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

Principal Investigator: Kate Keahey
CHAMELEON: A POWERFUL AND FLEXIBLE EXPERIMENTAL INSTRUMENT

- Large-scale instrument
  - Targeting 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 instrument
  - Bare metal reconfiguration, operated as single instrument, graduated approach for ease-of-use

- Connected instrument
  - Workload and Trace Archive
  - Partnerships with production clouds: CERN, OSDC, Rackspace, Google, and others
  - Partnerships with users

- Complementary instrument
  - Complementing GENI, Grid’5000, and other testbeds

- Sustainable instrument
  - Industry connections
CHAMELEON HARDWARE

Switch
Standard Cloud Unit
42 compute
4 storage
x2

Core Services
Front End and Data Mover Nodes
3 PB Central File Systems, Front End and Data Movers

Chameleon Core Network
100Gbps uplink public network (each site)

Switch
Standard Cloud Unit
42 compute
4 storage
x10

Heterogeneous Cloud Units
Alternate Processors and Networks
504 x86 Compute Servers
48 Dist. Storage Servers
102 Heterogeneous Servers
16 Mgt and Storage Nodes

To UTSA, GENI, Future Partners

Chicago
Austin

SCUs connect to core and fully connected to each other
# CAPABILITIES AND SUPPORTED RESEARCH

<table>
<thead>
<tr>
<th>Development of new models, algorithms, platforms, auto-scaling HA, etc., innovative application and educational uses</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Persistent, reliable, shared clouds</em></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Repeatable experiments in new models, algorithms, platforms, auto-scaling, high-availability, cloud federation, etc.</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Isolated partition, pre-configured images reconfiguration</em></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Virtualization technology (e.g., SR-IOV, accelerators), systems, networking, infrastructure-level resource management, etc.</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Isolated partition, full bare metal reconfiguration</em></td>
</tr>
</tbody>
</table>
SOFTWARE: CORE CAPABILITIES

Pre-configured Image Catalog
Bare metal images

User-Deployed Clouds

Persistently Deployed Clouds
OpenStack

Provisioning, Network, Scheduling and Orchestration
Linux Operating System Framework (LosF), (TACC)
KaDeploy, KaVLAN, OAR2, (Grid'5000)
Ironic, Neuron, OnMetal (OpenStack, Rackspace)
Orchestration: Nimbus, Interactive Experiment Management
EXPERIMENT WORKFLOW

- User interface: log in, manage profile
- Find Resources
  - Machine-parsable description (JSON)
  - Versioning (hardware upgrades, etc.)
  - Verification (maintenance, failures, etc.)
- Reserve Resources (browsing vs matching)
- Reconfigure testbed
- Shape experimental conditions
- Monitoring and metrics
  - Including fine-grain and energy monitoring
- Integration with workload generators, simulation, etc.
OUTREACH AND ENGAGEMENT

- Early User Program
  - Committed users, driving and testing new capabilities, enhanced level of support

- Chameleon Workshop
  - Annual workshop to inform, share experimental techniques solutions and platforms, discuss upcoming requirements, and showcase research

- Advisory Bodies
  - Research Steering Committee: advise on capabilities needed to investigate upcoming research challenges
  - Industry Advisory Board: provide synergy between industry and academia
PROJECT SCHEDULE

- **Fall 2014**: FutureGrid resources at UC and TACC available as OpenStack clouds
- **Spring 2015**: Initial bare metal reconfiguration capabilities available on FutureGrid UC&TACC resources for Early Users
- **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
- **Fall 2016**: Heterogeneous hardware available
TEAM

Kate Keahey  
Chameleon PI  
Science Director

Joe Mambretti  
Programmable networks

DK Panda  
High-performance networks

Paul Rad  
Industry Liason

Dan Stanzione  
Facilities Director

Warren Smith  
Director of Operations
PARTING THOUGHTS

- Large-scale, responsive experimental testbed
  - Targeting critical research problems at scale
  - Evolve with the community input
- Reconfigurable environment
  - Support use cases from bare metal to production clouds
  - Support for repeatable and reproducible experiments
- One-stop shopping for experimental needs
  - Trace and Workload Archive, user contributions, requirement discussions
- Engage the community
  - Network of partnerships and connections with scientific production testbeds and industry
  - Partnerships with existing experimental testbeds
  - Outreach activities
- Come visit us at www.chameleoncloud.org!