



[www.chameleoncloud.org](http://www.chameleoncloud.org)



## CHAMELEON: BUILDING A RECONFIGURABLE EXPERIMENTAL TESTBED FOR CLOUD RESEARCH

Principal Investigator: Kate Keahey

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

*BOF: Development of a Next-Generation, Interoperable, Federated  
Cyberinfrastructure, XSEDE 2015*

*July 28*

*St Louis, MO*

SEPTEMBER 15, 2015

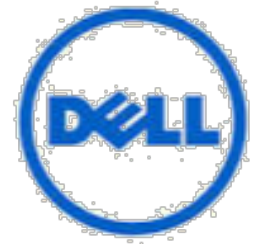
I



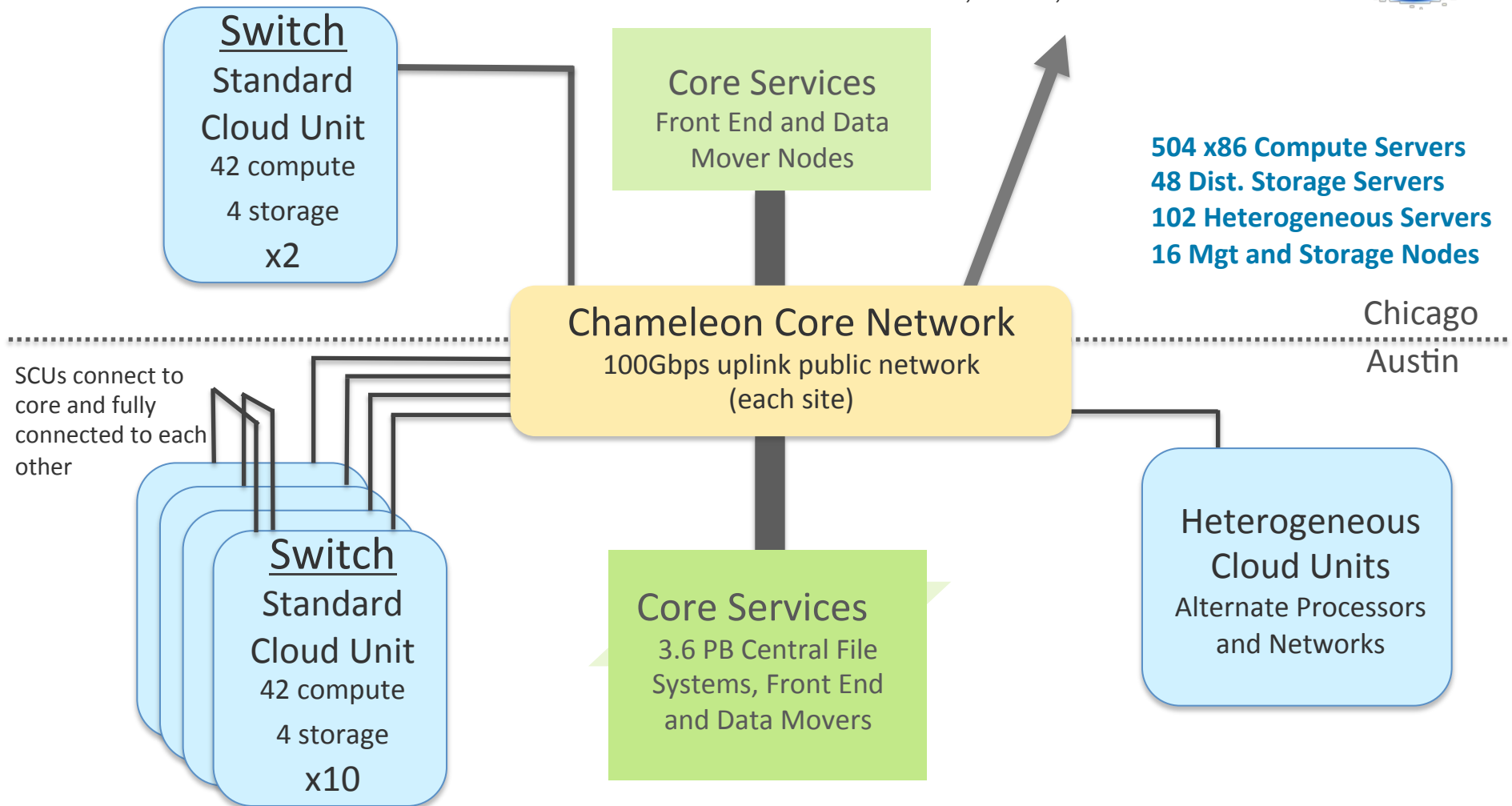
# CHAMELEON: A FLEXIBLE AND POWERFUL EXPERIMENTAL INSTRUMENT

- ▶ **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”
  - ▶ From bare metal reconfiguration to clouds
  - ▶ 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

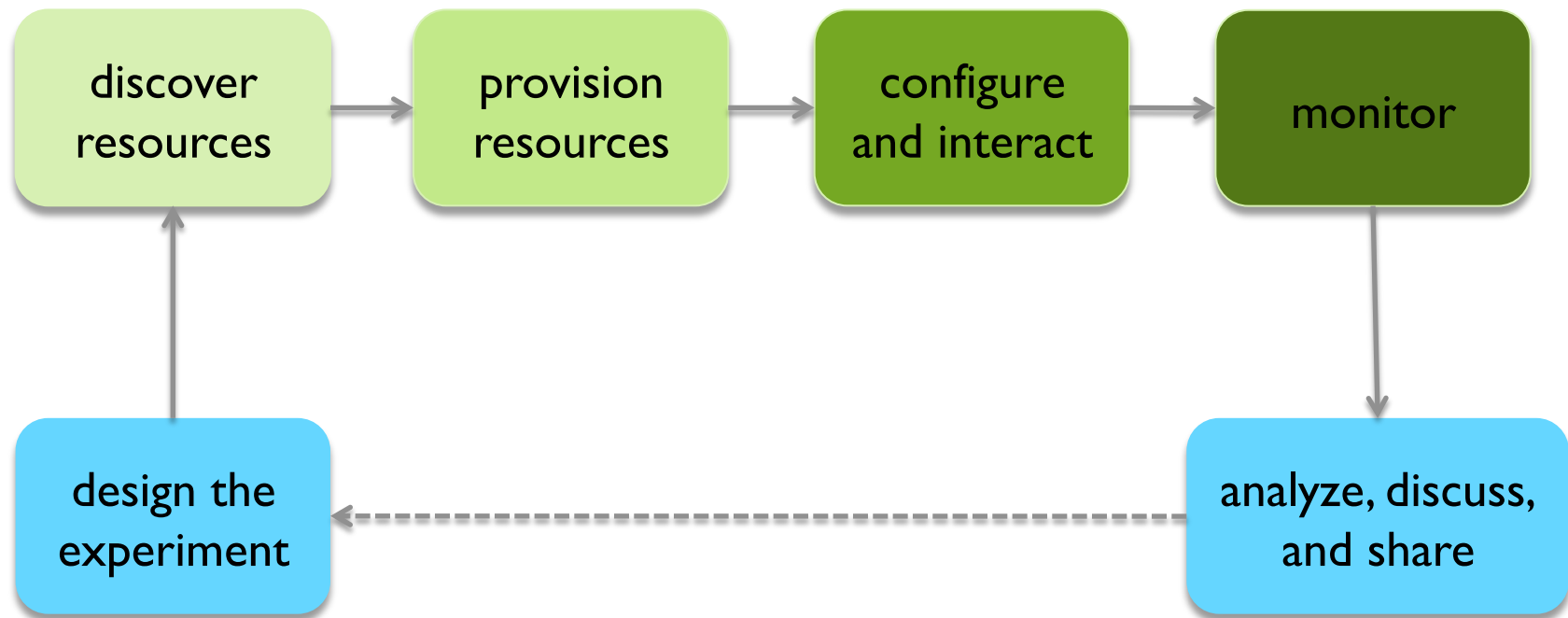
# CHAMELEON HARDWARE



To UTSA, GENI, Future Partners



# EXPERIMENTAL WORKFLOW

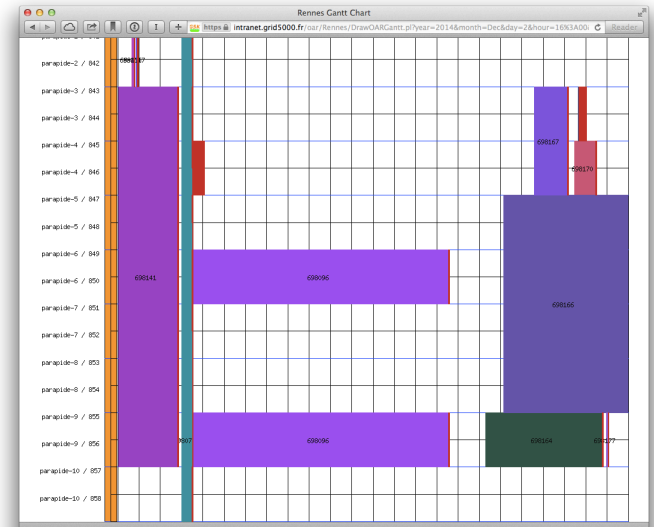


# CHI: SELECTING AND VERIFYING RESOURCES

- ▶ Complete, fine-grained and up-to-date representation
  - ▶ Machine parsable, enables match making
  - ▶ Versioned
    - ▶ “What was the drive on the nodes I used 6 months ago?”
  - ▶ Dynamically Verifiable
    - ▶ Does reality correspond to description? (e.g., failures)
- 
- ▶ Grid’5000 Registry
    - ▶ Automated resource description, automated export to RM
  - ▶ G5K-checks
    - ▶ Run at boot, acquire information, compare with resource catalog description

# CHI: PROVISIONING RESOURCES

- ▶ Resource leases
- ▶ Allocating a range of resources
  - ▶ Different node types, switches, etc.
- ▶ Multiple environments in one lease
- ▶ Advance reservations (AR)
  - ▶ Sharing resources across time
- ▶ Extensions: match making, Gantt chart displays



- 
- ▶ OpenStack Nova/Blazar
  - ▶ Extensions to support working with more resources, match making, and displays

# CHI: CONFIGURE AND INTERACT

- ▶ Map multiple appliances to a lease
- ▶ Allow deep reconfiguration (incl. BIOS)
- ▶ Snapshotting
- ▶ 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



# PROJECT STATUS AND TIMELINE

- ▶ 09/14: Project Start
- ▶ 12/14: NatureGrid@Chameleon (OpenStack/KVM)
- ▶ 04/15: Chameleon Technology Preview (bare metal)
- ▶ 06/15: Early Uses availability (new hardware)
- ▶ Today: **Public availability!**
- ▶ 2015/2016: Hardware and software upgrades
- ▶ Fall 2016: Heterogeneous hardware available

**Overall: 63 projects, 114 users, 44 institutions**

# PARTING THOUGHTS

- ▶ Work on your next research project @ [www.chameleoncloud.org](http://www.chameleoncloud.org)!

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

- ▶ Tell your colleagues about us
- ▶ Here today -- better tomorrow
- ▶ Creating a community
  - ▶ Workshops, traces, appliances, and other forms of engagement