

Deploy Big Data Extensions on vSphere Standard Edition

You can deploy Big Data Extensions 2.1.1 Fling on VMware® vSphere™ Standard Edition for the purpose of experimentation and proof-of-concept projects to explore the potential of running Hadoop, Mesos, and Kubernetes in a virtualized environment.

Intended Audience

This document is for vSphere system administrators and developers who want to deploy and use Big Data Extensions 2.1.1 Fling on vSphere Standard Edition. To successfully work with Big Data Extensions, you should be familiar with Hadoop and VMware vSphere.

Deploying Big Data Extensions on vSphere Standard Edition

To deploy Big Data Extensions on vSphere Standard Edition, you must deploy the CentOS 6 Node Template virtual machine or Debian 7 Node Template virtual machine and Big Data Extensions Management Server virtual machine separately within a vSphere datacenter object (In contrast, when deploying on vSphere Enterprise or Enterprise Plus, you deploy a single OVA file which, when deployment completes, creates both the Node Template virtual machine and Big Data Extensions Management Server virtual machine.). These instructions assume you are familiar with deploying virtual machines in vSphere using the vSphere Web Client. For information on deploying virtual machines, see the vSphere documentation.

Prerequisites

- Configure a vSphere Datacenter for use with Big Data Extensions. For information on setting up a datacenter, see “Getting Started with Big Data Extensions” in the *VMware vSphere Big Data Extensions Administrator’s and User’s Guide*.

NOTE vSphere Standard Edition does not include Distributed Resources Scheduler (DRS). You can ignore the steps to enable DRS when deploying Big Data Extensions on vSphere Standard Edition.

- Download the Big Data Extensions management server and node templates virtual machines for vSphere Standard Edition from <http://labs.vmware.com/flings/big-data-extensions-for-vsphere-standard-edition>.
- Verify that you have at least 40 GB of free storage space for the virtual machines. You will need additional storage resources for the Hadoop cluster.

Procedure

- 1 Using the vSphere Web Client, create a folder in which to store the Node Template virtual machine and Management Server virtual machine under a valid vCenter Datacenter object.
- 2 Right-click the datacenter under which you created the folder, and select **Deploy OVF Template**. Following the instructions in the **Deploy OVF Template** wizard, use the template OVA file to deploy the Node Template virtual machine to the folder you created in the previous step. If you want to deploy Hadoop, HBase or Mesos cluster, use the CentOS 6 node template ova file. If you want to deploy Kubernetes cluster, use the Debian 7 node template ova file.

- 3 Right-click the same datacenter object as in the previous step (step 2), and select **Deploy OVF Template**. Use the management server OVA file to deploy the Management Server virtual machine within the same folder in which you deployed the Node Template virtual machine.

When the wizard prompts you to type the name of the Node Template virtual machine, use the same name you used to label the Node Template virtual machine in the previous step (step 2). The Management Server uses this Node Template virtual machine to clone new virtual machines as the nodes in the cluster.

- 4 For information on configuring storage and networks for use with Big Data Extensions, see “Getting Started with Big Data Extensions” in the *VMware vSphere Big Data Extensions Administrator’s and User’s Guide*.

NOTE vSphere Standard Edition does not support the use of resource pools. You can ignore the steps in the Admin Guide to choose a resource pool in which to deploy the Big Data Extensions vApp.

When deployment completes, two virtual machines are in the datacenter:

- The Management Server virtual machine, which is powered on as part of the Big Data Extensions deployment.
- The Node Template virtual machine, which is powered off. Big Data Extensions clones virtual machines from this template when provisioning a cluster.

What to do next

- If you are using vSphere 5.1 or later, you can install and use the Big Data Extensions graphical user interface to perform management and monitoring tasks. See “Install the Big Data Extensions Plug-in” in the *VMware vSphere Big Data Extensions Administrator’s and User’s Guide*.
- If you are using vSphere 5.0, you must use the Serengeti Command-Line Interface Client to perform additional configuration tasks. See the *VMware vSphere Big Data Extensions Command-Line Interface Guide*.
- If you did not leave the **Initialize Resources** check box selected (checked), you must add a datastore and network for use by the cluster you intend to create. You also need to add local and shared datastores into Big Data Extensions Management Server before creating a cluster. See “Managing vSphere Resources for Hadoop and HBase Clusters,” in the *VMware vSphere Big Data Extensions Administrator’s and User’s Guide*, or the *VMware vSphere Big Data Extensions Command-Line Interface Guide*.

- Create a cluster with your preferred software distribution.
 - If you want to create a Hadoop or HBase cluster, see “Creating Hadoop and HBase Clusters” in the *VMware vSphere Big Data Extensions Administrator’s and User’s Guide*, or the *VMware vSphere Big Data Extensions Command-Line Interface Guide*.
 - If you want to create a Kubernetes cluster, see “Creating a Kubernetes cluster”.
 - If you want to create a Mesos cluster, see “Creating a Mesos cluster”.

Creating a Kubernetes cluster

Creating a Kubernetes cluster using BDE CLI

1. Login BDE Server as user serengeti.
2. Launch BDE CLI and connect to BDE Server.

```
$ serengeti
serengeti> connect --host localhost:8443
Enter the username: <vc_username>
Enter the password: <vc_password>
Connected
```

3. Open sample cluster spec file for Kubernetes cluster `/opt/serengeti/samples/kubernetes_cluster.json`, and add the vCenter Server connection info in the ‘configuration’ section of this file.

```
"configuration": {
  "kubernetes": {
    "env": {
      "GOPATH": "/opt/kubernetes/src/go",
      "GOVC_INSECURE": "1",
      "GOVC_URL": "https://<user:password>@<vCenter_Server_IP>/sdk",
      "GOVC_DATACENTER": "<Datacenter_Name>",
      "GOVC_RESOURCE_POOL": "/${GOVC_DATACENTER}/host/<vCenter_Cluster_Name>/Resources",
      "INSTANCE_RESOURCE_POOL": "/${GOVC_DATACENTER}/vm/<Folder_of_BDE_Server_VM>",
      // The vlan name for the VMs of the cluster created by Serengeti
      "GOVC_NETWORK": "<VLAN_Name>",
      // The password for the VMs to be created
      "GOVC_GUEST_LOGIN": "serengeti.yourpassword",
      "KUBE_USER": "serengeti"
    }
  }
}
```

4. Create the Kubernetes cluster.

```
serengeti> cluster create --name <unique_cluster_name> --distro kubernetes --specFile
/opt/serengeti/samples/kubernetes_cluster.json --password --yes
Enter the password: <yourpassword>
Confirm the password: <yourpassword>
```

Creating a Kubernetes cluster using BDE GUI (i.e. vSphere Web Client)

We will add a link here to the video demo.

Creating a Mesos cluster

Creating a Mesos cluster using BDE CLI

1. Login BDE Server as user serengeti.
2. (Optional) Before creating a Mesos cluster, you can add some big packages required by Mesos cluster into BDE Management Server, then the VMs don't need to download the packages from Internet. This can speed up the cluster creation. Please execute these commands on BDE Server only once:

```
cd /opt/serengeti/www/yum/repos/mesos/0.21.0/RPMS/
wget http://mirror.centos.org/centos/6/updates/x86_64/Packages/java-1.6.0-openjdk-1.6.0.33-
1.13.5.1.el6_6.x86_64.rpm
createrepo ..
mkdir -p /opt/serengeti/www/files
cd /opt/serengeti/www/files
wget http://downloads.mesosphere.io/master/centos/6/mesos-0.20.1-py2.6.egg
chmod +w /opt/serengeti/chef/cookbooks/mesos/attributes/default.rb
echo "default['mesos']['python_egg'] = 'https://$(hostname)/files/mesos-0.20.1-py2.6.egg'" >>
/opt/serengeti/chef/cookbooks/mesos/attributes/default.rb
knife cookbook upload mesos -V
cd ~
```

3. Launch BDE CLI and connect to BDE Server.

```
$ serengeti
serengeti> connect --host localhost:8443
Enter the username: <vc_username>
Enter the password: <vc_password>
Connected
```

4. Create the Mesos cluster.

```
serengeti> cluster create --name <unique_cluster_name> --distro mesos --specFile
/opt/serengeti/samples/mesos_full_cluster.json --password
```

Enter the password: <yourpassword>
Confirm the password: <yourpassword>

Creating a Mesos cluster using BDE GUI (i.e. vSphere Web Client)

We will add a link here to the video demo.