



CenturyLinkTM

Big Data Operations Guide for Cloudera Manager v5.x Hadoop

Logging into the Enterprise Cloudera Manager

1. On the server where you have installed 'Cloudera Manager', make sure that the server is running, if not start it. You can stop the Cloudera Manager server (for example, to perform maintenance on its host) without affecting the other services running on your cluster. Statistics data used by Activity Monitoring and Service Monitoring will continue to be collected during the time the server is down.

Controlling the Cloudera Manager service

[?](#)

```
To stop the Cloudera Manager server without affecting other services: servicecloudera-scm-server stop  
To restart it: servicecloudera-scm-server start
```

2. In a web browser, type the following URL:

`http(s)://<Server host>:<port>`

where:

<Server host> is the name or IP address of the host machine where the Cloudera Manager Server is installed.

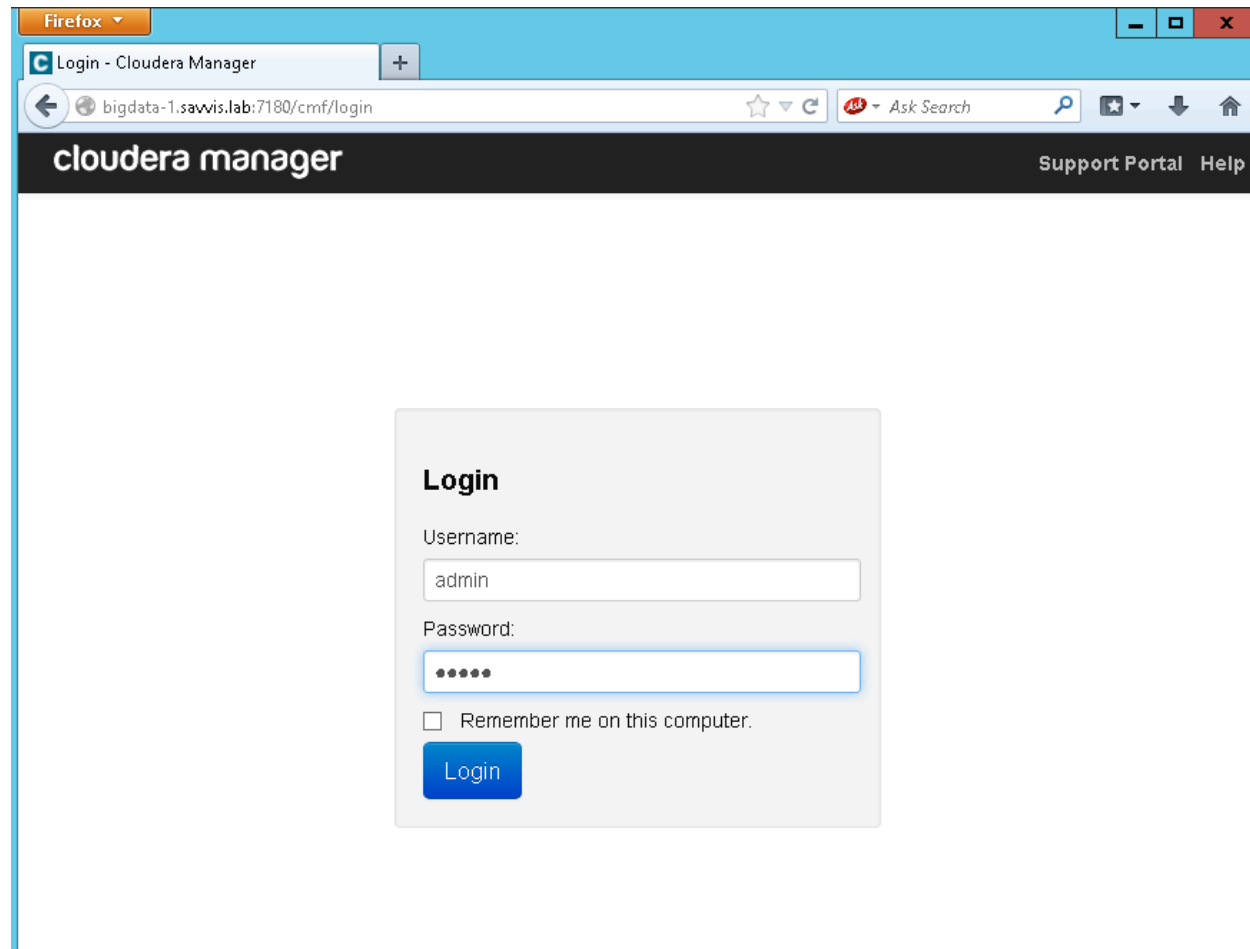
<port> is the port configured for the Cloudera Manager Server. The default port is 7180.

3. Log into the Cloudera Manager Admin Console. For **Enterprise Basic** and **Enterprise Basic With HBase** installations (managed), the review the following [link](#) to obtain the admin password.

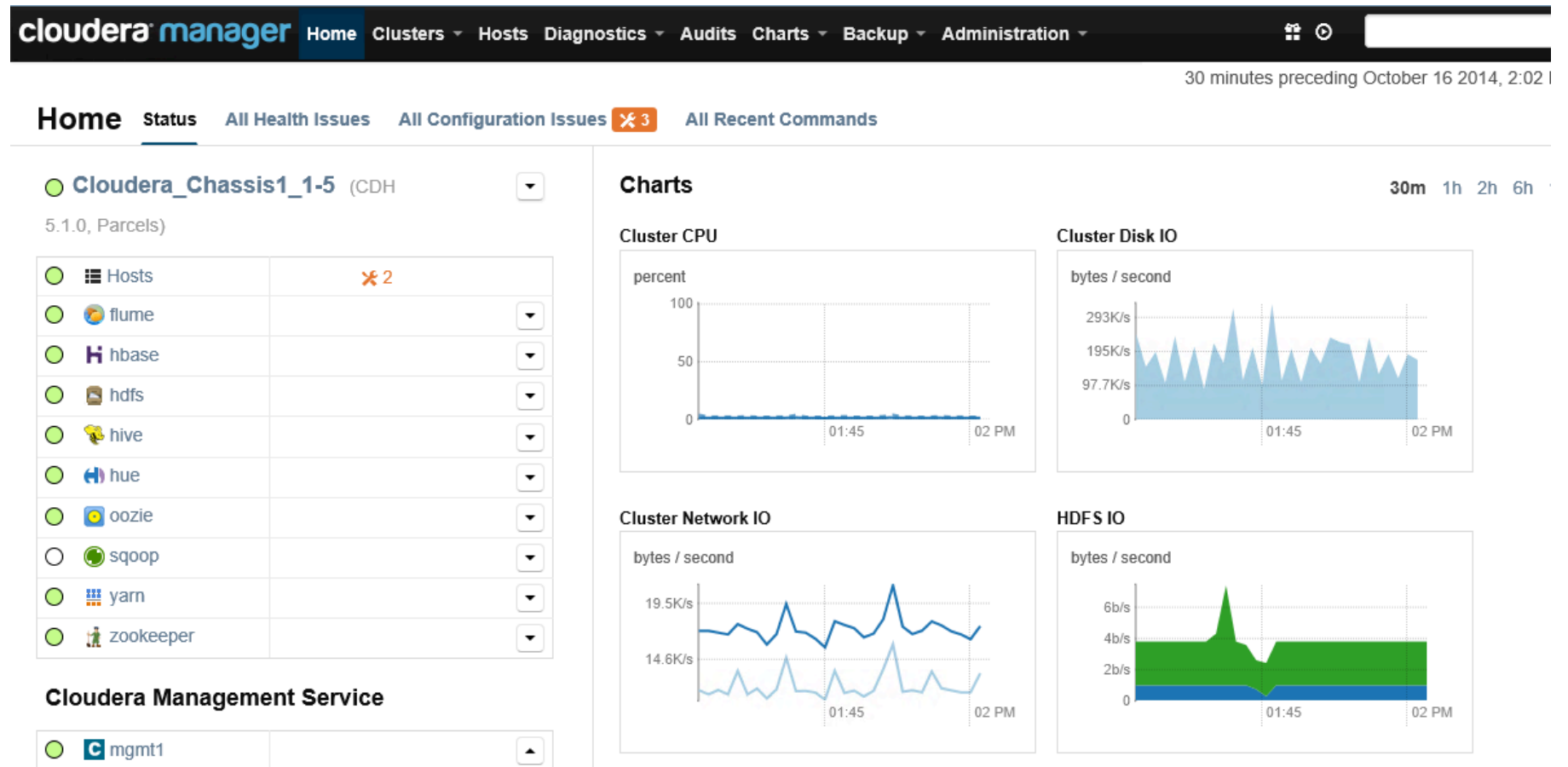
4. For **Cloudera Manager Express** (unmanaged), the admin user credentials are:

Username: admin

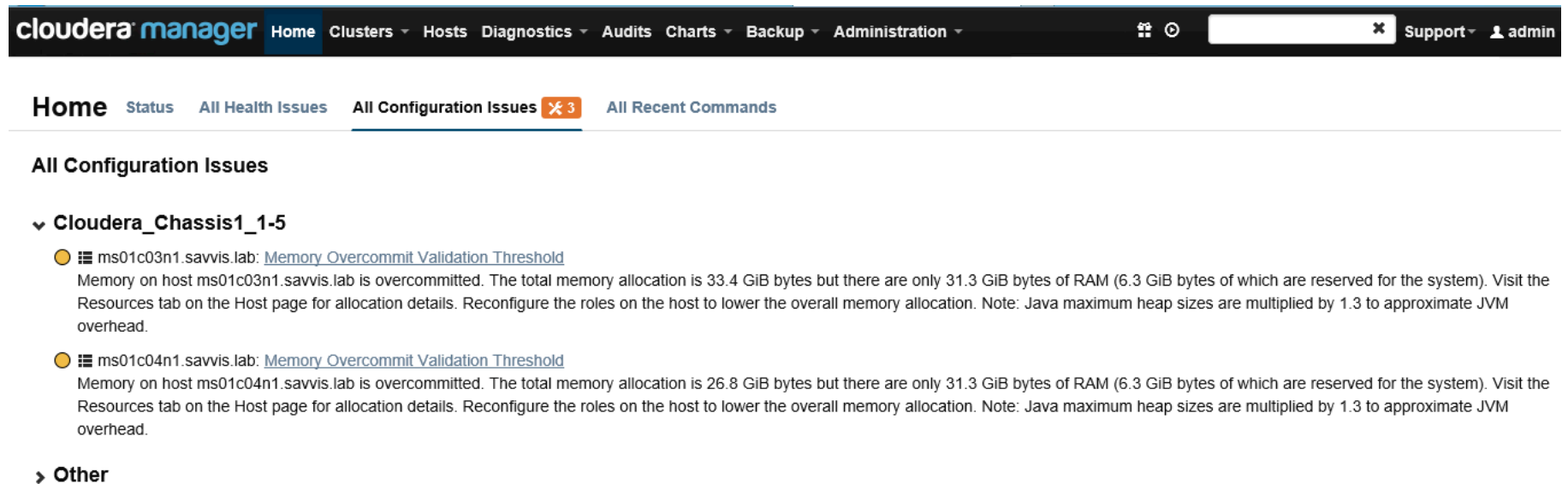
Password: admin



The Cloudera Manager 'dashboard' or 'home' shows Status and overall Cluster Health; note the subtabs for "Status", "All Health Issues", "All Configuration Issues" and "All Recent Commands"



Health and Configuration Issues can be determined by clicking "All Configuration Issues" and "All Recent Commands".



The screenshot shows the Cloudera Manager web interface. The top navigation bar includes links for Home, Clusters, Hosts, Diagnostics, Audits, Charts, Backup, and Administration. The main navigation bar highlights 'All Configuration Issues' with a red badge indicating 3 issues. Below this, the 'All Configuration Issues' section is expanded, showing a list of issues for 'Cloudera_Chassis1_1-5'. Two issues are listed, both related to 'Memory Overcommit Validation Threshold' on hosts 'ms01c03n1.savvis.lab' and 'ms01c04n1.savvis.lab'. Each issue description states that memory is overcommitted and provides details on total allocation versus available RAM, along with a note about Java heap sizes. A link to 'Memory Overcommit Validation Threshold' is provided for each issue. A link to 'Other' is also visible.

cloudera manager Home Clusters Hosts Diagnostics Audits Charts Backup Administration Support admin

Home Status All Health Issues All Configuration Issues 3 All Recent Commands

All Configuration Issues

▼ Cloudera_Chassis1_1-5

- ms01c03n1.savvis.lab: [Memory Overcommit Validation Threshold](#)
Memory on host ms01c03n1.savvis.lab is overcommitted. The total memory allocation is 33.4 GiB bytes but there are only 31.3 GiB bytes of RAM (6.3 GiB bytes of which are reserved for the system). Visit the Resources tab on the Host page for allocation details. Reconfigure the roles on the host to lower the overall memory allocation. Note: Java maximum heap sizes are multiplied by 1.3 to approximate JVM overhead.
- ms01c04n1.savvis.lab: [Memory Overcommit Validation Threshold](#)
Memory on host ms01c04n1.savvis.lab is overcommitted. The total memory allocation is 26.8 GiB bytes but there are only 31.3 GiB bytes of RAM (6.3 GiB bytes of which are reserved for the system). Visit the Resources tab on the Host page for allocation details. Reconfigure the roles on the host to lower the overall memory allocation. Note: Java maximum heap sizes are multiplied by 1.3 to approximate JVM overhead.

► Other

Check Hadoop cluster health from the CLI

On the server that is the 'namenode', execute the following command:

```
hadoop dfsadmin -report
```

?

```
[root@bigdata-2 ~]# su - hdfs -c "hadoop dfsadmin -report"
DEPRECATED: Use of this script to execute hdfs command is deprecated.
Instead use the hdfs command for it.
Configured Capacity: 26580896243712 (24.18 TB)
Present Capacity: 26580896243712 (24.18 TB)
DFS Remaining: 26580895305594 (24.18 TB)
DFS Used: 938118 (916.13 KB)
```

```
DFS Used%: 0.00%
Under replicated blocks: 0
Blocks with corrupt replicas: 0
Missing blocks: 0
-----
Datanodes available: 3 (3 total, 0 dead)
Live datanodes:
Name: 10.12.49.215:50010 (bigdata-6.savvis.lab)
Hostname: bigdata-6.savvis.lab
Rack: /default
Decommission Status : Normal
Configured Capacity: 8860298747904 (8.06 TB)
DFS Used: 315392 (308 KB)
Non DFS Used: 0 (0 B)
DFS Remaining: 8860298432512 (8.06 TB)
DFS Used%: 0.00%
DFS Remaining%: 100.00%
Last contact: Wed Feb 05 19:36:15 UTC 2014

Name: 10.12.49.214:50010 (bigdata-5.savvis.lab)
Hostname: bigdata-5.savvis.lab
Rack: /default
Decommission Status : Normal
Configured Capacity: 8860298747904 (8.06 TB)
DFS Used: 307267 (300.07 KB)
Non DFS Used: 0 (0 B)
DFS Remaining: 8860298440637 (8.06 TB)
DFS Used%: 0.00%
DFS Remaining%: 100.00%
Last contact: Wed Feb 05 19:36:14 UTC 2014

Name: 10.12.49.217:50010 (bigdata-8.savvis.lab)
Hostname: bigdata-8.savvis.lab
Rack: /default
Decommission Status : Normal
Configured Capacity: 8860298747904 (8.06 TB)
```

```
DFS Used: 315459 (308.07 KB)
Non DFS Used: 0 (0 B)
DFS Remaining: 8860298432445 (8.06 TB)
DFS Used%: 0.00%
DFS Remaining%: 100.00%
Last contact: Wed Feb 05 19:36:15 UTC 2014
```

To view individual Cludera Services, Click **Clusters** and select the Service Name under **Services**.

The screenshot displays the Cloudera Manager web interface. The top navigation bar includes links for Home, Clusters, Hosts, Diagnostics, Audits, Charts, Backup, and Administration. The main content area shows the 'Home' tab for a cluster named 'Cloudera_Chassis1_1-5 (CDH 5.1.0)'. A dropdown menu is open under the 'Clusters' tab, showing a list of services: flume, hbase, hdfs, hive, hue, oozie, sqoop, yarn, and zookeeper. The 'Cloudera Management Service' is also listed. The interface includes various charts and graphs, such as a 'Network IO' chart and a 'Disk IO' chart, which show data over time. The 'Cloudera Management Service' section shows a 'mgmt1' instance.

cloudera manager Home Clusters Hosts Diagnostics Audits Charts Backup Administration

Home Status All Health Issues

Cloudera_Chassis1_1-5 (CDH 5.1.0)

30 minutes preceding October 16, 2014, 3

Services

- flume
- hbase
- hdfs
- hive
- hue
- oozie
- sqoop
- yarn
- zookeeper

General

- Hosts
- Reports

Activities

- YARN Applications

Resource Management

- Dynamic Resource Pools
- Static Service Pools

Cloudera Management Service

- mgmt1

Network IO

24.4K/s
19.5K/s
14.6K/s

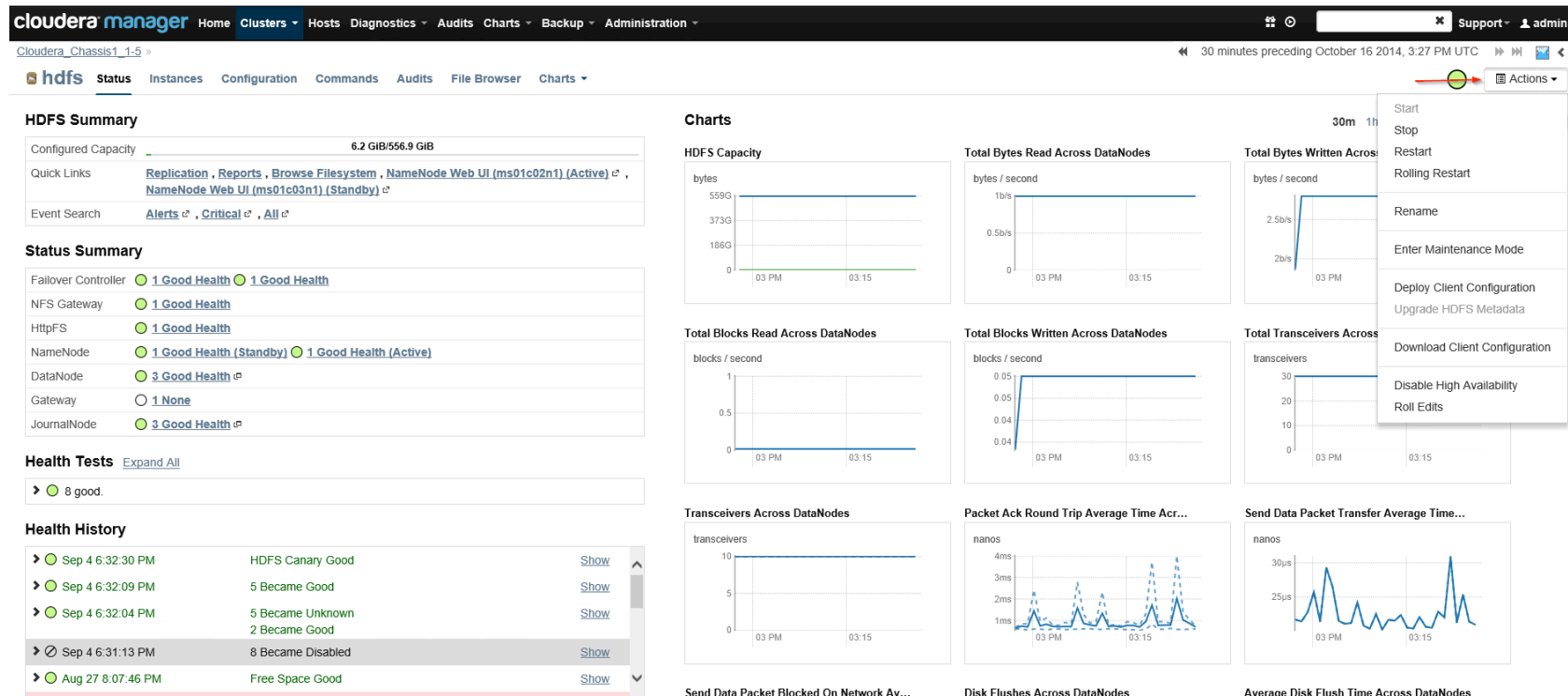
03 PM 03:15

Disk IO

6b/s
4b/s
2b/s
0

03 PM 03:15

You can drill into individual services by clicking the Service from Cluster > Services drop-down menu. The Actions drop-down provides restart options and additional functions.



Note: Your actions are now available to execute on the 'hdfs1' service.

Check hadoop service listing from the CLI

Using jps command to list running services

?

```
[root@bigdata-2 ~]# /usr/java/default/bin/jps
17209 DFSZKFailoverController
50220 Jps
```



```
14500 QuorumPeerMain
17107 JournalNode
17165 NameNode
```

Click on the 'Hosts' tab in the upper menu to the left, note all of the hostnames are listed in the cluster with other vital information, and note the sub-menu's under "Hosts" for 'Status','Configuration','Templates', and 'Parcels'

Hosts

Status Configuration Templates Disks Overview Parcels

Status

Add New Hosts to Cluster

Host Inspector

Re-run Upgrade Wizard

Filters

SEARCH

STATUS

All5

Good Health5

CDH VERSION

CLUSTERS

DECOMMISSIONED

LAST HEARTBEAT

LOAD (1 MINUTE)

LOAD (5 MINUTES)

LOAD (15 MINUTES)

MAINTENANCE MODE

RACK

SERVICES

Actions for Selected

Display 25 Entries

		Name	IP	Roles	Load Average	Disk Usage	Physical Memory	Swap Space
<input type="checkbox"/>		ms01c01n1.savvis.lab	10.12.48.101	➤ 19 Role(s)	0.06 0.12 0.12	<div>43.6 GiB / 1 TiB</div>	<div>9.3 GiB / 31.3 GiB</div>	<div>0 B / 4 GiB</div>
<input type="checkbox"/>		ms01c02n1.savvis.lab	10.12.48.102	➤ 5 Role(s)	0.56 0.33 0.16	<div>28.8 GiB / 485.8 GiB</div>	<div>8.6 GiB / 31.3 GiB</div>	<div>0 B / 4 GiB</div>
<input type="checkbox"/>		ms01c03n1.savvis.lab	10.12.48.103	➤ 8 Role(s)	0.06 0.06 0.01	<div>23.9 GiB / 485.8 GiB</div>	<div>4.2 GiB / 31.3 GiB</div>	<div>0 B / 4 GiB</div>
<input type="checkbox"/>		ms01c04n1.savvis.lab	10.12.48.104	➤ 11 Role(s)	0.00 0.05 0.04	<div>24.8 GiB / 485.8 GiB</div>	<div>6.4 GiB / 31.3 GiB</div>	<div>0 B / 4 GiB</div>
<input type="checkbox"/>		ms01c05n1.savvis.lab	10.12.48.105	➤ 3 Role(s)	0.11 0.10 0.02	<div>22.8 GiB / 485.8 GiB</div>	<div>1.9 GiB / 31.3 GiB</div>	<div>0 B / 4 GiB</div>

First

Previous

1

Next

Last

Click on 'Activities', to the left of 'Hosts' on the upper menu tab.

The screenshot displays the Cloudera Manager web interface. At the top, a navigation bar includes the Cloudera Manager logo and several menu items: Home, Clusters, Hosts, Diagnostics, Audits, Charts, Backup, and Administration. A red arrow points to the 'Clusters' menu item. Below the navigation bar, the main content area is divided into several sections. On the left, there is a 'Home' section with a 'Status' tab and a list of services for 'Cloudera_Chassis1_1-5 (CDH 5.1.0, Parcels)'. This list includes Hosts, flume, hbase, hdfs, hive, hue, oozie, sqoop, yarn, and zookeeper. Below this list is the 'Cloudera Management Service' section, which shows 'mgmt1'. In the center, a dropdown menu is open for the 'Clusters' menu item. This menu contains two columns: 'Services' and 'General'. The 'Services' column lists flume, hbase, hdfs, hive, hue, oozie, sqoop, yarn, and zookeeper. The 'General' column lists Hosts, Reports, Activities, YARN Applications, Dynamic Resource Pools, and Static Service Pools. A red arrow points to the 'Activities' option in the 'General' column. To the right of the dropdown menu, there are two line charts. The top chart is titled '30 minutes preceding October 16, 2014, 3' and shows a blue line graph with a peak around 03:15. The bottom chart is titled '30m 1h 2h' and shows a green line graph with a peak around 03:15. The charts are labeled 'Network IO' and 'Disk IO'.

cloudera manager Home Clusters Hosts Diagnostics Audits Charts Backup Administration

Home Status All Health Issues

Cloudera_Chassis1_1-5 (CDH 5.1.0, Parcels)

- Hosts
- flume
- hbase
- hdfs
- hive
- hue
- oozie
- sqoop
- yarn
- zookeeper

Cloudera Management Service

- mgmt1

Services

- flume
- hbase
- hdfs
- hive
- hue
- oozie
- sqoop
- yarn
- zookeeper

General

- Hosts
- Reports
- Activities
- YARN Applications
- Dynamic Resource Pools
- Static Service Pools

30 minutes preceding October 16, 2014, 3

Network IO

Disk IO

30m 1h 2h

03 PM 03:15

Click the 'YARN Applications' to view current and past mapreduce jobs.

cloudera manager

Home Clusters Hosts Diagnostics Audits Charts Backup Administration

Cloudera_Chassis1_1-5

30 days preceding October 16 2014, 3:39 PM UTC

Support admin

yarn

Status Instances Configuration Commands Audits Applications Charts Library

Actions

Applications

30m 1h 2h 6h 12h 1d 7d 30d

Search Select Attributes

Workload Summary

(For Completed Applications)

CPU Time

✓ 14.5s (1)

Duration

✓ 26.5s (1)

File Bytes Read

✓ 103 B (1)

File Bytes Written

✓ 1 MiB (1)

HDFS Bytes Read

✓ 2.6 KiB (1)

Results Charts

10/16/2014 3:37 PM - 10/16/2014 3:38 PM

Name: QuasiMonteCarlo Pool: root.mfelicia

Mapper: QuasiMonteCarlo\$QmcMapper Reducer: QuasiMonteCarlo\$QmcReducer

Type: MapReduce ID: job_1409855593126_0004 Duration: 26.5s

mfelicia CPU Time: 14.5s File Bytes Read: 103 B File Bytes Written: 1 MiB

HDFS Bytes Read: 2.6 KiB HDFS Bytes Written: 215 B Memory Allocated: 124.1M Pool: root.mfelicia

Actions Details

Similar MR2 Jobs

User's YARN applications

View on JobHistory Server

Check Hadoop mapreduce jobs from the CLI

Checking hadoop mapreduce jobs status

[?](#)

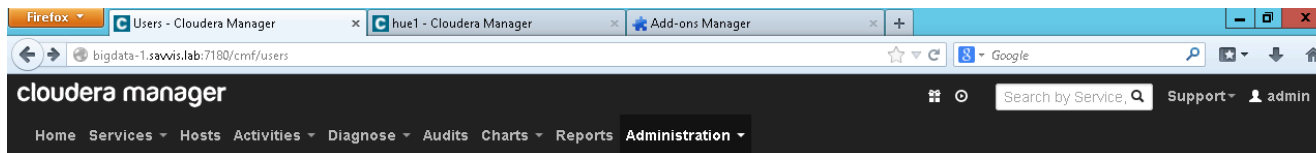
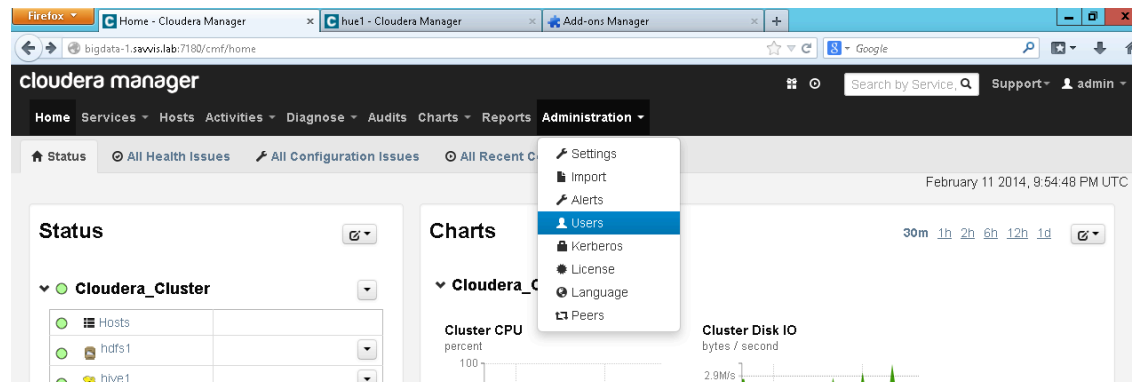
```

[root@bigdata-2 ~]# su - hdfs -c "hadoop job -list"
DEPRECATED: Use of this script to execute mapred command is deprecated.
Instead use the mapred command for it.
14/02/05 19:49:34 WARN conf.Configuration: session.id is deprecated. Instead, use dfs.metrics.session-id
14/02/05 19:49:34 INFO jvm.JvmMetrics: Initializing JVM Metrics with processName=JobTracker, sessionId=
0 jobs currently running
JobId State StartTime Username Priority SchedulingInfo

```

Cloudera Manager Users and Authentication

To add a Cloudera Manager user account: Administration tab -> Users -> Add User button -> Enter a username and password -> Select User Role -> Click Add.



Users

Checked Users:			Add Administrator Privileges	Remove Administrator Privileges	Delete	Add User
Username	Admin	User Type				
<input type="checkbox"/> admin	Yes	Cloudera Manager				Change Password

cloudera manager Home Clusters Hosts Diagnostics Audits Charts Backup Administration

Add User

Username

Password

Repeat Password

Role

Read-Only

Read-Only Users can:

- View data in Cloudera Manager

Upgrading Cloudera Manager

You can upgrade an existing Cloudera Manager to the latest version of Cloudera Manager. Upgrading preserves existing data and settings, while enabling the use of the new features provided with the latest product versions. To enable new features, some new settings are added, and some additional steps may be required, but nothing is removed.

The former Cloudera Manager Standard Editions is now known as Cloudera Express, and includes a number of features that were previously available only with Cloudera Manager Enterprise Edition. Specifically, service and activity, monitoring features are now available, and require databases to be set up for their use. Thus, upon upgrading Cloudera Manager, you will be asked for database information for these services. (You will have the option to use the embedded PostgreSQL database for this).

Understanding Upgrades

The process for upgrading to Cloudera Manager varies based on the starting point. The categories of tasks to be completed include the following:

- Install any databases that are newly required for this release. (If you are upgrading a Free Edition installation, you are asked to configure databases for the monitoring features that are now part of Cloudera Standard).
- Upgrade the Cloudera Manager server.
- Upgrade the hosts in the cluster.

Before Upgrading

- The Cloudera Manager Server must have SSH access to the cluster hosts and you must log in using a root account or an account that has password-less sudo permission. See [Requirements for Cloudera Manager](#) for more information.
- Ensure there are no running commands. Use the Admin Console's main navigation bar to check for any running commands. You can either wait for commands to complete or abort any running commands. For more information on viewing and aborting running commands, see [Viewing Running and Recent Commands](#).
- Ensure you have completed any required process for preparing databases, as described in [Database Considerations for Cloudera Manager Upgrades](#).

During the Upgrade

During the upgrade process, the following changes occur:

- The database schemas are modified for any databases storing information for Cloudera Manager Server, Activity Monitor, Service Monitor, Report Manager, and Host Monitor.
- Configuration information is reorganized.

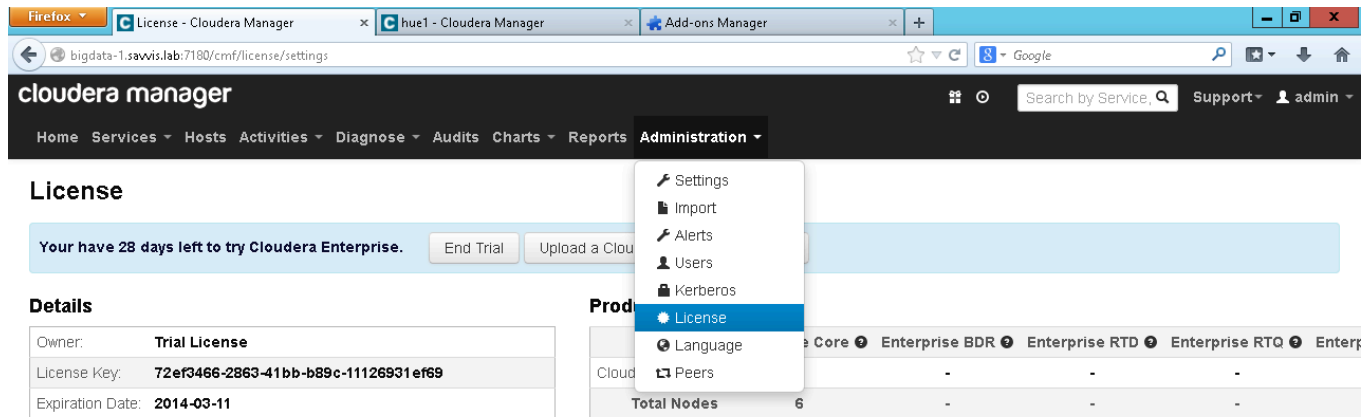
After Upgrading

After completing an upgrade to the latest Cloudera Manager (4.6 or later), the following is true:

- You have re-deployed client configurations to ensure client services have the most current configuration.
- Required databases are established to store information for Cloudera Manager Server, Hive Metastore, Activity Monitor, Service Monitor, Report Manager, and Host Monitor.
- The database schemas reflect the current version.
- The Host Monitor service is added and active.
- The Cloudera Manager Server and all supporting services, such as the Activity Monitor, Service Monitor, Report Manager, and Host Monitor are updated.

Managing Licenses

To access the **License** page, pull down the Administration menu and click License.



The screenshot shows the Cloudera Manager web interface in a Firefox browser. The browser tabs include 'License - Cloudera Manager', 'hue1 - Cloudera Manager', and 'Add-ons Manager'. The address bar shows the URL 'bigdata-1.sawvis.lab:7180/cm/lcense/settings'. The Cloudera Manager header includes a search bar and a user profile 'admin'. The 'Administration' menu is open, showing options: Settings, Import, Alerts, Users, Kerberos, License (highlighted), Language, and Peers. The main content area shows the 'License' page with a trial notice and a details table.

License

Your have 28 days left to try Cloudera Enterprise. [End Trial](#) [Upload a Cloud](#)

Details

Owner:	Trial License
License Key:	72ef3466-2863-41bb-b89c-11126931ef69
Expiration Date:	2014-03-11

Prod

	Core	Enterprise BDR	Enterprise RTD	Enterprise RTQ	Enterprise
Cloud	-	-	-	-	-
Total Nodes	6	-	-	-	-

Configuring Alert Delivery

Configuring Alert Email Delivery

When you install the Cloudera Manager Management Services, it asks you for information about the mail server you will use with the Alert Publisher.

However, if you need to change these settings, you can do so under the Alert Publisher section of the Management Services configuration tab.

Note that if you just want to add to or modify the list of alert recipient email addresses, you can do from the **Alerts** page, accessed under the Administration tab.

You can also send a test alert e-mail from the **Alerts** page under the **Administration** tab.

You can enable and disable email alerts delivery entirely (without changing the other email settings) with the **Enable email alerts** property.

To enable, disable, or configure email alerts:

1. From the **Clusters > Services** tab, select the **Cloudera Management Services** service instance.
2. Select Configuration > View and Edit.
3. Select the **Alert Publisher** and click the **Configuration tab** to see the list of properties. In order to receive email alerts you must set (or verify) the following settings:
 - Email protocol to use.
 - Your mail server hostname and port.
 - The username and password of the email user that will be logged into the mail server as the "sender" of the alert emails.
 - A comma-separated list of email addresses that will be the recipients of alert emails.
 - The format of the email alert message. Select **json** if you need the message to be parsed by a script or program.
4. Click the **Save Changes** button at the top of the page to save your settings.
5. You will need to restart the Alert Publisher role to have these changes take effect.

Firefox

All Services - Cloudera Manager

bigdata-1.savvis.lab:7180/cmfs/services/status

Google

cloudera manager

Search by Service,

Support

admin

Home

Services

Hosts

Activities

Diagnose

Audits

Charts

Reports

Administration

February 19 2014, 4:03:24 PM UTC

Tue 04

Wed 05

Thu 06

Fri 07

Sat 08

Feb 09

Mon 10

Tue 11

Wed 12

Thu 13

Fri 14

Sat 15

Feb 16

Mon 17

Tue 18

Now

Settings

Import

Alerts

Users

Kerberos

License

Language

Peers

All Services

Cloudera_Cluster

Name	Status	Role Counts	
hdfs1	Good Health	2 Failover Controllers, 1 HttpFS, 2 NameNodes, 3	JournalNodes
hive1	Good Health	1 Hive Metastore Server, 1 Gateway	
hue1	Good Health	1 Beeswax Server, 1 Hue Server	
mahout1	Good Health	1 JobTracker, 3 TaskTrackers, 1 Gateway	

Firefox

Alerts - Cloudera Manager

bigdata-1.sawvis.lab:7180/cm/alerts/config

Google

cloudera manager

Home

Services

Hosts

Activities

Diagnose

Audits

Charts

Reports

Administration

Search by Service,

Support

admin

Recipient(s): root@localhost [Send Test Alert](#)

Alert On Transitions Out of Alerting Health: No

Health Alert Threshold: Bad

Firefox automatically sends some data to Mozilla so that we can improve your experience.

[Choose What I Share](#)

Configuring Management Services Database Limits

Each Cloudera Management Service maintains a database for retaining the data it monitors. These databases (as well as the log files maintained by these services) can grow quite large.

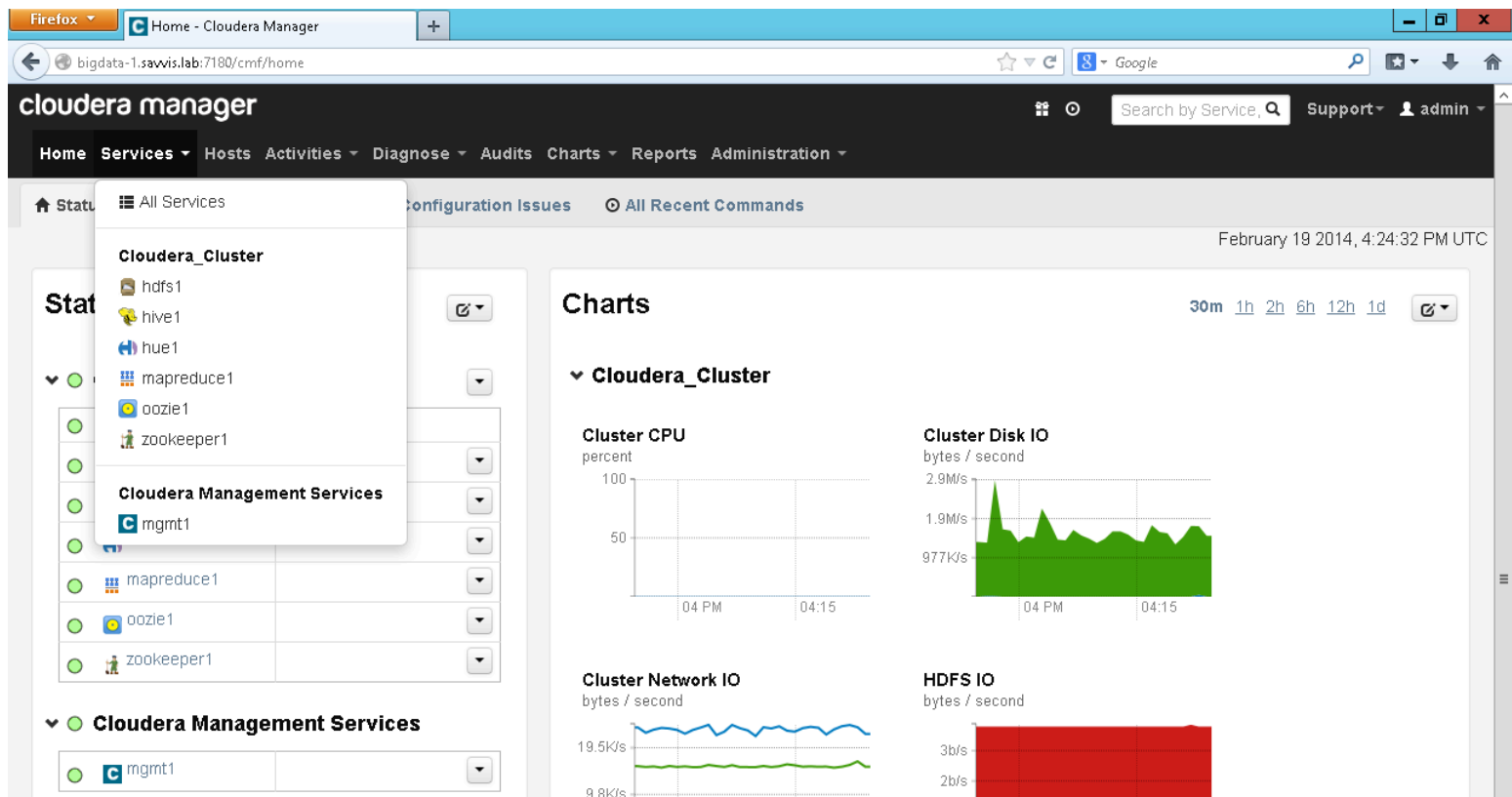
For example, the Activity Monitor maintains data at the service level, the activity level (MapReduce jobs and aggregate activities), and at the task attempt level.

Limits on these data sets are configured when you install your management services, but you can modify these parameters through the Configuration settings in the Cloudera Manager Admin console, for each management service.

For example, the Event Server lets you set a total number of events you want to store. Host Monitor and Service Monitor let you set data expiration thresholds (in hours), and Activity Monitor gives you "purge" settings (also in hours) for the data it stores. There are also settings for the logs that these various services create. You can throttle how big the logs are allowed to get and how many previous logs to retain.

To change any of the data retention or log size settings:

1. From the Services tab, select the Cloudera Management Services service instance.
2. Select **Configuration > View** and **Edit**.
3. In the left-hand column, select the **role group** for the role whose configurations you want to modify. (Note that the management services are singleton roles so there will be only a Base role group for the role.)
4. For some services, such as the **Activity Monitor**, **Service Monitor**, or **Host Monitor**, the purge or expiration period properties are found in the top-level settings for the role. Typically, **Log file size** settings will be under the **Logs** category under the **role group**.



Other Cloudera Manager Settings

From the Administration tab you can select options for configuring settings that affect how Cloudera Manager interacts with your cluster.

The Administration Settings Page

The **Settings** page provides a number of categories as follows:

- **Performance** — Set the Cloudera Manager Agent heartbeat here.
- **Advanced** — Enable API debugging and other advanced options.
- **Thresholds** — Set Agent Health status parameters. For configuration instructions, see [Configuring Agent Heartbeat and Health Status Options](#).
- **Security** — Set TLS encryption settings to enable TLS encryption between the Cloudera Manager Server, Agents, and clients. For configuration instructions, see [Configuring TLS Security for Cloudera Manager](#).
You can also:
 - Set the realm for Kerberos security and point to a custom keytab retrieval script. For configuration instructions, see [Configuring Hadoop Security with Cloudera Manager](#).
 - Specify session timeout and a "Remember Me" option.
- **Ports and Addresses** — Set ports for the Cloudera Manager Admin Console and Server. For configuration instructions, see [Configuring the Ports for the Admin Console and Agents](#).
- **Other** — To enable Cloudera usage data collection for configuration instructions, see [Sending Usage and Diagnostic Data to Cloudera](#). You can also:
 - Set a custom header color and banner text for the Admin console.
 - Set an "Information Assurance Policy" statement – this statement will be presented to every user before they are allowed to access the login dialog. The user must click "I Agree" in order to proceed to the login dialog.
 - Disable/enable the auto-search for the Events panel at the bottom of a page.
- **Support** — Enable access to online Help files from the Cloudera web site rather than from locally installed files. (See [Opening the Help Files from the Cloudera Web Site](#)), and enable automatic sending of diagnostic data to Cloudera when you trigger a data collection (see [Sending Diagnostic Data to Cloudera](#))
- **External Authentication** — Specify the configuration to use LDAP, Active Directory, or an external program for authentication. See [Configuring External Authentication](#) for instructions.
- **Parcels**— Configure settings for parcels, including the location of remote repositories that should be made available for download, and other settings such as the frequency with which Cloudera Manager will check for new parcels, limits on the number of downloads or concurrent distribution uploads. See [Using Parcels](#) for more information.

Firefox Home - Cloudera Manager

bigdata-1.savvis.lab:7180/cmf/home

cloudera manager

Search by Service, Q Support admin

Home Services Hosts Activities Diagnose Audits Charts Reports Administration

Status All Health Issues 4 All Configuration Issues All Recent Settings

February 19 2014, 4:24:32 PM UTC

Status

▼ Cloudera_Cluster

Hosts	
hdfs1	
hive1	
hue1	
mapreduce1	
oozie1	
zookeeper1	

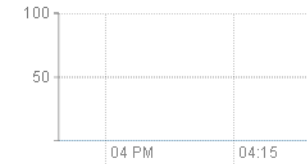
▼ Cloudera Management Services

mgmt1	
-------	--

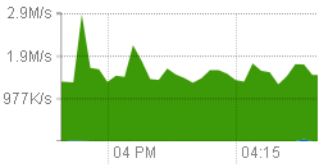
Charts

▼ Cloudera_Cluster

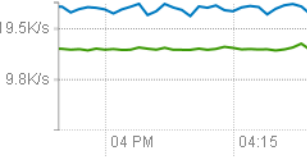
Cluster CPU
percent



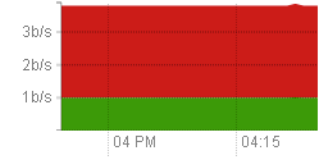
Cluster Disk IO
bytes / second




Cluster Network IO
bytes / second



HDFS IO
bytes / second



Running MapReduce Jobs
jobs



Firefox

Settings - Cloudera Manager

+

bigdata-1.savis.lab:7180/cmf/settings

☆

Google

Search by Service


Support

admin

cloudera manager

Home Services Hosts Activities Diagnose Audits Charts Reports Administration

Settings

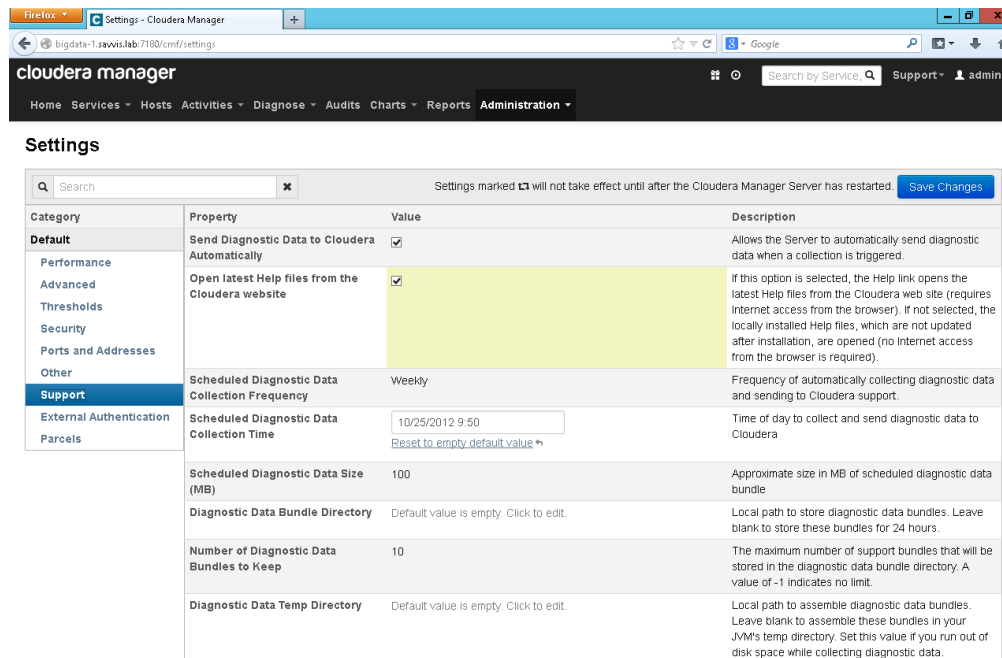
<div><div>Search</div><div>Settings marked  will not take effect until after the Cloudera Manager Server has restarted.</div><div>Save Changes</div></div>			
Category	Property	Value	Description
Default	Send Agent Heartbeat every	15 second(s)	The interval between each heartbeat that is sent from Agents to the Server.
Performance			
Advanced			
Thresholds			
Security			
Ports and Addresses			
Other			
Support			
External Authentication			
Parcels			

Opening the Help Files from the Cloudera Web Site


By default, when you click the Help link under the Support menu in the Cloudera Manager Admin console, Help files from the Cloudera web site are opened. This is because local Help files are not updated after installation. You can configure Cloudera Manager to open either the latest Help files from the Cloudera web site (this option requires Internet access from the browser) or locally-installed Help files.

To configure **Cloudera Manager** to open the Help files from the Cloudera web site (or local Help files):

1. From the **Administration** tab, select **Settings**.
2. Under the **Support** category, enable the **Open latest Help files from the Cloudera website**. This setting will be enabled by default and you can uncheck this option to open the locally-installed Help documents.
3. Click **Save Changes**.



The screenshot shows the Cloudera Manager web interface in a Firefox browser. The address bar shows the URL `bigdata-1.sawis.lab:7180/cm/settings`. The page title is "cloudera manager". The navigation bar includes links for Home, Services, Hosts, Activities, Diagnose, Audits, Charts, Reports, and Administration. The Administration tab is selected, and the Settings page is displayed.

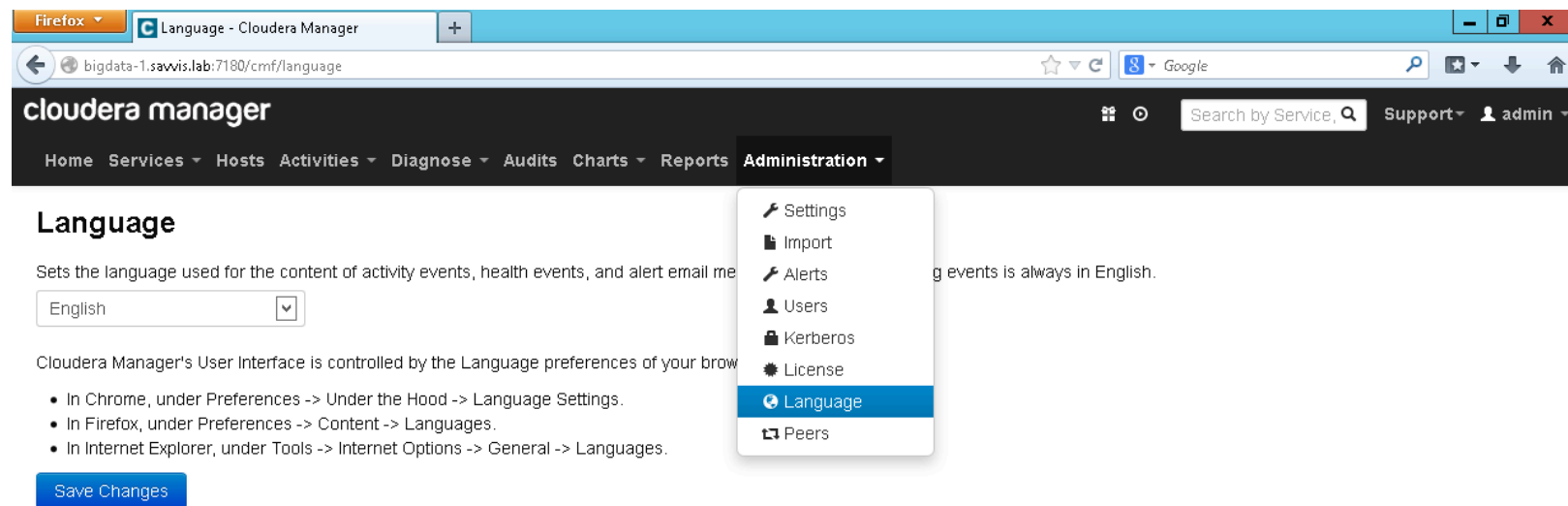
The Settings page has a search bar and a message: "Settings marked  will not take effect until after the Cloudera Manager Server has restarted." A "Save Changes" button is in the top right corner.

Category	Property	Value	Description
Default	Send Diagnostic Data to Cloudera Automatically	<input checked="" type="checkbox"/>	Allows the Server to automatically send diagnostic data when a collection is triggered.
Performance	Open latest Help files from the Cloudera website	<input checked="" type="checkbox"/>	If this option is selected, the Help link opens the latest Help files from the Cloudera web site (requires Internet access from the browser). If not selected, the locally installed Help files, which are not updated after installation, are opened (no Internet access from the browser is required).
Advanced	Scheduled Diagnostic Data Collection Frequency	Weekly	Frequency of automatically collecting diagnostic data and sending to Cloudera support.
Thresholds	Scheduled Diagnostic Data Collection Time	10/25/2012 9:50 Reset to empty default value	Time of day to collect and send diagnostic data to Cloudera
Security	Scheduled Diagnostic Data Size (MB)	100	Approximate size in MB of scheduled diagnostic data bundle
Ports and Addresses	Diagnostic Data Bundle Directory	Default value is empty. Click to edit.	Local path to store diagnostic data bundles. Leave blank to store these bundles for 24 hours.
Other	Number of Diagnostic Data Bundles to Keep	10	The maximum number of support bundles that will be stored in the diagnostic data bundle directory. A value of -1 indicates no limit.
Support	Diagnostic Data Temp Directory	Default value is empty. Click to edit.	Local path to assemble diagnostic data bundles. Leave blank to assemble these bundles in your JVM's temp directory. Set this value if you run out of disk space while collecting diagnostic data.
External Authentication			
Parcels			

User Interface Language Settings

You can change the language of the Cloudera Manager Admin Console User Interface through the language preference in your browser. Information on how to do this for the browsers supported by Cloudera Manager is shown under the **Administration > Language** page. You can also change the language for the information provided with activity and health events, and for alert email messages.

To change the language of the activity and health event information and alert email messages, select the language you want from the drop-down list on this page, and then click **Save Changes**.



Sending Usage and Diagnostic Data to Cloudera

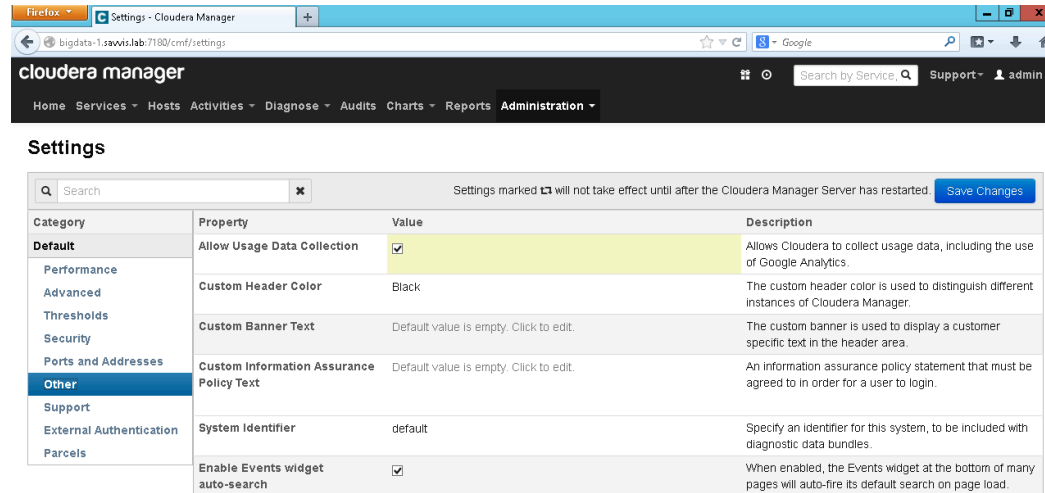
Cloudera Manager collects anonymous usage information and takes regularly scheduled snapshots of the state of your cluster and automatically sends them anonymously to Cloudera.

This helps Cloudera improve and optimize Cloudera Manager. If you are a Cloudera Enterprise user, you can also trigger the collection of diagnostic data and send it to Cloudera Support to aid in resolving a problem you may be having.

Anonymous Usage Data Collection

Cloudera Manager sends anonymous usage information using Google Analytics to Cloudera. The information helps Cloudera improve Cloudera Manager. By default anonymous usage data collection is *enabled*.

1. From the Administration tab, select Settings.
2. Under the **Other** category, set the **Allow Usage Data Collection** property.
3. Click **Save Changes**.



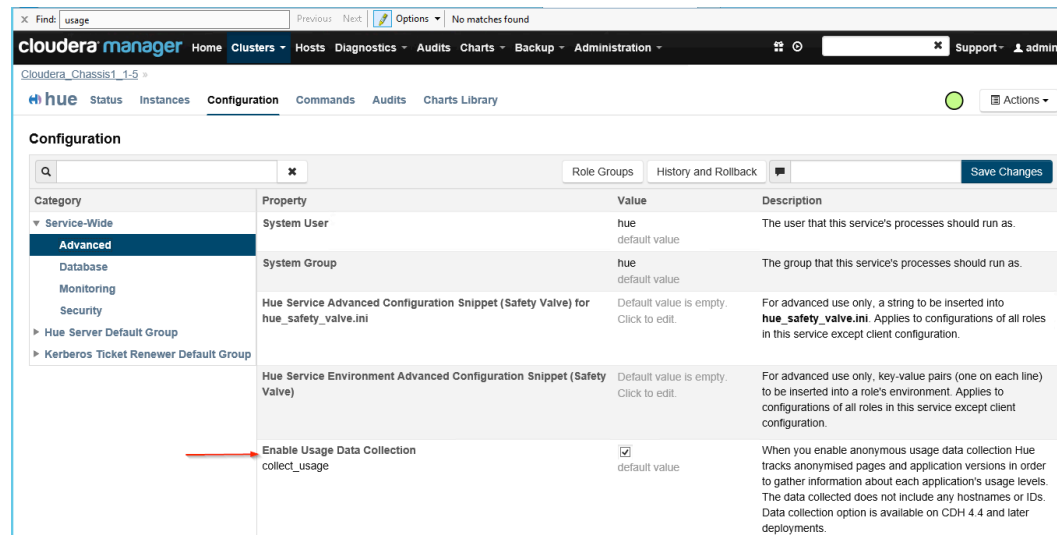
The screenshot shows the Cloudera Manager web interface. The top navigation bar includes links for Home, Services, Hosts, Activities, Diagnose, Audits, Charts, Reports, and Administration. The 'Administration' tab is selected. Below the navigation bar, the 'Settings' page is displayed. The left sidebar shows various categories: Default, Performance, Advanced, Thresholds, Security, Ports and Addresses, Other (selected), Support, External Authentication, and Parcels. The main content area shows a table of settings. The 'Allow Usage Data Collection' property under the 'Default' category is checked. The 'Enable Events widget auto-search' property under the 'Other' category is also checked. A 'Save Changes' button is visible at the top right of the settings table.

Category	Property	Value	Description
Default	Allow Usage Data Collection	<input checked="" type="checkbox"/>	Allows Cloudera to collect usage data, including the use of Google Analytics.
Performance	Custom Header Color	Black	The custom header color is used to distinguish different instances of Cloudera Manager.
Advanced	Custom Banner Text	Default value is empty. Click to edit.	The custom banner is used to display a customer specific text in the header area.
Thresholds	Custom Information Assurance Policy Text	Default value is empty. Click to edit.	An information assurance policy statement that must be agreed to in order for a user to login.
Security	System Identifier	default	Specify an identifier for this system, to be included with diagnostic data bundles.
Ports and Addresses	Enable Events widget auto-search	<input checked="" type="checkbox"/>	When enabled, the Events widget at the bottom of many pages will auto-fire its default search on page load.

Managing Hue Analytics Data Collection

Hue tracks anonymised pages and application versions in order to help compare each application's usage levels. The data collected does not include any hostnames or IDs. For example, the data is of the form: /2.3.0/pig, /2.5.0/beeswax/execute. You can restrict data collection as follows:

1. Go to the **Clusters > Hue** service.
2. Select **Configuration**.
3. Expand the **Service-Wide** category. Click **Advanced**.
4. Uncheck the **Enable Usage Data Collection** checkbox.
5. Click **Save Changes**.
6. Restart the Hue service.



The screenshot shows the Cloudera Manager interface for the Hue service configuration. The 'Configuration' tab is active, and the 'Advanced' category is expanded. The 'Enable Usage Data Collection' checkbox is checked, and a red arrow points to it. The 'collect_usage' property is visible below the checkbox.

Category	Property	Value	Description
Service-Wide	System User	hue default value	The user that this service's processes should run as.
Advanced	System Group	hue default value	The group that this service's processes should run as.
	Hue Service Advanced Configuration Snippet (Safety Valve) for hue_safety_valve.ini	Default value is empty. Click to edit.	For advanced use only, a string to be inserted into hue_safety_valve.ini . Applies to configurations of all roles in this service except client configuration.
	Hue Service Environment Advanced Configuration Snippet (Safety Valve)	Default value is empty. Click to edit.	For advanced use only, key-value pairs (one on each line) to be inserted into a role's environment. Applies to configurations of all roles in this service except client configuration.
	Enable Usage Data Collection	<input checked="" type="checkbox"/>	When you enable anonymous usage data collection Hue tracks anonymised pages and application versions in order to gather information about each application's usage levels. The data collected does not include any hostnames or IDs. Data collection option is available on CDH 4.4 and later deployments.
	collect_usage	default value	

Diagnostic Data Collection

To help with solving problems when using Cloudera Manager on your cluster, Cloudera Manager collects diagnostic data on a regular schedule, and automatically sends it to Cloudera.

By default Cloudera Manager is configured to collect data weekly and to send it *automatically*. You can schedule the frequency of data collection on a daily, weekly, or monthly schedule, or disable the scheduled collection of data entirely. You can also send a collected data set [manually](#).