CHAMELEON: RECONFIGURABLE TESTBED FOR COMPUTER SCIENCE EXPERIMENTS

Kate Keahey
Mathematics and CS Division, Argonne National Laboratory
CASE, University of Chicago
keahey@anl.gov

September 4, 2019
Computational Research Leadership Council meeting
CHAMELEON IN A NUTSHELL

- We like to change: testbed that adapts itself to your experimental needs
  - Deep reconfigurability (bare metal) and isolation (CHI) – but also ease of use (KVM)
  - CHI: power on/off, reboot, custom kernel, serial console access, etc.

- We want to be all things to all people: balancing large-scale and diverse
  - Large-scale: ~large homogenous partition (~15,000 cores), 5 PB of storage distributed over 2 sites (now +1!) connected with 100G network...
  - ...and diverse: ARM, Atoms, FGPAs, GPUs, Corsa switches, etc.

- Sustainable operations: leveraging mainstream cloud technologies
  - Powered by OpenStack with bare metal reconfiguration (Ironic) + “special sauce”
  - Chameleon team contribution recognized as official OpenStack component

- We live to serve: open, production testbed for Computer Science Research
  - Started in 10/2014, available since 07/2015, renewed in 10/2017
  - Currently 3,500+ users, 500+ projects, 100+ institutions
CHAMELEON HARDWARE

SkyLake
Standard Cloud Unit
32 compute
Corsa Switch
x2

SkyLake
Standard Cloud Unit
32 compute
Corsa Switch
x1

100Gbps uplink public network (each site)

Core Services
3.5PB Storage System

Core Services
0.5 PB Storage System

Chameleon Associate Site
Northwestern

GENI and other partners

Heterogeneous Cloud Units
GPUs (K80, M40, P100),
FPGAs, NVMe, SSDs, IB,
ARM, Atom, low-power Xeon

Chicago
Austin
CHAMELEON HARDWARE (DETAILS)

- “Start with large-scale homogenous partition”
  - 12 Haswell Standard Cloud Units (48 node racks), each with 42 Dell R630 compute servers with dual-socket Intel Haswell processors (24 cores) and 128GB RAM and 4 Dell FX2 storage servers with 16 2TB drives each; Force10 s6000 OpenFlow-enabled switches 10Gb to hosts, 40Gb uplinks to Chameleon core network
  - 3 SkyLake Standard Cloud Units (32 node racks); Corsa (DP2400 & DP2200) switches, 100Gb uplinks to Chameleon core network
  - Allocations can be an entire rack, multiple racks, nodes within a single rack or across racks (e.g., storage servers across racks forming a Hadoop cluster)

- Shared infrastructure
  - 3.6 + 0.5 PB global storage, 100Gb Internet connection between sites

- “Graft on heterogeneous features”
  - Infiniband with SR-IOV support, High-mem, NVMe, SSDs, GPUs (22 nodes), FPGAs (4 nodes)
  - ARM microservers (24) and Atom microservers (8), low-power Xeons (8)

- Coming soon: more nodes (CascadeLake), and more accelerators
EXPERIMENTAL WORKFLOW

- discover resources
  - Fine-grained
  - Complete
  - Up-to-date
  - Versioned
  - Verifiable

- allocate resources
  - Allocatable resources: nodes, VLANs, IPs
  - Advance reservations and on-demand
  - Isolation

- configure and interact
  - Deeply reconfigurable
  - Appliance catalog
  - Snapshotting
  - Orchestration
  - Networks: stitching and BYOC

- monitor
  - Hardware metrics
  - Fine-grained data
  - Aggregate
  - Archive

CHI = 65%*OpenStack + 10%*G5K + 25%*”special sauce”
LEAVING NO EXPERIMENT BEHIND...

Supporting research projects in architecture, operating systems design, virtualization, power management, real-time analysis, security, storage systems, databases, networking, machine learning, neural networks, data science, and many others.
BEYOND THE PLATFORM: BUILDING AN ECOSYSTEM

- Interacting with hardware providers
  - Bring Your Own Hardware (BYOH)
  - CHI-in-a-Box: deploy your own Chameleon site

- Helping our user interact – with us but primarily with each other
  - Facilitating contributions of appliances, tools, and other artifacts: appliance catalog, blog as a publishing platform, and eventually notebooks
  - Integrating tools for experiment management
  - Making reproducibility easier
CHI-IN-A-BOX

- CHI-in-a-box: packaging a commodity-based testbed
  - First released in summer 2018, continuously improving
- CHI-in-a-box scenarios
  - Independent testbed: package assumes independent account/project management, portal, and support
  - Chameleon extension: join the Chameleon testbed (currently serving only selected users), and includes both user and operations support Part-time extension: define and implement contribution models
  - Part-time Chameleon extension: like Chameleon extension but with the option to take the testbed offline for certain time periods (support is limited)
- Adoption
  - New Chameleon Associate Site at Northwestern since fall 2018 – new networking!
  - Two organizations working on independent testbed configuration
REPRODUCIBILITY DILEMMA

Should I invest in making my experiments repeatable?

Should I invest in more new research instead?

- Reproducibility as side-effect
  - Example: Linux “history” command
- Reproducibility by default: documenting the process via interactive papers
CHAMELEON JUPYTER INTEGRATION

- Combining the ease of notebooks and the power of a shared platform
  - Storytelling with Jupyter: ideas/text, process/code, results
  - Chameleon shared experimental platform

- JupyterLab server for our users
  - Just go to jupyter.chameleoncloud.org and log in with your Chameleon credentials

- Chameleon/Jupyter integration
  - Interfaces: python and bash for all the main testbed functions

- Templates of existing experiments

Screencast of a complex experiment: https://vimeo.com/297210055
SHARING, EXPERIMENTING, LEVERAGING

- Sharing Jupyter notebooks in Chameleon
  - Sharing with your project members via Chameleon object storage
  - Publish to github for versioning and sharing in wider circle
  - Informally: send via email
  - Challenges: more flexible sharing policy implementation, better integration with github to support more publishing and sharing, finding relevant research, publishing

- Integration with Zenodo
  - Import from GitHub
  - Publish to Zenodo
  - Sharing platform: thin layer of discovery, indexing, filtering, etc.
PARTING THOUGHTS

- Chameleon is a rapidly evolving experimental platform
  - Originally: “Adapts to the needs of your experiment”
  - Now also: “Adapts to the needs of its community and the changing research frontier”
- Towards an Ecosystem: a meeting place of users and providers sharing resources and research
  - Testbeds are more than just experimental platforms
  - Common/shared platform is a “common denominator” that can eliminate much complexity that goes into systematic experimentation, sharing, and reproducibility
  - A critical element in building a sustainable repeatability and replicability platform