



Cisco UCS Director Express for Big Data Deployment and Management Guide, Release 3.5

First Published: 2017-07-11

Last Modified: 2017-10-12

Americas Headquarters

Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134-1706
USA
<http://www.cisco.com>
Tel: 408 526-4000
800 553-NETS (6387)
Fax: 408 527-0883

THE SPECIFICATIONS AND INFORMATION REGARDING THE PRODUCTS IN THIS MANUAL ARE SUBJECT TO CHANGE WITHOUT NOTICE. ALL STATEMENTS, INFORMATION, AND RECOMMENDATIONS IN THIS MANUAL ARE BELIEVED TO BE ACCURATE BUT ARE PRESENTED WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED. USERS MUST TAKE FULL RESPONSIBILITY FOR THEIR APPLICATION OF ANY PRODUCTS.

THE SOFTWARE LICENSE AND LIMITED WARRANTY FOR THE ACCOMPANYING PRODUCT ARE SET FORTH IN THE INFORMATION PACKET THAT SHIPPED WITH THE PRODUCT AND ARE INCORPORATED HEREIN BY THIS REFERENCE. IF YOU ARE UNABLE TO LOCATE THE SOFTWARE LICENSE OR LIMITED WARRANTY, CONTACT YOUR CISCO REPRESENTATIVE FOR A COPY.

The Cisco implementation of TCP header compression is an adaptation of a program developed by the University of California, Berkeley (UCB) as part of UCB's public domain version of the UNIX operating system. All rights reserved. Copyright © 1981, Regents of the University of California.

NOTWITHSTANDING ANY OTHER WARRANTY HEREIN, ALL DOCUMENT FILES AND SOFTWARE OF THESE SUPPLIERS ARE PROVIDED "AS IS" WITH ALL FAULTS. CISCO AND THE ABOVE-NAMED SUPPLIERS DISCLAIM ALL WARRANTIES, EXPRESSED OR IMPLIED, INCLUDING, WITHOUT LIMITATION, THOSE OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NON-INFRINGEMENT OR ARISING FROM A COURSE OF DEALING, USAGE, OR TRADE PRACTICE.

IN NO EVENT SHALL CISCO OR ITS SUPPLIERS BE LIABLE FOR ANY INDIRECT, SPECIAL, CONSEQUENTIAL, OR INCIDENTAL DAMAGES, INCLUDING, WITHOUT LIMITATION, LOST PROFITS OR LOSS OR DAMAGE TO DATA ARISING OUT OF THE USE OR INABILITY TO USE THIS MANUAL, EVEN IF CISCO OR ITS SUPPLIERS HAVE BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.

Any Internet Protocol (IP) addresses and phone numbers used in this document are not intended to be actual addresses and phone numbers. Any examples, command display output, network topology diagrams, and other figures included in the document are shown for illustrative purposes only. Any use of actual IP addresses or phone numbers in illustrative content is unintentional and coincidental.

Cisco and the Cisco logo are trademarks or registered trademarks of Cisco and/or its affiliates in the U.S. and other countries. To view a list of Cisco trademarks, go to this URL: <http://www.cisco.com/go/trademarks>. Third-party trademarks mentioned are the property of their respective owners. The use of the word partner does not imply a partnership relationship between Cisco and any other company. (1110R)

© 2017 Cisco Systems, Inc. All rights reserved.



CONTENTS

Preface

Preface ix

Audience ix

Conventions ix

Related Documentation xi

Documentation Feedback xi

Obtaining Documentation and Submitting a Service Request xi

CHAPTER 1

New and Changed Information for this Release 1

New and Changed Information in Release 3.5 1

CHAPTER 2

Overview 3

Cisco UCS Director Express for Big Data 4

Cisco UCS Integrated Infrastructure for Big Data 4

Managing Cisco UCS Director and Cisco UCS Director Express for Big Data Personalities 4

 Creating User Roles 5

 Hadoop Administrator Permissions 5

 Hadoop User Permissions 8

Installing Cisco UCS Director Express for Big Data on VMware vSphere 11

Installing Cisco UCS Director Express for Big Data Bare Metal Agent on VMware vSphere 13

NTP Server Configuration on Cisco UCS Director Express for Big Data Bare Metal Agent 14

Adding Oracle JDK Software Versions to Bare Metal Agent 15

Cisco Server Support for Hadoop Distribution Deployments 15

Cisco Server Support for Splunk Enterprise Deployments 17

Adding a New Red Hat Version for Hadoop Cluster 18

 New RHEL Version 19

Supported Hadoop Distributions 19

Supported Splunk Distribution 19

Supported Oracle JDK Software Versions	20
Supported Oracle JDK Versions for Cloudera and Upgrade Scenarios	20
Supported Oracle JDK Versions for MapR and Upgrade Scenarios	21
Supported Oracle JDK Software Versions for Hortonworks and Upgrade Scenarios	21
Upgrade Hadoop Distribution Software	22
Digitally Signed Images	23
Requirements for Verifying Digitally Signed Images	23
Verifying a Digitally Signed Image	24
Upgrade of Bare Metal Agent	25
Configuring Cisco UCS Manager Accounts	25
High-level Workflow to Create an Instant Hadoop Cluster	25
High-level Workflow to Create a Customized Hadoop Cluster	26

CHAPTER 3**Licenses for Cisco UCS Director Express for Big Data** 27

About Licenses	27
Fulfilling the Product Access Key	28
Updating the License	29
Standard License Features	29

CHAPTER 4**Managing Hadoop Accounts** 31

Adding a Hadoop Account	31
Running a Cluster Inventory for a Hadoop Account	33
Purging Big Data Cluster Account Details	33
Rolling Back a Hadoop Cluster for a Hadoop Account	34
Access to Hadoop Managers from Cisco UCS Director Express for Big Data	34

CHAPTER 5**Managing Splunk Accounts** 35

Cisco UCS Director Express for Big Data with Splunk Enterprise	35
Adding a Splunk Account	36
Running a Cluster Inventory for a Splunk Account	37
Rolling Back a Cluster for a Splunk Account	38
Access Splunk Enterprise Monitoring Console User Interface from Cisco UCS Director Express for Big Data	38

CHAPTER 6**Configuring Big Data IP Pools** 39

Big Data IP Pools	39
Adding a Big Data IP Pool	39
Managing Big Data IP Pools	40

CHAPTER 7**Configuring Cisco UCS Service Profile Templates for Big Data 43**

Cisco UCS Service Profile Templates for Big Data	43
Creating a Cisco UCS Service Profile Template for Big Data	44
Creating a QoS Policy	45
Creating a VLAN Policy	46
Creating a vNIC Policy	47
Creating a Boot Order Policy	50
Creating a BIOS Policy	50
Creating a Local Disk Configuration Policy	53
Editing RAID Policy	54
Editing RAID Policy for Splunk	56
Configuring Local Disk Partitions	58
Creating a Customized Service Profile Template	59
Cloning a Cisco UCS Service Profile Template	60

CHAPTER 8**Configuring and Deploying Hadoop Cluster Deployment Templates 61**

Hadoop Cluster Profile Templates	61
Creating a Hadoop Cluster Profile Template	63
Creating a Services Selection Policy	64
Configuring the Rack Assignment Policy	65
Configuring the HDFS Policy	66
Configuring the CLDB Policy	67
Configuring the YARN Policy	67
Configuring the ZooKeeper Policy	68
Configuring the Kafka Policy	68
Configuring the HBase Policy	69
Configuring the Hive Policy	69
Configuring the Oozie Policy	70
Configuring the Hue Policy	70
Configuring the Spark Policy	71
Configuring the Key-Value Store Indexer Policy	71

Configuring the Solr Policy	72
Configuring the Sqoop Policy	72
Configuring the Impala Policy	72
Configuring the Flume Policy	73
Configuring the PIG Policy	73
Configuring the MAHOUT Policy	74
Configuring the Falcon Policy	74
Configuring the Tez Policy	75
Configuring the Storm Policy	75
Configuring the Ganglia Policy	76
Cloning a Hadoop Cluster Profile Template	76
Creating a Cluster Deployment Template	76

CHAPTER 9**Managing Hadoop Clusters 79**

Creating an Instant Hadoop Cluster	79
Creating a Customized Hadoop Cluster	83
Creating a Hadoop Cluster Using Workflow	86
Provisioning an Instant and Customized Hadoop Cluster	87
Managing a Hadoop Cluster	89
View Hadoop Cluster Details	91
Viewing a Cluster Snapshot	93
Adding a New Hadoop Service	93
Managing Nodes in a Cluster	94
Delete Node and Delete Node to Bare Metal Actions in Cloudera and Hortonworks	95
Adding Managed Nodes to the Hadoop Cluster	96
Adding Live Nodes to the Hadoop Cluster	97
Adding Bare Metal Nodes to the Hadoop Cluster	97
Adding Disks to the Hadoop Cluster	100
Service Roles	101

CHAPTER 10**Managing Splunk Clusters 103**

Creating an Instant Splunk Cluster	103
Creating a Splunk Cluster Using Workflow	107
Customizing Splunk Cluster Creation	108
Adding Bare Metal Nodes to the Splunk Cluster	112

Deploying Splunk Cluster with Archival Node and NFS Support	115
Managing a Splunk Cluster	115

CHAPTER 11**Big Data Cluster Configuration Settings 119**

Creating an External Database Configuration	119
Default Databases Used in Hadoop Distribution Services	120
Creating a Hadoop Cluster Configuration Parameters Template	121
Updating Hadoop Cluster Configuration Parameters Template - Post Hadoop Cluster Creation	122
Quality of Service System Classes	123
Editing QoS System Classes	124
Pre Cluster Performance Testing Settings	126
Approving Hadoop Cluster and Splunk Deployment Workflows	126
Adding NTP Server Details	127
Uploading Required OS and Big Data Software to Cisco UCS Director Bare Metal Agent	128
Supported Oracle JDK Software Versions	130
Supported Oracle JDK Versions for Cloudera and Upgrade Scenarios	130
Supported Oracle JDK Versions for MapR and Upgrade Scenarios	131
Supported Oracle JDK Software Versions for Hortonworks and Upgrade Scenarios	132
Supported Oracle JDK Software Versions for BigInsights	132
Cloudera, MapR, and Hortonworks RPMs on Cisco UCS Director Express for Big Data Bare Metal Agent	132
Cloudera and MapR RPMs for Upgrading Hadoop Cluster Distributions	141
Installation of User-Defined Software Post Hadoop Cluster Creation	142
Configuration Check Rules	143
Checking Hadoop Cluster Configuration	144
Fixing Configuration Violations	144

CHAPTER 12**Cisco UCS CPA Workflows 145**

Workflows for Big Data	145
About Service Requests for Big Data	147
Monitoring Service Requests for Big Data	149
Viewing UCS CPA Workflow Tasks	149
Workflow Customization to Deploy a Hadoop or Splunk Cluster	152
Deploying a Hadoop or Splunk Cluster Through Workflow Customization	152
Assigning Big Data Accounts to User Groups	153

Unassigning Big Data Accounts 154

Cloning UCS CPA Workflows 154

CHAPTER 13**Monitoring and Reporting 159**

About Monitoring and Reporting 159

Cisco UCS Director Express for Big Data Dashboard 159

Viewing a Deployed Cluster Report 160

Reports 160

Cluster-specific Metrics Supported per Hadoop Distribution 161

Host-specific Metrics Supported per Hadoop Distribution 162

CHAPTER 14**Proactive Status Monitoring and Diagnostics 165**

Aggregate CPU, Disk, and Network Bandwidth Utilization 165

Monitoring Aggregate CPU, Disk, and Network Bandwidth Utilization 166

Monitoring Top Jobs Based on CPU Utilization and Time 166

Performance Metrics for CPU, Disk, and Network 167

Viewing CPU, Disk, and Network Statistics for a Hadoop Cluster 168

Analyzing Performance Bottlenecks Through Historical Metrics 169

Setting Alerts for Hadoop Cluster Service Failures 169

Types of Disk and Network Failure Alerts 170

Setting Alerts for Disk and Network Failures 172

Setting Disk Utilization Threshold Alerts 173



Preface

- [Audience, page ix](#)
- [Conventions, page ix](#)
- [Related Documentation, page xi](#)
- [Documentation Feedback, page xi](#)
- [Obtaining Documentation and Submitting a Service Request, page xi](#)

Audience

This guide is intended primarily for data center administrators who use Cisco UCS Director Express for Big Data and who have responsibilities and expertise in one or more of the following:

- Server administration
- Storage administration
- Network administration
- Network security
- Virtualization and virtual machines

Conventions

Text Type	Indication
GUI elements	GUI elements such as tab titles, area names, and field labels appear in this font . Main titles such as window, dialog box, and wizard titles appear in this font .
Document titles	Document titles appear in <i>this font</i> .
TUI elements	In a Text-based User Interface, text the system displays appears in <i>this font</i> .

Text Type	Indication
System output	Terminal sessions and information that the system displays appear in <i>this font</i> .
CLI commands	CLI command keywords appear in this font . Variables in a CLI command appear in <i>this font</i> .
[]	Elements in square brackets are optional.
{x y z}	Required alternative keywords are grouped in braces and separated by vertical bars.
[x y z]	Optional alternative keywords are grouped in brackets and separated by vertical bars.
string	A nonquoted set of characters. Do not use quotation marks around the string or the string will include the quotation marks.
<>	Nonprinting characters such as passwords are in angle brackets.
[]	Default responses to system prompts are in square brackets.
!, #	An exclamation point (!) or a pound sign (#) at the beginning of a line of code indicates a comment line.

**Note**

Means *reader take note*. Notes contain helpful suggestions or references to material not covered in the document.

**Caution**

Means *reader be careful*. In this situation, you might perform an action that could result in equipment damage or loss of data.

**Tip**

Means *the following information will help you solve a problem*. The tips information might not be troubleshooting or even an action, but could be useful information, similar to a Timesaver.

**Timesaver**

Means *the described action saves time*. You can save time by performing the action described in the paragraph.

**Warning****IMPORTANT SAFETY INSTRUCTIONS**

This warning symbol means danger. You are in a situation that could cause bodily injury. Before you work on any equipment, be aware of the hazards involved with electrical circuitry and be familiar with standard practices for preventing accidents. Use the statement number provided at the end of each warning to locate its translation in the translated safety warnings that accompanied this device.

SAVE THESE INSTRUCTIONS

Related Documentation

Cisco UCS Director Documentation Roadmap

For a complete list of Cisco UCS Director documentation, see the *Cisco UCS Director Documentation Roadmap* available at the following URL: http://www.cisco.com/en/US/docs/unified_computing/ucs/ucs-director/doc-roadmap/b_UCSDirectorDocRoadmap.html.

Cisco UCS Documentation Roadmaps

For a complete list of all B-Series documentation, see the *Cisco UCS B-Series Servers Documentation Roadmap* available at the following URL: <http://www.cisco.com/go/unifiedcomputing/b-series-doc>.

For a complete list of all C-Series documentation, see the *Cisco UCS C-Series Servers Documentation Roadmap* available at the following URL: <http://www.cisco.com/go/unifiedcomputing/c-series-doc>.

**Note**

The *Cisco UCS B-Series Servers Documentation Roadmap* includes links to documentation for Cisco UCS Manager and Cisco UCS Central. The *Cisco UCS C-Series Servers Documentation Roadmap* includes links to documentation for Cisco Integrated Management Controller.

Documentation Feedback

To provide technical feedback on this document, or to report an error or omission, please send your comments to ucs-director-docfeedback@cisco.com. We appreciate your feedback.

Obtaining Documentation and Submitting a Service Request

For information on obtaining documentation, using the Cisco Bug Search Tool (BST), submitting a service request, and gathering additional information, see [What's New in Cisco Product Documentation](#).

To receive new and revised Cisco technical content directly to your desktop, you can subscribe to the [What's New in Cisco Product Documentation RSS feed](#). RSS feeds are a free service.



New and Changed Information for this Release

- [New and Changed Information in Release 3.5, page 1](#)

New and Changed Information in Release 3.5

The following table provides an overview of the significant changes to this guide for this current release. The table does not provide an exhaustive list of all changes made to this guide or of all new features in this release.

Table 1: New Features and Changed Behavior in Cisco UCS Director Express for Big Data, Release 3.5

Feature	Description	Where Documented
Updates to Overview chapter	Added the Splunk Distribution and Hadoop Distribution support details.	Overview
Updates to Configuring and Deploying Hadoop Cluster Deployment Templates chapter	Added the following: <ul style="list-style-type: none"> • Updated the support for Spark and Kafka services for Cloudera and Hortonworks clusters • Updated the new option to enable security for the MapR clusters • Updated the new option for the service end users to create Hadoop clusters using Advanced Catalog. 	Configuring and Deploying Hadoop Cluster Deployment Templates, on page 61
Updates to Managing Splunk Clusters chapter	Provides an option for the service end users to create Splunk clusters using Advanced Catalog.	Managing Splunk Clusters, on page 103

Feature	Description	Where Documented
Updates to Managing Hadoop Clusters chapter	Added the following: <ul style="list-style-type: none">• Provides an option to enable high availability support for Hortonworks cluster.• Provides an option to add and replace disks in Hadoop clusters.	Managing Hadoop Clusters
Updates to Big Data Cluster Configuration Settings chapter	Support for RHEL 7.3, CentOS 7.2, and CentOS 7.3 for deploying big data clusters.	Big Data Cluster Configuration Settings, on page 119



Overview

This chapter contains the following sections:

- [Cisco UCS Director Express for Big Data, page 4](#)
- [Cisco UCS Integrated Infrastructure for Big Data, page 4](#)
- [Managing Cisco UCS Director and Cisco UCS Director Express for Big Data Personalities, page 4](#)
- [Installing Cisco UCS Director Express for Big Data on VMware vSphere, page 11](#)
- [Installing Cisco UCS Director Express for Big Data Bare Metal Agent on VMware vSphere, page 13](#)
- [NTP Server Configuration on Cisco UCS Director Express for Big Data Bare Metal Agent, page 14](#)
- [Adding Oracle JDK Software Versions to Bare Metal Agent, page 15](#)
- [Cisco Server Support for Hadoop Distribution Deployments, page 15](#)
- [Cisco Server Support for Splunk Enterprise Deployments, page 17](#)
- [Adding a New Red Hat Version for Hadoop Cluster, page 18](#)
- [Supported Hadoop Distributions, page 19](#)
- [Supported Splunk Distribution, page 19](#)
- [Supported Oracle JDK Software Versions, page 20](#)
- [Upgrade Hadoop Distribution Software , page 22](#)
- [Digitally Signed Images, page 23](#)
- [Upgrade of Bare Metal Agent, page 25](#)
- [Configuring Cisco UCS Manager Accounts, page 25](#)
- [High-level Workflow to Create an Instant Hadoop Cluster, page 25](#)
- [High-level Workflow to Create a Customized Hadoop Cluster, page 26](#)

Cisco UCS Director Express for Big Data

Cisco UCS Director Express for Big Data is a single-touch solution within Cisco UCS Director that automates deployment of Big Data infrastructure. Cisco UCS Director Express for Big Data provides a single management pane across physical infrastructure and across Hadoop and Splunk Enterprise software. It supports key Hadoop distributions, including Cloudera, MapR, and Hortonworks.

Cisco UCS Director Express for Big Data delivers end-to-end automation of Hadoop cluster deployment, allowing you to spin up and expand clusters on-demand. The physical infrastructure configuration is handled automatically, with minimal user input. The configuration includes compute, internal storage, network, and installation of operating system, Java packages, and Hadoop, along with the provisioning of Hadoop services. This is achieved through Cisco UCS service profiles wherein both the physical infrastructure and Hadoop configuration are incorporated into a Hadoop cluster deployment profile.

Cisco UCS Director Express for Big Data also delivers end-to-end automation of Splunk cluster deployment, with minimal user input. This is achieved through Cisco UCS service profiles wherein both the physical infrastructure and Splunk configuration are incorporated into a Splunk cluster deployment profile.

Cisco UCS Integrated Infrastructure for Big Data

Cisco UCS Integrated Infrastructure for Big Data is an industry leading architecture designed to meet various Big Data workloads. It scales as processing and storage needs grow without increasing management challenges and delivers predictable performance along with reduced total cost of ownership (TCO).

Cisco UCS Integrated Infrastructure consists of the following components:

- Cisco UCS Fabric Interconnects
- Cisco UCS Fabric Extenders
- Cisco UCS C-Series Rack-Mount Servers
- Cisco UCS S-Series Storage Servers
- Cisco UCS Virtual Interface Cards (VICs)
- Cisco UCS Manager

You can read more about the Cisco UCS Integrated Infrastructure for Big Data in the [Data Center Designs Cloud Computing - Design Zone for Big Data](#).

Managing Cisco UCS Director and Cisco UCS Director Express for Big Data Personalities

Cisco UCS Director is the default personality made available after deployment, but you can choose to use only Cisco UCS Director Express for Big Data, or use both Cisco UCS Director and Cisco UCS Director Express for Big Data.

You can manage personalities here: **Administration > License > License Keys > Manage Personalities**.

Table 2: Personality Switch Behavior

Personality Selection	Cisco UCS Director Features	Cisco UCS Director Express for Big Data Features
Cisco UCS Director (Default)	Yes	No
Cisco UCS Director Express for Big Data	No	Yes
Cisco UCS Director and Cisco UCS Director Express for Big Data	Yes	Yes



Note Depending on the personality, you start with and the personality selection, Cisco UCS Director and Cisco UCS Director Express for Big Data features are enabled or disabled with the restart of services on the appliance.

Creating User Roles

You can create user roles that are specific to Cisco UCS Director Express for Big Data, and define menu settings and permissions for those users. Ensure that you create a group before you add users to any role.



Note You can determine the default roles only if the **Default Role** column on the **User Roles** page is marked with **Yes** in the system. Navigate to **Administration > System > User Roles**.

For example, you can create the following user roles, and then create users with those role:

- HadoopUser—A Hadoop user
- HadoopAdmin—A Hadoop administrator

For more information on Managing Users and Groups, see the latest *Cisco UCS Director Administration Guide*.

Hadoop Administrator Permissions

A Hadoop administrator can:

- Read—Permission to only read a file.
- Write—Permission to read, write, and modify a file.
- Read/Write—Permission to read and/or write to a file

The following table shows a list of operations that a Hadoop administrator can do:

Operations	Permissions	
	Read	Write
Virtual Computing	Yes	Yes (Only VM Management Actions)
VM Label	Yes	—
Assign VM to VDC	Yes	—
Virtual Storage	Yes	Yes
Virtual Network	Yes	Yes
Physical Computing	Yes	Yes
Physical Storage	Yes	Yes
Physical Network	Yes	Yes
Group Service Request	Yes	Yes
Approver Service Request	Yes	Yes
Budgeting	Yes	Yes
Resource Accounting	Yes	—
Chargeback	Yes	—
System Admin	Yes	Yes
Users and Groups	Yes	Yes
Virtual Accounts	Yes	Yes
Catalogs	Yes	Yes
VDC	Yes	Yes
Computing Policy	Yes	Yes
Storage Policy	Yes	Yes
Network Policy	Yes	Yes
Service Delivery	Yes	Yes
Resource Limit Report	Yes	Yes

Operations	Permissions	
	Read	Write
Group Users	Yes	Yes
Cloudsense Reports	Yes	Yes
Cloudsense Assessment Reports	Yes	Yes
Orchestration	Yes	Yes
Open Automation Modules	Yes	Yes
CS Shared Reports	Yes	Yes
Remote VM Access	—	Yes
Mobile Access Settings	Yes	Yes
End User Chargeback	Yes	—
Resource Groups	Yes	Yes
Tag Library	Yes	Yes
Big Data Infra	Yes	—
Big Data Accounts	—	Yes
Big Data Cluster Management	—	Yes
Big Data Node Management	—	Yes
Big Data Performance Test	—	Yes
Big Data Service Management	—	Yes
Big Data Role Management	—	Yes
Big Data UCS SP Template	—	Yes
Big Data Hadoop Profile Template	—	Yes
Big Data Hadoop Deploy Template	—	Yes
Big Data Cluster Deployment	—	Yes
Big Data License Upload	—	Yes

Operations	Permissions	
	Read	Write
Big Data Configuration Parameters Template	—	Yes
Big Data Faults	—	Yes
Big Data Settings - QoS	—	Yes
Big Data Settings - IP Pool	—	Yes
Big Data Settings - Pre_Cluster Sanity	—	Yes
Big Data Settings - Hadoop Software Upload	—	Yes
Big Data Settings - Configuration Check Rules	—	Yes
REST API access	Yes	Yes
Allow Change Password - Users	Yes	Yes

Hadoop User Permissions

A Hadoop user can:

- Read—Permission to only read a file.
- Write—Permission to read, write, and modify a file.
- Read/Write—Permission to read and/or write to a file.

The following table shows a list of operations that a Hadoop user can do:

Operations	Permissions	
	Read	Write
Virtual Computing	Yes	—
VM Label	Yes	—
Assign VM to VDC	Yes	—
Virtual Storage	Yes	—
Virtual Network	Yes	—

Operations	Permissions	
	Read	Write
Physical Computing	Yes	—
Physical Storage	Yes	—
Physical Network	Yes	—
Group Service Request	Yes	Yes
Approver Service Request	Yes	Yes
Budgeting	Yes	—
Resource Accounting	Yes	—
Chargeback	Yes	—
System Admin	Yes	—
Users and Groups	Yes	—
Virtual Accounts	Yes	—
Catalogs	Yes	—
VDC	Yes	—
Computing Policy	Yes	—
Storage Policy	Yes	—
Network Policy	Yes	—
Service Delivery	Yes	—
Resource Limit Report	Yes	—
Group Users	Yes	—
Cloudsense Reports	Yes	—
Cloudsense Assessment Reports	Yes	—
Orchestration	—	—
Open Automation Modules	—	—

Operations	Permissions	
	Read	Write
CS Shared Reports	—	—
Remote VM Access	—	—
Mobile Access Settings	—	—
End User Chargeback	—	—
Resource Groups	—	—
Tag Library	—	—
Big Data Infra	Yes	—
Big Data Accounts	—	—
Big Data Cluster Management	—	—
Big Data Node Management	—	—
Big Data Performance Test	—	—
Big Data Service Management	—	—
Big Data Role Management	—	—
Big Data UCS SP Template	—	—
Big Data Hadoop Profile Template	—	—
Big Data Hadoop Deploy Template	—	—
Big Data Cluster Deployment	—	—
Big Data License Upload	—	—
Big Data Configuration Parameters Template	—	—
Big Data Faults	—	—
Big Data Settings - QoS	—	—
Big Data Settings - IP Pool	—	—

Operations	Permissions	
	Read	Write
Big Data Settings - Pre_Cluster Sanity	—	—
Big Data Settings - Hadoop Software Upload	—	—
Big Data Settings - Configuration Check Rules	—	—
REST API access	Yes	—
Allow Change Password - Users	—	—

Installing Cisco UCS Director Express for Big Data on VMware vSphere

The Cisco UCS Director, Release 6.5 OVF file includes Cisco UCS Director Express for Big Data, Release 3.5.



Note

We recommend that you use VMware vCenter for OVF deployment. VMware vCenter versions 5.x and above are supported. OVF deployment wizards support only IPv4 addresses. If you require IPv6, deploy the OVF with IPv4 addresses and then use the ShellAdmin to configure IPv6 addresses.

Before You Begin

You need administrator privileges to connect to VMware vCenter. Cisco UCS Director requires a user account with system administrator privileges to discover, manage and automate VMware vCenter configuration from Cisco UCS Director. These operations include creating, deleting and modifying VMs, ESXi hosts and clusters, datastores and datastore clusters, standard and DV switches, and virtual network port groups.



Note

If you do not want to use DHCP, you need the following information: IPv4 address, subnet mask, and default gateway.

Step 1

Log in to VMware vSphere Client.

Step 2

In the **Navigation** pane, choose the **Data Center** where you want to deploy Cisco UCS Director. See [Cisco UCS Director Installation on VMware vSphere](#).

- Step 3** Choose **File > Deploy OVF Template**.
- Step 4** In the **Source** pane, do one of the following to choose your OVF source location:
- Click **Browse**, navigate to the location where you downloaded the OVF, choose the file, and click **Open**.
 - Replace *FQDN* (Fully Qualified Domain Name) with the path to the URL on your local area network where the OVF is stored, including the IP address or domain name, and click **Next**.
- Step 5** In the **OVF Template Details** pane, verify the details, and click **Next**.
- Step 6** In the **Name and Location** pane, do the following:
- a) In the **Name** field, edit the default VM name.
 - b) From the **Inventory Location** area, choose the inventory location where Cisco UCS Director Express for Big Data is being deployed, and click **Next**.
Note If you chose a Data Center in Step 2, option b might not be available.
 - c) Click **Next**.
- Step 7** In the **Resource Pool** pane, choose the required host, cluster, or resource pool, and click **Next**.
- Step 8** In the **Disk Format** pane, choose one of the following options and click **Next**:
- **Thick Provisioned (Lazy Zeroed)** format—To allocate storage immediately in thick format. This is the recommended format. All Cisco UCS Director Express for Big Data performance data is verified with this format.
 - **Thick Provisioned (Eager Zeroed)** format—To allocate storage in thick format. It might take longer to create disks using this option.
 - **Thin Provisioned** format—To allocate storage on demand as data is written to disk.
Note We recommend that you do not choose the **Thin Provisioned** format.
- Step 9** In the **Properties** pane, enter the following information and click **Next**:
- **Management IP Address**—The management IP address to be used for eth0. If your network uses DHCP, leave the default value of 0.0.0.0.
 - **Management IP Subnet Mask**—The management IP subnet mask to be used for eth0. If your network uses DHCP, leave the default value of 0.0.0.0.
 - **Gateway IP Address**
- Step 10** In the **Ready to Complete** pane, do the following:
- a) Verify the options that you chose in the previous panes.
 - b) Check **Power on after deployment**.
If you do not check this box, you must power on the VM manually after deployment.
 - c) Click **Finish**.
- Step 11** After the appliance has booted up, copy and paste the Cisco UCS Director Express for Big Data management IP address (from the IP address that is shown) into a supported web browser to access the **Login** page.
- Step 12** On the **Login** page, enter `admin` as the username and `admin` for the login password.
Note We recommend that you change the default admin password after this initial login.

- Step 13** Choose **Administration > License**.
- Step 14** On the **License** page, click **License Keys**.
- Step 15** Click **Manage Personalities**.
- Step 16** On the **Personality Configuration** screen, check the required personalities. You can check either **UCSD** or **Big Data** or both personalities if required.
- Step 17** Click **Submit**.
- Step 18** Log in to the Cisco UCS Director VM console with the default shelladmin credentials (for example, shelladmin/changeme) to apply the selected personalities.
- Follow the prompts to change the default password.
 - From the **Cisco UCS Director Shell Menu**, choose Stop Services and press Enter.
 - Press Enter to return to the main menu.
 - From the **Cisco UCS Director Shell Menu**, choose Start Services and press Enter.
 - Press Enter to return to the main menu.
 - To verify that all services have started, choose `Display services status`.
 - Choose Quit.
-

Installing Cisco UCS Director Express for Big Data Bare Metal Agent on VMware vSphere

Before You Begin

- You must have system administrator privileges for VMware vSphere or vCenter.
- If you want to use a static IP address rather than DHCP, you must know the following information:
 - IP address
 - Subnet mask
 - Default gateway

-
- Step 1** On the Cisco.com download site for Cisco UCS Director, download Cisco UCS Director Bare Metal Agent and unzip the OVF file.
- Step 2** Log in to VMware vSphere Client.
- Step 3** In the **Navigation** pane, click the vSphere host on which you want to deploy Cisco UCS Director Express for Big Data Bare Metal Agent.
- Step 4** Choose **File > Deploy OVF Template**.
- Step 5** On the **Source** screen of the **Deploy OVF Template** window, do one of the following to choose your OVF source location and then click **Next**:
- If the OVF file is stored on your local computer, browse to the location, choose the file, and click **Open**.

- If the OVF file is stored on a server on your local area network, enter the location of the file including the IP address or fully qualified domain name of the server.

- Step 6** On the **OVF Template Details** screen, verify the details and click **Next**.
- Step 7** On the **End User License Agreement** screen, review the license agreement and click **Accept**.
- Step 8** On the **Name and Location** screen, do the following:
- a) In the **Name** field, enter a unique name for the VM.
 - b) In the **Inventory Location** area, choose the location where you want the VM to reside.
 - c) Click **Next**.
- Step 9** On the **Storage** screen, choose the storage location for the VM and click **Next**.
- Step 10** On the **Disk Format** screen, click **Next** to accept the default radio button for **Thick Provision (Lazy Zeroed)** format.
- Step 11** On the **Network Mapping** screen, choose the network for the VM and click **Next**.
- Step 12** On the **Properties** screen, do the following:
- a) Configure the IP addresses for both the NICs (eth0 and eth1) that you want to assign, as follows:
 - To use DHCP to assign the IP addresses, leave the default of 0.0.0.0 in the IP address fields.
 - To use static IP addresses, enter the desired IP addresses in the IP address fields. If you only want to configure one NIC, only complete one set of IP addresses and leave the second set at the default.
 - b) Click **Next**.
- Step 13** On the **Ready to Complete** screen, verify the settings and click **Finish**.
A message appears to indicate that Cisco UCS Director Express for Big Data Bare Metal Agent is being deployed.
- Step 14** Log in to the Cisco UCS Director Express for Big Data Bare Metal Agent server with root privileges, and check if you are able to ping the Cisco UCS Director Express for Big Data Bare Metal Agent server.
- Step 15** In the **Navigation** pane, right-click the Cisco UCS Director Express for Big Data Bare Metal Agent server and choose **Edit Settings**.
- a) Choose the **Resources** tab.
 - b) In the **Resource Allocation** window, set CPU and Memory **Reservation** settings to the maximum.
 - c) Click **OK**.
- Step 16** Power on the VM.
-

NTP Server Configuration on Cisco UCS Director Express for Big Data Bare Metal Agent

You can configure Cisco UCS Director Express for Big Data Bare Metal Agent to have its clock synchronized to an external NTP server. This ensures that the correct calendar time is maintained on the Bare Metal Agent.

Ensure that you synchronize Cisco UCS Director Express for Big Data and Cisco UCS Director Express for Big Data Bare Metal Agent before you deploy any Hadoop cluster. It is recommended that you also synchronize Cisco UCS Manager and Cisco UCS Director Express for Big Data on the configured NTP server.

Follow the steps to locate the `ntp_server_config.sh`:

- Locate `/opt/cnsaroot/bigdata_templates/ntp_server_config.sh`
- Add execute permissions (`chmod+x ntp_server_config.sh`)
- Execute (`./ntp_server_config.sh <ntp_server_ip or hostname>`) on the Cisco UCS Director Express for Big Data Bare Metal Agent server.

Adding Oracle JDK Software Versions to Bare Metal Agent

You can upload Oracle JDK software and use Oracle JDK for all Hadoop distributions (Cloudera, MapR, and Hortonworks) through an instant Hadoop cluster and customized Hadoop cluster creation actions. You can also add new nodes to the existing cluster and support upgrading existing clusters.

-
- Step 1** Log in to the Cisco UCS Director Bare Metal Agent server.
 - Step 2** Navigate to **Solutions > Big Data > Settings** and click the **Software Catalogs** tab to create JDK software version folder names for each version.
 - Step 3** Copy JDK files in `.rpm` or `.gz` format to the respective version folders.
 - Step 4** On the menu bar, choose **Administration > Integration** to track software uploads.
 - Step 5** Click the **Change Record** tab to track the software upload in progress and verify if completed, failed, or timeout.
-

Cisco Server Support for Hadoop Distribution Deployments

The table shows Cisco UCS Director Express for Big Data compatibility with Cisco UCS hardware and software. This table does not reflect the compatibility between Cisco UCS hardware and software.

For information regarding Cisco UCS compatibility, see the [Cisco UCS Hardware and Software interoperability Matrices](#) for the appropriate releases.

**Note**

All Cisco UCS Director Express for Big Data functionality may not be available across all supported Cisco UCS software versions. Certain features may not be available in older versions of Cisco UCS software.

Components	Supported Models	Supported Software	
		From Version	To Version
Cisco UCS Manager	Software	Cisco UCS Infrastructure Bundle and Cisco UCS Manager Software Bundle, Release 2.2	Cisco UCS Infrastructure Bundle and Cisco UCS Manager Software Bundle, Release 3.1(2b), Release 3.1(2f), and Release 3.2(0.204a) for M3 Rack servers Cisco UCS Infrastructure Bundle and Cisco UCS Manager Software Bundle, Release 3.1(2b), Release 3.1(2f), and Release 3.2(0.204a) for M4 Rack servers and Storage servers
Cisco UCS 6200 Series Fabric Interconnects	<ul style="list-style-type: none"> • 6248UP • 6296UP 	Cisco UCS Infrastructure Bundle and Cisco UCS Manager Software Bundle, Release 2.2	Cisco UCS Infrastructure Bundle and Cisco UCS Manager Software Bundle, Release 3.1(2b)
Cisco UCS 2000 Series FEX Modules	2232PP 10GigE	Cisco UCS Infrastructure Bundle and Cisco UCS Manager Software Bundle, Release 2.2	Cisco UCS Infrastructure Bundle and Cisco UCS Manager Software Bundle, Release 3.1(2b)
Cisco UCS C-Series Rack-Mount Servers (Managed by Cisco UCS Manager)	<ul style="list-style-type: none"> • C240 M3 Rack Server • C240 M4 Rack Server • C220 M3 Rack Server • C220 M4 Rack Server • S3260 Storage Server 	Cisco UCS C-Series Rack-Mount Server Software Bundle, Release 2.2	Cisco UCS C-Series Rack-Mount Server Software Bundle, Release 3.1(2b) Cisco UCS Infrastructure Bundle and Cisco UCS Manager Software Bundle, Release 3.1(2b) for M4 Rack servers and Storage servers
Cisco UCS 6300 series Fabric interconnects	<ul style="list-style-type: none"> • 6324 • 6332 • 6332-16UP 	Cisco UCS Infrastructure Bundle and Cisco UCS Manager Software Bundle, Release 3.1	Cisco UCS Infrastructure Bundle and Cisco UCS Manager Software Bundle, Release 3.1(3b)

Cisco Server Support for Splunk Enterprise Deployments

The table shows Cisco UCS Director Express for Big Data compatibility with Cisco UCS hardware and software. This table does not reflect the compatibility between Cisco UCS hardware and software.

For information regarding Cisco UCS compatibility, see the [Cisco UCS Hardware and Software interoperability Matrices](#) for the appropriate releases.


Note

All Cisco UCS Director Express for Big Data functionality may not be available across all supported Cisco UCS software versions. Certain features may not be available in older versions of Cisco UCS software.

Components	Supported Models	Supported Software	
		From Version	To Version
Cisco UCS Manager	Software	Cisco UCS Infrastructure Bundle and Cisco UCS Manager Software Bundle, Release 2.2	Cisco UCS Infrastructure Bundle and Cisco UCS Manager Software Bundle, Release 3.1(2b), Release 3.1(2f), and Release 3.2(0.204a) for M3 Rack servers Cisco UCS Infrastructure Bundle and Cisco UCS Manager Software Bundle, Release 3.1(2b), Release 3.1(2f), and Release 3.2(0.204a) for M4 Rack servers
Cisco UCS 6200 Series Fabric Interconnects	6296UP	Cisco UCS Infrastructure Bundle and Cisco UCS Manager Software Bundle, Release 2.2	Cisco UCS Infrastructure Bundle and Cisco UCS Manager Software Bundle, Release 3.1(2b)

Components	Supported Models	Supported Software	
		From Version	To Version
Cisco UCS C-Series Rack-Mount Servers (Managed by Cisco UCS Manager)	<ul style="list-style-type: none"> • C220 M4 Rack Server for the Master Node, Deployment Node, and the Search Heads. • C240 M4 Rack Server and S3260 Storage Server for the Indexers. 	Cisco UCS C-Series Rack-Mount Server Software Bundle, Release 2.2	Cisco UCS C-Series Rack-Mount Server Software Bundle, Release 3.1(2b) Cisco UCS Infrastructure Bundle and Cisco UCS Manager Software Bundle, Release 3.1(2b) for M4 Rack servers and Storage servers
Cisco UCS 6300 series Fabric interconnects	<ul style="list-style-type: none"> • 6324 • 6332 • 6332-16UP 	Cisco UCS Infrastructure Bundle and Cisco UCS Manager Software Bundle, Release 3.1	Cisco UCS Infrastructure Bundle and Cisco UCS Manager Software Bundle, Release 3.1(3b)

Adding a New Red Hat Version for Hadoop Cluster

For more information on uploading supported Red Hat Enterprise Linux versions, see Chapter 12, Managing Hadoop and Splunk Clusters in *Cisco UCS Director Express for Big Data Management Guide*.

-
- 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**.
- Step 5** Choose the target Cisco UCS Director Express for Big Data Bare Metal Agent from the **Target BMA** drop-down list.
- Step 6** Check the **Restart BMA Services** check box to restart Cisco UCS Director Express for Big Data Bare Metal Agent after uploading the required files.
- Step 7** To verify that the new Red Hat version (operating system software) is available in the Cisco UCS Director Express for Big Data server, perform the following:
- Log in to the Cisco UCS Director Express for Big Data user interface.
 - On the menu bar, choose **Administration > Physical Accounts**.
 - Click the **Bare Metal Agents** tab.
- You can find the new Red Hat version listed in the **Image Catalogs** column of the **Bare Metal Agents** report.

New RHEL Version

Red Hat Enterprise Linux 7.2

Download the following file:

- rhel-server-7.2-x86_64-dvd.iso from [Red Hat Enterprise Linux](#)

Supported Hadoop Distributions

Cisco UCS Director Express for Big Data supports the following Hadoop distributions:

- Cloudera 5.4, 5.5, 5.6, 5.8, 5.10, and 5.11
- MapR 3.1, 4.0, 4.1, 5.0, 5.1, and 5.2
- Hortonworks 2.1, 2.2, 2.3, 2.4, and 2.5.3
- BigInsights 4.2

Table 3: Java and JDK Software for Hadoop Distributions

Hadoop Distribution	Supported Hadoop Distribution Versions	Installed Java and JDK
Cloudera	5.0.1, 5.0.6, 5.2.0, 5.2.1, 5.3.0, 5.4.1, 5.5.0, 5.6.0, 5.8.0, 5.10, and 5.11	oracle-j2sdk1.7
MapR	3.1.1, 4.0.1, 4.0.2, 4.1.0, 5.0.0, 5.1.0, and 5.2.1	java-1.7.0-openjdk
Hortonworks	2.1, 2.2, 2.3, 2.4, and 2.5.3	java-1.7.0-openjdk
BigInsights	4.2	java-1.8.0-openjdk



Note

You can choose and install the required Oracle JDK version.

Supported Splunk Distribution

Cisco UCS Director Express for Big Data supports the following Splunk distribution:

Splunk Distribution	Supported Splunk Distribution Version
Splunk	6.3.2, 6.4.0, 6.5.2, and 6.6.1

Splunk Distribution Version to Upgrade	Supported Upgrade Version
Splunk 6.4	Splunk 6.5
Splunk 6.4	Splunk 6.6.1

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 Enterprise 5.11.1	—	JDK 1.7.0_25	JDK 1.8.0_121
Cloudera Enterprise 5.10.0	—	JDK 1.7.0_25	JDK 1.8.0_121
Cloudera Enterprise 5.8.0	—	JDK 1.7.0_25	JDK 1.8.0_60
Cloudera Enterprise 5.6.x	—	JDK 1.7.0_25	JDK 1.8.0_60
Cloudera Enterprise 5.5.0	—	JDK 1.7.0_25	JDK 1.8.0_60
Cloudera Enterprise 5.4.x	—	JDK 1.7.0_25	JDK 1.8.0_60
Cloudera Enterprise 5.3.x	—	—	JDK 1.8.0_11

Hadoop Distribution Version to Upgrade	Supported Upgrade Version
Cloudera Enterprise 5.4.x, JDK 1.8	Cloudera Enterprise 5.5.0, JDK 1.8
Cloudera Enterprise 5.4.x, JDK 1.8	Cloudera Enterprise 5.6.x, JDK 1.8
Cloudera Enterprise 5.4.x, JDK 1.8	Cloudera Enterprise 5.8.x, JDK 1.8
Cloudera Enterprise 5.6.x, JDK 1.8	Cloudera Enterprise 5.8.x, JDK 1.8

Hadoop Distribution Version to Upgrade	Supported Upgrade Version
Cloudera Enterprise 5.8.0, JDK 1.8	Cloudera Enterprise 5.10.0, JDK 1.8
Cloudera Enterprise 5.8.0, JDK 1.8	Cloudera Enterprise 5.11.1, JDK 1.8

Supported Oracle JDK Versions for MapR and Upgrade Scenarios

Hadoop Distribution	JDK 1.7	JDK 1.8
MapR 5.2.1	Yes	Yes
MapR 5.2.0	Yes	Yes
MapR 5.1.0	Yes	Yes
MapR 5.0.0	Yes	Yes
MapR 4.1.0	No	Yes

Hadoop Distribution Version to Upgrade	Supported Upgrade Version
MapR 5.0.0, JDK 1.8	MapR 5.1.0, JDK 1.8
MapR 4.0.2, JDK 1.8	MapR 5.2.0, JDK 1.8

Supported Oracle JDK Software Versions for Hortonworks and Upgrade Scenarios

Hadoop Distribution	JDK 1.7	JDK 1.8
Hortonworks 2.5.3	—	Yes
Hortonworks 2.4	—	Yes
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
Hortonworks 2.2, JDK 1.7	Hortonworks 2.4, JDK 1.8

Upgrade Hadoop Distribution Software

You can upgrade to the latest Hadoop distributions from the following Hadoop distributions:

Table 4: Cloudera

Hadoop Distribution Version to Upgrade	Supported Upgrade Version
Cloudera-5.0.1	Cloudera-5.4.1
Cloudera-5.0.6	Cloudera-5.4.1
Cloudera-5.2.0	Cloudera-5.4.1
Cloudera-5.2.1	Cloudera-5.4.1
Cloudera-5.3.0	Cloudera-5.4.1
Cloudera-5.4.x	Cloudera-5.6.x
Cloudera-5.8.0	Cloudera-5.10.0
Cloudera-5.8.0	Cloudera-5.11.1

Table 5: MapR

Hadoop Distribution Version to Upgrade	Supported Upgrade Version
MapR-4.0.2	MapR-4.1.0
MapR-4.0.2	MapR-5.0.0
MapR-4.1.0	MapR-5.0.0

Table 6: Hortonworks

Hadoop Distribution Version to Upgrade	Supported Upgrade Version
Hortonworks-2.2 (ambari-1.7.0-centos6.tar.gz)	Hortonworks-2.3 Note Download ambari-2.1.1-centos6.tar.gz from http://public-repo-1.hortonworks.com/ambari/centos6/2.x/updates/2.1.1 .

Digitally Signed Images

Cisco UCS Director Express for Big Data images are delivered in digitally signed zip files. These signed zip files are wrapped in a container zip file that includes the following:

- Digitally signed zip file—Contains the Cisco UCS Director Express for Big Data installation or upgrade image
- Verification program—Verifies the certificate chain and signature. During certificate chain validation, the program verifies the authenticity of the end-entity certificate using Cisco's SubCA and root CA certificates. Then, the authenticated end-entity certificate is used to verify the signature.
- Digital signature file—Contains the signature that you can verify before installation or upgrade.
- Certificate file—Enables you to verify the digital signature. This Cisco-signed x.509 end-entity certificate contains a public key that can be used to verify the signature. This certificate is chained to the Cisco root posted on <http://www.cisco.com/security/pki/certs/crcam2.cer>.
- ReadMe file—Provides the information and instructions required to verify the digitally signed zip file.

Verify the image offline. Once the image is verified, you can begin the installation or upgrade of Cisco UCS Director Express for Big Data.

Requirements for Verifying Digitally Signed Images

Before you verify a Cisco UCS Director Express for Big Data digitally signed image, ensure that you have the following on your local machine:

- Connectivity to <https://www.cisco.com> during the verification process
- Python 2.7.4 or later
- OpenSSL

Verifying a Digitally Signed Image

Before You Begin

Download the Cisco UCS Director Express for Big Data image from Cisco.com.

Step 1 Unzip the file you downloaded from Cisco.com and verify that it contains the following files:

- ReadMe file
- Digitally signed zip file, for example CUCSD_6_5_0_0_65365_VMWARE_GA.zip, CUCSD_6_5_0_0_65717_HYPERV_GA.zip, or cucsd_patch_6_5_0_0_65365.zip
- Certificate file, for example UUCS_GENERIC_IMAGE_SIGNING-CCO_RELEASE.cer
- Digital signature generated for the image, for example CUCSD_6_5_0_0_65365_VMWARE_GA.zip.signature, CUCSD_6_5_0_0_65717_HYPERV_GA.zip.signature, or cucsd_patch_6_5_0_0_65365.zip.signature
- Signature verification program, for example cisco_x509_verify_release.py

Step 2 Review the instructions in the ReadMe file.

Note If there are any differences between these instructions and those in the ReadMe, follow the ones in the ReadMe.

Step 3 Run the signature verification program.

Example: Signature Verification for Upgrade Patch

```
python cisco_x509_verify_release.py -e UCS_GENERIC_IMAGE_SIGNING-CCO_RELEASE.cer
-i cucsd_patch_6_5_0_0_65365.zip -s cucsd_patch_6_5_0_0_65365.zip.signature -v dgst -sha512
```

Example: Signature Verification for VMware OVF Installation

```
python cisco_x509_verify_release.py -e UCS_GENERIC_IMAGE_SIGNING-CCO_RELEASE.cer
-i CUCSD_6_5_0_0_65365_VMWARE_GA.zip -s CUCSD_6_5_0_0_65365_VMWARE_GA.zip.signature -v dgst -sha512
```

Example: Signature Verification for Hyper-V VHD Installation

```
python cisco_x509_verify_release.py -e UCS_GENERIC_IMAGE_SIGNING-CCO_RELEASE.cer
-i CUCSD_6_5_0_0_65717_HYPERV_GA.zip -s CUCSD_6_5_0_0_65717_HYPERV_GA.zip.signature -v dgst -sha512
```

Step 4 Review the output and ensure that the verification has succeeded.

Example: Expected Output for Upgrade

```
Downloading CA certificate from http://www.cisco.com/security/pki/certs/crcam2.cer ...
Successfully downloaded and verified crcam2.cer.
Downloading SubCA certificate from http://www.cisco.com/security/pki/certs/innerspace.cer ...
Successfully downloaded and verified innerspace.cer.
Successfully verified root, subca and end-entity certificate chain.
Successfully verified the signature of cucsd_patch_6_5_0_0_65365.zip using
UCS_GENERIC_IMAGE_SIGNING-CCO_RELEASE.cer
```

Example: Expected Output for VMware OVF Installation

```
Downloading CA certificate from http://www.cisco.com/security/pki/certs/crcam2.cer ...
Successfully downloaded and verified crcam2.cer.
Downloading SubCA certificate from http://www.cisco.com/security/pki/certs/innerspace.cer ...
Successfully downloaded and verified innerspace.cer.
```

```
Successfully verified root, subca and end-entity certificate chain.  
Successfully verified the signature of CUCSD_6_5_0_0_65365_VMWARE_GA.zip using  
UCS_GENERIC_IMAGE_SIGNING-CCO_RELEASE.cer  
Example: Expected Output for Hyper-V VHD Installation  
  
Downloading CA certificate from http://www.cisco.com/security/pki/certs/crcam2.cer ...  
Successfully downloaded and verified crcam2.cer.  
Downloading SubCA certificate from http://www.cisco.com/security/pki/certs/innerspace.cer ...  
Successfully downloaded and verified innerspace.cer.  
Successfully verified root, subca and end-entity certificate chain.  
Successfully verified the signature of CUCSD_6_5_0_0_65717_HYPERV_GA.zip using  
UCS_GENERIC_IMAGE_SIGNING-CCO_RELEASE.cer
```

What to Do Next

Install or upgrade Cisco UCS Director Express for Big Data.

Upgrade of Bare Metal Agent

For detailed information on upgrading Bare Metal Agent, see the [Cisco UCS Director Upgrade Guide](#).

Configuring Cisco UCS Manager Accounts

Each Cisco UCS Manager account represents a single Cisco UCS domain that has to be managed by Cisco UCS Director Express for Big Data. See [Cisco UCS Director Management Guide for Cisco UCS Manager](#).

High-level Workflow to Create an Instant Hadoop Cluster

-
- Step 1** Create a Cisco UCS Manager account. See [Cisco UCS Director Management Guide for Cisco UCS Manager](#).
 - Step 2** Configure Big Data IP pools. See [Adding a Big Data IP Pool](#).
 - Step 3** Create an Instant Hadoop Cluster. See [Creating an Instant Hadoop Cluster](#).
-

High-level Workflow to Create a Customized Hadoop Cluster

-
- Step 1** Configure a Cisco UCS Service Profile template for Big Data. For more information, see [Creating a Cisco UCS Service Profile Template for Big Data](#).
 - Step 2** Create a Hadoop cluster configuration parameters template. See [Creating a Hadoop Cluster Configuration Parameters Template](#).
 - Step 3** Configure a Hadoop cluster profile template. See [Creating a Hadoop Cluster Profile Template](#).
 - Step 4** Configure a Hadoop cluster deployment template. See [Creating a Cluster Deployment Template](#).
 - Step 5** Create a customized Hadoop cluster. See [Creating a Customized Hadoop Cluster](#).
-



Licenses for Cisco UCS Director Express for Big Data

This chapter contains the following sections:

- [About Licenses, page 27](#)
- [Fulfilling the Product Access Key, page 28](#)
- [Updating the License, page 29](#)
- [Standard License Features, page 29](#)

About Licenses

You must obtain a license to use Cisco UCS Director Express for Big Data, as follows:

- 1 Before you install Cisco UCS Director Express for Big Data, generate the Cisco UCS Director Express for Big Data license key and claim a certificate (Product Access Key).
- 2 Register the Product Access Key (PAK) on the Cisco software license site, as described in [Fulfilling the Product Access Key, on page 28](#).
- 3 After you install Cisco UCS Director Express for Big Data, update the license in Cisco UCS Director Express for Big Data as described in [Updating the License, on page 29](#).
- 4 After the license has been validated, you can start to use Cisco UCS Director Express for Big Data.

Fulfilling the Product Access Key

Before You Begin

You need the PAK number.

-
- Step 1** Navigate to the [Cisco Software License website](#).
- Step 2** If you are directed to the Product License Registration page, you can take the training or click **Continue to Product License Registration**.
- Step 3** On the Product License Registration page, click **Get New Licenses from a PAK or Token**.
- Step 4** In the **Enter a Single PAK or TOKEN to Fulfill** field, enter the PAK number.
- Step 5** Click **Fulfill Single PAK/TOKEN**.
- Step 6** Complete the additional fields in **License Information** to register your PAK:

Name	Description
Organization Name	The organization name.
Site Contact Name	The site contact name.
Street Address	The street address of the organization.
City or Town	The city or town.
State or Province	The state or province.
Zip or Postal Code	The zip code or postal code.
Country	The country name.

- Step 7** Click **Issue Key**.
The features for your license appear, and you receive an email with the Digital License Agreement and a zipped license file.
-

Updating the License

Before You Begin

If you received a zipped license file by an email, extract and save the license (.lic) file to your local machine.

Step 1 Choose **Administration > License**.

Step 2 On the **License** page, click **License Keys**.

Step 3 Click **Update License**.

Step 4 On the **Update License** screen, do the following:

- a) Drop the `.lic` file from your local system or click **Select a File** and navigate to the location where you stored the `.lic` file.
To enter license text instead of file upload, check the **Enter License Text** checkbox and enter the license text in the **License Text** field.
 - b) Click **Submit**.
The license file is processed, and a message appears confirming the successful update.
-

Standard License Features

The following table lists the features supported for the standard licenses:

Features	Standard License
Operating system and Hadoop software installation	X
Server, Network, and Storage provisioning	X
On-demand cluster creation and expansion	X
Customized cluster creation	X
Automated cluster node addition and deletion	X
Add pre-existing Hadoop nodes	X
Start and stop cluster services	X
Start, stop, and restart Cluster	X
Dashboard for health and status monitoring	X

Features	Standard License
Support for latest Cloudera, MapR, and Hortonworks releases	X
Monitoring Storage and Network utilization	X
Monitoring Top Active and Long Running Jobs	X
On-Demand inventory collection	X
DIMM, Disk, Node, and Service Failure Alerts	X
Capacity Planning Alerts	X
HDFS rebalancing	X
LDAP integration	X
Hadoop parameter configuration	X
Globalization and localization Support	X
Cusotmizable workflows	X
North Bound REST API support	X
Cluster configuration consistency checks	X
Cluster performance analysis	X
Historical performance analysis	X
Automated install and setup of new Hadoop services	X
Automated Hadoop version upgrade	X
Role-Based Access Control (RBAC)	X
Approval workflows	X



Managing Hadoop Accounts

This chapter contains the following sections:

- [Adding a Hadoop Account, page 31](#)
- [Running a Cluster Inventory for a Hadoop Account , page 33](#)
- [Purging Big Data Cluster Account Details, page 33](#)
- [Rolling Back a Hadoop Cluster for a Hadoop Account, page 34](#)
- [Access to Hadoop Managers from Cisco UCS Director Express for Big Data, page 34](#)

Adding a Hadoop Account

If you want to manage your Hadoop cluster using Cisco UCS Director Express for Big Data, add a Hadoop account.

Step 1 Choose **Solutions > Big Data > Accounts**.

Step 2 Click **Hadoop Accounts**.

Step 3 Click **Add**.

Step 4 On the **Add Account** screen, complete the following fields:

Name	Description
Pod drop-down list	The pod to which you add the Hadoop account.
Account Name field	The account name.

Name	Description
Account Type drop-down list	The type of Hadoop distribution used for the cluster. Choose one of the following: <ul style="list-style-type: none"> • Cloudera • MapR • Hortonworks • BigInsights
Management Console IP field	The management console IP address.
SSH (root) Password field	The password associated with the SSH username.
Confirm Password field	The password associated with the SSH username.
Hadoop Manager Password field	The password associated with the Hadoop Manager for that account type.
Confirm Password field	The password associated with the Hadoop Manager for that account type.

Step 5 Click **Submit**.

Step 6 For the following actions, select a Hadoop account.

Name	Description
Edit	Allows you to edit a Hadoop account.
Delete	Deletes a Hadoop account.
Check Configuration	Allows you to validate an existing cluster configuration.
Rollback Cluster	Allows you to roll back a cluster and make all the nodes in the cluster available for bare metal servers. Roll back is not supported for a derived cluster account.
Launch Hadoop Manager	Allows you to launch the Hadoop manager from Cisco UCS Director Express for Big Data
View Details	Provides details of a Hadoop account.
Run Inventory	Collects the inventory of the Hadoop cluster for the selected Hadoop Account and the data persists in the Cisco UCS Director Express for Big Data database.
Configure Cluster	Allows you to customize the Hadoop cluster after creation.

Name	Description
Modify Credentials	Allows you to modify the SSH and admin Console credentials for a Hadoop account.

Running a Cluster Inventory for a Hadoop Account

When you create a Hadoop Account, a new system task (inventory collector) is created within the account. Navigate to the Big Data Tasks folder here: **Administration > System > System Tasks**. The system task collects the inventory of the Hadoop cluster for the selected Hadoop Account and establishes data in the Cisco UCS Director database. This collector adds to the system scheduler so that the system scheduler can be called at the interval configured in the collector (for example, 30 minutes).

For more information on how to manage system tasks in Cisco UCS Director, see the latest *Cisco UCS Director Administration Guide*.

-
- Step 1** Choose **Solutions > Big Data > Accounts**.
 - Step 2** Click **Hadoop Accounts**.
 - Step 3** On the **Accounts** screen, choose the Hadoop Account for which you want to run the inventory.
 - Step 4** Click **Run Inventory**.
 - Step 5** Click **Submit**.
-

Purging Big Data Cluster Account Details

A new system task (purging) is created within the account. Navigate to the Big Data Tasks folder here: **Administration > System > System Tasks**. The system scheduler deletes the monitoring and metric data that are older than 18 days. When the Cisco UCS Director Express for Big Data is up:

- the stale entries are deleted for the valid accounts that do not have add nodes.
- the account details are deleted for the invalid accounts.

Rolling Back a Hadoop Cluster for a Hadoop Account

You can roll back a Hadoop cluster and make all the nodes in the cluster available for a bare metal server. However, roll back is not supported for a derived cluster account.

-
- Step 1** Choose **Solutions > Big Data > Accounts**.
 - Step 2** Click **Hadoop Accounts**.
 - Step 3** On the **Accounts** screen, choose the Hadoop account for which you want to roll back the cluster.
 - Step 4** Click **Rollback Cluster**.
 - Step 5** On the **Rollback Cluster** screen, click **Submit**.
-

Access to Hadoop Managers from Cisco UCS Director Express for Big Data

You can access Hadoop managers for all the Hadoop accounts that you create in Cisco UCS Director Express for Big Data. Choose **Solutions > Big Data > Accounts**, and then click **Hadoop Accounts**. You can launch the Hadoop manager in supported browsers by clicking **Launch Hadoop Manager**.



Managing Splunk Accounts

This chapter contains the following sections:

- [Cisco UCS Director Express for Big Data with Splunk Enterprise](#) , page 35
- [Adding a Splunk Account](#), page 36
- [Running a Cluster Inventory for a Splunk Account](#), page 37
- [Rolling Back a Cluster for a Splunk Account](#), page 38
- [Access Splunk Enterprise Monitoring Console User Interface from Cisco UCS Director Express for Big Data](#), page 38

Cisco UCS Director Express for Big Data with Splunk Enterprise

Cisco UCS Director Express for Big Data with Splunk Enterprise deployment reliably collects and indexes machine data, from a single source to tens of thousands of sources, all in real time. Splunk Enterprise deployments expand to terabytes of operational data. Cisco UCS Director supports the massive scalability that Splunk Enterprise deployments to deliver exceptional performance.

Splunk Enterprise deployments consist of Cisco UCS as indexer and C220 M4 Server as search heads, along with administrative functions.

Splunk Enterprise deployments include the following:

- Cisco UCS
- Cisco UCS C-Series Rack-Mount Servers
- Cisco UCS Manager

Adding a Splunk Account

If you want to manage your Splunk cluster using Cisco UCS Director Express for Big Data, add a Splunk account.

Step 1 Choose **Solutions > Big Data > Accounts**.

Step 2 Click **Splunk Accounts**.

Step 3 Click **Add**.

Step 4 On the **Add Account** screen, complete the following fields:

Name	Description
Pod drop-down list	The pod to which the Splunk account to be added.
Account Name field	The Splunk account name.
Management Console IP field	The management console IP address.
SSH (Root) Password field	The password associated with the SSH username.
Confirm Password field	The password associated with the SSH username.
Splunk Manager Password field	The password associated with the Splunk Enterprise.
Confirm Password field	The password associated with the Splunk Enterprise.

Step 5 Click **Submit**.

Step 6 For the following actions, select a Splunk account.

Name	Description
Edit	Allows you to edit a Splunk account.
Delete	Deletes a Splunk account.
Check Configuration	Allows you to validate an existing cluster configuration.
Rollback Cluster	Allows you to roll back a cluster and make all the nodes in the cluster available for a bare metal server. Roll back is not supported for a derived cluster account.
Launch Splunk DMC	Allows you to launch the Splunk Enterprise from Cisco UCS Director Express for Big Data.
View Details	Provides details of a Splunk account.

Name	Description
Run Inventory	Collects the inventory of the Splunk cluster for the selected Splunk account and establishes data in the Cisco UCS Director Express for Big Data database.
Modify Credentials	Allows you to modify the SSH and admin Console credentials for a Splunk account.

Running a Cluster Inventory for a Splunk Account

With each new Splunk Account, a new system task (inventory collector) is created. Navigate to the Big Data Tasks folder here: **Administration > System > System Tasks**. The system task collects the inventory of the Splunk cluster for the selected Splunk Account and establishes data in the Cisco UCS Director database. This collector adds to the system scheduler so that it can be called at the interval configured in the collector (for example, 30 minutes).

For more information on how to manage system tasks in Cisco UCS Director, see the latest *Cisco UCS Director Administration Guide*.

-
- Step 1** Choose **Solutions > Big Data > Accounts**.
 - Step 2** Click **Splunk Accounts**.
 - Step 3** Choose the Splunk Account for which you want to run the inventory.
 - Step 4** Click **Run Inventory**.
 - Step 5** Click **Submit**.
-

Rolling Back a Cluster for a Splunk Account

You can roll back a cluster and make all the nodes in the cluster available for a bare metal server. However, roll back is not supported for a derived cluster account.

-
- Step 1** Choose **Solutions > Big Data > Accounts**.
 - Step 2** Click **Splunk Accounts**.
 - Step 3** Choose the Splunk account for which you want to roll back the cluster.
 - Step 4** Click **Rollback Cluster**.
 - Step 5** On the **Rollback Cluster** screen, click **Submit**.
-

Access Splunk Enterprise Monitoring Console User Interface from Cisco UCS Director Express for Big Data

You can access the Splunk Enterprise user Interface from Cisco UCS Director Express for Big Data. On the menu bar, choose **Solutions > Big Data > Accounts**, and then click **Splunk Accounts**. You can launch the Splunk Enterprise user interface in supported browsers by clicking the **Launch Splunk DMC**.



Configuring Big Data IP Pools

This chapter contains the following sections:

- [Big Data IP Pools, page 39](#)
- [Adding a Big Data IP Pool, page 39](#)
- [Managing Big Data IP Pools, page 40](#)

Big Data IP Pools

Big Data IP pools contain blocks of IP addresses that Cisco UCS Director Express for Big Data uses during the creation of Hadoop clusters. The IP addresses in each block must belong to the same subnet mask, default gateway, primary domain name server (DNS), and secondary DNS.



Note All IP addresses in a Big Data IP pool must be IPv4 addresses.

Adding a Big Data IP Pool

Step 1 Choose **Solutions > Big Data > Settings**.

Step 2 Click **Big Data IP Pools**.

Step 3 Click **Add**.

Step 4 In the **IP Pool Management Specification** page of the **Create an IP Pool** wizard, complete the following fields:

Name	Description
IP Pool Name field	A unique name for the IP Pool.
Description field	A short description that identifies the purpose of the pool.

Name	Description
Assignment Order drop-down list	The assignment order. Choose one of the following: <ul style="list-style-type: none"> • Default—A random identity is selected from the pool. • Sequential—The lowest available identity is selected from the pool.

Step 5 Click **Next**.

Step 6 In the **IPv4 Addresses** page of the **Create an IP Pool** wizard, complete the following fields:

- In the IPv4 Blocks table, click **Add (+)**.
- On the **Add Entry to IPv4 Blocks** screen, enter the IPv4 addresses to be included in the IP pool in the **Static IP Pool** field.
This can be a range of IP addresses, or a series of IP addresses separated by commas (,).
- Enter the subnet mask.
- Enter the default gateway.
- Enter the primary DNS server IPv4 address.
- Enter the server DNS server IPv4 address.
- Click **Submit** to save and exit.

Step 7 Click **Submit**.

Managing Big Data IP Pools

Step 1 Choose **Solutions > Big Data > Settings**.

Step 2 Click **Big Data IP Pools**. The following buttons are displayed.

Name	Description
Refresh	Refreshes the current page.
Add	Adds a new IP pool.

Step 3 For the following actions, choose an IP Pool from the table:

Name	Description
Edit	Modifies the IP pool specification.
Delete	Deletes the IP pool specification.

Name	Description
View Details	Allows you to view the IPv4 addresses in the IP pool and to view more reports. Note If you see a License Status tab, it indicates a licensing issue.

Step 4 With an IP Pool selected, click **View Details**.

Step 5 Click **IPv4 Addresses** to display the IP Addresses associated with the selected IP Pool.

Step 6 Select an IP Address from the table. The following buttons are displayed.

Name	Description
Release IP Address	The IP address is made available to be assigned against any bare metal server.
Release Multiple IP Addresses	Enables you to choose and release more than one IP address from the Release Multiple IP Addresses dialog box.



Configuring Cisco UCS Service Profile Templates for Big Data

This chapter contains the following sections:

- [Cisco UCS Service Profile Templates for Big Data](#), page 43
- [Creating a Cisco UCS Service Profile Template for Big Data](#), page 44
- [Creating a Customized Service Profile Template](#), page 59
- [Cloning a Cisco UCS Service Profile Template](#), page 60

Cisco UCS Service Profile Templates for Big Data

Cisco Unified Computing System (Cisco UCS) service profiles are a powerful means for streamlining the configuration and management of Cisco UCS servers. They provide a mechanism for rapidly provisioning servers and their associated network connections with consistency in all details of the environment. They can be set up in advance before physically installing the servers.

Service profiles are built on policies—Administrator-defined sets of rules and operating characteristics such as the server identity, interfaces, and network connectivity. Every active server in your Hadoop cluster must be associated with a service profile.

The Cisco UCS service profile template for the Big Data enables you to set up the configuration for the servers in your Hadoop cluster. The service profile template for Big Data is included in a cluster deploy template. When the cluster deploy template is applied to the servers, the service profile template configures one or more service profiles that are applied to the servers.



Note

The service profile template for Big Data wizard gathers the information required to create a service profile template. You can only apply this service profile template through the cluster deploy template.

For more information about service profiles and service profile templates, see the [Cisco UCS Manager configuration guides](#).

Creating a Cisco UCS Service Profile Template for Big Data

Before You Begin

Add a Cisco UCS Manager account.

Step 1 Choose **Solutions > Big Data > Containers**.

Step 2 Click **UCS SP Templates**.

Step 3 Click **Add (+)**.

Step 4 On the **UCS SP Template Specification** page of the **Create UCS SP Template for Big Data** wizard, complete the following fields:

Name	Description
Template Name field	A unique name for the template.
Template Description field	The description of the template.
Template Type drop-down list	Service profiles created from an initial template inherit all the properties of the template. Service profiles created from an initial service profile template are bound to the template. However, changes to the initial template do not automatically propagate to the bound service profiles. If you want to propagate changes to bound service profiles, unbind and rebind the service profile to the initial template.
Container Type drop-down list	The type of container for the cluster.
Use vNIC Bonding check box	Check the check box to have vNIC bonding. Eight vNICs are available in the Create vNIC Policy page of the Create UCS SP Template for Big Data wizard.
Use Multiple vNIC check box	Check the check box to have more than one vNIC interface and some policies for creating a template.

Step 5 Click **Next**.

What to Do Next

Create a QoS policy.

Creating a QoS Policy

The quality of service (QoS) policy assigns a system class to the outgoing traffic for a vNIC or vHBA. This system class determines the quality of service for that traffic. For certain adapters, you can also specify more controls on the outgoing traffic, such as burst and rate.

Step 1 On the **Create QoS Policy** page of the **Create UCS SP Template for Big Data** wizard, do one of the following:

- To accept the default QoS policies, click **Next**.
- To create one or more custom QoS policies, click **Add (+)** and continue with Step 2.
- To review or modify one of the default QoS policies, choose the policy in the table and click **Edit**. For information about the fields on the **Edit QoS Policy Entry** screen, see Step 2.

Step 2 On the **Add Entry to QoS Policy** screen, complete the following fields:

Name	Description
Name field	A unique name for the policy.
Priority drop-down list	Choose the priority assigned to this QoS policy. Choose one of the following: <ul style="list-style-type: none"> • Fe—Use this priority for QoS policies that control only vHBA traffic. • Platinum—Use this priority for QoS policies that control only vNIC traffic. • Gold—Use this priority for QoS policies that control only vNIC traffic. • Silver—Use this priority for QoS policies that control only vNIC traffic. • Bronze—Use this priority for QoS policies that control only vNIC traffic. • Best Effort—It is reserved for the Basic Ethernet traffic lane. If you assign this priority to a QoS policy and configure another system class as CoS 0, Cisco UCS does not default to this system class. It defaults to the priority with CoS 0 for that traffic.
Burst(Bytes) field	The normal burst size for servers that use this policy. This field determines the maximum size of traffic bursts beyond which the traffic is considered to exceed the rate limit. The default is 10240. The minimum value is 0, and the maximum value is 65535. This setting is not applicable to all adapters.

Name	Description
Rate drop-down list	<p>Choose the expected average rate of traffic. Choose one of the following:</p> <ul style="list-style-type: none"> • Line-rate—Equals a value of 0 and specifies no rate limiting. This is the default value. • Specify Manually—Enables you to specify the rate in a field. The minimum value is 0, and the maximum value is 40,000,000. <p>The granularity for rate limiting on a Cisco UCS M81KR Virtual Interface Card adapter is 1 Mbps. The adapters treat the requested rate as a "not-to-exceed" rate. Therefore, a value of 4.5 Mbps is interpreted as 4 Mbps. Any requested rate of more than 0 and less than 1 Mbps is interpreted as 1 Mbps, which is the lowest supported hardware rate limit.</p> <p>Rate limiting is not applicable to all adapters. For example, this setting is not supported on the Cisco UCS VIC-1240 Virtual Interface Card.</p>
Host Control drop-down list	<p>Determines whether Cisco UCS controls the class of service (CoS) for a vNIC. This setting has no effect on a vHBA.</p> <p>The default setting is None. Cisco UCS uses the CoS value associated with the priority regardless of the CoS value assigned by the host.</p>

Step 3 Click **Submit**.

Step 4 Click **Next**.

What to Do Next

Create a VLAN policy.

Creating a VLAN Policy

The VLAN policy creates a connection to a specific external LAN in the underlying infrastructure of a Hadoop cluster that is within a single Cisco UCS domain. The VLAN isolates traffic to that external LAN, including broadcast traffic.

Step 1 On the **Create VLAN Policy** page of the **Create UCS SP Template for Big Data** wizard, do one of the following:

- To accept the default VLAN policies, click **Next**.

- To create one or more custom VLAN policies, click **Add (+)** and continue with Step 2.
- To review or modify one of the default VLAN policies, choose the policy in the table and click **Edit**. For information about the fields on the **Edit VLAN Policy Entry** screen, see Step 2.

Step 2 On the **Add Entry to VLAN Policy** screen, complete the following fields:

Name	Description
VLAN Name field	The name of the VLAN policy.
Fabric ID drop-down list	Choose how to configure the VLAN. The setting can be one of the following: <ul style="list-style-type: none"> • Common or Global—The VLAN maps to the same VLAN ID in all available fabrics. • Fabric A—The VLAN maps to a VLAN ID that exists only in fabric A. • Fabric B—The VLAN maps to a VLAN ID that exists only in fabric B.
Sharing drop-down list	The default setting is None .

Step 3 Click **Submit**.

Step 4 Click **Next**.

What to Do Next

Create a vNIC policy.

Creating a vNIC Policy

The vNIC policy defines how a vNIC on a server connects to the LAN. Each server in a Hadoop cluster requires a vNIC policy for each of the following NICs:

- MGMT
- DATA



Note In addition to MGMT NIC, DATA NIC is available if you have checked **Multiple vNIC** in the **Create UCS SP Template for Big Data** wizard.

Step 1 On the **Create vNIC Policy** page of the **Create UCS SP Template for Big Data** wizard, do one of the following:

- To accept the default vNIC policies, click **Next**.
- To create one or more custom vNIC policies, click **Add (+)** and continue with Step 2.
- To delete a vNIC policy, choose a policy in the table and click **Delete**.
Note When you delete the vNICs, ensure that at least one vNIC is available per fabric interconnect.
- To review or modify one of the default vNIC policies, choose the policy in the table and click **Edit**. For information about the fields on the **Edit vNIC Policy Entry** screen, see Step 2.

Step 2 On the **Add Entry to vNIC Policy** screen, complete the following fields:

Name	Description
vNIC Name field	Name of the vNIC.
Fabric ID drop-down list	<p>Choose the fabric interconnect with which the vNICs created with this policy are associated.</p> <p>If you want vNICs created from this policy to access the second fabric interconnect when the default choice is unavailable, check the Enable Failover check box. When the Use vNIC Bonding check box is checked in the UCS SP Template Specification page of the Create UCS SP Template for Big Data wizard, the Enable Failover check box is disabled.</p> <p>Do not enable vNIC fabric failover under the following circumstances:</p> <p>Note</p> <ul style="list-style-type: none"> • If the Cisco UCS domain is running in Ethernet Switch Mode, vNIC fabric failover is not supported in that mode. If all Ethernet uplinks on one fabric interconnect fail, the vNICs do not fail over to other fabric interconnect. • If you associate one or more vNICs created from this template and associate the service profile with the server that has an adapter without the fabric failover support (For example, Cisco UCS 82598KR-CI 10-Gigabit Ethernet Adapter) Cisco UCS Manager generates a configuration fault.

Name	Description
VLANs area	<p>In the VLANs area, do the following to select the VLAN to be assigned to vNICs created from this policy:</p> <ol style="list-style-type: none"> 1 Click Add. 2 On the Add Entry to VLANs screen, complete the following fields: <ul style="list-style-type: none"> • Name drop-down list—Choose the VLAN that you want to associate with the vNIC template. • Set as Native VLAN check box—Check the check box if you want this VLAN to be the native VLAN for the port. 3 Click Submit.
MTU field	<p>The MTU, or packet size, to be used by the vNICs created from this vNIC policy. Enter an integer between 1500 and 9216.</p> <p>Note If the vNIC template has an associated QoS policy, the MTU specified has to be equal to, or less than the MTU specified in the associated QoS System class. If this MTU value exceeds the MTU value in the QoS system class, packets are dropped during data transmission.</p>
Pin Group drop-down list	This is a <i>display-only</i> field.
Adapter Policy field	This field is autopopulated with Linux .
Dynamic vNIC Connection Policy drop-down list	This is a <i>display-only</i> field.
QoS Policy drop-down list	Choose the quality of service policy that is used by the vNICs created from this vNIC policy.
Network Control Policy drop-down list	This is a <i>display-only</i> field.

Step 3 Click **Submit**.

Step 4 Click **Next**.

What to Do Next

Create a boot order policy.

Creating a Boot Order Policy

The Cisco UCS Manager boot policy overrides the boot order in the BIOS setup menu, and determines the following:

- Selection of the boot device
- Location from which the server boots
- Order in which boot devices are invoked

Step 1 On the **Create Boot Order Policy** page of the **Create UCS SP Template for Big Data** wizard, set the boot order for the following devices:

- **CD-ROM**
- **Storage**
- **LAN**

If you do not want to boot from a specific device, choose the blank space at the bottom of the drop-down list.

Note If you are booting for the first time, choose **1** for the LAN drop-down list to set it as the first boot device.

Step 2 In the **Select vNIC Policy for LAN Ethernet** table, click **Add (+)**.

Step 3 On the **Add Entry to Select vNIC Policy for LAN Ethernet** screen, do the following:

- a) From the **Select vNIC** drop-down list, choose the vNIC that you want to assign to the LAN.
- b) If you want the VLAN shown in the **VLAN** field to be the primary VLAN, check the **Set as Primary** check box.
- c) Click **Submit**.

Step 4 If you want to choose vNIC policies for other VLANs, repeat Steps 2 and 3 with a different vNIC from the **Select vNIC** drop-down list.

Step 5 Click **Next**.

What to Do Next

Create a BIOS policy.

Creating a BIOS Policy

The BIOS policy automates the configuration of certain BIOS settings for the servers in the cluster.



Note All the drop-down lists on the **Create UCS BIOS Policy** page are set to **Platform Default**—The BIOS uses the value for this attribute contained in the BIOS defaults for the server type and vendor. For more information, see [BIOS Parameters by Server Model](#).

Step 1

On the **Create UCS BIOS Policy** page of the **Create UCS SP Template for Big Data** wizard, complete the following fields:

Name	Description
Main	
Quiet Boot drop-down list	Determines what the BIOS displays during Power On Self-Test (POST).
POST Error Pause drop-down list	Determines what happens when the server encounters a critical error during POST.
Resume AC on Power Loss drop-down list	Determines the server behavior when the power is restored after an unexpected power loss.
From Panel Lockout drop-down list	Determines whether the server ignores the power and reset buttons on the front panel.
Processor	
Turbo Boost	Determines whether the processor uses Intel Turbo Boost Technology. This allows the processor to automatically increase its frequency if it is running below specifications for power, temperature, or voltage. Choose either Enabled or Disabled .
Enhanced Intel SpeedStep	Determines whether the processor uses Enhanced Intel SpeedStep Technology. This allows the system to dynamically adjust processor voltage and core frequency. Choose either Enabled or Disabled .
Hyper Threading	Determines whether the processor uses Intel Hyper-Threading Technology. This allows multithreaded software applications to execute threads in parallel within each processor. Choose either Enabled or Disabled .
Execute Disabled Bit	Classifies the memory areas on the server to specify where the application code can execute. Choose either Enabled or Disabled .

Name	Description
Virtualization Technology	<p>Determines whether the processor uses Intel Virtualization Technology (VT). This allows a platform to run multiple operating systems and applications in independent partitions.</p> <p>Note If you change this option, you must power cycle the server before the setting takes effect.</p> <p>Choose either Enabled or Disabled.</p>
Processor C State	<p>Determines whether the system enters into the power savings mode during idle periods.</p> <p>Choose either Enabled or Disabled.</p>
Processor C1E	<p>Determines whether the CPU transitions to its minimum frequency when entering the C1 state.</p> <p>Choose either Enabled or Disabled.</p>
Processor C3 Report	<p>Determines whether the BIOS sends the C3 report to the operating system. When the OS receives the report, it can transition the processor into the lower C3 power state to decrease energy usage while maintaining optimal processor performance. This option can be one of the following:</p> <ul style="list-style-type: none"> • Disabled—The BIOS does not send the C3 report. • ACPI C2—The BIOS sends the C3 report using the ACPI C2 format, allowing the OS to transition the processor to the C3 low-power state. • ACPI C3—The BIOS sends the C3 report using the ACPI C3 format, allowing the OS to transition the processor to the C3 low-power state.
Processor C6 Report	<p>Determines whether the BIOS sends the C6 report to the operating system. When the OS receives the report, it can transition the processor into the lower C6 power state to decrease energy usage while maintaining optimal processor performance.</p> <p>Choose either Enabled or Disabled.</p>
RAS Memory (reliability, availability, and serviceability of the memory)	

Name	Description
NUMA(nonuniform memory access)	<p>Determines the BIOS support for NUMA. Choose either Enabled or Disabled.</p> <ul style="list-style-type: none"> • Disabled—The BIOS does not support NUMA. • Enabled—The BIOS includes the ACPI tables that are required for NUMA-aware operating systems. If you enable this option, the system must disable Inter-Socket Memory interleaving on some platforms.

Step 2 Click Next.

What to Do Next

Create a Local Disk Configuration Policy.

Creating a Local Disk Configuration Policy

This policy defines disk configuration for disk partitions, and storage local to the server for the selected Hadoop and Splunk container types. It provides a flexible mechanism to define the RAID levels and JBOD configuration for name nodes and data nodes resident in a Hadoop cluster. This policy enables you to set a local disk configuration for all servers associated with a service profile.

Step 1 On the **Local Disk Configuration Policy** page of the **Create UCS SP Template for Big Data** wizard, review the following fields:

Name	Description
Configure RAID Policy for the Hadoop cluster table	RAID level configuration for Hadoop NameNode and DataNode (OS and data drives).
Configure Splunk RAID Policy table	RAID level configuration for Splunk Indexer, Search Heads, and Administrative nodes.
Use LVM For Disk Configuration check box	Create Logical Volume Manager (LVM) groups for disk partitions.
Partition Configuration table	Create partitions other than the /, /boot, swap, /tmp, /var/tmp, and /home partitions.

Step 2 Click Submit.

What to Do Next

- Create a Hadoop cluster profile template to deploy a Hadoop cluster.
- Create Cluster Deploy Templates for Hadoop and Splunk Enterprise clusters.

Editing RAID Policy

Use this procedure to edit the local disk drive configuration associated with the hardware RAID controller for C220 and C240 M3/M4 servers, and for the standalone boot drives on C240 M4 servers.

Before You Begin

Create a Local Disk Configuration Policy.

Step 1 On the **Local Disk Configuration Policy** page of the **Create UCS SP Template for Big Data** wizard, choose the node type in the **Configure RAID Policy** table.

Step 2 In the **Configure RAID Policy** table, click the **Edit selected entry in the table below** icon.

Step 3 On the **Edit Configure RAID Policy Entry** screen, complete the following fields:

Name	Description
Node Type	The selected node type. Note The following are the supported nodes: <ul style="list-style-type: none"> • Cluster node is applicable only for MapR cluster • Master and Data Nodes are applicable only for Cloudera, Hortonworks, and BigInsights clusters. • Kafka Node is applicable only for Cloudera and Hortonworks clusters. • Edge Node is applicable for all the Hadoop clusters.
OS Disks	
Use HDD drives on M4 servers with insufficient standalone boot drives check box	Allows you to enable an alternate policy for M4 servers with insufficient standalone boot drives.
Use standalone boot drives for M4 Servers check box	By default, this option is checked and disabled. Check this check box for M4 servers with sufficient standalone boot drives.

Name	Description
Use JBOD (Just a Bunch Of Disks) for M4 Servers check box	Check this check box for M4 servers with multiple hard drives.
RAID Level [OS] drop-down list	Choose the RAID level for the OS (operating system) disks.
Disks Per Group field	Specifies the number of disks that can exist per group during RAID configuration.
Write Mode drop-down list	Choose either Write through to write data without the RAID controller cache, or Write back to write data with the cache.
Read Mode drop down-list	Choose the method to read data from the disks.
Use Cache check box	Check the Use Cache check box to use the RAID controller cache to read and write operations.
Use Cache if Bad BBU check box	Check the Use Cache if Bad BBU check box to ensure that if the Battery Backup Unit (BBU) is not available for any reason, Write back will be disabled and Write Through will be enabled.
Strip Size (MB) drop- down list	Allows specification, in MB, of the strip size for each disk within a stripe.
Data Disks	
RAID Level [Data] drop-down list	Choose the RAID level for the Hadoop data disks.
Disks Per Group field	Specifies the number of disks that can exist per group during RAID configuration.
Write Mode drop-down list	Choose either Write through to write data without the RAID controller cache, or Write back to write data with the cache.
Read Mode drop down-list	Choose the method to read data from the disks.
Use Cache check box	Check the Use Cache check box to use the RAID controller cache to read and write operations.
Use Cache if Bad BBU check box	Check the Use Cache if Bad BBU check box to ensure that if the Battery Backup Unit (BBU) is not available for any reason, Write back will be disabled and Write Through will be enabled.

Name	Description
Strip Size (MB) drop- down	Allows specification, in MB, of the strip size for each disk within a stripe.

Step 4 Click **Submit**.

Editing RAID Policy for Splunk

Use this procedure to edit the local disk drive configuration associated with the hardware RAID controller for C220 and C240 M4 servers, and for the standalone boot drives on C240 M4 servers.

Before You Begin

- Create a UCS Service Profile Policy for the Splunk cluster deployment.
- Create a Local Disk Configuration Policy.

Step 1 On the **Local Disk Configuration Policy** page of the **Create UCS SP Template for Big Data** wizard, go to the **Configure Splunk RAID Policy** table and choose the node type.

Step 2 Click the **Edit selected entry in the table below** icon.

Step 3 On the **Edit Configure Splunk RAID Policy Entry** screen, complete the following fields:

Name	Description
Node Type	The selected node type.
OS Disks	
Use standalone boot drives for M4 Servers check box	By default, this option is checked for M4 servers with sufficient standalone boot drives.
Use HDD drives on M4 servers with insufficient standalone boot drives check box	Allows you to enable an alternate policy for M4 servers with insufficient standalone boot drives.
RAID Level [OS] drop-down list	Choose the RAID level for the OS (operating system) disks.
Disks Per Groups field	Specifies the number of disks that can exist per group during RAID configuration.
Write Mode drop-down list	Choose either Write through to write data without the RAID controller cache, or Write back to write data with the cache.

Name	Description
Read Mode drop down-list	Choose the method to read data from the disks.
Use Cache check box	Check the Use Cache check box to use the RAID controller cache to read and write operations.
Use Cache if Bad BBU check box	Check the Use Cache if Bad BBU check box to ensure that if the Battery Backup Unit (BBU) is not available for any reason, Write back will be disabled and Write Through will be enabled.
Strip Size (MB) drop- down list	Allows specification, in MB, of the strip size for each disk within a stripe.
Data Disks for Hot Data	
RAID Level [DATA] drop-down list	Choose the RAID level for the splunk data disks.
Disks Per Groups field	Specifies the number of disks that can exist per group during RAID configuration.
Write Mode drop-down list	Choose either Write through to write data without the RAID controller cache, or Write back to write data with the cache.
Read Mode drop down-list	Choose the method to read data from the disks.
Use Cache check box	Check the Use Cache check box to use the RAID controller cache to read and write operations.
Use Cache if Bad BBU check box	Check the Use Cache if Bad BBU check box to ensure that if the Battery Backup Unit (BBU) is not available for any reason, Write back will be disabled and Write Through will be enabled.
Strip Size (MB) drop- down	Allows specification, in MB, of the strip size for each disk within a stripe.
Data Disks for Cold Data	
Same Disk As Hot Data check box	Check the check box to enable that the same disk is configured for Hot Data to also host Cold Data. Note By default, all RAID configurations are available in Data Disks for Cold Data . Checking this check box hides all RAID configuration for cold data.
Data Disks for Frozen Data	

Name	Description
Same Disk As Cold Data check box	<p>Check the check box to enable that the same disk is configured for Cold Data to also host Frozen Data.</p> <p>Note By default, all RAID configurations are available in Data Disks for Frozen Data. Checking this check box hides all RAID configuration for frozen data.</p>

- Note**
- The RAID configuration of Data disks for hot, cold, or frozen data is applicable only for Indexer nodes. The RAID configuration of OS disks applies to Search Head and Administrative nodes.
 - One RAID group is created for each RAID configuration for hot, cold, or frozen data (for example, /data/disk1 for Hot, /data/disk2 for Cold, and /data/disk3 for Frozen).

Configuring Local Disk Partitions

Each partition in the disk has two attributes, mount point and size. You can create, edit, and delete partitions as per your requirements. The **Partition Configuration** table displays the /, /boot, swap, /tmp, /var/tmp, and /home partitions by default. You can update the size allocated for these partitions except for the root (/) and boot (/boot) partitions, but cannot delete the entries.

Table 7: Sample Partitions Table

Mount Point	Size	Editable
/	1 GB with grow	Not Editable
/boot	1024 MB	Not Editable
Swap	Any value	Editable
/tmp	5 GB	Editable
/var/tmp	5 GB	Editable
/home	5 GB	Editable

-
- Step 1** On the **Local Disk Configuration Policy** page of the **Create UCS SP Template for Big Data** wizard, click **Add** in the **Partition Configuration** table.
- Step 2** On the **Add Partition Configuration Entry** screen, enter the mount name in the **Mount Point** field.
- Step 3** Enter the size in the **Size** field.
- Step 4** Click **Submit**.
-

Creating a Customized Service Profile Template

- Step 1** Choose **Physical > Compute**.
- Step 2** In the left pane, expand the pod and then click the Cisco UCS Manager account.
- Step 3** In the right pane, click **VLANs**.
- Step 4** Click **Add (+)** to add three VLANs for creating three vNIC templates.
- Step 5** In the right pane, click **Organizations**.
- Step 6** Select the organization in which you want to create the service profile template and then click **View Details**.
- Step 7** Click **vNIC Template** and click **Add (+)**.
- Step 8** Create three vNIC templates using the created VLANs.
- Step 9** Click **Boot Policy** and click **Add (+)** to create a Boot policy by adding a local disk.
- Step 10** Click **QoS Policy** and click **Add (+)** to create a QoS policy.
- Step 11** Choose **Policies > Physical Infrastructure Policies > UCS Manager**.
- Step 12** Click **Storage Policy** and specify the Cisco UCS connections for the storage policy.
- Step 13** Click **Network Policy** and specify the required details.
- Step 14** Click **vNIC** and create three vNICs by mapping the vNIC templates. The name of the vNIC should be eth0, eth1, and eth2 respectively. After the service profile template creation, vNIC name will be generated as eth0-1, eth1-2, and eth2-3 respectively.
- Step 15** Click **Placement Policy** and specify the required details.
- Step 16** Choose **Physical > Compute > Organizations**.
- Step 17** Open the root and click **Service Profile Template**.
- Step 18** Click **Add (+)** and choose the created policies and provide the required information.
- Step 19** Click **Submit**.
- Step 20** Log on to UCSM. Launch the created service profile template and update the desired placement as 1, 2, and 3 for the three vNIC templates respectively.
- Step 21** Click **Save Changes**.
- Note** In the Storage module, select the **No vHBAs** option in the **How would you like to configure SAN connectivity?** field. This option does not allow you to create any vHBAs. If you choose this option, any server associated with a service profile that includes this policy is not connected to the SAN.

Cloning a Cisco UCS Service Profile Template

- Step 1** Choose **Solutions > Big Data > Containers**.
- Step 2** Click **UCS SP Templates for Big Data**.
- Step 3** Click the row for the template that you want to clone.
- Step 4** Click **Clone**.
- Step 5** In the **UCS SP Template Specification** page of the Clone UCS SP Template for Big Data wizard, do the following:
- Enter a unique name and description for the new service profile template.
 - Choose the Template Type from the drop-down list.
 - Click **Next**, review the information on each page, and modify if necessary.
 - Click **Submit**.
-



CHAPTER 8

Configuring and Deploying Hadoop Cluster Deployment Templates

This chapter contains the following sections:

- [Hadoop Cluster Profile Templates, page 61](#)
- [Creating a Hadoop Cluster Profile Template, page 63](#)
- [Cloning a Hadoop Cluster Profile Template, page 76](#)
- [Creating a Cluster Deployment Template, page 76](#)

Hadoop Cluster Profile Templates

The Hadoop cluster profile template specifies the number of nodes in the cluster. The template also takes care of provisioning and configuring the Hadoop cluster services. Apache Software Foundation projects around the world develop services for Hadoop deployment and integration. Some Hadoop distributions support only a subset of these services, or have their own distribution-specific services.

Each of the following supplies a dedicated function:



Note

You cannot uncheck some of the services because they are necessary to create a Hadoop cluster. All mandatory services are checked by default.

Hadoop Services	Cloudera	MapR	Hortonworks	BigInsights
HDFS	Yes	—	Yes	Yes
CLDB	—	Yes	—	—
YARN/MapReduce	Yes	Yes	Yes	Yes
ZooKeeper	Yes	Yes	Yes	Yes
HBase	Yes	Yes	Yes	Yes

Hadoop Services	Cloudera	MapR	Hortonworks	BigInsights
Hive	Yes	Yes	Yes	Yes
Oozie	Yes	Yes	Yes	Yes
Hue	Yes	—	—	Yes
Spark	Yes	Yes	Yes	Yes
Key-Value Store Indexer	Yes	—	—	Yes
Solr	Yes	—	—	Yes
Sqoop	Yes	Yes	Yes	Yes
Impala	Yes	—	—	Yes
Flume	Yes	Yes	—	Yes
PIG	—	Yes	Yes	Yes
MAHOUT	—	Yes	—	Yes
Falcon	—	—	Yes	Yes
Tez	—	—	Yes	Yes
Storm	—	—	Yes	Yes
Ganglia/Ambari Metrics	—	—	Yes	Yes
Drill	—	Yes	—	—
Kafka	Yes	—	Yes	—

Creating a Hadoop Cluster Profile Template

Before You Begin

Create a Hadoop cluster Configuration Template.

- Step 1** On the menu bar, choose **Solutions > Big Data > Containers**.
- Step 2** On the **Containers** page, click **Hadoop Cluster Profile Templates**.
- Step 3** Click **Add (+)**.
- Step 4** On the **Hadoop Cluster Profile Template** page of the **Create Hadoop Cluster Profile Template** wizard, complete the following fields:

Name	Description
Template Name field	A unique name for the template.
Template Description field	A short description for the template.
Node Count field	The number of nodes in the cluster. The default is four nodes.
Hadoop Distribution drop-down list	The type of Hadoop distribution. The Hadoop cluster services are displayed based on the selected Hadoop distribution.
Hadoop Distribution Version	Choose the Hadoop distribution version.
Hadoop Cluster Configuration Parameters Template drop-down list	Choose the cluster configuration parameters template.
Secure check box	Check this check box if you want to enable security for the cluster. This option is applicable only for MapR cluster.
Storage Pool Disk Grouping Configuration	
No.of disks per pool for the servers with disks 1-12 field	The number of disks per pool for the servers with disks ranging from 1 to 12. This field is displayed only for MapR cluster. You need to enter the value before creating a new node in the MapR cluster.
No.of disks per pool for the servers with disks 13-24 field	The number of disks per pool for the servers with disks ranging from 13 to 24. This field is displayed only for MapR cluster. You need to enter the value before creating a new node in the MapR cluster.

Name	Description
No.of disks per pool for the servers with disks >24 field	The number of disks per pool for the servers with disks ranging from 24 to 96. This field is displayed only for MapR cluster. You need to enter the value before creating a new node in the MapR cluster.

Step 5 Click Next.

What to Do Next

Create a Services Selection policy.

Creating a Services Selection Policy

The cluster policy contains the Hadoop cluster services that you want to enable in the Hadoop cluster.



Note The **Service Selection Page** displays the Hadoop cluster services, depending on the Hadoop distribution already selected on the **Hadoop Cluster Profile Template** page.

Step 1 On the **Services Selection Policy** page of the **Create Hadoop Cluster Profile Template** wizard, do the following:

Step 2 Check the check box for the optional Hadoop cluster services that you want to enable in your cluster. Some Hadoop cluster services are required for the distribution and cannot be disabled. The available Hadoop cluster services include the following:

- **HDFS**—A file system that spans all nodes in a Hadoop cluster for data storage. This service replicates data across multiple nodes to avoid data loss.
- **YARN**—A resource-management platform responsible for managing compute resources in clusters and using them for scheduling your applications.
- **HBase**—A high-speed read and write column-oriented database.
- **Hive**—The query engine framework for Hadoop that facilitates easy data summarization, ad-hoc queries, and the analysis of large data sets stored in HDFS and HBase. With SQL-like semantics, Hive makes it easy for RDBMS users to transition into querying unstructured data in Hadoop.
- **Oozie**—A workflow environment for coordinating complex data processing operations.
- **ZooKeeper**—An infrastructure for cross-node synchronization. The applications ensure that tasks across the cluster are serialized or synchronized.
- **Hue**—An interface that aggregates the most common Hadoop components to improve user experience. This allows you to avoid the underlying complexity of the system, and bypasses the command-line interface.
- **Spark**—An open-source data analytics engine.

- **Key-Value Store Indexer**—A method for indexing data across the cluster.
- **SOLR**—A method for searching data across the cluster.
- **Sqoop**—A client-server tool that transfers bulk data between Hadoop and structured data stores, such as relational databases.
- **Impala**—A massively parallel processing (MPP) SQL query engine that runs natively in Apache Hadoop.
- **Flume**—A distributed, reliable, and available service for efficiently collecting, aggregating, and moving large amounts of streaming data into the Hadoop Distributed File System (HDFS).
- **Kafka**—A fast, scalable, durable, and fault-tolerant publish-subscribe messaging system.

Step 3 Click Next.

What to Do Next

Configure the Rack Assignment policy.

Configuring the Rack Assignment Policy

Step 1 On the **Rack Assignment Policy** page of the **Create Hadoop Cluster Profile Template** wizard, you can:

- Create one or more Hadoop node configuration policies. Click **Add (+)**, and continue with Step 2.
- Modify the default node configuration policy. Choose the default policy in the table. Click **Edit**, and continue with Step 2.

Step 2 On the **Add Entry to Hadoop Node Configuration Policy** screen, do the following:

- a) In the **Rack Name** field, enter the name of the rack server.
- b) In the **DataNodes** field, click **Select** and check the check box for each node that you want to configure on that server.
Note Some Hadoop cluster services require a minimum number of nodes. For example, ZooKeeper requires a minimum of three nodes.
- c) Click **Submit**.

Step 3 Click Next.

What to Do Next

Configure the HDFS policy.

Configuring the HDFS Policy

-
- Step 1** On the **HDFS Policy** page of the **Create Hadoop Cluster Profile Template** wizard, do the following:
- Step 2** Click the row in the table with the node for which you want to change the HDFS policy configuration, and click **Edit**.
- Step 3** On the **Edit HDFS Policy Entry** screen, review and, if necessary, change the following fields:
- a) Choose **Yes** from the **DataNode** drop-down list if you want the node to act as the DataNode for HDFS. The data nodes store and retrieve data on request by the name node or by the client.

Note The node that act as a DataNode for HDFS is not allocated to Node1, Node2, and Node3 when the

 - node count is greater than 5 for Cloudera
 - node count is greater than 3 for Hortonworks.
 - b) Choose **Yes** from the **Primary NameNode** drop down-list if you want the node to act as the primary name node for HDFS. The primary name node maintains all the operations of the HDFS cluster. There can be only one primary name node for the HDFS cluster.
 - c) Choose **Yes** from the **Secondary NameNode** drop down-list if you want the node to act as a secondary name node for HDFS. The secondary name node is not a direct replacement for the primary name node. The main role of a secondary name node is to periodically merge the FSImage and edit log, to prevent the edit log from becoming too large. A secondary name node runs on a separate physical system because it requires more memory to merge two files. It keeps a copy of the merged file in its local file system so that it is available for use if the primary name node fails.
 - d) Choose **Yes** from the **Balancer** drop down-list if you want the node to act as a balancer for HDFS.
 - e) Choose **Yes** from the **HTTPFS** drop down-list if you want the node to act as HTTPFS for HDFS. This service provides HTTP access to HDFS.
 - f) Choose **Yes** from the **Fail Over Controller** drop down-list if you want the node to act as Fail Over Controller for HDFS.
 - g) Choose **Yes** from the **Gateway** drop down-list if you want the node to act as Gateway for HDFS.
- Step 4** Click **Submit**.
- Step 5** Repeat Steps 1 and 2 to configure the other nodes for HDFS.
- Step 6** Click **Next**.
-

What to Do Next

Configure the CLDB policy.

Configuring the CLDB Policy

- Step 1** On the **CLDB Policy** page of the **Create Hadoop Cluster Profile Template** wizard, do the following:
- Step 2** Click the row in the table with the node for which you want to change the CLDB policy configuration, and click **Edit**.
- Step 3** On the **Edit CLDB Policy Entry** screen, choose **Yes** if you want the node to act as a CLDB agent.
- Step 4** Click **Submit**.
- Step 5** Repeat Steps 1 and 2 to configure the other nodes for CLDB.
- Step 6** Click **Next**.
-

What to Do Next

Configure the YARN policy.

Configuring the YARN Policy

- Step 1** On the **YARN Policy** page of the **Create Hadoop Cluster Profile Template** wizard, do the following:
- Step 2** Click the row in the table with the node for which you want to change the YARN policy configuration, and click **Edit**.
- Step 3** On the **Edit YARN Policy Entry** screen, review and, if necessary, change the following fields:
- Choose **Yes** from the **Resource Manager** drop-down list if you want the node to act as a Resource Manager. The Resource Manager is the ultimate authority that allocates resources among all the applications in the system.
 - Choose **Yes** from the **Node Manager** drop down-list if you want the node to act as a task Node Manager. The Node Manager is responsible for launching the applications' containers, monitoring their resource usage (CPU, memory, disk, network), and reporting to the Resource Manager.
Note The node that act as a DataNode for HDFS is not allocated to Node1, Node2, and Node3 when the
 - node count is greater than 5 for Cloudera
 - node count is greater than 3 for Hortonworks.
 - Choose **Yes** from the **Gateway** drop down-list if you want the node to act as a Gateway.
 - Choose **Yes** from the **JobHistory** drop down-list if you want the node to preserve the Job History.
 - Click **Submit**.
- Step 4** Click **Submit**.
- Step 5** Repeat Steps 1 and 2 to configure the other nodes for Yarn.
- Step 6** Click **Next**.
-

What to Do Next

Configure the ZooKeeper policy.

Configuring the ZooKeeper Policy



Note Configure a minimum of three nodes for ZooKeeper.

-
- Step 1** On the **ZooKeeper Policy** page of the **Create Hadoop Cluster Profile Template** wizard, do the following:
 - Step 2** Click the row in the table with the node for which you want to change the ZooKeeper policy configuration, and click **Edit**
 - Step 3** On the **Edit ZooKeeper Policy Entry** screen, choose **Yes** to make the node to act as a ZooKeeper.
 - Step 4** Click **Submit**.
 - Step 5** Repeat Steps 1 and 2 to configure the other nodes for ZooKeeper.
 - Step 6** Click **Next**.
-

What to Do Next

Configure the HBase policy.

Configuring the Kafka Policy

-
- Step 1** On the **Kafka Policy** page of the **Create Hadoop Cluster Profile Template** wizard, do the following:
 - Step 2** Click the row in the table with the node for which you want to change the Kafka policy configuration, and click **Edit**.
 - Step 3** On the **Edit Kafka Policy Entry** screen, choose **Yes** from the **Kafka Broker** drop-down list if you want the node to act as a Kafka broker.
 - Note** All nodes are Kafka nodes, only if **Kafka** and **Zookeeper** check boxes are selected in the **Services Selection Policy** page while creating Cloudera or Hortonworks clusters. If it is a HDFS and Kafka cluster, then last two nodes of the template are selected for kafka by default.
 - Note** Addition of Kafka role is not supported in Clouder and Hortonworks clusters.
 - Step 4** Click **Submit**.
 - Step 5** Repeat Steps 1 and 2 to configure the other nodes for Kafka.
 - Step 6** Click **Next**.
-

Configuring the HBase Policy

-
- Step 1** On the **HBase Policy** page of the **Create Hadoop Cluster Profile Template** wizard, do the following:
- Step 2** Click the row in the table with the node for which you want to change the HBase policy configuration, and click **Edit**.
- Step 3** On the **Edit HBase Policy Entry** screen, review and, if necessary, change the following fields:
- Choose **Yes** from the **HBase Master** drop-down list if you want the node to act as a HBase master.
 - Choose **Yes** from the **Region Server** drop down-list if you want the node to act as a region server.
Note The node that act as a DataNode for HDFS is not allocated to Node1, Node2, and Node3 when the
 - node count is greater than 5 for Cloudera
 - node count is greater than 3 for Hortonworks.
 - Choose **Yes** from the **HBase Thrift Server** drop down-list if you want the node to act as a HBase Thrift.
 - Click **Submit**.
- Step 4** Click **Submit**.
- Step 5** Repeat Steps 1 and 2 to configure the other nodes for HBase.
- Step 6** Click **Next**.
-

What to Do Next

Configure the Hive policy.

Configuring the Hive Policy

-
- Step 1** On the **Hive Policy** page of the **Create Hadoop Cluster Profile Template** wizard, do the following:
- Step 2** Click the row in the table with the node for which you want to change the Hive policy configuration, and click **Edit**.
- Step 3** On the **Edit Hive Policy Entry** screen, review and, if necessary, change the following fields:
- Choose **Yes** from the **HiveServer2** drop-down list if you want the node to act as a HiveServer2.
 - Choose **Yes** from the **Hive Metastore Server** drop down-list if you want the node to act as a Hive metastore.
 - Choose **Yes** from the **WebHCat** drop down-list if you want the node to act as a WebHCat. WebHCat is the REST API for HCatalog, a table and storage management layer for Hadoop.
 - Choose **Yes** from the **Gateway** drop down-list if you want the node to act as a Gateway for Hive.
 - Click **Submit**.
- Step 4** Click **Submit**.
- Step 5** Repeat Steps 1 and 2 to configure the other nodes for Hive.
- Step 6** Click **Next**.
-

What to Do Next

Configure the Oozie policy.

Configuring the Oozie Policy

- Step 1** On the **Oozie Policy** page of the **Create Hadoop Cluster Profile Template** wizard, do the following:
- Step 2** Click the row in the table with the node for which you want to change the Oozie policy configuration, and click **Edit**.
- Step 3** On the **Edit Oozie Policy Entry** screen, choose **Yes** to make the node to act as an Oozie server.
- Step 4** Repeat Steps 1 and 2 to configure the other nodes for Oozie.
- Step 5** Click **Next**.
-

What to Do Next

Configure the Hue policy.

Configuring the Hue Policy

- Step 1** On the **Hue Policy** page of the **Create Hadoop Cluster Profile Template** wizard, do the following:
- Step 2** Click the row in the table with the node for which you want to change the Hue policy configuration, and click **Edit**.
- Step 3** On the **Edit Hue Policy Entry** screen, do the following:
- Choose **Yes** from the **Hue Server** drop-down list if you want the node to act as a Hue server.
 - Choose **Yes** from the **BeesWax Server** drop down-list if you want the node to act as a BeesWax server.
 - Choose **Yes** from the **Kt Renewer** drop down-list if you want the node to act as a Kt Renewer.
 - Click **Submit**.
- Step 4** Repeat Steps 1 and 2 to configure the other nodes for Hue.
- Step 5** Click **Next**.
-

What to Do Next

Configure the Spark policy.

Configuring the Spark Policy

- Step 1** On the **Spark Policy** page of the **Create Hadoop Cluster Profile Template** wizard, do the following:
- Step 2** Click the row in the table with the node for which you want to change the Spark policy configuration, and click **Edit**.
- Step 3** On the **Edit Spark Policy Entry** screen, review and, if necessary, change the following fields:
- Choose **Yes** from the **History Server** drop-down list if you want the node to act as a History Server.
 - Choose **Yes** from the **Gateway** drop down-list if you want the node to act as a gateway.
- Step 4** Click **Submit**.
- Step 5** Repeat Steps 1 and 2 to configure the other nodes for Spark.
- Note** When Spark applications run on a YARN cluster manager, resource management, scheduling, and security are controlled by YARN. In Cloudera cluster, configuring th Spark on Yarn policy is supported.
- Step 6** Click **Next**.
-

What to Do Next

Configure the Key-Value Store Indexer policy.

Configuring the Key-Value Store Indexer Policy

- Step 1** On the **Key-Value Store Indexer Policy** page of the **Create Hadoop Cluster Profile Template** wizard, do the following:
- Step 2** Click the row in the table with the node for which you want to change the Key-Value Store Indexer policy configuration, and click **Edit**.
- Step 3** On the **Edit KSIndexer Policy Entry** scree, choose **Yes** if you want the node to act as a KSIndexer server.
- Step 4** Click **Submit**.
- Step 5** Repeat Steps 1 and 2 to configure the other nodes for KSIndexer.
- Step 6** Click **Next**.
-

What to Do Next

Configure the Solr policy.

Configuring the Solr Policy

- Step 1** On the **Solr Policy** page of the **Create Hadoop Cluster Profile Template** wizard, do the following:
- Step 2** Click the row in the table with the node for which you want to change the Solr policy configuration, and click **Edit**.
- Step 3** On the **Edit Solr Policy Entry** screen, choose **Yes** if you want the node to act as a Solr server.
- Step 4** Click **Submit**.
- Step 5** Repeat Steps 1 and 2 to configure the other nodes for Solr.
- Step 6** Click **Next**.
-

What to Do Next

Configure the Sqoop policy.

Configuring the Sqoop Policy

- Step 1** On the **Sqoop Policy** page of the **Create Hadoop Cluster Profile Template** wizard, do the following:
- Step 2** Click the row in the table with the node for which you want to change the Sqoop policy configuration, and click **Edit**.
- Step 3** On the **Edit Sqoop Policy Entry** screen, choose **Yes** if you want the node to act as a Sqoop server.
- Step 4** Click **Submit**.
- Step 5** Repeat Steps 1 and 2 to configure the other nodes for Sqoop.
- Step 6** Click **Next**.
-

What to Do Next

Configure the Impala policy.

Configuring the Impala Policy

- Step 1** On the **Impala Policy** page of the **Create Hadoop Cluster Profile Template** wizard, do the following:
- Step 2** Click the row in the table with the node for which you want to change the Impala policy configuration, and click **Edit**.
- Step 3** On the **Edit Impala Policy Entry** screen, do the following:
- Choose **Yes** from the **Impala Daemon** drop-down list if you want the node to act as an Impala daemon.
 - Choose **Yes** from the **Impala StateStore** drop-down list if you want the node to act as an Impala Statestore.
 - Choose **Yes** from the **Impala Catalog Server** drop-down list if you want the node to act as an Impala catalog server. The other fields in this dialog box are for your information only.

d) Click **Submit**.

Step 4 Repeat Steps 1 and 2 to configure the other nodes for Impala.

Step 5 Click **Submit**.

What to Do Next

Configure the Flume policy.

Configuring the Flume Policy

Step 1 On the **Flume Policy** page of the **Create Hadoop Cluster Profile Template** wizard do the following:

Step 2 Click the row in the table with the node for which you want to change the Flume policy configuration, and click **Edit**.

Step 3 On the **Edit Flume Policy Entry** screen, choose **Yes** if you want the node to act as a Flume agent.

Step 4 Click **Submit**.

Step 5 Repeat Steps 1 and 2 to configure the other nodes for Flume.

Step 6 Click **Next**.

What to Do Next

Configure the PIG Policy.

Configuring the PIG Policy

Step 1 On the **Pig Policy** page of the **Create Hadoop Cluster Profile Template** wizard, do the following:

Step 2 Click the row in the table with the node for which you want to change the Pig policy configuration, and click **Edit**.

Step 3 On the **Edit Pig Policy Entry** screen, choose **Yes** if you want the node to act as a Pig agent.

Step 4 Click **Submit**.

Step 5 Repeat Steps 1 and 2 to configure the other nodes for Pig.

Step 6 Click **Next**.

What to Do Next

Configure the MAHOUT Policy.

Configuring the MAHOUT Policy

- Step 1** On the **MAHOUT Policy** page of the **Create Hadoop Cluster Profile Template** wizard, do the following:
- Step 2** Click the row in the table with the node for which you want to change the MAHOUT policy configuration, and click **Edit**.
- Step 3** On the **Edit MAHOUT Policy Entry** screen, choose **Yes** if you want the node to act as a MAHOUT agent.
- Step 4** Click **Submit**.
- Step 5** Repeat Steps 1 and 2 to configure the other nodes for MAHOUT.
- Step 6** Click **Submit**.
-

What to Do Next

Configure a Falcon Policy.

Configuring the Falcon Policy

- Step 1** On the **Falcon Policy** page of the **Create Hadoop Cluster Profile Template** wizard, do the following:
- Step 2** Click the row in the table with the node for which you want to change the Falcon policy configuration, and click **Edit**.
- Step 3** On the **Edit Falcon Policy Entry** dialog box, choose **Yes**. You can make the node act as a Falcon server from the **Falcon Server** and the Falcon client from **Falcon Client** drop-down lists.
- Step 4** Click **Submit**.
- Step 5** Repeat Steps 1 and 2 to configure the other nodes for Falcon.
- Step 6** Click **Submit**.
-

What to Do Next

Configure the Tez Policy.

Configuring the Tez Policy

- Step 1** On the **Tez Policy** page of the **Create Hadoop Cluster Profile Template** wizard, do the following:
- Step 2** Click the row in the table with the node for which you want to change the Tez policy configuration, and click **Edit**.
- Step 3** On the **Edit Tez Policy Entry** screen, choose **Yes** if you want the node to act as a Tez agent.
- Step 4** Click **Submit**.
- Step 5** Repeat Steps 1 and 2 to configure the other nodes for Tez.
- Step 6** Click **Submit**.
-

What to Do Next

Configure the Storm Policy.

Configuring the Storm Policy

- Step 1** On the **Storm Policy** page of the **Create Hadoop Cluster Profile Template** wizard, do the following:
- Step 2** Click the row in the table with the node for which you want to change the Storm policy configuration, and click **Edit**.
- Step 3** On the **Edit Storm Policy Entry** screen, do the following:
- Choose **Yes** in the **DRPC Server** drop-down list if you want the node to act as a DRPC server.
 - Choose **Yes** in the **Nimbus** drop-down list if you want the node to act as a Nimbus server.
 - Choose **Yes** in the **Storm REST API Server** drop-down list if you want the node to act as a Storm REST API server.
 - Choose **Yes** in the **Storm UI Server** drop-down list if you want the node to act as a Storm UI server.
 - Choose **Yes** in the **Supervisor** drop-down list if you want the node to act as a supervisor.
- Step 4** Click **Submit**.
- Step 5** Repeat Steps 1 and 2 to configure the other nodes for Storm.
- Step 6** Click **Submit**.
-

What to Do Next

Configure the Ganglia Policy.

Configuring the Ganglia Policy

-
- Step 1** On the **Ganglia Policy** page of the **Create Hadoop Cluster Profile Template** wizard, do the following:
- Step 2** Click the row in the table with the node for which you want to change the Ganglia policy configuration, and click **Edit**.
- Step 3** On the **Edit Ganglia Policy Entry** screen, choose **Yes**. You can make the node to act as a Ganglia server from the **Ganglia Server** and the Ganglia monitor from the **Ganglia Monitor** drop-down lists.
- Step 4** Click **Submit**.
- Step 5** Repeat Steps 1 and 2 to configure the other nodes for Ganglia.
- Step 6** Click **Submit**.
-

Cloning a Hadoop Cluster Profile Template

-
- Step 1** On the menu bar, choose **Solutions > Big Data > Containers**.
- Step 2** On the **Containers** page, click the **Hadoop Cluster Profile Templates**.
- Step 3** Click the row for the template that you want to clone.
- Step 4** Click **Clone**.
- Step 5** On the **Clone Hadoop Cluster Profile Template** screen, do the following:
- Enter a unique name and description for the new Hadoop cluster profile template.
 - Click **Next**, review the information on each page, and modify, if necessary.
 - Click **Submit**.
-

Creating a Cluster Deployment Template

Before You Begin

- Create a Cisco UCS Service Profile Template for Big Data

- Create a Hadoop Cluster Profile Template

Step 1 Choose **Solutions > Big Data > Containers**.

Step 2 Click **Cluster Deploy Templates**.

Step 3 Click **Add (+)**.

Step 4 On the **Add Cluster Deploy Template** screen, complete the following fields:

Name	Description
Template Name field	Enter a unique name for the Hadoop cluster deployment template.
Description field	Enter a short description of the template.
Container Type drop-down list	Choose the type of container for the cluster.
Select UCS Template drop-down list	Choose the UCS service profile template for Big Data that you want to use in the Hadoop cluster. Note If you choose Splunk as the container type, choose the UCS service profile template for Big Data with Splunk software to create a Splunk cluster.
Hadoop Cluster Profile Template drop-down list	Choose the Hadoop cluster profile template that you want to use. This option is displayed when you choose the container type as Hadoop.

Step 5 Click **Add**.



Managing Hadoop Clusters

This chapter contains the following sections:

- [Creating an Instant Hadoop Cluster, page 79](#)
- [Creating a Customized Hadoop Cluster, page 83](#)
- [Creating a Hadoop Cluster Using Workflow, page 86](#)
- [Provisioning an Instant and Customized Hadoop Cluster, page 87](#)
- [Managing a Hadoop Cluster, page 89](#)
- [Managing Nodes in a Cluster, page 94](#)
- [Delete Node and Delete Node to Bare Metal Actions in Cloudera and Hortonworks, page 95](#)
- [Adding Managed Nodes to the Hadoop Cluster, page 96](#)
- [Adding Live Nodes to the Hadoop Cluster, page 97](#)
- [Adding Bare Metal Nodes to the Hadoop Cluster, page 97](#)
- [Adding Disks to the Hadoop Cluster, page 100](#)
- [Service Roles, page 101](#)

Creating an Instant Hadoop Cluster

Before You Begin

- Create a service profile template .
- Create a server pool in the Cisco UCS Manager account that you plan to use for this cluster. See [Cisco UCS Director Management Guide for Cisco UCS Manager](#).

- Create a MAC address pool .

Step 1 Choose **Solutions > Big Data > Containers**.

Step 2 Click **Cluster Deploy Templates**.

Step 3 Click **Instant Hadoop Cluster**.

Step 4 On the **Instant Hadoop Cluster Creation** screen, complete the following fields.

Name	Description
Big Data Account Name field	Enter the name of the Big Data account.
UCSM Policy Name Prefix field	Enter the UCSM Policy Name prefix.
Hadoop Cluster Name field	Enter a unique name for the Hadoop cluster.
Hadoop Node Count field	Enter the number of nodes in the Hadoop cluster.
SSH (root) Password field	Enter the SSH root password. Note The SSH username pertains to the root user.
Confirm SSH Password field	Enter the SSH root password.
Hadoop Manager Password field	Enter the management console password.
Confirm Hadoop Manager Password field	Enter the management console password.
Host Node Prefix field	Enter the Host Node prefix for the cluster.
OS Version drop-down list	Choose the operating system to be installed on the servers in this cluster.
Hadoop Distribution drop-down list	Choose the Hadoop distribution to be used for this cluster.
Hadoop Distribution Version drop-down list	Choose the Hadoop distribution version.
Oracle JDK drop-down list	Choose the Oracle JDK version.
External Database drop-down list	Choose an external database. You can also configure a new database from here.

Name	Description
Multi-UCSM check box	<p>Check the Multi-UCSM check box if you use multiple UCSM accounts.</p> <p>Note If you use the multiple UCSM accounts option, you can configure the Hadoop Server Roles as described in Step 5. You can add UCSM Specific Inputs in the Add Entry to UCSM Specific Inputs table.</p> <p>The following workflows are established during an Instant and Customized Hadoop Cluster creation:</p> <ul style="list-style-type: none"> • UCS CPA Multi-UCSM Hadoop Cluster WF • Single UCSM Server Configuration WF. (This WF is triggered per UCSM Account. For example, UCSM 120, UCSM121.) • UCS CPA Node Bare Metal. (This WF is triggered per Node.) <p>The UCSM Specific Input area is displayed when the Multi-UCSM check box is checked.</p>
UCS Manager Account drop-down list	Choose the Cisco UCS Manager account for this cluster.
Organization drop-down list	Choose the organization in which the servers for this cluster are located.
UCS SP Template table	Choose an existing UCS Service Profile Template for cluster creation.
SSD Boot Drives Available for OS check box	<p>Check this check box if you do not want to validate the server disk availability for RAID level OS disks. Ensure that the servers contain Solid-State Drive (SSD).</p> <p>If the check box is not selected, the disk availability for both the OS disk and data disk are validated based on their RAID level.</p>
PXE VLAN ID field	Enter the PXE VLAN ID.
UCS SP Template table	Choose an existing UCS Service Profile Template for Hadoop cluster creation.

Step 5 If you want to edit a Hadoop Server Role, select the row for that role, and click **Edit**.

Step 6 On the **Edit Hadoop Server Roles Entry** screen, complete the following fields and click **Submit**.

Name	Description
Node Type field	Displays the Hadoop node role. Note Kafka node is supported only for Cloudera and Hortonworks clusters.
Node Count field	The number of nodes in the Hadoop cluster for the selected node type.
SSD Boot Drives Available for OS check box	Check this check box if you do not want to validate the server disk availability for RAID level OS disks. Ensure that the servers contain SSD. If the check box is not selected, the disk availability for both the OS disk and data disk are validated based on their RAID level. Note This check box is not displayed when the UCSM version is greater than or equal to 3.
Server Pool table	Enter the server pool that you want to use for the cluster for the selected node type. The Cisco UCS Manager account and the organization that you choose determine which server pools are displayed in this area.

Step 7 In the **vNIC Template** table, review and, if necessary, edit the vNIC templates available for the cluster.

Step 8 If you want to edit a vNIC template, select the row for that template and click **Edit**.

Step 9 On the **Edit vNIC Template Entry** screen, complete the following fields and click **Submit**.

Name	Description
vNIC Name drop-down list	The vNIC name in the selected template. This field is for your information only.
IP Pool drop-down list	Choose the big data IP pool that you want to use for IP addresses assigned to this vNIC.
MAC Address Pool drop-down list	Choose the MAC address pool that you want to use for this cluster. (This drop-down list is disabled if an existing UCS SP Template is selected.)
First MAC Address field	Enter the MAC address pool that you want to use for this cluster.
Size field	Enter the size. (This field is disabled if an existing UCS SP Template is selected.)

Name	Description
VLAN ID field	Enter the VLAN ID for this cluster. (This field is disabled if an existing UCS SP Template is selected.)

Step 10 Click **Submit**.

What to Do Next

You can view and monitor the workflow that is triggered after you create an instant Hadoop cluster.

Creating a Customized Hadoop Cluster

Before You Begin

- Create a service profile template.
- Create a Hadoop cluster profile template.
- Setup the details for Hadoop Config Parameters.
- Create a Hadoop cluster deployment template that Cisco UCS Director Express for Big Data uses to create the Hadoop cluster.
- Create a server pool in the Cisco UCS Manager account you plan to use for this cluster. See [Cisco UCS Director Management Guide for Cisco UCS Manager](#).
- Create a MAC address pool.

Step 1 Choose **Solutions > Big Data > Containers**.

Step 2 Click **Cluster Deploy Templates**.

Step 3 Select the template that you want to use for the Hadoop cluster and click **Customized Hadoop Cluster**.

Step 4 On the **Customized Hadoop Cluster Creation** screen, complete the following fields.

Name	Description
Big Data Account Name field	Enter the name of the Big Data account.
UCSM Policy Name Prefix field	Enter the UCSM Policy Name prefix.
Hadoop Cluster Name field	Enter a unique name for the Hadoop cluster.
Hadoop Node Count field	Enter the number of nodes in the Hadoop cluster.

Name	Description
SSH (root) Password field	Enter the SSH root password. Note The SSH username pertains to the root user.
Confirm SSH Password field	Enter the SSH root password.
Hadoop Manager Password field	Enter the management console password.
Confirm Hadoop Manager Password field	Enter the management console password.
Host Node Prefix field	Enter the Host Node prefix for the cluster.
OS Version drop-down list	Choose the operating system to be installed on the servers in this cluster.
Hadoop Distribution drop-down list	Choose the Hadoop distribution to be used for this cluster.
Hadoop Distribution Version drop-down list	Choose the Hadoop distribution version.
Oracle JDK drop-down list	Choose the Oracle JDK version.
External Database drop-down list	Choose an external database. You can also configure a new database from here.
Multi-UCSM check box	Click the Multi-UCSM check box if you use multiple UCSM accounts. Note If you use the multiple UCSM accounts option, you can configure the Hadoop Server Roles as described in Step 5. You can add UCSM Specific Inputs in the Add Entry to UCSM Specific Inputs table. The following workflows are created during an Instant Hadoop cluster creation and Customized Hadoop cluster creation: <ul style="list-style-type: none">• UCS CPA Multi-UCSM Hadoop Cluster WF• Single UCSM Server Configuration WF. (This WF is triggered per UCSM Account. For example, UCSM 120, UCSM121.)• UCS CPA Node Bare Metal. (This WF is triggered per Node.) The UCSM Specific Input area is displayed when the Multi-UCSM check box is checked.
UCS Manager Account drop-down list	Choose the Cisco UCS Manager account for this cluster.

Name	Description
Organization drop-down list	Choose the organization in which the servers for this cluster are located.
SSD Boot Drives Available for OS check box	Click this check box if you do not want to validate the server disk availability for RAID level OS disks. Ensure that the servers contain SSD. If the check box is not selected, the disk availability for both the OS disk and data disk are validated based on their RAID level.
PXE VLAN ID field	Enter the PXE VLAN ID.
UCS SP Template table	Choose an existing UCS Service Profile Template for Hadoop cluster creation.

Step 5 If you want to edit a Hadoop Server Role, select the row for that role, and click **Edit**.

Step 6 On the **Edit Hadoop Server Roles Entry** screen, complete the following fields and click **Submit**.

Name	Description
Node Type field	Displays the Hadoop node role.
Node Count field	The number of nodes in the Hadoop cluster for the selected node type.
SSD Boot Drives Available for OS check box	Check this check box if you do not want to validate the server disk availability for RAID level OS disks. Ensure that the servers contain SSD. If the check box is not selected, the disk availability for both the OS disk and data disk are validated based on their RAID level. Note This check box is not displayed when the UCSM version is greater than or equal to 3.
Server Pool table	Enter the server pool that you want to use for the cluster for the selected node type. The Cisco UCS Manager account and the organization that you choose determine which server pools are displayed in this area.

Step 7 In the **vNIC Template** table, review and, if desired, edit the vNIC templates available for the cluster.

Step 8 If you want to edit a vNIC template, select the row for that template and click **Edit**.

Step 9 On the **Edit vNIC Template Entry** screen, complete the following fields and click **Submit**.

Name	Description
vNIC Name drop-down list	The vNIC name in the selected template. This field is for your information only.
IP Pool field	Choose the big data IP pool that you want to use for IP addresses assigned to this vNIC.
MAC Address Pool drop-down list	Choose the MAC address pool that you want to use for this cluster. (This drop-down list is disabled if an existing UCS SP Template is selected.)
First MAC Address field	Enter the MAC address.
Size field	Enter the size.
VLAN ID field	The VLAN ID for this cluster. (This field is disabled if an existing UCS SP Template is selected.)

Note When you use vNIC bonding, ensure that you assign IP Pool, MAC Address Pool, and VLAN ID to the first vNIC in the **vNIC Template** table.

Step 10 Click **Submit**.

What to Do Next

You can view and monitor the workflow that gets triggered after creating a customized Hadoop cluster.

Creating a Hadoop Cluster Using Workflow

In Cisco UCS Director Express for Big Data, administrator can map the advanced catalog option to Hadoop cluster creation workflow, with limited user inputs, so that the service end user can trigger cluster creation. See [Cisco UCS Director End User Portal Guide](#).

Before You Begin

- Create a service profile template
- Create a server pool in the Cisco UCS Manager account that you plan to use for this cluster. See [Cisco UCS Director Management Guide for Cisco UCS Manager](#).
- Create a MAC address pool

- Create a user with user role as service end user.

-
- Step 1** Log into Cisco UCS Director Express for Big Data using admin credentials.
- Step 2** Choose **Orchestration** and click **Workflows**.
- Step 3** Click **Add Workflow**.
- Step 4** On the **Add Workflow Details** page, enter the workflow name and choose a folder. Click **Next**.
- Step 5** On the **Add User Inputs** page, enter the required details and click **Next**.
- Step 6** On the **Add User Outputs** page, enter the required details and click **Submit**.
- Step 7** Double-click the workflow in the **Workflow Designer**.
- Step 8** Add the Initiate Hadoop Cluster task.
- Step 9** Select the attributes that you want to map to the workflow input fields. Check the **Map to User Input** check box to provide user inputs, if required.
- Step 10** Enter required details in the **Hadoop Service Role** table and **vNIC Template** table, and click **Submit**.
- Step 11** Choose **Policies > Catalogs** and click **Add Catalog**.
- Step 12** On the **Add Catalog page**, choose the catalog type as Advanced and select a workflow. Click **Submit** to map the workflow to the catalog.
- Step 13** Log into Cisco UCS Director Express for Big Data using service end user credentials.
- Step 14** Choose **Catalogs**. The **Catalogs** page displays the list of catalogs available for the service end user.
- Step 15** Select a catalog and click **Create Request**. The **Create Server Request** page displays the mapped user inputs.
- Step 16** Specify the required details.
- Step 17** Click **Next** and enter the cluster details in the **Customize Workflow** page.
- Step 18** Click **Next** and view the cluster details in the **Summary** page.
- Step 19** Click **Submit** to trigger a workflow for creating a Hadoop cluster.
-

Provisioning an Instant and Customized Hadoop Cluster

Create and customize a Cluster Deploy Template to trigger the workflow.

Before You Begin

- Create a UCS Service Profile template for a Customized Hadoop Cluster
- Create a Hadoop Cluster Profile template for a Customized Hadoop Cluster

-
- Step 1** Choose **Policies > Orchestration**.
- Step 2** Click the **UCS CPA** folder from the **Workflows** tab.
- Step 3** Double-click the workflow to open the workflow designer and execute the workflow.
- a) When you open the workflow designer for an instant Hadoop Cluster, you get the following tasks, which are processed sequentially.

Task Name	Description
Instant Hadoop Cluster UCS SP	Cisco UCS Director Express for Big Data automatically specifies parameters for installing the OS and Hadoop distribution software at the back end.
Instant Hadoop Cluster Profile	Cisco UCS Director Express for Big Data automatically configures Hadoop cluster services at the back end.
Setup Hadoop Cluster Env	Sets up the environment for cluster-specific scripts and software files.
Muti Bare Metal OS Install WF	Attaches the UCS profile and sets up all boot files required to boot the operating system (Linux). When the Power ON task is executed, the boot files are picked up, and the operating system is installed successfully.
Multi Bare Metal WF Monitor	Checks the status of bare metal OS install workflow.
Synchronized Command Execution	—
Custom SSH Command	Installs and configures the Hadoop distribution software.
Provision Hadoop Cluster	Sends the Hadoop cluster properties to the Web Console.
Completed	The Hadoop cluster is provisioned successfully. Note If any of the tasks fail, you are informed that the provisioning has failed. For more information on how to monitor the workflow, see Monitoring Service Requests for Big Data , on page 149.

- b) When you open the workflow designer for a customized Hadoop Cluster, you get the following tasks that get processed sequentially.

Task Name	Description
Create UCS Service Profile Template	Specifies parameters for installing the OS and Hadoop distribution software.
Create Hadoop Cluster Profile	Configures Hadoop cluster services.
Setup Hadoop Cluster Env	Sets up the environment for cluster-specific scripts and software files.
Muti Bare Metal OS Install WF	Attaches the UCS profile and sets up all boot files required to boot the operating system (Linux). When the Power ON task is executed, the boot files are picked up, and the operating system is installed successfully.

Task Name	Description
Multi Bare Metal WF Monitor	Checks the status of bare metal OS install workflow.
Synchronized Command Execution	—
Custom SSH Command	Installs and configures the Hadoop distribution software.
Provision Hadoop Cluster	Sends the Hadoop cluster properties to the Web Console.
Completed	The Hadoop cluster is provisioned successfully. Note If any of the tasks fail, you are informed that the provisioning has failed. For more information on how to monitor the workflow, see Monitoring Service Requests for Big Data , on page 149.

Managing a Hadoop Cluster

You can manage an existing cluster.

-
- Step 1** Choose **Solutions > Big Data > Accounts**.
- Step 2** Click **Hadoop Accounts**.
- Step 3** Select an account and click **View Details**.
- Step 4** Click **Summary** to view the statistics data report for the selected Hadoop Account and the high-level report on the cluster and node account.
- Step 5** Click **Hosts** to view node details. For more information on **Hosts** page, see [Managing Nodes in a Cluster](#).
- Step 6** Click **Hadoop Clusters** to view list of actions that you can perform on an existing Hadoop cluster.
- Click **Role Topology** to view the topology of the nodes. (This tab is not editable.)
 - Click **View Details** to view the inputs for the Hadoop cluster you have created, and to view the virtual network interface configuration. (This information is not editable.)
 - Click **Shutdown Cluster** to stop all the services and power off all the nodes in the Hadoop cluster.
 - Click **Start Cluster** to power up all the nodes in the Hadoop cluster and start all services in the Hadoop cluster.
 - Click **Restart Cluster** to power off and then power up all nodes in the Hadoop cluster.
 - Click **Rebalance** to configure the threshold percentage to rebalance Hadoop clusters. For MapR cluster, configure the threshold percentage using CLDB Balancer **Disk Paused** and CLDB Balancer **Disk Max Switches in Nodes**.
 - Click **Upgrade Cluster** to upgrade Hadoop distributions from the current version, if available.
- Note** For a derived account, install and configure pssh and clush for the nodes in the Hadoop cluster and password less between nodes.
- Click **Upgrade Cluster**.

- 2 Choose the JDK version from the **Oracle JDK** drop-down list.
- 3 Choose the Hadoop distribution that you want to upgrade from the current version from the **Available Version** drop-down list.
- 4 Check the **Enable HA** check box to enable high availability for the Hadoop cluster, if Cloudera or Hortonworks is the Hadoop distribution.

Note To enable high availability in a Hortonworks cluster, you require a minimum of four nodes and node 1, 2, and 3 should be selected as Journal nodes.
- 5 Click **Submit**.

h) Enable Hadoop cluster high availability.

- 1 Click **Enable High Availability**.
- 2 From the **Enable High Availability** screen, access the **Standby Name Node** drop-down list and choose the Standby Name Node.
- 3 Check a minimum of three nodes from the **Journal Nodes** table.

Note Journal nodes selection is recommended to be on first three Master nodes and should be increased by the node count in odd number only.
- 4 Click **Submit**.

i) Disable Hadoop cluster high availability. This action is supported only for Cloudera and Hortonworks.

- 1 Click **Disable High Availability**.
- 2 From the **Disable High Availability** screen, access the **Standby Name Node** drop-down list and choose the Standby Name Node.
- 3 Check a minimum of three nodes from the **Journal Nodes** table.

Note Journal nodes selection is recommended to be on first three Master nodes and should be increased by the node count in odd number only.
- 4 Click **Submit**.

j) Click **Cluster Snapshot** to view the snapshot.

k) Click **View Reports** to view performance and monitoring reports.

Step 7 Click **Hadoop Services** to view the list of Hadoop services and their status. You can do the following:

- a) Click **Start All Services** to start all Hadoop services, depending on their status.
- b) Click **Stop All Services** to stop all Hadoop services, depending on their status.
- c) Click **Add New Service** to add a new Hadoop service.
- d) Click **Edit Service** to start and stop a particular Hadoop service.

Step 8 Click **Hadoop Service Roles** to view the list of Hadoop services. You can do the following:

- a) To add a role to the cluster, do the following on the **Add Role** screen:
 - 1 From the **Hadoop Service Name** drop-down list, choose the Hadoop service.
 - 2 From the **Role Type** drop-down list, choose the role type.
 - 3 From the **Node Name** drop-down list, choose the node name.

- 4 From the **Role Name** drop-down list, choose the role name.
 - 5 Click **Submit**.
- b) To start or stop any role that you have created, do the following on the **Start/Stop Role** screen:
- 1 From the **Hadoop Service Roles** tab, choose the Hadoop service.
 - 2 Click **Start/Stop Role**.
 - 3 Click **Submit**.
- c) To delete a role in the cluster, do the following on the **Delete Role** screen:
- 1 From the **Hadoop Service Roles** tab, choose the Hadoop service.
 - 2 Click **Delete**.
 - 3 Click **Submit**.

Step 9 Click **More Reports** to view the list of additional reports that you can generate about data usage and CPU utilization.

View Hadoop Cluster Details

For each Big Data Account, use the **Hadoop Clusters** tab to view details of all Hadoop clusters associated with the account. See [Managing a Hadoop Cluster](#)

You can view the following details by clicking **Hadoop Clusters**.

Name	Description
Big Data Account name	The name of the Big Data account.
UCS SP Template for Big Data	The UCS SP Template for Big Data that you used to create service profiles for the servers in the Hadoop cluster.
Hadoop Cluster Profile Template	The Hadoop cluster profile template that you used to configure the cluster services.
Hadoop Cluster Deploy Template	A unique name that you used for the Hadoop cluster deployment template.
UCSM Policy Name Prefix	The UCSM Policy Name prefix.
Hadoop Cluster Name	A unique name that you used for the Hadoop cluster.
Hadoop Node Count	The number of nodes in the Hadoop cluster.
Hadoop Node Prefix	The Host Node prefix for the cluster.

Name	Description
OS Version	The operating system that you installed on the servers for the Hadoop cluster.
Hadoop Distribution	The Hadoop distribution that you used for this cluster.
Hadoop Distribution Version	The Hadoop distribution version that used for this cluster.
PXE VLAN ID	The VLAN ID used for PXE boot of the servers.
UCS Service Profile Template	The UCS Service Profile Template that you specified parameters for installing the OS and Hadoop distribution software.
Host Maintenance Policy	–
Host Firmware Package	–
UCSM Version	Displays the Cisco UCS Manager version
vNIC: eth0	Displays IPv4 network information for the management interface and the management VLAN ID.
vNIC: eth1	Displays IPv4 network information for the DATA1 interface and the VLAN ID. This field is not displayed when the Use one vNIC check box is selected while creating a Cisco UCS Service Profile Template for Big Data.
vNIC: eth2	Displays IPv4 network information for the DATA2 interface and the VLAN ID. This field is not displayed when the Use one vNIC check box is selected while creating a Cisco UCS Service Profile Template for Big Data.

Viewing a Cluster Snapshot

A cluster snapshot displays configuration details of a Hadoop cluster, such as hosts, roles, and services. To view the current snapshot of a cluster, do the following:

-
- Step 1** Choose **Solutions > Big Data > Accounts**.
 - Step 2** Click **Hadoop Accounts**.
 - Step 3** Choose the Hadoop Account for which you want to view the snapshot and click **View Details**.
 - Step 4** Click **Hadoop Clusters**.
 - Step 5** Choose the Hadoop cluster for which you want to view the snapshot and click **Cluster Snapshot**.
 - Step 6** Click **Submit**.
You can view the snapshot for the selected Hadoop cluster.
-

Adding a New Hadoop Service

-
- Step 1** Choose **Solutions > Big Data > Accounts**.
 - Step 2** Choose a Cloudera account for which you can add a new Hadoop service.
Adding a new Hadoop service is not supported for MapR and Hortonworks distributions.
 - Step 3** Click **View Details**.
 - Step 4** Click **Hadoop Services**.
 - Step 5** Click **Add New Service**.
On the **Add New Service** screen, complete the following fields:

Name	Description
Workflow Inputs	
(Hadoop) Account Name	Choose the Cloudera Account.
Service Type	Enter the Hadoop service for Cloudera.
Role Assignment Pairs	Enter Role Assignment Pairs, separated by commas. For example, RoleType1:hostname1, RoleType2:hostname2
Dependant Services	Enter the list of dependent services. Use commas to separate list entries.
Pre Install Commands	Enter the list of commands. Use the "\n" command to insert a new line after each list entry.

Name	Description
Post Install Commands	Enter the list of commands. Use the "\n" command to insert a new line after each list entry.

Step 6 Click **Submit**.

Managing Nodes in a Cluster

You can add, delete, decommission, and recommission nodes in a cluster.

- **Managed Node**—Any node that was already a member (managed) in the cluster and deleted, which can be added again in the cluster.
- **Live Node**—Any node that has the operating system installed and is reachable from the Hadoop cluster.
- **Bare Metal Node**—Any node that is available and is not associated with the Hadoop cluster.

Step 1 Choose **Solutions > Big Data > Accounts**.

Step 2 Click **Hadoop Accounts**.

Step 3 Select an account and click **View Details**.

Step 4 Click **Hosts** to perform the following actions:

Name	Description
Refresh	Refreshes the page.
Favorite	Adds the page to Favorites.
Add Managed Node	Adds managed nodes to the Hadoop cluster.
Add Live Node	Adds live nodes to the Hadoop cluster.
Add Bare Metal Nodes	Adds bare metal nodes to the Hadoop cluster.
Add New Disks	Adds disks to the Hadoop cluster.
Remove Disks	Remove disks from the Hadoop cluster. Note This function is supported only in HDFS Data Node and Yarn Node Manager for Hortonworks, Cloudera, and BigInsights clusters only. However, there is no restriction for MapR clusters and you can delete any disks from any nodes.

Name	Description
Disk Locator	Locates the disks in Cisco UCS Manager and turns on the LED on the selected disk of the server.
Report Metadata	

Step 5

Select a host that allows you to perform the following actions:

Name	Description
View Details	Displays the summary of the CPU usage, the I/O status of the hosts disks, and so on. Note If you see a License Status tab, it indicates a licensing issue.
Delete Node	Deletes node from the cluster.
Assign Rack	Assigns the node to the rack server.
Recommission Node/ Decommission Node	Decommissioning or Recommissioning a node depends on its status. Note When the node is in decommissioned status, it means that all the roles for that node have been withdrawn.
Delete Node to Bare Metal	The node is removed from the cluster and disassociated from the service profile. The node becomes a bare metal server.
Host Mappings	Lists DNS entries of all the hosts in the Hadoop cluster.
Run Inventory	Collects the hardware inventory for the selected server. For example, disk details.

Delete Node and Delete Node to Bare Metal Actions in Cloudera and Hortonworks

When you perform the Delete Node or Delete Node to Bare Metal actions, the UCS CPA Delete Node (a new workflow) is created for Cloudera and Hortonworks. You can also execute this workflow for MapR to perform the delete node operation. If you execute the UCS CPA Delete Node workflow for Cloudera and Hortonworks, this workflow also provides the functionality of Delete Node to Bare Metal action. This functionality is based

on the Delete Node to Bare Metal flag setting as true or false. In addition, the Rollback UCS CPA Node Bare Metal workflow is created.

Adding Managed Nodes to the Hadoop Cluster

Add managed nodes to the Hadoop cluster.



Note

This **Add Managed Nodes** functionality is not supported for Hortonworks 2.3 and later versions, and MapR distribution, but you can use the Add Live Nodes functionality.

This feature allows you to add nodes that are available only from the following URLs, but not the members of the cluster.

- Cloudera—<http://serverip:7180/api/v6/hosts>, where the serverIP is the IPv4 address of the administration node.
- Hortonworks—<http://serverIP:8080/api/v1/hosts>, where the serverIP is the IPv4 address of the administration node



Note

The **Add Managed Nodes** functionality is not supported in Cloudera and BigInsights when you select the node type as **Edge Node**.

-
- Step 1** Choose **Solutions > Big Data > Accounts**.
 - Step 2** Click **Hadoop Accounts**.
 - Step 3** Double-click the Hadoop account.
 - Step 4** Click **Hosts**.
 - Step 5** Click **Add Managed Node**.
 - Step 6** From the **Host Name** drop-down list, choose the host name.
 - Step 7** Click **Submit**.
-

Adding Live Nodes to the Hadoop Cluster

Add live nodes to the Hadoop cluster.

-
- Step 1** Choose **Solutions > Big Data > Accounts**.
 - Step 2** Click **Hadoop Accounts**.
 - Step 3** Double-click the Hadoop account.
 - Step 4** Click **Hosts**.
 - Step 5** Click **Add Live Node**.
 - Step 6** Enter the IPv4 address in the **Host Management IPv4 Address** field.
 - Step 7** Enter the name of the rack server in the **Rack Name** field.
 - Step 8** Enter the password in the **(New Node) Password** field for that rack server.
 - Step 9** Choose the Cluster Manager Version for the Hadoop distribution from the **Cluster Management Version** drop-down list.
 - Step 10** Choose the operating system to be installed on the servers in this cluster from the **OS Version** drop-down list.
 - Step 11** Choose the node type from the **Node Type** drop-down list.
 - Step 12** Choose a MapR cluster template from the **Hadoop Template Name** drop-down list. This field is displayed only when you select the MapR cluster.
 - Step 13** Click **Submit**.
-

Adding Bare Metal Nodes to the Hadoop Cluster

Add bare metal nodes to the Hadoop cluster.

-
- Step 1** Choose **Solutions > Big Data > Accounts**.
 - Step 2** Click **Hadoop Accounts**.
 - Step 3** Double-click the Hadoop account.
 - Step 4** Click **Hosts**.
 - Step 5** Click **Add Bare Metal Nodes**.
 - Step 6** On the **Add Bare Metal Nodes** screen, complete the following fields:

Name	Description
Big Data Account Name field	The name of the Big Data account.
UCSM Policy Name Prefix field	The UCSM Policy Name prefix.
Hadoop Cluster Name field	A unique name for the Hadoop cluster.

Name	Description
Hadoop Node Count field	The number of nodes in the Hadoop cluster.
Host Node Prefix field	The Host Node prefix for the cluster.
OS Version drop-down list	Choose the operating system to be installed on the servers in this cluster.
Hadoop Distribution drop-down list	Choose the Hadoop distribution to be used for this cluster.
Hadoop Distribution Version drop-down list	Choose the Hadoop distribution version.
Oracle JDK Version drop-down list	Choose the Oracle JDK version.
External Database drop-down list	Choose an external database. You can also configure a new database from here.
Multi-UCSM check box	<p>Check the Multi-UCSM check box if you use multiple UCSM accounts.</p> <p>Note If you use the multiple UCSM accounts option, you can configure the Hadoop Server Roles as described in Step 7. You can add UCSM Specific Inputs in the Add Entry to UCSM Specific Inputs table.</p> <p>The following workflows are created during an Instant and Customized Hadoop Cluster creation:</p> <ul style="list-style-type: none"> • UCS CPA Multi-UCSM Hadoop Cluster WF. • Single UCSM Server Configuration WF. (This WF is triggered per UCSM Account. For example, UCSM 120, UCSM121.) • UCS CPA Node Bare Metal. (This WF is triggered per Node.) <p>The UCSM Specific Input area is displayed when the Multi-UCSM check box is checked.</p>
UCS Manager Account drop-down list	Choose the Cisco UCS Manager account for this cluster.
Organization drop-down list	Choose the organization in which the servers for this cluster are located.

Name	Description
SSD Boot Drives Available for OS check box	Check this check box if you do not want to validate the server disk availability for RAID level OS disks. Ensure that the servers contain Solid-State Drive (SSD). If the check box is not selected, the disk availability for both the OS disk and data disk are validated based on their RAID level.
Hadoop Template Name drop-down list	Choose a template for a MapR cluster. This field is displayed only when you select the MapR cluster.
UCS SP Template table	Choose an existing UCS Service Profile Template for Hadoop cluster creation.
PXE VLAN ID field	Enter the PXE VLAN ID.
UCSTemplate Name table	Check the UCS Service Profile Template check box that you want to use and click Submit to confirm the selection.

Step 7

If you want to edit a Hadoop Server Role, select the row for that role, and click **Edit**.

Step 8

On the **Edit Hadoop Server Roles Entry** screen, complete the following fields and click **Submit**.

Name	Description
Node Type field	Displays the Hadoop node role. Note Kafka node is supported only for Cloudera and Hortonworks clusters.
Node Count field	The number of nodes in the Hadoop cluster for the selected node type.
SSD Boot Drives Available for OS check box	Check this check box if you do not want to validate the server disk availability for RAID level OS disks. Ensure that the servers contain SSD. If the check box is not selected, the disk availability for both the OS disk and data disk are validated based on their RAID level. Note This check box is not displayed when the UCSM version is greater than or equal to 3.
Server Pool table	Enter the server pool that you want to use for the cluster for the selected node type. The Cisco UCS Manager account and the organization that you choose determine which server pools are displayed in this area.

Step 9 In the **vNIC Template** table, verify the vNIC templates available for the cluster.

Step 10 If you want to edit a vNIC template, select the row for that template and click **Edit**.

Step 11 On the **Edit vNIC Template Entry** screen, complete the following fields and click **Submit**.

Name	Description
vNIC Name drop-down list	The vNIC name in the selected template. This field is for your information only.
IP Pool drop-down list	Choose the Big Data IP pool that you want to use for IP addresses assigned to this vNIC.
MAC Address Pool drop-down list	Choose the MAC address pool that you want to use for this cluster.
First MAC Address field	Enter the MAC address.
Size field	Enter the size.
VLAN ID field	The VLAN ID for this cluster.

Step 12 Click **Submit**.

Adding Disks to the Hadoop Cluster

Step 1 Choose **Solutions > Big Data > Accounts**.

Step 2 Click **Hadoop Accounts**.

Step 3 Double-click the Hadoop account.

Step 4 Click **Hosts**.

Step 5 From the **More Actions** drop-down list, choose **Add New Disks**.

Step 6 Expand **Select disks** and check each disk that you want to use.

Step 7 Choose the method to write the data in the disks from the **Write Mode** drop-down list.

Step 8 Choose the method to read data from the disks from the **Read Mode** drop-down list.

Step 9 Check **Use Cache** to use the RAID controller cache to read and write operations.

Step 10 Check **Use Cache if Bad BBU** to ensure that if the Battery Backup Unit (BBU) is not available for any reason, **Write back** will be disabled and **Write Through** will be enabled.

Step 11 Choose a strip size for each disk within a stripe from the **Stripe Size (MB)** drop-down list.

Step 12 Click **Submit**.

Service Roles

If you use Add Node Bare Metal, Add Managed Node, and Add Live Node actions, the following nodes-specific roles are added for each Hadoop distribution.

Service Roles	Cloudera	MapR	Hortonworks
FileServer	No	Yes	No
DataNode	Yes	No	Yes
NodeManager	Yes	Yes	Yes
Ganglia Monitor	No	No	Yes
NFS Gateway	No	Yes	No



Managing Splunk Clusters

This chapter contains the following sections:

- [Creating an Instant Splunk Cluster, page 103](#)
- [Creating a Splunk Cluster Using Workflow, page 107](#)
- [Customizing Splunk Cluster Creation, page 108](#)
- [Adding Bare Metal Nodes to the Splunk Cluster, page 112](#)
- [Deploying Splunk Cluster with Archival Node and NFS Support, page 115](#)
- [Managing a Splunk Cluster, page 115](#)

Creating an Instant Splunk Cluster

Use this procedure to create an instant Splunk cluster with the predefined values for the UCS Service Profile template. The system creates the QUICK_UCS_SPLUNK template, a new UCS SP Template of container type splunk while creating the instant splunk cluster.

Step 1 Choose **Solutions > Big Data > Containers**.

Step 2 Click **Cluster Deploy Template**.

Step 3 Click **Instant Splunk Cluster Creation**.

Step 4 On the **Instant Splunk Cluster Creation** screen, complete the following fields:

Name	Description
Big Data Account Name field	The name of the Big Data account.
UCSM Policy Name Prefix field	The UCSM Policy Name prefix.
SSH (root) Password field	The SSH root password. Note The SSH username pertains to the root user.

Name	Description
Confirm SSH Password field	Enter the SSH root password.
Splunk Manager Password field	The management console password.
Confirm Splunk Manager Password field	Enter the management console password.
OS Version drop-down list	Choose the operating system to be installed on the servers in this cluster.
Splunk Distribution Version drop-down list	Choose the Splunk Enterprise version to be used for this cluster.
Multi-UCSM check box	<p>Check the Multi-UCSM check box if you use multiple UCSM accounts.</p> <p>Note If you use the multiple UCSM accounts option, you can configure the Splunk Server Roles as described in Step 7. You can add UCSM Specific Inputs in the Add Entry to UCSM Specific Inputs table.</p> <p>The following workflows are created during an Instant Splunk Cluster creation and Customized Splunk Cluster creation:</p> <ul style="list-style-type: none"> • UCS CPA Multi-UCSM Splunk Cluster WF • Single UCSM Server Configuration WF (This WF is triggered per UCSM Account. For example, UCSM 120, UCSM121) • UCS CPA Node Bare Metal (This WF is triggered per Node)
UCS Manager Account drop-down list	Choose the Cisco UCS Manager account for this cluster.
Organization drop-down list	Choose the organization in which the servers for this cluster are located.
UCS SP Template table	Choose an existing UCS Service Profile Template for cluster creation.
PXE VLAN ID field	Enter the PXE VLAN ID.

Step 5 Enter **Replication Factor**.

Step 6 Enter **Search Factor**. The search factor must be equal to the replication factor.

Step 7 In the **Splunk Server Roles** table, if you want to edit a Splunk Server Role, select the row for that role, and click **Edit**.

Step 8 On the **Edit Splunk Server Roles Entry** screen, complete the following fields and click **Submit**. The fields displayed in the **Edit Splunk Server Roles Entry** screen is based on the server role selection.

Name	Description
Node Type field	Displays the Splunk node role.
Node Count field	The number of nodes in the splunk cluster for the selected node type.
Host Name Prefix drop-down list	Choose the host name prefix for this splunk cluster.
SSD Boot Drives Available for OS check box	<p>Check this check box if you do not want to validate the server disk availability for RAID level OS disks. Ensure that the servers contain SSD.</p> <p>If the check box is not selected, the disk availability for both the OS disk and data disk are validated based on their RAID level.</p> <p>Note This check box is not displayed when the UCSM version is greater than or equal to 3.</p>
Search Head to be part of cluster	By default, this option is checked and disabled. The search head role is added to all Search Head cluster.
Validate Page check box	Check Validate Page to recalculate admin hostnames per given hostname prefix and node count.
Deploy roles on Bare Metal check box	Check Deploy roles on Bare Metal to deploy roles on a bare metal agent. By default, this option is checked. Uncheck this option to deploy admin roles on Live Nodes.
Use Existing License Master check box	Check Use Existing License Master to use the existing license master.
License Master BM drop-down list	Choose the license master bare metal.
Monitoring Console BM drop-down list	Choose the monitoring console bare metal.
Cluster Master BM drop-down list	Choose the cluster master bare metal.
Deployer BMs table	Choose the bare metal of the deployer server.
Deployment Server BMs table	Choose the bare metal of the deployment server.

Name	Description
Current License Master Live IPs	Enter the IP addresses of the current license master. This field is displayed when Use Existing License Master is checked.
New License Master Live IP	Enter the IP address of the new license master. This field is displayed when Deploy roles on Bare Metal is unchecked.
Monitoring Console Live IP	Enter the IP address of the monitoring console. This field is displayed when Deploy roles on Bare Metal is unchecked.
Cluster Master Live IP	Enter the IP address of the new license cluster master. This field is displayed when Deploy roles on Bare Metal is unchecked.
Deployer Live IPs	Enter the IP addresses of the deployer server. This field is displayed when Deploy roles on Bare Metal is unchecked.
Deployment Server Live IPs	Enter the IP addresses of the deployment server. This field is displayed when Deploy roles on Bare Metal is unchecked.
Server Pool table	Enter the server pool that you want to use for the cluster for the selected node type. The Cisco UCS Manager account and the organization that you choose determine which server pools are displayed in this area.

Note All Live IPs provided for admin roles of a Splunk cluster except for Existing Licensing server and running OS should be same as the Splunk Indexer or Search Head cluster.

Note Hostnames separated by comma or IP addresses can be provided and the hostname resolution should happen from the Cisco UCS Director appliance.

Step 9 In the **vNIC Template** table, review and, if desired, edit the vNIC templates available for the cluster.

Step 10 If you want to edit a vNIC template, select the row for that template and click **Edit**.

Step 11 On the **Edit vNIC Template Entry** screen, complete the following fields and click **Submit**.

Name	Description
vNIC Name drop-down list	The vNIC name in the selected template. This field is for your information only.
IP Pool field	Choose the Big Data IP pool that you want to use for IP addresses assigned to this vNIC.

Name	Description
MAC Address Pool drop-down list	Choose the MAC address pool that you want to use for this cluster. (This drop-down list is disabled if an existing UCS SP Template is selected.)
VLAN ID field	The VLAN ID for this cluster. (This field is disabled if an existing UCS SP Template is selected.)

Step 12 Click **Submit**.

Creating a Splunk Cluster Using Workflow

In Cisco UCS Director Express for Big Data, administrator can map the advanced catalog option to Splunk cluster creation workflow, with limited user inputs, so that the service end user can trigger cluster creation. See [Cisco UCS Director End User Portal Guide](#).

Before You Begin

- Create a service profile template
- Create a server pool in the Cisco UCS Manager account that you plan to use for this cluster. See [Cisco UCS Director Management Guide for Cisco UCS Manager](#).
- Create a MAC address pool

- Create a user with user role as service end user.

-
- Step 1** Log into Cisco UCS Director Express for Big Data using admin credentials.
- Step 2** Choose **Orchestration** and click **Workflows**.
- Step 3** Click **Add Workflow**.
- Step 4** On the **Add Workflow Details** page, enter the workflow name and choose a folder. Click **Next**.
- Step 5** On the **Add User Inputs** page, enter the required details and click **Next**.
- Step 6** On the **Add User Outputs** page, enter the required details and click **Submit**.
- Step 7** Double-click the workflow in the **Workflow Designer**.
- Step 8** Add the Initiate Splunk Cluster task.
- Step 9** Select the attributes that you want to map to the workflow input fields. Check the **Map to User Input** check box to provide user inputs, if required.
- Step 10** Enter required details in the **Splunk Service Role** table and **vNIC Template** table, and click **Submit**.
- Step 11** Choose **Policies > Catalogs** and click **Add Catalog**.
- Step 12** On the **Add Catalog** page, choose the catalog type as Advanced and select a workflow. Click **Submit** to map the workflow to the catalog.
- Step 13** Log into Cisco UCS Director Express for Big Data using service end user credentials.
- Step 14** Choose **Catalogs**. The **Catalogs** page displays the list of catalogs available for the service end user.
- Step 15** Select a catalog and click **Create Request**. The **Create Server Request** page displays the mapped user inputs.
- Step 16** Specify the required details.
- Step 17** Click **Next** and enter the cluster details in the **Customize Workflow** page.
- Step 18** Click **Next** and view the cluster details in the **Summary** page.
- Step 19** Click **Submit** to trigger a workflow for creating a Splunk cluster.
-

Customizing Splunk Cluster Creation

Before You Begin

- Create a UCS Service Profile Template.
- Create a Cluster Deploy Template.

-
- Step 1** Choose **Solutions > Big Data > Containers**.
- Step 2** Click **Cluster Deploy Templates**.
- Step 3** Click **Add** to create a cluster deploy template for the Splunk cluster. See [Creating a Cluster Deployment Template](#).
- Step 4** Click **Customized Splunk Cluster Creation**.
- Step 5** On the **Customized Splunk Cluster Creation** screen, complete the following fields.

Name	Description
Big Data Account Name field	The name of the Big Data account.
UCSM Policy Name Prefix field	The UCSM Policy Name prefix.
SSH (root) Password field	The SSH root password. Note The SSH username pertains to the root user.
Confirm SSH Password field	Enter the SSH root password.
Splunk Manager Password field	The management console password.
Confirm Splunk Manager Password field	Enter the management console password.
OS Version drop-down list	Choose the operating system to be installed on the servers in this cluster.
Splunk Distribution Version drop-down list	Choose the Splunk distribution version to be used for this cluster.
Multi-UCSM check box	Check the Multi-UCSM check box if you use multiple UCSM accounts. Note If you use the multiple UCSM accounts option, you can configure the Splunk Server Roles as described in the Step 8. You can add UCSM Specific Inputs in the Add Entry to UCSM Specific Inputs table. The following workflows are created during an Instant Splunk Cluster creation and Customized Splunk Cluster creation: <ul style="list-style-type: none"> • UCS CPA Multi-UCSM Splunk Cluster WF • Single UCSM Server Configuration WF (This WF is triggered per UCSM Account. For example, UCSM 120, UCSM121) • UCS CPA Node Bare Metal (This WF is triggered per Node)
UCS Manager Account drop-down list	Choose the Cisco UCS Manager account for this cluster.
Organization drop-down list	Choose the organization in which the servers for this cluster are located.
UCS SP Template table	Choose an existing UCS Service Profile Template for cluster creation.

Name	Description
PXE VLAN ID field	Enter the PXE VLAN ID.

Step 6 Enter **Replication Factor**.

Step 7 Enter **Search Factor**. The search factor must be equal to the replication factor.

Step 8 In the **Splunk Server Roles** table, if you want to edit a Splunk Server Role, select the row for that role, and click **Edit**.

Step 9 On the **Edit Splunk Server Roles Entry** screen, complete the following fields and click **Submit**. The fields displayed in the **Edit Splunk Server Roles Entry** screen is based on the server role

Name	Description
Node Type field	Displays the Splunk node role.
Node Count field	The number of nodes in the splunk cluster for the selected node type.
Host Name Prefix drop-down list	Choose the host name prefix for this splunk cluster.
SSD Boot Drives Available for OS check box	<p>Check this check box if you do not want to validate the server disk availability for RAID level OS disks. Ensure that the servers contain SSD.</p> <p>If the check box is not selected, the disk availability for both the OS disk and data disk are validated based on their RAID level.</p> <p>Note This check box is not displayed when the UCSM version is greater than or equal to 3.</p>
Search Head to be part of cluster	By default, this option is checked and disabled. The search head role is added to all Search Head cluster.
Validate Page check box	Check Validate Page to recalculate admin hostnames per given hostname prefix and node count.
Deploy roles on Bare Metal check box	Check Deploy roles on Bare Metal to deploy roles on a bare metal agent. By default, this option is checked. Uncheck this option to deploy admin roles on Live Nodes.
Use Existing License Master check box	Check Use Existing License Master to use the existing license master.
License Master BM drop-down list	Choose the license master bare metal.
Monitoring Console BM drop-down list	Choose the monitoring console bare metal.
Cluster Master BM drop-down list	Choose the cluster master bare metal.
Deployer BMs table	Choose the bare metal of the deployer server.

Name	Description
Deployment Server BMs table	Choose the bare metal of the deployment server.
Current License Master Live IPs	Enter the IP addresses of the current license master. This field is displayed when Use Existing License Master is checked.
New License Master Live IP	Enter the IP address of the new license master. This field is displayed when Deploy roles on Bare Metal is unchecked.
Monitoring Console Live IP	Enter the IP address of the monitoring console. This field is displayed when Deploy roles on Bare Metal is unchecked.
Cluster Master Live IP	Enter the IP address of the new license cluster master. This field is displayed when Deploy roles on Bare Metal is unchecked.
Deployer Live IPs	Enter the IP addresses of the deployer server. This field is displayed when Deploy roles on Bare Metal is unchecked.
Deployment Server Live IPs	Enter the IP addresses of the deployment server. This field is displayed when Deploy roles on Bare Metal is unchecked.
Server Pool table	Enter the server pool that you want to use for the cluster for the selected node type. The Cisco UCS Manager account and the organization that you choose determine which server pools are displayed in this area.

Step 10 In the **vNIC Template** table, review and, if desired, edit the vNIC templates available for the cluster.

Step 11 If you want to edit a vNIC template, select the row for that template and click **Edit**.

Step 12 On the **Edit vNIC Template Entry** screen, complete the following fields and click **Submit**.

Name	Description
vNIC Name drop-down list	The vNIC name in the selected template. This field is for your information only.
IP Pool field	Choose the big data IP pool that you want to use for IP addresses assigned to this vNIC.
MAC Address Pool drop-down list	Choose the MAC address pool that you want to use for this cluster. (This drop-down list is disabled if an existing UCS SP Template is selected.)

Name	Description
VLAN ID field	The VLAN ID for this cluster. (This field is disabled if an existing UCS SP Template is selected.)

Note When you use vNIC bonding, ensure that you assign IP Pool, MAC Address Pool, and VLAN ID to the first vNIC in the **vNIC Template** table.

Step 13 Click **Submit**.

Adding Bare Metal Nodes to the Splunk Cluster

Step 1 Choose **Solutions > Big Data > Accounts**.

Step 2 Click **Splunk Accounts**.

Step 3 Double-click the Splunk account.
You can see only the **Hosts** tab.

Step 4 Click **Add Bare Metal Nodes**.

Step 5 On the **Add Bare Metal Nodes** screen, complete the following fields:

Name	Description
Big Data Account Name field	The name of the Big Data account.
UCSM Policy Name Prefix field	The UCSM Policy Name prefix.
OS Version drop-down list	Choose the operating system to be installed on the servers in this cluster.
Splunk Version drop-down list	Choose the Splunk version.

Name	Description
<p>Multi-UCSM check box</p>	<p>Check the Multi-UCSM check box if you use multiple UCSM accounts.</p> <p>Note If you use the multiple UCSM accounts option, you can configure the Splunk Server Roles as described in the Step 6. You can add UCSM Specific Inputs in the Add Entry to UCSM Specific Inputs table.</p> <p>The following workflows are created during an Instant Hadoop Cluster creation and Customized Hadoop Cluster creation:</p> <ul style="list-style-type: none"> • UCS CPA Multi-UCSM Splunk Cluster WF. • Single UCSM Server Configuration WF (This WF is triggered per UCSM Account. For example, UCSM 120, UCSM121). • UCS CPA Node Bare Metal (This WF is triggered per Node).
<p>UCS Manager Account drop-down list</p>	<p>Choose the Cisco UCS Manager account for this cluster.</p>
<p>Organization drop-down list</p>	<p>Choose the organization in which the servers for this cluster are located.</p>
<p>UCS SP Template</p>	<p>Choose an existing UCS Service Profile Template for the cluster creation.</p>
<p>PXE VLAN ID field</p>	<p>Enter the PXE VLAN ID.</p>
<p>UCSTemplate Name table</p>	<p>Choose the UCS Service Profile Template for Splunk.</p>
<p>Replication Factor</p>	<p>The number of copies of data that you want the cluster to maintain. The default value is 2.</p>
<p>Search Factor</p>	<p>The number of searchable copies of data the cluster maintains. The default value is 2. The search factor must be equal to the replication factor.</p>

Step 6 In the **Splunk Server Roles** table, if you want to edit a Splunk Server Role, select the row for that role, and click **Edit**.

Step 7 On the **Edit Splunk Server Roles Entry** screen, complete the following fields and click **Submit**.

Name	Description
<p>Node Type field</p>	<p>Displays the Splunk node role.</p>

Name	Description
Node Count field	The number of nodes in the splunk cluster for the selected node type.
Host Name Prefix drop-down list	Choose the host name prefix for this splunk cluster.
SSD Boot Drives Available for OS check box	<p>Check this check box if you do not want to validate the server disk availability for RAID level OS disks. Ensure that the servers contain SSD.</p> <p>If the check box is not selected, the disk availability for both the OS disk and data disk are validated based on their RAID level.</p> <p>Note This check box is not displayed when the UCSM version is greater than or equal to 3.</p>
Search Head to be part of cluster	By default, this option is checked and disabled. The search head role is added to all Search Head cluster.
Server Pool table	<p>Enter the server pool that you want to use for the cluster for the selected node type.</p> <p>The Cisco UCS Manager account and the organization that you choose determine which server pools are displayed in this area.</p>

Step 8 In the **vNIC Template** table, review and, if desired, edit the vNIC templates available for the cluster.

Step 9 If you want to edit a vNIC template, select the row for that template and click **Edit**.

Step 10 On the **Edit vNIC Template Entry** screen, complete the following fields and click **Submit**.

Name	Description
vNIC Name drop-down list	This field is for your information only.
IP Pool drop-down list	Choose the Big Data IP pool that you want to use for IP addresses assigned to this vNIC.
MAC Address Pool drop-down list	Choose the MAC address pool that you want to use for this cluster. (This drop-down list is disabled if an existing UCS SP Template is selected.)
First MAC Address field	Choose the MAC address pool that you want to use for this cluster.
Size field	Enter the size.
VLAN ID field	The VLAN ID for this cluster.

Step 11 Click **Submit**.

Note By default, the hardware default is used as UUID pool for the servers in the cluster.

Deploying Splunk Cluster with Archival Node and NFS Support

The following are the scenarios to deploy a Splunk cluster along with Archival node:

- Configure Archival Node along with a Splunk Cluster—Archival node is configured along with the cluster automatically.
- Configure Archival Node on a Bare Metal—You can use the add Bare Metal option along with archival node settings like hostname prefix, number of archival nodes, and server pool. When the node comes up, the UCS CPA Splunk Add Live Archival Node workflow is used to configure NFS related setting on the node and configure mount point on indexers.
- Configuring Archival Node on a Live Node—You can use this to configure NFS related setting on the node and add it to the cluster.

For more information on how archival node disks are allocated to indexers, see the latest *Cisco UCS Integrated Infrastructure for Big Data with Splunk Enterprise*.

Managing a Splunk Cluster

You can manage the Splunk cluster from the **Hosts** tab.

Step 1 Choose **Solutions > Big Data > Accounts**.

Step 2 Click **Splunk Accounts**.

Step 3 Click **Hosts** to perform the following actions:

Name	Description
Refresh	Refreshes the page.
Favorite	Adds the page to Favorites.
Add Bare Metal Nodes	Add bare metal nodes to the Splunk cluster. You can add Indexer, Search Head, or Administrative node through Add Bare Metal workflow. You need to provide the Replication Factor based on the Indexer count.

Note You can also start, stop, or restart the Splunk cluster.

Step 4 Select a host that allows you to perform the following actions:

Name	Description
View Details	Displays the summary of the CPU usage, the I/O status of the hosts disks, and so on. Note If you see a License Status tab, it indicates a licensing issue.
Start	Starts the services on the node.
Stop	Stops the services on the node.
View Details	Restarts the services on the node.
Restart	Deletes node from the cluster.
Delete Node to Bare Metal	The node is removed from the cluster and disassociated from the service profile. The node becomes a Bare Metal server.

Step 5 Select an account and click **View Details**.
You can start, stop, or restart the Splunk cluster.

Step 6 Click the **Performance** tab.

Step 7 Click **Run Test**.

The Performance tab displays a default Big Data Metrics Report. This report shows the statistics collected for each host before the Splunk cluster creation and the reports post Splunk cluster creation only when you check the **Memory Test**, **Network Test**, and **Disk Test** check boxes in the **Pre Cluster Performance Tests** section of the **Management** tab. If you enable the precluster disk test, it impacts Splunk cluster creation.

Step 8 Click **Submit**, and then click **OK**.
For the following actions, choose the performance report:

Name	Description
View	Displays the metrics in the Big Data Metrics Report.
Compare	Compares and displays the metrics in the Big Data Metrics Report.
View Graph Report	Displays graphically the following reports from the Summary tab: <ul style="list-style-type: none"> • Average TRIAD Rate (MB/Sec) • Average Network Bandwidth (MB/Sec)
Delete	Deletes the Big Data Metrics Report.

Name	Description
More Reports	Displays the metrics on an hourly, daily, weekly, or monthly basis.

- Step 9** Click **Monitoring**.
Every time an inventory collection cycle is triggered, an entry listing the aggregate CPU, network bandwidth, and disk utilization metrics appears on the Monitoring Page.
- Step 10** Select the entry you want to analyze and click **View Details**.
- Step 11** Click **Back** to return to the **Monitoring** page.
-



CHAPTER 11

Big Data Cluster Configuration Settings

This chapter contains the following sections:

- [Creating an External Database Configuration, page 119](#)
- [Creating a Hadoop Cluster Configuration Parameters Template, page 121](#)
- [Updating Hadoop Cluster Configuration Parameters Template - Post Hadoop Cluster Creation, page 122](#)
- [Quality of Service System Classes, page 123](#)
- [Pre Cluster Performance Testing Settings, page 126](#)
- [Approving Hadoop Cluster and Splunk Deployment Workflows, page 126](#)
- [Adding NTP Server Details, page 127](#)
- [Uploading Required OS and Big Data Software to Cisco UCS Director Bare Metal Agent , page 128](#)
- [Cloudera, MapR, and Hortonworks RPMs on Cisco UCS Director Express for Big Data Bare Metal Agent, page 132](#)
- [Cloudera and MapR RPMs for Upgrading Hadoop Cluster Distributions, page 141](#)
- [Installation of User-Defined Software Post Hadoop Cluster Creation, page 142](#)
- [Configuration Check Rules, page 143](#)
- [Checking Hadoop Cluster Configuration, page 144](#)
- [Fixing Configuration Violations, page 144](#)

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.



Note MySQL is the only supported external database in Cisco UCS Director Express for Big Data.

Step 1 Choose **Solutions > Big Data > Settings**.

Step 2 Click **External Database**.

Step 3 Click **Add**.

You can modify or delete any database you have previously created from the external database table.

Step 4 On the **Create External Database Configurations** screen, 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 Choose **Solutions > Big Data > Settings**.

Step 2 Click **Hadoop Config Parameters**.

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 - Zookeeper Service** page of the **Create Hadoop Cluster Configuration Parameters Template** wizard, configure the parameters.
- Step 11** On the **Hadoop Config Parameters - Miscellaneous Parameters** page of the **Create Hadoop Cluster Configuration Parameters Template** wizard, configure the (ServiceLevel and RoleLevel) parameters.
- Step 12** Click **Submit**.
-

Updating Hadoop Cluster Configuration Parameters Template - Post Hadoop Cluster Creation

- Step 1** Choose **Solutions > Big Data > Accounts**.
- Step 2** Click the **Hadoop Accounts** tab and choose an existing Hadoop 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	<p>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.</p> <p>For example, this class has a drop policy that allows it to drop data packets if necessary. You cannot disable this system class.</p>
<ul style="list-style-type: none"> • Platinum • Gold • Silver • Bronze 	<p>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.</p>

System Class	Description
Fibre Channel	<p>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.</p> <p>For example, this class has a no-drop policy that ensures it never drops data packets. You cannot disable this system class.</p> <p>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</p>

Editing QoS System Classes

For more information on Quality of Service and System Classes, see [QoS System Classes](#).

Step 1 Choose **Solutions > Big Data > Settings**.

Step 2 Click **QoS System Class**.

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 On the **Modify QoS System Class** screen, complete the following fields:

Name	Description
Enabled check box	<p>If checked, the associated QoS class is configured on the fabric interconnect and can be assigned to a QoS policy.</p> <p>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.</p> <p>This check box is checked for Best Effort and Fibre Channel.</p>

Name	Description
CoS drop-down list	<p>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.</p> <p>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.</p>
Packet Drop check box	<p>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.</p>
Weight drop-down list	<p>A choice may be one of the following:</p> <ul style="list-style-type: none"> • 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.
Multicast Optimized check box	<p>If checked, the class is optimized to send packets to multiple destinations simultaneously. This option is not applicable to the Fibre Channel.</p>
MTU drop-down list	<p>The maximum transmission unit for the channel. This can be one of the following:</p> <ul style="list-style-type: none"> • 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. <p>This field is always set to fc for Fibre Channel.</p>

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 Choose **Solutions > Big Data > Settings**.

Step 2 Click **Management**.

Step 3 On the **Pre Cluster Performance Tests** page, 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 pre cluster 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** and click **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 Choose **Solutions > Big Data > Settings**.

Step 2 Click **Management**.

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.

Adding NTP Server Details

Step 1 Choose **Solutions > Big Data > Settings**.

Step 2 Click **Management**.

Step 3 Click **Add (+)**.

Step 4 On the **Add Entry to Servers** screen, complete the following fields:

Name	Description
Server Name field	The IP address of NTP server.
Is Primary Server check box	Click the check box if you want the server to be a primary server.

Step 5 Click **Submit**.

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

You can upload (add) required RHEL or CentOS ISO files, Big Data software and common software, and Oracle JDKs to Cisco UCS Director Bare Metal 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. Click the **Submit** button in the **Create Software Catalogs** dialog box to move the required files to the target Cisco UCS Director Bare Metal Agent.

Supported file formats:

- Linux OS— rhel-server-7.2-x86_64-dvd.iso, CentOS-7-x86_64-DVD-1511.iso
- Big Data software—[Big Data-Distro]-[Major version].[Minor version].[Patch version].zip (.gz or .tgz or .tar) For example, MapR-5.2.1.zip, cloudera-5.8.0, and Splunk-6.5.1.zip.
- Common software—bd-sw-rep.zip (.gz or .tgz or .tar)
- JDK software—jdk-8u60-linux-x64 (.rpm or gz)

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



Tip

If the required software column is empty for Big Data, then Cisco UCS Director Bare Metal Agent already contains all the files required.

Step 1 Choose **Solutions > Big Data > Settings**.

Step 2 Click **Software Catalogs**.

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 Bare Metal Agent from the **Target BMA** drop-down list.

Step 6 Check the **Restart BMA Services** to restart Cisco UCS Director Bare Metal 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	
OS Type drop-down list	Choose the OS image type that you want to install on the server. The drop-down list includes all OS types supported by the Cisco UCS Director Bare Metal Agent.
Catalog Name field	Operating System Name (for example, RHEL, CentOS)

Name	Description
Upload Type drop-down list	Choose one of the following: <ul style="list-style-type: none"> • Desktop file • The web server path that is reachable by the Cisco UCS Director Bare Metal Agent • Mountpoint in Cisco UCS Director Bare Metal Agent (For example, /root/iso) • Path to ISO in Cisco UCS Director Bare Metal Agent (For example, /temp/rhel65/iso) <p>Note If you select the Desktop file or web server path option, the .iso file is uploaded to BMA and then it is mounted. If you select the Mountpoint or Path to ISO option, the .iso file is directly mounted since it is already available in the Bare Metal Agent.</p>
Big Data Software Upload	
Catalog Name field	Big Data software. For example, Hadoop Distribution (for example, Distribution_name-Major version.Minor version.Patch version) or splunk enterprise software (Splunk-.Major version.Minor version.Patch version)
OS Type drop-down list	Choose the required OS type. For MapR and Splunk, choose the type as Any option. The file path for the OS-specific software for Cloudera and Hontonwork are modified. For example, the file paths for the repository in Cisco UCS Director Bare Metal Agent are /opt/cnsaroot/bd-sw-rep/cloudera-5.8.0/RHEL6 and /opt/cnsaroot/bd-sw-rep/cloudera-5.8.0/RHEL7.
Upload Type drop-down list	Choose one of the following: <ul style="list-style-type: none"> • Desktop file • The web server path that is reachable by the Cisco UCS Director Bare Metal Agent to upload remote software to Bare Metal Agent <p>Note If you select the Desktop file or web server path option, the big data software is copied to respective the folders in /opt/cnsaroot/bd-sw-rep/ location.</p>
Common Software Upload	

Name	Description
Upload Type drop-down list	Choose one of the following: <ul style="list-style-type: none"> • Desktop file • The web server path that is reachable by the Cisco UCS Director Bare Metal Agent to upload remote software to Bare metal Agent
JDK Upload	
JDK Version field	JDK version. For example, jdk-8u60-linux-x64.rpm
Upload Type drop-down list	Choose one of the following: <ul style="list-style-type: none"> • Desktop file • The web server path that is reachable by the Cisco UCS Director Bare Metal Agent to upload remote software to Bare metal Agent

Step 7 Click **Submit**.

What to Do Next

You can track software uploads here: **Administration** > **Integration**. Click **Change Record** 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 Enterprise 5.11.1	—	JDK 1.7.0_25	JDK 1.8.0_121
Cloudera Enterprise 5.10.0	—	JDK 1.7.0_25	JDK 1.8.0_121
Cloudera Enterprise 5.8.0	—	JDK 1.7.0_25	JDK 1.8.0_60
Cloudera Enterprise 5.6.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 Enterprise 5.5.0	—	JDK 1.7.0_25	JDK 1.8.0_60
Cloudera Enterprise 5.4.x	—	JDK 1.7.0_25	JDK 1.8.0_60
Cloudera Enterprise 5.3.x	—	—	JDK 1.8.0_11

Hadoop Distribution Version to Upgrade	Supported Upgrade Version
Cloudera Enterprise 5.4.x, JDK 1.8	Cloudera Enterprise 5.5.0, JDK 1.8
Cloudera Enterprise 5.4.x, JDK 1.8	Cloudera Enterprise 5.6.x, JDK 1.8
Cloudera Enterprise 5.4.x, JDK 1.8	Cloudera Enterprise 5.8.x, JDK 1.8
Cloudera Enterprise 5.6.x, JDK 1.8	Cloudera Enterprise 5.8.x, JDK 1.8
Cloudera Enterprise 5.8.0, JDK 1.8	Cloudera Enterprise 5.10.0, JDK 1.8
Cloudera Enterprise 5.8.0, JDK 1.8	Cloudera Enterprise 5.11.1, JDK 1.8

Supported Oracle JDK Versions for MapR and Upgrade Scenarios

Hadoop Distribution	JDK 1.7	JDK 1.8
MapR 5.2.1	Yes	Yes
MapR 5.2.0	Yes	Yes
MapR 5.1.0	Yes	Yes
MapR 5.0.0	Yes	Yes
MapR 4.1.0	No	Yes

Hadoop Distribution Version to Upgrade	Supported Upgrade Version
MapR 5.0.0, JDK 1.8	MapR 5.1.0, JDK 1.8
MapR 4.0.2, JDK 1.8	MapR 5.2.0, JDK 1.8

Supported Oracle JDK Software Versions for Hortonworks and Upgrade Scenarios

Hadoop Distribution	JDK 1.7	JDK 1.8
Hortonworks 2.5.3	—	Yes
Hortonworks 2.4	—	Yes
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
Hortonworks 2.2, JDK 1.7	Hortonworks 2.4, JDK 1.8

Supported Oracle JDK Software Versions for BigInsights

Hadoop Distribution	JDK 1.7	JDK 1.8
BigInsights 4.2	Yes	—

Cloudera, MapR, and Hortonworks RPMs on Cisco UCS Director Express for Big Data Bare Metal 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-1.7-1.el6.noarch.rpm from http://dl.fedoraproject.org/pub/epel/6/x86_64/clustershell-1.6-1.el6.noarch.rpm
- clustershell-1.7-1.el7.noarch.rpm from http://dl.fedoraproject.org/pub/epel/7/x86_64/c/clustershell-1.7.2-1.el7.noarch.rpm

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 Bare Metal Agent)
- mysql-connector-java-5.1.39.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>

Cloudera 5.5.0 Packages and Parcels

Download the following packages to `/opt/cnsaroot/bd-sw-rep/cloudera-5.5.1`:

- cm5.5.0-centos6.tar.gz from <http://archive.cloudera.com/cm5/repo-as-tarball/5.5.0/cm5.5.0-centos6.tar.gz>
- CDH-5.5.0-1.cdh5.5.0.p0.8-el6.parcel from <http://archive.cloudera.com/cdh5/parcels/5.5.0/CDH-5.5.0-1.cdh5.5.0.p0.8-el6.parcel>
- CDH-5.5.0-1.cdh5.5.0.p0.8-el6.parcel.sha1 from <http://archive.cloudera.com/cdh5/parcels/5.5.0/CDH-5.5.0-1.cdh5.5.0.p0.8-el6.parcel.sha1>
- manifest.json from <http://archive.cloudera.com/cdh5/parcels/5.5.0/manifest.json>

Cloudera 5.6.0 Packages and Parcels

Download the following packages to `/opt/cnsaroot/bd-sw-rep/cloudera-5.6.0`:

- cm5.6.0-centos6.tar.gz from <http://archive.cloudera.com/cm5/repo-as-tarball/5.6.0/cm5.6.0-centos6.tar.gz>
- CDH-5.6.0-1.cdh5.6.0.p0.45-el6.parcel from <http://archive.cloudera.com/cdh5/parcels/5.6.0/CDH-5.6.0-1.cdh5.6.0.p0.45-el6.parcel>
- CDH-5.6.0-1.cdh5.6.0.p0.45-el6.parcel.sha1 from <http://archive.cloudera.com/cdh5/parcels/5.6.0/CDH-5.6.0-1.cdh5.6.0.p0.45-el6.parcel.sha1>
- manifest.json from <http://archive.cloudera.com/cdh5/parcels/5.6.0/manifest.json>

Cloudera 5.8.0 Packages and Parcels

For RHEL 6.x, download the following packages to `/opt/cnsaroot/bd-sw-rep/cloudera-5.8.0`:

- `cm5.8.1-centos6.tar.gz` from <http://archive.cloudera.com/cm5/repo-as-tarball/5.8.1/cm5.8.1-centos6.tar.gz>
- `CDH-5.8.0-1.cdh5.8.0.p0.45-el6.parcel` from <http://archive.cloudera.com/cdh5/parcels/5.8.0/CDH-5.8.0-1.cdh5.8.0.p0.42-el6.parcel>
- `CDH-5.8.0-1.cdh5.8.0.p0.45-el6.parcel.sha1` from <http://archive.cloudera.com/cdh5/parcels/5.8.0/CDH-5.8.0-1.cdh5.8.0.p0.42-el6.parcel.sha1>
- `manifest.json` from <http://archive.cloudera.com/cdh5/parcels/5.8.0/manifest.json>

For RHEL 7.x, download the following packages to `/opt/cnsaroot/bd-sw-rep/cloudera-5.8.0`:

- `cm5.8.1-centos7.tar.gz` from <http://archive.cloudera.com/cm5/repo-as-tarball/5.8.1/cm5.8.1-centos7.tar.gz>
- `CDH-5.8.0-1.cdh5.8.0.p0.45-el7.parcel` from <http://archive.cloudera.com/cdh5/parcels/5.8.0/CDH-5.8.0-1.cdh5.8.0.p0.42-el7.parcel>
- `CDH-5.8.0-1.cdh5.8.0.p0.45-el7.parcel.sha1` from <http://archive.cloudera.com/cdh5/parcels/5.8.0/CDH-5.8.0-1.cdh5.8.0.p0.42-el7.parcel.sha1>
- `manifest.json` from <http://archive.cloudera.com/cdh5/parcels/5.8.0/manifest.json>

Cloudera 5.11.0 Packages and Parcels

Download the following packages to `/opt/cnsaroot/bd-sw-rep/cloudera-5.11.0`:

- `cm5.11.0-centos6.tar.gz` from <https://archive.cloudera.com/cm5/repo-as-tarball/5.11.0/cm5.11.0-centos7.tar.gz>
- `CDH-5.11.0-1.cdh5.11.0.p0.34-el7.parcel` from <http://archive.cloudera.com/cdh5/parcels/5.11.0/CDH-5.11.0-1.cdh5.11.0.p0.34-el7.parcel>
- `CDH-5.11.0-1.cdh5.11.0.p0.34-el7.parcel.sha1` from <http://archive.cloudera.com/cdh5/parcels/5.11.0/CDH-5.11.0-1.cdh5.11.0.p0.34-el7.parcel.sha1>
- `KAFKA-2.1.1-1.2.1.1.p0.18-el7.parcel` from <https://archive.cloudera.com/kafka/parcels/2.1/KAFKA-2.1.1-1.2.1.1.p0.18-el7.parcel>
- `KAFKA-2.1.1-1.2.1.1.p0.18-el7.parcel.sha1` from <https://archive.cloudera.com/kafka/parcels/2.1/KAFKA-2.1.1-1.2.1.1.p0.18-el7.parcel.sha1>
- `manifest.json` from <http://archive.cloudera.com/cdh5/parcels/5.11.0/manifest.json>

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/depoy/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 Bare Metal 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 Bare Metal 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>

MapR 5.1.0 Packages

Download the following packages to /opt/cnsaroot/bd-sw-rep/MapR-5.1.0

- mapr-ecosystem-5.x-20160304.rpm.tgz from <http://archive.mapr.com/releases/ecosystem-all/redhat/mapr-ecosystem-5.x-20160304.rpm.tgz>
- mapr-v5.1.0GA.rpm.tgz from <http://archive.mapr.com/releases/v5.1.0/redhat/mapr-v5.1.0GA.rpm.tgz>
- nss-3.21.0-8.el6.x86_64.rpm from ftp://mirror.switch.ch/pool/4/mirror/scientificlinux/6.8/x86_64/os/Packages/nss-3.21.0-8.el6.x86_64.rpm
- nss-softokn-3.14.3-23.el6_7.x86_64.rpm from http://vault.centos.org/6.7/updates/x86_64/Packages/nss-softokn-3.14.3-23.el6_7.x86_64.rpm
- nss-softokn-freebl-3.14.3-23.el6_7.x86_64.rpm from http://vault.centos.org/6.7/updates/x86_64/Packages/nss-softokn-freebl-3.14.3-23.el6_7.x86_64.rpm
- nss-sysinit-3.21.0-8.el6.x86_64.rpm from <http://archive.mapr.com/releases/ecosystem-all/redhat/mapr-whirr-0.8.1.18380-GA.noarch.rpm>
- nspr-4.11.0-1.el6.x86_64.rpm from ftp://mirror.switch.ch/pool/4/mirror/scientificlinux/6.3/x86_64/updates/security/nspr-4.11.0-1.el6.x86_64.rpm
- nss-tools-3.21.0-8.el6.x86_64.rpm from http://pkg-updates.fuel-infra.org/centos6/nss-tools-3.21.0-8.el6.x86_64.rpm
- nss-util-3.21.0-2.el6.x86_64.rpm from http://vault.centos.org/6.7/cr/x86_64/Packages/nss-util-3.21.0-2.el6.x86_64.rpm
- mapr-whirr-0.8.1.18380-GA.noarch.rpm from <http://archive.mapr.com/releases/ecosystem-all/redhat/mapr-whirr-0.8.1.18380-GA.noarch.rpm>

MapR 5.2.0 Packages

For RHEL 6.x, download the following packages to /opt/cnsaroot/bd-sw-rep/MapR-5.2.0:

- mapr-ecosystem-5.x-20160816.rpm.tgz from <http://archive.mapr.com/releases/ecosystem-all/redhat/mapr-ecosystem-5.x-20160816.rpm.tgz>
- mapr-v5.2.0GA.rpm.tgz from <http://archive.mapr.com/releases/v5.2.0/redhat/mapr-v5.2.0GA.rpm.tgz>
- nss-3.21.0-8.el6.x86_64.rpm from ftp://mirror.switch.ch/pool/4/mirror/scientificlinux/6.8/x86_64/os/Packages/nss-3.21.0-8.el6.x86_64.rpm
- nss-softokn-3.14.3-23.el6_7.x86_64.rpm from http://vault.centos.org/6.7/updates/x86_64/Packages/nss-softokn-3.14.3-23.el6_7.x86_64.rpm
- nss-softokn-freebl-3.14.3-23.el6_7.x86_64.rpm from http://vault.centos.org/6.7/updates/x86_64/Packages/nss-softokn-freebl-3.14.3-23.el6_7.x86_64.rpm
- mnss-sysinit-3.21.0-8.el6.x86_64.rpm from <http://archive.mapr.com/releases/ecosystem-all/redhat/mapr-whirr-0.8.1.18380-GA.noarch.rpm>
- nspr-4.11.0-1.el6.x86_64.rpm from ftp://mirror.switch.ch/pool/4/mirror/scientificlinux/6.3/x86_64/updates/security/nspr-4.11.0-1.el6.x86_64.rpm
- nss-tools-3.21.0-8.el6.x86_64.rpm from http://vault.centos.org/6.7/cr/x86_64/Packages/nss-tools-3.21.0-8.el6.x86_64.rpm
- nss-util-3.21.0-2.el6.x86_64.rpm from http://vault.centos.org/6.7/cr/x86_64/Packages/nss-util-3.21.0-2.el6.x86_64.rpm
- mapr-whirr-0.8.1.18380-GA.noarch.rpm from <http://archive.mapr.com/releases/ecosystem-all/redhat/mapr-whirr-0.8.1.18380-GA.noarch.rpm>

For RHEL 7.x, download the following packages to /opt/cnsaroot/bd-sw-rep/MapR-5.2.0:

- mapr-ecosystem-5.x-20160816.rpm.tgz from <http://archive.mapr.com/releases/ecosystem-all/redhat/mapr-ecosystem-5.x-20160816.rpm.tgz>
- libgenders-1.22-2.el7.x86_64.rpm from http://dl.fedoraproject.org/pub/epel/7/x86_64/l/libgenders-1.22-2.el7.x86_64.rpm
- libgenders-devel-1.22-2.el7.x86_64.rpm from http://dl.fedoraproject.org/pub/epel/7/x86_64/l/libgenders-devel-1.22-2.el7.x86_64.rpm
- mapr-whirr-0.8.1.18380-GA.noarch.rpm from <http://archive.mapr.com/releases/ecosystem-all/redhat/mapr-whirr-0.8.1.18380-GA.noarch.rpm>
- mapr-v5.2.0GA.rpm.tgz from <http://package.mapr.com/releases/v5.2.0/redhat/mapr-v5.2.0GA.rpm.tgz>
- mysql-connector-java-5.1.26.tar.gz from <http://cdn.mysql.com/archives/mysql-connector-java-5.1/mysql-connector-java-5.1.26.tar.gz>
- mysql-connector-java-5.1.26.tar.gz from <http://cdn.mysql.com/archives/mysql-connector-java-5.1/mysql-connector-java-5.1.26.tar.gz>
- sshpass-1.05-5.el7.x86_64.rpm from http://ftp.pbone.net/mirror/download.fedora.redhat.com/pub/fedora/epel/7/x86_64/s/sshpass-1.05-5.el7.x86_64.rpm
- soci-3.2.3-1.el7.x86_64.rpm from http://ftp.pbone.net/mirror/download.fedora.redhat.com/pub/fedora/epel/7/x86_64/s/soci-3.2.3-1.el7.x86_64.rpm

- soci-mysql-3.2.3-1.el7.x86_64.rpm from http://ftp.pbone.net/mirror/download.fedora.redhat.com/pub/fedora/epel/7/x86_64/soci-mysql-3.2.3-1.el7.x86_64.rpm
- mapr-setup from <http://package.mapr.com/releases/v5.2.0/redhat/mapr-setup>
- ext-2.2.zip from <http://dev.sencha.com/deploy/ext-2.2.zip>
- pdsh-2.31-1.el7.x86_64.rpm from https://dl.fedoraproject.org/pub/epel/7/x86_64/p/pdsh-2.31-1.el7.x86_64.rpm

MapR 5.2.1 Packages

Download the following packages to /opt/ensaroot/bd-sw-rep/MapR-5.2.1

- mapr-ecosystem-5.x-20170405.rpm.tgz from <http://archive.mapr.com/releases/ecosystem-5.x/redhat/mapr-ecosystem-5.x-20170405.rpm.tgz>
- mapr-v5.2.1GA.rpm.tgz from <http://archive.mapr.com/releases/v5.2.1/redhat/mapr-v5.2.1GA.rpm.tgz>
- ext-2.2.zip from <http://172.26.225.242:9000/mount/MapR-5.2.0/ext-2.2.zip>
- mapr-setup from <http://archive.mapr.com/releases/v5.2.1/redhat/mapr-setup>
- libgenders-1.22-2.el7.x86_64.rpm from http://10.29.160.57/MapR-5.2.0/libgenders-1.22-2.el7.x86_64.rpm
- libgenders-devel-1.22-2.el7.x86_64.rpm from http://10.29.160.57/MapR-5.2.0/libgenders-devel-1.22-2.el7.x86_64.rpm
- pdsh-2.31-1.el7.x86_64.rpm from http://10.29.160.57/MapR-5.2.0/pdsh-2.31-1.el7.x86_64.rpm
- pdsh-rcmd-rsh-2.31-1.el7.x86_64.rpm from http://10.29.160.57/MapR-5.2.0/pdsh-rcmd-rsh-2.31-1.el7.x86_64.rpm
- rpmrebuild-2.4-1.noarch.rpm from <http://10.29.160.57/MapR-5.2.0/rpmrebuild-2.4-1.noarch.rpm>
- scala-2.10.3.rpm from <http://10.29.160.57/MapR-5.2.0/scala-2.10.3.rpm>
- soci-3.2.3-1.el7.x86_64.rpm from http://10.29.160.57/MapR-5.2.0/soci-3.2.3-1.el7.x86_64.rpm
- soci-mysql-3.2.3-1.el7.x86_64.rpm from http://10.29.160.57/MapR-5.2.0/soci-mysql-3.2.3-1.el7.x86_64.rpm
- sshpass-1.05-5.el7.x86_64.rpm from http://10.29.160.57/MapR-5.2.0/sshpass-1.05-5.el7.x86_64.rpm

Common Package for Hortonworks

Download the following common package to /opt/ensaroot/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 Bare Metal Agent)

Hortonworks 2.1 Packages

Download the following packages to /opt/ensaroot/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>

Hortonworks 2.4 Packages

For RHEL 6.x, download the following packages to `/opt/cnsaroot/bd-sw-rep/Hortonworks-2.4`:

- HDP-2.4.0.0-centos6-rpm.tar.gz from <http://public-repo-1.hortonworks.com/HDP/centos6/2.x/updates/2.4.0.0/HDP-2.4.0.0-centos6-rpm.tar.gz>
- ambari-2.2.2.0-centos6.tar.gz from <http://public-repo-1.hortonworks.com/ambari/centos6/2.x/updates/2.2.2.0/ambari-2.2.2.0-centos6.tar.gz>
- HDP-UTILS-1.1.0.20-centos6.tar.gz from <http://public-repo-1.hortonworks.com/HDP-UTILS-1.1.0.20/repos/centos6/HDP-UTILS-1.1.0.20-centos6.tar.gz>

For RHEL 7.x, download the following packages to `/opt/cnsaroot/bd-sw-rep/Hortonworks-2.4`:

- HDP-2.4.0.0-centos7-rpm.tar.gz from <http://public-repo-1.hortonworks.com/HDP/centos7/2.x/updates/2.4.2.0/HDP-2.4.2.0-centos7-rpm.tar.gz>
- ambari-2.2.2.0-centos7.tar.gz from <http://public-repo-1.hortonworks.com/ambari/centos7/2.x/updates/2.2.2.0/ambari-2.2.2.0-centos7.tar.gz>
- HDP-UTILS-1.1.0.20-centos7.tar.gz from <http://public-repo-1.hortonworks.com/HDP-UTILS-1.1.0.20/repos/centos7/HDP-UTILS-1.1.0.20-centos7.tar.gz>

Hortonworks 2.5 Packages

For RHEL 6.x, download the following packages to `/opt/cnsaroot/bd-sw-rep/Hortonworks-2.5`:

- HDP-2.5.0.0-centos6-rpm.tar.gz from <http://public-repo-1.hortonworks.com/HDP/centos6/2.x/updates/2.5.0.0/HDP-2.5.0.0-centos6-rpm.tar.gz>
- ambari-2.4.1.0-centos6.tar.gz from <http://public-repo-1.hortonworks.com/ambari/centos6/2.x/updates/2.4.1.0/ambari-2.4.1.0-centos6.tar.gz>
- HDP-UTILS-1.1.0.21-centos6.tar.gz from <http://public-repo-1.hortonworks.com/HDP-UTILS-1.1.0.21/repos/centos6/HDP-UTILS-1.1.0.21-centos6.tar.gz>

For RHEL 7.x, download the following packages to `/opt/cnsaroot/bd-sw-rep/Hortonworks-2.5`:

- HDP-2.5.0.0-centos7-rpm.tar.gz from <http://public-repo-1.hortonworks.com/HDP/centos7/2.x/updates/2.5.0.0/HDP-2.5.0.0-centos7-rpm.tar.gz>
- ambari-2.4.1.0-centos7.tar.gz from <http://public-repo-1.hortonworks.com/ambari/centos7/2.x/updates/2.4.1.0/ambari-2.4.1.0-centos7.tar.gz>
- HDP-UTILS-1.1.0.21-centos7.tar.gz from <http://public-repo-1.hortonworks.com/HDP-UTILS-1.1.0.21/repos/centos7/HDP-UTILS-1.1.0.21-centos7.tar.gz>

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>

MapR 5.2.0 Packages

- mapr-setup from <http://package.mapr.com/releases/v5.2.0/redhat/mapr-setup>
- mapr-v5.2.0GA.rpm.tgz: from <http://archive.mapr.com/releases/v5.2.0/redhat/mapr-v5.2.0GA.rpm.tgz>

**Note**

mapr-v5.2.0GA.rpm.tgz contains the following
 mapr-client-5.2.0.39122.GA-1.x86_64.rpm,
 mapr-posix-client-platinum-5.2.0.39122.GA-1.x86_64.rpm,
 mapr-posix-client-basic-5.2.0.39122.GA-1.x86_64.rpm,shalsum.txt,
 mapr-upgrade-5.2.0.39122.GA-1.x86_64.rpm, mapr-nfs-5.2.0.39122.GA-1.x86_64.rpm,
 and mapr-core-5.2.0.39122.GA-1.x86_64.rpm files.

- mysql-connector-java-5.1.26.tar.gz from <https://downloads.mysql.com/archives/get/file/mysql-connector-java-5.1.26.tar.gz>
- mapr-ecosystem-5.x-20160816.rpm.tgz from <http://archive.mapr.com/releases/ecosystem-all/redhat/mapr-ecosystem-5.x-20160816.rpm.tgz>

Installation of User-Defined Software Post Hadoop Cluster Creation

Cisco UCS Director Express for Big Data provides an option to add user-defined installation packages (RPMs) post Hadoop cluster creation specific to a version. In Cisco UCS Director Express for Big Data, you cannot install more Hadoop related software other than what is required for the selected type of Hadoop distribution when creating an instant Hadoop cluster or customizing a Hadoop cluster.

To install user-defined Hadoop related software, you can specify a list of RPMs in the HadoopDistributionRPM.txt file. This modifiable list defines the required packages for each version of a Hadoop distribution. You can locate the HadoopDistributionRPM.txt here:

`/opt/cnsaroot/bigdata_templates/common_templates` in the

Cisco UCS Director Express for Big Data Bare Metal Agent server.

For example, you can specify a list of RPMs in the HadoopDistributionRPM.txt file for MapR-5.0.0:

- mapr-v5.0.0GA.rpm.tgz
- mapr-ecosystem-5.x-20150709.rpm.tgz
- mapr-whirr-0.8.1.18380-GA.noarch.rpm

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** page 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** Choose **Solutions > Big Data > Accounts**.
 - Step 2** Click **Hadoop Accounts**.
 - Step 3** Choose the account for which you want to run the configuration check and click **Check Configuration**.
 - Step 4** Click **Submit**.
A page 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** page for the selected Hadoop Account.
-

What to Do Next



Note

You can track configuration checks here: **Administration > Integration**. Click **Change Record** 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** page 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** Choose **Solutions > Big Data > Accounts**.
 - Step 2** Click **Faults**.
 - 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.
-



Cisco UCS CPA Workflows

This chapter contains the following sections:

- [Workflows for Big Data, page 145](#)
- [About Service Requests for Big Data, page 147](#)

Workflows for Big Data

Cisco UCS Director Express for Big Data defines a set of workflows in the **UCS CPA** folder under **Orchestration** page.

- **UCS CPA Multi-UCSM Hadoop Cluster WF**—This workflow is triggered if you use multiple UCSM Hadoop accounts.
- **UCS CPA Multi-UCSM Splunk Cluster WF**—This workflow is triggered if you use multiple UCSM Splunk accounts.
- **UCS CPA Single UCSM Server Configuration WF**—This workflow is triggered for every UCSM account.
- **UCS CPA Node Bare Metal**—This workflow is triggered per node in the cluster.
- **UCS CPA Node Bare Metal with Storage Profile WF**—This workflow is triggered using Cisco UCS S3260 storage server having SSD Server Boot Drive under SAS RAID controller to deploy OS.

Table 8: Disk Group Policy

Servers Details	Disk Group Policy Required?	Disk Policy Allocation Details
Servers having PCH Controller and and Data drive slots with SSD disks	No	Not applicable

Servers Details	Disk Group Policy Required?	Disk Policy Allocation Details
SAS controller with dedicated boot drive slots (20x) and Data drive slots with SSD disks	Yes	For OS boot, if you want to manually allocate the SSD slot numbers to RAID 0 or RAID 1, BD_V1_SSM_RAID0 or BD_V1_SSM_RAID1 should be created. Currently dedicated boot drive slots 201 and 202 are supported only by Cisco UCS S3260 Storage Server.
SAS controller with dedicated boot drive slots (20x) with SSD disks and Data drive slots without SSD disks	Yes	For OS boot, if you want to automatically allocate the SSD slot numbers to RAID 0 or RAID 1, Boot_SSD_RAID0 or Boot_SSD_RAID1 should be created.
SAS controller and Data drive slots with HDD only	Yes	For OS boot, if you want to automatically allocate the HDD slot numbers to RAID 0 or RAID 1, Boot_HDD_RAID0 or Boot_HDD_RAID1 should be created. Note You can create Boot_HDD_RAID0 or Boot_HDD_RAID1 if SSD disks are available in Data drive slots in addition to HDD.

**Note**

In some cases, Wait for Complete Association might fail for Cisco UCS 326x storage servers during UCS CPA Node Bare Metal with Storage Profile after a time out. In such cases, association process might have already failed after Bind UCS Service Profile to Template post 90% of FSM with Storage System initialisation error. You need to perform a re-acknowledgement on the Cisco UCS server and resubmit the Bare Metal workflow.

- UCS CPA Delete Node—This workflow is triggered if you delete a node from the Hadoop cluster.
- UCS CPA Cloudera Add Live Node—This workflow is triggered if you add a Live Node for Cloudera.
- UCS CPA MapR Add Live Node—This workflow is triggered if you add a Live Node for MapR.
- UCS CPA Hortonworks Add Live Node—This workflow is triggered if you add a Live Node for Hortonworks.
- UCS CPA BigInsights Add Live Node—This workflow is triggered if you add a Live Node for BigInsights.

- UCS CPA Instant Hadoop Cluster WF—This workflow is triggered if you create an instant Hadoop cluster based on the node count and other mandatory inputs such as the IP address range, memory, and number of interfaces. Cisco UCS Director Express for Big Data automatically creates one UCS service profile and a Hadoop cluster profile template at the back end that are required to create an instant Hadoop cluster. This saves you the effort of manually creating a service profile and a cluster profile.
- UCS CPA Customized Hadoop Cluster WF—This workflow is triggered if you choose to use a specific UCS service profile. A Hadoop cluster profile template with the specified number of nodes to be created in the Hadoop cluster.
- UCS CPA Disable High Availability WF—This workflow is triggered when you disable high availability.
- UCS CPA Enable High Availability WF—This workflow is triggered when you enable high availability.
- UCS CPA Shutdown Big Data Cluster WF—This workflow is triggered when you shut down the Hadoop cluster.
- UCS CPA Start Big Data Cluster WF—This workflow is triggered when you power up the Hadoop cluster.
- UCS CPA Splunk Add Live Archival Node—This workflow is triggered when if you add a Live Archival Node for Splunk.
- UCS CPA Upgrade Cluster WF—This workflow is triggered when you upgrade the Hadoop cluster.
- UCS CPA Cloudera Add New Service WF—This workflow is triggered when you add a new service for Cloudera.
- UCS CPA MapR Add New Service WF—This workflow is triggered when you add a new service for MapR.
- UCS CPA Hortonworks Add New Service WF—This workflow is triggered when you add a new service for Hortonworks.
- UCS CPA BigInsights Add New Service WF—This workflow is triggered when you add a new service for BigInsights.
- UCS CPA Wait For Complete UCS SP Association WF—This workflow is triggered .

About Service Requests for Big Data

Cisco UCS Director Express for Big Data leverages Cisco UCS Director service requests and workflow orchestration for the overall deployment of Hadoop clusters. Each service request is a provisioning workflow that is created during a cluster creation.

For example, one UCS CPA Multi-UCSM Hadoop Cluster W/F, one Single UCSM Server Configuration W/F, and four UCS CPA Node Bare Metal W/Fs are created for a four-node Hadoop cluster. When the workflows are complete, the cluster is created under **Solutions > Big Data > Accounts** for that UCSM account.

A set of service requests is created under **Organizations > Service Requests** during a cluster creation.

- UCS CPA Multi-UCSM Hadoop Cluster WF—This workflow is triggered if you use multiple UCSM Hadoop accounts. It also applies when you create an instant or customized Hadoop cluster.
- UCS CPA Multi-UCSM Splunk Cluster WF—This workflow is triggered if you use multiple UCSM Splunk accounts. It also applies to when you create an instant or customized Splunk cluster.

- UCS CPA Single UCSM Server Configuration WF—This workflow is triggered for every UCSM account.
- UCS CPA Node Bare Metal—This workflow is triggered per node in the cluster.

The following service requests are created when you add a Bare Metal Node or a Live Node to the cluster.

- UCS CPA Cloudera Add Live Node—This workflow is triggered if you add a Live Node for Cloudera.
- UCS CPA MapR Add Live Node—This workflow is triggered if you add a Live Node for MapR.
- UCS CPA Hortonworks Add Live Node—This workflow is triggered if you add a Live Node for Hortonworks.
- UCS CPA BigInsights Add Live Node—This workflow is triggered if you add a Live Node for BigInsights.

The following service requests are created as explained, below:

- UCS CPA Disable High Availability WF—This workflow is triggered when you disable high availability.
- UCS CPA Enable High Availability WF—This workflow is triggered when you enable high availability.
- UCS CPA Shutdown Big Data Cluster WF—This workflow is triggered when you shut down the Hadoop cluster.
- UCS CPA Start Big Data Cluster WF—This workflow is triggered when you power up the Hadoop cluster.
- UCS CPA Upgrade Cluster WF—This workflow is triggered when you upgrade the Hadoop cluster.
- UCS CPA Cloudera Add New Service WF—This workflow is triggered when you add a new service for Cloudera.
- UCS CPA Hortonworks Add New Service WF—This workflow is triggered when you add a new service for Hortonworks.
- UCS CPA MapR Add New Service WF—This workflow is triggered when you add a new service for MapR.
- UCS CPA BigInsights Add New Service WF—This workflow is triggered when you add a new service for BigInsights.

For more information on the service requests and workflows, see the following guides:

- *Cisco UCS Director Self-Service Portal Guide*
- *Cisco UCS Director Orchestration Guide*

Monitoring Service Requests for Big Data

Before You Begin

Create and customize a Cluster Deploy Template to monitor a service request.

-
- Step 1** Choose **Organizations > Service Requests**.
- Step 2** Click **Service Requests**.
- Step 3** Select the service request that you want to monitor and click **View Details**. One of the following Request Status is displayed:
- Complete
 - In Progress
 - Cancelled
 - Failed
-

Viewing UCS CPA Workflow Tasks

From the **Service Request Status** screen, you can view the following:

- Workflow Status
- Log
- Objects Created and Modified
- Input/Output



Note You can only modify inputs for failed service requests.

-
- Step 1** Choose **Organizations > Service Requests**. You can see the list of user-specific service requests added to a specific group. For example, All User Groups.
- Step 2** Choose the **Service Request ID** that you want to view.
- Step 3** Double-click the **Service Request ID** that opens the **Service Request Status** screen. (You can also choose the Service Request ID by the workflow name associated with it, and click **View Details**. For example, choose the **UCS CPA Node Bare Metal** workflow and click **View Details**). On the **Service Request Status** screen, you can view the following tasks for the workflow:

UCS CPA Multi-UCSM Hadoop Cluster WF/UCS CPA Multi-UCSM Splunk Cluster WF	UCS CPA Single UCSM Server Configuration WF	UCS CPA Node Bare Metal
<p>The following tasks are associated with the UCS CPA Multi-UCSM Hadoop Cluster WF/UCS CPA Multi-UCSM Splunk Cluster WF:</p> <ol style="list-style-type: none"> 1 Initiated by Admin 2 Multi-UCSM Hadoop Cluster profile/Multi-UCSM Splunk Cluster profile 3 Setup Big Data Cluster Env 4 User (Compute) Approval Required 5 User (OS) Approval Required 6 Multi-UCSM Configuration WF 7 Multi Bare Metal WF Monitor 8 Synchronized Command Execution 9 User (Hadoop) Approval Required 10 Custom SSH Command 11 Monitor Shell Script Status 12 Provision Hadoop Cluster/Provision Splunk Cluster 13 Complete 	<p>The following tasks are associated with the Single UCSM Server Configuration WF:</p> <ol style="list-style-type: none"> 1 Initiated by Admin 2 Create UCS Service Profile Template 3 Change Maintenance Policy UCS SP Template 4 Muti-Bare Metal OS Install WF 5 Multi-Bare Metal WF Monitor 6 Change Maintenance Policy for UCS SP Template 7 Complete 	

UCS CPA Multi-UCSM Hadoop Cluster WF/UCS CPA Multi-UCSM Splunk Cluster WF	UCS CPA Single UCSM Server Configuration WF	UCS CPA Node Bare Metal
		<p>The following tasks are associated with the UCS CPA Node Bare Metal:</p> <ol style="list-style-type: none"> 1 Initiated by Admin 2 Modify Workflow Priority (High) 3 Manage DB Lock 4 Assign Bare Metal SR ID 5 Manage DB Lock 6 Create UCS Service Profile from Template 7 Unbind UCS Service Profile from Template 8 Modify UCS Service Profile Boot Policy 9 Unbind UCS ServiceProfile vNIC from Template 10 Bind UCS vNIC to Template 11 Associate UCS Service Profile 12 Manage DB Lock 13 Assign ServerIdentity 14 Manage DB Lock 15 Setup PXE Boot (OS Type: CentOS Live) 16 Manage DB Lock 17 Setup RAID Commands 18 Manage DB Lock 19 UCS Blade Reset Action 20 Monitor PXE Boot 21 Monitor RAID Configuration 22 Archive PXE Boot Request 23 UCS Blade Power OFF Action 24 Setup PXE Boot (OS Type: RHEL6.5) 25 Manage DB Lock 26 Setup RAID Commands 27 Manage DB Lock 28 Wait for complete Association 29 UCS Blade Reset Action 30 Monitor PXE Boot 31 Modify UCS Service Profile Boot Policy 32 Bind UCS Service Profile to

UCS CPA Multi-UCSM Hadoop Cluster WF/UCS CPA Multi-UCSM Splunk Cluster WF	UCS CPA Single UCSM Server Configuration WF	UCS CPA Node Bare Metal
		Template 33 Wait for complete Association 34 UCS Blade Reset Action 35 Assign IP Status 36 Custom SSH Command 37 Custom SSH Command 38 Synchronized Command Execution 39 Linux Shell Script Execution 40 UCS Blade Power OFF Action 41 UCS Blade Power ON Action 42 Synchronized Command Execution 43 Complete

Step 4 Click **Close**.

Workflow Customization to Deploy a Hadoop or Splunk Cluster

You can customize the following UCS CPA workflows and use them to deploy a Hadoop or Splunk cluster. You can add installation packages (RPMs) required for your cluster environment in the Cisco UCS Director Express for Big Data Bare Metal Agent.

- Rename the UCS CPA Multi-UCSM Hadoop Cluster WF
- Rename the UCS CPA Multi-UCSM Splunk Cluster WF
- Rename the Single UCSM Server Configuration WF
- Rename the UCS CPA Node Bare Metal

Deploying a Hadoop or Splunk Cluster Through Workflow Customization

Before You Begin

For more information on workflow orchestration, see the *Cisco UCS Director Orchestration Guide*.

- Customize UCS CPA Node Bare Metal workflows that you want to use in the cloned Single UCSM Server Configuration WF.

- Customize the Single UCSM Server Configuration WF that you want to use in the UCS CPA Multi-UCSM Hadoop Cluster WF or UCS CPA Multi-UCSM Splunk Cluster WF.

-
- Step 1** Choose **Orchestration**.
- Step 2** Click the **UCS CPA** folder from **Workflows**.
- Step 3** Double-click the workflow that you want to customize in the **Workflow Designer**. For instance, double-click the **UCS CPA Multi-UCSM Hadoop Cluster WF**.
- Step 4** Double-click the **Muti-UCSM Configuration WF** task in the **Workflow Designer**.
- Step 5** Click **Next** on the **Workflow Task Basic Information** page.
- Step 6** On the **User Input Mappings to Task Input Attributes** page, select the attributes that you want to map to the workflow input fields or provide values in the next step. If necessary, check the **Map to User Input** check box to provide user inputs.
- Step 7** Click **Next** on the **User Input Mappings to Task Input Attributes** page.
- Step 8** Enter the task values which are not mapped to workflow inputs. For example, enter the name of the cloned Single UCSM Server Configuration WF in the **Workflow Name** field.
- Step 9** Click **Revalidate** to validate task input values.
- Step 10** Click **Next**.
- Step 11** On the **User Output mappings to Task Output Attributes** page, select the attributes that you want to map to the workflow output fields.
- Step 12** Check the **Map to User Output** check box and choose the value from the **User Output** drop-down list.
- Step 13** Click **Submit**.
-

Assigning Big Data Accounts to User Groups

-
- Step 1** Choose **Orchestration** and click **Workflows**.
- Step 2** Click **Add Workflow**.
- Step 3** On the **Add Workflow Details** page, enter the workflow name and choose a folder. Click **Next**.
- Step 4** On the **Workflow Task Basic Information** page, enter the required details and click **Next**.
- Step 5** On the **User Input Mappings to Task Input Attributes** page, select the attributes that you want to map to the workflow input fields or provide values in the next step. If necessary, check the **Map to User Input** check box to provide user inputs..
- Step 6** On the **User Input Mappings to Task Input Attributes** page, choose big data account type and big data account.
- Step 7** Check the **Assign Users** checkbox and click **Submit**. The big data accounts are assigned to specified users if the **Allow resource assignment to users** option is enabled in the **Users Groups** page. When the **Assign Users** checkbox is not checked, the big data accounts are assigned to a specified group.
-

Unassigning Big Data Accounts

Before You Begin

-
- Step 1** Choose **Orchestration** and click **Workflows**.
- Step 2** Click **Add Workflow**.
- Step 3** On the **Add Workflow Details** page, enter the workflow name and choose a folder. Click **Next**.
- Step 4** On the **Workflow Task Basic Information** page, enter the required details and click **Next**.
- Step 5** On the **User Input Mappings to Task Input Attributes** page, select the attributes that you want to map to the workflow input fields or provide values in the next step. If necessary, check the **Map to User Input** check box to provide user inputs..
- Step 6** On the **User Input Mappings to Task Input Attributes** page, choose big data account type and big data account.
- Step 7** Check the **Unassign Users** checkbox and click **Submit**. The big data accounts are unassigned from the group.
-

Cloning UCS CPA Workflows

To customize cluster deployment through bare metal workflows, you can clone the following workflows in the UCS CPA folder:

- Clone the UCS CPA Node Bare Metal workflows.
- Rename the Single UCSM Server Configuration WF using the cloned UCS CPA Node Bare Metal workflows.
- Rename the UCS CPA Multi-UCSM Hadoop Cluster WF using the cloned Single UCSM Server Configuration WF.

-
- Step 1** Choose **Orchestration**.
- Step 2** Click the **UCS CPA** folder from **Workflows** and choose the workflow that you want to clone.
- Step 3** Click **Clone Workflow**.
- Step 4** On the **Workflow Details** page of the **Clone Workflow** wizard, complete the following fields:

Name	Description
Workflow Name field	A unique name for the workflow.
Version field	The current version of the workflow that you are cloning. This is a display-only field.
Description field	The description of the workflow.

Name	Description
Workflow Context drop-down list	<p>The workflow context. Workflow Orchestration supports the following options:</p> <ul style="list-style-type: none"> • Any—Enables you to use the workflow in any context. • Selected VM—Enables you to use the execute workflow. This option can be selected only when you choose a VM. • Check the Save As Compound Task check box to define the workflow as a compound task. • Check the Place in New Folder check box, and enter the folder name in the Folder Name field, to assign the workflow to a new folder other than the UCS CPA folder.
Select Folder drop-down list	Choose a folder. UCS CPA is the default folder for Big Data.
Notify status of execution to initiator User check box	Check the check box to notify the user through email, then enter appropriate email addresses in the Additional User(s) to send Email Notification field.

Step 5 Click Next.

Step 6 On the **Workflow User Inputs** page of the **Clone Workflow** wizard, complete the following fields:

Name	Description
Associate to Activity check box	If the check box is checked then any existing workflow's user input is overridden by any selected activity user input.
Activity drop-down list.	Choose an activity. The user-input table is updated based on the selected activity.

Name	Description
Workflow User Inputs table	<p>On the Workflow User Inputs page:</p> <ol style="list-style-type: none"> 1 Click the + icon to add workflow input properties. 2 On the Add Entry to screen, complete the following fields: <ol style="list-style-type: none"> 1 Enter the name for the activity workflow input in the Input Label field. 2 Enter the description for the activity workflow input in the Input Description field. 3 Check the Optional check box to set the input as optional during workflow execution. 4 Click Select. On the Select screen, click Input Type. 5 Click Submit.

Step 7 Click **Next**.

Step 8 On the **Workflow User Outputs** page of the **Clone Workflow** wizard, do the following:

Name	Description
Workflow User Outputs table	<p>On the Workflow User Outputs page:</p> <ol style="list-style-type: none"> 1 Click the + icon to add workflow output properties. 2 On the Add Entry to screen, complete the following fields: <ol style="list-style-type: none"> 1 Enter the name for the activity workflow output in the Output Label field. 2 Enter the description for the activity workflow output in the Output Description field. 3 Check the Optional check box to set the output as optional during workflow execution. 4 Click Select. On the Select screen, click Output Type. 5 Click Submit.

Step 9 Click **Submit**.



Monitoring and Reporting

This chapter contains the following sections:

- [About Monitoring and Reporting](#), page 159
- [Cisco UCS Director Express for Big Data Dashboard](#), page 159
- [Viewing a Deployed Cluster Report](#), page 160
- [Reports](#), page 160

About Monitoring and Reporting

Cisco UCS Director Express for Big Data can monitor virtual infrastructure and system resources, and provide a wide array of reports.

Cisco UCS Director Express for Big Data monitors a range of cluster events:

- High CPU usage
- Memory usage
- Disk capacity
- Disk IO utilization

Cisco UCS Director Express for Big Data displays statistics from the respective pages for selected Big Data Account and Hosts. You can also generate reports that itemize system details.

Cisco UCS Director Express for Big Data Dashboard

Cisco UCS Director Express for Big Data provides complete system visibility through real-time and historical monitoring. See [Reports](#)



Note

The **Dashboard** tab shows up in the menu bar only after a summary report is added to the dashboard.

The customizable dashboard displays processing, memory, storage, and network utilization metrics.

- Per-node statistics: CPU, memory, and disk
- Health of Hadoop cluster components: HDFS, MapReduce jobs
- Graphs based on historical data



Note Drag and Drop summary report icons from the **Customize** drawer to the **Summary** and **Hosts** tabs, where you can expand and close reports.

You can:

- 1 Add summary reports to the dashboard from the Big Data Account **Summary** tab. You can customize the summary reports to display the statistics for a specific time period, or export these reports from the dashboard.
- 2 Add summary reports to the dashboard from the Big Data Account **Hosts** tab. You can customize the summary reports to display the statistics for a specific time period, or export these reports from the dashboard.



Note Not all summary reports apply to MapR.

- 3 Add UCSM Accounts summary reports to the dashboard from **Physical > Compute**.
- 4 Add Data Center summary reports to the dashboard from **Physical > Compute**.

Viewing a Deployed Cluster Report

You can generate a **Big Data Account Summary Report** with or without credentials, to view the details of the deployed clusters.

- Step 1** On the menu bar, choose **Solutions > Big Data > Containers**.
- Step 2** Click the **Deployed Clusters** tab.
- Step 3** Choose the deployed cluster, and click **View Report**.
- Step 4** In the **View Report** dialog box, choose the report type that you want to generate.
- Step 5** Click **Submit**.

Reports

UCSM Account Summary Reports

- Overview
- UCS Chassis Inventory

- UCS Server Inventory
- UCS Fabric Interconnect Inventory
- UCS Servers Associated vs Unassociated
- Rack Server Discovery Policy
- Power Policy
- Global Power Allocation Policy

Big Data Account Summary Reports

- Overview

Cluster-specific Metrics Supported per Hadoop Distribution

Metrics	Cloudera Distribution	Hortonworks Distribution	MapR Distribution	Remarks
Average CPU IO Wait (%)	Yes	Yes	Yes	—
Average CPU idle (%)	Yes	Yes	Yes	—
Average CPU Nice (%)	Yes	Yes	Yes	—
Average CPU System (%)	Yes	Yes	Yes	—
Average CPU User	Yes	Yes	Yes	—
CPU Percentage Across Hosts	Yes	—	—	Metrics for Host CPU usage across hosts
Cluster Disk IO	Yes	—	—	Metrics for Total Disk Write and Read Bytes Across Disks
Cluster Network IO	Yes	—	—	Metrics for Total Bytes Transmitted and Received Across Network Interfaces
HDFS IO	Yes	—	—	Metrics for Total Bytes written and Read Across Data Nodes

Metrics	Cloudera Distribution	Hortonworks Distribution	MapR Distribution	Remarks
Total Space Utilization	—	Yes	—	—
CPU	—	Yes	—	—
Load	—	Yes	—	—
Memory	—	Yes	—	—
Network	—	Yes	—	—
Process	—	Yes	—	—

Host-specific Metrics Supported per Hadoop Distribution

Metrics	Cloudera Distribution	Hortonworks Distribution	MapR Distribution
Average CPU IO Wait (%)	Yes	Yes	Yes
Average CPU Idle (%)	Yes	Yes	Yes
Average CPU Nice (%)	Yes	Yes	Yes
Average CPU System (%)	Yes	Yes	Yes
Average CPU User	Yes	Yes	Yes
Load Average(%)	Yes	—	—
Host CPU Usage	Yes	—	—
Host Memory Usage	Yes	—	—
Host Network Throughput	Yes	—	—
Disk Latency	Yes	—	—
Aggregate Disk Throughput	Yes	—	—
Role-CPU Usage	Yes	—	—
HOST CPU	—	Yes	—
DISK Usage	—	Yes	—

Metrics	Cloudera Distribution	Hortonworks Distribution	MapR Distribution
LOAD Usage	—	Yes	—
Memory-CPU Usage	—	Yes	—
Network Usage	—	Yes	—
Process Usage	—	Yes	—



Proactive Status Monitoring and Diagnostics

This chapter contains the following topics:

- [Aggregate CPU, Disk, and Network Bandwidth Utilization, page 165](#)
- [Monitoring Aggregate CPU, Disk, and Network Bandwidth Utilization, page 166](#)
- [Monitoring Top Jobs Based on CPU Utilization and Time, page 166](#)
- [Performance Metrics for CPU, Disk, and Network, page 167](#)
- [Viewing CPU, Disk, and Network Statistics for a Hadoop Cluster, page 168](#)
- [Analyzing Performance Bottlenecks Through Historical Metrics, page 169](#)
- [Setting Alerts for Hadoop Cluster Service Failures, page 169](#)
- [Types of Disk and Network Failure Alerts, page 170](#)
- [Setting Alerts for Disk and Network Failures, page 172](#)
- [Setting Disk Utilization Threshold Alerts, page 173](#)

Aggregate CPU, Disk, and Network Bandwidth Utilization

You can monitor the aggregate CPU, disk, and network bandwidth utilization across all the hosts in a cluster. The metrics are collected in the following ways:

- **Aggregate CPU and Disk metrics:** For every host that is running the job, the PID collects the percentage of CPU and memory used by the job. The sum of all these percentages gives the aggregate CPU and disk metrics.
- **Aggregate network bandwidth metrics:** For aggregate network bandwidth of one node, obtain the network bandwidth on each network interface, and then add them. Similarly network bandwidths are measured for all the nodes in the cluster. The sum of all these bandwidths provides the aggregate network bandwidth metrics for the cluster.
- **Duration of long-running jobs:** A Rest API collects the start time, elapsed time, and end time for each job identified on the cluster. The difference between start time and end time provides the duration of completed jobs. The elapsed time reports the duration of the jobs running currently.

Monitoring Aggregate CPU, Disk, and Network Bandwidth Utilization

-
- Step 1** On the menu bar, choose **Solutions > Big Data > Accounts**.
- Step 2** Click the **Big Data Accounts** tab.
- Step 3** Choose the Big Data Account and click **View Details**.
- Step 4** Click the **Hadoop Clusters** tab.
- Step 5** Choose the big data cluster and click **View Reports**.
- Step 6** Click the **Monitoring** tab.
Every time an inventory collection cycle is triggered, an entry listing the aggregate CPU, network bandwidth, and disk utilization metrics appears on the Monitoring Page.
- Note** For Splunk cluster, on clicking the **View Details** button, the **Monitoring** tab is displayed. Step 4 and Step 5 are specific to Hadoop cluster only.
- Step 7** Select the entry you want to analyze and click **View Details**.
- Click the **Aggregate CPU** tab to view the aggregate CPU utilization of all nodes for a particular time period.
 - Click the **Aggregate Disks** tab to view the aggregate disk utilization and available memory across the cluster.
 - Click the **Aggregate Network Bandwidth Utilization** to view the aggregated network bandwidth across the cluster.
- Step 8** Click **Back** to return to the **Monitoring** page.
-

Monitoring Top Jobs Based on CPU Utilization and Time

To monitor top jobs based on CPU utilization or time (both active and completed long-running jobs), 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 Big Data Account and click **View Details**.
- Step 4** Click the **Hadoop Clusters** tab.
- Step 5** Choose the Hadoop cluster and click **View Reports**.
- Click the **Top 10 High CPU Jobs** tab to view the top ten jobs, based on CPU utilization.
 - Click the **Top 10 Long Running Active Jobs** tab to view the current top ten long-running jobs.
 - Click the **Top 10 Long Duration Jobs** tab to view the completed top ten long-running jobs.
- Step 6** Click **Back** to return back to the **Hadoop Clusters** page.
-

Performance Metrics for CPU, Disk, and Network

You can find performance bottlenecks that occur in the compute, network, or Hadoop setup across the cluster. You can collect CPU, disk, and network metrics and analyze these metrics to fix bottlenecks.

The metrics reports are of the following types:

- **Pre-Cluster:** This metrics report is generated automatically for a server that has been installed with Red Hat Linux. This report is created before the server becomes part of a cluster.
- **Post-Cluster:** This metrics report is generated on demand when you run the performance test for a Hadoop cluster.

When you run the performance test for a Hadoop cluster, the following metrics are shown in detail:

- **Memory metrics:** Memory metrics measure the memory utilization of each host on the Hadoop cluster. The report includes the triad rate, which is the average rate at which read, write, and copy operations take place. The triad rate is a standard measure of memory bandwidth.
- **Network metrics:** Network metrics measure the network bandwidth of the Hadoop cluster. The report displays the rates at which network packets are transferred between the client and the server in the Hadoop cluster.
- **Disk metrics:** Disk metrics identify how fast a disk can perform. The disk metrics are included only in the pre-cluster report. The report lists the following:
 - The time taken to read and write a file.
 - The time taken to rewrite to an existing a file.
 - The time to randomly (nonsequentially) read and write files.
- **DFSIO metrics:** The DFSIO test is a Hadoop benchmark that stress-tests the storage I/O (read and write) capabilities of the cluster. The report measures the bytes processed, execution time, the average I/O rate, and throughput to read and write multiple files. The DFSIO metrics report is included only in the post-cluster report.
- **TeraSort metrics:** The TeraSort test is a Hadoop benchmark that tests the memory of the cluster. The report lists the counters for generating input, sorting the generated input, and validating the sorted output. The TeraSort metrics report is included only in the post-cluster report.

Viewing CPU, Disk, and Network Statistics for a Hadoop Cluster

You can collect and compare CPU, disk, and network metrics with the pre-cluster creation and post-cluster creation reports for a Hadoop cluster.

-
- Step 1** On the menu bar, choose **Solutions > Big Data >> Accounts**.
- Step 2** Click the **Big Data Accounts** tab.
- Step 3** Choose the Big Data Account and click **View Details**.
- Step 4** Click the **Hadoop Clusters** tab.
- Step 5** Choose the Hadoop cluster and click **View Reports**.
- Step 6** Click the **Performance** tab.
- Step 7** Click **Run Test**.
The Performance tab displays a default Big Data Metrics Report. This report shows the statistics collected for each host before the Hadoop cluster creation and the reports collected after Hadoop cluster creation.
- Step 8** Click **Submit**, and then click **OK**.
For the following actions, choose the performance report:

Name	Description
View	Displays the metrics in the Big Data Metrics Report.
Compare	Compares and displays the metrics in the Big Data Metrics Report.
View Graph Report	Displays graphically the following reports from the Summary tab: <ul style="list-style-type: none"> • Average TRIAD Rate (MB/Sec) • Average Network Bandwidth (MB/Sec) • Average DFSIO Write (MB/Sec) • Average DFSIO Read (MB/Sec)
Delete	Deletes the Big Data Metrics Report.
More Reports	Displays the metrics as hourly, daily, weekly, or monthly values.

Analyzing Performance Bottlenecks Through Historical Metrics

You can compare a metrics report generated while the cluster was performing well with a report generated during poor performance. It helps you identify a cause or causes of a performance bottleneck in the Hadoop cluster.

To compare and analyze two metrics reports, 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 Big Data Account and click **View Details**.
 - Step 4** Click the **Hadoop Clusters** tab.
 - Step 5** Choose the Hadoop cluster, and click **View Reports**.
 - Step 6** Click the **Performance** tab.
 - Step 7** Click **Run Test**.
The Performance tab displays a default Big Data Metrics Report. This report shows the statistics collected for each host before the Hadoop cluster creation and the reports collected after Hadoop cluster creation.
 - Step 8** Choose two reports that you want to compare, and click **Compare**.
You can compare a report generated while the cluster was performing well and a report generated during poor performance.
 - Step 9** Click **Submit**.
-

Setting Alerts for Hadoop Cluster Service Failures

You can create an alert to monitor the health of the Hadoop cluster whenever Hadoop services go down. Based on the trigger conditions, you can also activate customized workflows that automatically take corrective action.

-
- Step 1** On the menu bar, choose **Policies > Orchestration**.
 - Step 2** Click the **Triggers** tab.
 - Step 3** Click **Add**.
On the **Trigger Information** page of the **Add Trigger** wizard, complete the following fields:

Name	Description
Trigger Name field	Name of the trigger.
Is Enabled check box	Check this box to enable the trigger.
Description	Description of the trigger.
Frequency	Choose the trigger rule validation frequency.

Name	Description
Trigger Type	Choose the type of trigger. <ul style="list-style-type: none"> • Stateful • Stateless

Step 4 Click **Next**.

Step 5 On the **Specify Conditions** page of the **Add Trigger** wizard, click **Add a new entry to the table below (+)**.

Step 6 In the **Add Entry to Conditions** dialog box, complete the following fields:

- a) From the **Type of Object to Monitor** drop-down list, choose **BigData Cluster**.
- b) From the **Object** drop-down list, choose the Hadoop cluster to be monitored.
- c) From the **Parameter** drop-down list, choose the parameter to use in validation.
- d) From the **Operation** drop-down list, choose **Equals** or **Not Equals**.
- e) From the **Value** drop-down list, choose **All Services Up** or **Any Service Down**.
- f) Click **Submit**.
- g) From the **Trigger When** drop-down list, make a choice to satisfy all the conditions, or any individual condition.

Step 7 Click **Next**.

Step 8 On the **Specify Workflow** page of the **Add Trigger** wizard, do the following when the Hadoop cluster service is down and when the trigger is reset:

- a) Choose the maximum number of invocations from the **Maximum Number of Invocations** drop-down list.
- b) Select a workflow for execution when the trigger state becomes active, and check the **Pass Monitored Object** check box, if necessary.
- c) Select the workflow input.
- d) Select a workflow for execution when the trigger state becomes clear, and check the **Pass Monitored Object** check box, if necessary.
- e) Select the workflow input.

Step 9 Click **Next**.

Step 10 On the **Specify Workflow Inputs** page of the **Add Trigger** wizard, enter the inputs for the selected workflows, and then click **Submit**.

Types of Disk and Network Failure Alerts

You can create alerts to detect faults related to disks and networks in a cluster.

The alerts that you can create for memory faults are as follows:

- **fltMemoryUnitInoperable:** Triggers when the number of correctable or uncorrectable errors have reached a threshold on a DIMM. The DIMM becomes inoperable.
- **fltMemoryUnitThermalThresholdNonRecoverable:** Triggers when the memory unit temperature on a server is out of the operating range. The issue is not recoverable.

- **fltMemoryArrayVoltageThresholdCritical:** Triggers when the memory array voltage exceeds the specified hardware voltage rating.
- **fltMemoryArrayVoltageThresholdNonRecoverable:** Triggers when the memory array voltage exceeds the specified hardware voltage rating, with potential memory hardware damage.
- **fltMemoryBufferUnitThermalThresholdCritical:** Triggers when the temperature of a memory buffer unit on a blade or rack server exceeds a critical threshold value.
- **fltMemoryBufferUnitThermalThresholdNonRecoverable:** Triggers when the temperature of a memory buffer unit on a blade or rack server is out of the operating range. The issue is not recoverable.
- **fltMemoryUnitDisabled:** Triggers when the server BIOS disables a DIMM. The BIOS could disable a DIMM for several reasons, including incorrect location of the DIMM or incompatible speed.

The alerts that you can create for disk faults are as follows:

- **fltStorageItemCapacityExceeded:** Triggers when the partition disk usage exceeds 70% but is less than 90%.
- **fltStorageItemCapacityWarning:** Triggers when the partition disk usage exceeds 90%.
- **fltStorageLocalDiskInoperable:** Triggers when the local disk has become inoperable.
- **fltStorageLocalDiskSlotEpUnusable:** Triggers when the server disk drive is in a slot that the storage controller does not support.
- **fltStorageLocalDiskMissing:** Triggers when a disk is missing.
- **fltStorageLocalDiskDegraded:** Triggers when the local disk has degraded. The fault description contains the physical drive state, which indicates the reason for the degradation.

The alerts that you can create for network faults are as follows:

- **fltAdaptorUnitMissing:** Triggers when the network adapter is missing, or the server cannot detect or communicate with the adapter.
- **fltAdaptorHostIfLink-down:** Triggers—
 - When the fabric interconnect is in End-Host mode and all uplink ports have failed.
 - When the server port to which the adapter is pinned have failed.
 - When a transient error causes the link to fail.
- **fltAdaptorExtIfLink-down:** Triggers—
 - When the adapter's connectivity to any of the fabric interconnects cannot be validated.
 - When a node reports that a vNIC is down, or reports a link-down event on the adapter link.

Setting Alerts for Disk and Network Failures

You can create alerts for disk or network failures in the Hadoop cluster. Alerts help you in proactive cluster maintenance. Based on the trigger conditions, you can activate customized workflows that automatically take corrective action.

Step 1 On the menu bar, choose **Policies > Orchestration**.

Step 2 Click the **Triggers** tab.

Step 3 Click **Add**.

On the **Trigger Information** page of the **Add Trigger** wizard, complete the following fields:

Name	Description
Trigger Name field	Name of the trigger.
Is Enabled check box	Check this box to enable the trigger.
Description	Description of the trigger.
Frequency	Choose the trigger rule validation frequency.
Trigger Type	Choose the type of trigger. <ul style="list-style-type: none"> • Stateful • Stateless

Step 4 Click **Next**.

Step 5 On the **Specify Conditions** page of the **Add Trigger** wizard, click **Add a new entry to the table below (+)**, and complete the following fields in the **Add Entry to Conditions** dialog box:

- a) From the **Type of Object to Monitor** drop-down list, choose **BigData Nodes**.
- b) From the **Object** drop-down list, choose the disk to be monitored.
- c) From the **Parameter** drop-down list, choose the parameter to use in validation.
- d) From the **Operation** drop-down list, choose the type of operation.
- e) From the **Value** drop-down list, choose the value to use in validation.
- f) Click **Submit**.
- g) From the **Trigger When** drop-down list, make a choice to satisfy all conditions, or any individual condition.

Step 6 Click **Next**.

Step 7 On the **Specify Workflow** page of the **Add Trigger** wizard, do the following when there is a network or disk failure and when the trigger is reset:

- a) From the **Maximum Number of Invocations** drop-down list, choose the maximum number of invocations.
- b) Select a workflow for execution when the trigger state becomes active and check the **Pass Monitored Object** check box, if necessary.
- c) Select the workflow input.

- d) Select a workflow for execution when the trigger state becomes clear, and check the **Pass Monitored Object** check box, if necessary.
- e) Select the workflow input.

Step 8 Click **Next**.

Step 9 On the **Specify Workflow Inputs** page of the **Add Trigger** wizard, enter the inputs for the selected workflows, and then click **Submit**.

Setting Disk Utilization Threshold Alerts

You can set an alert to be delivered when the disk capacity reaches a threshold. This helps you to proactively plan for capacity expansions.

Step 1 On the menu bar, choose **Policies > Orchestration**.

Step 2 Click the **Triggers** tab.

Step 3 Click **Add**.

On the **Trigger Information** page of the **Add Trigger** wizard, complete the following fields:

Name	Description
Trigger Name field	Name of the trigger.
Is Enabled check box	Check this box to enable the trigger.
Description	Description of the trigger.
Frequency	Choose the trigger rule validation frequency.
Trigger Type	Choose the type of trigger. <ul style="list-style-type: none"> • Stateful • Stateless

Step 4 Click **Next**.

Step 5 On the **Specify Conditions** page of the **Add Trigger** wizard, click **+**, and complete the following fields in the **Add Entry to Conditions** dialog box:

- a) From the **Type of Object to Monitor** drop-down list, choose **BigData Cluster**.
- b) From the **Object** drop-down list, choose the disk to be monitored.
- c) From the **Parameter** drop-down list, choose the **Disk Utilization (%)**.
- d) From the **Operation** drop-down list, choose the type of operation.
- e) From the **Value** drop-down list, choose the threshold value to use for validation.

- f) Click **Submit**.
- g) From the **Trigger When** drop-down list, make a choice to satisfy all conditions, or any individual condition.

Step 6 Click **Next**.

Step 7 On the **Specify Workflow** page of the **Add Trigger** wizard, do the following when the disk utilization reaches the threshold value and when the trigger is reset:

- a) From the **Maximum Number of Invocations** drop-down list, choose the maximum number for invocations.
- b) Select a workflow for execution when the trigger state becomes active, and check the **Pass Monitored Object** check box, if necessary.
- c) Select the workflow input.
- d) Select a workflow for execution when the trigger state becomes clear, and check the **Pass Monitored Object** check box, if necessary.
- e) Select the workflow input.

Step 8 Click **Next**.

Step 9 On the **Specify Workflow Inputs** page of the **Add Trigger** wizard, enter the inputs for the selected workflows, and then click **Submit**.
