Virtualizing Big Data/Hadoop Workloads

Update for vSphere 6

Justin Murray
VMware
Agenda

• The Hadoop Customer Journey
• Why Virtualize Hadoop?
• vSphere Big Data Extensions and Project Serengeti
• Performance and Reference Architectures
• References
• Conclusion
Customer Stages on the Hadoop Journey

**Stage 1: Piloting**
- Often start with a line of business
- Try 1 or 2 use cases to explore the value of Hadoop

**Stage 2: Hadoop in Production**
- Serves a few departments/BUs
- A few more use cases
- Core Hadoop + components

**Stage 3: Cloud Analytics Platform**
- Serve many departments
- Often part of mission critical workflow
- Integrated with other big data services
Why Virtualize Hadoop?
Executive Summary

- Increased hardware resource utilization
- Scale out and scale in a cluster at will
- Hadoop cluster isolation (same as the hardware does, using resource pools)
- No degradation in performance

- Certified by Cloudera and Hortonworks
Increase Utilization to Control Costs

- Consolidated cluster has access to entire pool of physical resources
- Take advantage of multi-tenancy to increase utilization during non-peak hours
- Reduce latency on priority jobs on consolidated cluster
Virtualizing Big Data - Value Propositions

**Operational Simplicity with Flexibility**
- Rapid deployment of clusters
- Self service tools
- Avoid dedicated hardware

**Maximize Resource Utilization**
- Scale out and scale in
- VM-based isolation
- Increase resource utilization
- Resource pool-based prioritization

**Architect Scalable Platform**
- Deployment choice
- Maintain management flexibility at scale
- Control Costs
- Leverage vSphere features
Hadoop 2.0 – Yet Another Resource Negotiator (YARN)
A Virtualized Hadoop 2.0 Cluster
Skyscape

- A UK company that provides cloud computing services to the UK Government’s G-Cloud initiative.
- Skyscape offers IaaS, PaaS, SaaS.
- 5 customers lined up at the first day of GA.
- Expect to expand to 140 servers very soon.
- Skyscape Hadoop in the Cloud is built on top of BDE.
- Used BDE API extensively.

vSphere Big Data Extensions
Big Data Extensions - Highlights

- Open source project
- Tool to simplify virtualized Hadoop deployment & operations
- Virtualization changes for core Hadoop
- Contributed back to Apache Hadoop
- Complements resource management on vSphere
vSphere Big Data Extensions
Hadoop Cluster Deployment on VMware

On physical machines

1. Server Preparation
2. OS Installation
3. Network Configuration
4. Hadoop Installation and Configuration

On VMware

- Big Data Extensions for VM creation, configuration, start-up
- Big Data Extensions or other Hadoop Management Tool
One Click to Scale out the Cluster on the Fly
BDE Allows Flexible Configurations

```
{
  "name": "master",
  "roles": [
    "hadoop_namenode",
    "hadoop_resourcemanager"
  ],
  "instanceNum": 1,
  "instanceType": "LARGE",
  "cpuNum": 2,
  "memCapacityMB": 4096,
  "storage": {
    "type": "SHARED",
    "sizeGB": 20
  },
  "haFlag": "on",
  "rpNames": [
    "rpl"
  ]
}
{
  "name": "data",
  "roles": [
    "hadoop_datanode"
  ],
  "instanceNum": 3,
  "instanceType": "MEDIUM",
  "cpuNum": 2,
  "memCapacityMB": 2048,
  "storage": {
    "type": "LOCAL",
    "sizeGB": 50
  },
  "placementPolicies": {
    "instancePerHost": 1,
    "groupRacks": {
      "type": "ROUNDROBIN",
      "racks": ["rack1", "rack2", "rack3"]
    }
  }
}
```

- **Storage configuration**: Choice of shared or local
- **High Availability option**: HA Flag
- **Number of nodes and resource configuration**: Instance number, type, CPU, memory capacity
- **VM placement policies**: Instant per host, group racks (round robin)
Deployment Options with Big Data Extensions

BDE Original Style

BDE provisions VMs and installs the Hadoop software from a local YUM repo

BDE 2.0

BDE provisions base VMs

Hadoop management tool installs software

BDE 2.1 (shipped Oct ‘14)

BDE creates VMs and calls management tool API

Hadoop management tool installs software under the hood
Enhancements in BDE 2.2

GA : 4th June 2015
Future Improvements

• Better **infrastructure management**
  – Environment Checking
  – FQDN management
  – Centralized user management
  – Shrink clusters
  – InstantClone

• Further 3rd Party App Manager integration
Environment Checking

• Problem
  – Pre-requisite requirements that Hadoop and BDE depend on.
  – When the pre-requisites are not set up correctly, especially network related items, problems can occur.
  – It can take a while to troubleshoot these issues

• Solution
  – We are providing a list of items to check with specific steps **before you provision a cluster**
  – May become a script that can be run to diagnose the environment.
Shrink a Cluster

• Problem
  – BDE did not provide a straightforward way to reduce the number of (compute) nodes in a cluster. May want to shrink the cluster after it finishes processing a known spike in the workload.

• Solution
  – Use cluster resize command or UI to reduce the number of virtual machines in a specified group.
  – Targeting stateless nodes (NodeManager, JobTracker etc.)
  – 3rd party App Mgr will be notified that this is happening
Add the Cloudera Manager AppManager into BDE
Add the Ambari App Manager into BDE
InstantClone – vSphere 6
InstantClone - “Linked Clone” of Memory

- COW Memory Clone
- VMFork
- Same CPU Configuration
- Linked Clone to create Delta Disks

Parent VM

Child1 VM

Child2 VM
InstantClone: Value Proposition

- Instant provisioning of ready-to-go virtual machines
  - Linux VMs in ~0.5s
  - Windows VMs in ~5s
  - Ongoing work to reduce these times even further
- Significant scale-out with little overhead
  - 60 Linux VMs instantiated in ~7.5s
  - Scales with number of cores
- Memory consolidation
  - If many VMs share common applications
  - Launch common applications then clone
BDE Provisioning Optimizations

- **Overarching principle: Make the common case much faster**

- One parent template virtual machine per host
  - Steady state: No templates are cloned

- Any desired virtual machines created as forked children
  - Potentially different CPU, Memory, disks
  - Some persistent (e.g., master) some non-persistent (e.g., workers)

- Possible other optimization
  - Create parent template hierarchy for each “type” of VM (e.g., master, compute)
A BDE R&D “Fling” for Container Managers

- BDE for provisioning Mesos and Kubernetes
  

- Not part of the BDE product – unsupported so far, but very interesting to some users
Events Coming Up – Big Data Team present

• Vmworld 2015
  – US in late August
  – Europe, mid October
  – Good supply of talks, demos and HOLs there
• Strata+HadoopWorld in the Fall in New York – VMware will be there
• Hope to see you all at one or more of these events
Conclusions

- Hadoop workloads work very well on VMware vSphere
  - Various performance studies have shown that any difference between virtualized performance and native performance is minimal
  - Follow the general best practice guidelines that VMware has published

- vSphere Big Data Extensions enhances your Hadoop experience on the VMware virtualization platform
  - Rapid provisioning tool for deployment of Hadoop components in virtual machines
  - Design patterns such as data-compute separation can be used to provide elasticity of your Hadoop cluster on demand.
  - User self service available with Hadoop using tools such as vCloud Automation Center integrated with BDE
VMware vSphere BDE and Hadoop Resources

- VMware vSphere BDE web site
  - http://www.vmware.com/bde

- Virtualized Hadoop Performance with VMware vSphere 6 on High-Performance Servers
  - http://www.vmware.com/resources/techresources/10452

- Virtualized Hadoop Performance with VMware vSphere 5.1
  - http://www.vmware.com/resources/techresources/10360

- Benchmarking Case Study of Virtualized Hadoop Performance on vSphere 5

- Hadoop Virtualization Extensions (HVE):

- Apache Hadoop High Availability Solution on VMware vSphere 5.1

- Container Orchestration on vSphere with Big Data Extensions (Mesos and Kubernetes)