Hadoop Cluster Configuration Settings

This chapter contains the following sections:

Creating an External Database Configuration

You can deploy each Hadoop cluster with its own external database for all Hadoop distributions (Cloudera, MapR, and Hortonworks) using instant Hadoop cluster and customized Hadoop cluster creation actions.

You can configure a new database or use an existing database in Cisco UCS Director Express for Big Data. The Oozie, Hive, and Hue services use configured database information that you have created using the **Create External Database Configurations** dialog.

Step 1 On the menu bar, choose **Solutions** > **Big Data** > **Settings**.

Step 2 Click the **External Database** tab.

Step 3 Click **Add**.
You can modify or delete any database you have previously created from the external database table.

Step 4 In the **Create External Database Configurations** dialog, complete the following fields:

Name	Description
Database Name field	Enter a unique name for the database type you want to create.
Database Type field	Choose the database type from the list.
Server Name field	Enter an IPv4 address for the database server.
Port field	Enter a port number based on the database type.
User Name field	Enter a username to access the database server.
Password field	Enter the password to access the database server.
Confirm Password field	Confirm the password to access the database server.

Step 5 Click **Submit.**

What to Do Next

Deploy Hadoop clusters through instant Hadoop cluster and customized Hadoop cluster creation actions.

Default Databases Used in Hadoop Distribution Services

Default Databases for Cloudera (Service Names):

- Cloudera Manager—mysql

- Oozie—mysql

- Hive—mysql

- Hue—mysql

Default Databases for MapR (Service Names):

- Oozie—Derby

- Hive—mysql

- Hue—SQlite

Default Databases for Hortonworks (Service Names):

- Ambari—PostGres

- Oozie—Derby

- Hive—mysql

Creating a Hadoop Cluster Configuration Parameters Template

You can create the Hadoop Cluster Configuration Parameters Template only from the **Hadoop Config Parameters** tab on the menu bar here: **Solutions** > **Big Data** > **Settings** before triggering a Hadoop cluster. You can select the Hadoop cluster configuration parameters template to edit, clone, or delete.

Step 1 On the menu bar, choose **Solutions** > **Big Data** > **Settings**.

Step 2 Click the **Hadoop Config Parameters** tab.

Step 3 Click **Add**.

Step 4 On the **Hadoop Config Parameters** page of the **Create Hadoop Cluster Configuration Parameters Template** wizard, complete the following fields:

Name	Description
Template Name field	A unique name for the Hadoop cluster configuration parameter template.
Template Description field	The description for the Hadoop cluster configuration parameter template.
Hadoop Distribution drop-down list	Choose the Hadoop distribution.
Hadoop Distribution Version drop-down list	Choose the Hadoop distribution version.

Step 5   Click **Next**.

Step 6   On the **Hadoop Config Parameters - HDFS Service** page of the **Create Hadoop Cluster Configuration Parameters Template** wizard, specify the Hadoop cluster HDFS service parameter name, value, and the minimum supported Hadoop distribution.

Step 7   On the **Hadoop Config Parameters - YARN Service** page of the **Create Hadoop Cluster Configuration Parameters Template** wizard, configure the parameters.

Step 8   On the **Hadoop Config Parameters - HBase Service** page of the **Create Hadoop Cluster Configuration Parameters Template** wizard, configure the parameters.

Step 9   On the **Hadoop Config Parameters - MapReduce Service** page of the **Create Hadoop Cluster Configuration Parameters Template** wizard, configure the parameters.

Step 10   On the **Hadoop Config Parameters - Miscellaneous Parameters** page of the **Create Hadoop Cluster Configuration Parameters Template** wizard, configure the (ServiceLevel and RoleLevel) parameters.

Step 11   Click **Submit**.

Updating Hadoop Cluster Configuration Parameters Template - Post Hadoop Cluster Creation

Step 1   On the menu bar, choose **Solutions** > **Big Data** > **Accounts**.

Step 2   Click the **Big Data Accounts** tab and choose an existing Big Data Account.

Step 3   Click **Configure Cluster**.

Step 4   On the **Hadoop Config Parameters** page of the **Update Hadoop Cluster Configuration Parameters Template** wizard, choose the Hadoop distribution.

Step 5   Click **Next**.

Step 6   On the **Hadoop Config Parameters - HDFS Service** page of the **Update Hadoop Cluster Configuration Parameters Template** wizard, specify the Hadoop cluster HDFS service parameter name, value, and the minimum supported Hadoop distribution version, if any.

Step 7   On the **Hadoop Config Parameters - YARN Service** page of the **Update Hadoop Cluster Configuration Parameters Template** wizard, update the parameters as required.

Step 8   On the **Hadoop Config Parameters - HBase Service** page of the **Update Hadoop Cluster Configuration Parameters Template** wizard, update the parameters as required.

Step 9   On the **Hadoop Config Parameters - MapReduce Service** page of the **Update Hadoop Cluster Configuration Parameters Template** wizard, update the parameters as required.

Step 10   On the **Hadoop Config Parameters - Miscellaneous Parameters** page of the **Update Hadoop Cluster Configuration Parameters Template** wizard, update the (ServiceLevel and RoleLevel) parameters as required.

Step 11   Click **Submit**.

Quality of Service System Classes

For more information on Quality of Service and System Classes, see QoS System Classes.

Quality of Service

Cisco Unified Computing System provides the following methods to implement quality of service (QoS):

- System classes that specify the global configuration for certain types of traffic across the entire system.
- QoS policies that assign system classes for individual vNICs.
- Flow control policies that determine how uplink Ethernet ports handle pause frames.

System Classes

Cisco UCS uses Data Center Ethernet (DCE) to handle all traffic inside a Cisco UCS domain. This industry-standard enhancement to Ethernet divides the bandwidth of the Ethernet pipe into eight virtual lanes. Two virtual lanes are reserved for internal system use and management traffic. You can configure quality of service (QoS) for the other six virtual lanes. System classes determine how the DCE bandwidth in these six virtual lanes is allocated across the entire Cisco UCS domain.

Each system class reserves a specific segment of the bandwidth for a specific type of traffic, which provides a level of traffic management, even in an oversubscribed system. For example, you can configure the Fibre Channel Priority system class to determine the percentage of DCE bandwidth allocated to FCoE traffic.

The following table describes the system classes that you can configure:

System Class	Description
Best Effort	A system class that sets the quality of service for the lane reserved for basic Ethernet traffic. Some properties of this system class are preset and cannot be modified. For example, this class has a drop policy that allows it to drop data packets if necessary. You cannot disable this system class.
• Platinum • Gold • Silver • Bronze	A configurable set of system classes that you can include in the QoS policy for a service profile. Each system class manages one lane of traffic. All properties of these system classes are available for you to assign custom settings and policies.

System Class	Description
Fibre Channel	A system class that sets the quality of service for the lane reserved for Fibre Channel over Ethernet traffic. Some properties of this system class are preset and cannot be modified.
	For example, this class has a no-drop policy that ensures it never drops data packets. You cannot disable this system class.
	Note FCoE traffic has a reserved QoS system class that cannot be used by any other type of traffic. If any other type of traffic has a CoS value that is used by FCoE, the value is re-marked to 0

Editing QoS System Classes

For more information on Quality of Service and System Classes, see QoS System Classes.

Step 1   On the menu bar, choose **Solutions** > **Big Data** > **Settings**.

Step 2   Click the **QoS System Class** tab.

Step 3   Choose the QoS System Class (by Priority) that you want to edit and click **Edit**.

- Best Effort

- Platinum

- Gold

- Silver

- Bronze

Step 4   In the **Modify QoS System Class** dialog box, complete the following fields:

Name	Description
Enabled check box	If checked, the associated QoS class is configured on the fabric interconnect and can be assigned to a QoS policy.
	If unchecked, the class is not configured on the fabric interconnect. Any QoS policies associated with this class default to Best Effort or, if a system class is configured with a CoS of 0, to the CoS 0 system class.
	This check box is checked for Best Effort and Fibre Channel.

Name	Description
CoS drop-down list	The class of service. You can enter an integer value between 0 and 6, with 0 being the lowest priority and 6 being the highest priority. We recommend that you do not set the value to 0, unless you want that system class to be the default system class for traffic if the QoS policy is deleted or the assigned system class is disabled. This field is set to 7 for internal traffic and to any for Best effort. Both of these values are reserved and cannot be assigned to any other priority.
Packet Drop check box	This field is unchecked for the Fibre Channel class, which never allows dropped packets, and is checked for Best Effort, which always allows dropped packets.
Weight drop-down list	A choice may be one of the following: • An integer between 1 and 10. If you select an integer, Cisco UCS determines the percentage of network bandwidth assigned to the priority level as described in the Weight (%) field. • Best-effort. • None.
Muticast Optimized check box	If checked, the class is optimized to send packets to multiple destinations simultaneously. This option is not applicable to the Fibre Channel.
MTU drop-down list	The maximum transmission unit for the channel. This can be one of the following: • An integer between 1500 and 9216. This value corresponds to the maximum packet size. • fc—A predefined packet size of 2240. • Normal—A predefined packet size of 1500. • Specify Manually—A packet size between 1500 to 9216. This field is always set to fc for Fibre Channel.

Step 5   Click **Submit**.

Pre Cluster Performance Testing Settings

You can analyze memory, network, and disk metrics. A default Big Data Metrics Report provides the statistics collected for each host before creating any Hadoop cluster.

Step 1 On the menu bar, choose **Solutions** > **Big Data** > **Settings**.

Step 2 Click the **Management** tab.

Step 3 In the **Pre Cluster Performance Tests** section, check the check boxes for the following:

- Memory Test

- Network Test

- Disk Test

Note By default, the check boxes to run the memory, network, and the disk tests are unchecked. If you enable the precluster disk test, it impacts Hadoop cluster creation.

Step 4 Click **Submit.**

Approving Hadoop Cluster and Splunk Deployment Workflows

Before You Begin

Choose **Administration** > **Users and Groups** > **Users**, and add users with the following user roles:

- Network Admin (system default user role)

- Computing Admin (system default user role)

- Big Data User

Step 1 On the menu bar, choose **Solutions** > **Big Data** > **Settings**.

Step 2 Click the **Management** tab.

Step 3 Check the **Require OS User Approval** check box.

a) From the **User ID** table, check the **Login Name** of the user against the Network Admin user role.

b) Enter the **Number of Approval Request Reminders**.

Note Set the number of approval request reminders to zero if the reminder email has to be sent at a specified interval until the Network Admin approves or rejects the request.

c) Enter the **Reminder Interval(s)** in hours.

Note Check the **Approval required from all the users** check box, if you want all users to approve or reject the request.

Step 4 Check the **Require Compute User Approval** check box.

a) From the **User ID** table, select the **Login Name** of the user against the Computing Admin user role.

b) Enter the **Number of Approval Request Reminders**.

    Note   Set the number of approval request reminders to zero if the reminder email has to be sent at a specified interval until the Computing Admin approves or rejects the request.

c) Enter the **Reminder Interval(s)** in hours.

    Note   Check the **Approval required from all the users** check box, if you want the users to approve or reject the request.

Step 5   Check the **Require Accounts User Approval** check box.

a) From the **User ID** table, select the **Login Name** of the user against the Hadoop User role.

b) Enter the **Number of Approval Request Reminders**.

    Note   Set the number of approval request reminders to zero if the reminder email has to be sent at a specified interval until the Hadoop User approves or rejects the request.

c) Enter the **Reminder Interval(s)** in hours.

    Note   Check the **Approval required from all the users** check box, if you want the users to approve or reject the request.

Step 6   Click **Submit**.

What to Do Next

Verify whether users of Network Admin, Computing Admin, and Big Data Accounts User roles have approved the request before deploying any Big Data software.

Uploading Required OS and Big Data Software to Cisco UCS Director Baremetal Agent

You can upload (add) required RHEL 6.x ISO files, Big Data software and common software, and Oracle JDKs to Cisco UCS Director Baremetal Agent. You can upload the required files from your local system or any remote system, and the files are first uploaded to Cisco UCS Director. You can click the **Submit** button in the **Create Software Catalogs** dialog box to move the required files to the target Cisco UCS Director Baremetal Agent.

Supported file formats:

- Linux OS—rhel-x.x.iso

- Big Data software—xxxx-x.y.z.zip (.gz or .tgz or .tar) For example, MapR-x.y.z.zip, splunk-x.y.z.zip

- Common software—bd-sw-rep.zip (.gz or .tgz or .tar)

- JDK software—(x.y.z.rpm or .gz)

The Software Catalogs page displays Big Data and other required software in the Cisco UCS Director Baremetal Agent.

> **Tip** If the required software column is empty for Big Data, then Cisco UCS Director Baremetal Agent already contains all the files required.

Step 1 On the menu bar, choose **Solutions** > **Big Data** > **Settings**.

Step 2 Click the **Software Catalogs** tab.

Step 3 Click **Add**.

Step 4 Click **Upload** to upload files from your local system.

> **Note** Create a folder to include all the required files for the Big Data software, and compress the folders before uploading in the format specified.

Step 5 Choose the target Cisco UCS Director Baremetal Agent from the **Target BMA** drop-down list.

Step 6 Check the **Restart BMA Services** to restart Cisco UCS Director Baremetal Agent after uploading the required files.

> **Note** Refresh the **Software Catalogs** page after 5 to 10 minutes to see new and modified catalogs.

Name	Description
Linux OS Upload	
Catalog Name field	Operating System Name (for example, RHEL.6.5)
Upload Type drop-down list	Choose one of the following: • Desktop file • The web server path that is reachable by the Cisco UCS Director Baremetal Agent • Mountpoint in Cisco UCS Director Baremetal Agent (For example, /root/iso) • Path to ISO in Cisco UCS Director Baremetal Agent (For example, /temp/rhel65/iso)
Big Data Software Upload	
Catalog Name field	Big Data software. For example, Hadoop Distribution (for example, distribution_name-x.y.z) or splunk enterprise software (splunk-x.y.z).
Upload Type drop-down list	Choose one of the following: • Desktop file • The web server path that is reachable by the Cisco UCS Director Baremetal Agent for uploading remote software to Baremetal Agent
Common Software Upload	

Name	Description
Upload Type drop-down list	Choose one of the following: • Desktop file • The web server path that is reachable by the Cisco UCS Director Baremetal Agent for uploading remote software to Baremetal Agent
JDK Upload	
JDK Version field	JDK version. For example, JDK 1.x
Upload Type drop-down list	Choose one of the following: • Desktop file • The web server path that is reachable by the Cisco UCS Director Baremetal Agent to upload remote software to Baremetal Agent

Linux OS Upload	
Catalog Name field	Operating System Name (for example, RHEL.6.5)
Upload Type drop-down list	Choose one of the following: • Desktop file • The web server path that is reachable by the Cisco UCS Director Baremetal Agent • Mountpoint in Cisco UCS Director Baremetal Agent (For example, /root/iso) • Path to ISO in Cisco UCS Director Baremetal Agent(For example, /temp/rhel65/iso)
Big Data Software Upload	
Catalog Name field	Big Data software. For example, Hadoop Distribution (for example, distribution_name-x.y.z) or splunk enterprise software (splunk-x.y.z)

Upload Type drop-down list	Choose one of the following: • Desktop file • The web server path that is reachable by the Cisco UCS Director Baremetal Agent to upload remote software to Baremetal Agent
Common Software Upload	
Upload Type drop-down list	Choose one of the following: • Desktop file • The web server path that is reachable by the Cisco UCS Director Baremetal Agent to upload remote software to Baremetal Agent
JDK Upload	
JDK Version field	JDK version. For example, JDK 1.x
Upload Type drop-down list	Choose one of the following: • Desktop file • The web server path that is reachable by the Cisco UCS Director Baremetal Agent to upload remote software to Baremetal Agent

Step 7   Click **Submit**.

What to Do Next

You can track software uploads here: **Administration** > **Integration**. Click the **Change Record** tab to track the software upload in progress and verify its status.

Supported Oracle JDK Software Versions

This section lists the supported Oracle JDK software versions:

Supported Oracle JDK Versions for Cloudera and Upgrade Scenarios

Hadoop Distribution	**Oracle JDK 1.6**	**Oracle JDK 1.7**	**Oracle JDK 1.8**
Cloudera Manager 5.4.x	—	JDK 1.7.0_25	JDK 1.8.0_60

Hadoop Distribution	Oracle JDK 1.6	Oracle JDK 1.7	Oracle JDK 1.8
Cloudera Manager 5.3.x	—	—	JDK 1.8.0_11

Hadoop Distribution Version to Upgrade	Supported Upgrade Version
Cloudera Manager 5.3.x, JDK 1.7.0_25	Cloudera Manager 5.4.x, JDK 1.8

Supported Oracle JDK Versions for MapR

Hadoop Distribution	JDK 1.7	JDK 1.8
MapR 5.0.0	Yes	Yes
MapR 4.1.0	No	Yes

Hadoop Distribution Version to Upgrade	Supported Upgrade Version
MapR 4.1.0, JDK 1.7	MapR 5.0.0, JDK 1.8

Supported Oracle JDK Software Versions for Hortonworks and Upgrade Scenarios

Hadoop Distribution	JDK 1.7	JDK 1.8
Hortonworks 2.3	—	Yes
Hortonworks 2.2	Yes	Yes

Hadoop Distribution Version to Upgrade	Supported Upgrade Version
Hortonworks 2.2, JDK 1.7	Hortonworks 2.3, JDK 1.8

Cloudera, MapR, and Hortonworks RPMs on Cisco UCS Director Express for Big Data Baremetal Agent

Common Packages for Cloudera, MapR, and Hortonworks

Note   For any Hadoop software that is not available, update the
`/opt/cnsaroot/bigdata_templates/common_templates/HadoopDistributionRPM.txt`
file with an appropriate file from the online repository of the vendor.

Note   We recommend that you verify the supported versions from the Hadoop Vendor Support Documentation.

Download the following common packages to `/opt/cnsaroot/bd-sw-rep/`:

- pssh-2.3.1.tar.gz from https://pypi.python.org/packages/source/p/pssh
- clustershell-1.6-1.el6.noarch.rpm from http://dl.fedoraproject.org/pub/epel/6/x86_64

Common Packages for Cloudera

Download the following packages to `/opt/cnsaroot/bd-sw-rep/cloudera-5.X.X`:

- ClouderaEnterpriseLicense.lic—Get the license keys from Cloudera
- userrpmlist.txt—For more packages lists
- catalog.properties—Provides the label name for the Cloudera version (x represents the Cloudera version on the Cisco UCS Director Express for Big Data Baremetal Agent)
- mysql-connector-java-5.1.26.tar.gz from http://cdn.mysql.com/archives/mysql-connector-java-5.1

Cloudera 5.0.1 Packages and Parcels

Download the following packages to `/opt/cnsaroot/bd-sw-rep/cloudera-5.0.1`:

- cm5.0.1-centos6.tar.gz from http://archive.cloudera.com/cm5/repo-as-tarball/5.0.1
- CDH-5.0.1-1.cdh5.0.1.p0.47-el6.parcel from http://archive.cloudera.com/cdh5/parcels/5.0.1
- CDH-5.0.1-1.cdh5.0.1.p0.47-el6.parcel.sha1 from http://archive.cloudera.com/cdh5/parcels/5.0.1
- manifest.json from http://archive.cloudera.com/cdh5/parcels/5.0.1

Cloudera 5.0.6 Packages and Parcels

Download the following packages to `/opt/cnsaroot/bd-sw-rep/cloudera-5.0.6`:

- cm5.0.6-centos6.tar.gz from http://archive.cloudera.com/cm5/repo-as-tarball/5.0.6
- CDH-5.0.6-1.cdh5.0.6.p0.7-el6.parcel from http://archive.cloudera.com/cdh5/parcels/5.0.6

- CDH-5.0.6-1.cdh5.0.6.p0.7-el6.parcel.sha1 from http://archive.cloudera.com/cdh5/parcels/5.0.6

- manifest.json from http://archive.cloudera.com/cdh5/parcels/5.0.6

Cloudera 5.2.0 Packages and Parcels

Download the following packages to `/opt/cnsaroot/bd-sw-rep/cloudera-5.2.0`:

- cm5.2.0-centos6.tar.gz from http://archive.cloudera.com/cm5/repo-as-tarball/5.2.0

- CDH-5.2.0-1.cdh5.2.0.p0.36-el6.parcel from http://archive.cloudera.com/cdh5/parcels/5.2.0

- CDH-5.2.0-1.cdh5.2.0.p0.36-el6.parcel.sha1 from http://archive.cloudera.com/cdh5/parcels/5.2.0

- manifest.json from http://archive.cloudera.com/cdh5/parcels/5.2.0

Cloudera 5.2.1 Packages and Parcels

Download the following packages to `/opt/cnsaroot/bd-sw-rep/cloudera-5.2.1`:

- cm5.2.1-centos6.tar.gz from http://archive.cloudera.com/cm5/repo-as-tarball/5.2.1

- CDH-5.2.1-1.cdh5.2.1.p0.12-el6.parcel from http://archive.cloudera.com/cdh5/parcels/5.2.1

- CDH-5.2.1-1.cdh5.2.1.p0.12-el6.parcel.sha1 from http://archive.cloudera.com/cdh5/parcels/5.2.1

- manifest.json from http://archive.cloudera.com/cdh5/parcels/5.2.1

Cloudera 5.3.0 Packages and Parcels

Download the following packages to `/opt/cnsaroot/bd-sw-rep/cloudera-5.3.0`:

- cm5.3.0-centos6.tar.gz from http://archive.cloudera.com/cm5/repo-as-tarball/5.3.0

- CDH-5.3.0-1.cdh5.3.0.p0.30-el6.parcel from http://archive.cloudera.com/cdh5/parcels/5.3.0

- CDH-5.3.0-1.cdh5.3.0.p0.30-el6.parcel.sha1 from http://archive.cloudera.com/cdh5/parcels/5.3.0

- manifest.json from http://archive.cloudera.com/cdh5/parcels/5.3.0

Cloudera 5.4.1 Packages and Parcels

Download the following packages to `/opt/cnsaroot/bd-sw-rep/cloudera-5.4.1`:

- cm5.4.1-centos6.tar.gz from http://archive.cloudera.com/cm5/repo-as-tarball/5.4.1

- CDH-5.4.1-1.cdh5.4.1.p0.6-el6.parcel from http://archive.cloudera.com/cdh5/parcels/5.4.1

- CDH-5.4.1-1.cdh5.4.1.p0.6-el6.parcel.sha1 from http://archive.cloudera.com/cdh5/parcels/5.4.1

- manifest.json from http://archive.cloudera.com/cdh5/parcels/5.4.1

Common Packages for MapR 3.1.1, 4.0.1, and 4.0.2

Download the following common packages to `/opt/cnsaroot/bd-sw-rep/MapR-3.1.1 and MapR-4.0.x directories`:

- libgenders-devel-1.14-2.el6.rf.x86_64.rpm from http://pkgs.repoforge.org/libgenders/

- libgenders-1.14-2.el6.rf.x86_64.rpm from http://pkgs.repoforge.org/libgenders/

- ext-2.2.zip from `http://dev.sencha.com/deploy/ext-2.2.zip`

- sshpass-1.05-1.el6.x86_64.rpm from
  http://ftp.pbone.net/mirror/download.fedora.redhat.com/pub/fedora/epel/6/x86_64

- soci-mysql-3.2.1-1.el6.x86_64.rpm from http://ftp.is.co.za/mirror/fedora.redhat.com/epel/6/x86_64

- soci-3.2.1-1.el6.x86_64.rpm from http://ftp.is.co.za/mirror/fedora.redhat.com/epel/6/x86_64

- pdsh-2.27-1.el6.rf.x86_64.rpm from http://pkgs.repoforge.org/pdsh

- mapr-whirr-0.7.0.16780-1.noarch.rpm from http://archive.mapr.com/releases/ecosystem-all/redhat

- mapr-drill-0.7.0.29434-1.noarch.rpm from http://archive.mapr.com/releases/ecosystem/redhat

- catalog.properties—Provides the label name for the MapR version (x represents the MapR version on the Cisco UCS Director Express for Big Data Baremetal Agent)

- license.txt

Common Packages for MapR 4.1.0 and 5.0.0

Download the following common packages to `/opt/cnsaroot/bd-sw-rep/MapR-4.1.0 and MapR-5.0.0 directories`:

- libgenders-devel-1.14-2.el6.rf.x86_64.rpm from http://pkgs.repoforge.org/libgenders/

- libgenders-1.14-2.el6.rf.x86_64.rpm from http://pkgs.repoforge.org/libgenders/

- ext-2.2.zip from `http://dev.sencha.com/deploy/ext-2.2.zip`

- sshpass-1.05-1.el6.x86_64.rpm from
  http://ftp.pbone.net/mirror/download.fedora.redhat.com/pub/fedora/epel/6/x86_64

- soci-mysql-3.2.2-2.el6.x86_64.rpm from http://ftp.univie.ac.at/systems/linux/fedora/epel/6/x86_64

- soci-3.2.2-2.el6.x86_64.rpm from http://ftp.univie.ac.at/systems/linux/fedora/epel/6/x86_64

- pdsh-2.27-1.el6.rf.x86_64.rpm from http://pkgs.repoforge.org/pdsh

- mapr-whirr-0.8.1.18380-GA.noarch.rpm from http://archive.mapr.com/releases/ecosystem-all/redhat

- catalog.properties—Provides the label name for the MapR version (x represents the MapR version on the Cisco UCS Director Express for Big Data Baremetal Agent)

- license.txt

MapR 3.1.1 Packages

Download the following packages to /opt/cnsaroot/bd-sw-rep/MapR-3.1.1

- mapr-v3.1.1GA.rpm.tgz from http://package.mapr.com/releases/v3.1.1/redhat

- mapr-ecosystem-20140617.rpm.tgz

MapR 4.0.1 Packages

Download the following packages to /opt/cnsaroot/bd-sw-rep/MapR-4.0.1

- mapr-v4.0.1GA.rpm.tgz from http://package.mapr.com/releases/v4.0.1/redhat

- mapr-ecosystem-4.x-20141105.rpm.tgz

MapR 4.0.2 Packages

Download the following packages to /opt/cnsaroot/bd-sw-rep/MapR-4.0.2

- mapr-v4.0.2GA.rpm.tgz from http://package.mapr.com/releases/v4.0.2/redhat

- mapr-ecosystem-20150205.rpm.tgz

MapR 4.1.0 Packages

Download the following packages to /opt/cnsaroot/bd-sw-rep/MapR-4.1.0

- mapr-v4.1.0GA.rpm.tgz from http://package.mapr.com/releases/v4.1.0/redhat

- mapr-ecosystem-4.x-20150610.rpm.tgz from http://archive.mapr.com/releases/ecosystem-all/redhat

MapR 5.0.0 Packages

Download the following packages to /opt/cnsaroot/bd-sw-rep/MapR-5.0.0

- mapr-v5.0.0GA.rpm.tgz: from http://package.mapr.com/releases/v5.0.0/redhat

- mapr-ecosystem-5.x-20150709.rpm.tgz from http://archive.mapr.com/releases/ecosystem-all/redhat

Common Package for Hortonworks

Download the following common package to `/opt/cnsaroot/bd-sw-rep/Hortonworks-2.X:`

- openssl-1.0.1e-30.el6.x86_64.rpm

- catalog.properties—Provides the label name for the Hortonworks version (x represents the Hortonworks version on the Cisco UCS Director Express for Big Data Baremetal Agent)

Hortonworks 2.1 Packages

Download the following packages to `/opt/cnsaroot/bd-sw-rep/Hortonworks-2.1:`

- HDP-2.1.5.0-centos6-rpm.tar.gz from http://public-repo-1.hortonworks.com/HDP/centos6

- ambari-1.6.1-centos6.tar.gz from http://public-repo-1.hortonworks.com/ambari/centos6

- HDP-UTILS-1.1.0.17-centos6.tar.gz from
  http://public-repo-1.hortonworks.com/HDP-UTILS-1.1.0.17/repos/centos6

Hortonworks 2.2 Packages

Download the following packages to `/opt/cnsaroot/bd-sw-rep/Hortonworks-2.2:`

- HDP-2.2.0.0-centos6-rpm.tar.gz from http://public-repo-1.hortonworks.com/HDP/centos6

- ambari-1.7.0-centos6.tar.gz from http://public-repo-1.hortonworks.com/ambari/centos6

- HDP-UTILS-1.1.0.20-centos6.tar.gz from
  http://public-repo-1.hortonworks.com/HDP-UTILS-1.1.0.20/repos/centos6

Hortonworks 2.3 Packages

Download the following packages to `/opt/cnsaroot/bd-sw-rep/Hortonworks-2.3`:

- HDP-2.3.0.0-centos6-rpm.tar.gz from
  http://public-repo-1.hortonworks.com/HDP/centos6/2.x/updates/2.3.0.0

- ambari-2.1.1-centos6.tar.gz from http://public-repo-1.hortonworks.com/ambari/centos6/2.x/updates/2.1.1

- HDP-UTILS-1.1.0.20-centos6.tar.gz from
  http://public-repo-1.hortonworks.com/HDP-UTILS-1.1.0.20/repos/centos6

Cloudera and MapR RPMs for Upgrading Hadoop Cluster Distributions

Cloudera 5.3.0 Packages and Parcels

- cm5.3.0-centos6.tar.gz from http://archive.cloudera.com/cm5/repo-as-tarball/5.3.0

- CDH-5.3.0-1.cdh5.3.0.p0.30-el6.parcel from http://archive.cloudera.com/cdh5/parcels/5.3.0

- CDH-5.3.0-1.cdh5.3.0.p0.30-el6.parcel.sha1 from http://archive.cloudera.com/cdh5/parcels/5.3.0

- manifest.json from http://archive.cloudera.com/cdh5/parcels/5.3.0

Cloudera 5.4.1 Packages and Parcels

- cm5.4.1-centos6.tar.gz from http://archive.cloudera.com/cm5/repo-as-tarball/5.4.1

- CDH-5.4.1-1.cdh5.4.1.p0.6-el6.parcel from http://archive.cloudera.com/cdh5/parcels/5.4.1

- CDH-5.4.1-1.cdh5.4.1.p0.6-el6.parcel.sha1 from http://archive.cloudera.com/cdh5/parcels/5.4.1

- manifest.json from http://archive.cloudera.com/cdh5/parcels/5.4.1

MapR 4.1.0 Packages

- mapr-setup from http://package.mapr.com/releases/v4.1.0/redhat

- mapr-v4.1.0GA.rpm.tgz from http://package.mapr.com/releases/v4.1.0/redhat

- mysql-connector-java-5.1.26.tar.gz from http://cdn.mysql.com/archives/mysql-connector-java-5.1

MapR 5.0.0 Packages

- mapr-setup from http://package.mapr.com/releases/v5.0.0/redhat

- mapr-v5.0.0GA.rpm.tgz: from http://package.mapr.com/releases/v5.0.0/redhat

- mysql-connector-java-5.1.26.tar.gz from http://cdn.mysql.com/archives/mysql-connector-java-5.1

Configuration Check Rules

You can validate an existing cluster configuration by running a configuration check. The configuration check process involves comparing the current cluster configuration with reporting violations and configuration check rules.

Configuration check rules are predefined Cisco Validated Design (CVD) parameters for Hadoop clusters. Configuration check rules appear under **Solutions > Big Data > Settings**. After the configuration check is complete, violations appear in the **Faults** tab under **Solutions > Big Data > Accounts**. You can enable or disable configuration check rules at any time, but you cannot add new rules.

Configuration Check Rule	Description
Parameter	The predefined CVD parameter of the configuration.
Enabled	The state of the configuration check rule, either enabled (true) or disabled (false).
Expected value	The value expected for a parameter as defined in the Cisco Validated Design (CVD).
Description	The description of the parameter of the configuration.
Distribution	The Hadoop distribution.
Minimum Supported Distribution	The minimum supported version of Hadoop distribution.
Service	The Hadoop service.
Role	The Hadoop service role.
Type	The type of violation, either CVD or Inconsistent.
Fix Workflow	The reference to the workflow that can be triggered for fixing violations.

When the actual cluster configuration values differ from the expected values defined in the configuration check rules, then those configuration values are reported as violations. For example, CVD mandates that the NameNode heap size is 4 GB. But if the NameNode heap size in the cluster configuration is found to be 1 GB, then this is reported as a CVD violation. Also, inconsistent configuration parameters are reported. For example, NameNode heap size on both the primary and secondary nodes must be of the same size. If there is a mismatch in the size, then this parameter is reported as inconsistent.

Checking Hadoop Cluster Configuration

To validate the configuration of a cluster, do the following:

Step 1   On the menu bar, choose **Solutions** > **Big Data** > **Accounts**.

Step 2   Click the **Big Data Accounts** tab.

Step 3   Choose the account for which you want to run the configuration check and click **Check Configuration**.

Step 4   Click **Submit**.
A dialog box appears with the information that the configuration check is in progress.

Step 5   Click **OK**.
After the configuration check is complete, the violations appear under the **Faults** tab for the selected Big Data Account.

What to Do Next

Note   You can track configuration checks here: **Administration** > **Integration**. Click the **Change Record** tab to track the configuration checks in progress and verify if completed or failed.

Fixing Configuration Violations

After the configuration check is complete, the configuration violations appear in the Faults tab for the selected Big Data Account. You can either choose to fix these configuration violations manually on the Big Data Cluster Configuration page, or trigger a workflow. To trigger a workflow to fix the violation, create a workflow with the same name as the code specified in the violation.

To fix a configuration violation through a workflow, do the following:

Step 1   On the menu bar, choose **Solutions** > **Big Data** > **Accounts**.

Step 2   Click the **Faults** tab.

Step 3   Choose the configuration violation you want to fix and click **Trigger Workflow**.
If a workflow exists with the same name as the code specified in the violation, then the workflow is triggered.

Step 4   Enter the required inputs for the workflow and click **Submit**.
A service request ID is generated after you submit the inputs. You can check the status of the service request on the **Service Requests** page.