



ExoGENI Racks GEC15 status update

Ilia Baldine ibaldin@renci.org

Chris Heermann ckh@renci.org

Jonathan Mills jonmills@renci.org

Victor Orlikowski vjo@cs.duke.edu

and of course

Jeff Chase chase@cs.duke.edu

ExoGENI Overview

Presentation title goes here

ExoGENI Testbed

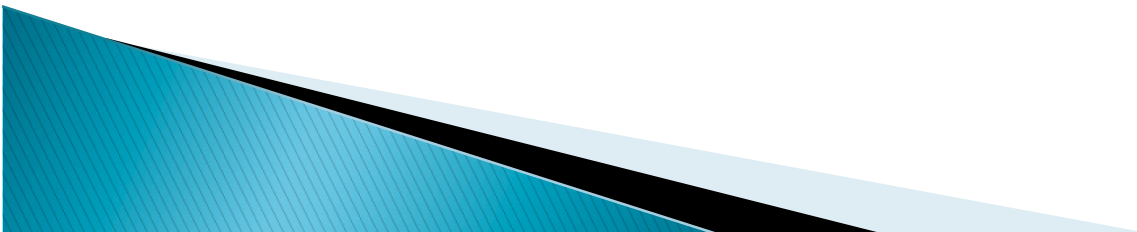


- ▶ 14 GPO-funded racks
 - Partnership between RENCi, Duke and IBM
 - IBM x3650 M4 servers (X-series 2U)
 - 1x146GB 10K SAS hard drive + 1x500GB secondary drive
 - 48G RAM 1333Mhz
 - Dual-socket 8-core CPU
 - Dual 1Gbps adapter (management network)
 - 10G dual-port Chelseo adapter (dataplane)
 - BNT 8264 10G/40G OpenFlow switch
 - DS3512 6TB sliverable storage
 - iSCSI interface for head node image storage as well as experimenter slivering
- ▶ Each rack is a small networked cloud
 - OpenStack-based
 - EC2 node sizes (m1.small, m1.large etc)
- ▶ <http://www.exogeni.net>

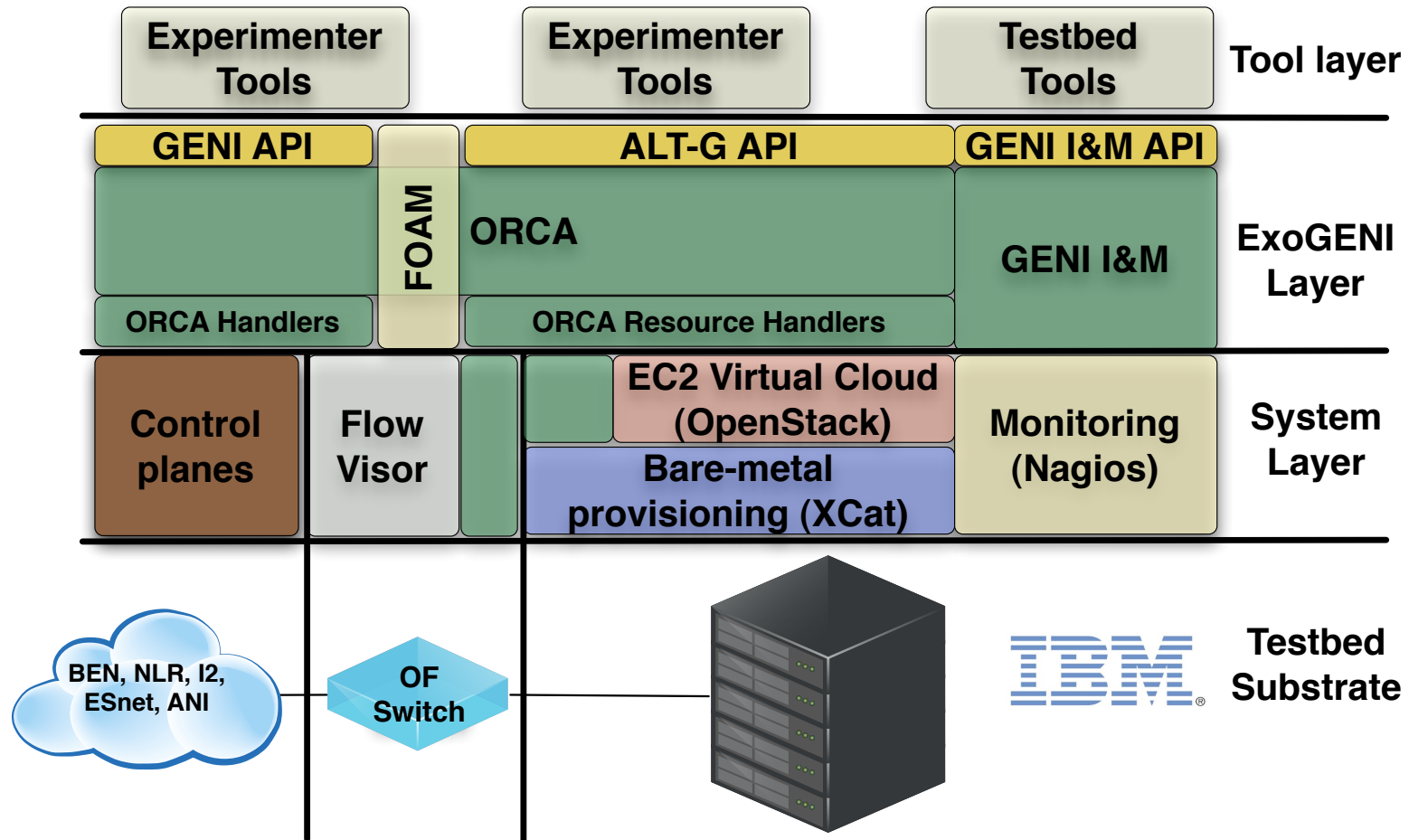


ExoGENI

- ▶ ExoGENI is a collection of off-the shelf institutional clouds
 - With a GENI federation on top
 - xCAT – IBM product
 - OpenStack– RedHat product
- ▶ Operators decide how much capacity to delegate to GENI and how much to retain for yourself
- ▶ Familiar industry–standard interfaces (EC2)
- ▶ GENI Interface

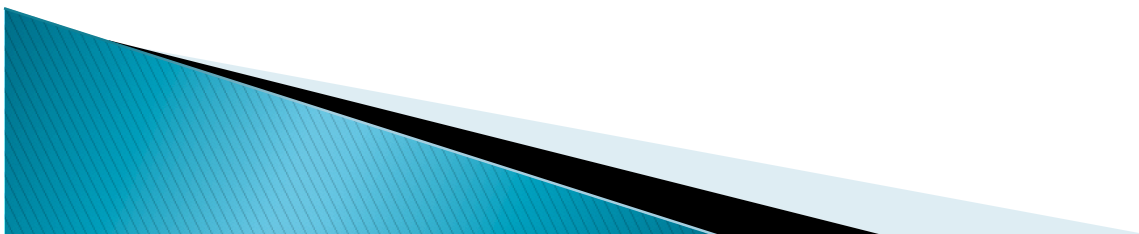


Rack Software Stack

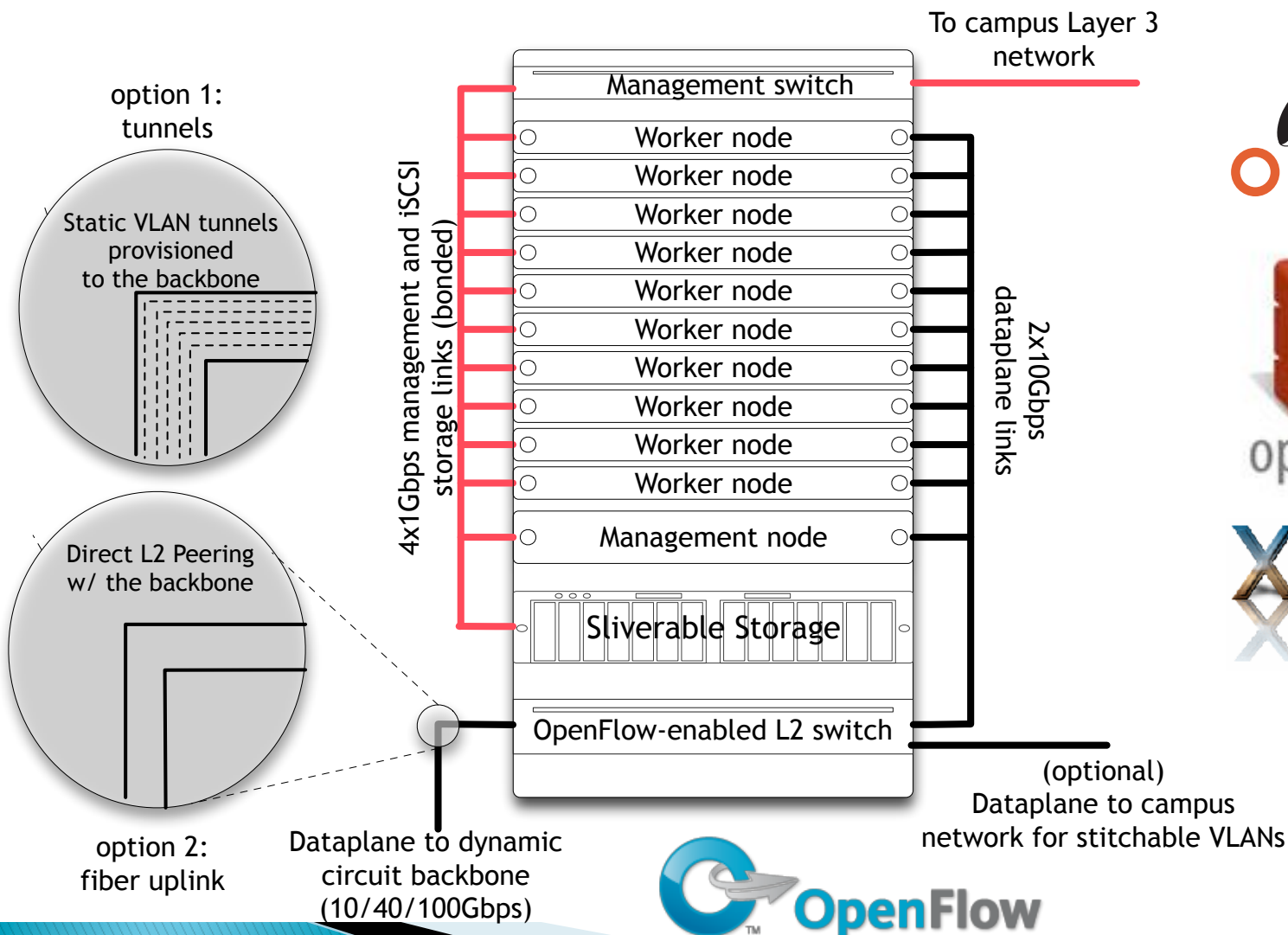


Rack software

- ▶ CentOS 6.X base install
- ▶ Resource Provisioning
 - xCAT for bare metal provisioning
 - OpenStack + NEuca for VMs
 - FlowVisor
 - Floodlight used internally by ORCA
- ▶ GENI Software
 - ORCA for VM, baremetal and OpenFlow
 - FOAM for OpenFlow experiments
- ▶ Worker and head nodes can be reinstalled remotely via IPMI + KickStart
- ▶ Monitoring via Nagios (Check_MK)
 - ExoGENI ops staff can monitor all racks
 - Site owners can monitor their own rack
- ▶ Syslogs collected centrally



An ExoGENI cloud “rack site”



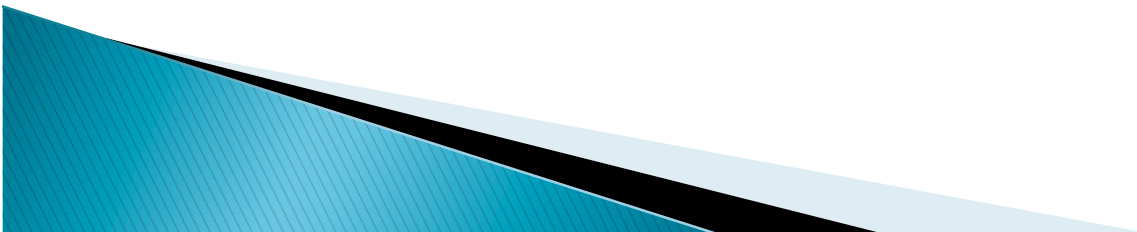
ExoGENI Status

- ▶ 3 new racks deployed
 - RENCi, GPO and NICTA
- ▶ 2 more racks delivered, being configured
 - FIU and UH
- ▶ Connected via BEN (<http://ben.renci.org>), LEARN and NLR FrameNet, (eventually I2)



Since GEC14

- ▶ Improvements in stability (in support of OpenStack/Diablo and OpenFlow)
- ▶ GMOC reporting
- ▶ GENI AM API compliance
- ▶ Acceptance testing



ExoGENI unique features

- ▶ ExoGENI racks are separate aggregates but also act as a single aggregate
 - Transparent stitching of resources from multiple racks
- ▶ ExoGENI racks are targeted at computing applications as well as experimentation
 - Already running HPC workflows linked to OSG and national supercomputers
 - Strong performance isolation is one of key goals
- ▶ A model for deeply reconfigurable federated compute/storage/network infrastructure for campuses and labs.



ExoGENI near term enhancements

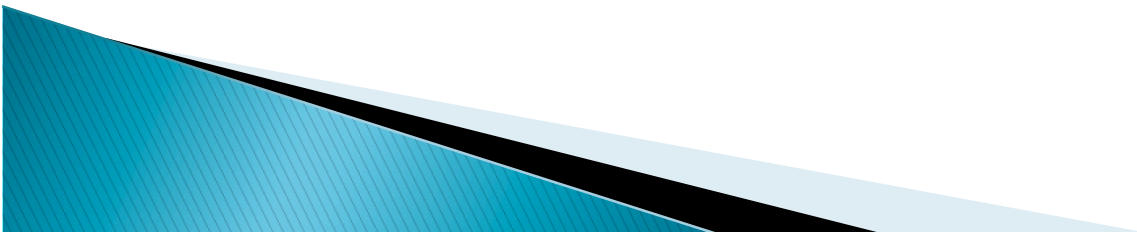
- ▶ Upgrade from OpenStack/Diablo to OpenStack/Essex-Folsom (the RedHat distribution)
 - With NEuca enhancements to support adding removing network interfaces on VMs at run-time
 - <http://www.networkedclouds.net>
 - Uses OVS and a custom Quantum plugin
 - Support for link QoS using OVS
- ▶ Expected OpenFlow firmware upgrade
 - Support for hybrid mode
 - Native VLAN provisioning and OpenFlow
 - Will improve scalability

ExoGENI details

Presentation title goes here

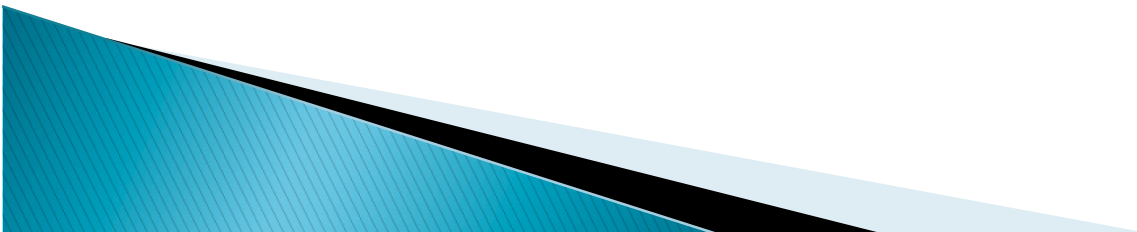
Rack Connectivity

- ▶ Rack has a management connection to campus network
- ▶ A connection to FrameNet or I2 ION
 - Direct peering
 - Via a pool of vlans with static tags through a regional
- ▶ It may have an optional connection to the OpenFlow campus network for experiments

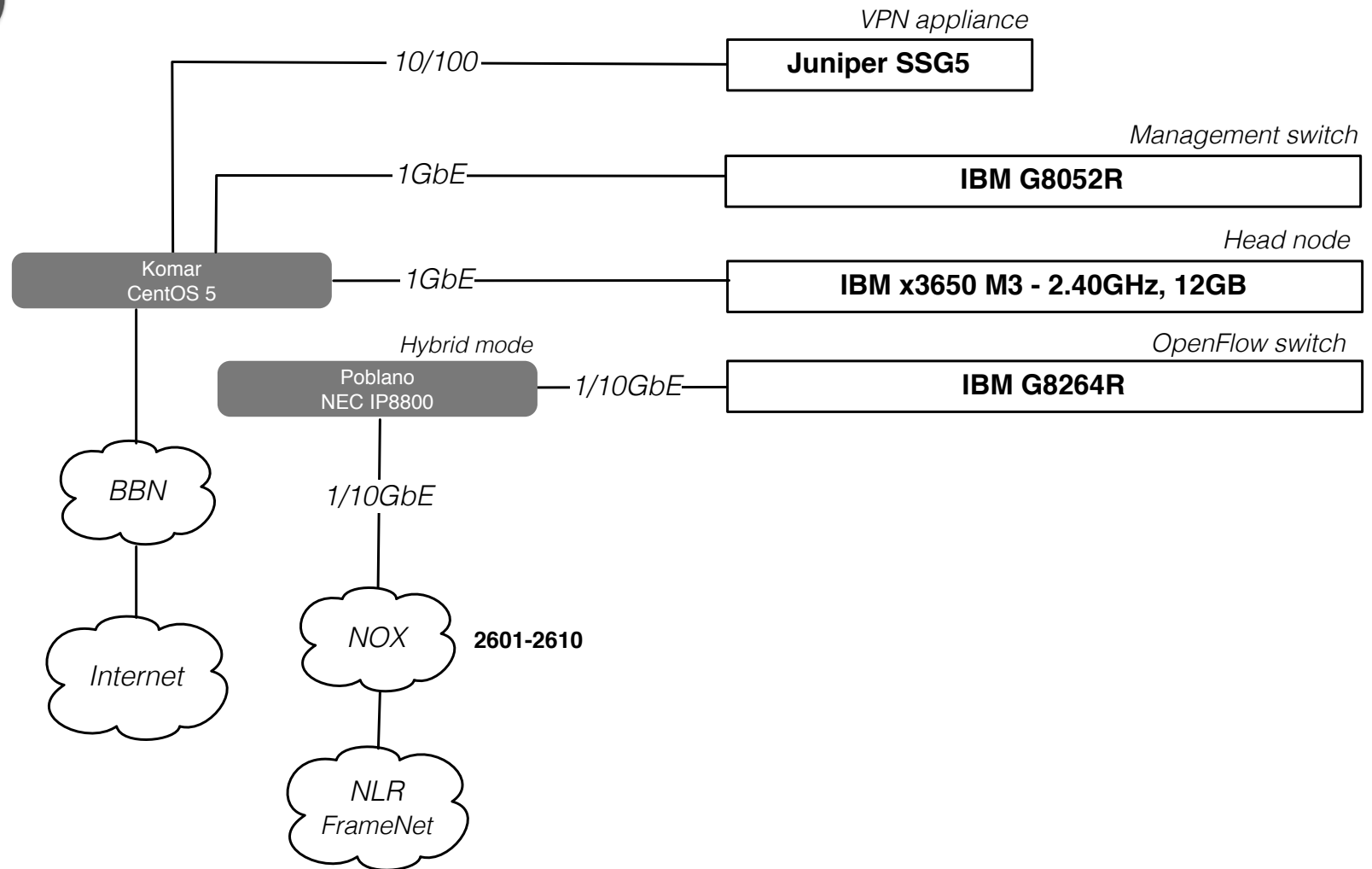


Rack IP address assignment

- ▶ /24 of publicly routable IP addresses is the best choice
- ▶ 2 are assigned to elements of the rack
 - Management/Head node
 - SSG5 VPN appliance (to create a secure mesh for management access between racks)
- ▶ The rest is used to assign IP addresses to experimenter instances
 - VMs and hardware nodes



Example rack connection (GPO/BBN)

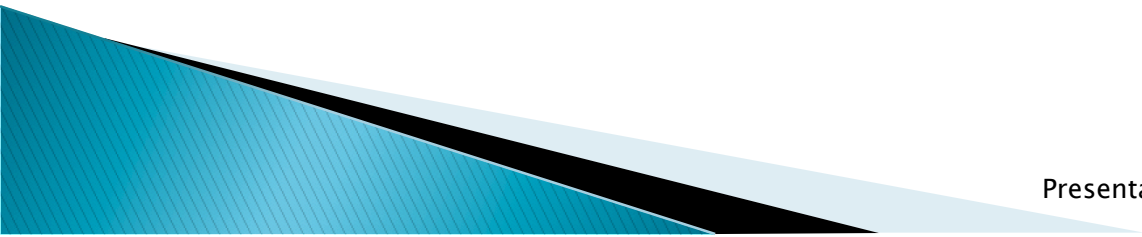


Rack installation

► Particulars:

- Power options include (negotiated ahead of time)
 - 208V 3Phase
 - 208V 1Phase
 - 120V 1Phase
 - Total of ~10kW of power needed.
- Space:
 - e1350 42U Rack Cabinet 79.5" H x 25.5" W x 43.5" D (2020 mm x 648 mm x 1105 mm)
- Racks arrive on-site pre-assembled by the vendor
 - IBM representative will need to come on-site to complete install and hookup
 - NBD hardware support
 - ExoGENI Ops finishes ORCA configuration
 - GPO acceptance testing

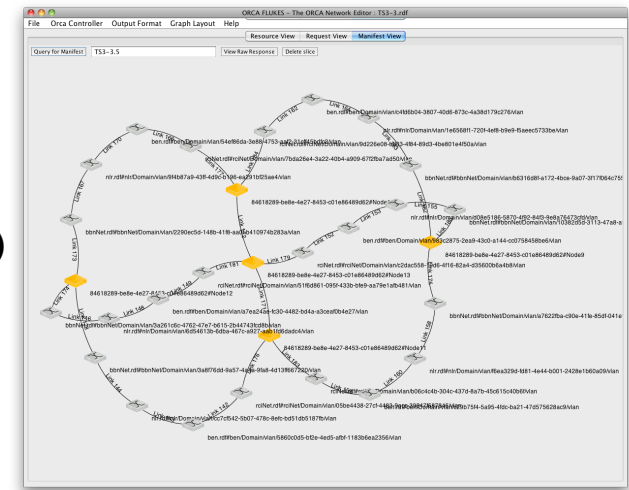
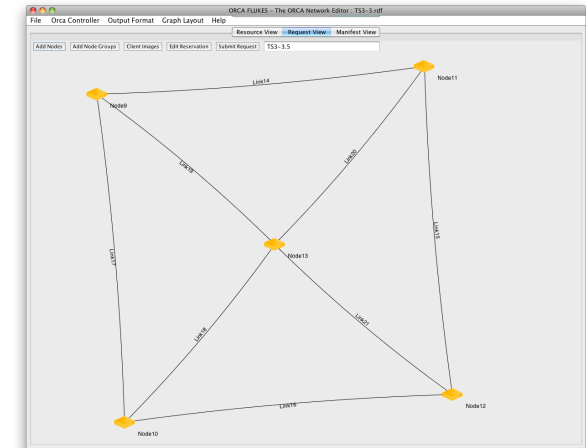
ExoGENI experiments



Presentation title goes here

Experimentation

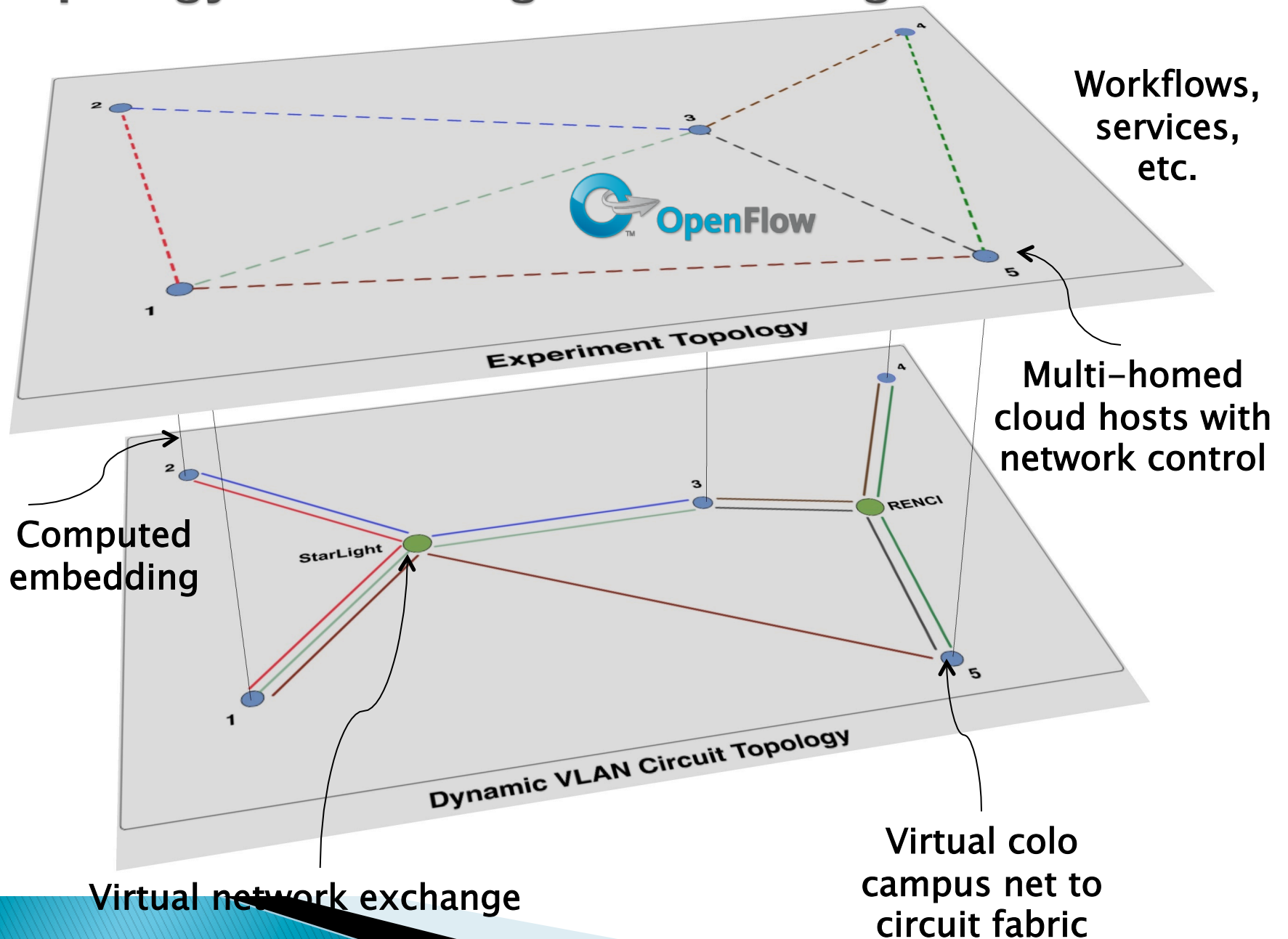
- ▶ Compute nodes
 - Up to 100 VMs in each full rack
 - A few (2) bare-metal nodes
 - BYOI (Bring Your Own Image)
- ▶ True Layer 2 slice topologies can be created
 - Within individual racks
 - Between racks
 - With automatic and user-specified resource binding and slice topology embedding
 - Dynamic circuit services (Sherpa, OSCARS) called where available
- ▶ OpenFlow experimentation
 - Within racks
 - Between racks
 - Include OpenFlow overlays in NLR (and I2)
 - On-ramp to campus OpenFlow network (if available)
- ▶ Experimenters are allowed and encouraged to use their own virtual appliance images



Tools/APIs

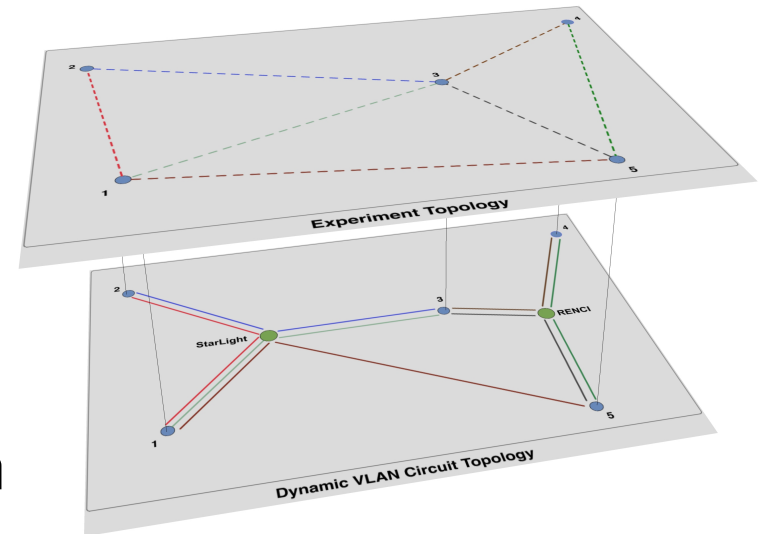
- ▶ GENI AM API
 - Omni
- ▶ ORCA Native API
 - Flukes

Topology embedding and stitching

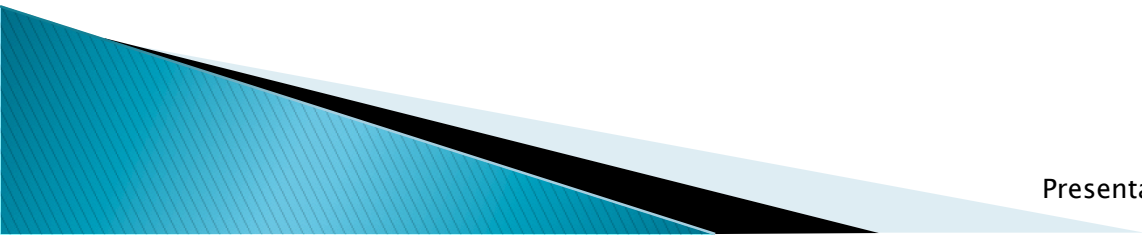


ExoGENI slice isolation

- ▶ Strong isolation is the goal
- ▶ Compute instances are KVM based and get a dedicated number of cores
- ▶ VLANs are the basis of connectivity
 - VLANs can be best effort or bandwidth-provisioned (within and between racks)



ORCA control framework



Presentation title goes here

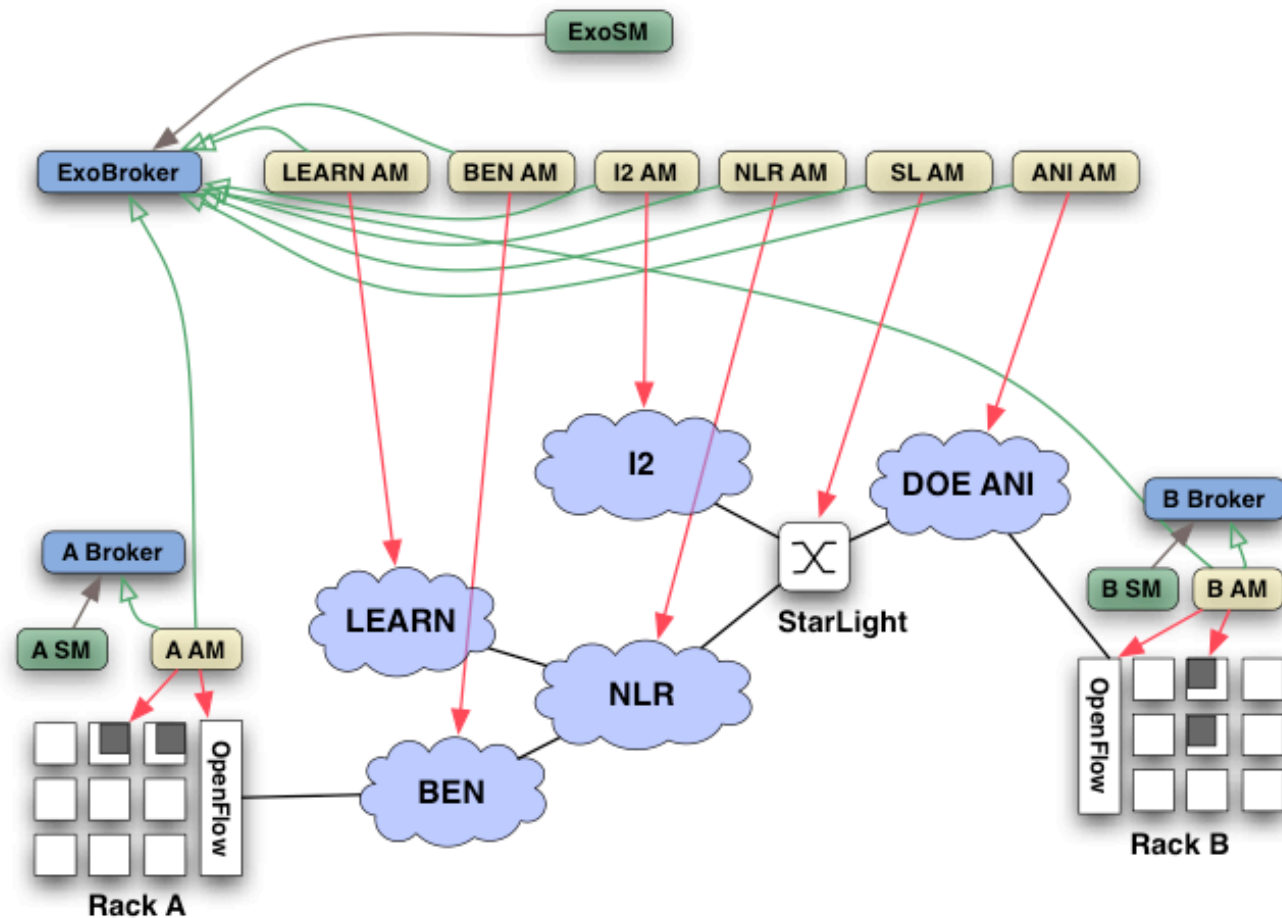
ORCA Overview

- ▶ Originally developed by Jeff Chase and his students at Duke
- ▶ Funded as Control Framework Candidate for GENI
 - Jointly developed by RENC1 and Duke for GENI since 2008.
- ▶ Supported under several current NSF and DOE grants to enable ORCA to run computational networked clouds as well as GENI
- ▶ Fully distributed architecture
- ▶ Federated with GENI
- ▶ Supports ORCA-native interface, resource specification and tools
 - Flukes
- ▶ Supports GENI AM API and GENI Rspec
 - Omni

ORCA Deployment in ExoGENI

- ▶ Each rack runs its own ORCA actor (SM) that exposes
 - ORCA native API
 - GENI AM API
- ▶ Rack-local SM
 - Can only create slice topologies with resources within that rack
- ▶ ‘ExoSM’ has global visibility
 - Has access to resources in all racks
 - Has access to network backbone resources for stitching topologies between racks
- ▶ Uniquely ExoGENI racks act both as
 - Independent GENI aggregates
 - A collective aggregate with intelligent topology embedding and stitching via dynamic or static circuits (NLR, ESnet, I2)

ExoGENI ORCA deployment



Thank you!

- ▶ <http://www.exogeni.net>
 - (see the section for operators in the wiki)
- ▶ <http://geni-orca.renci.org>

