



www.chameleoncloud.org

CHAMELEON: HOW TO BUILD A CLOUD++

Kate Keahey

Mathematics and CS Division, Argonne National Laboratory

CASE, University of Chicago

keahey@anl.gov

September 16, 2019

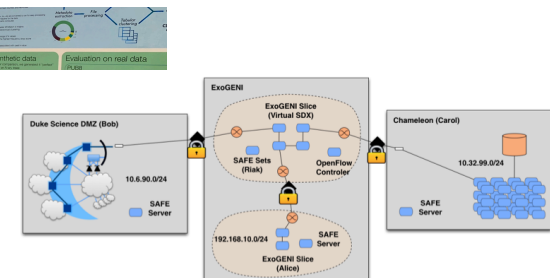
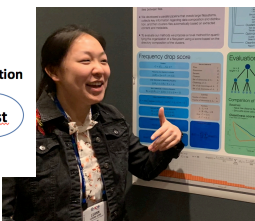
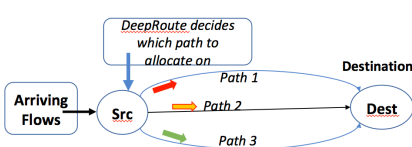
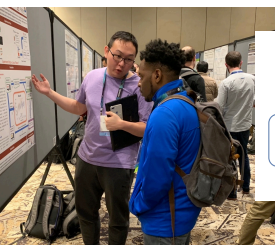
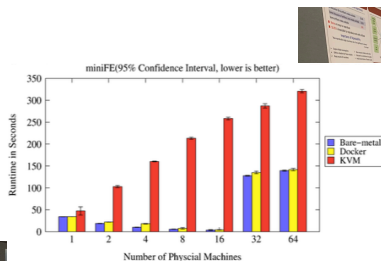
2019 NSF Workshop on Connecting Large Facilities and Cyberinfrastructure



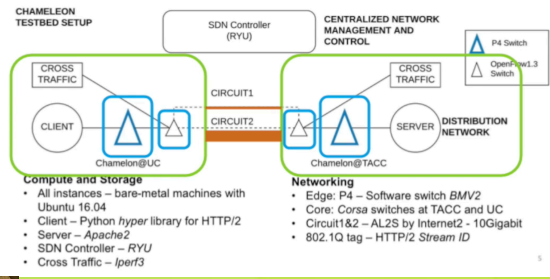
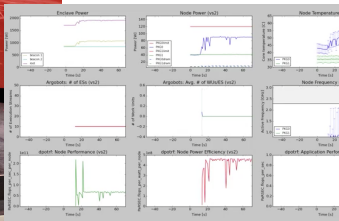
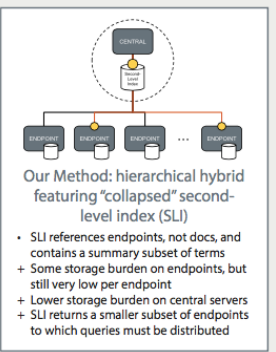
CHAMELEON IN A NUTSHELL

- ▶ 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: ARMs, Atoms, FPGAs, GPUs, Corsas switches, etc.
- ▶ Cloud++: 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, testbed available since 07/2015, renewed in 10/2017
 - ▶ Currently 3,500+ users, 500+ projects, 100+ institutions

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.



TOWARDS A REPRODUCIBILITY ECOSYSTEM

Should I invest in making my experiments repeatable?



Should I invest in more new research instead?

- ▶ Combining the ease of notebooks and the power of a shared platform
 - ▶ Storytelling with Jupyter: ideas/text, process/code, results – but limited containers
 - ▶ Chameleon: sophisticated experimental containers in need of “storytelling”
- ▶ Reproducibility by default: Chameleon + JupyterHub
 - ▶ Integrated Jupyter server
 - ▶ Python and bash interfaces for all the main testbed functions
 - ▶ Working with named containers
- ▶ Integration with Zenodo
 - ▶ Import/export of code/notebooks
 - ▶ Publishing via Zenodo: store your experiments and make them citable!



PARTING THOUGHTS

- ▶ Physical environment: a rapidly evolving platform implemented as cloud++
 - ▶ Specially adapted cloud with support for advanced cloud computing research
 - ▶ 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...
 - ▶ ... as well as education!
- ▶ Be part of the change: tell us what capabilities we should provide to help you share and leverage the contributions of others!