



www.chameleoncloud.org

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

Metropolitan Research and Education Network (MREN)

Executive Committee Meeting

National Center for Supercomputing Applications (NCSA)

Urbana Illinois

August 20 2015

JANUARY 12, 2016

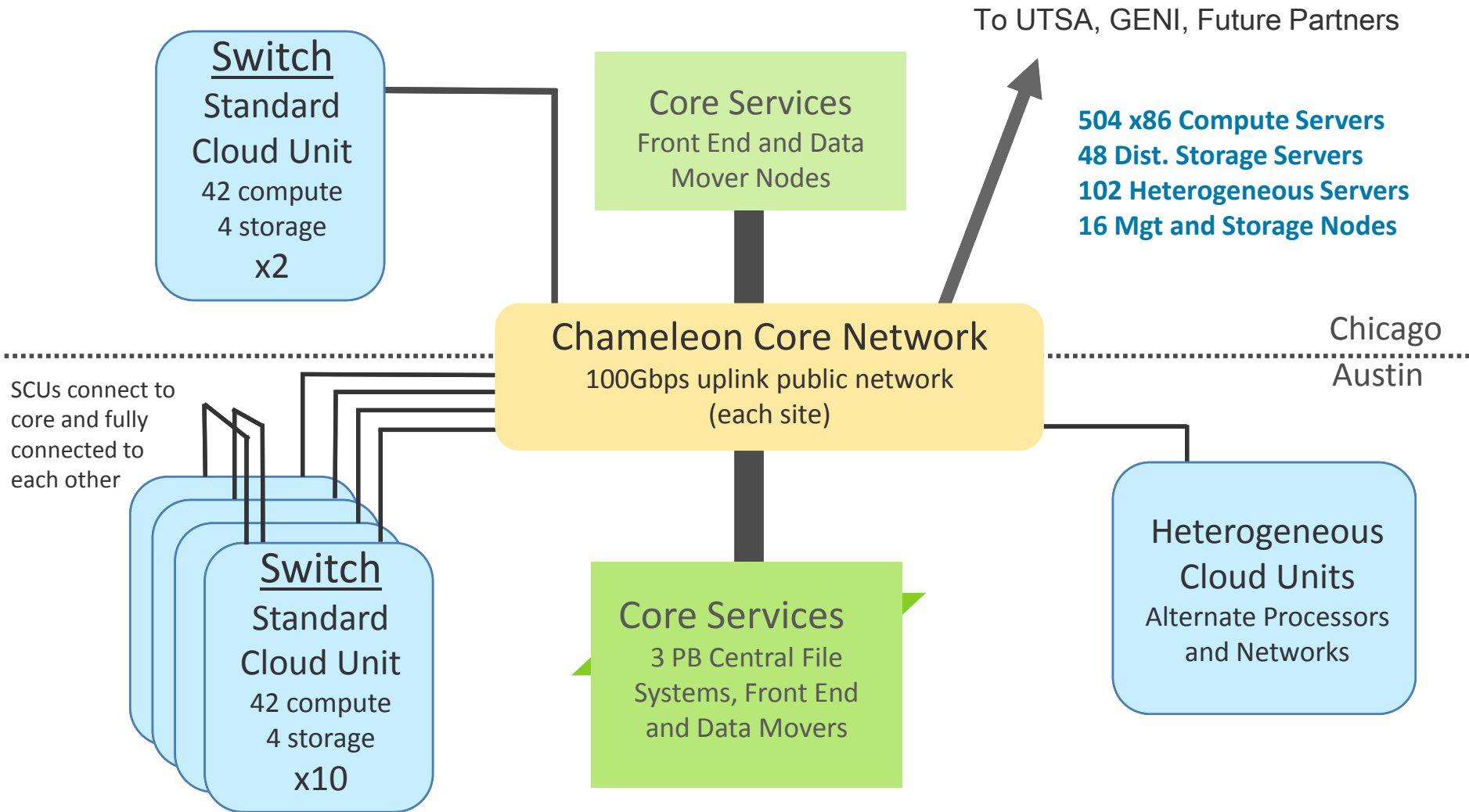
I



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



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, pre-configured images reconfiguration

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

Isolated partition, full bare metal reconfiguration

SOFTWARE: CORE CAPABILITIES

Persistent Clouds

OpenStack

User-Deployed Clouds

Pre-configured Image Catalog

Bare metal images

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

PARTNERSHIP WITH GENI COMMUNITY

- ▶ Chameleon Will Enable 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.
- ▶ The Architectural Options For Blending Chameleon and GENI Are Being Discussed

EXPERIMENTER RESOURCES

- ▶ Unique Deeply Programmable Blending Of Resources.
- ▶ Extremely Close Integration (Ideally, Seamless) of Programmable Networks With Programmable Clouds
- ▶ A Richly Resourced Platform For Experimentation With Exceptional Novel Architecture.
- ▶ Enables the Creation of New, Highly Innovative Distributed Environments, Including Specialized/Customized vs Generic Environments
- ▶ Highly Granulated Levels of Experimental Research Control, Measurement, Analytics, Visualization, and Reproducibility.

CURRENT PROJECT SCHEDULE

- ▶ Now: FutureGrid@Chameleon
- ▶
- ▶ Winter 2014: Meetings, Workshops, Planning Discussions
- ▶ Spring 2015: Maintain the momentum: Initial bare metal reconfiguration 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 generally available
- ▶ 2015/2016: Refinements to experiment management capabilities
- ▶ Fall 2016: Heterogeneous hardware available

TEAM

Kate Keahey
Chameleon PI
Science Director



Paul Rad
Industry Liason



Joe Mambretti
Programmable networks



Warren Smith
Director of Operations

DK Panda
High-performance
networks



Dan Stanzione
Facilities Director



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!