Software Defined Services: Exchange Points (SDX) and ScienceDMZs (SD-SDMZ)

Internet2 Technology Exchange
September 26, 2016
Miami, Florida

Tom Lehman
Director of Research
UMD/MAX

Xi Yang
Senior Research Scientist
UMD/MAX
Team and Research Projects

- **UMD/MAX Team**
  - Tom Lehman
  - Xi Yang
  - Alberto Jimenez
  - Multiple Students

- **Results from several research projects including:**
  - High Performance Computing with Data and Networking Acceleration (HPCDNA)
  - Resource Aware Intelligent Network Services (RAINS)
  - GENI Enabled Software Defined Exchange (SDX)
Today’s Topics

• Software Defined ScienceDMZ (SD-SDMZ)
• Software Defined Exchange (SDX)
• Deployments at UMD and the Washington International Exchange (WIX)
• Services provisioning using the model-based multi-resource orchestration VersaStack Software developed by UMD/MAX team
• “MAX AWS Service” available at WIX
• Hybrid Cloud Service Provisioning Demonstration
Why Software Defined?

• Researchers are increasingly interested in an emerging class of hybrid services that integrate on-campus instrumentation and infrastructure, regional and national cyberinfrastructure, and elastic private and public clouds, over high performance networks.

• Research and Education (R&E) Cyberinfrastructure needs to evolve towards an “as a Service” model based on Software Defined Infrastructure (SDI)

• This transformation needs to address the unique requirements of the domain science communities

• Technologies and open source from the commercial focused Software Defined Network (SDN), Cloud, and Data Center communities can be leveraged, but alone are insufficient
Software Defined ScienceDMZ

- Traditional SDMZ: Bare metal Data Transfer Nodes (DTNs, perfSONAR nodes, manual switch/router control)
- SD-SDMZ: Local Compute Cluster, OpenStack, Ceph, SDN Network Control. On Demand, scalable, traditional services and advanced hybrid cloud services
• Brocade MLXe
  ➢ 100G, OpenFlow, OpenDaylight SDN
• Cisco Unified Computing System (UCS)
  ➢ Integrated Compute and SR-IOV Networking
• Ceph High Performance Storage System
Advanced Hybrid Cloud (AHC) Service

On Demand, Application Specific, Hybrid Topologies which include one of more of the following:

✓ Local OpenStack Virtual Machines (with SRIOV interfaces to network and storage)
✓ Dedicated Local Ceph Storage Resources and Connections
✓ Integrated AWS Resources (Virtual Private Cloud (VPC) or Public)
  • User controlled AWS resources, or
  • SD-SDMZ facility provided AWS resources (EC2, Storage, S3 endpoints)
✓ Network Connections
  • AWS Direct Connect integration for access to AWS Public or VPC resources
  • Layer2/Layer2 External Connections across Internet2, ESnet, others
  • Customized topology templates for individual user requirements
✓ Future:
  • Service connections/integration with other public cloud infrastructures
  • Service connections/integration with other R&E cloud, HPC, data repositories, etc.
  • Schedulable Services
UMD SD-SDMZ Services

Data Transfer Node (DTN) Service

SDMZ DTN service for data movement to/from HPC and other systems

✓ Built using AHC Service (Local OpenStack Virtual Machines with SRIOV interfaces to network and Ceph storage)

✓ Globus Endpoints

✓ Dataplane integration with HPC file systems (via IB/Ethernet Gateway)

✓ HPC System compute nodes mount SD-SDMZ CephFS

✓ On-Demand Scalable DTN infrastructure (dedicated DTN nodes on a per project or user basis available)
Hybrid Topology Building

- Orchestration across a diverse set of resources can be challenging. As an example to build this:

- Required steps include:
  - Instantiate tenant VMs and networks on local OpenStack, attach SR-IOV interfaces to local VMs, create dedicated Ceph images for VM mount
  - Configure AWS Direct Connect for proper VPC access
  - Provision a Layer2 path across MAX regional network to AWS
  - Instantiate AWS VPC resources
  - Instantiate a local VM with BGP configured for AWS peering
  - Configure proper private IP addressing and external gateway functions
Our Approach and Solution

• **Multi-Resource Orchestration:** integrating and orchestrating the network and network services with the things that attach to the network – compute, storage, clouds, and instruments.

• **Model Driven:** using models to describe resources in order to allow integrated reasoning, abstraction, and user centered services

• **Intelligent Computation Services:** Model driven, multi-resource computation services to enable orchestration services in response to high level user requests.

• **We want to “Orchestrate the Automaters”**
Model Driven Orchestration

• Modeling schemas based on Network Markup Language (NML).
• Developed extensions to allow resources which are connected to the network to also be modeled: Multi-Resource Markup Language (MRML)
  – https://github.com/MAX-UMD/nml-mrml
• Utilizes Resource Description Framework (RDF) and OWL 2 Web Ontology Language W3C Specifications
VersaStack Software

- UMD/MAX developed VersaStack is an open source model driven orchestration system:
  - github.com/MAX-UMD/versastack.community
  - video: tinyurl.com/umdmax-versastack
- Native VersaStack Application (Northbound) API
- Access via GENI Aggregate Manager API
- Multiple Drivers, Southbound APIs
Software Defined Exchange (SDX)

- Same technology base can be used for SDX
- WIX is a production Exchange Point in McLean, Virginia (jointly operated by Internet2 and MAX)
- Includes OSCARS service enabling Dynamic Cross Connects
- MAX has made its AWS Direct Connect Service available at the WIX via Layer2 VLAN service
- MAX runs a GENI AM/ VersaStack instance with OSCARS and AWS drivers
WIX as a GENI Enabled SDX

• Would like be able to control Exchange Point resource utilization, in an automated fashion:
  – at the Federation (Clearinghouse) level, Virtual Organization (Project) Level, Slice Level, and User Level
  – also need to be able to adjust authorizations and access polices in near-real time
• Current SDX Functionality
  – Leverages GRAM GENI AM based policy features
  – Establish resource quotas on a Clearinghouse, Slice, or User basis
  – Enforce allocation by total bandwidth, number of VLANs in use
• Future Capabilities Desired
  – Additional parameters available for resource access
    • Interfaces, VLAN Ranges
    • at the Federation (Clearinghouse) level, Virtual Organization (Project) Level, Slice Level, and User Level
  – Ability for real-time resource utilization adjustments based on user priority and preemption
  – SDXs with compute and storage embedded
“MAX AWS Service” at WIX

• Authorized users can provision a Layer 2 path thru WIX and utilize the MAX AWS Direct Connection to the US East-1 Region

• Multiple options for Layer 2 provisioning to and thru WIX for this service:
  – GENI Stitching
  – AL2S via OESS
  – OSCARS/NSI Service

• Depending on use case, there may be several advantages to accessing AWS via Direct Connect, and also to connecting to a specific AWS region

• Contact MAX for further information if interested in trying this service
Example SD-SDMZ Use Cases

• Global Land Cover Facility (GLCF)
  – Hybrid cloud topology to facilitate data download from R&E and AWS S3 locations to local HPC filesystem

• Pan-STARRS Astronomy
  – Local compute/storage resources to facilitate download and inline processing of telescope data

• Large Scale Geosimulation
  – Hybrid cloud topology to facilitate Hadoop cluster set up with local nodes and scalable bursting in to AWS
Demonstration

Advanced Hybrid Cloud Service

Layer 2 Path across MAX

Layer 3 Path

AWS VPC

Local OpenStack VMs with Ceph Storage
Thanks