Chameleon

www. chameleoncloud.org

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

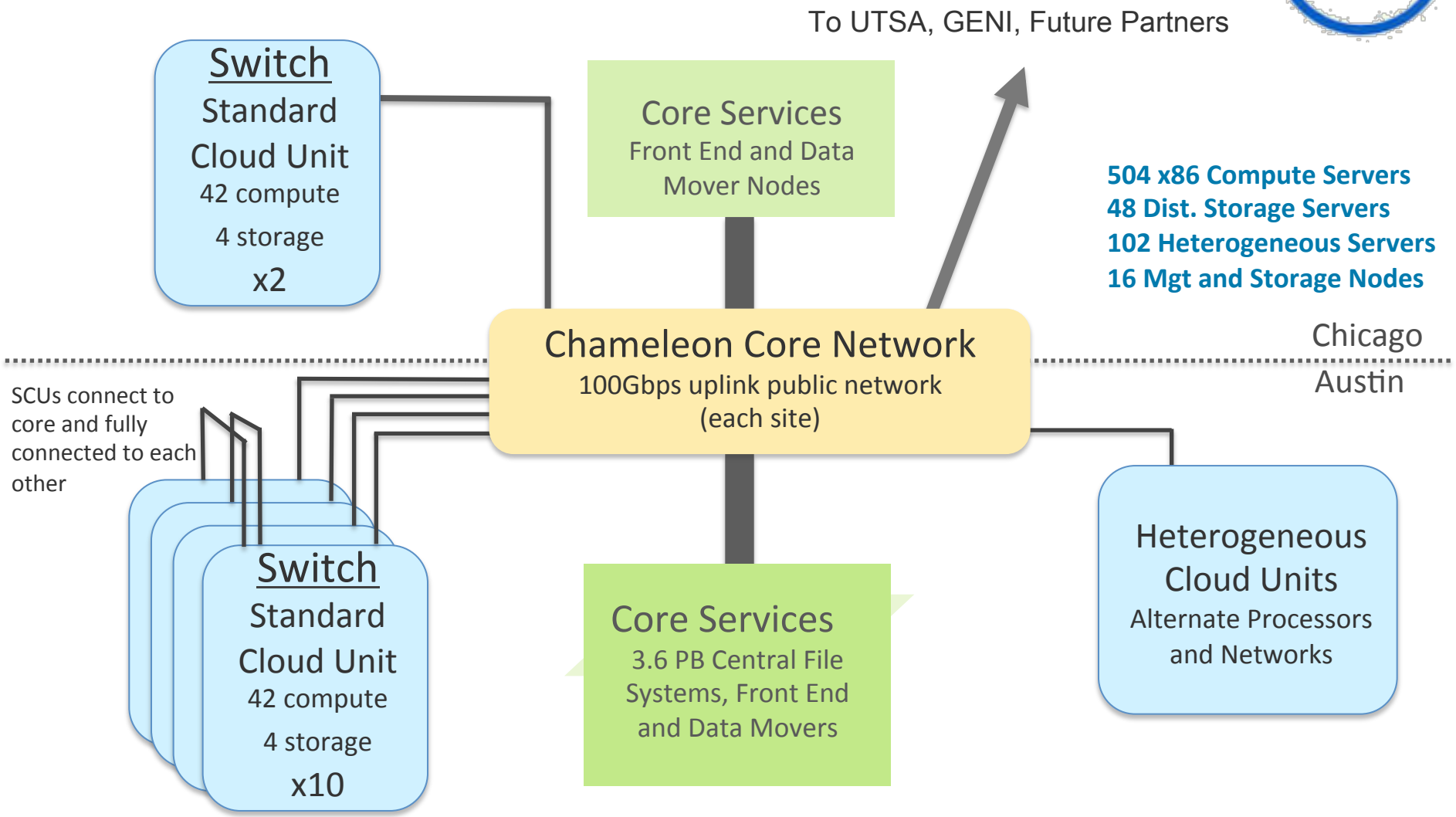
Kate Keahey

keahey@anl.gov

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



CAPABILITIES AND SUPPORTED RESEARCH

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

Persistent, reliable, shared clouds

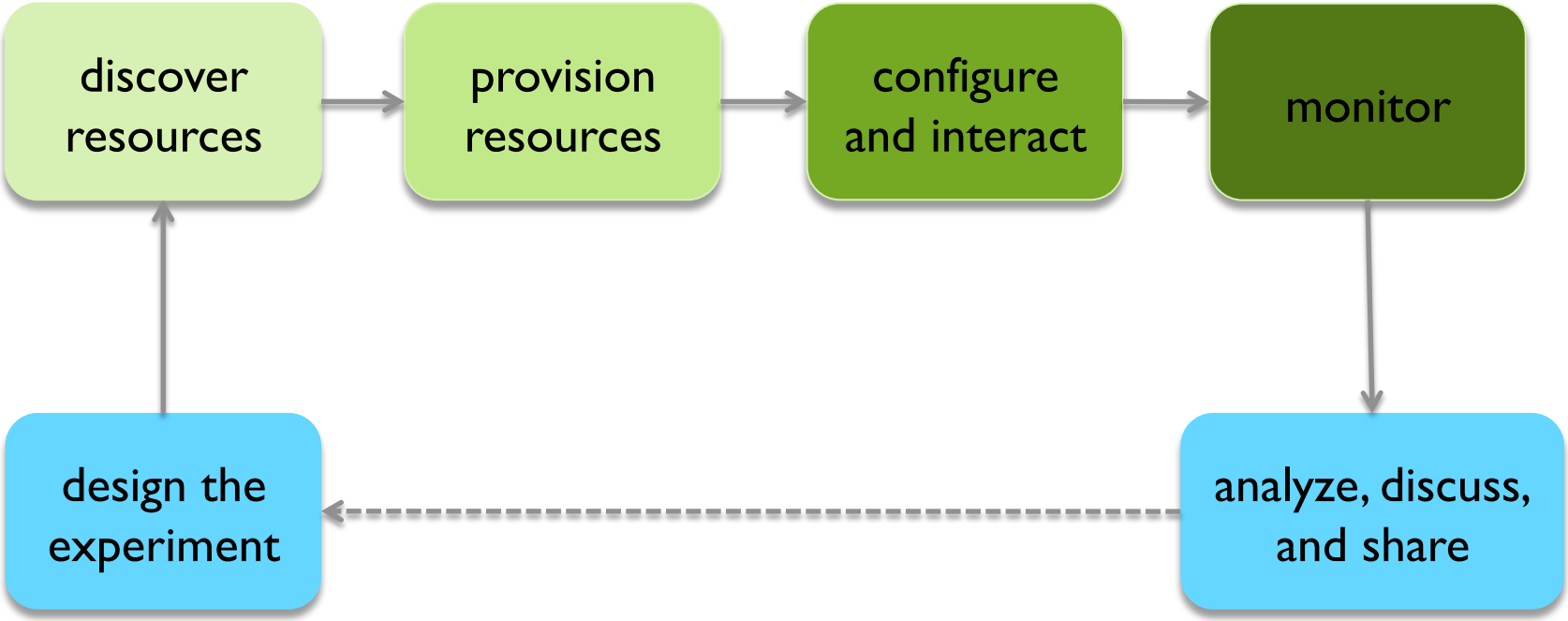
Repeatable experiments in new models, algorithms, platforms, auto-scaling, high-availability, cloud federation, etc.

Isolated partition, Chameleon Appliances

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

Isolated partition, bare metal reconfiguration: OpenStack and Grid'5000

EXPERIMENTAL WORKFLOW



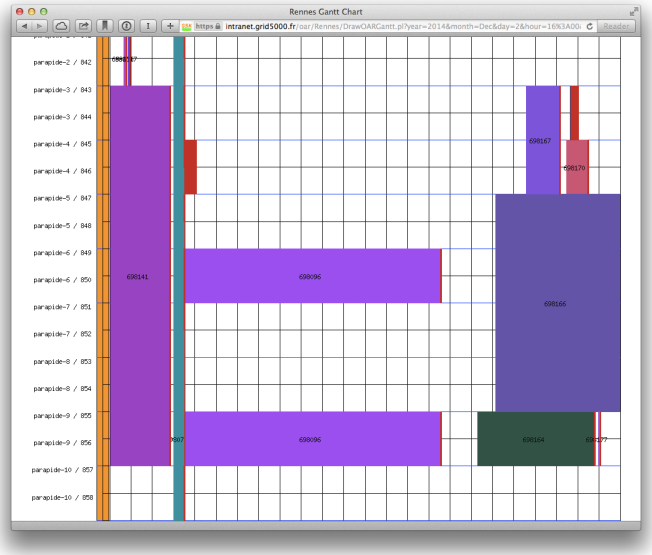
SELECTING AND VERIFYING RESOURCES

- 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

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
- Eventually: match making, Gantt chart displays



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

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

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 SCHEDULE

- Now: FutureGrid@Chameleon
  - Chameleon Technology Preview
  - OpenStack FutureGrid-style cloud
  - 43 projects, 81 users, 29 institutions
- Summer 2015: New hardware: large-scale homogenous partitions available to Early Users
- Fall 2015: Large-scale homogenous partitions and bare metal reconfiguration generally available
- 2015/2016: Refinements to experiment management capabilities, higher level capabilities
- Fall 2016: Heterogeneous hardware available

GET ENGAGED

- www.chameleoncloud.org
- Use the FutureGrid@Chameleon KVM cloud
- Technology Preview on FutureGrid hardware
- Early User Program
  - Committed users, driving and testing new capabilities, enhanced level of support
  - Sign up to get access to new hardware

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