ESnet SDN Experiences

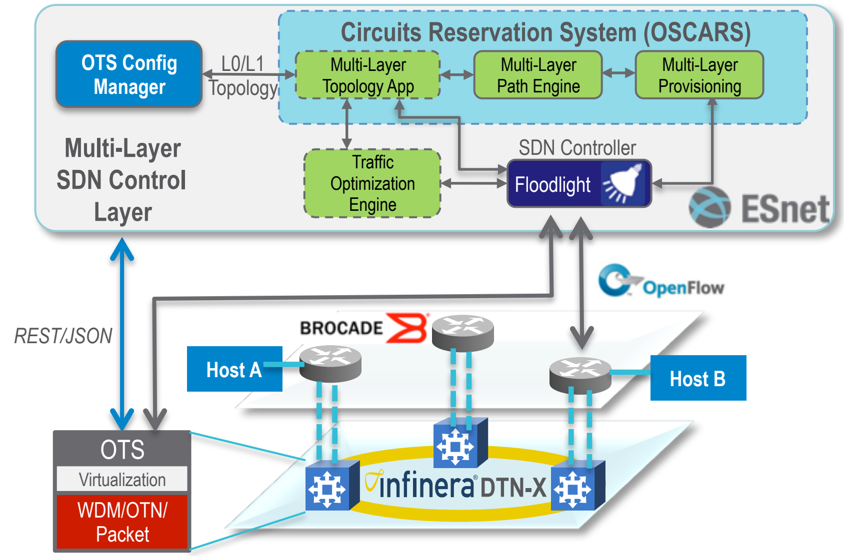
Roadmap to Operating SDN-based Networks Workshop
July 14-16, 2015
Berkeley, CA
C. Guok, B. Mah, I. Monga, E. Pouyoul

Things We Have Tried

- Multi-Layer SDN
  - Layer1 / Layer 2 modeling and provisioning
  - Layer2 / Layer 3 provisioning and routing
- SDX
  - BGP exchange and Layer 3 routing
- SDN Testbed
  - SDN deployment with overlay on production WAN infrastructure
- Network Operating System
  - ESnet Network Operating System (ENOS)

ESnet Multi-Layer SDN

- Unified layer 0-1 topology modeling and abstractions
- Dynamic hierarchical provisioning of layer 1-2 resources using OSCARS
- Ethernet transport SDN via Openflow
- Optical transport SDN (OTS) via Openflow (with extensions)



Impact

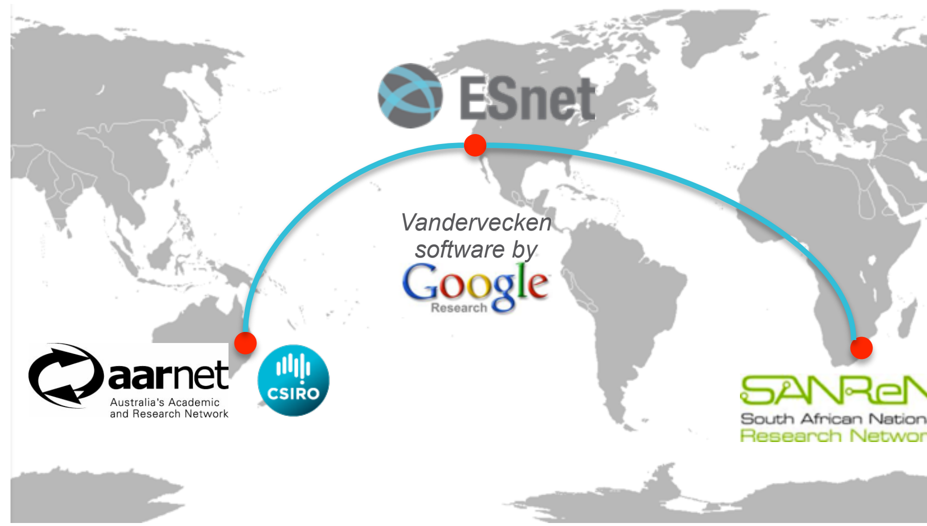
- Unified control of vendor and network layer agnostic resources using a standard protocol
- Flexible resource selection and programmability to meet complex service requirements, e.g. jitter-free, packet replication, etc
- Supports intelligent "layerless" networking decisions

Objective

- Cost savings
- Simplified control and management
- Enhanced services

ESnet SDX: Multi-continent SDN BGP peering

- Inter-operability with routing standards protocol (i.e. BGP) using SDN paradigms
- Physically distinct control plane (using off the rack Unix server) and data plane (using OpenFlow switch) functions
- Exchange of ~15K routes (R&E route table)
- Dynamic layer 2 setup (using OSCARS with NSI) for transport of layer 3 BGP protocol messages



Impact

- Replacement for classic single vendor monolithic router
- Programmable platform for custom routing functions without compromising backward compatibility with standard protocols
- Logically centralized control enabling intelligent decision making with the network as a single entity

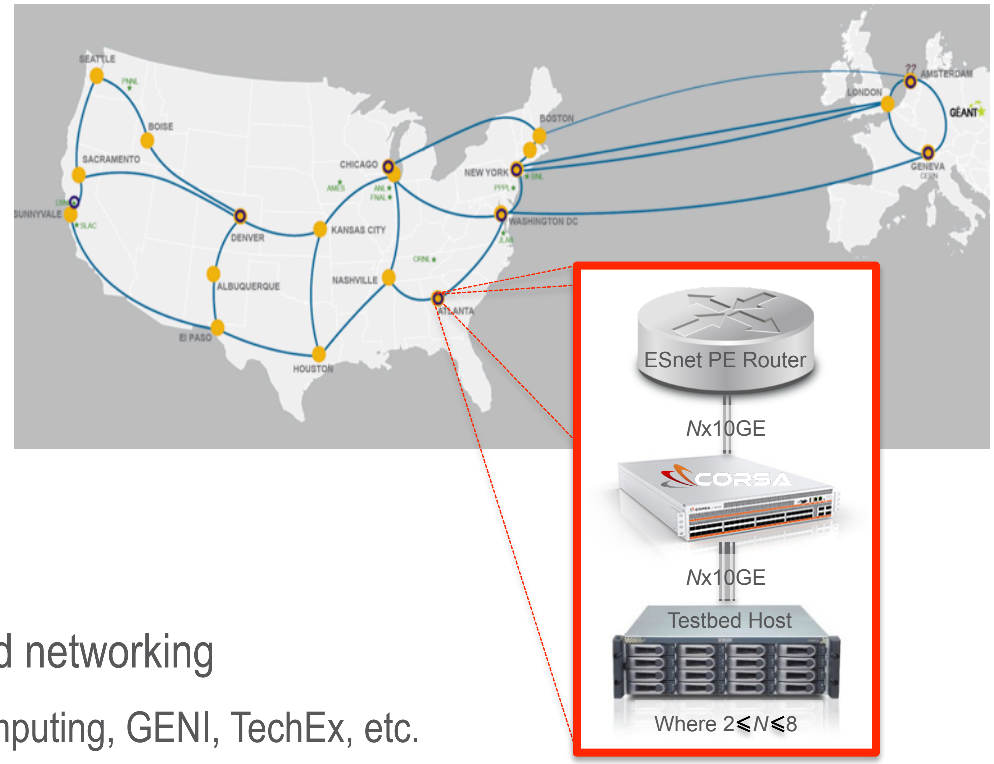
Objective

- Cost savings
- Simplified control and management

ESnet SANReN South African National Research Network

ESnet SDN Testbed

- 10G+ Overlay on ESnet5 via OSCARS circuits
  - Includes hardware at LBNL, NERSC, StarLight, Denver, Atlanta, Wash DC, NYC, Amsterdam, and Geneva
- Built using Corsa white boxes (www.corsa.com)
- Several test hosts, capable of running many VMs.
- Available to ESnet collaborators
- Ability to connect to other Testbeds using OSCARS



Impact

- Support ESnet SDN work: ENOS, Intent-based networking
- Support SDN demos of collaborators: Supercomputing, GENI, TechEx, etc.
- Support agile service creation and new SDN services such as multipoint-VPN service

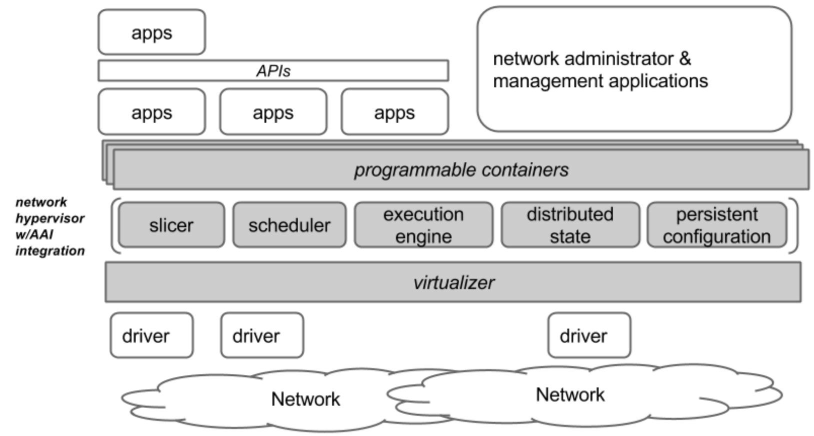
Objective

- Scalable and Flexible Architecture
- Support a wide range of SDN experiments
- Enough path diversity to do interesting experiments

ESnet

ESnet Operating System (ENOS)

- SDN platform for flexible service creation
- Multi-tenant operating system platform to support various science applications and customers
- Intent-based client interfaces and policy infrastructure to render intents to network provisioning (LDRD)
- SDN controller and vendor agnostic



Notional ESnet Network Operating System

Impact

- Programmable platform to build next-generation, custom ESnet services for science collaborations
- Network automation including programmatic provisioning of services by ESnet network engineers
- Potential new architecture based on white-box switches, potential for ESnet6 architecture

Objective

- New ESnet Architecture
- Cost savings
- Simplified control and management
- Enhanced services

Things We Have Learned*

- SDN Hardware
  - Not all OpenFlow switches are created equal
- SDN Management and Control
  - Control plane resiliency, consistency and scalability adds significant complexity
  - Secure in-band Application-Control Plane and Control-Data Plane is a necessity
- SDN Support
  - Monitoring and Troubleshooting can be painful
  - Commercial support model is unclear
- SDN (End-to-End) Dataplane
  - Multi-Domain SDN is necessary for a predictable end-to-end behavior

**These are our observations, your mileage may vary*

ESnet

SDN Hardware

- Flow table sizes range from several thousands to millions of entries, with different behavior
- Hardware design (ASIC, FPGA, CPU) are often driving software (e.g. single table vs multi-table)
- Switches are not supported equally in each controller, being OF compliant is not sufficient
- Production requirements such as graceful failure are not dictated by the OF and often overlooked
- More work needs to be done on QoS implementations (e.g. policing/shaping, multiple queues (4+), WRED)

SDN Management and Control

- Higher level intelligent functions should be centralized but “mundane” or “reactive” functions should be local (e.g. brain / body functions)
- Information consistency across a logically centralized control plane managing a very large geographic footprint network is hard
- Predicting failure behavior using a logically centralized control plane can be challenging, especially if the network is bi-partitioned
- Out-of-band application-control plane and control-data plane communications is cost prohibitive in a wide-area network, therefore secure in-band connections are a necessity
- Ability to provide different resource views/abstractions to different applications is important
- There must be support for standard protocols (at least for now) for interaction with external peers (e.g. eBGP)

ESnet

SDN Support

- Tools to troubleshoot issues are needed, e.g.
  - Check flow entry consistency between switch and controller
  - Determine flow rule overlaps
  - etc
- Switch local debugging is necessary (e.g. local counters, logs, etc), especially if you need to determine if the switch or controller is mis-behaving
- Separate commercial support of controller and Network Elements (NE) (i.e. switches) might be problematic (e.g. finger-pointing)

ESnet

SDN (End-to-End) Dataplane

- Networks are domain bound, but networking is end-to-end
- Multi-Domain SDN is necessary for a predictable end-to-end behavior (i.e. comparable service models)
- North-Bound, and East/West-Bound interfaces do not have to be distinct, but they may have to provide different resource views/abstractions
- Trust and policy control are critical to multi-domain SDN
- (Multi-domain) Topology modeling and sharing can get messy (esp. ensuring remote-IDs between peers)

Questions?

Contacts: chin, imonga, bmah, lomax @ es.net