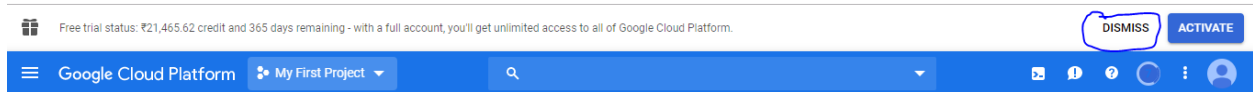


Hortonworks Data Platform : Automated Install with Ambari



<https://docs.cloudera.com/HDPDocuments/Ambari/Ambari-2.6.2.2/index.html>

Configuring iptables

For Ambari to communicate during setup with the hosts it deploys to and manages, certain ports must be open and available. The easiest way to do this is to temporarily disable iptables, as follows:

RHEL/CentOS/Oracle Linux 6

```
sudo chkconfig iptables off
sudo /etc/init.d/iptables stop
```

2.1. Download the Ambari Repository

2.1.1. RHEL/CentOS/Oracle Linux 6

On a server host that has Internet access, use a command line editor to perform the following steps:

1. Log in to your host with your username.
2. Download the Ambari repository file to a directory on your installation host.

```
sudo wget -nv http://public-repo-
1.hortonworks.com/ambari/centos6/2.x/updates/2.6.2.2/ambari.repo -O
/etc/yum.repos.d/ambari.repo
```

Important

Do not modify the ambari.repo file name. This file is expected to be available on the Ambari Server host during Agent registration.

3. Confirm that the repository is configured by checking the repo list.

yum repolist

You should see values similar to the following for Ambari repositories in the list. Version values vary, depending on the installation.

4. Install the Ambari bits. This also installs the default PostgreSQL Ambari database.

sudo yum install ambari-server

5. **Enter y** when prompted to confirm transaction and dependency checks. A successful installation displays output similar to the following:

```
Installing : postgresql-libs-8.4.20-3.el6_6.x86_64 1/4
Installing : postgresql-8.4.20-3.el6_6.x86_64 2/4
Installing : postgresql-server-8.4.20-3.el6_6.x86_64 3/4
Installing : ambari-server-2.2.0.0-1470.x86_64 4/4
Verifying : ambari-server-2.2.0.0-1470.x86_64 1/4
Verifying : postgresql-8.4.20-3.el6_6.x86_64 2/4
Verifying : postgresql-server-8.4.20-3.el6_6.x86_64 3/4
Verifying : postgresql-libs-8.4.20-3.el6_6.x86_64 4/4
Installed:
ambari-server.x86_64 0:2.2.0.0-1470
Dependency Installed:
postgresql.x86_64 0:8.4.20-3.el6_6
postgresql-libs.x86_64 0:8.4.20-3.el6_6
postgresql-server.x86_64 0:8.4.20-3.el6_6
Complete!
```

Ambari Server by default uses an embedded PostgreSQL database. When you install the Ambari Server, the PostgreSQL packages and dependencies must be available for install. These packages are typically available as part of your Operating System repositories. Please confirm you have the appropriate repositories available for the postgresql-server packages.

2.2. Set Up the Ambari Server

Before starting the Ambari Server, you **must** set up the Ambari Server. Setup configures Ambari to talk to the Ambari database, installs the JDK and allows you to customize the user account the Ambari Server daemon will run as. The ambari-server setup command manages the setup

process. Run the following command on the Ambari server host to start the setup process. You may also append [Setup Options](#) to the command.

sudo ambari-server setup

Using python /usr/bin/python2

Setup ambari-server

Checking SELinux...

SELinux status is 'enabled'

SELinux mode is 'enforcing'

Temporarily disabling SELinux

WARNING: SELinux is set to 'permissive' mode and temporarily disabled.

OK to continue [y/n] (y)? y

Customize user account for ambari-server daemon [y/n] (n)? n

Adjusting ambari-server permissions and ownership...

Checking firewall status...

Checking JDK...

[1] Oracle JDK 1.8 + Java Cryptography Extension (JCE) Policy Files 8

[2] Oracle JDK 1.7 + Java Cryptography Extension (JCE) Policy Files 7

[3] Custom JDK

=====

Enter choice (1): 1

To download the Oracle JDK and the Java Cryptography Extension (JCE) Policy Files you must accept the license terms found at <http://www.oracle.com/technetwork/java/javase/terms/license/index.html> and not accepting will cancel the Ambari Server setup and you must install the JDK and JCE files manually.

Do you accept the Oracle Binary Code License Agreement [y/n] (y)? y

Downloading JDK from <http://public-repo-1.hortonworks.com/ARTIFACTS/jdk-8u60-linux-x64.tar.gz> to /var/lib/ambari-server/resources/jdk-8u60-linux-x64.tar.gz

jdk-8u60-linux-x64.tar.gz... 100% (172.8 MB of 172.8 MB)

Successfully downloaded JDK distribution to /var/lib/ambari-server/resources/jdk-8u60-linux-x64.tar.gz

Installing JDK to /usr/jdk64/

Successfully installed JDK to /usr/jdk64/

Downloading JCE Policy archive from http://public-repo-1.hortonworks.com/ARTIFACTS/jce_policy-8.zip to /var/lib/ambari-server/resources/jce_policy-8.zip

Successfully downloaded JCE Policy archive to /var/lib/ambari-server/resources/jce_policy-8.zip

Installing JCE policy...

Completing setup...

Enter advanced database configuration [y/n] (n)? n

About to start PostgreSQL

Configuring local database...

Connecting to local database...done.

Configuring PostgreSQL...

Restarting PostgreSQL

Extracting system views...

...ambari-admin-2.2.0.0.1310.jar

Adjusting ambari-server permissions and ownership...

Ambari Server 'setup' completed successfully.

2.3. Start the Ambari Server

- Run the following command on the Ambari Server host:

sudo ambari-server start

- To check the Ambari Server processes:

ambari-server status

- To stop the Ambari Server:

ambari-server stop

Note

If you plan to use an existing database instance for Hive or for Oozie, you must complete the preparations described in [Using Non-Default Databases-Hive](#) and [Using Non-Default Databases-Oozie](#) **before** installing your Hadoop cluster.

Next Steps

[Install, configure and deploy an HDP cluster](#)

3. Installing, Configuring, and Deploying a HDP Cluster

Use the Ambari Install Wizard running in your browser to install, configure, and deploy your cluster, as follows:

- [Log In to Apache Ambari](#)
- [Name Your Cluster](#)
- [Select Stack](#)
- [Install Options](#)
- [Confirm Hosts](#)
- [Choose Services](#)
- [Assign Masters](#)
- [Assign Slaves and Clients](#)
- [Customize Services](#)
- [Review](#)
- [Install, Start and Test](#)
- [Complete](#)

3.1. Log In to Apache Ambari

After starting the Ambari service, open Ambari Web using a web browser.

1. Point your browser to <http://<your.ambari.server>:8080>, where <your.ambari.server> is the name of your ambari server host. For example, a default Ambari server host is located at <http://c6401.ambari.apache.org:8080>.
2. Log in to the Ambari Server using the default user name/password: admin/admin. You can change these credentials later.

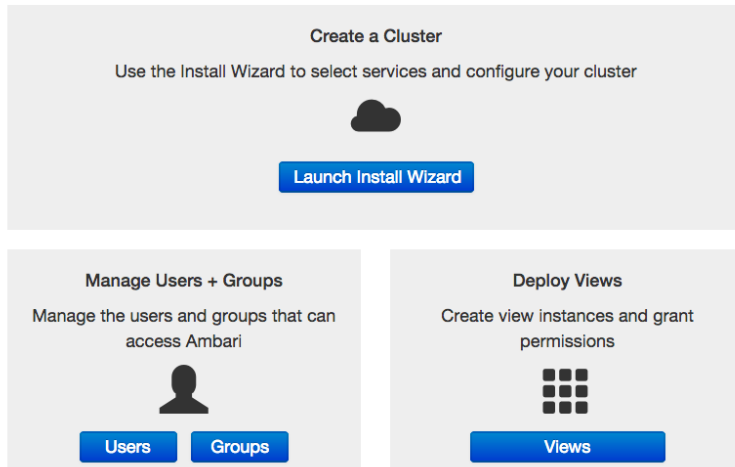
For a new cluster, the Ambari install wizard displays a Welcome page from which you [launch the Ambari Install wizard](#).

3.2. Launching the Ambari Install Wizard

From the Ambari Welcome page, choose Launch Install Wizard.

Welcome to Apache Ambari

Provision a cluster, manage who can access the cluster, and customize views for Ambari users.



3.3. Name Your Cluster

1. In Name your cluster, type a name for the cluster you want to create. Use no white spaces or special characters in the name.
2. Choose Next.

3.4. Select Stack

The Service Stack (the Stack) is a coordinated and tested set of HDP components. Use a radio button to select the Stack version you want to install. To install an HDP 2x stack, select the HDP 2.3, HDP 2.2, or HDP 2.1 radio button.

Select Stack

Please select the service stack that you want to use to install your Hadoop cluster.

Stacks

- ☒ HDP 2.3
- ☐ HDP 2.2
- ☐ HDP 2.1

▶ [Advanced Repository Options](#)

← Back

Next →

Expand Advanced Repository Options to select the Base URL of a repository from which Stack software packages download. Ambari sets the default Base URL for each repository, depending on the Internet connectivity available to the Ambari server host, as follows:

3.5. Install Options

In order to build up the cluster, the install wizard prompts you for general information about how you want to set it up. You need to supply the FQDN of each of your hosts. The wizard also needs to access the private key file you created in [Set Up Password-less SSH](#).

Using the host names and key file information, the wizard can locate, access, and interact securely with all hosts in the cluster.

1. Use the Target Hosts text box to enter your list of host names, one per line.

You can use ranges inside brackets to indicate larger sets of hosts. For example, for host01.domain through host10.domain use host[01-10].domain

```
[mohamedirfan@hmaster ~]$ hostname -f  
hmaster.c.hdpirfan.internal
```

```
[mohamedirfan@hslave1 ~]$ hostname -f  
hslave1.c.hdpirfan.internal
```

```
[mohamedirfan@hslave2 ~]$ hostname -f  
hslave2.c.hdpirfan.internal
```

2. If you want to let Ambari automatically install the Ambari Agent on all your hosts using SSH, select Provide your SSH Private Key and either use the Choose File button in the Host Registration Information section to find the private key file that matches the public key you installed earlier on all your hosts or cut and paste the key into the text box manually.

The screenshot shows the 'Host Registration Information' section of the Ambari installation wizard. It includes a text box for 'Target Hosts' with the value 'datanode2-demo.c.ultimate-bit-117015.internal'. Below this, there are two radio buttons: 'Provide your SSH Private Key to automatically register hosts' (selected) and 'Perform manual registration on hosts and do not use SSH'. The first option has a 'Choose File' button and an 'openSSH private' button. A text box for the SSH Private Key is shown with a sample key. Below this is an 'SSH User Account' field with a red error message 'User name is required'. At the bottom, there are 'Back' and 'Register and Confirm' buttons. A blue arrow points from the 'Target Hosts' text box to a blue box containing the text 'The key will be populated in this text area'. Another blue arrow points from the 'SSH Private Key' text box to the same blue box. A third blue arrow points from the 'SSH User Account' field to a blue box containing two bullet points: 'Enter your Login Id which you used when logging in into you're google cloud VM using putty.' and 'Remember your Login Id is your Gmail user name with .'s replaced with _'s'. A green arrow points from the 'Register and Confirm' button to a blue box containing the text 'Click Here'.

datanode2-demo.c.ultimate-bit-117015.internal

The key will be populated in this text area

Host Registration Information

☒ Provide your **SSH Private Key** to automatically register hosts

Choose File openSSH private

-----BEGIN RSA PRIVATE KEY-----
MIIEoQIBAAKCAQEAhNYYTlG2r1LyXQbOQRcV7p+1NjV/NtKZAtRr5ZsXrFQ
1nHoP

SSH User Account

User name is required

☐ Perform manual registration on hosts and do not use SSH

Back Register and Confirm

- Enter your Login Id which you used when logging in into you're google cloud VM using putty.
- Remember your Login Id is your Gmail user name with .'s replaced with _'s

Install Options

Enter the list of hosts to be included in the cluster and provide your SSH key.

Target Hosts

Enter a list of hosts using the Fully Qualified Domain Name (FQDN), one per line. Or use [Pattern Expressions](#)

```
hmaster.c.hdpirfan.internal  
hslave1.c.hdpirfan.internal  
hslave2.c.hdpirfan.internal
```

Host Registration Information

- ☒ Provide your [SSH Private Key](#) to automatically register hosts

opensshpvt

```
tnEpZOrJttDgP13JQmCXXEBrwNIVjyb1bKuEybUI8F2RpVuokQM=  
-----END RSA PRIVATE KEY-----
```

SSH User Account

mohamedirfan

- ☐ Perform [manual registration](#) on hosts and do not use SSH

4. Choose Register and Confirm to continue.

Confirm Hosts prompts you to confirm that Ambari has located the correct hosts for your cluster and to check those hosts to make sure they have the correct directories, packages, and processes required to continue the install.

If any hosts were selected in error, you can remove them by selecting the appropriate checkboxes and clicking the grey Remove Selected button. To remove a single host, click the small white **Remove** button in the Action column.

At the bottom of the screen, you may notice a yellow box that indicates some warnings were encountered during the check process. For example, your host may have already had a copy of `wget` or `curl`. Choose [Click here to see the warnings](#) to see a list of what was checked and what caused the warning. The warnings page also provides access to a python script that can help you clear any issues you may encounter and let you run Rerun Checks.

Confirm Hosts

Registering your hosts.

Please confirm the host list and remove any hosts that you do not want to include in the cluster.

Remove Selected		Show: All (3) Installing (0) Registering (0) Success (3) Fail (0)		
<input type="checkbox"/>	Host	Progress	Status	Action
<input type="checkbox"/>	hmaster.c.hdpirfan.internal	<div></div>	Success	<button>Remove</button>
<input type="checkbox"/>	hslave1.c.hdpirfan.internal	<div></div>	Success	<button>Remove</button>
<input type="checkbox"/>	hslave2.c.hdpirfan.internal	<div></div>	Success	<button>Remove</button>
Show: 25 1 - 3 of 3				

Some warnings were encountered while performing checks against the 3 registered hosts above [Click here to see the warnings.](#)

[← Back](#)

[Next →](#)

3.7. Choose Services

Based on the Stack chosen during Select Stack, you are presented with the choice of Services to install into the cluster. HDP Stack comprises many services. You may choose to install any other available services now, or to [add services](#) later. The install wizard selects all available services for installation by default.

1. Choose none to clear all selections, or choose all to select all listed services.
2. Choose or clear individual checkboxes to define a set of services to install now.
3. After selecting the services to install now, choose Next.

3.8. Assign Masters

The Ambari install wizard assigns the master components for selected services to appropriate hosts in your cluster and displays the assignments in Assign Masters. The left column shows services and current hosts. The right column shows current master component assignments by host, indicating the number of CPU cores and amount of RAM installed on each host.

1. To change the host assignment for a service, select a host name from the drop-down menu for that service.
2. To remove a ZooKeeper instance, click the green minus icon next to the host address you want to remove.
3. When you are satisfied with the assignments, choose Next.

Hive Remote Meta store Setting.

```
sudo yum install mysql-connector-java*
```

```
ls -al /usr/share/java/mysql-connector-java.jar
```

```
cd /var/lib/ambari-server/resources/
```

```
sudo ln -s /usr/share/java/mysql-connector-java.jar mysql-connector-java.jar
```

```
sudo ambari-server setup --jdbc-db=mysql --jdbc-driver=/usr/share/java/mysql-connector-java.jar
```

3.9. Assign Slaves and Clients

The Ambari installation wizard assigns the slave components (DataNodes, NodeManagers, and RegionServers) to appropriate hosts in your cluster. It also attempts to select hosts for installing the appropriate set of clients.

1. Use **all** or **none** to select all of the hosts in the column or none of the hosts, respectively. If a host has an asterisk next to it, that host is also running one or more master components. Hover your mouse over the asterisk to see which master components are on that host.

Hortonworks Data Platform Dec 21, 2015

38

2. Fine-tune your selections by using the checkboxes next to specific hosts.

3. When you are satisfied with your assignments, choose Next.

3.10. Customize Services

The Customize Services step presents you with a set of tabs that let you review and modify your HDP cluster setup. The wizard attempts to set reasonable defaults for each of the options. You are **strongly encouraged** to review these settings as your requirements might be slightly different.

Browse through each service tab and by hovering your cursor over each of the properties, you can see a brief description of what the property does. The number of service tabs shown depends on the services you decided to install in your cluster. **Any tab that requires input shows a red badge with the number of properties that need attention.** Select each service tab that displays a red badge number and enter the appropriate information.

Directories

The choice of directories where HDP will store information is critical. Ambari will attempt to choose reasonable defaults based on the mount points available in your environment but you are **strongly encouraged** to review the default directory settings recommended by Ambari. In particular, confirm directories such as /tmp and /var are **not** being used for HDFS NameNode directories and DataNode directories under the **HDFS** tab.

Passwords

You must provide database passwords for the Hive and Oozie services and the Master Secret for Knox. Using Hive as an example, choose the **Hive** tab and expand the Advanced section. In Database Password field marked in red, provide a password, then retype to confirm it.

Note

By default, Ambari will install a new MySQL instance for the Hive Metastore and install a Derby instance for Oozie. If you plan to use existing databases for MySQL, Oracle or PostgreSQL, modify these options before proceeding. Refer to [Using Non-Default Databases](#) for more information on using existing databases.

Important

Using the **Microsoft SQL Server** or **SQL Anywhere** database options are not supported.

Service Account Users and Groups

The service account users and groups are available under the **Misc** tab. These are the **operating system accounts the service components will run as**. If these users do not exist on your hosts, Ambari will automatically create the users and groups locally on the hosts. If these users already exist, Ambari will use those accounts.

Hortonworks Data Platform Dec 21, 2015

39

Depending on how your environment is configured, you might not allow groupmod or usermod operations. If this is the case, you **must** be sure all users and groups are already created and **be sure to** select the "Skip group modifications" option on the **Misc** tab. This tells Ambari to not modify group membership for the service users.

Refer to the Ambari Reference Guide [Customizing HDP Services](#) for more information on the service account users and groups that are needed for HDP.

After you complete Customizing Services, choose Next.

3.11. Review

The assignments you have made are displayed. Check to make sure everything is correct. If you need to make changes, use the left navigation bar to return to the appropriate screen. To print your information for later reference, choose Print.

When you are satisfied with your choices, choose Deploy.

3.12. Install, Start and Test

The progress of the install displays on the screen. Ambari installs, starts, and runs a simple test on each component. Overall status of the process displays in progress bar at the top of the screen and host-by-host status displays in the main section. Do not refresh your browser during this process. Refreshing the browser may interrupt the progress indicators.

To see specific information on what tasks have been completed per host, click the link in the Message column for the appropriate host. In the Tasks pop-up, click the individual task to see the related log files. You can select filter conditions by using the Show dropdown list. To see a larger version of the log contents, click the Open icon or to copy the contents to the clipboard, use the Copy icon.

When Successfully installed and started the services appears, choose Next.

3.13. Complete

The Summary page provides you a summary list of the accomplished tasks. Choose Complete. Ambari Web GUI displays.

Ambari Workouts:

Setting up user for hdfs access:

Onboarding on new user in Hadoop:

In the edge node:

Either sysadmin/hadoop admin will do this:

Admin create linux account:

```
sudo useradd irfan
```

```
sudo passwd irfan
```

Now Irfan become a linux user:

```
su irfan
```

```
cd /home/irfan
```

User is trying to become a hadoop user: (Not possible)

```
[irfan@instance-4 ~]$ hadoop fs -ls /user/irfan
```

ls: `/user/irfan': No such file or directory

```
[irfan@instance-4 ~]$ hadoop fs -mkdir -p /user/irfan
```

mkdir: Permission denied: user=irfan, access=WRITE, inode="/user/irfan":hdfs:hdfs:drwxr-xr-x

Admin balamurugan_shss will give access to the user for HDFS access

```
[balamurugan_shss@instance-4 ~]$
```

```
sudo su hdfs
```

```
hadoop fs -mkdir -p /user/irfan
```

```
hadoop fs -ls /user/
```

```
hadoop fs -chown irfan:hdfs /user/irfan
```

```
su irfan
cd ~
filename
hadoop fs -put filename /user/irfan/
sudo su hdfs
hadoop fs -mkdir -p /user/irfan
hadoop fs -chown irfan:hdfs /user/irfan
```

Adding irfan to the other groups:

```
sudo usermod -a -G hdfs irfan
```

```
sudo usermod -a -G hive irfan
```

```
sudo usermod -a -G hadoop irfan
```

End user will do this:

1. login into edgenode alone..
2. put get from hadoop fs in his /user/irfan directory

Login to hive using mr engine:

```
hive --hiveconf hive.execution.engine=mr
```

Sqoop:

In the Edge node:

```
sudo yum install mysql-connector-java*
```

```
drop user 'user1'@'localhost';
```

```
CREATE USER 'user1'@'localhost' IDENTIFIED BY 'user1';
```

```
grant all privileges on custdb.* to 'user1'@'localhost' identified by 'user1';
```

```
grant all privileges on custdb.* to 'user1'@'%' identified by 'user1';
```

```
hadoop fs -mkdir -p /user/balamurugan_shss
```

```
sqoop import --connect jdbc:mysql://10.128.0.3/custdb --username user1 --password user1 -table  
customer --driver com.mysql.jdbc.Driver -m 1
```

To work in Ambari Views:

Login to any node and create the ambari admin user:

```
hadoop fs -mkdir /user/admin
```

```
hadoop fs -chown admin:hadoop /user/admin
```

Oozie Configuration:

log in into server where mysql is installed by Ambari as a part of hive remote metastore config:

```
sudo su mysql
```

```
mysql -u root
```

```
mysql>
```

```
create databases oozie;
```

```
CREATE USER 'oozie'@'localhost' IDENTIFIED BY 'oozie';
```

```
grant all privileges on oozie.* to 'oozie'@'localhost' identified by 'oozie';
```

```
grant all privileges on oozie.* to 'oozie'@'%' identified by 'oozie';
```

If Oozie doesn't start kill the service in the node where oozie runs

```
ps -ef | grep oozie
```

```
kill -9 oozie_process_id
```

Start oozie server again in the Ambari

Phoenix :

`/usr/hdp/current/phoenix-client/bin/sqlline.py zookeeper-node-hostname:2181:/hbase-unsecure`

List and kill the yarn applications if running for long time or not running:

`yarn application -list`

`yarn application -kill application_1568821429313_0009`