Why Your Hadoop Cluster Needs a Cluster Manager

Robert Stober
Solutions Architect
A Hadoop Cluster is more than just Hadoop
There’s a lot more to managing a Hadoop cluster than just managing Hadoop.

- **Provisioning** - Hadoop can’t be used until an OS has been provisioned onto the cluster nodes.
- **Configuration** - Configure network interfaces, disks and users.
- **Updates** - Deploy software and OS patches.
- **Management** - Changing configurations, replacing failed hardware, and power management.
- **Health Management** - Metrics, health checks, notifications and automatic corrective actions.

Hadoop Manager tools only manage Hadoop, but they don’t provide cluster management.
Building a Hadoop Cluster is Difficult

- Most cluster management solutions use the “toolkit” approach (Linux distro + tools)
  - Tools typically used: Ganglia, Cacti, Nagios, CFEngine, System Imager, xCAT, Chef, Puppet, ROCKS, Warewulf, Knife, Crowbar, etc.

- The problem with the “toolkit” approach:
  - Tools rarely designed to work together
  - Each tool has its own command line interface and GUI
  - Tools rarely designed to scale
  - Each tool has its own daemon/agent and database
  - Tools rarely designed for HPC
  - Roadmap dependent on developers of the tools

- Making a collection of unrelated tools work together requires a lot of expertise and scripting.
- This complexity makes adoption difficult and expensive.
Bright Cluster Manager for Apache Hadoop makes it easy to install and manage Hadoop clusters.

Become productive immediately - no need to learn new cluster management frameworks.

Evaluate distributions quickly and easily.

Lower administration load and cost than other management solutions considered.

Manage hybrid HPC/Hadoop infrastructure using a single-pane-of-glass.
Questions?

Please contact us:
www.brightcomputing.com/Contact