Introducing Configuration Management Capabilities into CloudLab Experiments

Dmitry Duplyakin
(presenting)

Robert P. Ricci

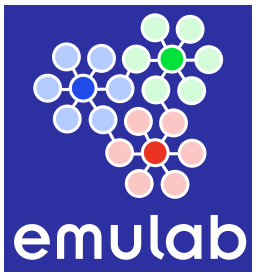
04/11/16

Computer and Network Experimental Research Using Testbeds, CNERT'16

This Talk Covers

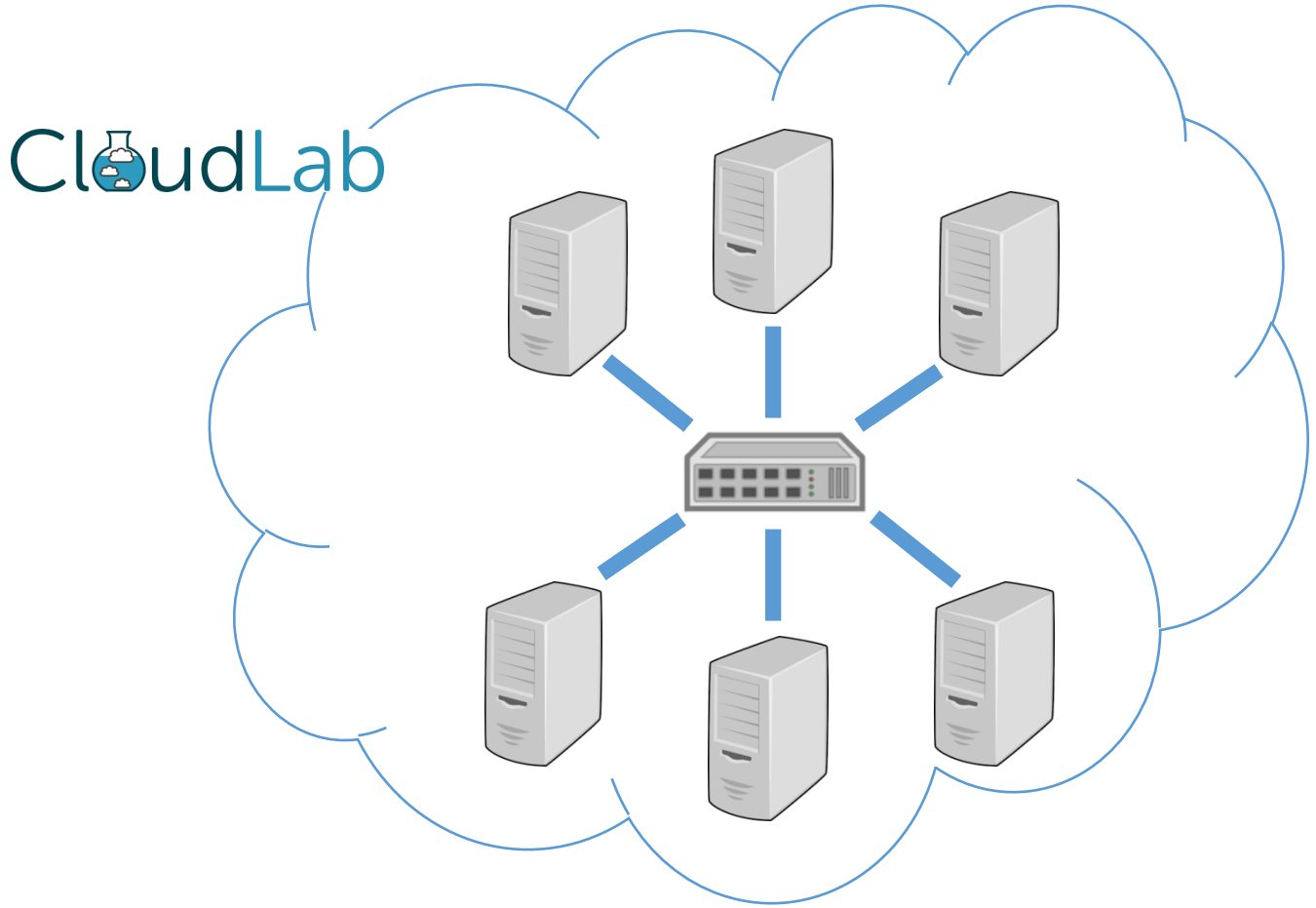
- Testbeds
- Experiments
- Configuration Management
- Demo
- Discussion

Long History of Testbed Research at Utah

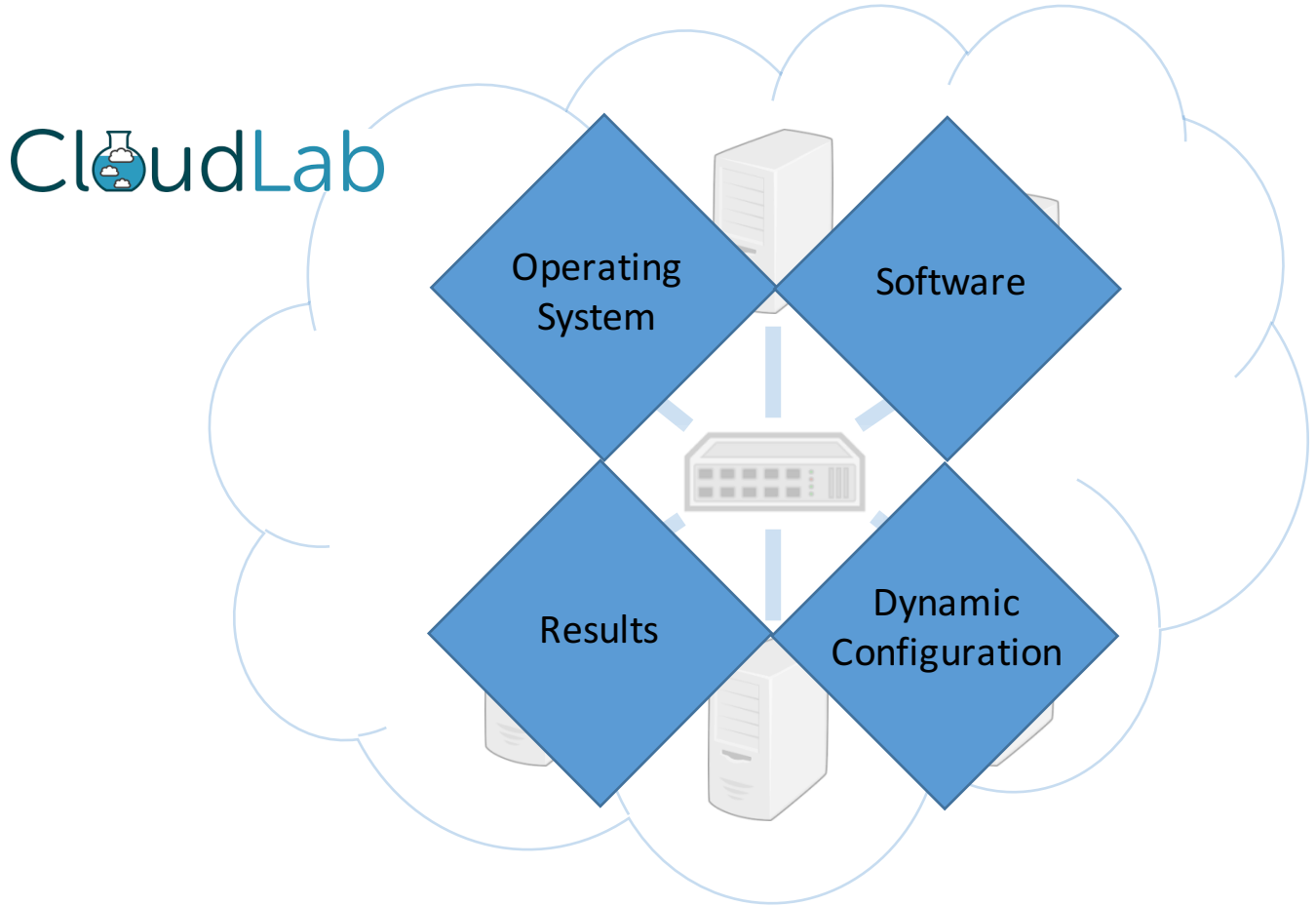


PhantomNet

Testbeds: Provide Isolated and Recreatable Environments

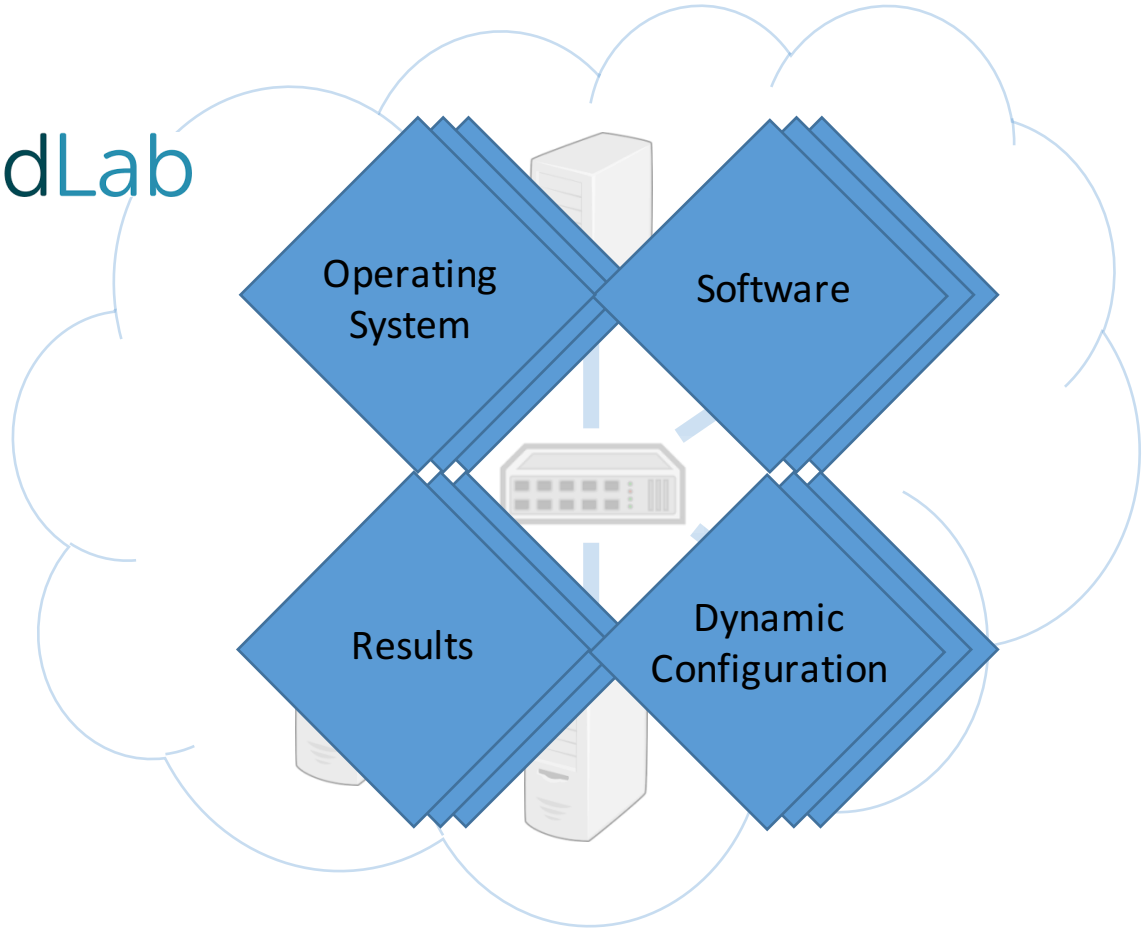


Experiments: Require Building Software Environments …

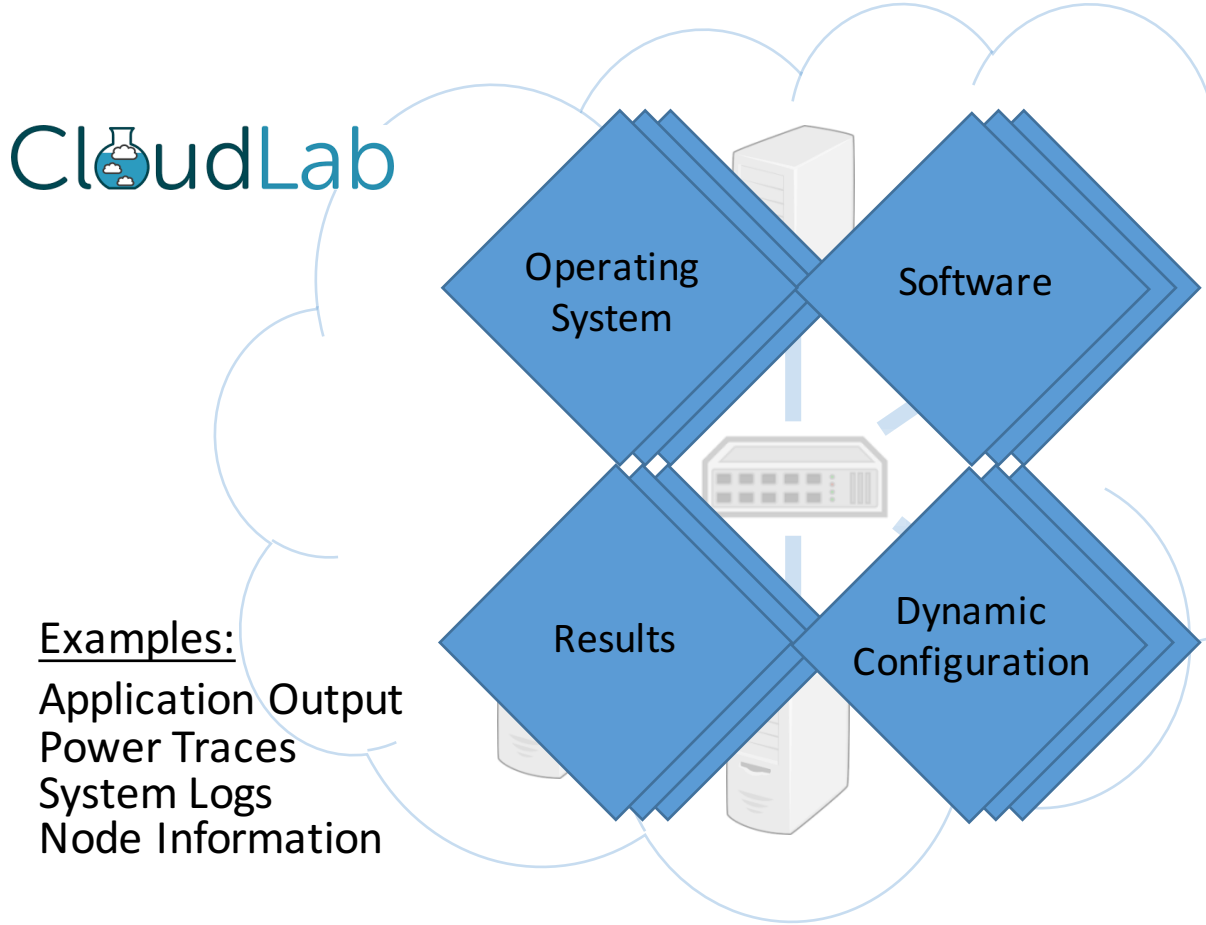


… on Many Nodes

CloudLab



… with Many Components



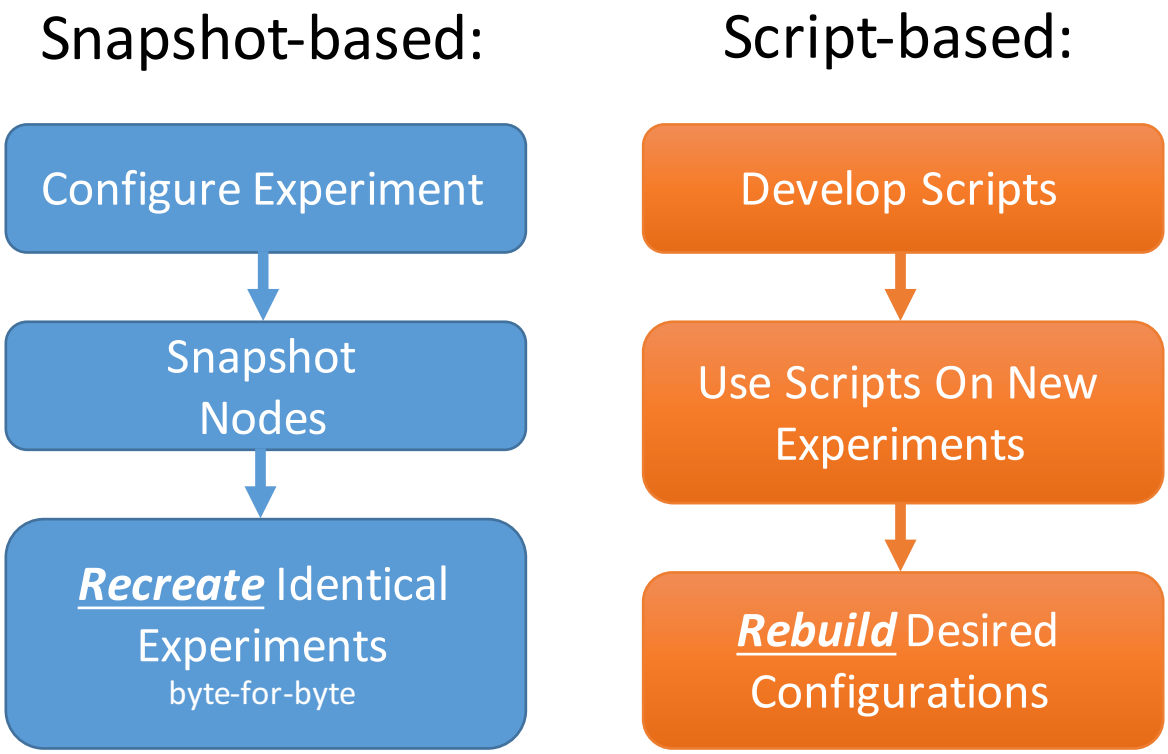
Examples:

Compiler
MPI library
Scheduler
Network File System
Numerical Libraries
Benchmarks

Examples:

Application Output
Power Traces
System Logs
Node Information

Common Workflows



Goal: Build Computing Clusters using Experiment Nodes

2 x 3 x 2 x … x 4

configurations

for controller and compute nodes

compiler versions

- GCC 4.8.2
- GCC 5.2
- GCC 5.3

library versions

- OpenMPI 1.8
- OpenMPI 1.10

architectures

- ARMv8 at Utah
- Haswell at Wisconsin
- Ivy Bridge at Clemson
- Sandy Bridge on Apt

Goal: Build Computing Clusters using Experiment Nodes

2 x 3 x 2 x … x 4

configurations

for controller and compute nodes

compiler versions

- GCC 4.8.2
- GCC 5.2
- GCC 5.3

library versions

- OpenMPI 1.8
- OpenMPI 1.10

new versions become available often

architectures

- ARMv8 at Utah
- Haswell at Wisconsin
- Ivy Bridge at Clemson
- Sandy Bridge on Apt

new machines are installed

new components are added

Goal: Build Computing Clusters using Experiment Nodes

2 x 3 x 2 x … x 4

configurations

for controller and compute nodes

compiler versions

- GCC 4.8.2
- GCC 5.2
- GCC 5.3

library versions

- OpenMPI 1.8
- OpenMPI 1.10

architectures

- ARMv8 at Utah
- Haswell at Wisconsin
- Ivy Bridge at Clemson
- Sandy Bridge on Apt

Problem: Explosion of Configurations!

(snapshot- and simple script-based approaches don't scale)

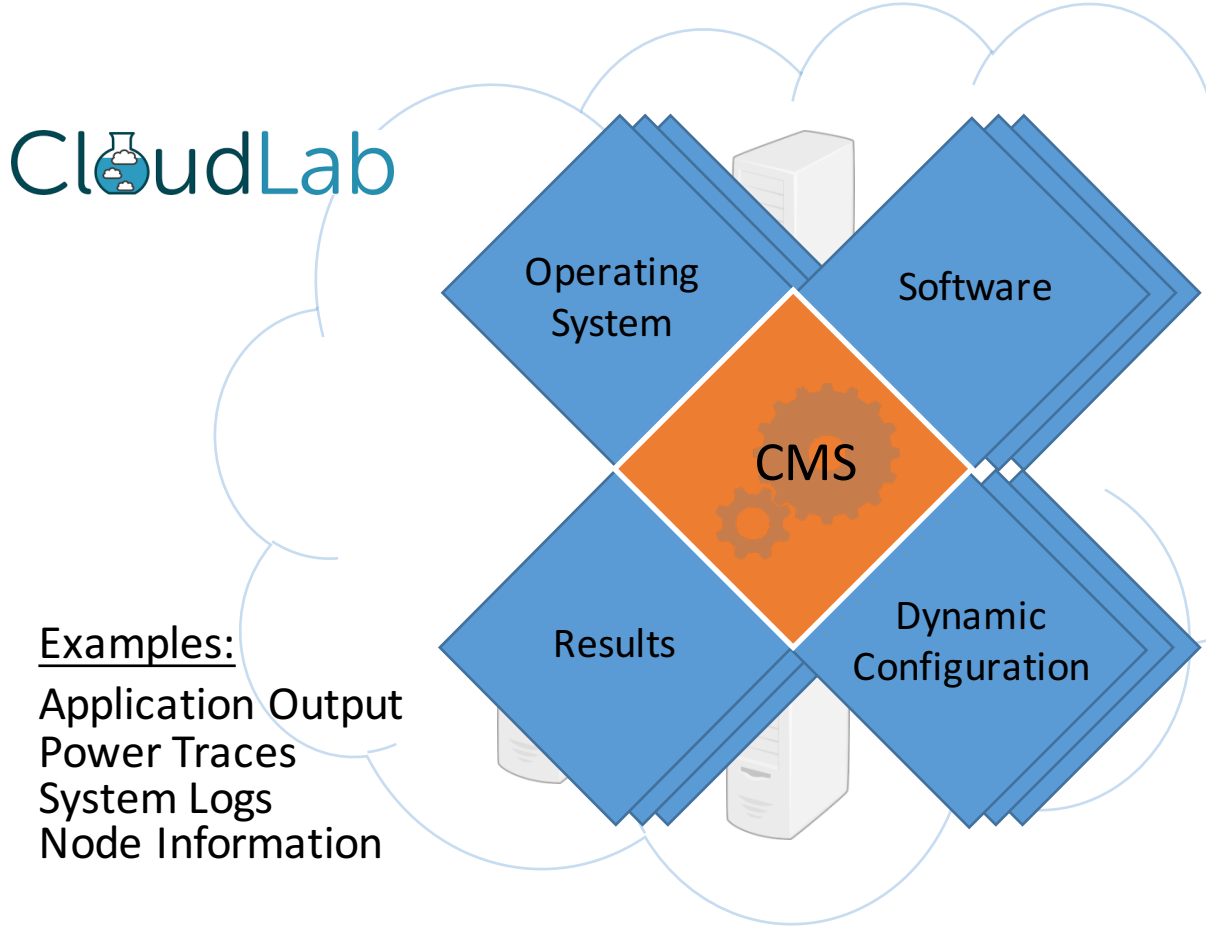
Proposed Strategies

- DevOps: treat infrastructure as code
- Develop modular, reusable infrastructure code
- Explore interactions between components
- Change perspective: from “bags of scripts” to “hierarchies of roles”
- Find balance between simplicity and flexibility
- Transparently support different hardware

Implementation

- Used Chef, a configuration management system (CMS)
- On CloudLab, built a profile that turns an experiment into a Chef "cluster"
  - https://www.cloudlab.us/p/utahstud/ChefCluster
- Enabled easy integration of public and private infrastructure code
  - Chef Supermarket and emulab/chef-repo on GitHub
- Developed *recipes*, *cookbooks*, and *roles* for building computing clusters

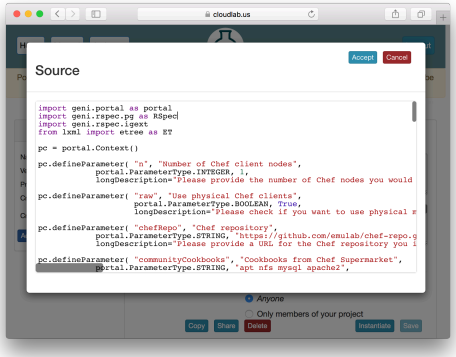
Configuration Management System: Enables "Orchestration"

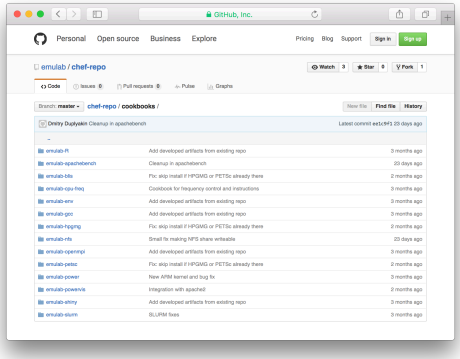


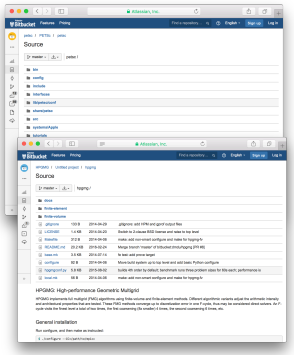
Demo

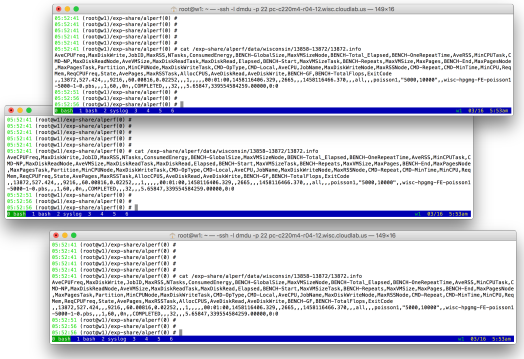
CloudLab, Chef, Supermarket

Putting Things Together









CloudLab

Libraries,
Benchmarks

Results,
Logs

More information: hands-on tutorial at http://cloudlab.us/chef
(presented at the GENI Regional Workshop at ASU in March 2016)

Thank you!

dmitry.duplyakin@colorado.edu

ricci@cs.utah.edu