External Login at Scale

Georg Rath, Doug Jacobsen, Tina Declerck
Cray User Group 2018
Problem Statement

- Two Cray XC systems, one CS system hosting two clusters
  - No single management solution provided by vendor
  - Rhine on XC, custom xCAT on CS
  - Cray-Ansible on XC, cfengine on CS
- Different userlands, kernels, user experience
  - Cost of keeping in sync prohibitive
Project Goals

• Reduced operational effort
  • Leverage SMWFlow process
  • Operator knowledge applicable to all systems
• Single user environment
  • Leveraging IMPS image building
• Federated systems
  • One center, one workload manager
• Quicker turnaround for new systems
  • Testbeds, Proof-of-Concepts, Customer driven one-offs
System View

smflow single repository:
Roles, Parameters, Node Groups, etc

provisions

Cori SMW
manages

Hosts

Edison SMW
manages

Hosts

Denovo SMW
manages

Hosts

PDSF SMW
manages

Hosts
eLogin Infrastructure

- Images are created by IMPS
- ConfigSets are created by CMF
- Fused by the provisioning system
- Elogin provisioning is OpenStack until CLE6-UP05
  - CLE6-UP06 OpenStack was replaced by custom Cray
The NERSC Way

- Modify ConfigSet (set up NodeGroups, no PE)
- Export image and ConfigSet from SMW
- Prepare Image
  - Merge ConfigSet with image
  - Remove assumptions about local disks
  - Remove OpenStack dependency
- Deploy image to compute node (stateful and manual)
- Boot stateless using xCAT
- Pivot into image using kexec
- 16 node test cluster (Gregor) running this now
Booted eLogin

Welcome at mc0801

Login:
Benefits

- All system have the same management
  - Pooling of efforts
  - Consistent experience for users (and admins!)
- Move from homegrown to supported*
- Fix once, deploy everywhere
- Template for rapid systems deployment
Work in Progress

- Adapting image to work with OpenSUSE
  - Licensing
  - Bootable prototype achieved
- Automate image deployment with xCAT
  - Evaluated udpcast, bittorrent
  - Most likely will use tiered HTTP (ubiquitous, simple)
- Move to UP06 SMW-integrated management
Roadmap

- Summer 2018: Move Denovo to OpenCray
- As soon as UP06 is deployed on test systems:
  - Move from xCAT to Cray
- Fall 2018: Test Cluster with external Lustre
- 2019: Adapt Slurm Federation
Thank You