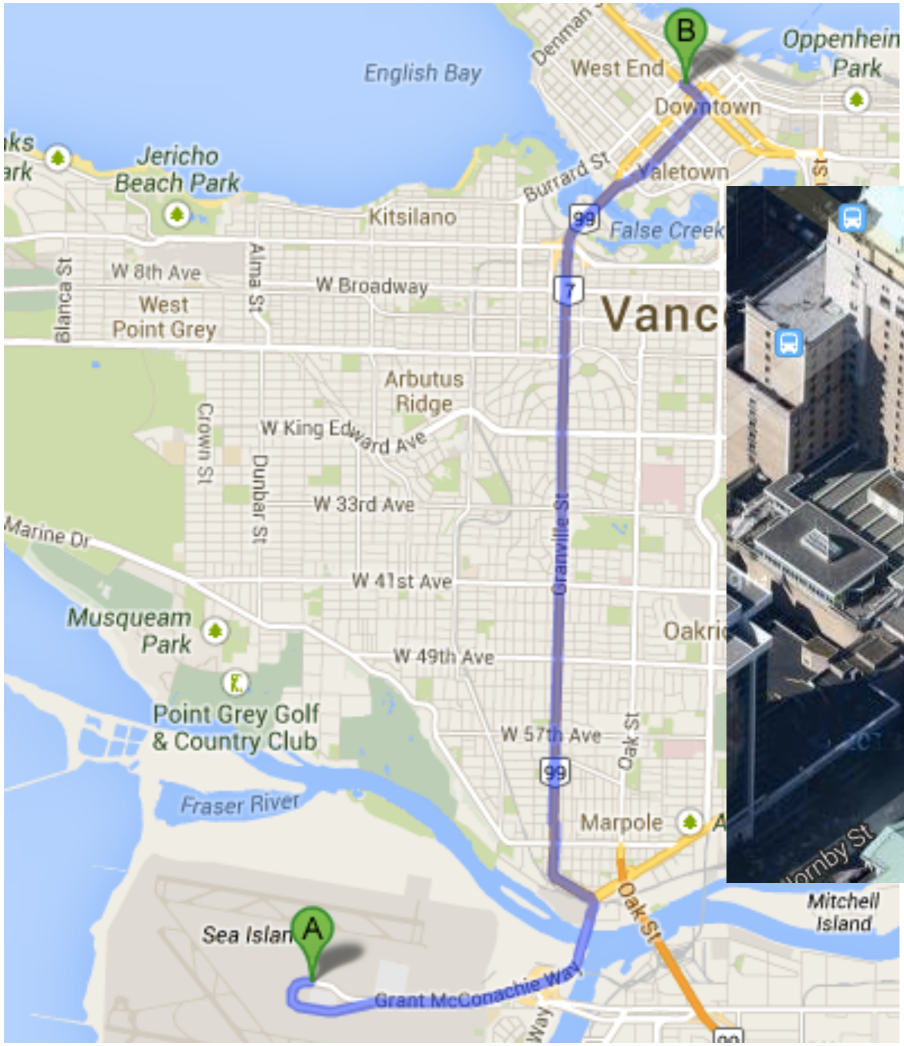


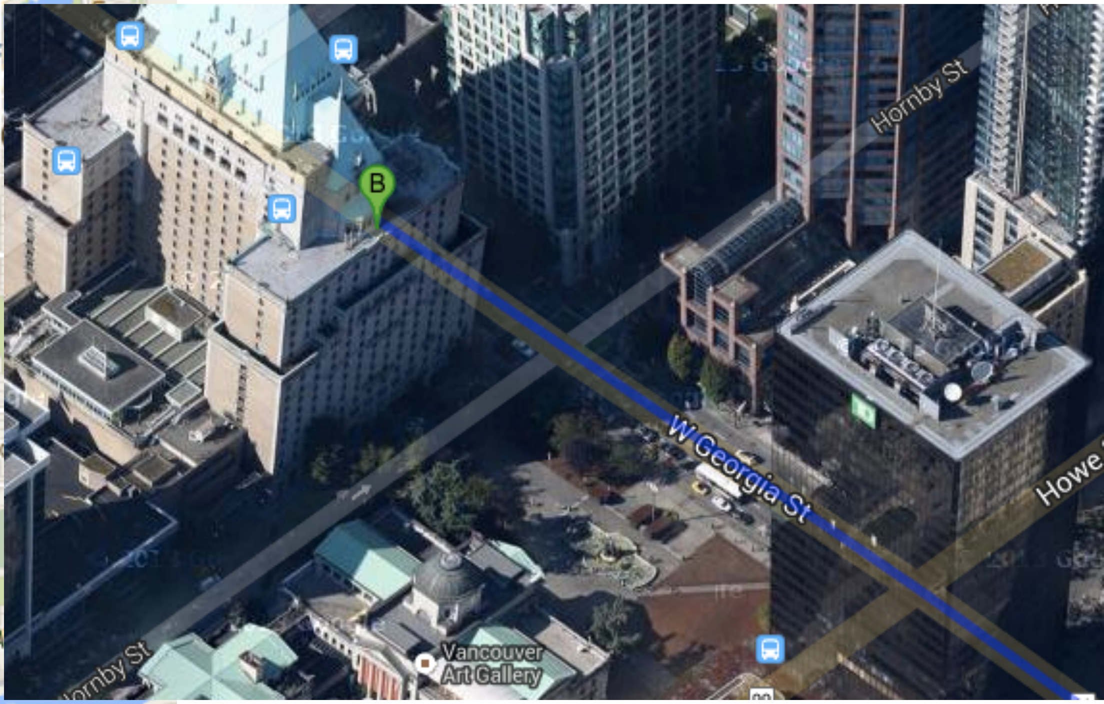
Use Case III: Two-Level Traffic Engineering

- 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

$\min_{(r,p)}$ MLU

[C1] r is a routing, p is a routing;

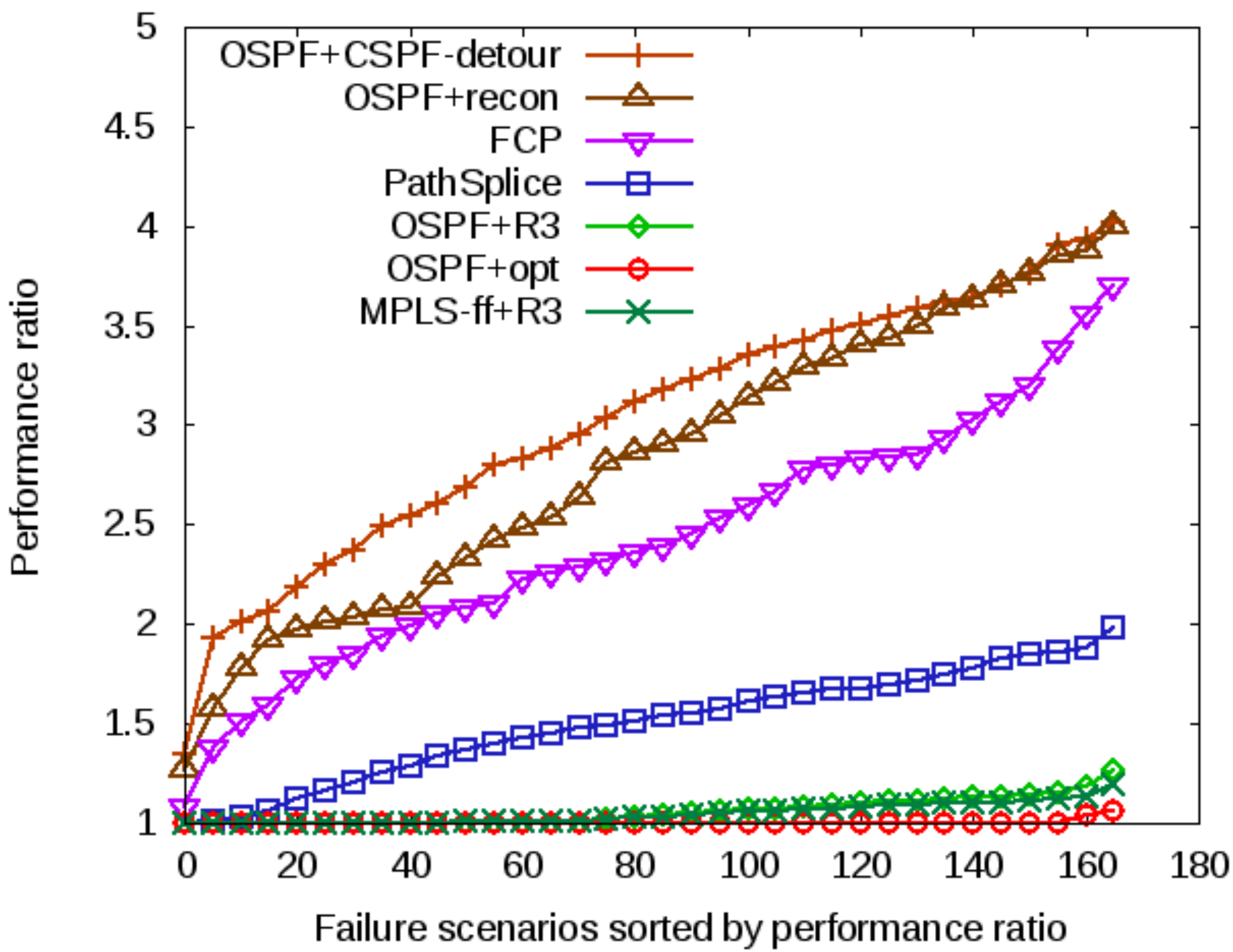
[C2] $\forall x \in X_F, \forall e$:

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

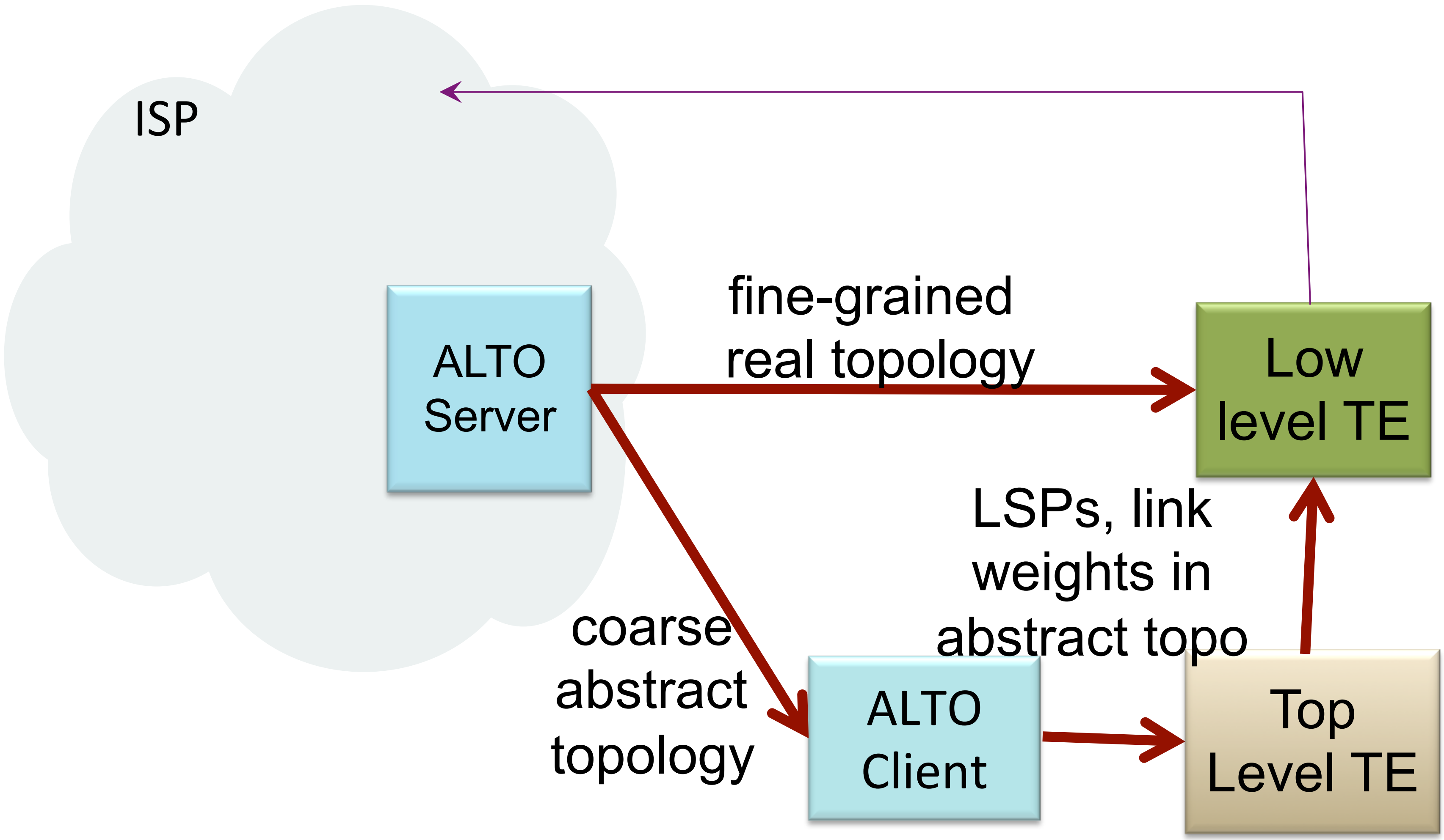
– [C2] contains infinite number of constraints, we apply LP duality to convert to polynomial number of constraints
– r may be given (e.g., as an outcome of OSPF)

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.

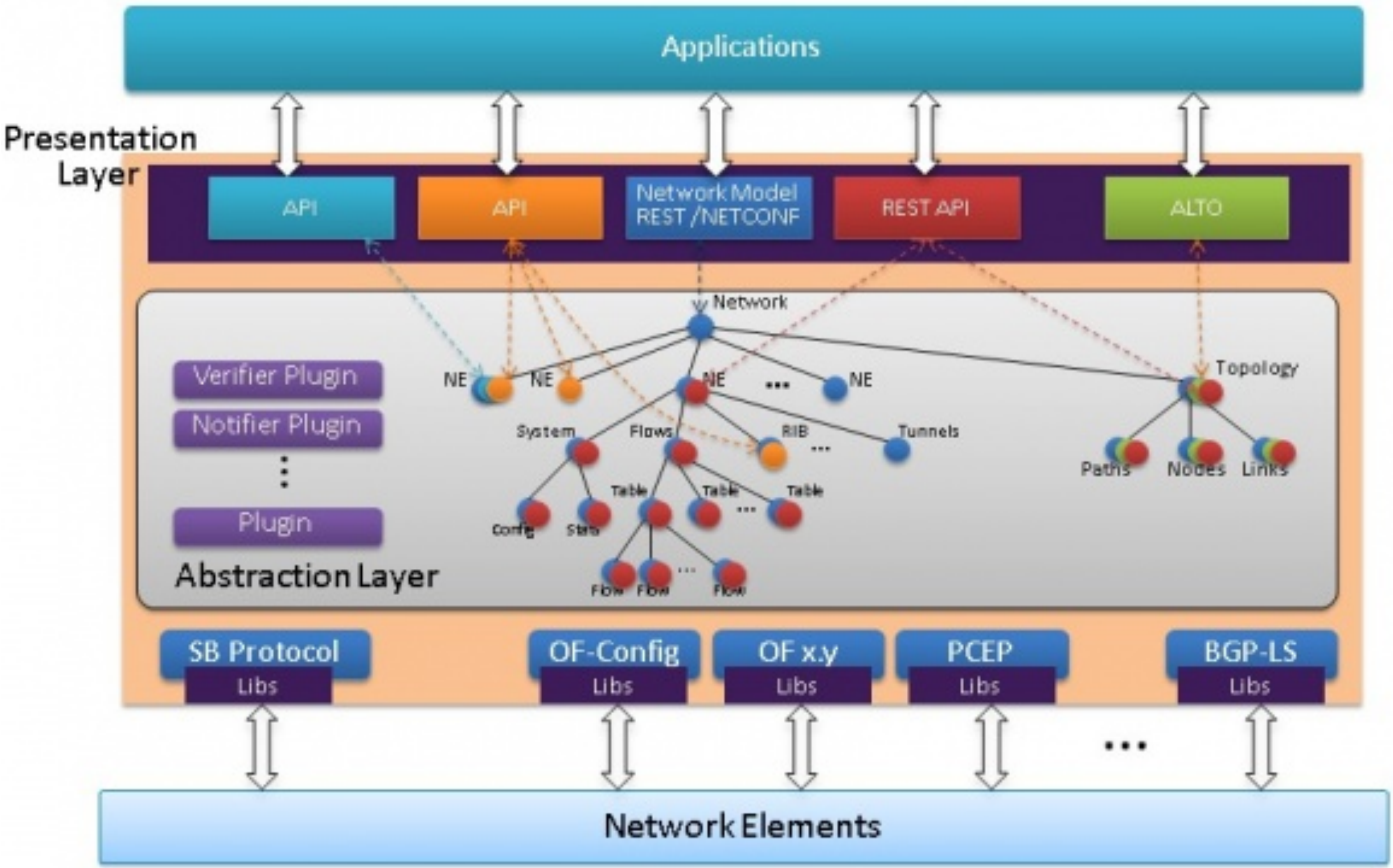
Use Case III Example: R3 TE



Use Case III: Example Workflow



Use Case IV: SDN



https://wiki.opendaylight.org/view/File:SAL_NB_Plugins.jpg

Use Case IV: SDN

Current programming model

Policy

Programming

p.TcpDst=22?

yes → drop

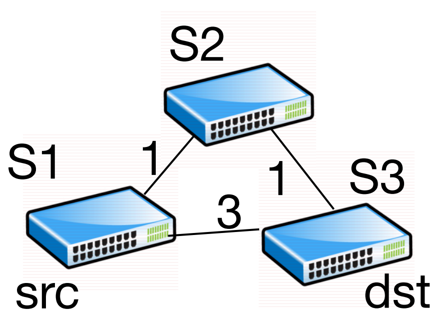
match:{TcpDst=22}
action:drop

no → send along lowest-weight path from p.EthSrc to p.EthDst

S1: match:{EthSrc=p.EthSrc, EthDst=p.EthDst}
action:outPort(1)

S2: match:{EthSrc=p.EthSrc, EthDst=p.EthDst}
action:outPort(1)

S3: match:{EthSrc=p.EthSrc, EthDst=p.EthDst}
action:outPort(1)



Use Case IV: SDN

Current programming model

Policy

Programming

p.TcpDst=22?

yes → drop

no → send along lowest-weight path from p.EthSrc to p.EthDst

match:{TcpDst=22}
action:drop

priority: HIGH

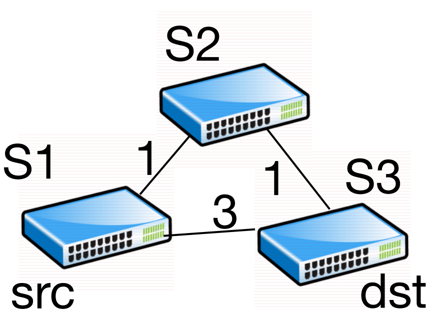
~~tcpDst!=22~~

S1: match:{EthSrc=p.EthSrc, EthDst=p.EthDst}
action:outPort(1)

priority: LOW

S2: match:{EthSrc=p.EthSrc, EthDst=p.EthDst}
action:outPort(1)

S3: match:{EthSrc=p.EthSrc, EthDst=p.EthDst}
action:outPort(1)

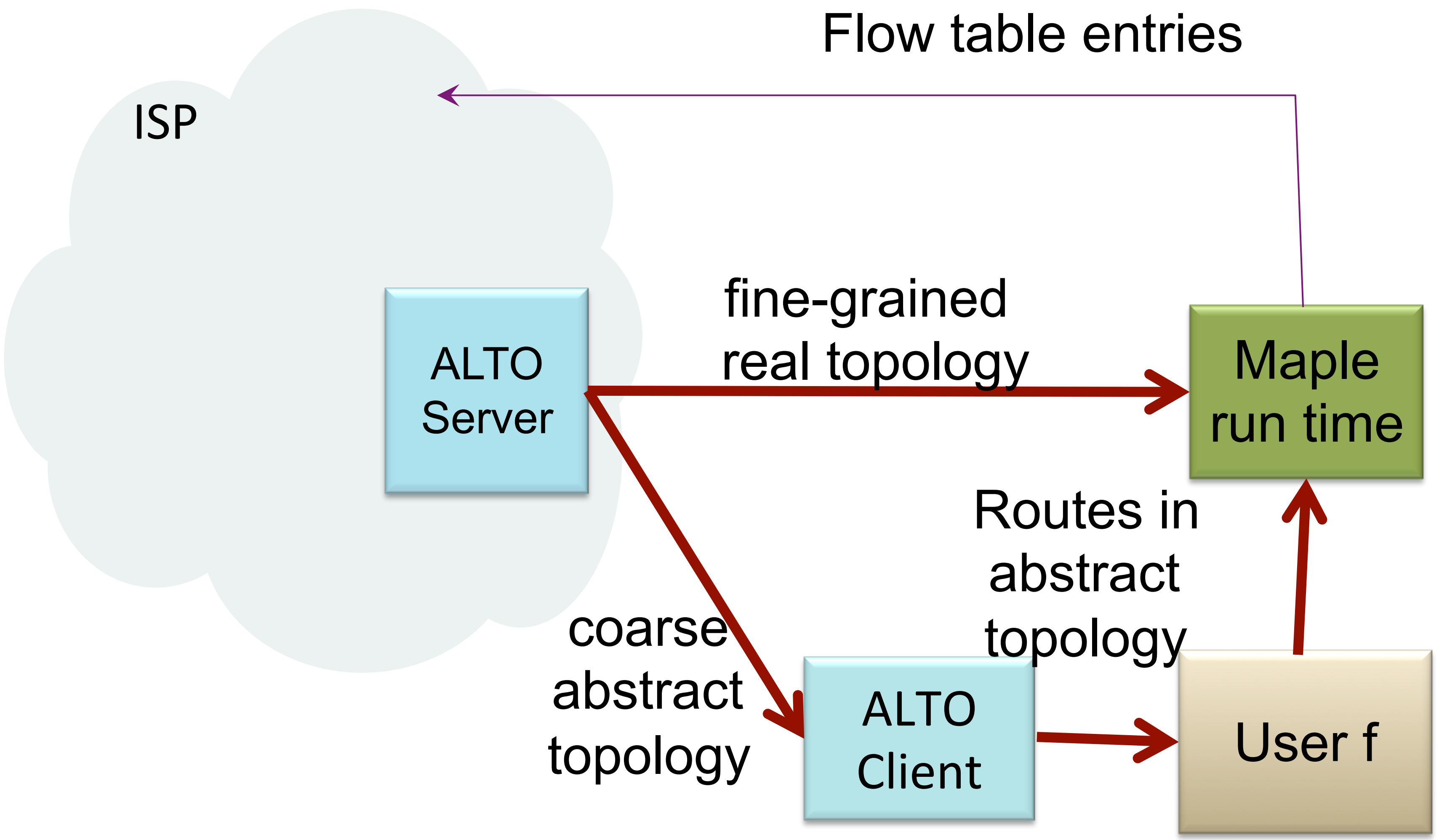


Use Case IV: SDN (Maple)

```
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);
  }
}
```

Maple: Simplifying SDN Programming using Algorithmic Policies, A. Voellmy, J. Wang, Y.R. Yang, B. Ford, and P. Hudak, in SIGCOMM'13.

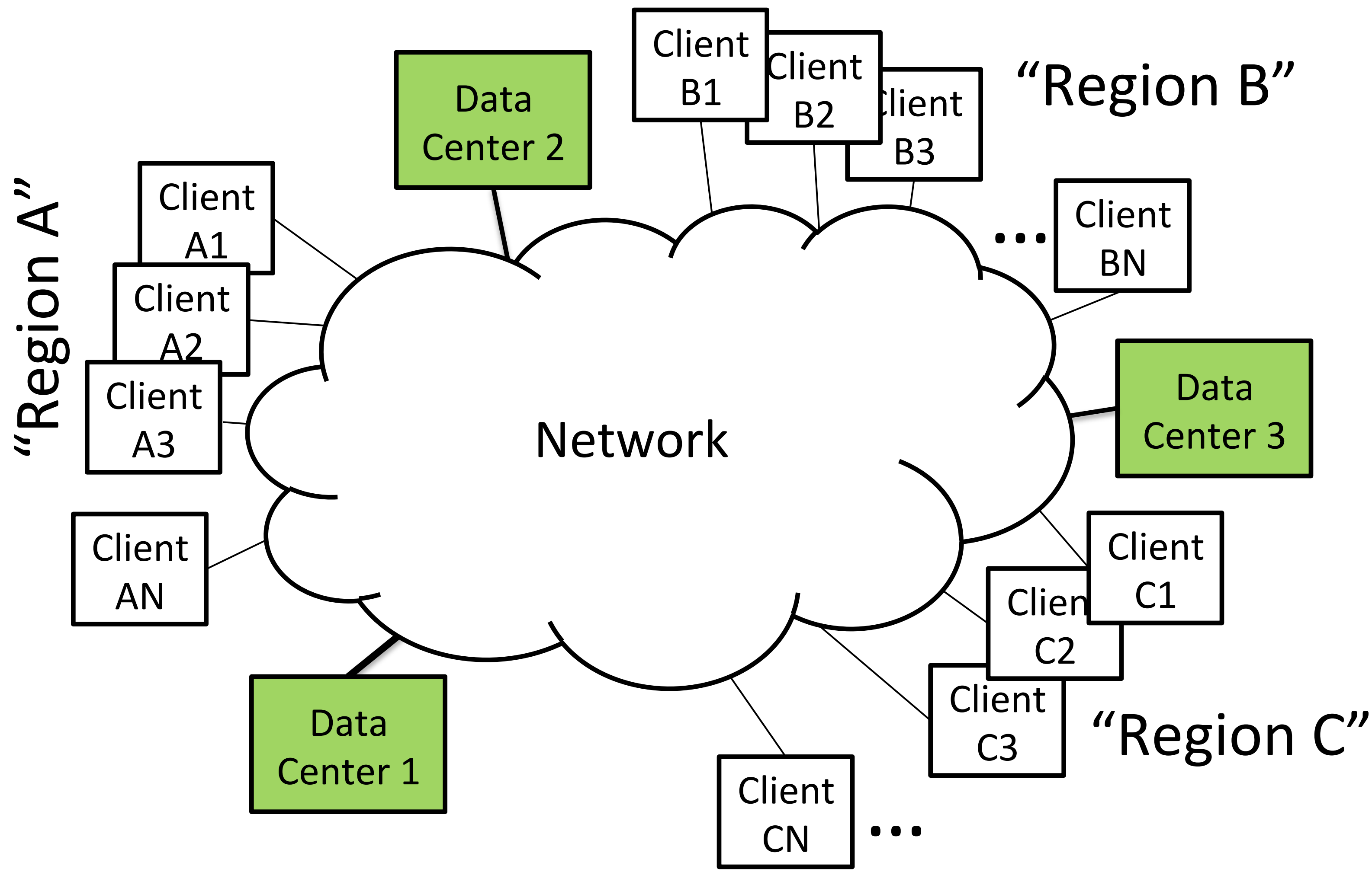
Use Case IV: Example Workflow



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).

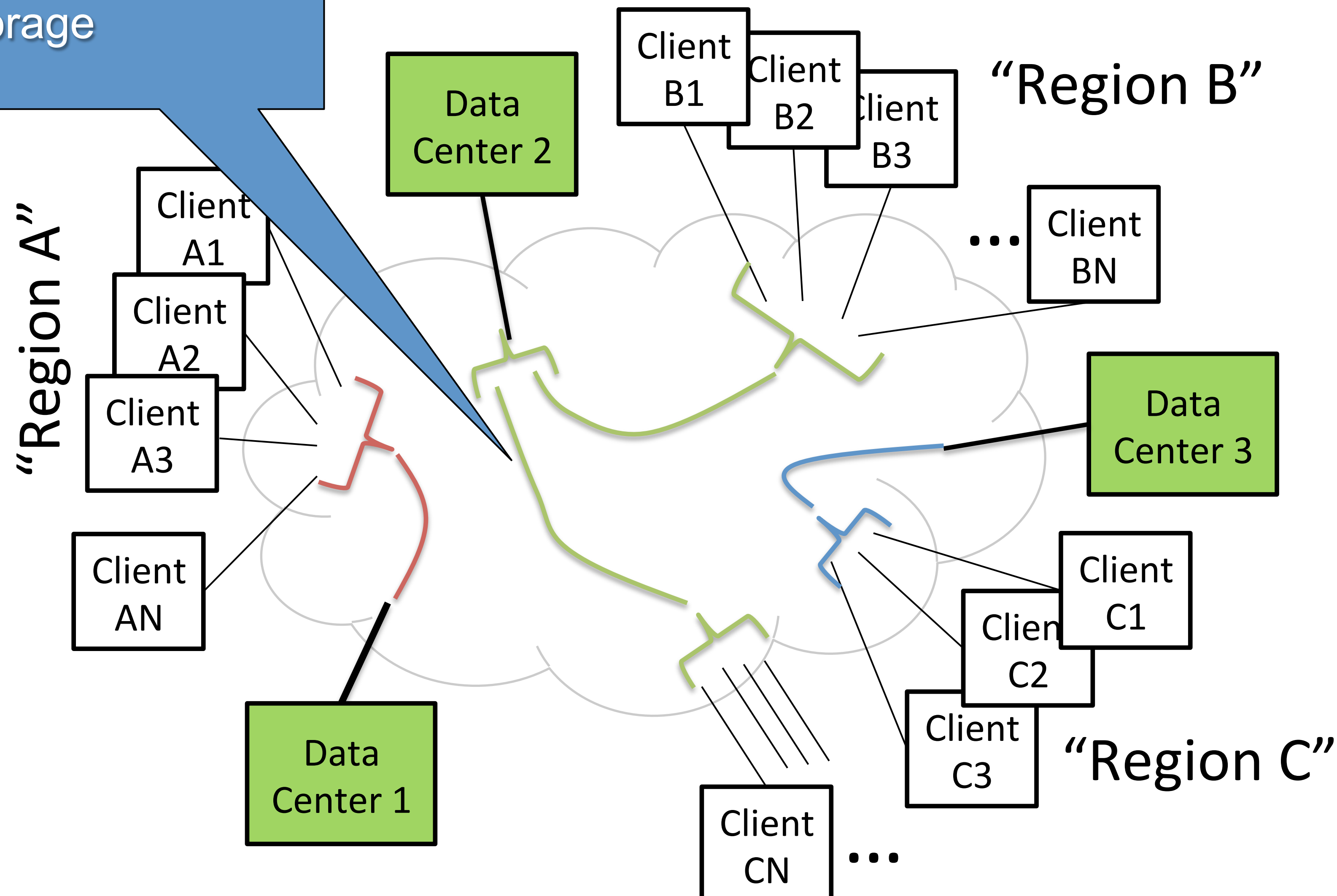
Use Case V: Example



Source: <draft-bernstein-alto-large-bandwidth-cases-01>

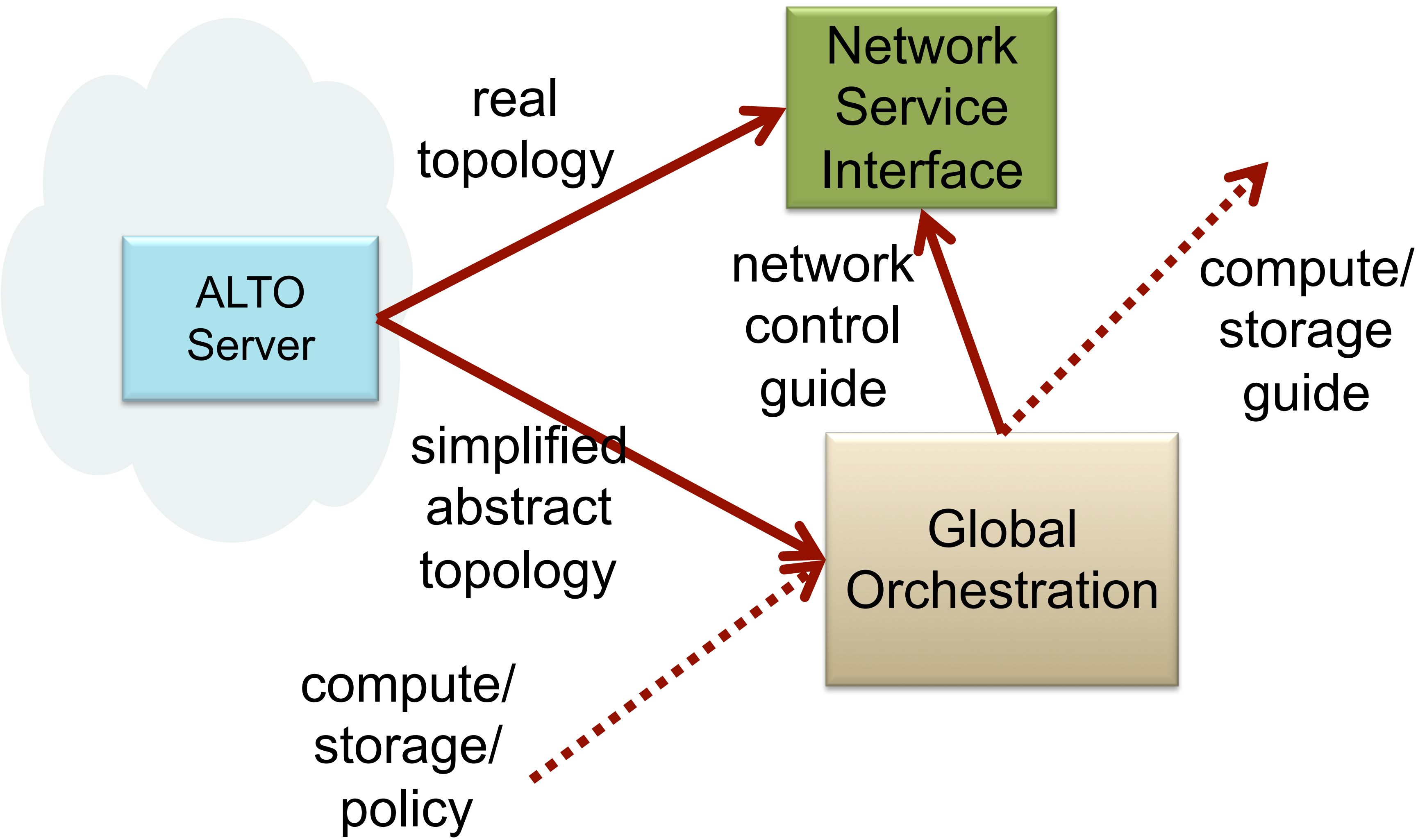
Use Case V: Example

Chosen paths considered both net and compute/ storage



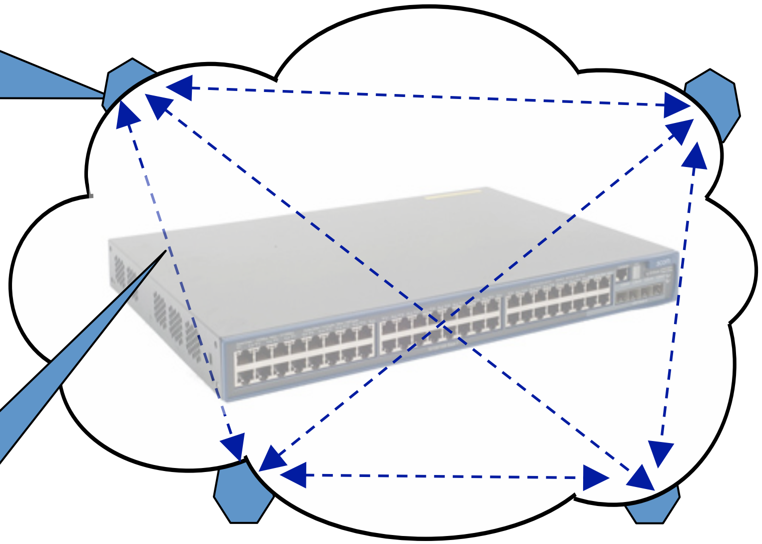
Source: <draft-bernstein-alto-large-bandwidth-cases-01>

Use Case V: Example Workflow



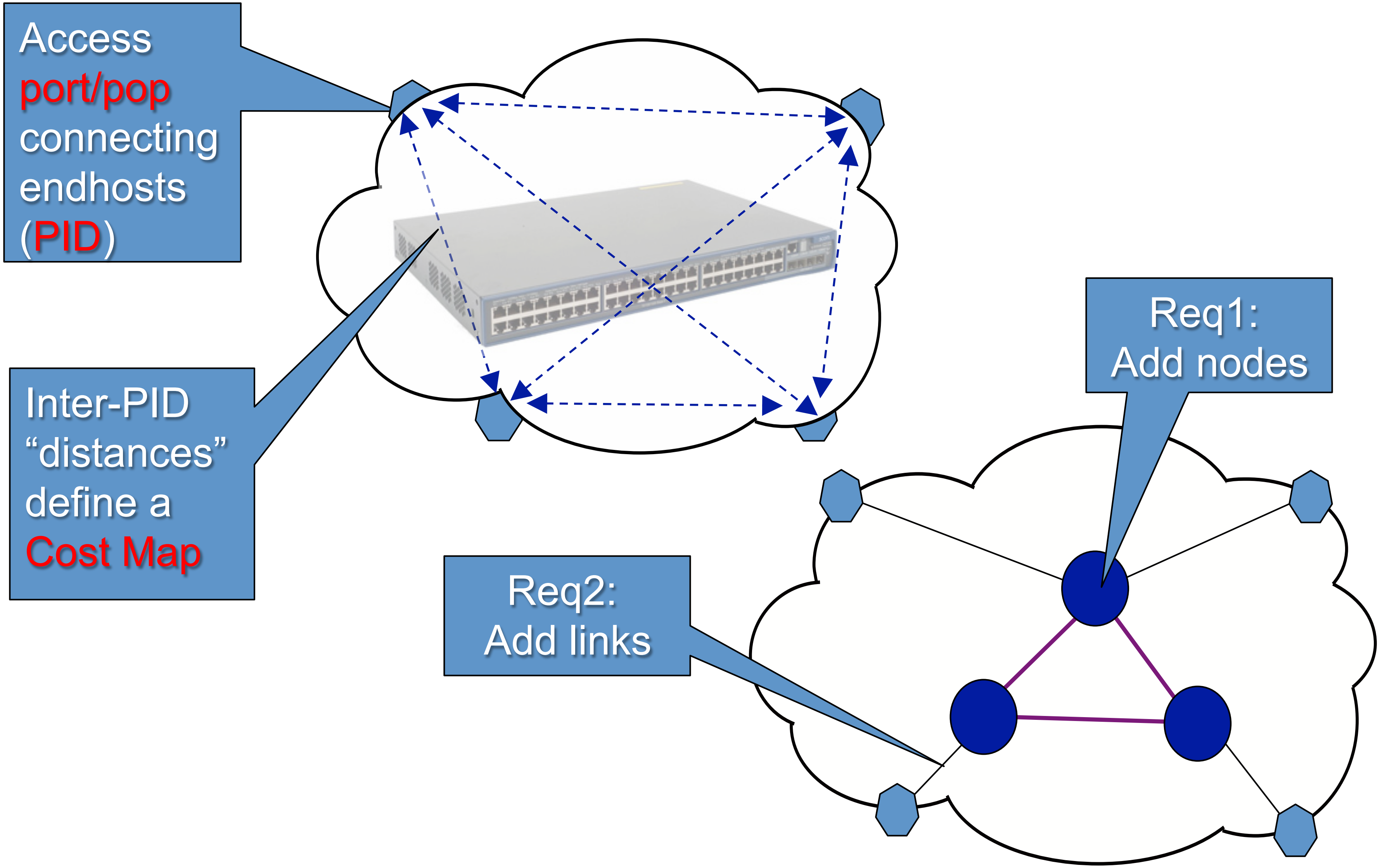
Basic Requirements:
Extend the Base ALTO "Single Switch" Abstraction

Access port/pop connecting endhosts (PID)



Inter-PID "distances" define a Cost Map

Basic Requirements:
Extend the Base ALTO "Single Switch" Abstraction



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
  - [I-D.lee-alto-app-net-info-exchange] & [I-D.yang-alto-topology] provide tentative models and encodings for 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

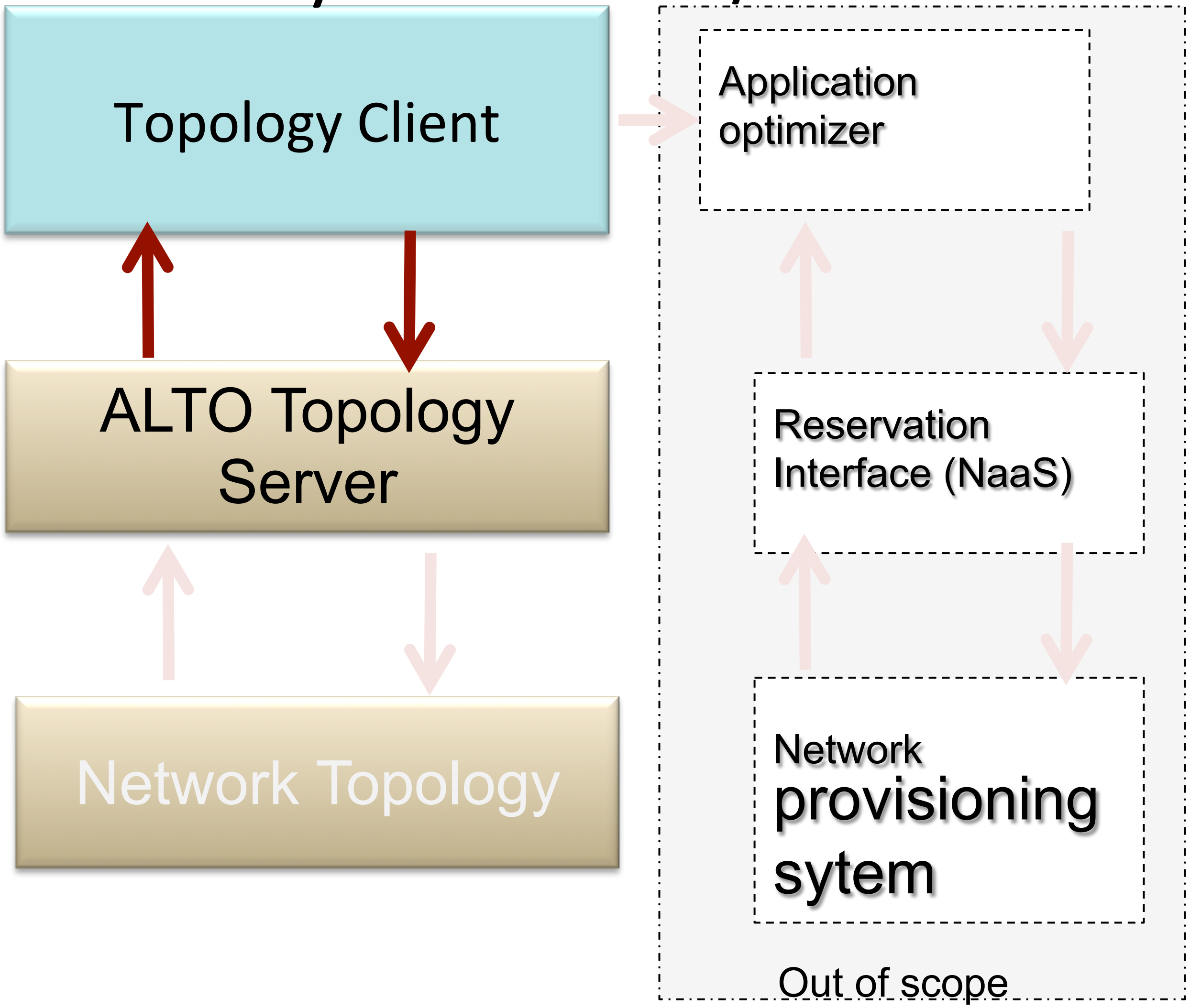
Next Step

- Specify a complete ALTO Topology Services Resource
- Propose WG re-chartering to include Topology Service and go beyond P2P

Backup Slides

Framework

- How how the service may interface w/ service interfaces



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.