ORCA-BEN: a distributed control framework for GENI
- A Inter-Cloud IaaS Provisioning System for Scientific Workflows

Yufeng Xin yxin@renci.org
The Project Team

Co-PI: Jeffrey Chase
Duke University

PI: Ilia Baldine
RENCI, UNC-Chapel Hill

Students: Prateek Jaipuria
Muzhi Zhao

Staff: Victor Orlikowski

Anirban Mandal
GENI Controller, application interface, actor registry

Aydan Yumerefendi
Core code enhancements

Chris Heermann
BEN Testbed operations, interconnections with national/international networks

Sincere thanks to our sponsors:
GPO/GENI, ASCR/DOE, SDCI/NSF
ORCA-BEN: one of several GENI Control Framework Candidates
- Lead of ‘Cluster D’ of GENI projects

Project goal: extend ORCA (Open Resource Control Architecture) to
- Cross cloud provisioning: connections, clusters, and topologies
- Provision multi-layered network domains (BEN, NLR Sherpa, I2 ION)
- Leveraging Semantic Web technologies
- Interoperate with other CF candidates (XMLRPC controller)
  - Connectivity
  - API and Resource description translations

Big Idea: create a variety of application-specific topologies out of collection of independent cloud providers connected by one or more multi-layered network providers
- GENI is a kind of IaaS.
- 40GB/100GB network will bring everything closer.
- Other uses: large-scale scientific applications on clouds interconnected by dynamic high speed networks
Outline

- **ORCA Architecture and Implementation**
  - Distributed actors and lease-based operation
    - Substrate Stitching
    - Semantic resource representations
  - Substrate components and interfaces
    - BEN and beyond
    - NEuca + ImageProxy
- **Interoperability**
  - Connectivity
  - Resource description translations
- **Scientific Application Support**
  - Wide area Hadoop based applications
  - Workflows
**Goal:** provision built-to-order *slices* of networked virtual infrastructure from multiple cloud or network providers, dynamically on the fly.
ORCA Design and Implementation

- Pervasive virtualization
- Brokering
- Pluggable
  - Policy
  - Controller
  - Interface
- Cloud technologies
- Semantic web resource representation
- Federated ID management
NEuca – networked extension to Eucalyptus

- Instantiate VM topologies
- Download/register user images at multiple sites
- Provide publicly routable IP with port number(s) to reach VM

ORCA
EUCAL
YPTUS
WORKER
WORKER
WORKER
WORKER
NEUCA
IMAGE PROXY
DNAT PROXY

Eucalyptus

RENCI
Eucalyptus/XCat ImageProxy
BEN Topology and Inter-Domain Connectivity

I2 ION

C-Wave

NLR FrameNet

RENCI

ATDNet, MAX, BOSSnet

Dark fiber to McLean, VA

L3 PoP, Raleigh

RENCI

UNC-CH

NCI BEN Fiber

UNCC/RENCI BEN Fiber

RNCI BEN Fiber

Infinera DTN

Polatis

Duke

6509

Potential connection

Existing connection

National/International research networks

StarLight

Cisco RTP

Cisco Research Wave

C-Wave

NLR 7606

Polatis OXC

To BEN

10Gbps
Distributed Stitching in ORCA

- Stitching is the process of connecting resources of different types of belonging to different providers to each other
- Stitching involves a process of agreeing on labels between two parties
  - Labels may denote bandwidth (VLAN tag, MPLS label, DWDM wavelength)
- Can be done through multi-step negotiation
  - Complex protocols may be required
- ORCA approach is based on analyzing dependencies of providers (label producer vs. label consumer; can you translate labels?)
  - A DAG of dependencies is formed based on the requested topology
  - Properties carrying labels are passed from the root(s) of the DAG along the branches
  - Certain resources wait to be instantiated until the label information needed to stitch them is available
Outline

• **ORCA Architecture**
  – Distributed actors and lease-based operation
  – Substrate components and interfaces
  – Substrate Stitching
  – **Semantic resource representations**
    • Semantic graph
    • Substrate: NDL-OWL, Cloud
    • Domain: Abstraction
    • Request: Cluster, connection, topology

• **Interoperability**
  – Connectivity
  – Resource description translations

• **Scientific Application Support**
  – Wide area Hadoop based applications
  – Workflows
Outline

• **ORCA Architecture and Implementation**
  – Distributed actors and lease-based operation
    • Substrate Stitching
    • Semantic resource representations
  – Substrate components and interfaces
    • BEN and beyond
    • NEuca + ImageProxy

• **Interoperability**
  – Connectivity
  – Resource description translations
Substrate Resource Description

- **Multilayer Network description**
  - Semantic graph: dynamical model with new RDF resources and properties (xc, virtual interface, connection hierarchy, reservations)
  - OWL vs. RDF
  - DTN network description
  - Collection (bag, set, list), labelSet, etc

- **Cloud resource description**
  - Virtualization as an adaptation
  - Substrate hosting capability
  - Shortcut to classified type
Request and Embedding

– Request Type
  • Bound vs. Unbound
  • Single cloud site vs. multiple cloud sites
    – Connection:
    – Virtual Topology:
    – Virtual cluster: Node group and dependency: Master/Slave groups,

– Properties:
  • Reservation ID
  • Reservation term: time.owl
  • VM type: EC2Small
    – Predefined (CPU, Memory, Disk)
  • IP addresses
    – Range per Node Group
  • VM Image
    – Image URL (XML metafile e.g. ImageProxy or OVF)
  • Post-boot NEuca script
  • Network QoS
    – Bandwidth

– **Embedding Policy**
  • Path computing
  • Topology embedding
  • Cluster and topology partitioning and embedding
ORCA Architecture and Information Life Cycle

Request RDF

Experiment Control Tools

Sliver reservation manifest

Redemption sub-request

Broker

delegate

lease

ticket

request

Web portal

Operators

Substrate abstraction and advertisement

Aggregate

Substrate description
Outline

• **ORCA Architecture**
  – Distributed actors and lease-based operation
  – Substrate components and interfaces
  – Substrate Stitching
  – Semantic resource representations
    • Semantic graph
    • Substrate: NDL-OWL, Cloud
    • Domain: Abstraction
    • Request: Cluster, connection, topology

• **Interoperability**
  – Connectivity
  – Resource description translations
  – User Interface: Web, XMLRPC API

• **Scientific Application Support**
  – Wide area Hadoop based applications
  – Workflows
ProtoGENI Interoperability

• Resource description conversion
  – From NS2/RSpec v1/v2 to NDL-OWL requests

• GENI AM API and ProtoGENI API support via ORCA XMLRPC controller in the SM
  – Validated via GPOs ‘omni’
  – ListResource(),createSliver(),sliverStatus(),deleteSliver()

• Next steps
  – Advertizing ORCA clouds to ProtoGENI
  – Creating ProtoGENI-friendly manifests
Cluster-D connectivity
Ongoing inter-cluster connectivity, interoperability, and application integration

- GENI AM API and ProtoGENI API support via ORCA XMLRPC controller in the SM.
- Developing I2 ION, and OSCAR interfaces
- Connectivity to ESNet
- Monitoring integration
- Hadoop based analytics
- Scientific workflow
Use case 1: Wide area Hadoop based scientific applications over clouds

• Map-reduce is an example of a step within a complex scientific workflow.
• Hadoop VM image, script to launch Hadoop cluster on Neuca/EC2 clouds.
• ORCA controller to provision a Hadoop cluster on multiple Neuca/EC2 clouds interconnected by multi-domain high speed network.
• Performance study with different distributions of Hadoop nodes.
• Application: Hadoop-Blast.
Use case 2: Scientific Workflow

• Multi-scale chemistry simulation pipeline (UNC – EFRC, DOE)
• MotifNetwork (an NSF CDI-II project)
• Integrated Microbial Genomes or IMG(DOE environmental genomics)
• Nearby Supernova Factory or SNfactory (DOE cosmological physics),
• Workflow for Collaborative Adaptive Sensing of the Atmosphere or CASA (an NSF ERC).
More Information

- [http://geni-orca.renci.org](http://geni-orca.renci.org)
- [http://www.networkedclouds.org](http://www.networkedclouds.org)