Enabling BigData @ MOC

Presenter: Ata Turk
Team: Ugur Kaynar, Ravi Santosh Gudimetla, Sahil Nikhil Tikale, Jason Hennessey, Ata Turk, Peter Desnoyers, Orran Krieger
Goals of the MOC

- a Big data project
  - part of Mass Big Data Initiative
  - BigData Platform (BDP) research in Commonwealth

- users & government to easily upload and share datasets

- exploit innovative infrastructure:
  - e.g., SSDs, low-latency networking, RDMA, accelerators…

- level field for BDP innovation
  - research/innovation in a production environment
Background and Motivation

• Large-scale BigData deployments are dedicated bare-metal
• Most BigData deployments are moving to the cloud:
  • Access to rich data sets
  • Per-tenant environments
  • Elasticity, on-demand usage
• All single provider public clouds serve BigData service as a base offering
  • e.g. Elastic Map Reduce (EMR) from AWS
• Innovation by third parties difficult
BigData @ MOC

• Explore Containerized, Virtualized, Bare-Metal

• Containers
  • problems in isolation & multi-tenancy

• OpenStack Sahara & Apache BigTop for OpenStack

• Focus on bare-metal setups
  • not investigated in detail, gaining traction
Advantages of Bare-Metal

- Better performance
  - underlying storage HDDs
- Predictable performance
  - I/O intensive workloads (e.g. Hadoop)
- Better isolation, multi-tenancy, security
- Support for special hardware configs…
- Where there is demand there will be supply
  - IBM SoftLayer, Rackspace, Internap, …
Design

0: number of nodes, user public key
1: reserve servers
2: clone golden image for servers
3: spin-up per-tenant Headnode & Services
4: chroot images for per-node config
5: PXE-boot servers
6: boot using cloned images
7: return public IP of master server
Current status

• A POC prototype that shows we can serve any size Bare-Metal BigData cluster within minutes (~5 mins for now)

• HaaS automation
  • dynamic node allocation and network isolation

• CEPH cloning & image management

• Per-tenant headnode & iSCSI gateway

• Full support for multi-tenant model

• NAT service for remote access
Check out BigData Demo

Team: Ugur Kaynar, Ravi Santosh Gudimetla, Sahil Nikhil Tikale, Jason Hennessey, Ata Turk
Positive Properties of this Design

• Avoid decoupling data and compute
  • avoid I/O redirect over network, use local storage for processed data

• Efficient support for multi-tenancy, isolation
  • HaaS for isolating per-tenant network / project

• Retain benefits of image management
  • image management for different BigData configurations

• Avoid long OS boots: Bare-metal BigData on-demand (mins)
  • network boot OS and BigData applications from shared storage
  • avoid initial OS deployment overhead