Abstract Topology and Cost Maps for Software-Defined Inter-Domain Circuits

Malathi Veeraraghavan, Christian Esteve Rothenberg
University of Virginia (UVA), Charlottesville, VA 22904, USA
University of Campinas, Sao Paulo, Brazil

Oct. 15-16, 2015
SwitchON Workshop

Outline:

- BGP software complex; Simpler solution for SDN?
- New dynamic Layer-2 rate-guaranteed service
  - Interdomain - need topology or reachability exchanges
- Application Layer Traffic Optimization (ALTO)

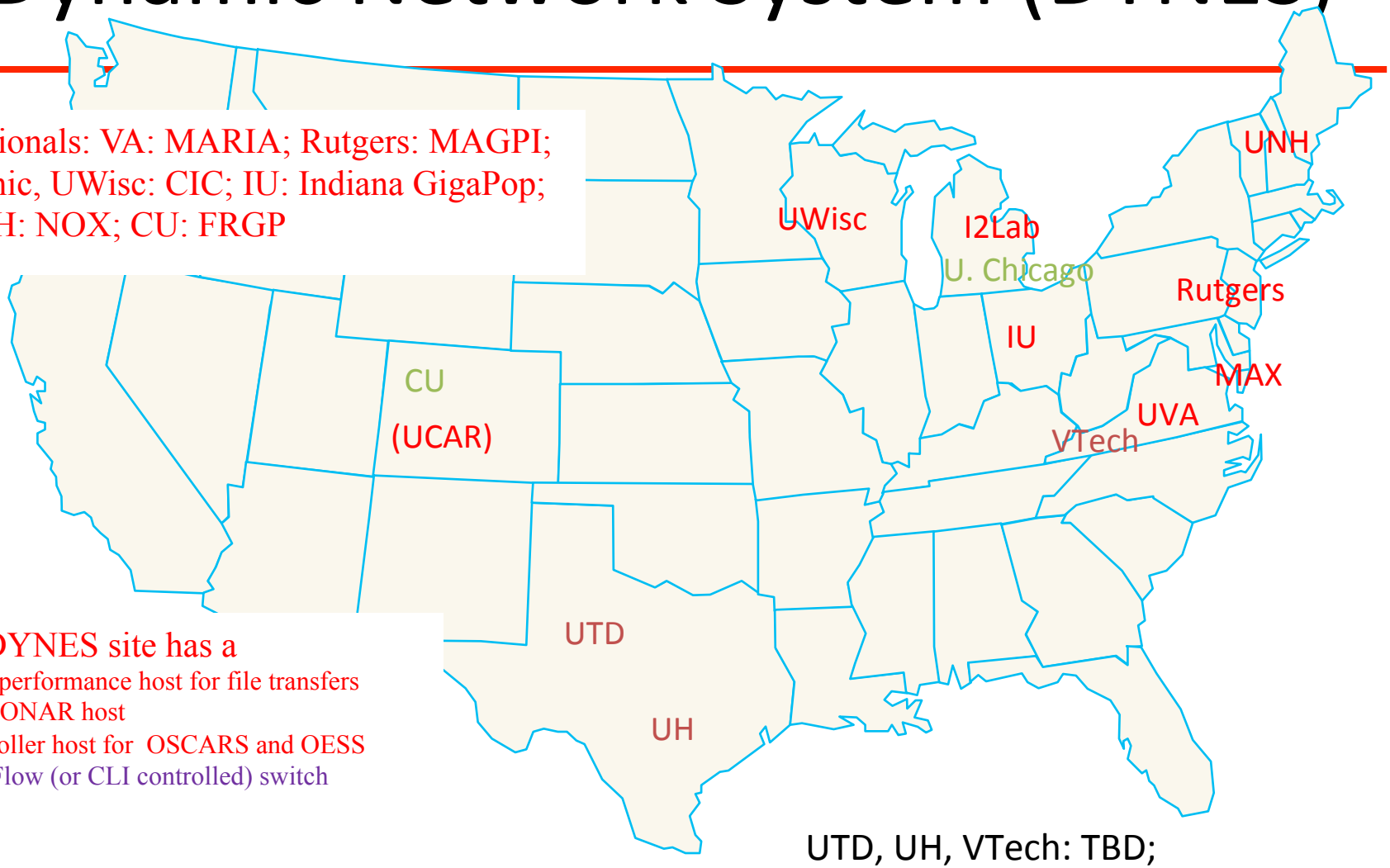
Acknowledgment

- UVA Graduate students: Fatma Al-Ali, Shuoshuo (Shawn) Chen, Xiang Ji, Sourav Maji, Fabrice Mizero, Reza Rahimi, Xiaoyu (Sherry) Wang
- NSF CC-NIE project collaborators: Steve Emmerson, (UCAR), Ivan Seskar, Steve Decker, Joe Slezak (Rutgers U), Jerrold Robaidek and Dale Carder (U. Wisc)
- DYNES effort collaborators: Brian Cashman, Eric Boyd (Internet2), A. J. Ragusa, Luke Fowler (IU), Chin Guok (ESnet), T. Lehman, X. Yang (MAX), Ezra Kissel (Indiana U), R. D. Russell and P. MacArthur (U. New Hampshire), Conan Moore (U.Colorado), and Ryan Harden (U. Chicago), Ron Withers (U. Virginia), John Lawson (MARIA), GRNOC, and several regional REN providers for their support.
- Thanks to NSF for CNS-1116081, OCI-1127340, ACI-1340910, and CNS-1405171, ACI-0958998, and DOE grant DE-SC0011358

Applications

- Slice isolation ala GENI; future Internets
- Large dataset transfers:
  - Instead of determining available capacity in the data-plane (TCP), reserve resources and blast
  - Predictable transfer delays for co-scheduling with compute resources (HPC jobs)
- Low-latency applications: networked robotics, remote instrument control (scientific), remote visualization - at layer-1, low-prop. delay paths

Dynamic Network System (DYNES)



Regionals: VA: MARIA; Rutgers: MAGPI;
UChic, UWisc: CIC; IU: Indiana GigaPop;
UNH: NOX; CU: FRGP

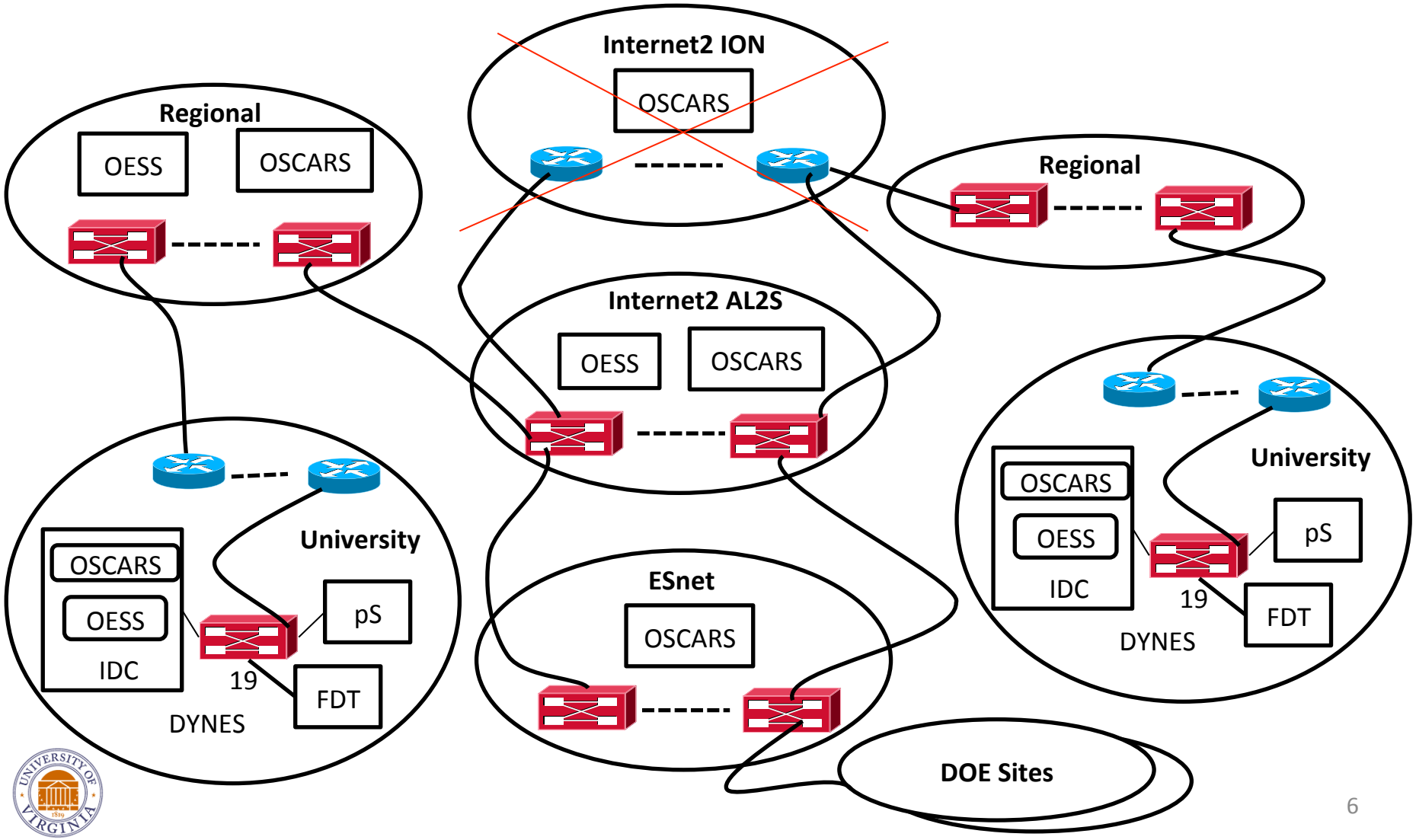
Each DYNES site has a
- high-performance host for file transfers
- perfSONAR host
- controller host for OSCARS and OESS
- OpenFlow (or CLI controlled) switch

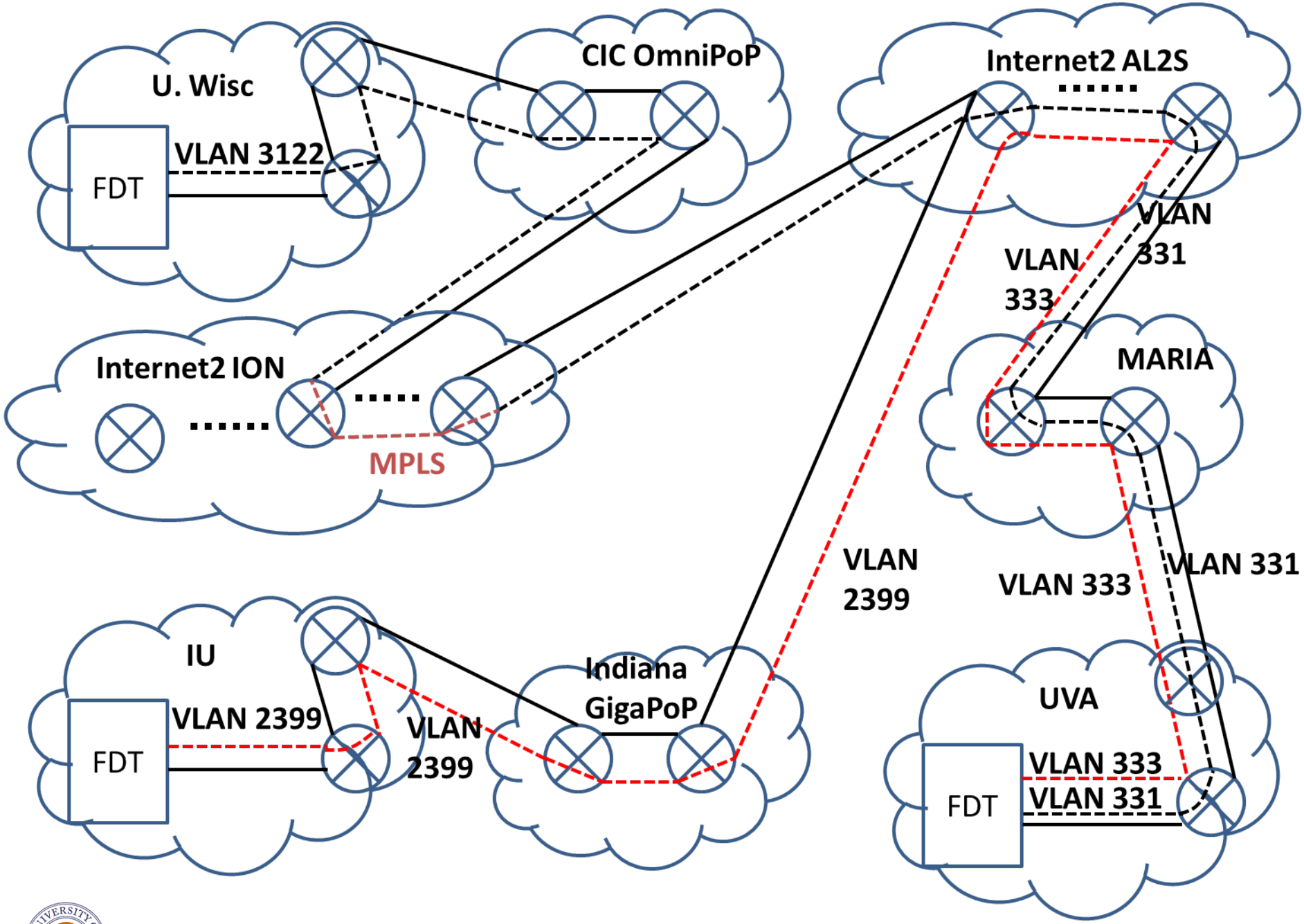
UTD, UH, VTech: TBD;
CU, Chicago: started

Multi-domain testbed

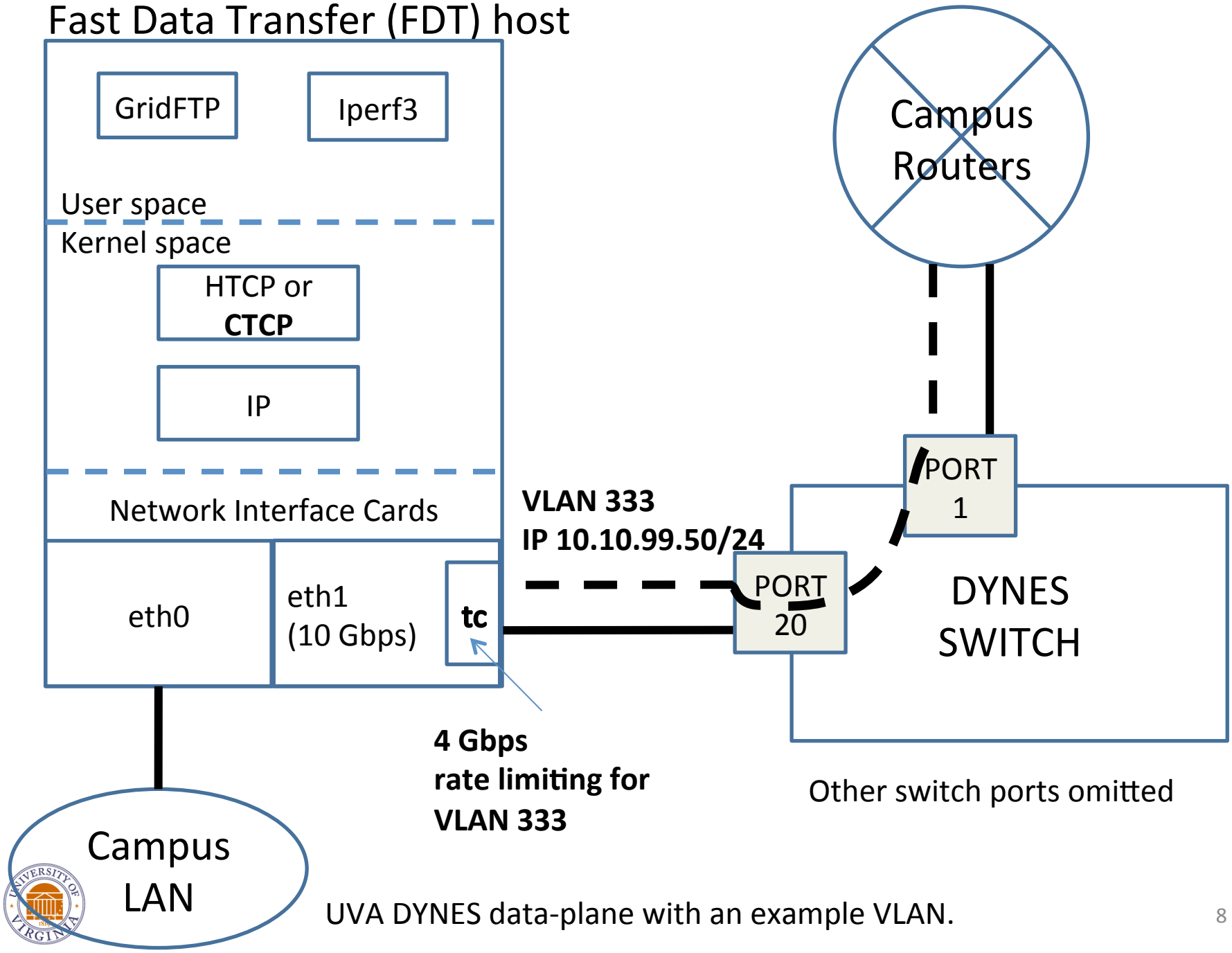
- DYNES: Dynamic Network System
  - Eric Boyd, Shawn McKee, Harvey Newman, Paul Sheldon: PIs
  - NSF MRI project : File Data Transfer (FDT) host + Switch (OpenFlow) + SDN Controller (IDC) + perfSONAR host
  - 40 universities and 11 regionals
- Configured DYNES equipment/software in 8 campuses
- Dynamically created **inter-domain** L2 paths via OESS GUI (running OSCARS on most DYNES IDCs)
- Configured FDT: vconfig, ifconfig, Linux tc
- Tested iperf3 and GridFTP: 0 loss?
  - Need Circuit TCP (CTCP)
  - Traffic Control (tc) Token Bucket Filter (TBF) at sender

Multi-domain deployment





Examples: End-to-end L2 paths between UVA and IU, and between UVA and UWisc.



UVA DYNES data-plane with an example VLAN.

Scalability
if ARPAnet --> Internet desired

- Topology approach:
  - OESS collects topology and passes to OSCARS
  - OSCARS pushes topology to pS Topology Service (pS-TS)
  - When another OSCARS receives a path-reservation request, it pulls the topology information from pS-TS in real-time
  - No information hiding between organizations
- OESS GUI shows all endpoints; Need DNS?
- AuthN/AuthZ:
  - Add DYNES to InCommon single-sign service?
  - GlobusOnline type service?
  - Scalability?

Control-plane models

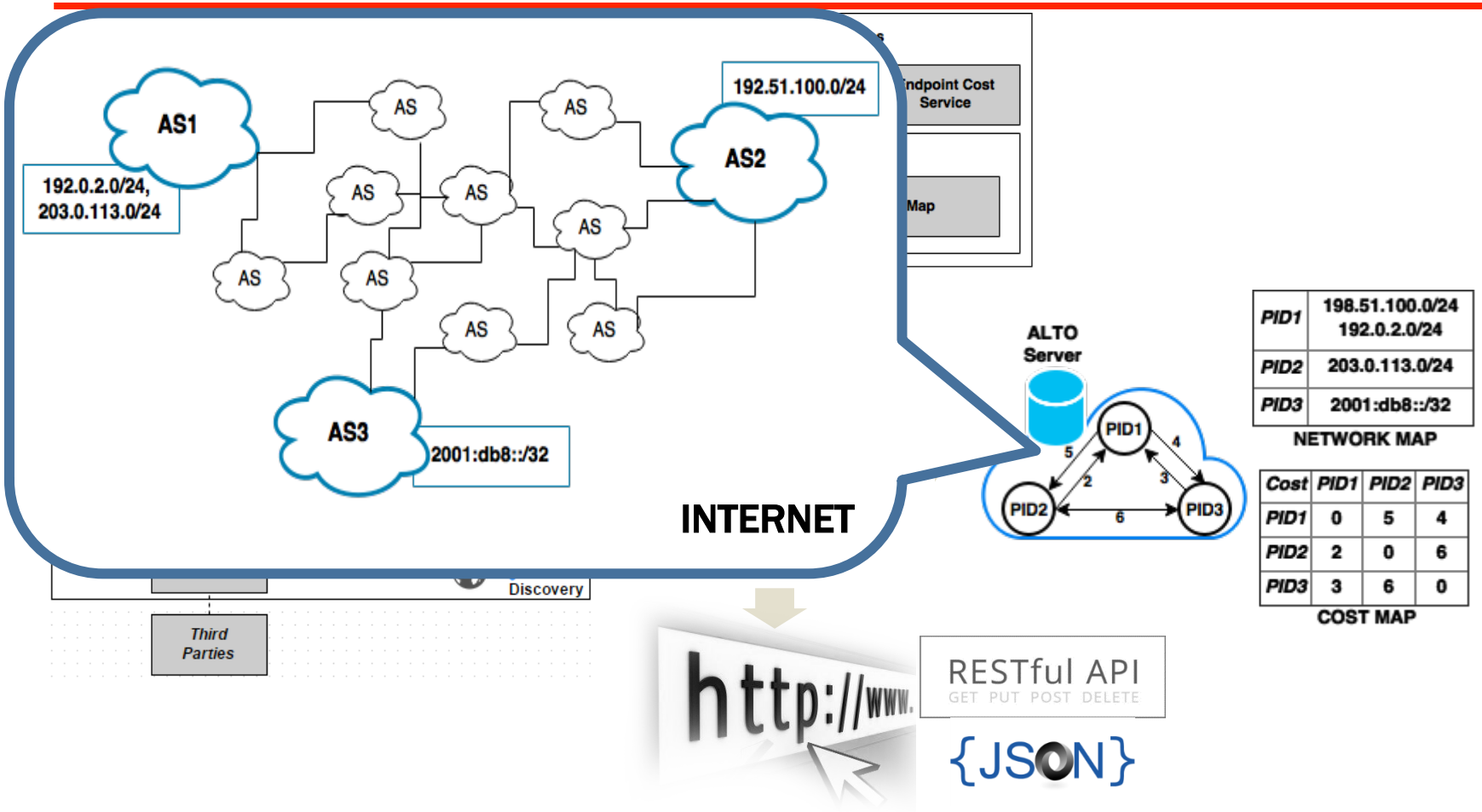
- Daisy-chain vs. tree-model
- Who owns and operates a multi-domain controller?
- Research literature and PCE IETF work
  - To avoid lockup of resources:
    - Daisy-chaining requires limited resource allocation on forward signaling path
    - Multiple start-time options to increase chance of success
    - Fast processing
- Tree-model AuthN needs?
  - Global PSTN: no customer-provider relationships required with providers more than two hops away in daisy-chain model. Not so in tree model
  - Testbed view (GENI) vs. ARPAnet growth view

International component

- Added Keio University, Yokohama, Japan
- Will similarly add a node in Brazil

Source	urn:ogf:network:domain=dynes.virginia.edu:node=dynes-sw1.dynes.virginia.edu:port=Te+0/24:link=*	
Destination	urn:ogf:network:domain=dcn.jgn-x.jp:node=kote-mx80-1:port=xe-0/0/0:link=starlight	
Path	VLAN	Hop
		urn:ogf:network:domain=dynes.virginia.edu:node=dynes-sw1.dynes.virginia.edu:port=Te+0/24:link=*
	338	urn:ogf:network:domain=dynes.virginia.edu:node=dynes-sw1.dynes.virginia.edu:port=Te+0/1:link=internet2
	338	urn:ogf:network:domain=al2s.net.internet2.edu:node=sdn-sw.ashb.net.internet2.edu:port=et-3/0/0.0:link=uva
	338	
	n/a*	urn:ogf:network:domain=al2s.net.internet2.edu:node=sdn-sw.ashb.net.internet2.edu:port=et-9/0/0.0:link=l2-ASHB-WASH-100GE-11823
	n/a*	
	3909	urn:ogf:network:domain=al2s.net.internet2.edu:node=sdn-sw.wash.net.internet2.edu:port=eth9/1:link=l2-ASHB-WASH-100GE-11823
	3909	
	n/a*	urn:ogf:network:domain=al2s.net.internet2.edu:node=sdn-sw.wash.net.internet2.edu:port=eth5/2:link=ION-AL2S-WASH
	n/a*	
	3121	urn:ogf:network:domain=ion.internet2.edu:node=rtr.wash:port=et-9/0/0:link=al2s
	3121	urn:ogf:network:domain=ion.internet2.edu:node=rtr.wash:port=et-9/0/0:link=et-9/0/0.201
	3121	urn:ogf:network:domain=ion.internet2.edu:node=rtr.losa:port=et-10/0/0:link=et-10/0/0.201
	3121	urn:ogf:network:domain=ion.internet2.edu:node=rtr.losa:port=xe-0/1/2:link=transpac
	n/a*	urn:ogf:network:domain=transpac.org:node=rtr.losa:port=xe-0/2/0:link=internet2
	n/a*	urn:ogf:network:domain=transpac.org:node=rtr.losa:port=xe-0/0/0:link=jgn-x
	n/a*	urn:ogf:network:domain=dcn.jgn-x.jp:node=kote-mx80-3:port=xe-0/0/0:link=transpac
	n/a*	urn:ogf:network:domain=dcn.jgn-x.jp:node=kote-mx80-3:port=xe-0/0/1:link=xe-0/0/1.0
		urn:ogf:network:domain=dcn.jgn-x.jp:node=kote-mx80-2:port=xe-0/0/2:link=xe-0/0/2.0

Application-Layer Traffic Optimization (ALTO)



PID1	198.51.100.0/24 192.0.2.0/24
PID2	203.0.113.0/24
PID3	2001:db8::/32

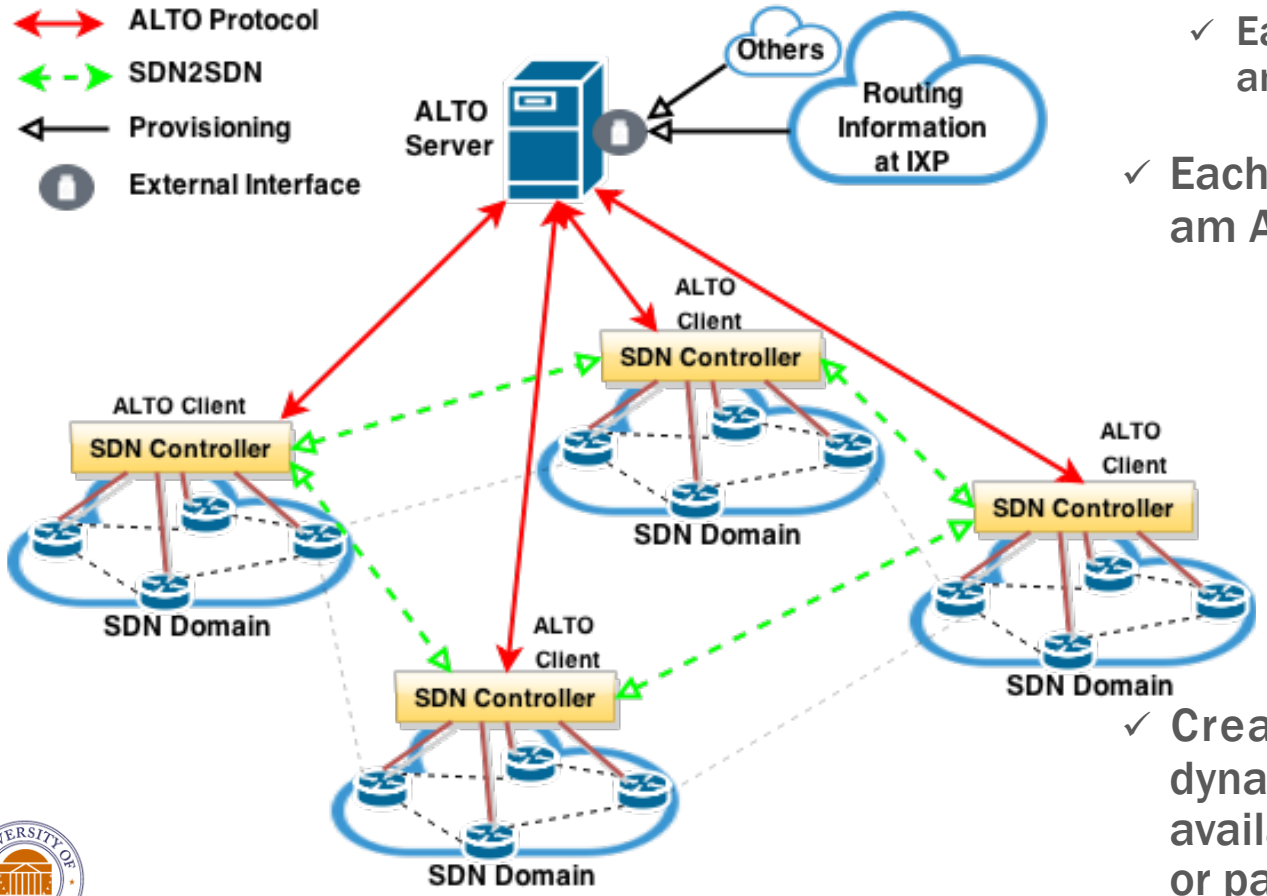
NETWORK MAP

Cost	PID1	PID2	PID3
PID1	0	5	4
PID2	2	0	6
PID3	3	6	0

COST MAP

Experimental Evaluation

- ALTO-SDN Use Case



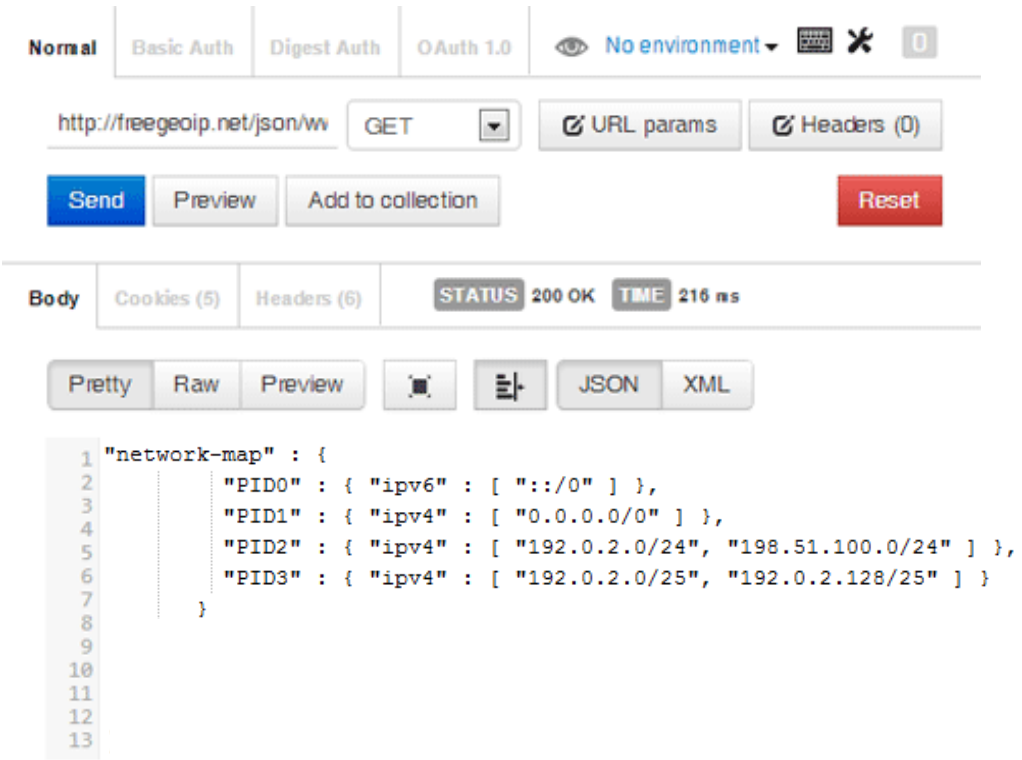
- ✓ A shared ALTO Server
  - ✓ Each SDN Controller acts an ALTO client
- ✓ Each SDN controller acts as am ALTO Server & Client
- ✓ Create Cost Maps with dynamic behavior such as available bandwidth, delay or packet loss rate.

Experimental Evaluation

Assess ALTO server delivering ALTO services in compliance with the RFC 7285.

- ✓ The network map service
- ✓ The cost map service
- ✓ The Filtered Map Service
- ✓ The Endpoint Cost Service

Postman - REST Client



Summary

- Lagopus and OVS to implement switches with support for QoS mechanisms: policing, scheduling
- Leverage Linux tc and DPDK features
  - Linux tc TBF and HTB queueing disciplines
- Interface OSCARS SDN controller for L2 circuits with ALTO server
- Test for scalability