Crowbar

Deploying Cloud Benchmarks

Nicholas Wakou
Dell Cloud Computing Solutions
5/09/2012
For the next 5 mins .......

- Big Data realities
- Crowbar
  - Overview
  - What is does
  - High-level components
- Dashboard
- Conclusion
Reality 1: We are experiencing massive data growth.

- Every 18 months, non-rich structured and unstructured enterprise data doubles.
- Traditional data warehouse data volumes, while growing, are a small fraction of the data challenge.
- Some 80%† of enterprise data resides in other data sources.

Petabytes


Reality 2: The amount of time required to benchmark big data will grow exponentially

- There is pain in running DW benchmarks
  - Cost
  - Time to Deploy
  - Time to Configure (baselining)
  - Time to Tune
  - Technical expertise
  - Market relevance
- Benchmark SUTs require more resources than production systems
- Big Data benchmark SUTs will require even more resources and need a lot more time to configure.
Crowbar : Overview

• Open source software framework
• Provides a modular platform for deploying large scale cloud infrastructure
• Automates required installation and configuration tasks from bare metal to deploying big data environments.
• Core capabilities:
  ➢ Hardware configuration – updating and configuring BIOS and BMC boards.
  ➢ Deployment of base operating system.
  ➢ Deployment of cloud/big data components.
  ➢ Providing core network infrastructure services (NTP, DNS, DHCP).
  ➢ Monitoring availability and performance of all deployed components.
Crowbar – What It Does

- Provides cloud foundation - Bare metal installation, configuration and ongoing configuration management
- Offers catalogue of services/ capabilities (barclamps) for reuse
- Tested on following browsers: Firefox 3.5+, Firefox 4.0, IE 7, Safari 5

Crowbar interfaces:

**Nodes**: Core barclamp that does initial configuration of physical or virtual servers to bring them online ‘unassigned’. Sets up
- Barclamp framework
- provisioning server,
- base environment and OS,
- network services (ntp, dns, dhcp, VLAN config)
- IP address location service,
- RAID controllers, BIOS and BMC boards.

**Barclamps**: Additional modular software capabilities, i.e. specific roles (titled active proposals), that can be installed on a selected ‘online’ nodes. Installs
- Cloudera Manager
- Pig
- Sqoop
- Hive
- Nagios
- Ganglia
- Other components will be added in future
Crowbar : High-level Components

- Nova
- Swift
- Glance
- Hadoop
- Nova Premium

Crowbar UI/API/CLI

Framework Applications (Nagios, Ganglia)

Chef Server / Crowbar Rails App

SledgeHammer
(Bios Update/Config, RAID Update/Config, BMC Update/Config)

Installation Services
(DHCP, DNS, NTP, TFTP)

Open Source Component

Dell IPEC Specific Component

Dell Solutions Specific Component
The main dashboard shows the nodes and if they are ready.

- The columns are the discovered switches and the nodes are under the switch they are connected to.
- Green icon means ready. Red icon means not ready. Next to the switch is a pie chart of the status of the underlying nodes.
- Selecting a node will bring up the details for that node. The stats and info about the node is displayed. Selecting a role in the role list will highlighting all nodes that also have that role.
Conclusion

• Crowbar is open-source - can be customized for benchmark environments
  • Many open-source developers are customizing it for their environments
  • Barclamps can easily be imported into Crowbar.
• Crowbar is being used in both physical and virtual environments
  • Developers test their code in virtual environments before testing on physical
  • Benchmark SUTs can be modeled on virtual environments
  • Performance parameter tuning on virtual clusters
• Remote access and management possible with Crowbar
• Crowbar reduces Time to Deploy (TtD)
  • Saves money
  • Ease of benchmarking
  • Repeatable and Reusable