Chameleon

www. chameleoncloud.org

GENI FEDERATION WITH CHAMELEON: A LARGE-SCALE, RECONFIGURABLE EXPERIMENTAL ENVIRONMENT FOR CLOUD RESEARCH

Principal Investigator: Kate Keahey

Co-PIs: J. Mambretti, D.K. Panda, P. Rad, W. Smith, D. Stanzione

Presented By Joe Mambretti, Director,

International Center for Advanced Internet Research, Northwestern University

GENI-FIRE Federation Workshop

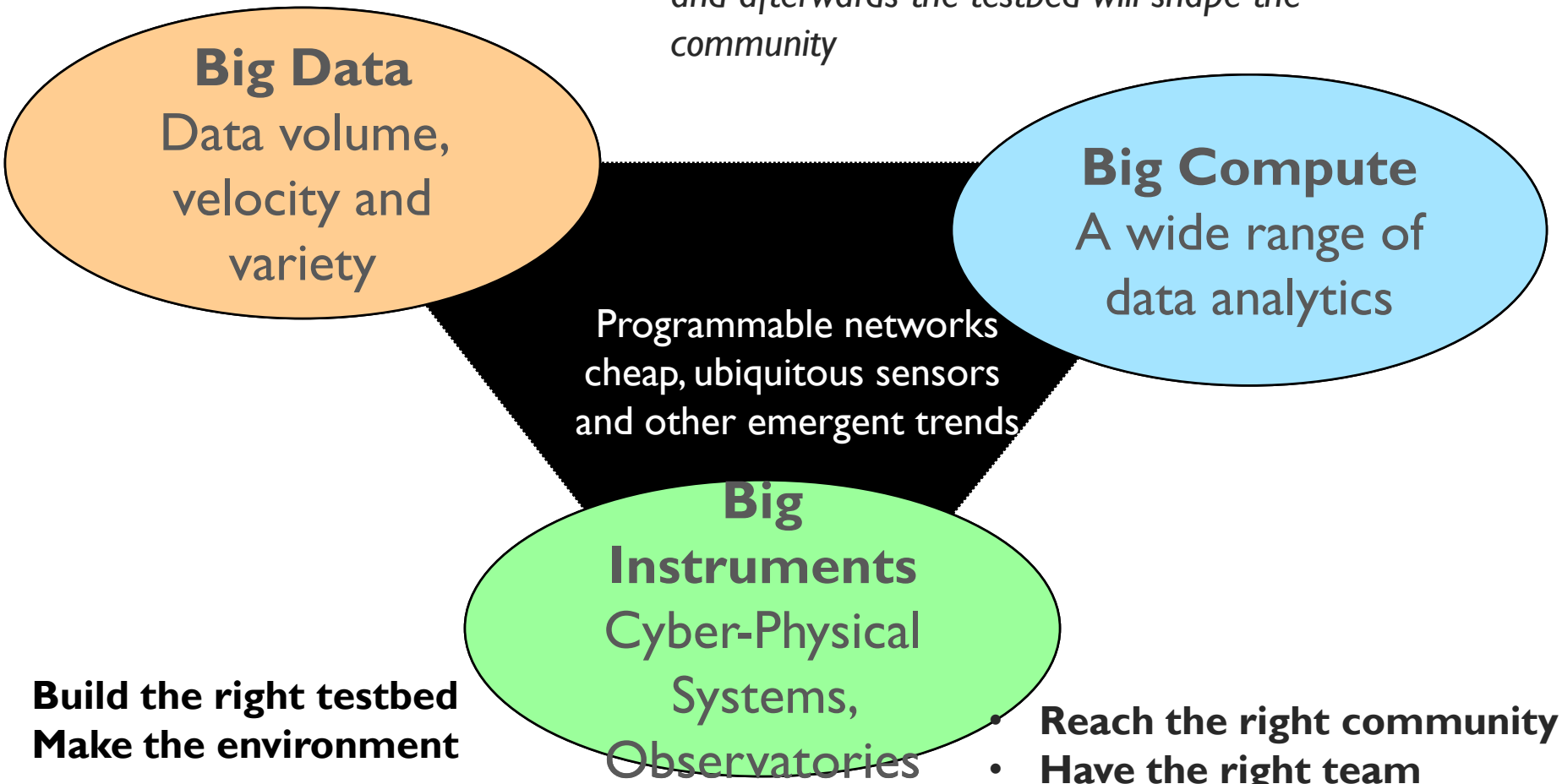
Washington DC

Sept 17-18, 2015

THE UNIVERSITY OF CHICAGO TACC NORTHWESTERN UNIVERSITY THE OHIO STATE UNIVERSITY UTSA NSF

TESTBED TO SUPPORT THE COMMUNITY'S RESEARCH CHALLENGES

The community builds the testbed, and afterwards the testbed will shape the community



- **Build the right testbed**
- **Make the environment**

- **Reach the right community**
- **Have the right team**

CHAMELEON: A POWERFUL AND FLEXIBLE EXPERIMENTAL INSTRUMENT

- Large-scale
  - Targeting Big Data, Big Compute, Big Instrument research
  - Over 650 nodes, 5 PB disk, 100G network
- Reconfigurable
  - Bare metal reconfiguration, single instrument, graduated approach for ease-of-use
- Connected
  - Workload and Trace Archive, partners with production clouds
- Complementary
  - Complementing GENI, Comet, Wrangler, XSEDE
  - Partnering with GENI, Grid'5000, OCC, OSDC et al (FIRE?)
- Sustainable
  - Strong industry connections

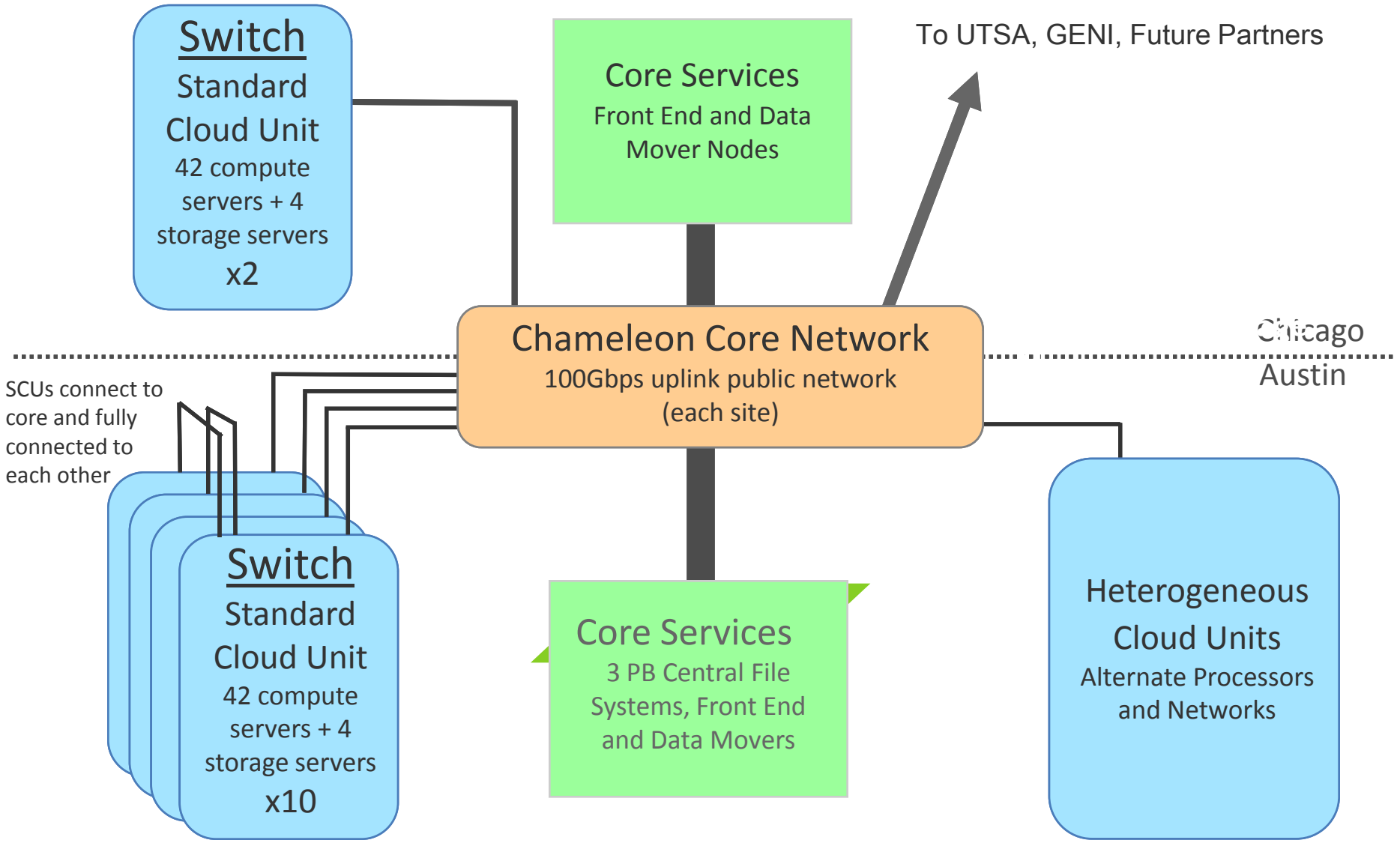
RESEARCH COMMUNITIES AND CAPABILITIES

Users	New models, algorithms, platforms, auto-scaling HA, etc., Application and educational uses	*Persistent, reliable, shared cloud*
Core Researchers and Users	Repeatable experiments in new models, algorithms, platforms, auto-scaling, HA, etc.	*Isolated partition, pre-configured images reconfiguration*
Core Researchers	Virtualization technology (SR-IOV, accelerators, etc.) Infrastructure-level resource management	*Isolated partition, full bare metal reconfiguration*

SUPPORTED APPLICATIONS

- CPS
  - Offloading, muti-criteria trade-off analysis (response time vs cost), auto-scaling, high availability, etc.
- Machine learning, data mining
  - Mix of Big Compute and Big Data simulations and models, design of novel data processing frameworks
- System Software/Virtualization
  - Hypervisors optimizing a range of qualities, SR-IOV, virtualizing accelerators, etc.
- Networking
  - Programmable networks & QoS, refinement and effects of SR-IOV, large dataflows, end-to-end QoS

ARCHITECTURE



SYSTEM SOFTWARE: CORE CAPABILITIES

Users

Persistent clouds
OpenStack

Pre-configured Image Catalog
VM images

Core Researchers and Users

Pre-configured Image Catalog
Bare metal images

Core Researchers

Provisioning
LosF (TACC)
OpenFlow (GENI)

Scheduling
OAR2 (Grid'5000)

Orchestration
Cloudinit.d

CHAMELEON SERVICES AND FEATURES

- User Services
  - Allocation management through reservations, automatic image deployment
  - Dedicated Web portal for reservations, docs, stats, etc.
- Experiment Support
  - Trace and Workload Archive
  - Experiment enhancement (e.g., load generators)
- Additional Features
  - Reconfigurable, connected instrument
  - Development-focused approach
  - Ease-of- use: one stop shopping for experimental needs
  - Distinct from off-the-shelf cloud services
  - Code ownership and collaboration

NETWORKING CAPABILITIES

- Expose SDN, OpenFlow, etc. to users
  - Isolation
  - Hybrid Network Capabilities
  - Programmable Topologies
  - Integration With Other Resources Within and External to the Testbed
- Pushing 100G Networks To Their Limit
  - Using 100G + SDN Optimally
  - Chameleon appliances and services allow experimenters a highly granulated view into -- and control -- over traffic flows
- Integration/Federation with GENI (Et Al…*N Testbeds)
- Within Common Policy Context

PARTNERSHIP WITH GENI COMMUNITY

- Chameleon Enables the GENI Virtual Laboratory For Networking and Distributed Systems Research and Education To Extended Significantly With Many New Types of Resources.
- This Blending of Resources Will Enable Investigations Of New Types Of Innovative Highly Distributed Environments at Scale.

GENI-CHAMELEON FEDERATION

- Federation: ~ Identity Federation (Authentication/Authorization)
- Goal: Experiments Should Be Able To Log Into Either The GENI or Chameleon Environments Through a Federated Identity Mechanism.
- After An Experimenter Logs In, Portals Should Be Able To Identify the Groups And Or Projects In Which That Experimenter is a Member To Verify Access To Resources That Belong To Specific Projects.

GENI AND OPENID: PHASE 1

- GENI Currently Supports an OpenID Provider (OP) Through Which Identity Can Be Federated to Services.
- Phase 1: GENI⇔ Chameleon Federation.
  - Chameleon Has Implemented an OpenID RP To Receive Identity Information From GENI.
  - However, A Project Association Is Required
  - Currently, GENI Experimenters Who Would Like To Use Chameleon Need To Be Added As Members Of the "GENI/Chameleon Federation" Project In The GENI Environment.

GENI AND OPENID: PHASE 1 (CONT.A)

- Afterward, When They Are Project Members, Chameleon Resources Are Shown As Available.
- They Can See And Select A “Use-Chameleon” Button That Brings Them To the Chameleon/OpenID Page Where They Will Be Authenticated And Then Automatically Added To The “GENI/Chameleon Federation” Project In the Chameleon Environment.
- Subsequently, They Have Log-In Access to Chameleon OpenStack interface(s) And Can Log Into Chameleon Using Their GENI OpenID directly from the Chameleon Portal.

GENI AND OPENID: PHASE 1 (CONT.B)

- This Project Association Provides A Means To Allow Initial Exploration and Evaluation of the Chameleon Environment By Experimenters
- If The Environment Proves To Be Useful, The Experimenters Can Establish Their Own Chameleon Projects To Enable Larger Scaling of Resources

PHASE 2

- **Phase 2: Chameleon⇔ GENI Federation**
  - **The GENI Portal Cannot Receive OpenID Federated Identity Information From Another Source (Relaying Party or RP, i.e., Chameleon)**
  - **This Issue Is Being Addressed Through the G⇔ C Federation Project.**
  - **Chameleon Will Soon Implement a Process to Transmit Identity Information to GENI Enable Federation from Chameleon -> GENI**
    - **1. Creating OpenID RP in GENI Portal**
    - **2. Creating OpenID OP in Chameleon Portal**
    - **3. Investigating Potential For Chameleon Shibboleth IdP to log In To GENI (Technique Used By SAVI, Which Uses OpenStack & ShibIdP To Send Project Data From Keystone To LDAP)**

PHASE 3

- Project Federation Between GENI and Chameleon
- A Mechanism Is Required To Enable An Experimenter To Bring a GENI Project to Chameleon
- GENI Designates A “Project Lead” Who Is Manually Approved (e.g., FT Faculty at Accredited Research Institution)

PHASE 4

- Plans Are To Enable GENI Experiments To Reserve And Use Chameleon Resources Without Using The CC User Interface
- This Requires Some Development Activities By The Chameleon Systems Team Addressing Issues At:
  - API Layer
  - Control Plane Layer
  - Orchestration Layer
  - Data Plane Layer
  - SDI Infrastructure Federation
  - Etc

PHASE 5: FEDERATION-AS-A-SERVICE

- Federation-as-a-Service
- Generalized Architectural Model For Federation With N Testbeds (GENI, FIRE, Chameleon, CloudLab, OSDC, CERN CT, SAVI, OCX, JGN-X, IOFT, NGN, GTS, and Many Others)
- SDX Implementations (e.g., StarLight SDX) Will Be Key Resources
- An Open Architecture API Would Be Useful
- Also, Mechanisms For Policy Implementation
- Architecture Should Incorporate Options For Policy Based Access to Other SDN Resources
- Policy Architecture For SDN Is Progressing Through Standardization Processes

APPLICATIONS REQUIRING PARTICULARLY STRICT FEDERATION POLICY RULES

- Sensitive Research Data
- Apps Based On Highly Restricted Data Access
- Computational Bioinformatics Research
- Various Medical Applications
- Access To Specialized Instrumentation

THANKS!