ORACLE®

Oracle® Public Cloud Machine

Using Oracle Big Data Cloud Machine

Release 17.1.2

E85986-01

April 2017

Documentation that describes how to use Oracle Big Data Cloud Machine to store and manage large amounts of data of various data types in the cloud, by using Hadoop HDFS file system and associated services and tools.

ORACLE®

Oracle Public Cloud Machine Using Oracle Big Data Cloud Machine, Release 17.1.2

E85986-01

Contents

3   Accessing Your Oracle Big Data Cloud Machine

4   Managing Oracle Big Data Cloud Machine

A   bda-oss-admin Command Reference

B   Pages for Administering Oracle Big Data Cloud Machine

Preface

Using Oracle Big Data Cloud Machine describes how to use Oracle Big Data Cloud Machine to store and manage large amounts of data of various data types in the cloud, by using Hadoop HDFS file system and associated services and tools.

Topics

- Audience

- Documentation Accessibility

- Related Resources

- Conventions

Audience

Using Oracle Big Data Cloud Machine is intended for administrators and users who want to provision a Hadoop cluster in the cloud and use it to manage big data.

Documentation Accessibility

For information about Oracle's commitment to accessibility, visit the Oracle Accessibility Program website at `http://www.oracle.com/pls/topic/lookup?ctx=acc&id=docacc`.

Access to Oracle Support

Oracle customers that have purchased support have access to electronic support through My Oracle Support. For information, visit `http://www.oracle.com/pls/topic/lookup?ctx=acc&id=info` or visit `http://www.oracle.com/pls/topic/lookup?ctx=acc&id=trs` if you are hearing impaired.

Related Documents

For more information, see these Oracle resources:

- About Oracle Cloud in *Getting Started with Oracle Cloud*.

- Getting Started with Oracle Exadata Cloud Machine in *Using Oracle Database Exadata Cloud Machine*.

Conventions

The following text conventions are used in this document:

Convention	Meaning
boldface	Boldface type indicates graphical user interface elements associated with an action, or terms defined in text or the glossary.
italic	Italic type indicates book titles, emphasis, or placeholder variables for which you supply particular values.
`monospace`	Monospace type indicates commands within a paragraph, URLs, code in examples, text that appears on the screen, or text that you enter.

1

Getting Started with Oracle Big Data Cloud Machine

This section describes how to get started with Oracle Big Data Cloud Machine.

Topics

- About Oracle Big Data Cloud Machine

- Before You Begin with Oracle Big Data Cloud Machine

- How to Begin with Oracle Big Data Cloud Machine Subscriptions

- About Oracle Big Data Cloud Machine Users and Roles

- Typical Workflow for Using Oracle Big Data Cloud Machine

About Oracle Big Data Cloud Machine

A subscription to Oracle Big Data Cloud Machine gives you access to the resources of a pre-configured Oracle Big Data environment, including a complete installation of the Cloudera Distribution Including Apache Hadoop (CDH) and Apache Spark. Use Oracle Big Data Cloud Machine to capture and analyze the massive volumes of data generated by social media feeds, e-mail, web logs, photographs, smart meters, sensors, and similar devices.

When you order a subscription to Oracle Big Data Cloud Machine, you can begin with a starter pack of three nodes for each cluster and then proceed to add as many nodes as desired per cluster (up to 60 total nodes per cluster). Oracle manages the whole hardware and networking infrastructure as well as the initial setup, while you have complete administrator's control of the software.

All nodes in an Oracle Big Data Cloud Machine instance form a cluster. Each Oracle Big Data Cloud Machine rack can host several clusters (instances), and large clusters can also span multiple Oracle Big Data Cloud Machine racks.

Software

A subscription to Oracle Big Data Cloud Machine includes the following:

- Oracle Linux operating system

- Cloudera Distribution Including Apache Hadoop (CDH)

  CDH has a batch processing infrastructure that can store files and distribute work across a set of computers. Data is processed on the same computer where it is stored. In a single Oracle Big Data Cloud Machine cluster, CDH distributes the files and workload across a number of servers, which compose a cluster. Each server is a node in the cluster.

CDH includes:

–   File system: The Hadoop Distributed File System (HDFS) is a highly scalable file system that stores large files across multiple servers. It achieves reliability by replicating data across multiple servers without RAID technology. It runs on top of the Linux file system.

–   MapReduce engine: The MapReduce engine provides a platform for the massively parallel execution of algorithms written in Java. Oracle Big Data Cloud Machine runs YARN by default.

–   Administrative framework: Cloudera Manager is a comprehensive administrative tool for CDH.

–   Apache projects: CDH includes Apache projects for MapReduce and HDFS, such as Hive, Pig, Oozie, ZooKeeper, HBase, Sqoop, and Spark.

–   Cloudera applications: Oracle Big Data Cloud Machine includes all products included in Cloudera Enterprise Data Hub Edition, including Impala, Search, and Navigator.

Several CDH utilities and other software available on Oracle Big Data Cloud Machine provide graphical, web-based, and other language interfaces for ease of use.

•   Oracle Big Data Connectors, which facilitate access to data stored in an Apache Hadoop cluster.

Included are:

–   Oracle SQL Connector for Hadoop Distributed File System

–   Oracle Loader for Hadoop

–   Oracle XQuery for Hadoop

–   Oracle R Advanced Analytics for Hadoop

–   Oracle Data Integrator Enterprise Edition

•   Oracle Big Data Spatial and Graph, which provides advanced spatial and graph analytic capabilities to supported Apache Hadoop and NoSQL Database Big Data platforms.

Before You Begin with Oracle Big Data Cloud Machine

Before you begin using Oracle Big Data Cloud Machine, you should be familiar with the following:

•   Oracle Cloud

See About Oracle Cloud for information about Oracle Cloud, especially Buying a Nonmetered Subscription to an Oracle Cloud Service.

> **Note:**
>
> Trial subscriptions are not available for Oracle Big Data Cloud Machine, and you can't purchase a subscription through the web site. You must purchase a subscription from an Oracle Sales Representative.
>
> When you subscribe to Oracle Big Data Cloud Machine, you can select a starter pack of three nodes plus up to 57 additional nodes, in one-node increments. Later, after you've created the service instance, you can add additional nodes to your subscription, up to 60 total.

- Cloudera Distribution Including Apache Hadoop (CDH)

    CDH provides the software infrastructure for working with big data. See the Cloudera documentation for the Cloudera release supported by this release of Oracle Big Data Cloud Machine.

Before you create your first Oracle Big Data Cloud Machine cluster:

- Get an Oracle Big Data Cloud Machine subscription. See How to Begin with Oracle Big Data Cloud Machine Subscriptions. Contact an Oracle Sales Representative for options and pricing.

- Create a Secure Shell (SSH) public/private key pair so you can provide the public key when you create the cluster. See Generating a Secure Shell (SSH) Public/Private Key Pair.

How to Begin with Oracle Big Data Cloud Machine Subscriptions

Here's how to get started with Oracle Big Data Cloud Machine subscriptions.

1. You must contact an Oracle sales representative to purchase a subscription to Oracle Big Data Cloud Machine. Trial subscriptions aren't available for Oracle Big Data Cloud Machine.

    See Buying a Nonmetered Subscription to an Oracle Cloud Service in *Getting Started with Oracle Cloud*.

2. Set up your Oracle Database Public Cloud Services account.

    See Setting Up an Oracle Cloud Services Account in *Getting Started with Oracle Cloud*.

3. Learn about Oracle Cloud users and roles. See Oracle Cloud User Roles and Privileges in *Getting Started with Oracle Cloud*.

4. Verify Oracle Big Data Cloud Machine is ready to use.

    See Verifying That a Service Is Running in *Getting Started with Oracle Cloud*.

5. Create accounts for your users and assign them appropriate privileges and roles.

    See Managing User Accounts and Managing User Roles in *Managing and Monitoring Oracle Cloud*.

6. Be sure to review the prerequisites described in Before You Begin with Oracle Big Data Cloud Service before you create your first Oracle Big Data Cloud Machine cluster.

Generating a Secure Shell (SSH) Public/Private Key Pair

Several tools exist to generate SSH public/private key pairs. The following sections show how to generate an SSH key pair on UNIX, UNIX-like and Windows platforms.

Generating an SSH Key Pair on UNIX and UNIX-Like Platforms Using the ssh-keygen Utility

UNIX and UNIX-like platforms (including Solaris and Linux) include the ssh-keygen utility to generate SSH key pairs.

To generate an SSH key pair on UNIX and UNIX-like platforms using the ssh-keygen utility:

1. Navigate to your home directory:

    ```
    $ cd $HOME
    ```

2. Run the ssh-keygen utility, providing as *filename* your choice of file name for the private key:

    ```
    $ ssh-keygen -b 2048 -t rsa -f filename
    ```

    The ssh-keygen utility prompts you for a passphrase for the private key.

3. Enter a passphrase for the private key, or press **Enter** to create a private key without a passphrase:

    ```
    Enter passphrase (empty for no passphrase): passphrase
    ```

    Note:

    While a passphrase is not required, you should specify one as a security measure to protect the private key from unauthorized use. When you specify a passphrase, a user must enter the passphrase every time the private key is used.

    The ssh-keygen utility prompts you to enter the passphrase again.

4. Enter the passphrase again, or press **Enter** again to continue creating a private key without a passphrase:

    ```
    Enter the same passphrase again: passphrase
    ```

5. The ssh-keygen utility displays a message indicating that the private key has been saved as *filename* and the public key has been saved as *filename*.pub. It also displays information about the key fingerprint and randomart image.

Generating an SSH Key Pair on Windows Using the PuTTYgen Program

The PuTTYgen program is part of PuTTY, an open source networking client for the Windows platform.

To generate an SSH key pair on Windows using the PuTTYgen program:

1. Download and install PuTTY or PuTTYgen.

To download PuTTY or PuTTYgen, go to http://www.putty.org/ and click the **You can download PuTTY here** link.

2. Run the PuTTYgen program.

   The PuTTY Key Generator window is displayed.

3. Set the **Type of key to generate** option to **SSH-2 RSA**.

4. In the **Number of bits in a generated key** box, enter **2048**.

5. Click Generate to generate a public/private key pair.

   As the key is being generated, move the mouse around the blank area as directed.

6. (Optional) Enter a passphrase for the private key in the **Key passphrase** box and reenter it in the **Confirm passphrase** box.

   > **Note:**
   >
   > While a passphrase is not required, you should specify one as a security measure to protect the private key from unauthorized use. When you specify a passphrase, a user must enter the passphrase every time the private key is used.

7. Click **Save private key** to save the private key to a file. To adhere to file-naming conventions, you should give the private key file an extension of `.ppk` (PuTTY private key).

8. Select all of the characters in the **Public key for pasting into OpenSSH authorized_keys file** box.

   Make sure you select all the characters, not just the ones you can see in the narrow window. If a scroll bar is next to the characters, you aren't seeing all the characters.

9. Right click somewhere in the selected text and select **Copy** from the menu.

10. Open a text editor and paste the characters, just as you copied them. Start at the first character in the text editor, and do not insert any line breaks.

11. Save the text file in the same folder where you saved the private key, using the `.pub` extension to indicate that the file contains a public key.

12. If you or others are going to use an SSH client that requires the OpenSSH format for private keys (such as the `ssh` utility on Linux), export the private key:

    a. On the **Conversions** menu, choose **Export OpenSSH key** .

    b. Save the private key in OpenSSH format in the same folder where you saved the private key in `.ppk` format, using an extension such as `.openssh` to indicate the file's content.

About Oracle Big Data Cloud Machine Users and Roles

Oracle Big Data Cloud Machine supports the following service roles and operating system roles.

Cloud Service Users and Roles

In addition to the roles and privileges described in Oracle Cloud User Roles and Privileges in *Getting Started with Oracle Cloud*, the following roles are created for Oracle Big Data Cloud Machine:

- Big Data Administrator

  A user assigned this role has complete administrative control over the service.

- Viewer

  A user assigned this role has read-only access to the service.

When the Oracle Big Data Cloud Machine account is first set up, the service administrator is given the Big Data Administrator role. User accounts must be added and assigned one of the above roles before anyone else can access and use Oracle Big Data Cloud Machine.

Only the identity domain administrator is allowed to create user accounts and assign roles. See Managing User Accounts in *Managing and Monitoring Oracle Cloud*.

The predefined Oracle Big Data Cloud Machine roles are associated with specific clusters. That is, if you have two clusters called **test123** and **production123**, four predefined roles are available to assign to users:

- test123 Big Data Administrator

- test123 Viewer

- production123 Big Data Administrator

- production123 Viewer

Users have access only to those clusters associated with the roles assigned them. For example, in the above case, a user might be assigned the role Big Data Administrator for test123 and the role Viewer for production123.

Operating System Users and Roles

Every Oracle Big Data Cloud Machine cluster node is provisioned with the following operating system user accounts.

- **opc**

  The system administrator account you use in conjunction with the `sudo` command to gain `root` user access to your nodes. By default, `opc` doesn't allow connection using a password; however, you may choose to connect using a password by assigning a known password to `opc` or by creating another user with a known password. See Managing User Accounts in *Oracle Big Data Appliance Software User's Guide*.

- **root**

  The root administrator for the system. You do not have direct access to this account. To perform operations that require root user access, execute `sudo -s` as the `opc` user. By default, `root` doesn't require a password.

- **oracle**

An operating system and cluster user account that is used to run jobs on the cluster during the validation of the cluster. This account is used by the system and has a randomly generated password.

Enabling IPsec VPN Access to Oracle Big Data Cloud Machine

Oracle Network Cloud Service — VPN for Engineered Systems is an add-on service available at an additional subscription fee. Using this service, you can create a secure virtual private network (VPN) tunnel over the Internet that connects your corporate network to Oracle Public Cloud services, such as Oracle Big Data Cloud Service. The service uses IPsec, which is a suite of protocols designed to authenticate and encrypt all IP traffic between two locations.

> **Note:** For information on IPSec standards, see the Internet Engineering Task Force (IETF) Request for Comments (RFC) 6071: *IP Security (IPsec) and Internet Key Exchange (IKE) Document Roadmap*.

Before you request VPN, ensure these requirements are met at your site:

- **VPN device requirements.** You need a VPN gateway device that uses current IPSec standards to establish a secure tunnel between your network and the Oracle Public Cloud. You will provide the details of your device to Oracle. The device must support:

    - IPv4 traffic with support for ICMP, TCP and UDP. Multicast traffic is not supported.

    - Tunnel mode sessions: Tunnel mode is used to create a virtual private network between your network and the Oracle Public Cloud, rather than between a specific set of hosts. It is used to protect all communications between both networks.

    - Authentication with pre-shared keys. The same pre-shared key is configured on each IPSec VPN gateway device.

    - Dynamic rekeying: IPsec uses dynamic rekeying to control how often a new key is generated during communication. Communication is sent in blocks and each block of data is secured with a different key.

- **Network requirements for an IPSec VPN connection.** Both sides must provide subnets:

    - On your side, dedicate subnets in your network for this VPN connection. You will indicate these subnets to Oracle. You will give the necessary information about these subnets to Oracle. To prevent an IP address conflict in the end-to-end network connection, mask your internal systems with a public or non-RFC 1918 address range.

    - On the Oracle side, the network engineers from the Oracle Cloud Operations will provide the destination subnets in a way that avoids IP address conflicts.

To request a VPN provisioning by Oracle Support:

1. Contact your sales representative and ask them to place an order for Oracle Network Cloud Service — VPN for Engineered Systems — Non-metered. This can

be a separate order, or it can be made in conjunction with an order for Oracle Big Data Cloud Service.

2. Once you have an active subscription to Oracle Network Cloud Service — VPN for Engineered Systems, go to the My Oracle Support Note 2056914.1 and follow its instructions.

Oracle engineers will receive your information and check that all prerequisites are met. Next, during an agreed maintenance window, Oracle together with your network engineers will provision the VPN service and run through a post-configuration checklist to ensure that the VPN connection is working and that the setup is completed.

Where Do Services Run On a Cluster?

Services are installed on all nodes of an Oracle Big Data Cloud Machine cluster, but individual services run only on designated nodes. An Oracle Big Data Cloud Machine cluster can have from three to 60 nodes, and the services are distributed differently, depending on the number of nodes:

- Services on a three-node cluster are distributed as shown in Where Do the Services Run on a Three-Node, Development-Only Cluster?.

- Services on a cluster of four or more nodes are distributed as shown in Where Do the Services Run on a Cluster of Four or More Nodes?.

Consider the following:

- If you have a three-node cluster and then expand it to four or more nodes, the services are automatically redistributed, as shown in Where Do the Services Run on a Cluster of Four or More Nodes?.

- Three- and four-node clusters are recommended for development purposes only.

- The minimum recommended size for a production cluster is five nodes.

- If you plan to install Oracle Big Data Discovery on Oracle Big Data Cloud Machine, a minimum of five nodes is required.

Where Do the Services Run on a Three-Node, Development-Only Cluster?

The following table shows how services are distributed on the nodes of Oracle Big Data Cloud Machine three-node clusters.

Service	Node1	Node2	Node3
Active Navigator Key Trustee Server (if secure cluster is enabled)	Yes	No	No
Cloudera Manager and Cloudera Manager roles	No	No	Yes
DataNode	Yes	Yes	Yes
Failover Controller	Yes	Yes	No

Service	Node1	Node2	Node3
Hive	No	**Yes**	No
Hive Metastore	No	**Yes**	No
HttpFS	No	**Yes**	No
Hue	No	**Yes**	No
JobHistory	No	No	**Yes**
JournalNode	**Yes**	**Yes**	**Yes**
Kerberos Master KDC (if secure cluster is enabled.)	**Yes**	No	No
Kerberos Slave KDC (if secure cluster is enabled)	No	**Yes**	No
MySQL Backup	No	**Yes**	No
MySQL Primary	No	No	**Yes**
NameNode	**Yes**	**Yes**	No
NodeManager	**Yes**	**Yes**	**Yes**
ODI	No	**Yes**	No
Oozie	No	**Yes**	No
Passive Navigator Key Trustee Server (if secure cluster is enabled)	No	**Yes**	No
ResourceManager	**Yes**	No	**Yes**
Sentry	**Yes**	No	No
Spark History	No	No	**Yes**
WebHCat	No	**Yes**	No
ZooKeeper	**Yes**	**Yes**	**Yes**

Where Do the Services Run on a Cluster of Four or More Nodes?

The following table shows how services are distributed on the nodes of Oracle Big Data Cloud Machine clusters with four, five, or more nodes.

Note: The minimum recommended size for a production cluster is five nodes.

Service	Node1	Node2	Node3	Node4	Node5 to Node*nn*
Active Navigator Key Trustee Server (if secure cluster is enabled)	**Yes**	No	No	No	No
Balancer	**Yes**	No	No	No	No
Cloudera Manager Agent	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**
Cloudera Manager Server	No	No	**Yes**	No	No
Cloudera Manager and Cloudera Manager roles	No	No	**Yes**	No	No
DataNode	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**
Failover Controller	**Yes**	**Yes**	No	No	No
Hive	No	No	No	**Yes**	No
Hive Metastore	No	**Yes**	No	No	No
HttpFS	No	**Yes**	No	No	No
Hue	No	No	No	**Yes**	No
JobHistory	No	No	**Yes**	No	No
JournalNode	**Yes**	**Yes**	**Yes**	No	No
Kerberos KDC (if secure cluster is enabled.)	**Yes**	**Yes**	No	No	No
MySQL Backup	No	**Yes**	No	No	No
MySQL Primary	No	No	**Yes**	No	No
NameNode	**Yes**	**Yes**	No	No	No
Navigator Audit Server and Navigator Metadata Server	No	No	**Yes**	No	No

Service	Node1	Node2	Node3	Node4	Node5 to Node*nn*
NodeManager	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**
Oozie	No	No	No	**Yes**	No
Oracle Data Integrator Agent	No	No	No	**Yes**	No
Passive Navigator Key Trustee Server (if secure cluster is enabled)	No	**Yes**	No	No	No
ResourceManager	No	No	**Yes**	**Yes**	No
Sentry Server (if enabled)	**Yes**	No	No	No	No
Solr	No	No	No	**Yes**	No
Spark History	No	No	**Yes**	No	No
ZooKeeper	**Yes**	**Yes**	**Yes**	No	No

Note: If Oracle Big Data Discovery is installed, the NodeManager and DataNode on Node05 of the cluster are decomissioned.

Typical Workflow for Using Oracle Big Data Cloud Machine

To start using Oracle Big Data Cloud Machine, refer to the following tasks as a guide:

Task	Description	More Information
Purchase a subscription to Oracle Big Data Cloud Machine.	Purchase a subscription to Oracle Big Data Cloud Machine.	How to Begin with Oracle Big Data Cloud Machine Subscriptions
Add and manage users and roles	Create accounts for your users and assign them appropriate privileges. Assign the necessary Oracle Big Data Cloud Machine roles.	About Oracle Big Data Cloud Machine Users and Roles
Create a service instance	Use the wizard to create a service instance, which allocates resources for a cluster.	Creating a Oracle Big Data Cloud Machine Instance

Task	Description	More Information
Create an SSH key pair	Create an SSH public/private key pair for use when creating clusters.	Generating a Secure Shell (SSH) Public/Private Key Pair
Create a cluster	Use a wizard to create a cluster, using the resources allocated to the service instance..	Creating a Cluster
Control network access to services	Use a *whitelist* to open or close the ports used by services such as Cloudera Manager and Hue.	Controlling Network Access to Services
Access and work with your cluster.	Access and work with your cluster using Secure Shell and by using graphical tools such as Cloudera Manager and Hue.	Accessing Your Oracle Big Data Cloud Machine
Add nodes to a cluster	Add nodes in one-node increments up to a total of 60 nodes in the cluster.	Adding Permanent Nodes To a Cluster
Patch a cluster	Apply a patch or roll back a patch.	Installing a One-Off Patch

2

Managing the Life Cycle of Oracle Big Data Cloud Machine

This section describes tasks to manage the life cycle of your service.

Topics

- Understanding the Workflow and Life Cycle of Service Instances and Clusters
- About Oracle Big Data Cloud Machine Nodes
- Creating a Oracle Big Data Cloud Machine Instance
- Creating a Cluster
- Adding Nodes to a Cluster
- Controlling Network Access to Services
- Viewing All Clusters
- Viewing Details About a Cluster
- Using HDFS Transparent Encryption

  Restarting a Cluster Node
- Restarting the Virtual Machines (VMs) Hosting a Cluster
- Updating the SSH Public Key for a Cluster
- Supporting Multiple Key Pairs for Secure Shell (SSH) Access
- Deleting a Cluster

Understanding the Workflow and Life Cycle of Service Instances and Clusters

For each Oracle Big Data Cloud Machine cluster you want to create, you must have a separate subscription.

The process is as follows:

1. **Purchase a subscription.** You have to contact an Oracle Sales Representative to buy an subscription; you can't buy one online. See How to Begin with Oracle Big Data Cloud Machine Subscriptions.

   Among the details you provide when buying a subscription is the number of permanent nodes you want for your cluster. Permanent nodes include Oracle

Processing Units (OCPUs), memory, and storage. A cluster must have at least a starter pack of three nodes and can have an additional 57 nodes for a total of 60.

After you've created a cluster, you can extend your subscription with additional nodes. If you want additional clusters, you have to purchase additional subscriptions. Each cluster requires a new starter pack of three nodes. Make sure all clusters needed are listed in the same order for an Oracle Big Data Cloud Machine.

2. **If this is a new Oracle Cloud account, provide account details,** such as account name and administrator details. An account can include subscriptions to multiple services. If you purchased the subscription as part of an existing account, you skip this step.

3. **Create a service instance.** This step allocates the resources for a cluster. It doesn't create the cluster itself. See Creating a Oracle Big Data Cloud Machine Instance.

4. **Create the cluster.** You can have only one cluster per service instance. See Creating a Cluster.

5. After the cluster is created, you can **extend** it by adding **permanent nodes** to your subscription, at additional cost. These nodes, which can be *permanent Hadoop nodes* or *edge nodes*, remain part of the cluster for its lifetime. See About Oracle Big Data Cloud Machine Nodes and Adding Permanent Nodes To a Cluster.

6. If needed, you can delete the cluster and create a new one, using the resources specified in your subscription and allocated to the service instance. In this case, you don't have to create a new service instance before creating the cluster. Because the resources are already allocated, you can create the cluster directly.

The above steps are similar to the steps for other Oracle Cloud services, but each service has its own variations. For generic instructions on subscribing to a service and instantiating it, see Getting Started with Oracle Cloud.

About Oracle Big Data Cloud Machine Nodes

All Oracle Big Data Cloud Machine clusters have *permanent Hadoop nodes* and may also have *edge nodes*.

Every cluster must have at least three permanent Hadoop nodes (a starter pack) and can have an additional 57 permanent nodes, which can be a combination of permanent Hadoop nodes and edge nodes.

When a subscription is ordered, only the total number of permanent nodes has to be specified. You allocate the nodes as permanent Hadoop nodes or as edge nodes when creating the cluster.

Permanent Hadoop Nodes

Permanent Hadoop nodes last for the lifetime of the cluster. Each node has:

*   32 Oracle Compute Units (OCPUs)

*   248 GB of available RAM

*   48 TB storage

*   Full use of the Cloudera Enterprise Data Hub Edition software stack, including licenses and support

When you subscribe, you're only required to get a starter pack of three nodes, although you can include up to 57 additional nodes (to total 60 nodes). If your cluster doesn't already have the maximum 60, you can also order and add additional permanent nodes after the cluster is up and running.

When planning the number of nodes you want for a cluster, be aware of the following:

- Three-node clusters are recommended for development only. A production cluster should have five or more nodes. This is to ensure that, if a node fails, you can migrate the node responsibilities to a different node and retain quorums in the high availability setup of the cluster.

- Services are distributed differently on three-node clusters than they are on clusters of four or more nodes. See Where Do the Services Run on a Three-Node, Development-Only Cluster?

- You must have at least four permanent Hadoop nodes before you can add edge nodes.

- Installing Oracle Big Data Discovery on a Oracle Big Data Cloud Machine cluster requires at least five nodes.

Edge Nodes

Edge nodes provide an interface between the Hadoop cluster and the outside network. They are commonly used to run client applications and cluster administration tools, keeping them separate from the nodes of the cluster that run Hadoop services. Like permanent Hadoop nodes, edge nodes last for the lifetime of the cluster. They have the same characteristics as permanent Hadoop nodes:

- 32 Oracle Compute Units (OCPUs)

- 248 GB of available RAM

- 48 TB storage

- Full use of the Cloudera Enterprise Data Hub Edition software stack, including licenses and support

When you create a cluster or expand a cluster, you can specify how many of the nodes will be edge nodes, as long as the first four nodes are permanent Hadoop nodes.

Creating a Oracle Big Data Cloud Machine Instance

When you create an Oracle Big Data Cloud Machine *instance*, you initiate a process that allocates resources for a cluster. You have to create a service instance before you can create a cluster.

Procedure

To create an Oracle Big Data Cloud Service instance:

1. Go to your My Services Dashboard by clicking the link in your Welcome e-mail or by logging in through `cloud.oracle.com` and navigating to My Services.
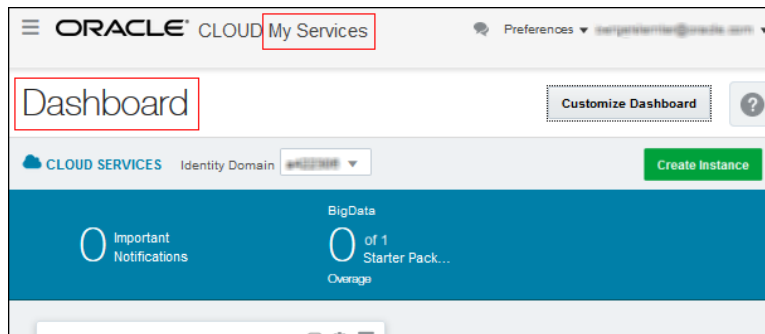
2. Click the **Create Instance** button and select **BigData**.
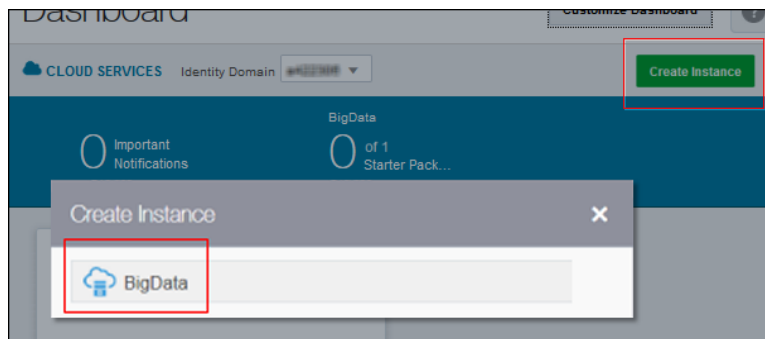
3. On the Instance Details page of the Create New Oracle Big Data Cloud Service Instance wizard, provide the details for the service instance and then click **Next**:

   Instance Details

   - **Name** – Enter a name for the service instance. You'll see this name later in the Create Cluster wizard, where it will appear on a list of instances available for creating a cluster.

   - **Plan** – There's currently only one plan for Oracle Big Data Cloud Service, and it's selected by default.

   - **Starter Pack** – Every service instance must include one (and only one) starter pack of three nodes.

   - **Additional Nodes**

     Enter the number of additional nodes you want. Make sure that you have enough cloud credits to pay for the additional nodes.

   Administrator Details

   - **E-mail** – The e-mail address of the administrator for this service instance.

   - **Use e-mail as user name** – Select this to use the administrator's e-mail address, above, as the user name for this service instance.

   - **User Name** — Enter a user name for the user, if you didn't select **Use e-mail as user name**.

   - **First/Last Name** — Enter the administrator's first and last name.
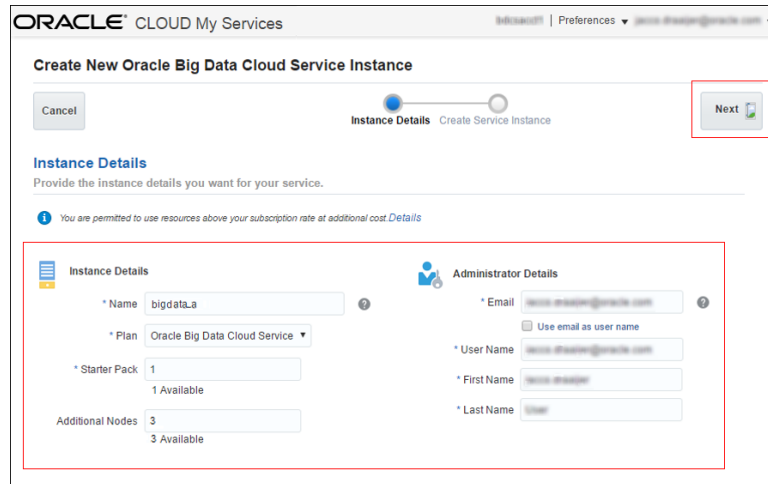
4. On the Confirmation page, if you're satisfied with the details, click **Create Service Instance** to initiate the allocation of resources.
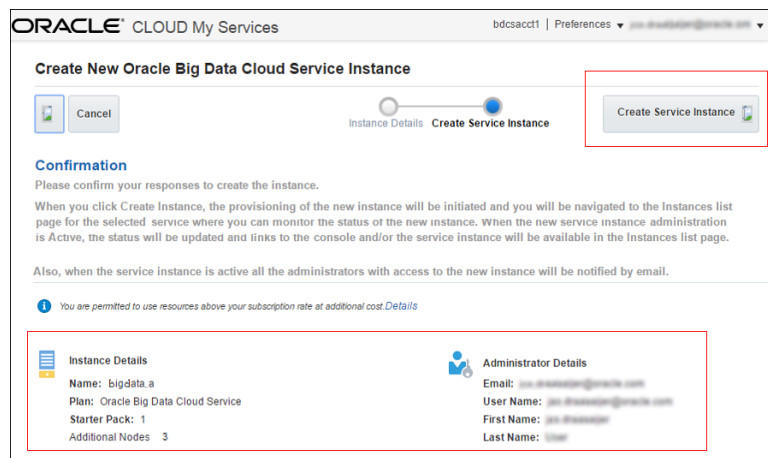
When the new service instance appears under **Service Instances** on the Service Details page, the instance's resources are available for creating a cluster.

> **Note:**
>
> The process takes a few minutes, so your new service instance won't appear right away. On the Overview tab, click the **Refresh** button occasionally until
>
> the service instance is shown:
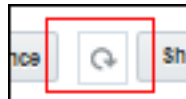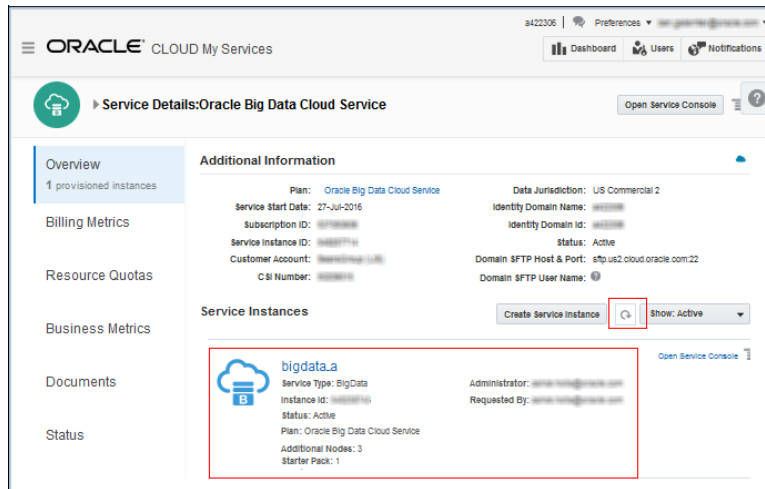
You'll also receive an "Action Required" e-mail announcing that the service instance is ready and you can create a cluster using those resources.
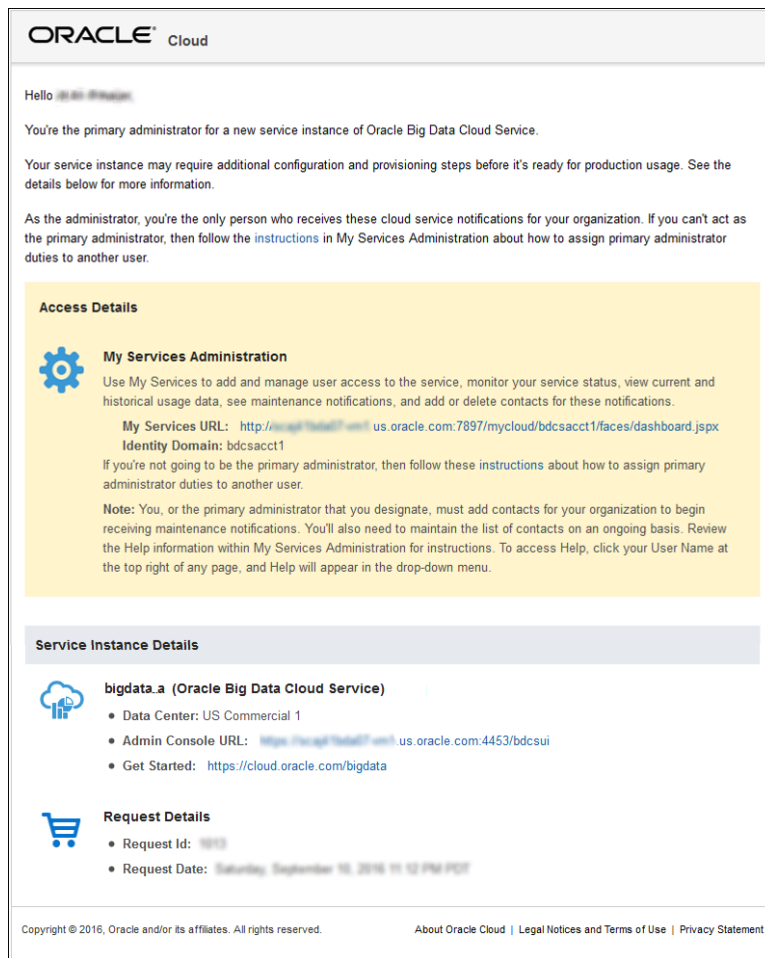
5. When the instance is ready, you can proceed to create a cluster. See Creating a Cluster.

Creating a Cluster

You can create a single cluster from the resources allocated in an Oracle Big Data Cloud Machine instance.

Before You Begin

Before you can create a cluster,

- You must have an Oracle Big Data Cloud Machine instance that hasn't yet been used for a cluster. See Understanding the Workflow and Life Cycle of Service Instances and Clusters and Creating a Oracle Big Data Cloud Machine Instance.

- You must have a secure shell (SSH) key pair. See Generating a Secure Shell (SSH) Public/Private Key Pair.

Procedure

To create an Oracle Big Data Cloud Machine cluster:

1. Go to the service console for your account, in either of the following ways:

    - In the "Action Required" e-mail you received when your service instance was ready, find the **Service Instance Details** section, and click the **Admin Console URL** link. If prompted, sign in.
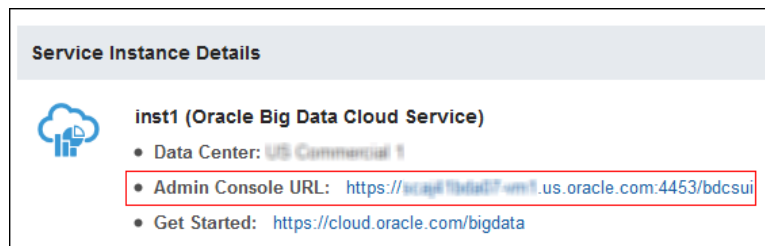
    - If you just created the service instance and are still on the Service Details page, click the **Open Service Console** link:
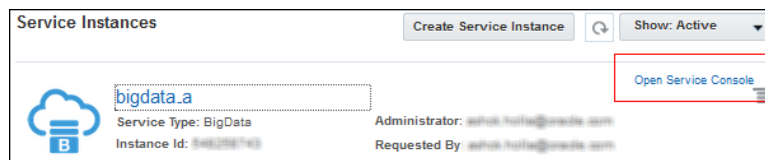
2. On the Clusters page of the service console, click the **Create Cluster** button.
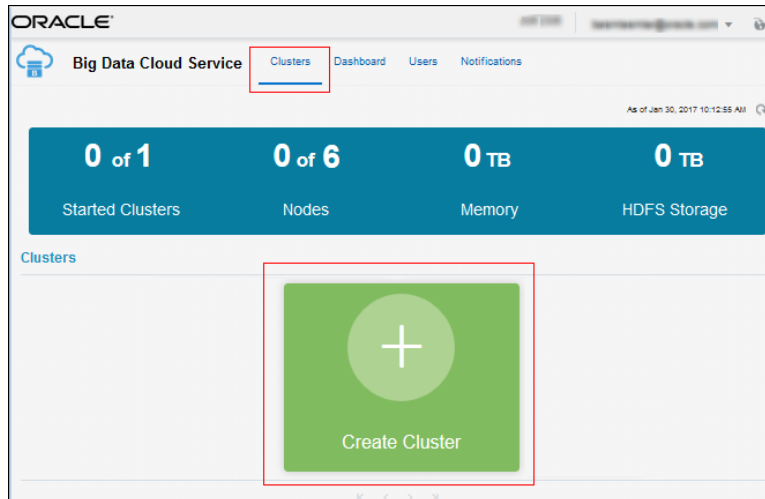
3. On the Cluster Details page of the Create Cluster wizard, configure details for your cluster, as described below, and then click **Next**:

    • **Name**—Enter a name for the cluster.

    • **Description**—Enter a description for the cluster.

    • **Big Data Appliance System**—Select the resource configuration to use for this cluster. The items on this list show the Oracle Big Data Cloud Machine instances that are available for creating clusters. Each item on the list shows the name of an instance and the number of nodes in it.
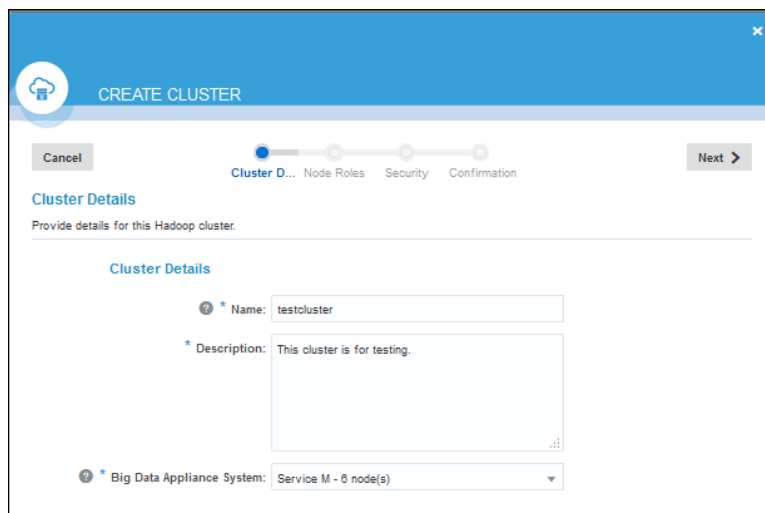
4. On the Node Roles page of the wizard, assign the roles of the available nodes, as described below, and then click **Next**.

    • **Permanent Hadoop Nodes** — Select the number of nodes to use as permanent Hadoop nodes. If you have only three or four nodes available, you must use them all as permanent Hadoop nodes. If you have any left after the first four are assigned, you can assign them to be either Hadoop permanent nodes or edge nodes.

- **Edge Nodes** — If you have any remaining nodes after assigning your permanent Hadoop nodes, that number appears here.

Important: You can't change the roles of nodes after the cluster is created.

---
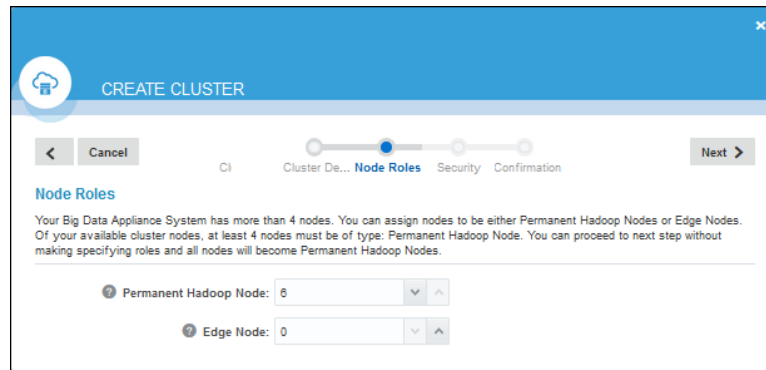
5. On the Security page, configure the security details for your cluster, as described below, and then click **Next**:

   - **SSH Key**—The SSH public key is used for authenticating the `opc` user when using an SSH client to connect to a node that is associated with this cluster.

     Use either of the following methods to specify the SSH public key:

     – Paste in the value of the SSH public key to be used by the `opc` user. Make sure the value doesn't contain extra spaces or line breaks and doesn't have a line break at the end.

     – Click **Select File** to select a file containing the new public key.

   - **Cloudera Administrator Password** —Enter a string for the password for accessing Cloudera Manager and other Cloudera tools. Don't use a dollar sign ($) in your password. It may be accepted, but it can cause problems later on. *Be sure to record the password and store it securely.*

   - **Secure Setup** — Select **Enabled** to enable the included security services. When you select this option, you enable:

     – MIT Kerberos

     – Apache Sentry

     – HDFS Encryption—Key Trustee Servers are installed on the master nodes

     – Network Encryption

     – Auditing

     When you create a cluster with security enabled, you can't disable any of these features for the lifetime of the cluster.
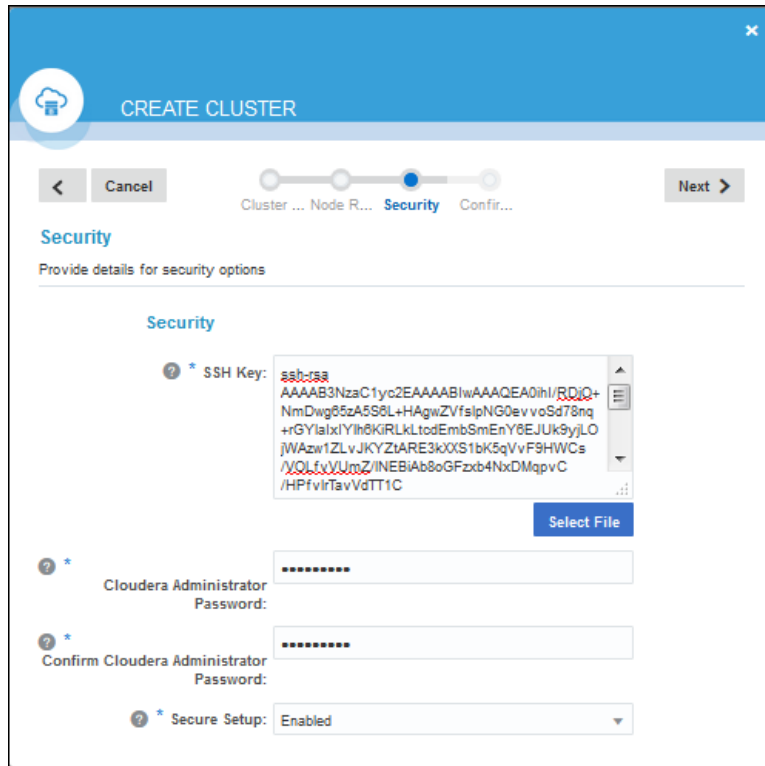
6. On the Confirmation page, review the information. If you're satisfied with the information, click **Provision**.
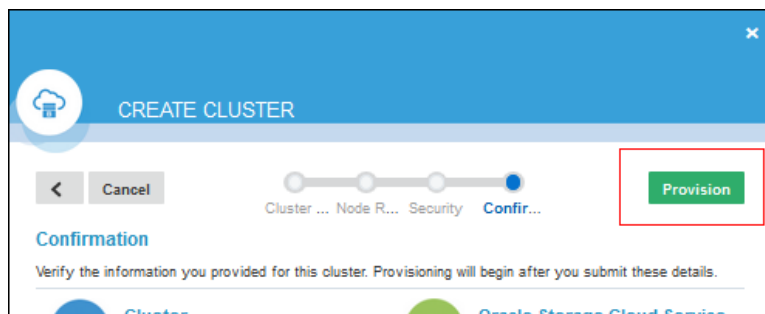
When the cluster is ready, the new cluster is shown on your service console's Clusters page, with the message **Ready**.
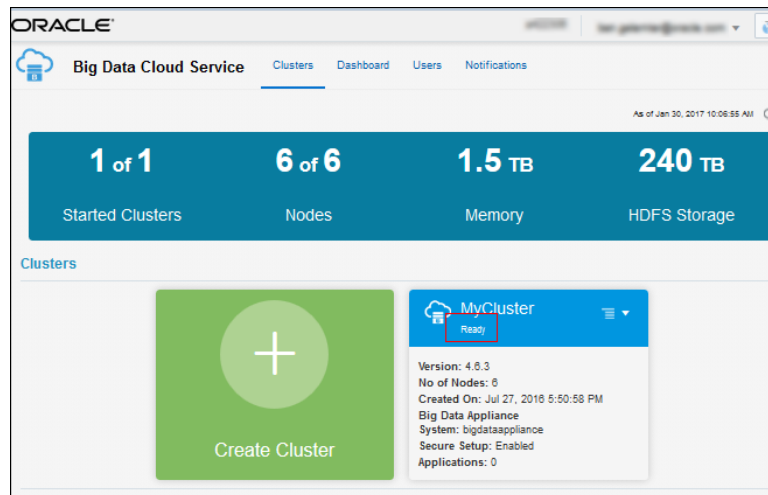
7. To see details about your new cluster, click the menu and select **View Details**
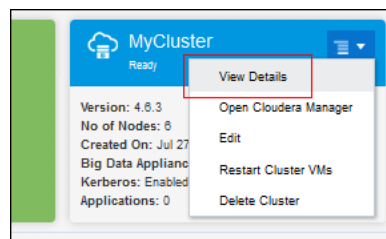
8. Review the details on the Cluster Details pages.
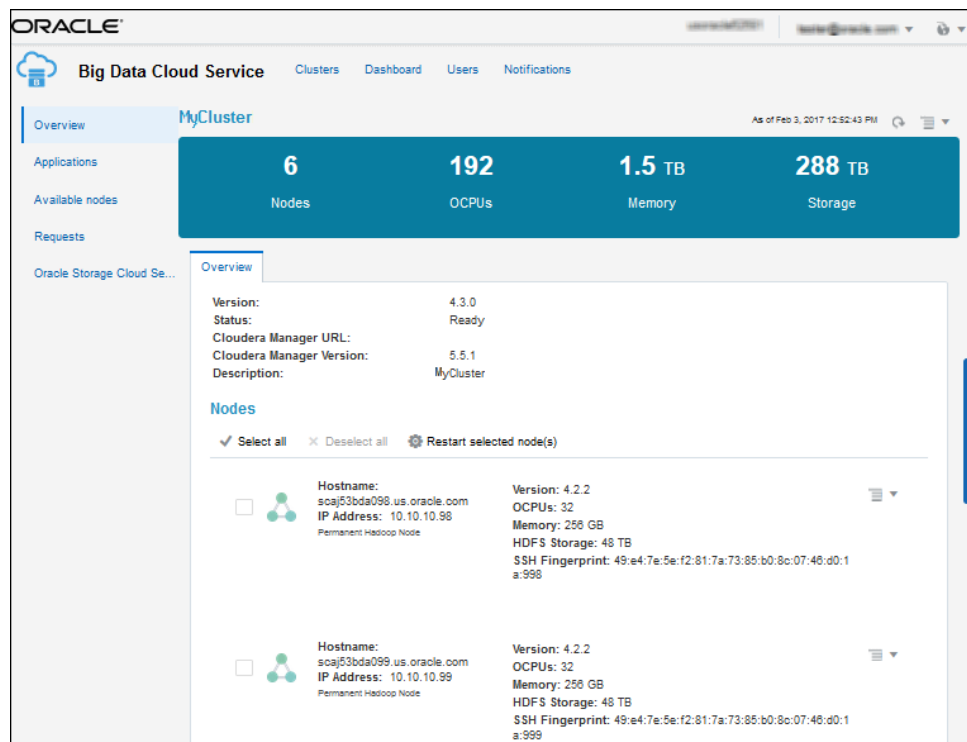
Adding Nodes to a Cluster

You can extend a cluster by adding *permanent Hadoop nodes*, *edge nodes*, and *cluster compute* nodes.

See About Oracle Big Data Cloud Machine Nodes for information about the different kinds of nodes.

Topics

- Adding Permanent Nodes To a Cluster
- Adding Nodes if the Regular Process Is Interrupted

Adding Permanent Nodes To a Cluster

You can add additional *permanent* nodes to a cluster after it was created and started. Permanent nodes include *permanent Hadoop nodes* and *edge nodes*.

To add permanent nodes to a cluster:

1. Contact an Oracle Sales Representative to extend your subscription by the number of permanent nodes you want to add. You don't have to say how many will be used as permanent Hadoop nodes and how many as edge nodes yet. You'll make that choice in the steps below.

   You'll receive an e-mail notification when the nodes have been added to your subscription.

2. Go to your Big Data Cloud Service console.

3. Find the cluster where you want to add the node(s), click ☰, and select **Extend**.
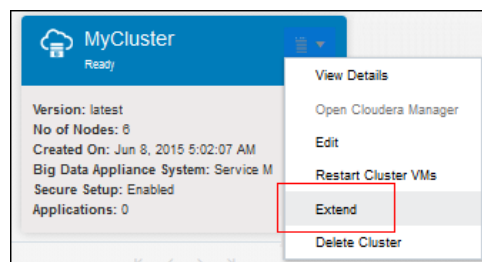
4. In the **Extend** *clustername* dialog box, select how many of the extra nodes will be permanent Hadoop nodes and how many will be edge nodes. Then enter your Cloudera administrator password, and click **Extend**.

   Four permanent Hadoop nodes are required before you can assign any additional nodes as edge nodes. For example, if your cluster has three permanent Hadoop nodes and you've purchased an entitlement to two more, you must allocate one as a permanent Hadoop node. You can then allocate the fifth node either as a permanent Hadoop node or an edge node. If your cluster already has four nodes and you purchase an entitlement to more nodes, you can allocate them as any combination of permanent Hadoop nodes and edge nodes.

When successful, the new status appears on the **Cluster Summary** page and on the **Cluster Details** page of the service console. If the process failed for any reason, see Adding Nodes if the Regular Process Is Interrupted.

Adding Nodes if the Regular Process Is Interrupted

It's possible for the process of adding nodes to a cluster be interrupted, for example by a problem with a background service. In that case, you might have *assigned* the nodes to the cluster, but they're prevented from being *added*. If so, you can perform the following tasks to complete the process.

To complete the process of adding nodes to a cluster, if the process is stuck, do the following:

1. Go to the Cluster Details page and click the **Available Nodes** tab. If you added nodes but the process isn't complete, or if the process failed, those nodes are listed on this page.

2. Click **Extend Cluster with Available Nodes**.
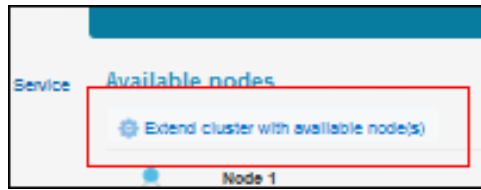
3. The **Extend** *clustername* **dialog** box shows the number of nodes that are available to be added. The dialog box shows different messages, depending on the nodes you're adding.

If the final number of permanent nodes in the cluster will be five or more, you're prompted to select how many to allocate as permanent Hadoop nodes and how many as edge nodes. The first four nodes in the cluster must be permanent Hadoop nodes. Select how many to allocate as each type, enter your Cloudera password, and click the **Extend** button.
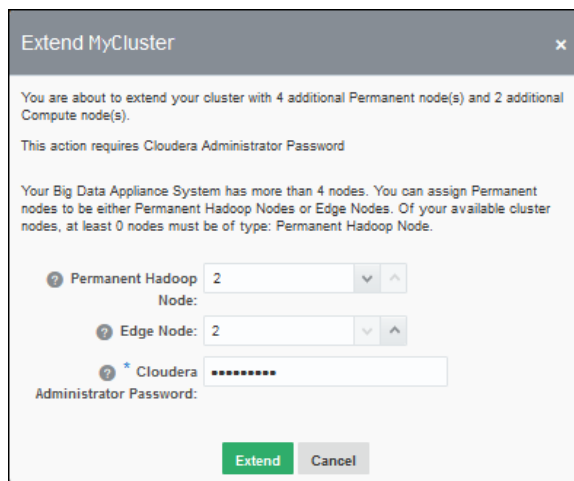
If you're just entering a fourth permanent Hadoop node, you aren't prompted to allocate permanent nodes as one or another type. Just enter your Cloudera password and click **Extend**.
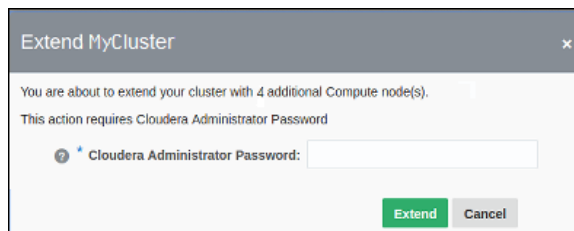
When successful, the new status appears on the Cluster Summary page of the Service Console and on the Cluster Details page.

Controlling Network Access to Services

Oracle Big Data Cloud Machine provides a utility that uses Linux `iptables` and configuration files called *whitelists* to filter network traffic to services in the cluster. The whitelist configuration specifies whether network requests from specified clients will be accepted or denied for specified services at specified ports. When a connection tries to establish itself, `iptables` looks for a matching client IP address or range of IP addresses in the whitelist. If it doesn't find one, it uses the default action.

Services That Can Be Added to the Whitelist

The following table shows the services that can be added to the whitelist, the files that store the configurations for those services, and their default behaviors; that is, when their configuration files are empty:

Service	Whitelist Configuration File	Default (Empty File)
Cloudera Manager	`/opt/oracle/bda/cloudera_manager_whitelist`	Deny access
Hue	`/opt/oracle/bda/hue_whitelist`	Deny access
Secure Shell (SSH)	`/opt/oracle/bda/ssh_whitelist`	Allow access

Don't edit the configuration files directly. Use the `bdacli bdcs_whitelist` command, as described below.

Managing the Whitelist With the bdacli bdcs_whitelist Command

Use the `bdacli bdcs_whitelist` command to manage the whitelist configuration for a cluster.

Run the command as the `root` user on the primary host of the cluster. To find out what your primary host is, connect to any node and enter `bdacli getinfo cluster_primary_host`. For example:

```
# bdacli getinfo cluster_primary_host
host1891
```

bdacli bdcs_whitelist Usage

Syntax

```
bdacli bdcs_whitelist parameters
```

Parameters

The following table describe the parameters for the `bdacli bdcs_whitelist` command.

Parameter	Description
`reset_default_config`	Resets the files to empty and applies the default configuration. Denies all incoming traffic to the server except SSH, which allows all traffic. This only affects ports controlled by the whitelist. All other `iptables` configurations aren't touched.
`reload_config`	Deletes all `iptables` rules on ports controlled by the whitelist files and reprocesses what is in the whitelist files. If the files are empty, the default configuration is applied. If there are one or more entries, then all traffic is denied except the whitelist in the files.
`allow service ip/range`	Adds an IP address or a range of IP addresses to the whitelist of the named service and runs an `iptables` command to allow access to that service from those IP addresses. See the **Variable** table, below, for descriptions of the `service` and `ip/range` variables.
`deny service ip/range`	Removes an IP address or a range of IP addresses to the whitelist of a specific service and runs an `iptables` command to deny access to that service from those IP addresses. See the **Variables** table, below, for descriptions of the `service` and `ip/range` variables.

Variables

The following table describe the variables for the `bdacli bdcs_whitelist allow` and `bdacli bdcs_whitelist deny` commands.

Variable	Description
`service`	One of the following: • `cloudera_manager` • `hue` • `ssh` • `all` — where the command applies to all the above services. We recommend that you do not use `all`, but rather run the command for just the services you need to open. This is recommended because it leaves the critical SSH settings alone (which default to open) unless you explicitly change them. Changing SSH settings from the default can lock you permanently out of the cluster.

Variable	Description
`ip/range`	One of the following: • Specify a single IP address by simply giving the address, for example: **192.0.2.48** • Specify a range of IP addresses by using either of the following: **192.0.2.0/24** **192.0.2.0/255.255.255.0** Both of the above ranges mean `192.0.2.0` to `192.0.2.255` inclusive

Example

```
# bdacli bdcs_whitelist allow cloudera_manager 198.51.100.48
BDCS Network Services Firewall & Whitelist
host1891.us.example.com
host1892.us.example.com
host1893.us.example.com
host1894.us.example.com
host1895.us.example.com
```

In the above example:

- `bdacli bdcs_whitelist allow cloudera_manager` specifies that Cloudera Manager will accept requests from the specified client IP addresses.

- `198.51.100.48` specifies that the client with that IP address will be allowed access to the service.

- `host1891.us.example.com` through `host1895.us.example.com` are the host names of the nodes of the cluster.

Reviewing the Configuration

After changing the configuration, you should check to see that it is as expected. There are two ways you can do this:

- Review the whitelist configuration files for the services, for example:

```
# more /opt/oracle/bda/cloudera_manager_whitelist
198.51.100.48
# more /opt/oracle/bda/hue_whitelist
198.51.100.48
```

- Use the `iptables -L` command to show the network firewall rules in effect, including whitelist rules. You must run this command as the `root` user.

Viewing All Clusters

From the Oracle Big Data Cloud Service console, you can:

- View the total resources allocated across all clusters in the service.

- View the details for each cluster.

To view all clusters, open the Oracle Big Data Cloud Service console. See Accessing the Oracle Big Data Cloud Service Console.

For information about the details provided on the Oracle Big Data Cloud Service console, see Exploring the Oracle Big Data Cloud Service Console Clusters Page.

Viewing Details About a Cluster

From the Oracle Big Data Cloud Service Cluster Details page, you can:

- View a summary of details for the cluster.

- View the resources allocated for the cluster.

- View the details and status information for each node in the cluster.

To view detailed information for a cluster:

1. Open the Oracle Big Data Cloud Service console. See Accessing the Oracle Big Data Cloud Service Console.

2. Click the ≡ icon on the row of the service instance for which you want to view more information, and select **View Details**.

   The Oracle Big Data Cloud Service Cluster Details page is displayed. See Exploring the Oracle Big Data Cloud Service Cluster Details Page for description and instructions on all the items on the page.

Using HDFS Transparent Encryption

HDFS Transparent Encryption protects Hadoop data that's at rest on disk. When the encryption is enabled for a cluster, data write and read operations on encrypted *zones* (HDFS directories) on the disk are automatically encrypted and decrypted. This process is "transparent" because it's invisible to the application working with the data. HDFS Transparent Encryption does not affect user access to Hadoop data, although it can have a minor impact on performance.

Prerequisite

The cluster where you want to use HDFS Transparent Encryption must have Kerberos enabled.

> **Important:**
>
> Security Setup must be enabled when creating the cluster. The person creating the cluster must choose the **Security Setup: Enabled** option on the Security page of the Create Cluster wizard, as described in Creating a Cluster. You can't enable Kerberos for a cluster after it's been created.

When you create a cluster with Security Setup enabled, the following takes place:

- HDFS Transparent Encryption is enabled on the cluster. You can verify this by entering the following at the command line:

  ```
  bdacli getinfo cluster_hdfs_transparent_encryption_enabled
  ```

- MIT Kerberos, Sentry, Network Firewall, Network Encryption, and Auditing are also enabled on the cluster.

- Two principals are created as part of the Kerberos configuration:

  - `hdfs/`*`clustername`*`@BDACLOUDSERVICE.ORACLE.COM` — The password for authenticating this principal is your Cloudera admin password.

  - `oracle/`*`clustername`*`@BDACLOUDSERVICE.ORACLE.COM` — The password for authenticating this principal is your Oracle operating system password.

  In both cases, *`clustername`* is the name of your cluster and `BDACLOUDSERVICE.ORACLE.COM` is the Kerberos realm for Oracle Big Data Cloud Machine.

- A Key Trustee Server is installed and configured on the cluster. This server is used for managing keys and certificates for HDFS Transparent Encryption. See Cloudera Navigator Key Trustee Server for more information about this server. (You should back up Key Trustee Server databases and configuration files on a regular schedule. See the Cloudera documentation topic, Backing Up and Restoring Key Trustee Server.)

Creating Encryption Zones on HDFS

An encryption zone is an HDFS directory in which the contents are encrypted on a write operation and decrypted on a read operation.

See Also:

Cloudera documentation Managing Encryption Keys and Zones.

Prerequisites:

1. Make sure services are healthy in Cloudera Manager. Especially make sure the Key Trustee service is healthy.

2. Make sure the two KMS hosts are in sync.

   On each KMS host run the commands below as the `root` user. The output should be the same on each host. If not, open a service request (SR) with Oracle Support, because that would indicate a problem synchronizing the two Key Management Servers.

```
# ls -l /var/lib/kms-keytrustee/keytrustee/.keytrustee
# cksum /var/lib/kms-keytrustee/keytrustee/.keytrustee/*
# gpg --homedir /var/lib/kms-keytrustee/keytrustee/.keytrustee --fingerprint;
```

Perform the following steps on any node of the cluster as the `root` user, unless otherwise specified.

To create an encryption zone:

1. Create an encryption key for the zone:

   a. Authenticate the `hdfs/`*`clustername`*`@BDACLOUDSERVICE.ORACLE.COM` principal using your Cloudera password, for example:

```
# kinit -p hdfs@BDACLOUDSERVICE.ORACLE.COM
Password for hdfs@BDACLOUDSERVICE.ORACLE.COM: ****
```

    b. Create the encryption key, using the following command::

```
hadoop key create keyname
```

For example:

```
# hadoop key create bdakey
bdakey has been successfully created with options Options{cipher='AES/CTR/
NoPadding', bitLength=128, description='null',
attributes=null}.
org.apache.hadoop.crypto.key.kms.LoadBalancingKMSClientProvider@4145bad8 has
been updated.
```

2. Create a new empty directory and make it an encryption zone using the key generated above with the following two commands:

```
# hadoop fs -mkdir path
# hdfs crypto -createZone -keyName keyname -path path
```

For example:

```
# hadoop fs -mkdir /zone
# hdfs crypto -createZone -keyName bdakey -path /zone
Added encryption zone /zone
```

> **Note:** Encryption zones must be created as the super user, but after that access to encrypted file data and metadata is controlled by normal HDFS file system permissions.

3. Verify creation of the new encryption zone by running the -listZones command; for example:

```
# hdfs crypto -listZones
/zone bdakey
```

Adding Files to Encryption Zones

Use the hadoop fs -put command to add files to the encryption zone.

For example:

```
# hadoop fs -put helloWorld /zone
```

Viewing Keys in Encryption Zones

Use the hadoop key list command to view keys in an encryption zone.

For example:

```
# hadoop key list
Listing keys for KeyProvider:
org.apache.hadoop.crypto.key.kms.LoadBalancingKMSClientProvider@xxxxxx
MYKEY1
MYKEY2
```

Restarting a Cluster Node

To restart a cluster node:

1. Open the Oracle Big Data Cloud Service console. See Accessing the Oracle Big Data Cloud Service Console.

2. Click the ☰ icon on the row of the cluster with the node you want to restart, and select **View Details**.

   The Oracle Big Data Cloud Service Cluster Overview page is displayed.

3. Click the ☰ icon on the row of the node you want to restart, and select **Restart Node**.

Restarting the Virtual Machines (VMs) Hosting a Cluster

To restart the virtual machines (VMs) hosting a Hadoop cluster:

1. Open the Oracle Big Data Cloud Service console. See Accessing the Oracle Big Data Cloud Service Console.

2. Click the ☰ icon on the row of the cluster whose VMs you want to restart, and select **Restart Cluster VMs**.

Updating the SSH Public Key for a Cluster

A cluster must have a Secure Shell (SSH) key pair associated with it to permit secure access for the opc user. When you create a cluster, you must specify the public key. After the cluster has been created, you can replace that key (or any subsequently assigned key) with a new one.

To replace the SSH public key:

1. Open the Service Console for the cluster. See Accessing the Oracle Big Data Cloud Service Console.

2. Find the cluster whose SSH public key you want to change.

3. From the ☰ menu for the cluster, select **Edit**.

   The **Edit** *Cluster name* dialog is displayed. The **SSH Key** field shows the cluster's most recent SSH public key.

4. Check the **Update SSH key** checkbox.

5. Use either of the following methods to specify the new public key:

   a. Select and delete the current key value shown in the text area; then paste in a new one. Make sure the value does not contain extra spaces or line breaks and does not have extra line breaks at the end.

      Note: Some tools generate public SSH keys with a line break at the end, and that is allowed here. However, you shouldn't add any additional line breaks.

   b. Click **Select File** to select a file containing the new public key.

The value of the new public key is pasted into the text area.

6. Click **Edit**.

Supporting Multiple Key Pairs for Secure Shell (SSH) Access

By default, the Oracle Big Data Cloud Machine `opc` user has Secure Shell (SSH) access to all the nodes of the cluster when using the SSH key pair that was provided when the cluster was provisioned. You can also provide SSH access from different clients and for other users. For example, you may want to provide `opc` access to an Oracle Big Data Discovery administrator who accesses the cluster from a different computer, or you may want to create other users with different access rights.

Adding SSH Support for the opc User Using a Different Key Pair

When anOracle Big Data Cloud Machine cluster is provisioned, `/home/opc/.ssh/authorized_keys` files are created on all the nodes of the cluster. The `authorized_keys` files contain the SSH public key that was provided when the cluster was provisioned.

To add an additional public key for the `opc` user,

1. Obtain the new SSH public key.

   The user who needs access to the cluster can create the SSH key pair, retain the private key, and transfer the public key to the Oracle Big Data Cloud Machine administrator. Or, the administrator can create the new key pair and transfer the private key to the other user. See Generating a Secure Shell (SSH) Public/Private Key Pair.

2. Connect as the `opc` user to a node to which you want to add the key. See Connecting to a Cluster Node Through Secure Shell (SSH).

3. On a new line of the `/home/opc/.ssh/authorized_keys` file, paste the contents of the new SSH public key file. Do not add extra lines or line breaks.

4. Repeat the process on every node to which you want to provide access by using the new key pair.

Adding SSH Support for Other User Accounts

To add an SSH key pair for a user other than `opc`:

1. Obtain the new SSH public key.

   The user who needs access to the cluster can create the SSH key pair, retain the private key, and transfer the public key to the Oracle Big Data Cloud Machine administrator. Or, the administrator can create the new key pair and transfer the private key to the other user. See Generating a Secure Shell (SSH) Public/Private Key Pair.

2. Connect as the `opc` user to a node to which you want to add the key. See Connecting to a Cluster Node Through Secure Shell (SSH).

3. Create a `/home/user/.ssh/authorized_keys` file, where `user` is the name of the user who will have SSH access.

4. Paste the contents of the new SSH public key file into`/home/user/.ssh/authorized_keys` file. Do not add extra lines or line breaks.

5. Repeat the process on every node to which you want to provide SSH access for the user.

Deleting a Cluster

To delete a cluster:

1. Open the Oracle Big Data Cloud Service console. See Accessing the Oracle Big Data Cloud Service Console.

2. Click the ☰ icon on the row of the cluster you want to delete, and select **Delete Cluster**.

The cluster is deleted.

3

Accessing Your Oracle Big Data Cloud Machine

This section describes how to access clusters and the tools, utilities and interfaces available in a cluster.

> **Note:** By default, the ports used for accessing Cloudera Manager and Hue are blocked. To control access to those ports, you must configure the network whitelist for the service. That configuration controls whether network requests from specified IP addresses (or ranges of addresses) will be accepted or denied at that port. See Controlling Network Access to Services.

Topics

- Accessing the Oracle Big Data Cloud Service Console

- Connecting to a Cluster Node Through Secure Shell (SSH)

- Accessing Cloudera Manager to Work with Hadoop Data and Services

- Accessing Cloudera Hue to Manage Hadoop Data and Resources

Accessing the Oracle Big Data Cloud Service Console

Use the Oracle Big Data Cloud Service console to create and manage Hadoop clusters.

1. Display the Sign In to Oracle Cloud page by clicking the **My Services URL** link in your Welcome e-mail or by following these instructions:

    a. Open your web browser and go to the Oracle Cloud website: `http://cloud.oracle.com`.

    b. Click **Sign In**.

    c. In the My Services box, select the data center where your services are located; for example, **Public Cloud Services - NA** or **Public Cloud Services - EMEA**.

    d. Click **Sign In to My Services**.

2. On the Sign In to Oracle Cloud page, enter your user name, your password and the name of your identity domain. Then, click **Sign In**.

    The My Services Dashboard is displayed.

3. In the list of services, locate the entry for Oracle Big Data Cloud Service and then click its name.

The Oracle Big Data Cloud Service console is displayed.

Accessing Cloudera Manager to Work with Hadoop Data and Services

You can access Cloudera Manager from the Oracle Big Data Cloud Service console, or you can access it directly from a browser.

Open Cloudera Manager from the Oracle Big Data Cloud Service Console

1. Open the Oracle Big Data Cloud Service console. See Accessing the Oracle Big Data Cloud Service Console.

2. Click the ☰ icon on the row of the cluster you want to manage, and select **Open Cloudera Manager**.

   The Cloudera Manager application is displayed.

Open Cloudera Manager from a Web Browser

> **Note:** By default, the port used for accessing Cloudera Manager, port 7183, is blocked. To control access to that port, you must configure the network whitelist for the service. That configuration controls whether network requests from specified IP addresses (or ranges of addresses) will be accepted or denied at that port. See Controlling Network Access to Services.

To open Cloudera Manager from a web browser:

1. Open the Oracle Big Data Cloud Service console. See Controlling Network Access to Services.

2. Click the ☰ icon on the row of the cluster you want to manage, and select **View Details**.

   The Cluster Overview page is displayed.

3. Find the URL for Cloudera Manager listed at the top of the page.

4. Open a browser and navigate to that URL.

5. Enter your credentials for logging in, as prompted.

Accessing Cloudera Hue to Manage Hadoop Data and Resources

Hue runs in a browser and provides an easy-to-use interface to several applications to support interaction with Hadoop and HDFS. You can use Hue to perform any of the following tasks:

- Query Hive data stores

- Create, load, and delete Hive tables

- Work with HDFS files and directories

- Create, submit, and monitor MapReduce jobs

- Monitor MapReduce jobs

- Create, edit, and submit workflows using the Oozie dashboard

- Manage users and groups

Hue is automatically installed and configured on Oracle Big Data Cloud Machine clusters. Hue runs on port 8888 of the ResourceManager node (node03).

> **Note:** By default, port 8888 is blocked. To control access to that port, you must configure the network whitelist for the service. That configuration controls whether network requests from specified IP addresses (or ranges of addresses) will be accepted or denied at that port. See Controlling Network Access to Services.

To use Hue:

1. Log in to Cloudera Manager and click the **hue** service on the Home page.

2. On the hue page under Quick Links, click `Hue Web UI`.

3. Bookmark the Hue URL, so that you can open Hue directly in your browser. The following URL is an example:

   `https://bda1node03.example.com:8888`

4. Log in with your Hue credentials.

   If Hue accounts haven't been created yet, you can log into the default Hue administrator account, by using the following credentials:

   - **Username**: `admin`

   - **Password:** *cm-admin-password*

   where *cm-admin-password* is the password that was specified when the cluster for the Cloudera Manager `admin` user was activated. You can then create other user and administrator accounts.

   > **See Also:**
   >
   > *Hue User Guide* at
   >
   > http://archive-primary.cloudera.com/cdh5/cdh/5/hue/user-guide/

Connecting to a Cluster Node Through Secure Shell (SSH)

To gain local access to the tools, utilities and other resources on an Oracle Big Data Cloud Machine cluster node, you use Secure Shell (SSH) client software to establish a secure connection and log in.

> **Note:** By default, the port used for accessing the cluster through SSH, port 22, is open. However, you can control access to that port by configuring the network whitelist for SSH. That configuration controls whether network requests from specified IP addresses (or ranges of addresses) will be accepted or denied at that port. See Controlling Network Access to Services.

Several SSH clients are freely available. The following sections show how to use SSH clients on UNIX, UNIX-like and Windows platforms to connect to a node.

The following instructions describe how to connect as the `opc` user and then use the `sudo` command to open a `root` shell. After you do this the first time, you may choose instead to connect using a password by assigning a known password to `opc` or creating another user with a known password. For information about creating users and setting passwords, see Managing User Accounts in *Oracle Big Data Appliance Software User's Guide*.

Connecting to a Node By Using PuTTY on Windows

PuTTY is a freely available SSH client program for Windows.

Before You Begin

Before you use the PuTTY program to connect to a node, you need the following:

- The IP address of the node

  The IP address of the node is listed on the Cluster Details page for the cluster containing the node. To display this page, see Viewing Details About a Cluster.

- The SSH private key file that pairs with the public key associated with the cluster

  The public key was associated with your cluster when it was created. See Creating a Cluster. If you don't have the private key that's paired with the public key, contact your administrator.

  The private key file must of the PuTTY `.ppk` format. If the private key file was originally created on the Linux platform, you can use the PuTTYgen program to convert it to the `.ppk` format.

  For instructions on creating an SSH key pair, see Generating a Secure Shell (SSH) Public/Private Key Pair.

Procedure

To connect to a node using the PuTTY program on Windows:

1. Download and install PuTTY.

   To download PuTTY, go to `http://www.putty.org/` and click the **You can download PuTTY here** link.

2. Run the PuTTY program.

   The PuTTY Configuration window is displayed, showing the Session panel.

3. In **Host Name (or IP address)** box, enter the IP address of the node.

4. Confirm that the **Connection type** option is set to **SSH**.

5. In the Category tree, expand **Connection** if necessary and then click **Data**.

   The Data panel is displayed.

6. In the **Auto-login username** box, enter **opc**. As the `opc` user, you can use the `sudo` command to gain root access to the node, as described in the last step, below.

7. Confirm that the **When username is not specified** option is set to **Prompt**.

8. In the Category tree, expand **SSH** and then click **Auth**.

    The **Auth** panel is displayed.

9. Click the **Browse** button next to the **Private key file for authentication** box. Then, in the **Select private key file** window, navigate to and open the private key file that matches the public key that is associated with the cluster.

10. In the **Category** tree, click **Session**.

    The **Session** panel is displayed.

11. In the **Saved Sessions** box, enter a name for this connection configuration. Then, click **Save**.

12. Click **Open** to open the connection.

    The PuTTY Configuration window is closed and the PuTTY window is displayed.

    If this is the first time you are connecting to the VM, the PuTTY **Security Alert** window is displayed, prompting you to confirm the public key. Click **Yes** to continue connecting.

13. To perform operations that require root access to the node—such as issuing `bda-ossadmin` commands—open a root command shell. Enter `sudo -s` at the command prompt:

    ```
    $ sudo -s
    # whoami
    # root
    ```

Connecting to a Node By Using SSH on UNIX

UNIX and UNIX-like platforms (including Solaris and Linux) include the `ssh` utility, an SSH client.

Before You Begin

Before you use the `ssh` utility to connect to a node, you need the following:

*   The IP address of the node

    The IP address of the node is listed on the Cluster Details page of the cluster containing the node. To display this page, see Viewing Details About a Cluster.

*   The SSH private key file that pairs with the public key associated with the cluster

    The public key was associated with your cluster when it was created. See Creating a Cluster. If you don't have the private key that's paired with the public key, contact your administrator.

Procedure

To connect to a node using the ssh utility on UNIX and UNIX-like platforms:

1. In a command shell, set the file permissions of the private key file so that only you have access to it:

    ```
    $ chmod 600 private-key-file
    ```

private-key-file is the path to the SSH private key file that matches the public key that is associated with the cluster.

2. Run the ssh utility:

```
$ ssh -i private-key-file opc@node-ip-address
```

where:

- *private-key-file* is the path to the SSH private key file.

- **opc** is the opc operating system user. As opc, you can use the sudo command to gain root access to the node, as described in the next step.

- *node-ip-address* is the IP address of the node in *x.x.x.x* format.

If this is the first time you are connecting to the node, the ssh utility prompts you to confirm the public key. In response to the prompt, enter **yes**.

3. To perform operations that require root access to the node—such as issuing bda-oss-admin commands—open a root command shell. Enter sudo -s at the command prompt:

```
$ sudo -s
# whoami
# root
```

4

Managing Oracle Big Data Cloud Machine

The Oracle Linux operating system and Cloudera's Distribution including Apache Hadoop (CDH) underlie all other software components installed on Oracle Big Data Cloud Machine .

Topics

- Oracle Big Data Cloud Machine Management Software

- Installing Oracle Big Data Cloud Machine Software

- Oracle Big Data Cloud Machine Utilities

- Executing Commands Across a Cluster Using the dcli Utility

Oracle Big Data Cloud Machine Management Software

Management software components are installed on all nodes of an Oracle Big Data Cloud Machine cluster. Oracle Linux, required drivers, firmware, and hardware verification utilities are factory installed. Among them are the following tools::

- The `bdacli` utility queries various configuration files to return information about the cluster, nodes, and software patches. See bdacli .

- The `setup-root-ssh` utility sets up passwordless SSH for the root user among all the nodes in an Oracle Big Data Cloud Machine cluster. See Setting Up Passwordless SSH.

- The `dcli` utility executes commands across a group of nodes on a Oracle Big Data Cloud Machine cluster and returns the output. See Executing Commands Across a Cluster Using the dcli Utility.

Installing Oracle Big Data Cloud Machine Software

This chapter explains how to reinstall, and reconfigure software on an Oracle Big Data Cloud Machine cluster.

> **Note:**
>
> Ensure that you know the current passwords for the operating system `root` and `oracle` users and the the Cloudera Manager `admin` user. If you're reinstalling Oracle Big Data Connectors, then you also need the password for Oracle Data Integrator.

Topics

- [Upgrading the Software on a Cluster](#)
- [Adding Support for Oracle Audit Vault and Database Firewall](#)
- [Installing a One-Off Patch](#)

Upgrading the Software on a Cluster

The procedure for upgrading the software is the same whether you are upgrading from one major release to another or just applying a patch set.

The process upgrades all components of the software stack including the firmware, Oracle Linux Unbreakable Enterprise Kernel (UEK), CDH, JDK, and Oracle Big Data Connectors.

To upgrade only Oracle Big Data Connectors, and no other components of the software stack, contact Oracle Support for assistance.

Software downgrades are not supported.

Upgrading the Software

Follow these procedures to upgrade the software on an Oracle Big Data Cloud Machine cluster to the current version.

Prerequisites

You must know the passwords currently in effect for the cluster, which the Mammoth utility will prompt you for:

- `oracle`
- `root`
- Cloudera Manager `admin`

Upgrading to the Current Software Version

Making sure cluster services are healthy before upgrade, and especially after reboot is very important. Manual steps will be needed to resume.

Upgrade the Oracle Big Data Cloud Machine software to the current software version as follows. This is a summary. Refer to My Oracle Support (MOS) Doc ID 2101906.1 for details, including prerequisites, further information on the steps below, and known issues.

> **Note:** All Oozie jobs should be stopped before the upgrade. Failure to do this may cause the upgrade to fail.

---

1. Download and unzip the Mammoth bundle, as described in Downloading the Mammoth Software Deployment Bundle.. Mammoth is a command-line utility for installing and configuring the Oracle Big Data Cloud Machine software.

   You must be logged in as root to node 1 of the cluster.

2. Change to the `BDAMammoth` directory.

   ```
   # cd /opt/oracle/BDAMammoth
   ```

3. Run the `mammoth` command with the -p option:

```
# ./mammoth -p
```

Mammoth automatically upgrades the base image if necessary.

4. After all nodes reboot, perform the following checks.

   a. Check uptime.

   ```
   # dcli -C uptime
   ```

   b. Check /root/BDA_REBOOT_SUCCEEDED.

   ```
   # dcli -C ls -ltr /root/BDA_REBOOT_SUCCEEDED
   ```

   Note: Note: If there is no `BDA_REBOOT_SUCCEEDED`, check for `/root/BDA_REBOOT_*` and `/root/bda_reboot_status`.

   c. Verify that the kernel and JDK are upgraded.

   ```
   # dcli -C uname -a
   # dcli -C java -version
   ```

   d. Check that all Cloudera Configuration Manager services are healthy. You may need to manually restart some services.

   Note: During any upgrade, it is **crucial that all services in Cloudera Manager are healthy after the reboot before continuing**. Failure to do so will result in upgrade failures.

5. After the reboot and the post reboot checks, log on to node 1 of the cluster and rerun `mammoth -p` in order to resume the upgrade.

   ```
   # cd /opt/oracle/BDAMammoth
   # ./mammoth -p
   ```

6. When the upgrade is complete, perform the post-upgrade validation steps described in the MOS document (Doc ID 2101906.1).

Adding Support for Oracle Audit Vault and Database Firewall

Only Release 12.1.1 of Oracle Audit Vault and Database Firewall Server is supported for use with Oracle Big Data Cloud Machine at this time.

Before installing support on Oracle Big Data Cloud Machine, ensure that Oracle Audit Vault and Database Firewall Server is up and running. The software must be installed on a separate server on the same network as Oracle Big Data Cloud Machine.

You must also have the following information about the Audit Vault Server installation:

• Audit Vault Server administration user name and password

• Database service name

• IP address

• Port number

To add support for Oracle Audit Vault and Database Firewall:

1. Log in to the first NameNode (`node01`) of the cluster.

2. Add support for Oracle Audit Vault and Database Firewall:

```
# bdacli enable auditvault
INFO: Logging all actions in /opt/oracle/BDAMammoth/bdaconfig/tmp/
bda1node01-20140805072714.log and traces in /opt/oracle/BDAMammoth/bdaconfig/tmp/
bda1node01-20140805072714.trc
INFO: This is the install of the primary rack
INFO: Checking if password-less ssh is set up
INFO: Executing checkRoot.sh on nodes

/opt/oracle/BDAMammoth/bdaconfig/tmp/all_nodes #Step -1#
SUCCESS: Executed checkRoot.sh on nodes

/opt/oracle/BDAMammoth/bdaconfig/tmp/all_nodes #Step -1#
INFO: Executing checkSSHAllNodes.sh on nodes

/opt/oracle/BDAMammoth/bdaconfig/tmp/all_nodes #Step -1#
SUCCESS: Executed checkSSHAllNodes.sh on nodes

/opt/oracle/BDAMammoth/bdaconfig/tmp/all_nodes #Step -1#
INFO: Reading component versions from

/opt/oracle/BDAMammoth/bdaconfig/COMPONENTS
INFO: Creating nodelist files...
Please enter the Audit Vault Server Admin Username: admin_username
Please enter the Audit Vault Server Admin Password: admin_password
Enter password again: admin_password
Please enter the Audit Vault Server Database Service Name: service_name
Please enter the Audit Vault Server IP Address: IP address
Please enter the Audit Vault Server Port: port_number
INFO: Creating environment.pp file ...
INFO: Making sure all puppet agents can be accessed.
INFO: Pinging puppet agents
INFO: Adding audit Vault Service. This will take some time ...
   .
   .
   .
```

Installing a One-Off Patch

One-off patch bundles provide a fix to specific bugs in one or more releases. You use Mammoth to apply the patch to your cluster.

To install a one-off patch bundle:

1. Download the patch bundle from the Automated Release Update (ARU) system to a directory such as /tmp on the first node of the Oracle Big Data Cloud Machine cluster.

   The file is named BDA-patch-release-patch.zip. The examples in this procedure use the name BDA-patch-4.3.1-123456.zip.

2. Unzip the file. For example:

```
# unzip BDA-patch-4.3.0-123456.zip
Archive:  BDA-patch-4.3.0-123456.zip
   creating: BDA-patch-4.3.0-123456/
  inflating: BDA-patch-4.3.0-123456/BDA-patch-4.3.0-123456.run
  inflating: BDA-patch-4.3.0-123456/README.txt
```

3. Change to the patch directory created in Step 2. For example:

```
$ cd BDA-patch-4.3.0-123456
```

4. Extract the contents of the run file. For example:

```
$ ./BDA-patch-4.3.0-123456.run
Big Data Appliance one-off patch 123456 for v4.3.0 Self-extraction

Removing existing temporary files

Generating /tmp/BDA-patch-4.3.0-123456.tar
Verifying MD5 sum of /tmp/BDA-patch-4.3.0-123456.tar
/tmp/BDA-patch-4.3.0-123456.tar MD5 checksum matches

Extracting /tmp/BDA-patch-4.3.0-123456.tar to /opt/oracle/BDAMammoth/patches/
123456
Removing temporary files

Please "cd /opt/oracle/BDAMammoth" before running "./mammoth -p 123456"
```

5. Change to the BDAMammoth directory:

```
$ cd /opt/oracle/BDAMammoth
```

6. Install the patch. For example:

```
$ ./mammoth -p 123456
```

Alternatively, you can use the bdacli command. See bdcli.

Oracle Big Data Cloud Machine Utilities

This chapter describes the utilities available on Oracle Big Data Cloud Machine.

Topics

- bdacli

- remove-root-ssh

- setup-root-ssh

bdacli

The Oracle Big Data Cloud Machine command-line interface (bdacli) queries various configuration files to return information about the rack, cluster, server, InfiniBand network, and software patches.

Syntax

```
bdacli action [parameters]
```

Actions

help
Displays general usage information for bdacli, a list of actions, and a list of supported parameters for the getinfo action.

{add | remove} patch *patch_number*
Adds or removes a software patch on Oracle Big Data Cloud Machine that matches *patch_number*. You must log in as `root` to use `add` or `remove`.

admin_cluster *parameter node_name*
Enables you to administer the nodes in a cluster in response to a failing server. The following table describes the parameters.

Parameter	Description
decommission	Removes the specified node from the cluster and decommissions the server in Cloudera Manager. It also updates the Mammoth files. You can decommission a failing, noncritical node. Note that critical services on the node must be moved first.
recommission	Removes the node from the list of decommissioned nodes, and recommissions the server in Cloudera Manager. Use this command after decommissioning and repairing a failing server.
migrate	Moves the services from a critical node to a noncritical node, and decommissions the failing server in Cloudera Manager. You specify the name of the failing critical node, and the utility selects the noncritical node for the migration. When migration is complete, the new node has all of the functionality of the original critical node. You can only migrate a critical node, and should do so only when it is failing.
reprovision	Restores a server to the cluster as a noncritical node, and recommissions the server in Cloudera Manager. Use this command after migrating the services of a critical node and repairing the failing server.

{enable | disable} service
Enables or disables an optional software component on Oracle Big Data Cloud Machine. You must log in as `root` to use `enable` or `disable`.

The following describes the component parameters.

Parameter	Description
asr	Auto Service Request
auditvault	Oracle Audit Vault and Database Firewall plugin
bdc	Oracle Big Data Connectors
https_cm_hue_oozie	HTTPS for Cloudera Manager, Hue, and Oozie. There is no `disable` option. Kerberos must be enabled. **Note:** You can't enable Kerberos on a cluster after it's been created. You must specify that Kerberos is to be enabled when creating the cluster.

Parameter	Description
hdfs_encrypted_data_transport	Encryption of data transfer between DataNodes and clients, and among DataNodes. Kerberos must be enabled. **Note:** You can't enable Kerberos on a cluster after it's been created. You must specify that Kerberos is to be enabled when creating the cluster.
hadoop_network_encryption	HTTPS encryption and Kerberos authentication for HDFS and YARN, also enables web shuffling for YARN. Kerberos must be enabled. **Note:** You can't enable Kerberos on a cluster after it's been created. You must specify that Kerberos is to be enabled when creating the cluster.
sentry	Apache Sentry authorization

{start | stop | restart | status} service

Starts, stops, restarts, or returns the current status of a service on a cluster or a specific node.

The following table describes the service parameters:

Parameter	Description
big_data_sql_cluster	Oracle Big Data SQL on all nodes of the cluster
big_data_sql_server *node_name*	Oracle Big Data SQL on a specified node of the cluster. Use bdacli with this parameter only from the first node of the cluster, where the current config.json file is stored.

getinfo [parameter]

Returns a list of getinfo parameters. If you include a parameter name in the command, then getinfo returns information about that system component:

- **InfiniBand parameters**: The bdacli command queries the InfiniBand fabric. See InfiniBand Parameters.

- **Cluster parameters**: Describes a logical Oracle Big Data Cloud Machine cluster. The bdacli command queries the current config.json file for the Hadoop cluster where the command executes. See Cluster Parameters.

- **Server parameters**: Describes a server. The bdacli command queries the operating system of the server where the bdacli command executes. See Server Parameters.

- **One-off patch parameters**: Provides information about one-off patches. See One-Off Patch Parameters.

Parameter names that end with an "s" return lists. Boolean parameters return a string of either true or false.

InfiniBand Parameters

The following table describes the InfiniBand parameters for `bdacli getinfo`.

Parameter	Returns
`ib_bda_servers`	A list of all Oracle Big Data Cloud Machine servers on the InfiniBand fabric. The list contains the unqualified client host names, which are sorted alphabetically. The list can include servers in multiple racks that are cabled together. See Examples.
`ib_switches`	Lists the switches on the InfiniBand fabric, including the switch name, Globally Unique Identifier (GUID), and type (`GTW` for a gateway switch, or `36P` for a spine switch). The list contains the unqualified administration names, which are sorted alphabetically by switch name. See See Examples.

Cluster Parameters

The following tables describe the cluster parameters.

- General Cluster Parameters

- Oracle Big Data Connectors Status Parameters

- Cluster Network Parameters

- Cluster Security Parameters

The following table describes the general cluster parameters for `bdacli getinfo`.

General Cluster Parameters

Parameter	Returns
`cluster_asr_installed`	`true` if Auto Service Request is configured for this cluster; `false` otherwise
`cluster_big_data_sql_enabled`	`true` if Oracle Big Data SQL is enabled for this cluster; `false` otherwise.
`cluster_cdh_version`	The version of Cloudera's Distribution including Apache Hadoop installed on this cluster, such as `4.5.0-ol6`.
`cluster_cm_server`	The Cloudera Manager address, including the server name and port number, such as `bda1node03.example.com:7180`.
`cluster_cm_version`	The version of Cloudera Manager running on this cluster, such as `4.8.0-ol6`.
`cluster_name`	The name of the cluster, such as `cluster-c`.

Parameter	Returns
`cluster_primary_host`	The unqualified host name of the server that hosts the puppet master. The Mammoth utility was deployed from this host, and any reconfiguration of the cluster must be done while logged in to that server.
`cluster_type`	The type of cluster (default: `Hadoop`).
`cluster_version`	The software version installed on this cluster by the Mammoth utility, such as 3.1.0.

The following table describes describes the cluster parameters related to Oracle Big Data Connectors for `bdacli getinfo`.

Oracle Big Data Connectors Status Parameters

Parameter	Returns
`cluster_bdc_installed`	true if Oracle Big Data Connectors is installed; false otherwise
`cluster_odi_enabled`	true if the Oracle Data Integrator agent is enabled; false otherwise.
`cluster_odi_version`	The version of Oracle Data Integrator agent installed on the cluster.
`cluster_oraah_version`	The version of Oracle R Advanced Analytics for Hadoop installed on the cluster
`cluster_oraloader_version`	The version of Oracle Loader for Hadoop installed on the cluster
`cluster_osch_version`	The version of Oracle SQL Connector for HDFS installed on the cluster
`cluster_oxh_version`	The version of Oracle XQuery for Hadoop installed on the cluster

The following table describes the cluster network parameters for `bdacli getinfo`.

Cluster Network Parameters

Parameter	Returns
`cluster_hosts_entries`	A list of `/etc/hosts` entries that you can append to the `/etc/hosts` file on any device on the same InfiniBand fabric as the Oracle Big Data Cloud Machine cluster, to ensure that Hadoop traffic uses the InfiniBand network. Do not add these entries to a device on a different fabric.
	Each entry is on a separate line with three parts: the InfiniBand IP address, the full client host name, and the short client host name.
`cluster_ilom_ips`	An ordered list of IP addresses for the Oracle ILOMs in the servers, starting with the first node in the cluster
`cluster_ilom_names`	A list of unqualified host names on the administrative network for the Oracle ILOMs in the servers, in order starting with the first server in the cluster
`cluster_node_ips`	The IP addresses on the client network of all nodes in this cluster
`cluster_node_names`	The host names on the client network of all nodes in the cluster, such as `bda1node01`

The following table describes the cluster security parameters for `bdacli getinfo`.

Cluster Security Parameters

Parameter	Returns
`cluster_av_admin`	The name of the Oracle Audit Vault and Database Firewall administration user. Returns an error if Audit Vault is not configured for this cluster.
`cluster_av_enabled`	`true` if Oracle Audit Vault and Database Firewall auditing is enabled; `false` otherwise
`cluster_av_port`	The port number that the Audit Vault server listens on. Returns an error if Oracle Audit Vault and Database Firewall is not configured on this cluster.
`cluster_av_server`	The IP address of the Audit Vault server. Returns an error if Oracle Audit Vault and Database Firewall is not configured on this cluster.

Parameter	Returns
cluster_av_service	The database service name for the Audit Vault server. Returns an error if Oracle Audit Vault and Database Firewall is not configured on this cluster.
cluster_hdfs_encryption_enabled	true if network encryption of Hadoop data is enabled for this cluster; false otherwise
cluster_hdfs_transparent_encryption_enabled	true if HDFS Transparent Encryption of Hadoop data at rest is enabled for this cluster; false otherwise
cluster_kerberos_enabled	true if Kerberos security is enabled; false otherwise.
cluster_kerberos_kdc_hosts	A list of key distribution center (KDC) hosts external to Oracle Big Data Appliance. Returns an error if Kerberos is not enabled.
cluster_kerberos_kdc_on_bda	true if the Kerberos KDC is on Oracle Big Data Appliance; false otherwise. Returns an error if Kerberos is not enabled.
cluster_kerberos_realm	The Kerberos realm for the cluster. Returns an error if Kerberos is not enabled.
cluster_sentry_enabled	true if Sentry is configured on the cluster; false otherwise.

Server Parameters

The following table describes the server parameters for bdacli getinfo.

Parameter	Returns
server_mammoth_installed	true if the Mammoth utility has deployed the Oracle Big Data Appliance software on this server; false otherwise.
server_name	The name of this server on the client network, such as bda1node01.
server_os_version	The version of Oracle Linux on this server, such as 6.4.

One-Off Patch Parameters

The following table describes the one-off patch parameters for bdacli getinfo.

Parameter	Returns
`available_patches`	A list of valid patches available for installation. A valid patch has a directory under `/opt/oracle/bda/patches` or `/opt/oracle/BDAMammoth/patches` that contains a file named `inventory`.
`installed_patches`	A list of patches already installed. An installed patch has a directory in both `/opt/oracle/bda/patches` and `/opt/oracle/BDAMammoth/patches`.

Examples

The following commands provide information about the optional software on the cluster:

```
# bdacli getinfo cluster_bdc_installed
true
# bdacli getinfo cluster_odi_version
11.1.1.7.0

# bdacli getinfo cluster_hdfs_transparent_encryption_enabled
true
```

The following command lists all switches on the current InfiniBand fabric. In this example, three Oracle Big Data Cloud Machine racks are on the fabric with the standard hardware configuration of one spine switch and two gateway switches each.

```
$ bdacli getinfo ib_switches
bda1sw-iba0 00:21:28:6c:c8:af:a0:a0 36P
bda1sw-ibb0 00:21:28:46:9e:3b:a0:a0 36P
bda1sw-ibs0 00:21:28:6c:c8:ae:a0:a0 36P
bda2sw-ib1 00:21:28:46:98:d3:a0:a0 36P
bda2sw-ib2 00:21:28:de:ae:4a:c0:a0 GTW
bda2sw-ib3 00:21:28:c3:70:9a:c0:a0 GTW
bda3sw-ib1 00:21:28:46:90:ee:a0:a0 36P
bda3sw-ib2 00:21:28:df:34:8a:c0:a0 GTW
bda3sw-ib3 00:21:28:df:0f:0a:c0:a0 GTW
bda4sw-ib1 00:21:28:e8:af:23:a0:a0 36P
bda4sw-ib2 00:10:e0:0c:48:a0:c0:a0 GTW
bda4sw-ib3 00:21:28:f4:82:ce:c0:a0 GTW
```

This example installs patch 1234:

```
$ bdacli add patch 1234
```

remove-root-ssh

Removes passwordless SSH previously established by the `setup-root-ssh` ommand.

Syntax

```
remove-root-ssh [-C | -c | -g | -j] [-p]
```

```
remove-root-ssh -h
```

Parameters

-C
Targets all servers in the cluster, using the list of servers in `/opt/oracle/bda/cluster-hosts-infiniband`.

-c *host1, host2,...*
Targets the servers specified as *host1*, *host2*, and so forth, on the command line.

-g *groupfile*
Targets a user-defined set of servers listed in *groupfile*. You can enter either server names or IP addresses in the file, one per line.

-j "etho0_ips[*range*]"
Specifies the range of servers in a starter pack [1-3] or a starter pack plus additional nodes [1–60]. This parameter is required in the 2.2.x base image when the utility is used before network configuration.

-h
Displays Help.

-p *password*
Specifies the `root` password on the command line.

Oracle recommends that you omit this parameter. You will be prompted to enter the password, which the utility does not display on your screen.

Usage Notes

You must know the `root` password to use this command.

If you don't specify the target servers, then `remove-root-ssh` uses all servers in the cluster.

See Also
setup-root-ssh, Setting Up Paswordless SSH

Example

This example shows passwordless SSH being removed:

```
# remove-root-ssh
calling /opt/oracle/bda/bin/dcli -c 192.168.42.37,192.168.42.38,192.168.42.39 --unkey
192.168.42.37: ssh key dropped
192.168.42.38: ssh key dropped
192.168.42.39: ssh key dropped
remove-root-ssh succeeded
```

setup-root-ssh

Establishes passwordless SSH for the `root` user.

Syntax

```
setup-root-ssh [-C | -c | -g | -j] [-p]

setup-root-ssh -h
```

Parameters

-C

Targets all servers in the cluster, using the list of servers in `/opt/oracle/bda/cluster-hosts-infiniband`.

-c *host1, host2,...*

Targets the servers specified as `host1`, `host2`, and so forth, on the command line.

-g *groupfile*

Targets a user-defined set of servers listed in *groupfile*. You can enter either server names or IP addresses in the file, one per line.

-j "etho0_ips[*range*]"

Specifies the range of servers in a starter pack [`1-3`] or a starter pack and additional servers [`1-60`]. This parameter is required in the 2.2.x base image when the utility is used before network configuration.

-h

Displays Help.

-p *password*

Specifies the root password on the command line.

Oracle recommends that you omit this parameter. You will be prompted to enter the password, which the utility does not display on your screen.

Usage Notes

You must know the `root` password to use this command.

If you don't specify the target servers, then `setup-root-ssh` uses all servers in the cluster.

See Also

remove-root-ssh , Setting Up Paswordless SSH

Example

This example shows passwordless SSH being set up for `root`:

```
# setup-root-ssh
Enter root password: password
spawn /opt/oracle/bda/bin/dcli -c 192.168.42.37,192.168.42.38... -k
root@192.168.42.37's password:
root@192.168.42.38's password:
    .
    .
    .
192.168.42.37: ssh key added
192.168.42.38: ssh key added
    .
    .
```

```
      .
setup-root-ssh succeeded
```

Executing Commands Across a Cluster Using the dcli Utility

The dcli utility executes commands across a group of servers on Oracle Big Data Cloud Machine and returns the output.

This chapter contains the following sections:

- Overview of the dcli Utility

- Setting Up Passwordless SSH

- Basic Use of dcli

- dcli Syntax

- dcli Return Values

- dcli Examples

Overview of the dcli Utility

The `dcli` utility executes commands across a group of nodes in an Oracle Big Data Cloud Machine cluster and returns the output. You use `dcli` to reinstall or reconfigure software on a cluster.

Setting Up Passwordless SSH

The `dcli` utility requires a passwordless Secure Shell (SSH) between the local server and all target servers. You run the `dcli` utility on the local server, and the commands specified in `dcli` execute on the target servers.

Two scripts facilitate the use of SSH on Oracle Big Data Cloud Machine: `setup-root-ssh` and `remove-root-ssh`. These scripts accept two options that are also used by `dcli`:

- `-C`: Targets all the servers in a Hadoop cluster

- `-g`: Targets a user-defined set of servers

See dcli Options for details about these options.

> **Note:** The Mammoth and Mammoth Reconfiguration utilities require remote `root` access. However, Oracle Big Data Cloud Machine doesn't require it for normal operation.

To set up passwordless SSH for root:

1. Connect to a node using PuTTY or a similar utility. Select an SSH connection type. See Connecting to a Cluster Node Through Secure Shell (SSH).

2. Log in as `root`.

3. Set up passwordless SSH for root across the cluster:

   ```
   # setup-root-ssh
   ```

You see the message `ssh key added` from each server.

4. You can now run any `ssh` command on any server in the rack without entering a password. In addition to `dcli` commands, you can use `scp` to copy files between servers.

5. To remove passwordless SSH from `root`:

remove-root-ssh

See also full parameter descriptions of setup-root-ssh and remove-root-ssh .

Basic Use of dcli

Getting Help

To see the `dcli` help page, enter the `dcli` command with the `-h` or `--help` options. You can see a description of the commands by entering the `dcli` command with no options.

Identifying the Target Servers

You can identify the servers where you want the commands to run either in the command line or in a file. For a list of default target servers, use the -t option. To change the target servers for the current command, use the `-c` or `-g` options described in #GUID-42B627A1-5B10-4458-86C2-36091B553AF1/GUID-E21C3C51-E084-4798-98D4-F1B61E1660C3.

You can manually create files with groups of servers to manage together. For example, you might manage servers 5 to 18 together, because they have no special functions like servers 1 to 4.

Specifying the Commands

You typically specify a command for execution on the target servers on the command line. However, you can also create a command file for a series of commands that you often use together or for commands with complex syntax. See the `-x` option in #GUID-42B627A1-5B10-4458-86C2-36091B553AF1/GUID-E21C3C51-E084-4798-98D4-F1B61E1660C3.

You can also copy files to the target servers without executing them by using the `-f` option.

Controlling the Output Levels

You can request more information with the `-v` option or less information with the `-n` option. You can also limit the number of returned lines with the `--maxlines` option, or replace matching strings with the `-r` option.

Following are examples of various output levels using a simple example: the Linux date command.

> **Note:** The output from only one server (node07) is shown. The syntax in these examples executes the `date` command on all servers.

This is the default output, which lists the server followed by the date.

```
# dcli date
bda1node07-adm.example.com: Tue Feb 14 10:22:31 PST 2016
```

The minimal output returns `OK` for completed execution:

```
# dcli -n date
OK: ['bda1node07.example.com']
```

Verbose output provides extensive information about the settings under which the command ran:

```
dcli -v dateoptions.nodes: Noneoptions.destfile: Noneoptions.file:
Noneoptions.group: dcserversoptions.maxLines: 100000options.listNegatives:
Falseoptions.pushKey: Falseoptions.regexp: Noneoptions.sshOptions:
Noneoptions.scpOptions: Noneoptions.dropKey: Falseoptions.serializeOps:
Falseoptions.userID: rootoptions.verbosity 1options.vmstatOps Noneoptions.execfile:
Noneargv: ['/opt/oracle/bda/bin/dcli', '-g', 'dcservers', '-v', 'date']Success
connecting to nodes: ['bda1node07.example.com']...entering thread for
bda1node07.example.com:execute: /usr/bin/ssh -l root bda1node07.example.com '
date' ...exiting thread for bda1node07.example.com status: 0bda1node07.example.com:
Tue Feb 14 10:24:43 PST 2012]
```

dcli Syntax

```
dcli [option] [command]
```

Parameters

option
An option described in the table below. You can omit all options to run a command on all servers in the cluster.

command
Any command that runs from the operating system prompt. If the command contains punctuation marks or special characters, then enclose the command in double quotation marks.

The backslash (\) is the escape character. Precede the following special characters with a backslash on the command line to prevent interpretation by the shell. The backslash is not needed in a command file. See the -x option for information about command files.

- $ (dollar sign)

- ' (quotation mark)

- < (less than)

- > (greater than)

- ( ) (parentheses)

dcli Options

Option	Description
-c *nodes*	Specifies a comma-separated list of Oracle Big Data Cloud Machine nodes where the command is executed

Option	Description
`-C`	Uses the list of servers in `/opt/oracle/bda/cluster-rack-infiniband` as the target. See Identifying the Target Servers.
`-d destfile`	Specifies a target directory or file name for the `-f` option
`-f file`	Specifies files to be copied to the user's home directory on the target nodes. The files are not executed. See the `-l` option.
`-g groupfile`	Specifies a file containing a list of Oracle Big Data Cloud Machine nodes where the command is executed. You can use either server names or IP addresses in the file.
`-h, --help`	Displays a description of the commands
`-k`	Pushes the `ssh` key to each node's `/root/.ssh/authorized_keys` file. See Setting Up Passwordless SSH for an easier alternative.
`-l userid`	Identifies the user ID for logging in to another server. The default ID is `root`.
`--maxlines=maxlines`	Identifies the maximum lines of output displayed from a command executed on multiple servers. The default is 10,000 lines.
`-n`	Abbreviates the output for non-error messages. Only the server name is displayed when a server returns normal output (return code `0`). You cannot use the `-n` and `-r` options together.
`-r regexp`	Replaces the output with the server name for lines that match the specified regular expression
`-s sshoptions`	Specifies a string of options that are passed to SSH
`--scp=scpoptions`	Specifies a string of options that are passed to Secure Copy (SCP), when these options are different from `sshoptions`
`--serial`	Serializes execution over the servers. The default is parallel execution.
`-t`	Lists the target servers

Option	Description
--unkey	Drops the keys from the authorized_key files of the target servers
-v	Displays the verbose version of all messages
--version	Displays the dcli version number
--vmstat=VMSTATOPS	Displays the syntax of the Linux Virtual Memory Statistics utility (vmstat). This command returns process, virtual memory, disk, trap, and CPU activity information.
	To enter a vmstat command, enclose its options in quotation marks. For example:
	`--vmstat="-a 3 5"`
	See your Linux documentation for more information about vmstat.
-x execfile	Specifies a command file to be copied to the user's home directory and executed on the target servers. See the -l option.

dcli Return Values

- 0: The command ran successfully on all servers.

- 1: One or more servers were inaccessible or remote execution returned a nonzero value. A message lists the unresponsive servers. Execution continues on the other servers.

- 2: A local error prevented the command from executing.

If you interrupt the local dcli process, then the remote commands may continue without returning their output or status.

dcli Examples

Following are examples of the dcli utility.

This example returns the default list of target servers:

```
# dcli -t
Target nodes: ['bda1node01-adm.example.com', 'bda1node02-adm.example.com',
'bda1node03-adm.example.com', 'bda1node04-adm.example.com', 'bda1node05-
adm.example.com', 'bda1node06-adm.example.com', 'bda1node07-adm.example.com',
'bda1node08-adm.example.com', 'bda1node09-adm.example.com']
```

The next example checks the temperature of all servers:

```
# dcli 'ipmitool sunoem cli "show /SYS/T_AMB" | grep value'

bda1node01-adm.example.com: value = 22.000 degree C
bda1node02-adm.example.com: value = 22.000 degree C
bda1node03-adm.example.com: value = 22.000 degree C
bda1node04-adm.example.com: value = 23.000 degree C
```

.
.
.

A

bda-oss-admin Command Reference

Use the Oracle Big Data Cloud Machine command line utility `bda-oss-admin` to manage users and resources of your cluster.

To issue `bda-oss-admin` commands, you must connect to a node as the `opc` user and then use the `sudo` command to switch to the `root` user. See Connecting to a Cluster Node Through Secure Shell (SSH).

Syntax

```
bda-oss-admin [options] subcommand [arguments]...
```

Options

Option	Description
`--version`	Show the `bda-oss-admin` version
`--cm-admin` *user_name*	Cloudera Manager administrator user name
`--cm-passwd` *password*	Cloudera Manager administrator password. The command will prompt for the password if it is not provided.
`--b64-cm-passwd` *password*	The Cloudera Manager password is Base64 encoded. It will be decoded before uploading.
`--cm-url` *url*	Cloudera Manager URL; for example, `https://servername.bigdata.oraclecloud.com:7183`
`-b` `--b64-encoded-pwds`	The password is Base64 encoded. It will be decoded before uploading.
`-h` `--help`	Show help for this command.

When you specify any of the above options on the command line, the options must be placed immediately after the `bda-oss-admin` command and before any of its subcommands. For example, this command is legal:

```
# bda-oss-admin --cm-passwd Welcome_123 list_swift_creds
```

However, the following command is not legal, because the `--cm-passwd` option is placed after the `list_swift_creds` subcommand:

```
# bda-oss-admin list_swift_creds --cm-passwd Welcome_123
```

Environment Variables

Instead of setting Cloudera Manager credentials as options on the command line, you can set them as environment variables.

This environment variable...	Is equivalent to this option...
CM_ADMIN	--cm-admin
CM_PASSWORD	--cm-passwd
CM_URL	--cm-url

Subcommands

Command	Task
add_bdcs_cp_extensions_mr	Add Oracle Big Data Cloud Machine `classpath` extensions to the MapReduce configuration.
print_yarn_mapred_cp	Display the YARN MapReduce default classpath.
remove_bdcs_cp_extensions_mr	Remove Oracle Big Data Cloud Machine `classpath` extensions from the MapReduce configuration.
restart_cluster	Restart the cluster (only stale services).

bda-oss-admin add_bdcs_cp_extensions_mr

Adds the Oracle Big Data Cloud Machine `classpath` extensions to the MapReduce configuration file, `mapred-site.xml`.

Syntax

```
bda-oss-admin add_bdcs_cp_extensions_mr [options]
```

Options

Option	Description
-h --help	Show help for this command

Example

```
# bda-oss-admin add_bdcs_cp_extensions_mr
Changes will not affect the cluster until services get restarted. See the
restart_cluster command
```

bda-oss-admin print_yarn_mapred_cp

Display the YARN MapReduce default classpath.

Syntax

```
bda-oss-admin print_yarn_mapred_cp [options]
```

Options

Option	Description
-h --help	Show help for this command.

Example

bda-oss-admin print_yarn_mapred_cp

```
YARN Mapreduce Conf mapreduce_application_classpath:
$HADOOP_MAPRED_HOME/*,$HADOOP_MAPRED_HOME/lib/*,
$MR2_CLASSPATH,/opt/oracle/orabalancer-2.4.0-h2/jlib/
orabalancer-2.4.0.jar,/opt/oracle/orabalancer-2.4.0-h2/jlib/
commons-math-2.2.jar,/opt/oracle/bda/bdcs/bdcs-rest-api-app/
current/lib-hadoop/*
```

bda-oss-admin remove_bdcs_cp_extensions_mr

Removes the Oracle Big Data Cloud Machine `classpath` extensions from the MapReduce configuration file, `mapred-site.xml`.

Syntax

```
bda-oss-admin remove_bdcs_cp_extensions_mr [options]
```

Options

Option	Description
-h --help	Show help for this command

Example

bda-oss-admin remove_bdcs_cp_extensions_mr
```
Changes will not affect the cluster until services get restarted. See the
restart_cluster command
```

bda-oss-admin restart_cluster

Restarts the Hadoop Cluster by interacting with Cloudera Manager.

Syntax

```
bda-oss-admin restart_cluster [options]
```

Options

Option	Description
-h --help	Show help for this command.

Example

```
# bda-oss-admin restart_cluster
Restarting the cluster...
*********************
Cluster restarted successfully
```

B

Pages for Administering Oracle Big Data Cloud Machine

This section provides information about what you can do and what you see on each of the Oracle Cloud pages for administering Oracle Big Data Cloud Machine.

Topics

- Exploring the Oracle Big Data Cloud Service Console Clusters Page

- Exploring the Oracle Big Data Cloud Service Cluster Details Page

Exploring the Oracle Big Data Cloud Service Console Clusters Page

The Oracle Big Data Cloud Service console lists all the clusters in the current identity domain.

What You See On the Page

The following table describes each item on the Oracle Big Data Cloud Service console page.

> **Note:** Depending on your role, you may not have access to all the options on this page. Only a user with Administrator privileges has access to all.

Element	Description
Clusters	The **Clusters** link displays the current page, the Oracle Big Data Cloud Service console.
Dashboard	Click the **Dashboard** link to displays your My Services Dashboard, where you can check the overall status of your active services for the current data center and identity domain.
Users	Click the **Users** link to display the My Services Users page, where you can manage users and perform access control.
Notifications	Click the **Notifications** link to display the My Services Notifications page, where you can monitor any notifications from the system, including service outages and planned service outages (for regular maintenance).
Started Clusters	Total number of clusters started, out of the total number in the service.
Nodes	Total number of nodes in all running clusters, out of the total number allocated for the service.

Element	Description
Memory	Total amount of memory allocated across all clusters in the service.
HDFS Storage	Total amount of Hadoop Distributed File System (HDFS) storage allocated across all clusters in the service.
Create Cluster	Click to create a new cluster. See Creating a Cluster.
Clusters	All the clusters in the account, including running, stopped, and failed clusters. .
Cluster name *Status*	Each box shows information about a single cluster. The icon changes, depending on the status of the cluster, and the status is displayed below the cluster name.
☰ (for the cluster)	Click the icon to open a menu with these options: • **View Details**—View details of the cluster on the Oracle Big Data Cloud Cluster Overview page. See Exploring the Oracle Big Data Cloud Service Cluster Details Page. • **Open Cloudera Manager**—Open the Cloudera Manager application to manage the Hadoop cluster. • **Edit**—Open the **Edit** *cluster name* dialog to edit the description and to update the SSH public key for the cluster. See Updating the SSH Public Key for a Cluster. • **Restart Cluster VMs**—Restart all virtual machines (VMs) hosting the nodes of the cluster. • **Delete Instance**—Delete the cluster.
Version	Version of Oracle Big Data Cloud Machine software running on this cluster.
No of Nodes	Number of nodes allocated to this cluster.
Secure Setup	Indicates whether or not security, Kerberos, is enabled for this cluster.
Applications	Number of applications installed on the cluster (not including CDH software); for example, Oracle Big Data Discovery Cloud Service.

Exploring the Oracle Big Data Cloud Service Cluster Details Page

The Oracle Big Data Cloud Service Cluster Overview page displays details about the cluster.

What You See on the Page

This page includes common information and controls that are always visible plus four tabs:

• Common Items

- Overview Tab

- Applications Tab

- Available Nodes Tab

- Requests Tab

> **Note:** Depending on your role, you may not have access to all the options on this page. Only a user with Administrator privileges has access to all.

Common Items

The items always visible on this page are described in the following table.

Item	Description
Clusters	Click the **Clusters** link to open the Oracle Big Data Cloud Service console, where you can see overview information about all of your clusters. Click any cluster name to see details about that cluster.
Dashboard	
Users	Click the **Users** icon to open the My Services Users page, where you can manage users and perform access control.
Notifications	Click the **Notifications** icon to open the My Services Notifications page, where you can monitor any notifications from the system, including service outages and planned service outages (for regular maintenance).
☰ (for the cluster)	Click the icon to open a menu with these options: • **Open Cloudera Manager**—Open the Cloudera Manager application to manage the Hadoop cluster. • **Edit**—Open the **Edit** *cluster name* dialog to edit the description and to update the SSH public key for the cluster. See Updating the SSH Public Key for a Cluster. • **Restart Cluster VMs**—Restart all virtual machines (VMs) hosting the the nodes of the cluster. • **Extend**—Extend the cluster with the additional nodes allocated to you. To see details about these allocated nodes, see Available Nodes Tab. This command is available only if you've added extra nodes to your subscription but haven't yet added them to a cluster. • **Delete Cluster**—Delete the cluster.
Overview	Click to display the **Cluster Overview** tab (when it isn't currently displayed).
Applications	Click to display the **Applications** tab (when it isn't currently displayed).

Item	Description
Available nodes	Click to display the **Available Nodes** tab (when it isn't currently displayed). The **Available Nodes** tab is available only if you've added extra nodes to your subscription but haven't yet added them to a cluster.
Requests	Click to display the **Requests** tab (when it isn't currently displayed)..

Overview tab

The **Overview** tab shows details about all the nodes in the cluster, as described in the following table:

Item	Description
NodesS	Total number of nodes in the cluster.
OCPUs	Total number of Oracle CPUs (OCPUs) allocated for this cluster.
Memory	Total amount of memory allocated for this cluster.
Storage	Total amount of Hadoop Distributed File System (HDFS) storage allocated for this cluster.
Version	The version of the Oracle Big Data Cloud Machine software used in this cluster.
Status	The operational status of the cluster, including Creating, Updating, Restarting, Terminating. On-boarding, Ready, etc.
Cloudera Manager URL	The URL for accessing Cloudera Manager for the cluster. Click the link to open Cloudera Manager
Cloudera Manager Version	The version of Cloudera Manager on this cluster.
Description	A short description of the cluster.
Nodes	All the nodes in the cluster are listed below this header.
Restart selected node(s)	Click to restart any nodes selected in the Nodes list on this page. You can select all the nodes by clicking **Select all** at the top of the list or individually by selecting the check box(es) for one or more nodes. This option is available only on a cluster with a status of Ready, Restart Failed, or Unknown

Item	Description
	Each row that begins with this icon shows details of one node in the cluster.
Hostname	The name of the host for this node.
IP Address	The IP address of this node.
Version	The version of the Big Data Cloud Service software installed on the host.
OCPUs	Total number of Oracle CPUs allocated for this node.
Memory	Total amount of memory allocated for this node.
HDFS Storage	Total amount of HDFS storage allocated for this node.
SSH Fingerprint	The fingerprint (short sequence of bytes) used to identify the Secure Shell (SSH) public key associated with this node.
≡ (for a node)	Click the icon to open a menu with this option: • **Restart Node**—Restart the node. This option is available only on a cluster with a status of Ready, Restart Failed, or Unknown.

Applications Tab

The **Applications** tab shows details about any applications that are installed in this cluster. Different applications may show different information and provide different commands that can be run.

Item	Description
Application Name	The name of the application, for example, **BDDCS** for Oracle Big Data Discovery.
Status	Status of the service, for example, enabled or disabled.
URL	The URL for accessing the application.

Available Nodes Tab

The **Available Nodes** tab shows details about nodes that have been allocated for this service but haven't yet been added to the cluster. The tab is available only if you have such nodes that are allocated but not yet added. Click the **Extend cluster with available node(s)** button to add them.

Item	Description
Available nodes	All of the nodes that have been allocated for this service are listed below this header.
Extend cluster with available node(s)	Click this button to extend this cluster with all of the nodes in the list.
OCPUs	Total number of Oracle CPUs allocated for this node.
Memory	Total amount of memory allocated for this node.
HDFS Storage	Total amount of HDFS storage allocated for this node.

Requests Tab

The **Requests** tab shows a list of requests made against the cluster, the status of the requests, plus some additional details, as described in the following table.

Item	Description
Filtering/ordering choices	Controls for filtering and ordering the list of requests. You can select from the following: • **Filter by request type** • **Filter by request status** • **Sort by** *ascending/descending start/end time*
Request type *status*	Each row in the table shows one request. The type of request is shown in the first column and the status of that request is shown beneath it. Request types include: • **Create cluster** • **Terminate cluster** • **Update cluster** • **Restart cluster** • **Extend cluster** Statuses include: • **Successful** • **Failed** • **Processing** For failed actions, click **Error Details** to see the error number and a short message.
Start Time	The date and time the action was initiated.
End Time	The date and time the action was completed. For actions in progress, this field is blank.

Item	Description
Request ID	An alphanumeric string that identifies the request. This ID may be useful when troubleshooting problems with the request.