www. chameleoncloud.org

# CHAMELEON: BUILDING A RECONFIGURABLE EXPERIMENTAL TESTBED FOR CLOUD RESEARCH

Kate Keahey

*keahey@anl.gov*

***LSN Meeting***
***December 8, 2015***
*National Science Foundation*



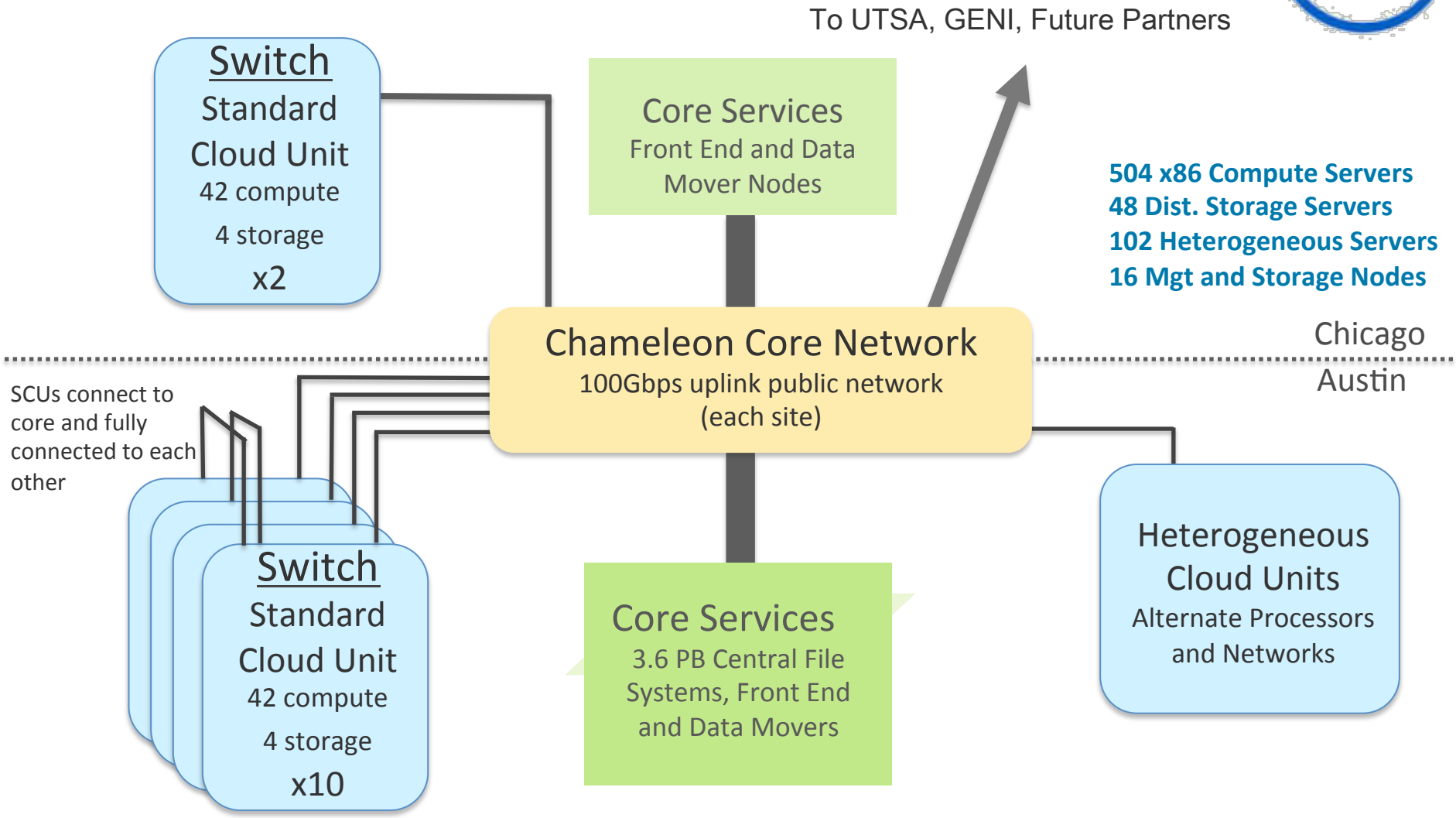





# CHAMELEON DESIGN STRATEGY

- Large-scale: "Big Data, Big Compute, Big Instrument research"
  - ~650 nodes (~14,500 cores), 5 PB disk over two sites, 2 sites connected with 100G network
- Reconfigurable: "As close as possible to having it in your lab"
  - Bare metal reconfiguration, operated as a single instrument
  - Support for repeatable and reproducible experiments
- Connected: "One stop shopping for experimental needs"
  - Workload and Trace Archive
  - Partnerships with production clouds: CERN, OSDC, Rackspace, Google, and others
  - Partnerships with users
- Complementary: "Can't do everything ourselves"
  - Complementing GENI, Grid'5000, and other experimental testbeds

# CHAMELEON HARDWARE



# CHAMELEON HARDWARE

- Standard Cloud Units (SCU) (deployed)
  - Each of the 12 Standard Cloud Units is a single 48U rack
  - 42 Dell R630 compute servers, each with dual-socket Intel Xeon (Haswell) processors and 128GB of RAM
  - 4 DellFX2 storage servers, each with a connected JBOD of 16 2TB drives (total of 128 TB per SCU)
  - Allocations can be an entire SCU, multiple SCUs, or within a single SCU, or across SCUs (e.g., storage servers for Hadoop configurations)
  - 48 port Force10 s6000 OpenFlow-enabled switches 10Gb to hosts, 40Gb uplinks to Chameleon core network
  - Connectx3 IB network
- Shared infrastructure (deployed)
  - 3.6 PB global storage, 100Gb Internet connection between sites
- Heterogeneous Cloud Units (to be procured in Y2)
  - ARM microservers, Atom microservers, SSDs, GPUs, FPGAs

# CAPABILITIES AND SUPPORTED RESEARCH

Development of new models, algorithms, platforms, auto-scaling HA, etc., innovative application and educational uses

*Persistent, reliable, shared clouds*

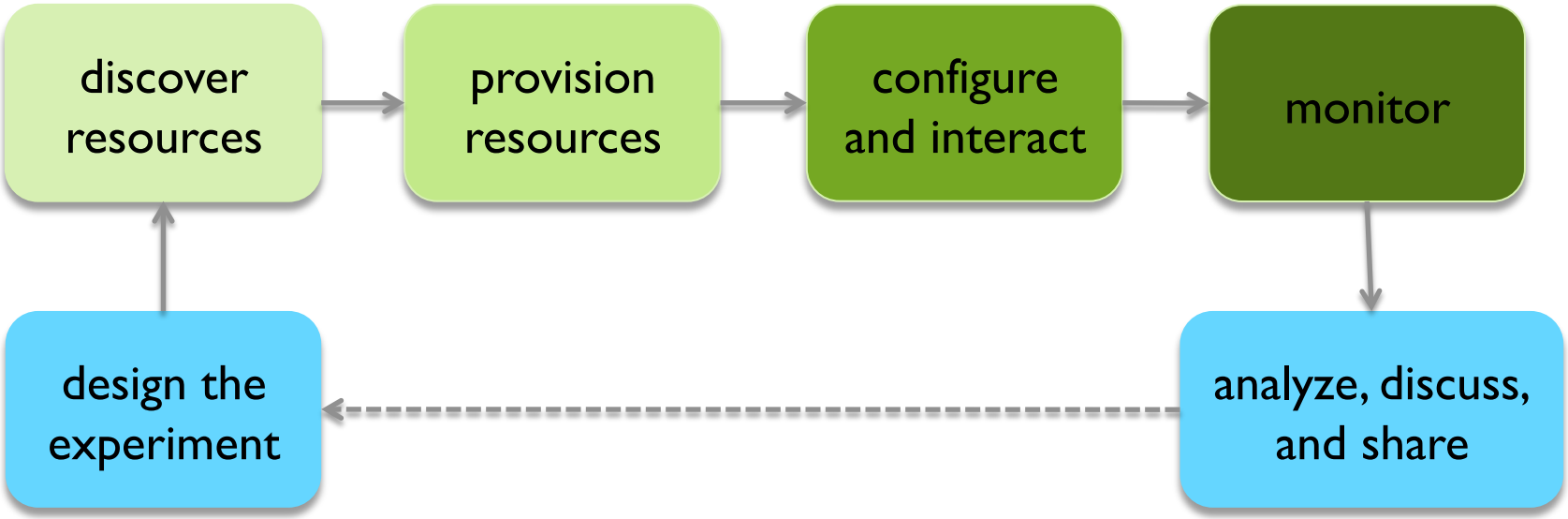
Repeatable experiments in new models, algorithms, platforms, auto-scaling, high-availability, cloud federation, etc.

*Isolated partition, Chameleon Appliances*

Virtualization technology (e.g., SR-IOV, accelerators), systems, networking, infrastructure-level resource management, etc.

*Isolated partition, full bare metal reconfiguration*

# IMPLEMENTING THE EXPERIMENTAL WORKFLOW




- Starting from scratch
- Sustainability as design criterion: can a CS testbed be built from commodity components?
- OpenStack + G5K + special sauce

# CHI: DISCOVERING AND VERIFYING RESOURCES

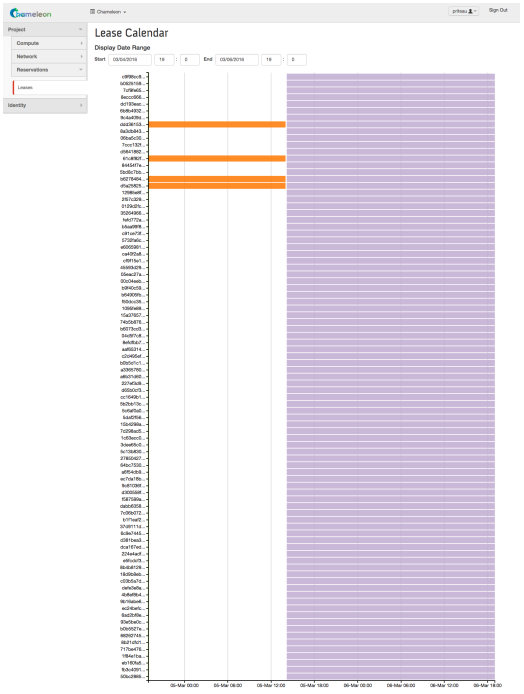
- Fine-grained, up-to-date, and complete representation
- Both machine parsable and user friendly representations
- Testbed versioning
  - "What was the drive on the nodes I used 6 months ago?"
- Dynamically verifiable
  - Does reality correspond to description? (e.g., failure handling)

---

- Grid'5000 registry toolkit + Chameleon portal
  - Automated resource description, automated export to RM/Blazar
- G5K-checks
  - Can be run after boot, acquires information and compares it with resource catalog description

# CHI: PROVISIONING RESOURCES

- Resource leases
- Advance reservations (AR)
  - Facilitates allocating at large scale
- Allocating a range of resources
  - Different node types, switches, etc.
- Multiple environments in one lease
- Future extensions: match making, internal management




- OpenStack Nova/Blazar
- Extensions to support Gantt chart displays and other features

# CHI: CONFIGURE AND INTERACT

- Allow deep reconfigurability (access to console)
- Map multiple appliances to a lease
- Snapshotting for image sharing
- Efficient appliance deployment
- Handle complex appliances
  - Virtual clusters, cloud installations, etc.
- Interact: reboot, power on/off, access to console
- Shape experimental conditions

---

- OpenStack Ironic, Glance, and meta-data servers

# CHI: MONITORING

- Enables users to understand what happens during the experiment
- Types of monitoring
  - User resource monitoring
  - Infrastructure monitoring (e.g., PDUs)
  - Custom user metrics
- High-resolution metrics
- Easily export data for specific experiments

---

- OpenStack Ceilometer

# BUILDING CHI: CHAMELEON BARE METAL

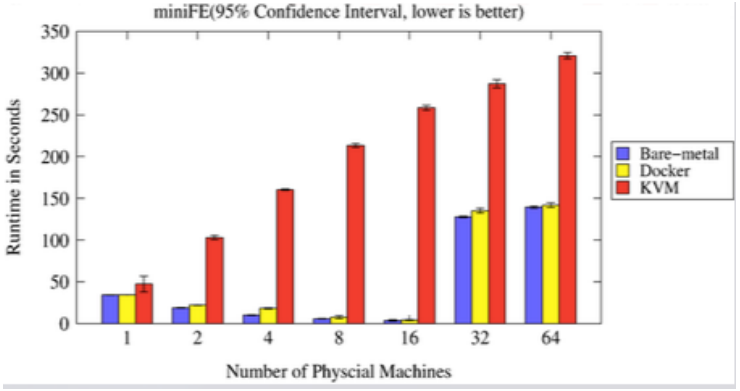
- Defining requirements (proposal stage)
- Developing architecture
- Technology Evaluation and Risk Analysis
  - Rough requirements based analysis
  - Technology evaluation: Grid'5000 and OpenStack
  - Implementation proposals
- Implementing CHI
- Technology Preview deployment
- Early User and public availability

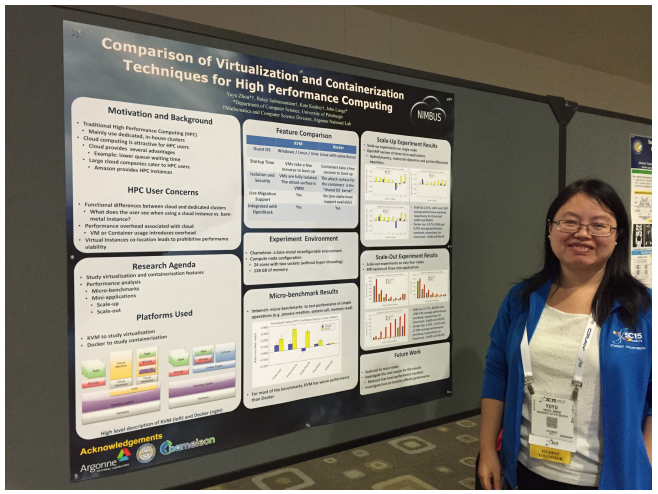
# CHAMELEON TIMELINE AND STATUS

- 10/14: Project starts
- 12/14: FutureGrid@Chameleon (OpenStack KVM cloud)
- 04/15: Chameleon Technology Preview on FG hardware
- 06/15: Chameleon Early User on new hardware
- 07/15: Chameleon Public availability (bare metal)
- 09/15: Chameleon KVM OpenStack cloud available
- 10/15: Interoperability with GENI (1st phase)
- Today: 500+ users/140+ projects
- 2016: Heterogeneous hardware available

# VIRTUALIZATION OR CONTAINERIZATION?

- Yuyu Zhou, University of Pittsburgh
- Research: lightweight virtualization
- Testbed requirements:
  - Bare metal reconfiguration
  - Boot from custom kernel
  - Console access
  - Up-to-date hardware
  - Large scale experiments

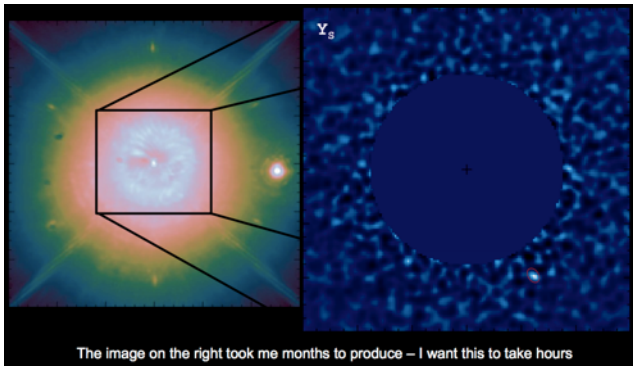


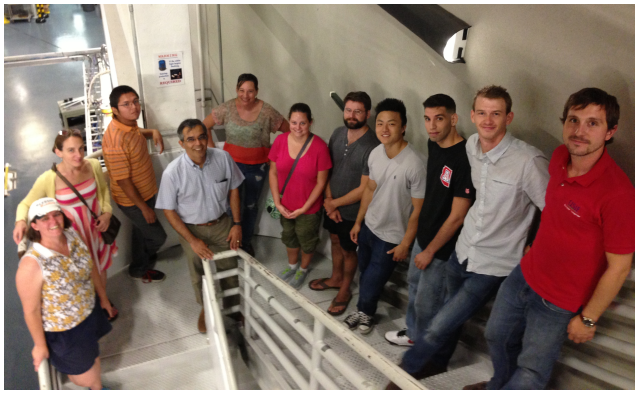




*SC15 Poster: "Comparison of Virtualization and Containerization Techniques for HPC"*

# TEACHING CLOUD COMPUTING

- Nirav Merchant and Eric Lyons, University of Arizona
- ACIC2015: project-based learning course
  - Data mining to find exoplanets
  - Scaled analysis pipeline by Jared Males
  - Develop a VM/workflow management appliance and best practice that can be shared with broader community
- Testbed requirements:
  - Easy to use IaaS/KVM installation
  - Minimal startup time
  - Support distributed workers
  - Block store: make copies of many 100GB datasets

# IN THE PIPELINE…

- Y2 Theme is Users
- Outreach
- Experiment management
  - Appliances: snapshotting, sharing, appliance marketplace, community
  - Experiment Blueprint: automation and preservation
- Functionality: from possible to easy
  - Better reconfiguration capabilities
  - Better isolation and networking capabilities
  - Better infrastructure monitoring (PDUs, etc.)
  - Allocation management
  - And others

# PARTING THOUGHTS

- Work on your next research project @ www.chameleoncloud.org!

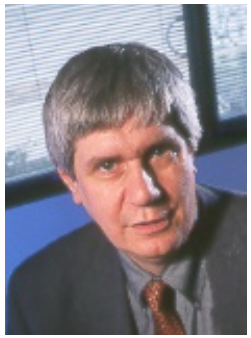
*The most important element of any experimental testbed is users and the research they work on*

- From vision to reality with Express Delivery
  - Operational testbed: 500+ users/140+ projects
  - Sustainability as a design criterion: building a CS testbed as an application of cloud computing
- Pursuing sustainability for us, for the broader community, and for testbeds and operation blueprints everywhere

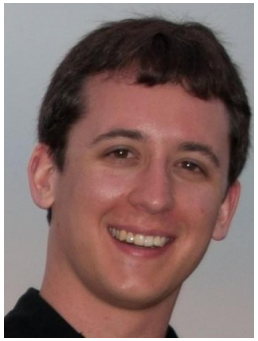
# CHAMELEON TEAM

Kate Keahey
Chameleon PI
Science Director
Architect
University of Chicago

Paul Rad
Industry Liason
Education and training
UTSA



Joe Mambretti
Programmable networks
Federation activities
Northwestern University



Pierre Riteau
Devops Lead
University of Chicago

DK Panda
High-perf networking
Ohio State University

Dan Stanzione
Facilities Director
TACC