Managing and Monitoring a Scalable Lustre® Infrastructure

LUG 2013
Makia Minich (makia_minich@xyratex.com)
What this presentation is not...

- A Lustre® presentation
  - We've had a lot of those, how about a break for a few minutes?
- An evil vendor presentation
  - Makia != Sales-Droid
What will we talk about then?

• What are we trying to manage?
• What are the important bits?
• What tools are available?
• Where do we go from here?
What are we managing?
What are we trying to manage?
What are we trying to manage?

- Clouds
- Clusters
- Filesystems
- Groups of Nodes
- Multiple Data Centers

What we need is a scalable toolset that allows us to quickly see what is going on and what is going wrong?
What is important?
What is important?

- Scalability
- Installation
- Monitoring
- Multi-node administration
- Multi-node configuration
What tools are available?
What tools are available?

- Basic Linux Tools
  - NTP
  - SSH
  - DHCP and PXE
  - cron and at
  - syslog
What tools are available? - Admin

- **Diskfull Installation**
  - Tools that manage image or package based installs
  - Kickstart
  - YACI

- **Diskless Administration**
  - Tools that manage images used for diskless deployments
  - OneSIS
  - GeDI

- **Configuration Management**
  - Centralized place to manage and distribute configuration files
  - puppet
  - chef
What tools are available? - Execution

- Parallel Shells
  - Allow you to easily run commands across the system
  - pdsh
  - dsh

- Execution Frameworks
  - Environment to create more complex jobs based on resources
  - ClusterShell (python)
  - SLURM (resource manager)
What tools are available? - Monitoring

- Tools that provide full system monitoring
- Nagios/Icinga - polling
- Ganglia - multicast
- LMT - lustre
What tools are available? - Advanced

• Failover
  o Adding fault tolerance to the system
  o Corosync/Heartbeat
  o Pacemaker

• Console/Power Access (IPMI)
  o Allowing low-level control over the system remotely
  o IPMI tools
  o Conman/Powerman
Minor Rant -- Firmware

• Not such a pretty story
  o Mixture of tools depending on the hardware
  o Switches
  o Disks
  o Enclosures
  o BIOS (coreboot?)
  o ...

• Varied solutions
  o Network Boot DOS
  o Local commands
  o Out of band tools
Minor Rant - Firmware

Suggestions (which apply to more than just firmware)

• NO MORE DOS!
  o Linux command line is always a preference
• Minimize Reboots
  o Downtime is at a premium
• If it can be done out of band, make the tools available and usable
Where do we go from here?
Where do we go from here?

• We have lots of tools to choose from...
• BUT, we're missing the "expertise"
  ○ "Update: What's Missing from HPC?"
  ○ http://youtu.be/v0G8g1p01QQ
  ○ Where do we get "new blood"?

Need more "Cluster Challenge" type activities ...
• Builds awareness and expertise
• Gets new eyes on the problems

YouTube: "Update: What's Missing from HPC?"
Where do we go from here? - Cluster Suites

Vendor options for toolsets:

- Bright Cluster Manager
- Chroma
- Clusterstor Manager / Sonexion Manager
- Redhat Cluster Suite
Thank You

makia_minich@xyratex.com