2016

CREATING A 3 NODE HADOOP CLUSTER

Hortonworks Data Platform

NAME :- ARNAB MUKHERJEE

dell
Microsoft
24-Sep-16

Acknowledgement

I have taken efforts in this project. However, it would not have been possible without the kind support and help of many individuals and organizations.

I would like to extend my sincere thanks to all of them.I am highly indebted to Acadgild for their guidance and constant supervision as well as for providing necessary information regarding the project & also for their support in completing the project.

I would like to express my gratitude towards my mentor Mr. Manas Puri & my parents for their kind co-operation and encouragement which help me in completion of this project.I would like to express my special gratitude and thanks to my mentor again for giving me such attention and time.

Contents

Hortonworks Data Platform : Automated Install with Ambari

The Hortonworks Data Platform, powered by Apache Hadoop, is a massively scalable and 100% open source platform for storing, processing and analyzing large volumes of data. It is designed to deal with data from many sources and formats in a very quick, easy and cost-effective manner. The Hortonworks Data Platform consists of the essential set of Apache Hadoop projects including MapReduce, Hadoop Distributed File System (HDFS), HCatalog, Pig, Hive, HBase, Zookeeper and Ambari. Hortonworks is the major contributor of code and patches to many of these projects. These projects have been integrated and tested as part of the Hortonworks Data Platform release process and installation and configuration tools have also been included.

1. Getting Ready

This section describes the information and materials we should get ready to install a HDP cluster using Ambari. Ambari provides an end-to-end management and monitoring solution for our HDP cluster. Using the Ambari Web UI and REST APIs, we can deploy, operate, manage configuration changes, and monitor services for all nodes in our cluster from a central point. Here we are going to do this in our local machine with three nodes.

1) VMware workstation
2) Operating Systems Requirements - CentOS v6.8
3) Browser Requirements –
   Linux (CentOS)
   • Firefox 18
   • Google Chrome 26
4) Software Requirements –
   - yum and rpm (CentOS)
   - scp, curl, unzip, tar, and wget
   - For CentOS 6: Python 2.6.*

5) JDK Requirements –
   • Oracle JDK 1.8 64-bit (minimum JDK 1.8_60) (default)
   • Oracle JDK 1.7 64-bit (minimum JDK 1.7_67)
6) Database Requirements
   - PostgreSQL 8
   - PostgreSQL 9.1.13+,9.3
   - MySQL 5.6
   - Oracle 11gr2
   - Oracle 12c*

Note:-By default, Ambari will install an instance of PostgreSQL on the Ambari Server host. Optionally, to use an existing instance of PostgreSQL, MySQL or Oracle

7) Memory Requirements - The Ambari host should have at least 1 GB RAM, with 500 MB free.

2. Collect Information

Before deploying an HDP cluster, we should collect the following information

The fully qualified domain name (FQDN) of each host in our system. The Ambari install wizard supports using IP addresses. We can use hostname -f to check or verify the FQDN of a host.
A list of components we want to set up on each host.

3. Set Up Password-Less SSH

To have Ambari Server automatically install Ambari Agents on all our cluster hosts, we must set up password-less SSH connections between the Ambari Server host and all other hosts in the cluster. The Ambari Server host uses SSH public key authentication to remotely access and install the Ambari Agent.

1. Generate public and private SSH keys on the Ambari Server host.

```
ssh-keygen
```

2. Copy the SSH Public Key (id_rsa.pub) to the root account on our target hosts

```
ssh-copy-id –i /root/.ssh/ id_rsa.pub root@FQDN
```

3. From the Ambari Server, make sure we can connect to each host in the cluster using SSH, without having to enter a password.

ssh root@<remote.target.host> where <remote.target.host> has the value of each host name in our cluster

4. Edit ssh_config file and add below section in the end of the file

```
vi /etc/ssh/ssh_config
StrictHostKeyChecking no
```

5. Depending on our version of SSH, we may need to set permissions on the .ssh directory (to 700) and the authorized_keys file in that directory (to 600) on the target hosts.

```
chmod 700 ~/.ssh
chmod 600 ~/.ssh/authorized_keys
```

6. Retain a copy of the SSH Private Key on the machine from which we will run the webbased Ambari Install Wizard.

4. Enable NTP On The Cluster And On The Browser Host

The clocks of all the nodes in our cluster and the machine that runs the browser through which we access the Ambari Web interface must be able to synchronize with each other.

To check that the NTP service will be automatically started upon boot, run the following command on each host:

CentOS

```
chkconfig --list ntpd
systemctl is-enabled ntpd
```

To set the NTP service to auto-start on boot, run the following command on each host:

```
chkconfig ntpd on
systemctl enable ntpd
```

To start the NTP service, run the following command on each host:

```
service ntpd start
systemctl start ntpd
```

5. Edit The Host File

1. Using a text editor, open the hosts file on every host in our cluster. For example:

```
vi /etc/hosts
```

2. Add a line for each host in our cluster. The line should consist of the IP address and the FQDN. For example:

```
1.2.3.4   <fully.qualified.domain.name>
```

6. Set The Hostname

1. Confirm that the hostname is set by running the following command:

```
hostname -f
```

This should return the <fully.qualified.domain.name> we just set.

2. Using a text editor, open the network configuration file on every host and set the desired network configuration for each host. For example

```
vi /etc/sysconfig/network
```

3. Modify the HOSTNAME property to set the fully qualified domain name

```
NETWORKING=yes
```

HOSTNAME=<fully.qualified.domain.name>

7. 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:

chkconfig iptables off

chkconfig ip6tables off

chkconfig NetworkManager off

chkconfig network on

reboot the system

Ambari checks whether iptables is running during the Ambari Server setup process. If iptables is running, a warning displays, reminding us to check that required ports are open and available. The Host Confirm step in the Cluster Install Wizard also issues a warning for each host that has iptables running.

8. Disable Ipv6 And Selinux

Edit /boot/grub/grub.conf and append below parameters in the end of the kernel line of default kernel & reboot the system.

ipv6.disable=1 selinux=0

9. Download Packages

http://public-repo-1.hortonworks.com/HDP/centos6/2.x/updates/2.4.2.0/HDP-2.4.2.0-centos6-rpm.tar.gz

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

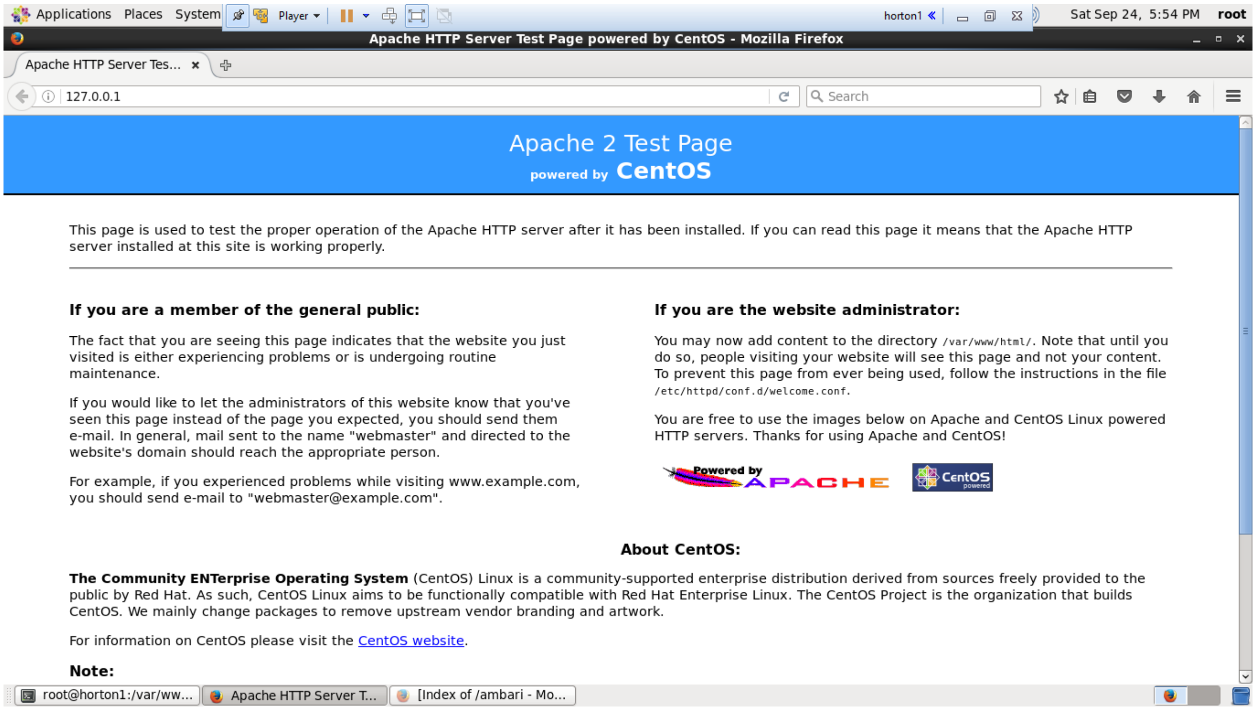
10. Getting Started Setting Up A Local Repository

1. Create an HTTP server.
a. On the mirror server, install an HTTP server (such as Apache httpd) .
b. Activate this web server.
c. Ensure that any firewall settings allow inbound HTTP access from our cluster nodes to our mirror server

```
yum install httpd
service http status
service httpd start
chkconfig http on
```

open web browser and check by typing these in url bar

127.0.0.1


Applications Places System
Player
horton1
Sat Sep 24, 5:54 PM
root
Apache HTTP Server Test Page powered by CentOS - Mozilla Firefox
Apache HTTP Server Tes...
127.0.0.1
Search
root@horton1:/var/ww...
Apache HTTP Server T...
[Index of /ambari - Mo...
Powered by APACHE
CentOS powered

Apache 2 Test Page
powered by **CentOS**

This page is used to test the proper operation of the Apache HTTP server after it has been installed. If you can read this page it means that the Apache HTTP server installed at this site is working properly.

If you are a member of the general public:

The fact that you are seeing this page indicates that the website you just visited is either experiencing problems or is undergoing routine maintenance.

If you would like to let the administrators of this website know that you've seen this page instead of the page you expected, you should send them e-mail. In general, mail sent to the name "webmaster" and directed to the website's domain should reach the appropriate person.

For example, if you experienced problems while visiting www.example.com, you should send e-mail to "webmaster@example.com".

If you are the website administrator:

You may now add content to the directory `/var/www/html/`. Note that until you do so, people visiting your website will see this page and not your content. To prevent this page from ever being used, follow the instructions in the file `/etc/httpd/conf.d/welcome.conf`.

You are free to use the images below on Apache and CentOS Linux powered HTTP servers. Thanks for using Apache and CentOS!

About CentOS:

The Community ENTerprise Operating System (CentOS) Linux is a community-supported enterprise distribution derived from sources freely provided to the public by Red Hat. As such, CentOS Linux aims to be functionally compatible with Red Hat Enterprise Linux. The CentOS Project is the organization that builds CentOS. We mainly change packages to remove upstream vendor branding and artwork.

For information on CentOS please visit the CentOS website.

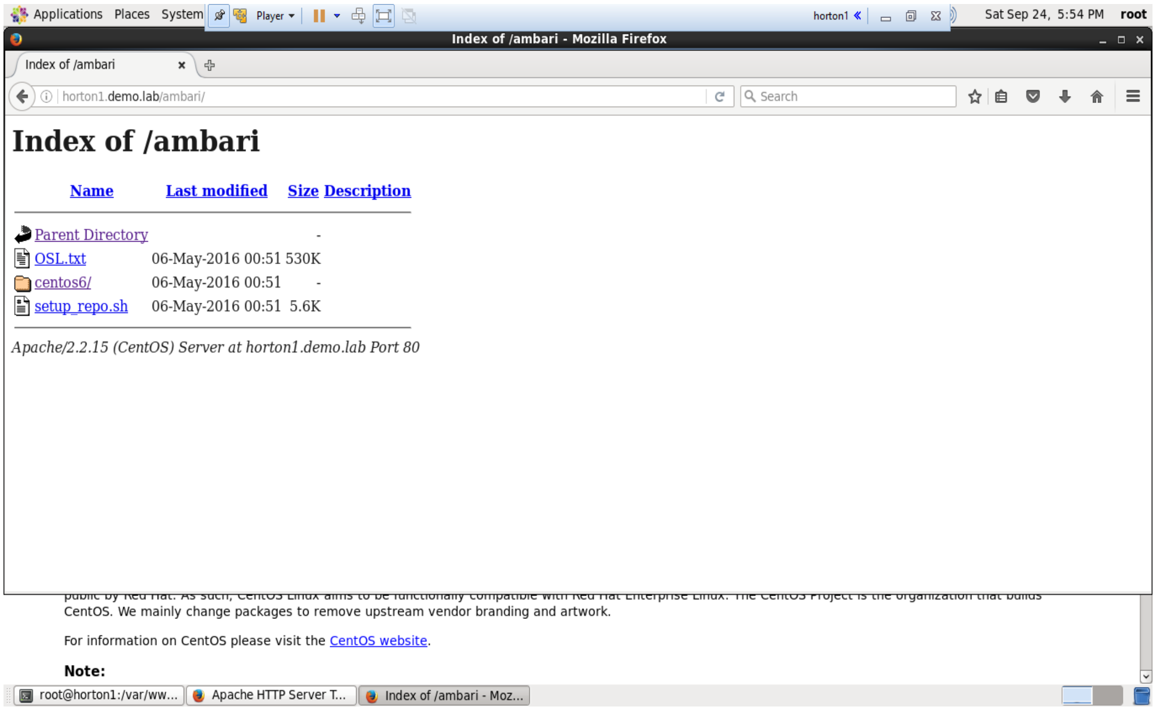
Note:

2. On our mirror server, create a directory for our web server.:

```
mkdir -p /var/www/html/
```

move the extracted tar files in this directori and create soft link

```
tar xzf ambari-2.2.2.0-centos6.tar.gz
tar xzf HDP-UTILS-1.1.0.20-centos6.tar.gz
tar xzf HDP-2.4.2.0-centos6-rpm.tar.gz
mv -v AMBARI-2.2.2.0   /var/www/html
mv -v HDP   /var/www/html
mv -v HDP-UTILS-1.1.0.20   /var/www/html
ln –s /var/www/html/ AMBARI-2.2.2.0   ambari
ln –s /var/www/html/ HDP  hdp
ln –s /var/www/html/ HDP-UTILS-1.1.0.20  hdp-utils
```



Index of /ambari

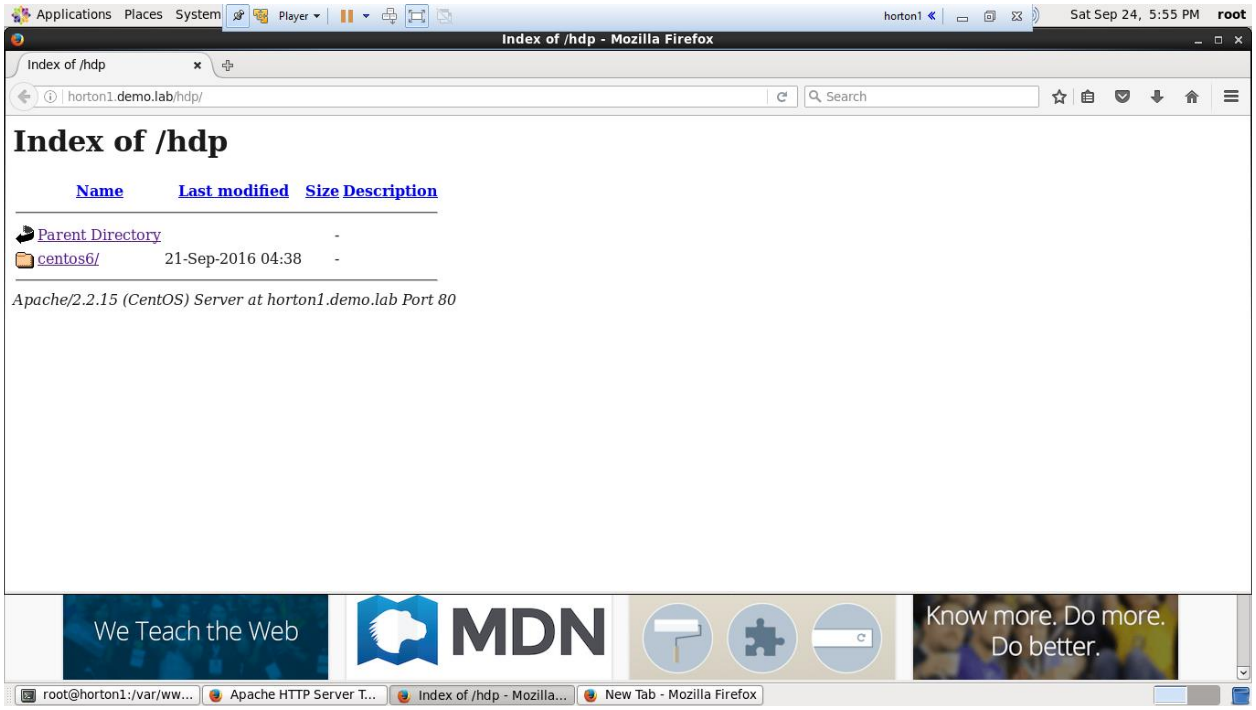
```
      Name          Last modified      Size Description
Parent Directory                          -
OSL.txt         06-May-2016 00:51  530K
centos6/        06-May-2016 00:51    -
setup_repo.sh   06-May-2016 00:51  5.6K
```

Apache/2.2.15 (CentOS) Server at horton1.demo.lab Port 80

public by Red Hat. As such, CentOS Linux aims to be functionally compatible with Red Hat Enterprise Linux. The CentOS Project is the organization that builds CentOS. We mainly change packages to remove upstream vendor branding and artwork.

For information on CentOS please visit the CentOS website.

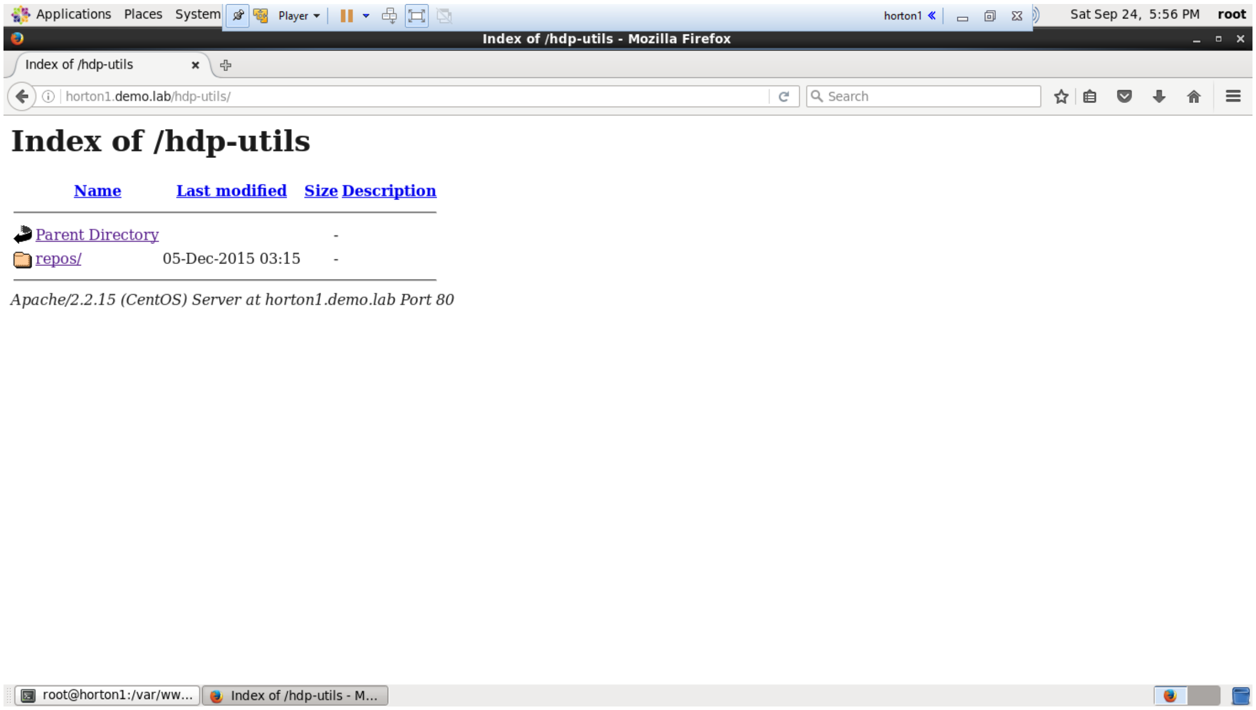
Note:



Index of /hdp

```
     Name            Last modified      Size  Description
Parent Directory                         -
centos6/        21-Sep-2016 04:38        -
```

Apache/2.2.15 (CentOS) Server at horton1.demo.lab Port 80


Applications Places System
Player
horton1
Sat Sep 24, 5:56 PM
root
Index of /hdp-utils - Mozilla Firefox
Index of /hdp-utils
horton1.demo.lab/hdp-utils/
Search
root@horton1:/var/ww...
Index of /hdp-utils - M...

Index of /hdp-utils

```
     Name            Last modified      Size  Description
Parent Directory                         -
repos/          05-Dec-2015 03:15        -
```

Apache/2.2.15 (CentOS) Server at horton1.demo.lab Port 80

3.Now open the file /etc/yum.repos.d/ambari.repo and enter the following information

[AMBARI-2.2.2.X]

name=ambari

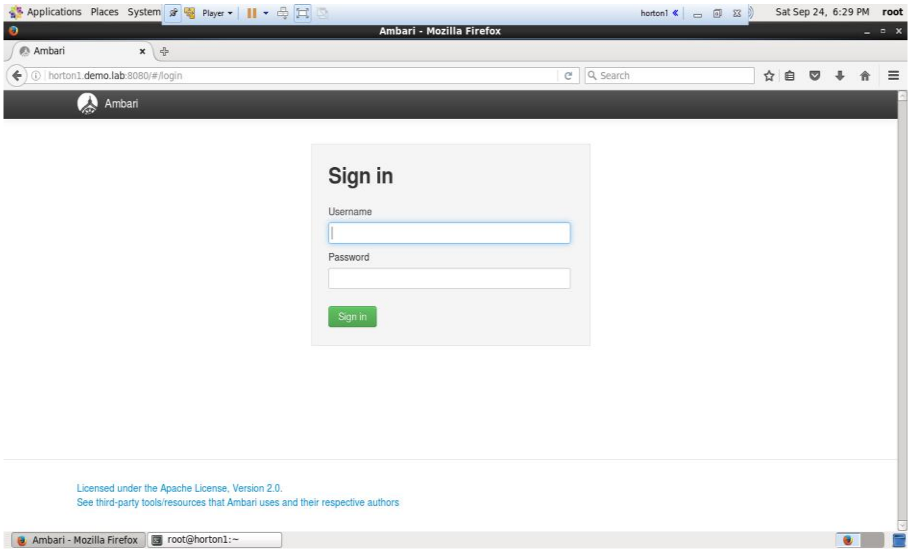
```
baseurl=http://horton1.demo.lab/ambari/centos6/2.2.2.0-460

gpgkey=http://horton1.demo.lab/ambari/centos6/2.2.2.0-460/RPM-GPG-KEY/RPM-GPG-KEY-Jenkins

enabled=1
```

4. Confirm availability of the repositories

```
yum repolist
```

11. Installing Ambari

1. Log in to our host as root.
2. Clean the repository
   ```
   yum clean all
   ```
3. Confirm that the repository is configured by checking the repo list.
   ```
   yum repolist
   ```
   We should see values similar to the following for Ambari repositories in the list.
4. Find ambary packages
   ```
   Yum list all | grep –I ambari
   ```
   ```
   ambari-agent.x86_64                  2.2.2.0-460         @AMBARI-2.2.2.X
   ambari-metrics-grafana.x86_64        2.2.2.0-460         @AMBARI-2.2.2.X
   ambari-metrics-hadoop-sink.x86_64    2.2.2.0-460         @AMBARI-2.2.2.X
   ambari-metrics-monitor.x86_64        2.2.2.0-460         @AMBARI-2.2.2.X
   ambari-server.x86_64                 2.2.2.0-460         @AMBARI-2.2.2.X
   ambari-metrics-collector.x86_64      2.2.2.0-460          AMBARI-2.2.2.X
   ambari-metrics-common.noarch         2.2.2.0-460          AMBARI-2.2.2.X
   smartsense-hst.x86_64                1.2.2.0-460          AMBARI-2.2.2.X
   ```
5. Install the Ambari bits. This also installs the default PostgreSQL Ambari database.
   ```
   yum install ambari-server.x86_64
   ```
6. Choose the default options.
7. To check the ambari status
   ```
   ambari-server status
   ```
8. To start ambary server
   ```
   ambari-server start
   ```
9. To stop ambari server
   ```
   ambary-server stop
   ```



Licensed under the Apache License, Version 2.0.
See third-party tools/resources that Ambari uses and their respective authors

12. Installing, Configuring, And Deploying A HDP Cluster

1. Log In to Apache Ambari

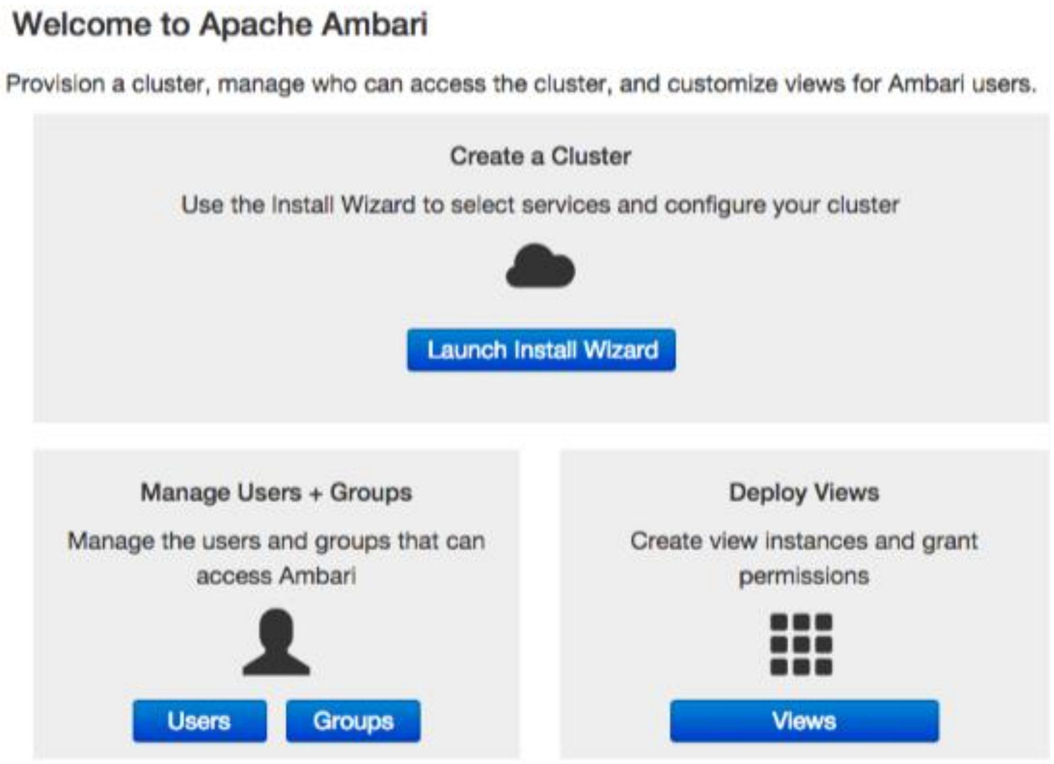
After starting the Ambari service, open Ambari Web using a web browser.

Point our browser to http://<our.ambari.server>:8080,where <our.ambari.server> is the name of our ambari server host. For example, a default Ambari server host is located at http://c6401.ambari.apache.org:8080.

Log in to the Ambari Server using the default user name/password: admin/admin. We can change these credentials later.

.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.

Create a Cluster

Use the Install Wizard to select services and configure your cluster

Manage Users + Groups

Manage the users and groups that can access Ambari

Deploy Views

Create view instances and grant permissions

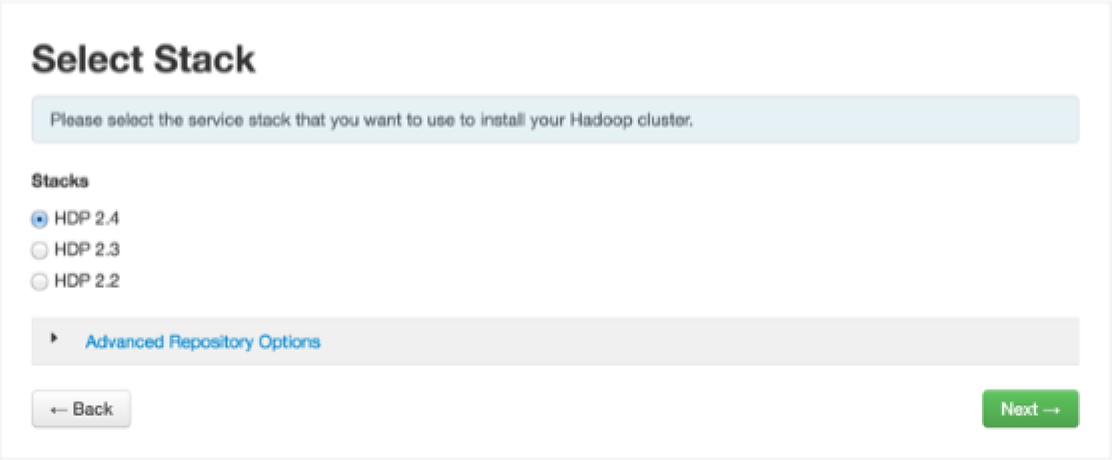
3. Name our Cluster

In Name our cluster, type a name for the cluster we want to create. Use no white spaces or special characters in the name.

Choose Next

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 we want to install.



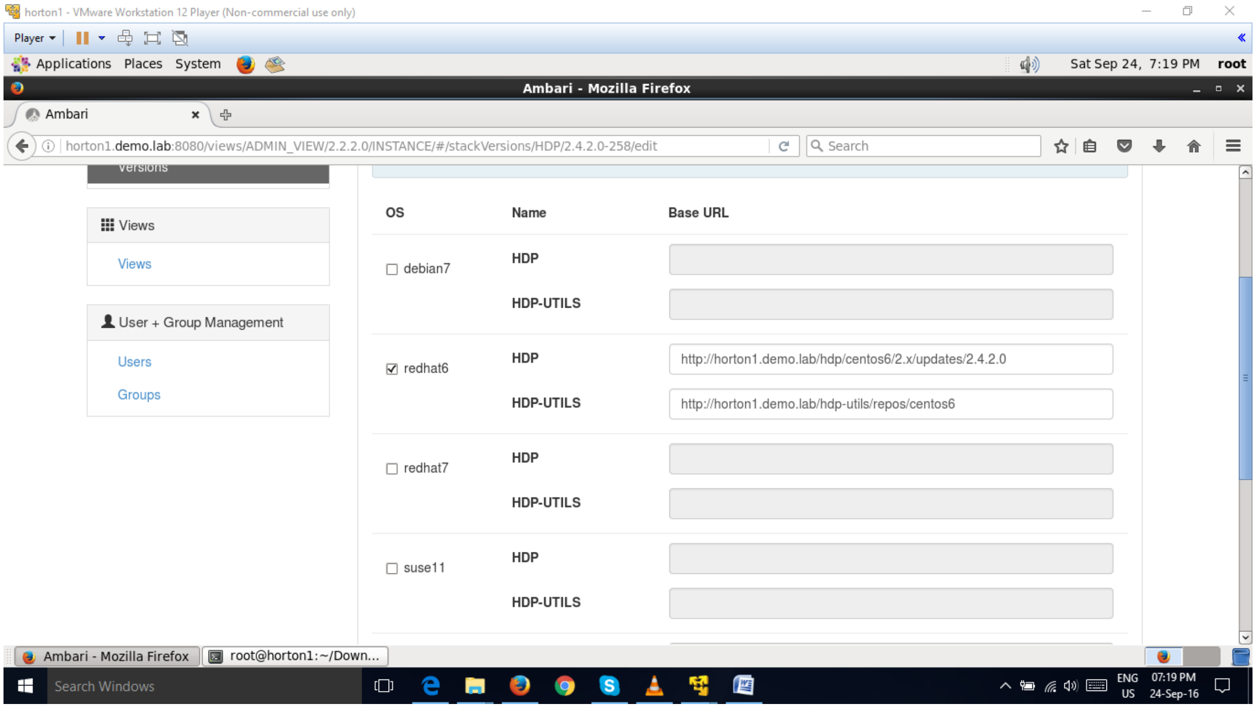
Select Stack

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

Stacks

● HDP 2.4
○ HDP 2.3
○ HDP 2.2

5. Expand Advanced Repository Options to select the Base URL of a repository from which Stack software packages download.



horton1.demo.lab:8080/views/ADMIN_VIEW/2.2.2.0/INSTANCE/#/stackVersions/HDP/2.4.2.0-258/edit

OS	Name	Base URL
☐ debian7	HDP	
	HDP-UTILS	
☑ redhat6	HDP	http://horton1.demo.lab/hdp/centos6/2.x/updates/2.4.2.0
	HDP-UTILS	http://horton1.demo.lab/hdp-utils/repos/centos6
☐ redhat7	HDP	
	HDP-UTILS	
☐ suse11	HDP	
	HDP-UTILS	

6. In order to build up the cluster, the install wizard prompts us for general information about how we want to set it up. We need to supply the FQDN of each of our hosts. The wizard also needs to access the private key file we 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.

7. Confirm Hosts prompts us to confirm that Ambari has located the correct hosts for our 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, we 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, we may notice a yellow box that indicates some warnings were encountered during the check process. For example, our 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 we clear any issues we may encounter and let we run Rerun Checks.

When we are satisfied with the list of hosts, choose Next

8. Based on the Stack chosen during Select Stack, we are presented with the choice of Services to install into the cluster. HDP Stack comprises many services. We 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.

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

9. The Ambari install wizard assigns the master components for selected services to appropriate hosts in our 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.

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

10. Assign Slaves and Clients

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

- 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 our mouse over the asterisk to see which master components are on that host.
- Fine-tune our selections by using the checkboxes next to specific hosts.
- When we are satisfied with our assignments, choose Next.

11. The Customize Services step presents we with a set of tabs that let we review and modify our HDP cluster setup. The wizard attempts to set reasonable defaults for each of the options. We are strongly encouraged to review these settings as our requirements might be slightly different.
    Browse through each service tab and by hovering our cursor over each of the properties, we can see a brief description of what the property does. The number of service tabs shown depends on the services we decided to install in our 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 our environment but we 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
    We 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.

12. The assignments we have made are displayed. Check to make sure everything is correct. If we need to make changes, use the left navigation bar to return to the appropriate screen.

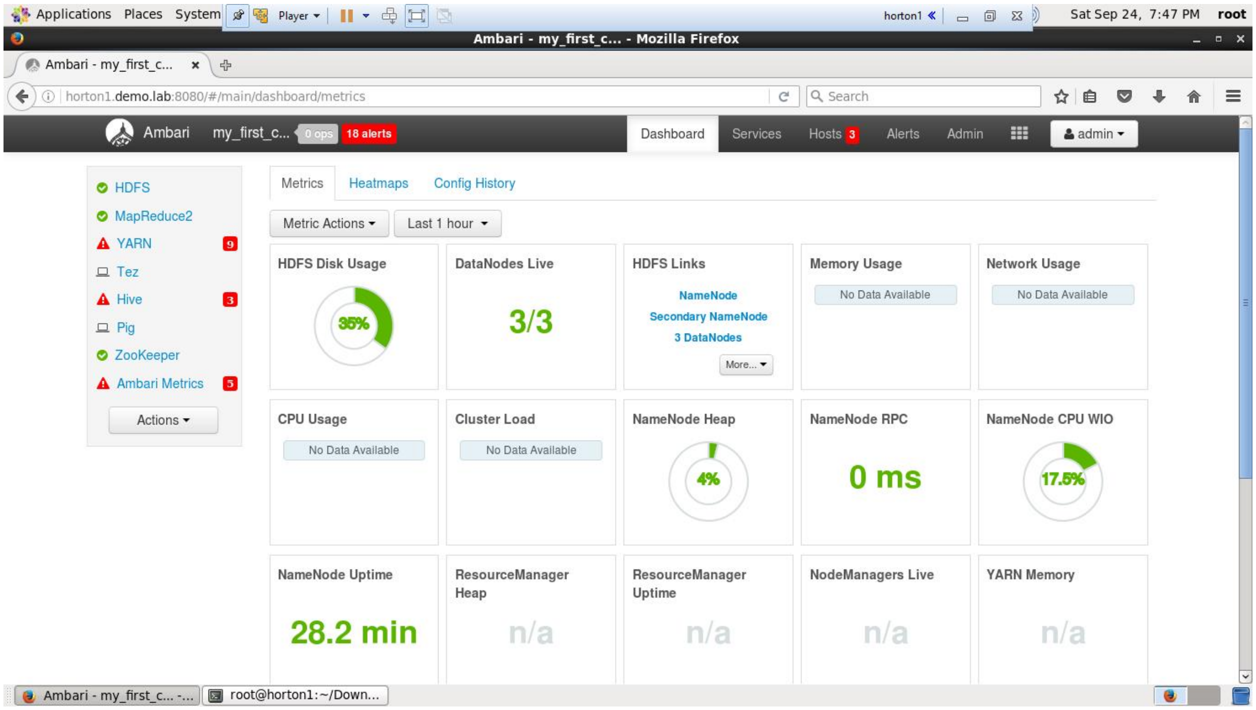
    To print our information for later reference, choose Print.

    When we are satisfied with our choices, choose Deploy.

13. 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 our 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. We 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.

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



13. Overview: Ambari User's

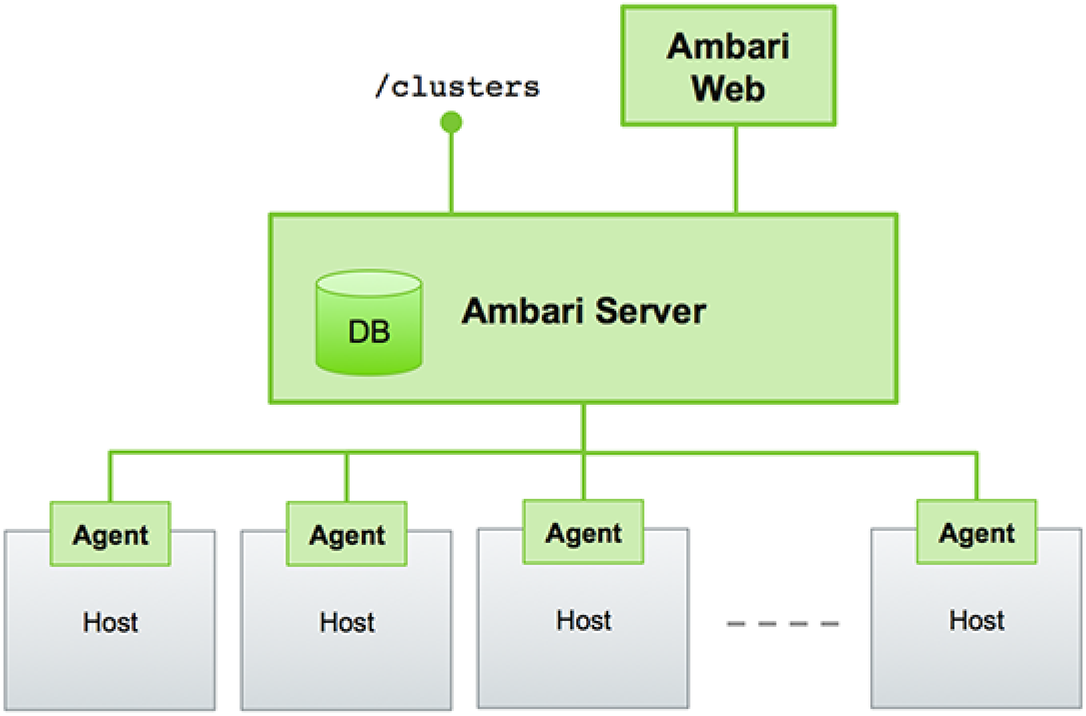
Hadoop is a large-scale, distributed data storage and processing infrastructure using clusters of commodity hosts networked together. Monitoring and managing such complex distributed systems is a non-trivial task. To help we manage the complexity, Apache Ambari collects a wide range of information from the cluster's nodes and services and presents it to we in an easy-to-read and use, centralized web interface, Ambari Web.

Ambari Web displays information such as service-specific summaries, graphs, and alerts. We use Ambari Web to create and manage our HDP cluster and to perform basic operational tasks such as starting and stopping services, adding hosts to our cluster, and updating service configurations. We also can use Ambari Web to perform administrative tasks for our cluster such as enabling Kerberos security and performing Stack upgrades.

14. Architecture

The Ambari Server serves as the collection point for data from across our cluster. Each host has a copy of the Ambari Agent - either installed automatically by the Install wizard or manually - which allows the Ambari Server to control each host.

Figure - Ambari Server Architecture



15. Accessing Ambari Web

Typically, we start the Ambari Server and Ambari Web as part of the installation process. If Ambari Server is stopped, we can start it using a command line editor on the Ambari Server host machine. Enter the following command:

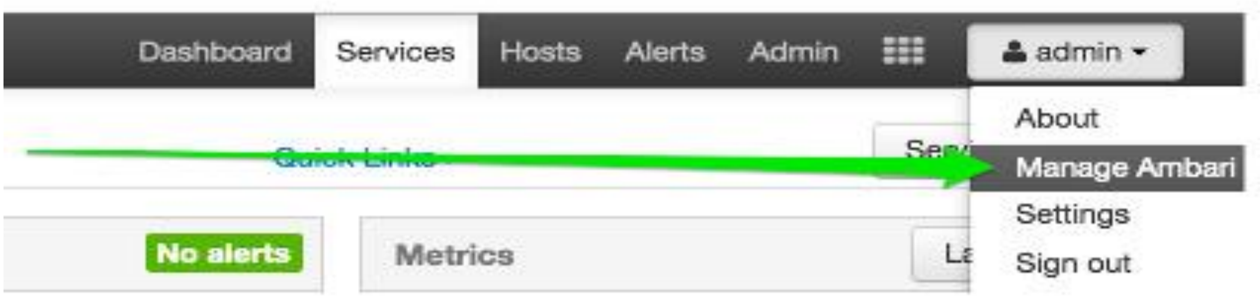
ambari-server start

To access Ambari Web, open a supported browser and enter the Ambari Web URL:

http://<our.ambari.server>:8080

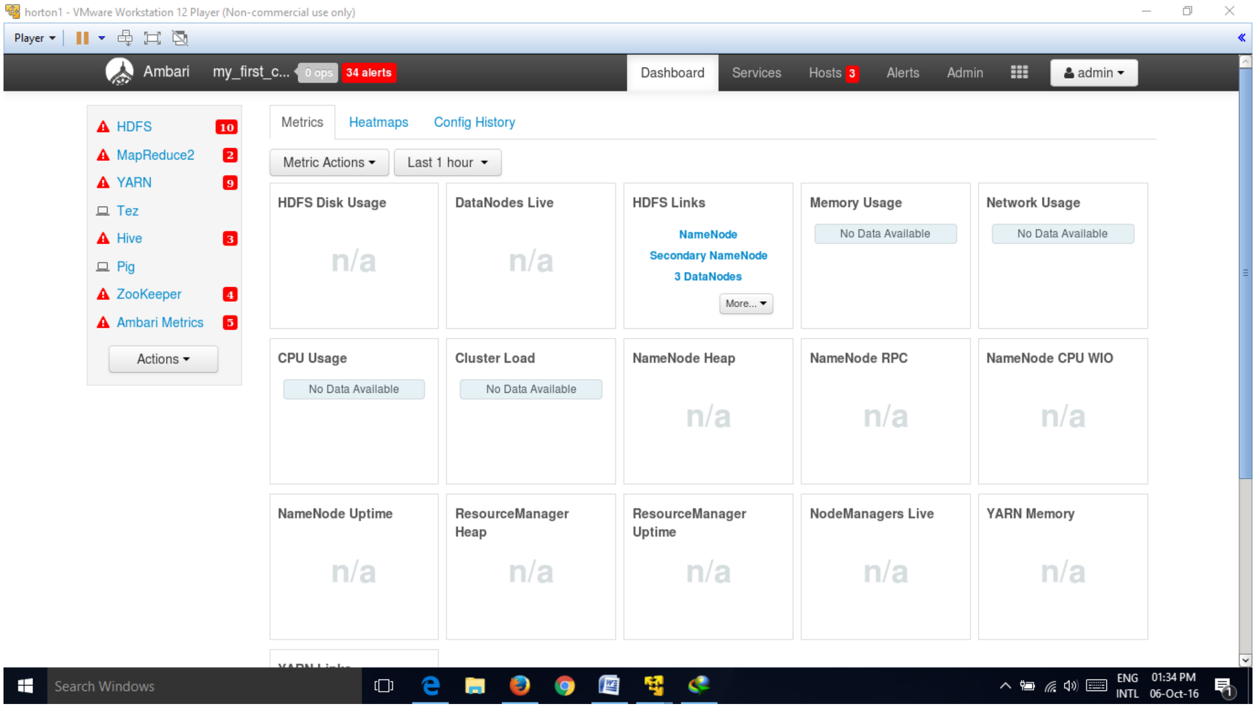
Enter our user name and password. If this is the first time Ambari Web is accessed, use the default values, admin/admin.

These values can be changed, and new users provisioned, using the Manage Ambari option.



16. Viewing the Cluster Dashboard

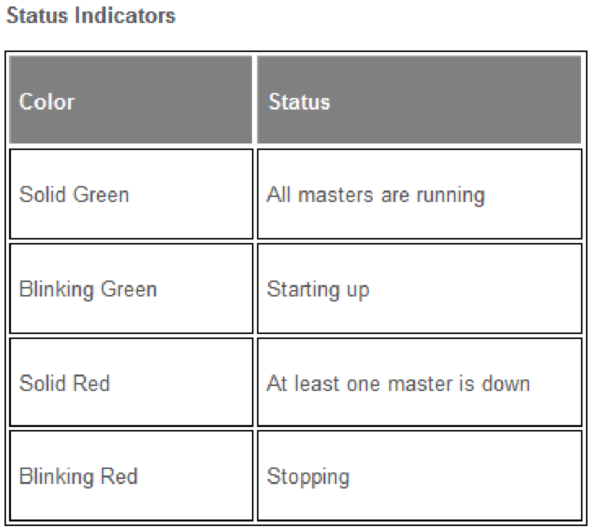
Ambari Web displays the **Dashboard** page as the home page. Use the **Dashboard** to view the operating status of our cluster. Each metrics widget displays status information for a single service in our HDP cluster. The Dashboard displays all metrics for the HDFS, YARN, HBase, and Storm services, and cluster-wide metrics by default. We can add and remove individual widgets, and rearrange the Dashboard by dragging and dropping each widget to a new location in the dashboard. Status information appears as simple pie and bar charts, more complex charts showing usage and load, sets of links to additional data sources, and values for operating parameters such as uptime and average RPC queue wait times. Most widgets display a single fact by default. For example, HDFS Disk Usage displays a load chart and a percentage figure.



Each Service installed in our cluster also has a Service-specific dashboard.

17. Scanning Service Status

Notice the color of the dot appearing next to each component name in a list of components, services or hosts. The dot color and blinking action indicates operating status of each component, service, or host.



Status Indicators

Color	Status
Solid Green	All masters are running
Blinking Green	Starting up
Solid Red	At least one master is down
Blinking Red	Stopping

18. Widget Descriptions

The Dashboard includes metrics for the following services:

View **Metrics** that indicate the operating status of our cluster on the Ambari Dashboard. Each metrics widget displays status information for a single service in our HDP cluster. The Ambari Dashboard displays all metrics for the HDFS, YARN, HBase, and Storm services, and cluster-wide metrics by default.

We can add and remove individual widgets, and rearrange the dashboard by dragging and dropping each widget to a new location in the dashboard.

Status information appears as simple pie and bar charts, more complex charts showing usage and load, sets of links to additional data sources, and values for operating parameters such as uptime and average RPC queue wait times. Most widgets display a single fact by default. For example, HDFS Disk Usage displays a load chart and a percentage figure. The Ambari Dashboard includes metrics for the following services:

Ambari Service Metrics and Descriptions

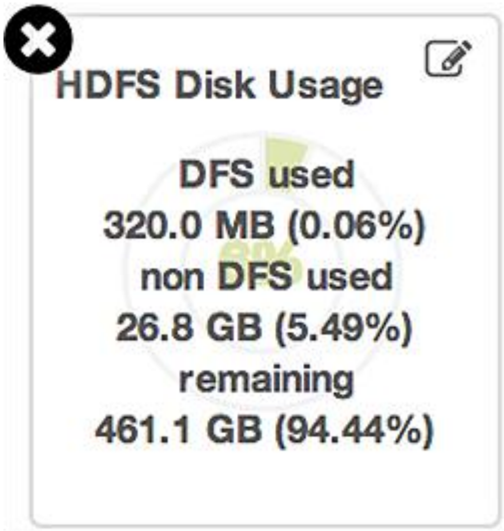
Metric:	Description:
HDFS	
HDFS Disk Usage	The Percentage of DFS used, which is a combination of DFS and non-DFS used.
Data Nodes Live	The number of DataNodes live, as reported from the NameNode.
NameNode Heap	The percentage of NameNode JVM Heap used.
NameNode RPC	The average RPC queue latency.
NameNode CPU WIO	The percentage of CPU Wait I/O.
NameNode Uptime	The NameNode uptime calculation.
YARN (HDP 2.1 or later Stacks)	
ResourceManager Heap	The percentage of ResourceManager JVM Heap used.
ResourceManager Uptime	The ResourceManager uptime calculation.
NodeManagers Live	The number of DataNodes live, as reported from the ResourceManager.
YARN Memory	The percentage of available YARN memory (used vs. total available).
HBase	
HBase Master Heap	The percentage of NameNode JVM Heap used.
HBase Ave Load	The average load on the HBase server.

Metric:	Description:
HBase Master Uptime	The HBase Master uptime calculation.
Region in Transition	The number of HBase regions in transition.
Storm (HDP 2.1 or later Stacks)	
Supervisors Live	The number of Supervisors live, as reported from the Nimbus server.

19. Widget Details

To see more detailed information about a service, hover our cursor over a Metrics widget.

More detailed information about the service displays, as shown in the following example:



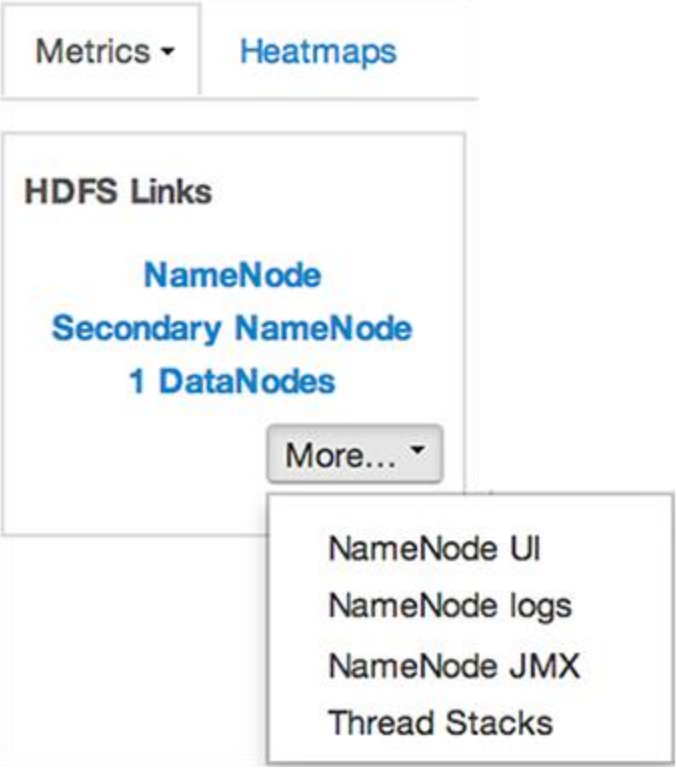
HDFS Disk Usage

DFS used
320.0 MB (0.06%)
non DFS used
26.8 GB (5.49%)
remaining
461.1 GB (94.44%)

- To remove a widget from the mashup, click the white X.
- To edit the display of information in a widget, click the pencil icon.

20. Linking to Service UIs

The HDFS Links and HBase Links widgets list HDP components for which links to more metrics information, such as thread stacks, logs and native component UIs are available. For example, we can link to NameNode, Secondary NameNode, and DataNode components for HDFS, using the links shown in the following example:



Choose the More drop-down to select from the list of links available for each service. The Ambari Dashboard includes additional links to metrics for the following services:

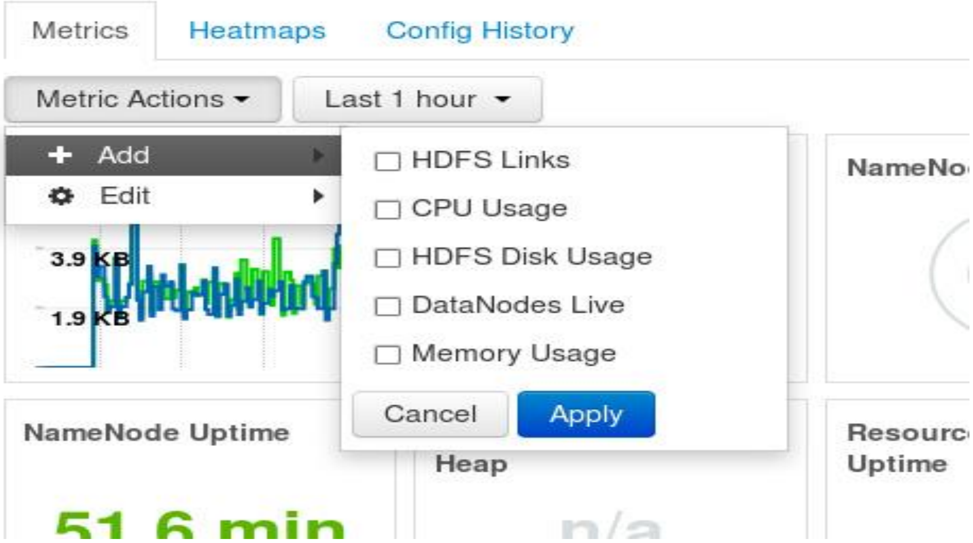
Links to More Metrics for HDP Services

Service:	Metric:	Description:
HDFS		
	NameNode UI	Links to the NameNode UI.
	NameNode Logs	Links to the NameNode logs.
	NameNode JMX	Links to the NameNode JMX servlet.
	Thread Stacks	Links to the NameNode thread stack traces.
HBase		
	HBase Master UI	Links to the HBase Master UI.
	HBase Logs	Links to the HBase logs.
	ZooKeeper Info	Links to ZooKeeper information.
	HBase Master JMX	Links to the HBase Master JMX servlet.
	Debug Dump	Links to debug information.
	Thread Stacks	Links to the HBase Master thread stack traces.

21. Adding a Widget to the Dashboard

To replace a widget that has been removed from the dashboard:

1. Select the Metrics drop-down, as shown in the following example:

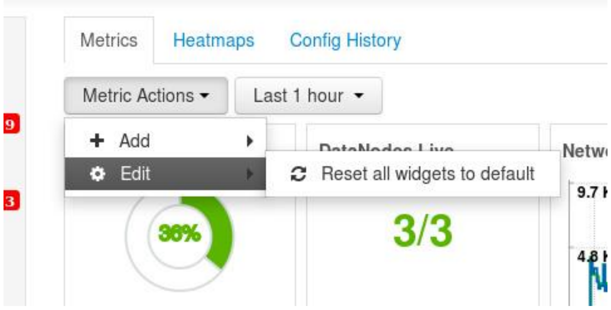


2. Choose Add.
3. Select a metric, such as Region in Transition.
4. Choose Apply.

22. Resetting the Dashboard

To reset all widgets on the dashboard to display default settings:

1. Select the Metrics drop-down, as shown in the following example:



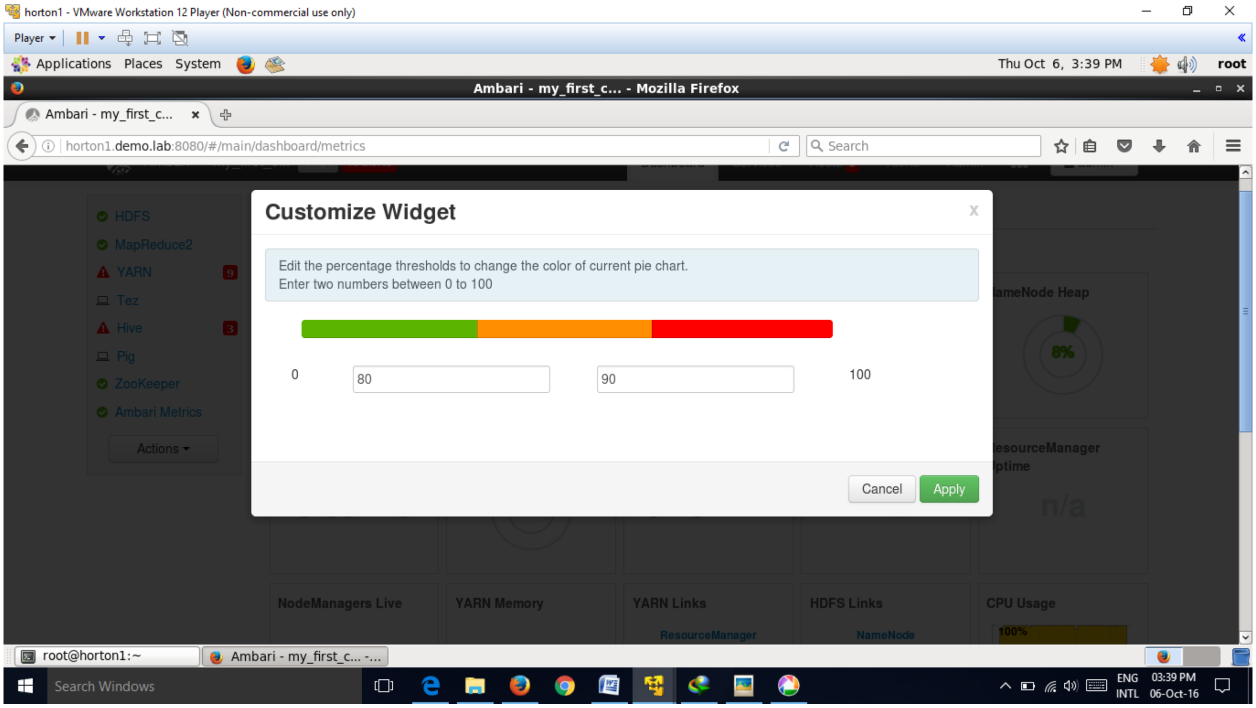
2. Choose Edit.
3. Choose Reset all widgets to default.

23. Customizing Widget Display

To customize the way a service widget displays metrics information:

1. Hover our cursor over a service widget.
2. Select the pencil-shaped, edit icon that appears in the upper-right corner.

   The Customize Widget pop-up window displays properties that we can edit, as shown in the following example.



Customize Widget

Edit the percentage thresholds to change the color of current pie chart.
Enter two numbers between 0 to 100

0 | 80 | 90 | 100

3. Follow the instructions in the Customize Widget pop-up to customize widget appearance.

   In this example, we can adjust the thresholds at which the HDFS Capacity bar chart changes color, from green to orange to red.

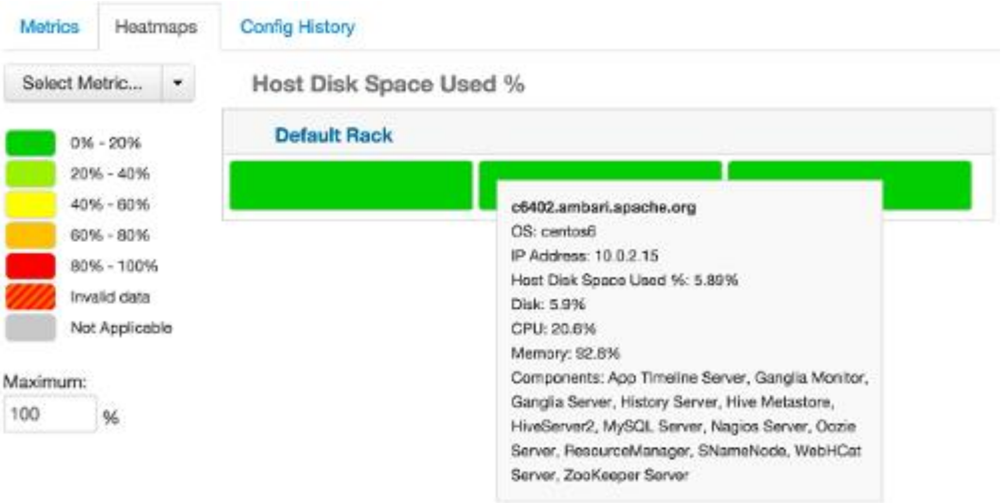
4. To save our changes and close the editor, choose `Apply`.
5. To close the editor without saving any changes, choose `Cancel`.

Note

Not all widgets support editing.

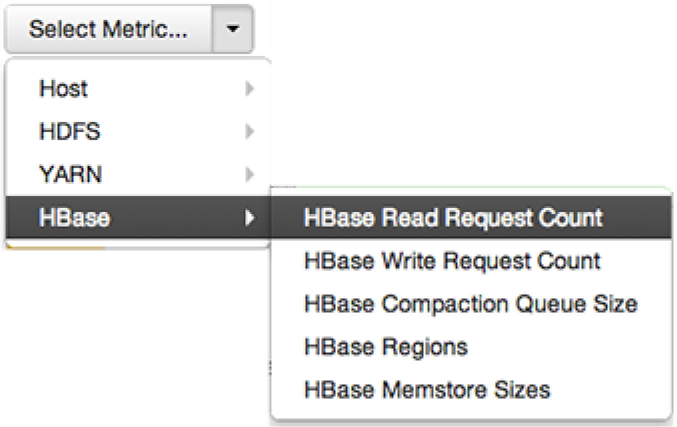
24. Viewing Cluster Heatmaps

Heatmaps provides a graphical representation of our overall cluster utilization using simple color coding.



c6402.ambari.apache.org
OS: centos6
IP Address: 10.0.2.15
Host Disk Space Used %: 5.89%
Disk: 5.9%
CPU: 20.6%
Memory: 92.6%
Components: App Timeline Server, Ganglia Monitor, Ganglia Server, History Server, Hive Metastore, HiveServer2, MySQL Server, Nagios Server, Oozie Server, ResourceManager, SNameNode, WebHCat Server, ZooKeeper Server

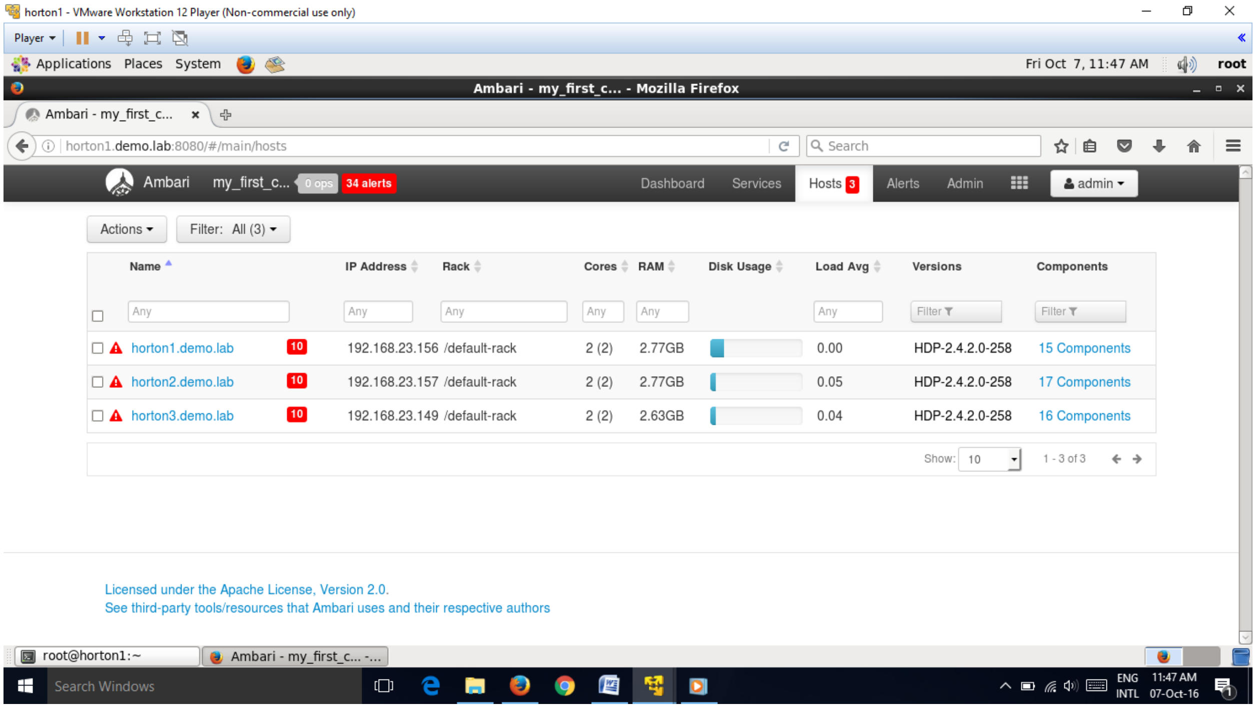
A colored block represents each host in our cluster. To see more information about a specific host, hover over the block representing the host in which we are interested. A pop-up window displays metrics about HDP components installed on that host. Colors displayed in the block represent usage in a unit appropriate for the selected set of metrics. If any data necessary to determine state is not available, the block displays "Invalid Data". Changing the default maximum values for the heatmap lets we fine tune the representation. Use the Select Metric drop-down to select the metric type.



25. Working with Hosts

Use Hosts to view hosts in our cluster on which Hadoop services run. Use options on **Actions** to perform actions on one or more hosts in our cluster.

View individual hosts, listed by fully-qualified domain name, on the Hosts landing page.


horton1 - VMware Workstation 12 Player (Non-commercial use only)
Player
Applications Places System
Fri Oct 7, 11:47 AM
root
Ambari - my_first_c... - Mozilla Firefox
horton1.demo.lab:8080/#/main/hosts
Search
Ambari my_first_c... 0 ops 34 alerts
Dashboard Services Hosts 3 Alerts Admin
admin
Actions
Filter: All (3)
root@horton1:~
Ambari - my_first_c...
Search Windows
ENG INTL 11:47 AM 07-Oct-16

	Name		IP Address	Rack	Cores	RAM	Disk Usage	Load Avg	Versions	Components
□	Any		Any	Any	Any	Any		Any	Filter	Filter
□	horton1.demo.lab	10	192.168.23.156	/default-rack	2 (2)	2.77GB		0.00	HDP-2.4.2.0-258	15 Components
□	horton2.demo.lab	10	192.168.23.157	/default-rack	2 (2)	2.77GB		0.05	HDP-2.4.2.0-258	17 Components
□	horton3.demo.lab	10	192.168.23.149	/default-rack	2 (2)	2.63GB		0.04	HDP-2.4.2.0-258	16 Components

Show: 10 1 - 3 of 3

Licensed under the Apache License, Version 2.0.
See third-party tools/resources that Ambari uses and their respective authors

26. Determining Host Status

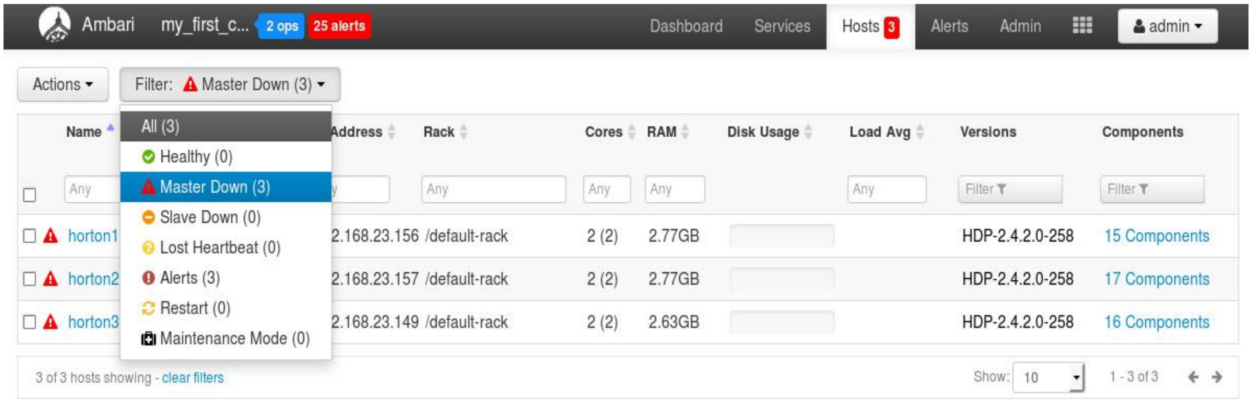
A colored dot beside each host name indicates operating status of each host, as follows:

- Red - At least one master component on that host is down. Hover to see a tooltip that lists affected components.
- Orange - At least one slave component on that host is down. Hover to see a tooltip that lists affected components.
- Yellow - Ambari Server has not received a heartbeat from that host for more than 3 minutes.
- Green - Normal running state.

A red condition flag overrides an orange condition flag, which overrides a yellow condition flag. In other words, a host having a master component down may also have other issues. The following example shows three hosts, one having a master component down, one having a slave component down, and one healthy. Warning indicators appear next to hosts having a component down.

27. Filtering the Hosts List

Use Filters to limit listed hosts to only those having a specific operating status. The number of hosts in our cluster having a listed operating status appears after each status name, in parenthesis. For example, the following cluster has one host having healthy status and three hosts having Maintenance Mode turned on.



	Name	Address	Rack	Cores	RAM	Disk Usage	Load Avg	Versions	Components
□	Any		Any	Any	Any		Any	Filter	Filter
□	horton1	2.168.23.156	/default-rack	2 (2)	2.77GB			HDP-2.4.2.0-258	15 Components
□	horton2	2.168.23.157	/default-rack	2 (2)	2.77GB			HDP-2.4.2.0-258	17 Components
□	horton3	2.168.23.149	/default-rack	2 (2)	2.63GB			HDP-2.4.2.0-258	16 Components

3 of 3 hosts showing - clear filters Show: 10 1 - 3 of 3

For example, to limit the list of hosts appearing on Hosts home to only those with Healthy status, select Filters, then choose the Healthy option. In this case, one host name appears on Hosts home. Alternatively, to limit the list of hosts appearing on Hosts home to only those having Maintenance Mode on, select Filters, then choose the Maintenance Mode option. In this case, three host names appear on Hosts home.

Use the general filter tool to apply specific search and sort criteria that limits the list of hosts appearing on the Hosts page.

28. Performing Host-Level Actions

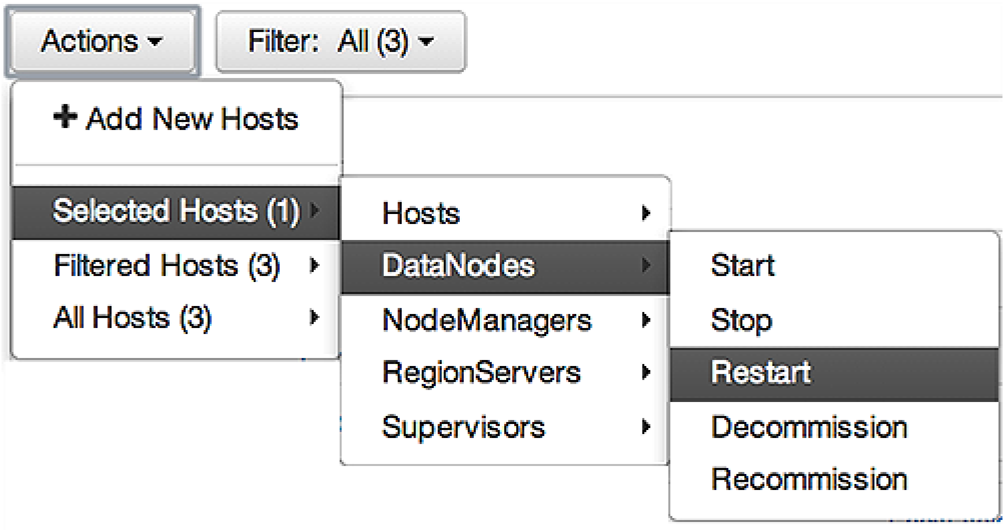
Use Actions to act on one, or multiple hosts in our cluster. Actions performed on multiple hosts are also known as bulk operations.

Actions comprises three menus that list the following option types:

- Hosts - lists selected, filtered or all hosts options, based on our selections made using Hosts home and Filters.
- Objects - lists component objects that match our host selection criteria.
- Operations - lists all operations available for the component objects we selected.

For example, to restart DataNodes on one host:

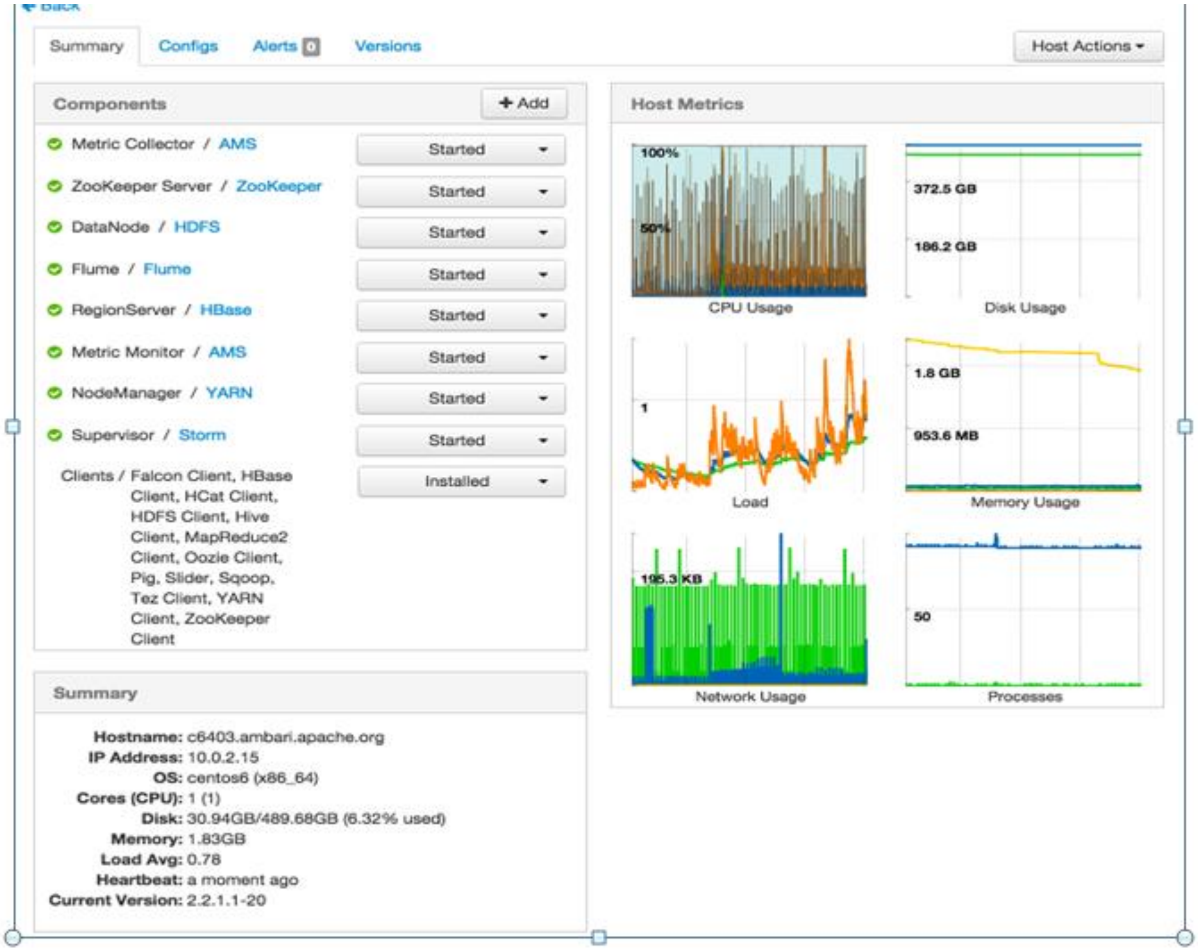
1. In Hosts, select a host running at least one DataNode.
2. In Actions, choose `Selected Hosts > DataNodes > Restart`, as shown in the following image.



3. Choose OK to confirm starting the selected operation.
4. Optionally, use Monitoring Background Operations to follow, diagnose or troubleshoot the restart operation.

29. Viewing Components on a Host

To manage components running on a specific host, choose a FQDN on the Hosts page.Summary-Components lists all components installed on that host.



Metric Collector / AMS — Started

ZooKeeper Server / ZooKeeper — Started

DataNode / HDFS — Started

Flume / Flume — Started

RegionServer / HBase — Started

Metric Monitor / AMS — Started

NodeManager / YARN — Started

Supervisor / Storm — Started

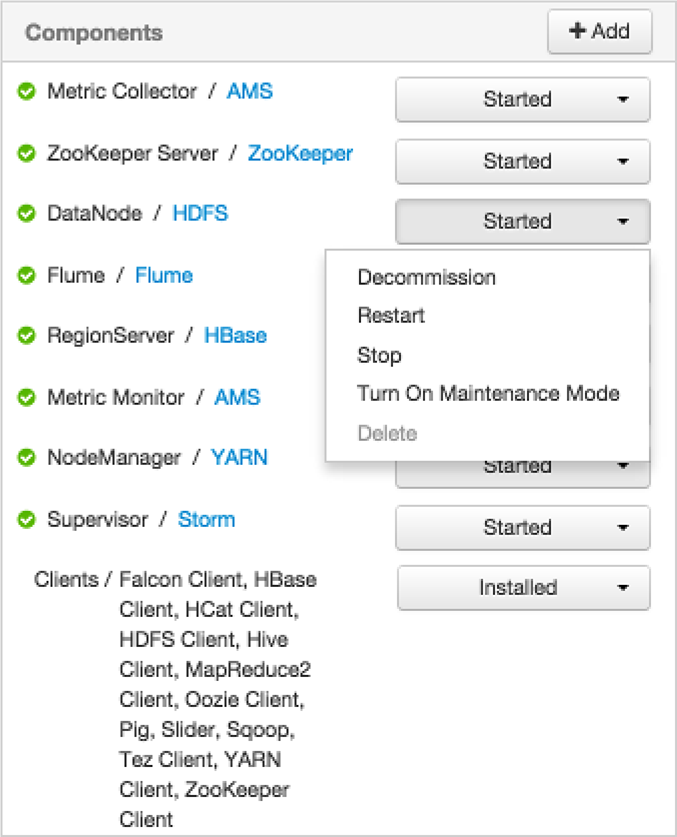
Clients / Falcon Client, HBase Client, HCat Client, HDFS Client, Hive Client, MapReduce2 Client, Oozie Client, Pig, Slider, Sqoop, Tez Client, YARN Client, ZooKeeper Client — Installed

Summary

Hostname: c6403.ambari.apache.org
IP Address: 10.0.2.15
OS: centos6 (x86_64)
Cores (CPU): 1 (1)
Disk: 30.94GB/489.68GB (6.32% used)
Memory: 1.83GB
Load Avg: 0.78
Heartbeat: a moment ago
Current Version: 2.2.1.1-20

Choose options in `Host Actions`, to start, stop, restart, delete, or turn on maintenance mode for all components installed on the selected host.

Alternatively, choose action options from the drop-down menu next to an individual component on a host. The drop-down menu shows current operation status for each component, For example, we can decommission, restart, or stop the DataNode component (started) for HDFS, by selecting one of the options shown in the following example:

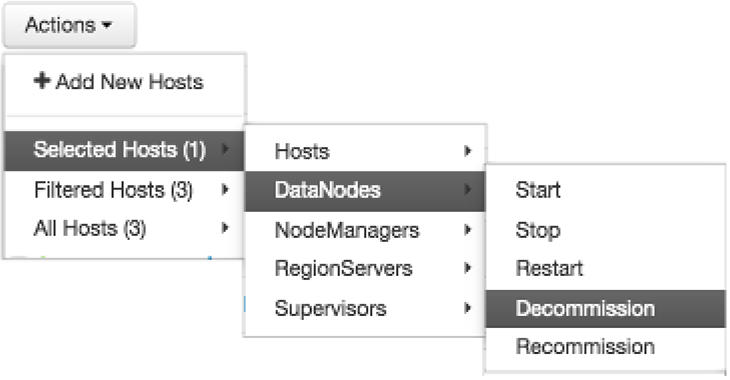


30. How to Decommission a Component

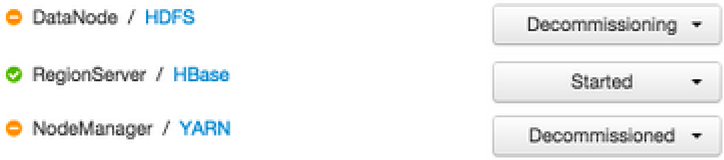
To decommission a component using Ambari Web browse **Hosts** to find the host FQDN on which the component resides.

Using **Actions**, select **HostsComponent Type**, then choose **Decommission**.

For example:



The UI shows "Decommissioning" status while steps process, then "Decommissioned" when complete.



DataNode / HDFS — Decommissioning

RegionServer / HBase — Started

NodeManager / YARN — Decommissioned

31. How to Delete a Component

To delete a component using Ambari Web, on `Hosts` choose the host FQDN on which the component resides.

1. In `Components`, find a decommissioned component.
2. Stop the component, if necessary.

   Note

   A decommissioned slave component may restart in the decommissioned state.

3. For a decommissioned component, choose **Delete** from the component drop-down menu.

   Note

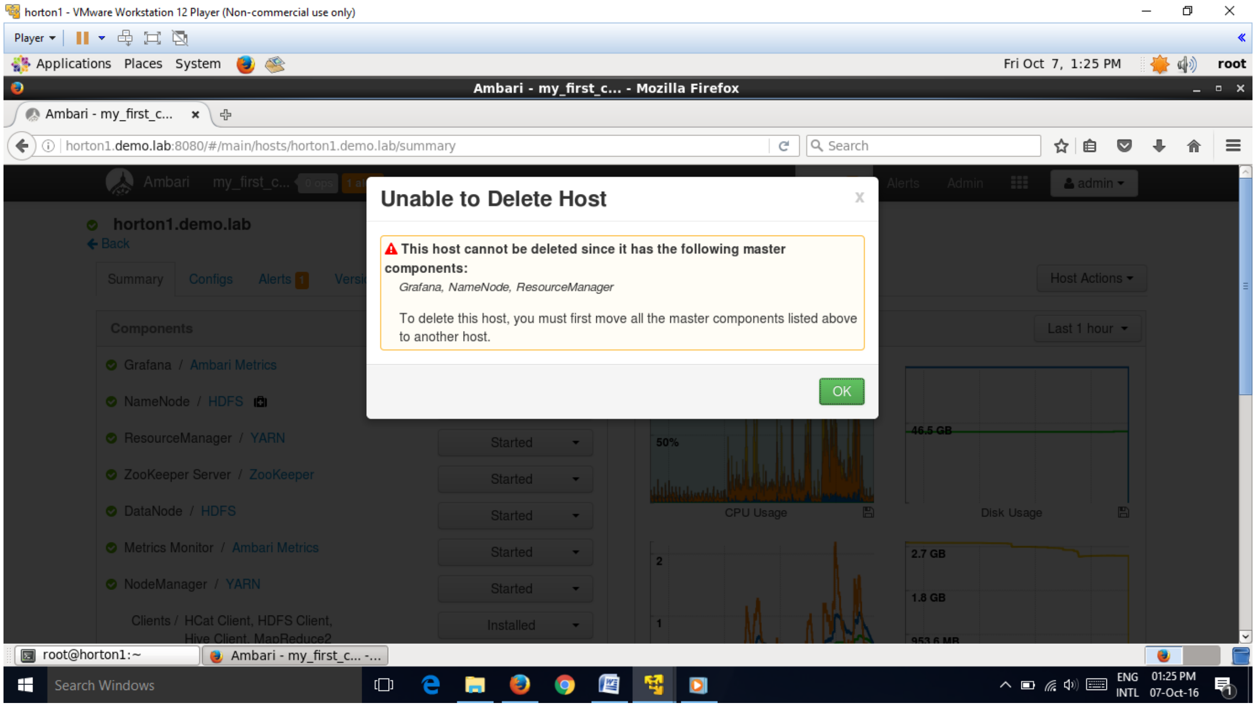
   Restarting services enables Ambari to recognize and monitor the correct number of components.

4. Deleting a slave component, such as a DataNode does not automatically inform a master component, such as a NameNode to remove the slave component from its exclusion list. Adding a deleted slave component back into the cluster presents the following issue; the added slave remains decommissioned from the master's perspective. Restart the master component, as a work-around.

32. How to Delete a Host from a Cluster

1. In Hosts, click on a host name.
2. On the Host-Details page, select Host Actions drop-down menu.
3. Choose Delete.

If we have not completed prerequisite steps, a warning message similar to the following one appears:


horton1 - VMware Workstation 12 Player (Non-commercial use only)
Ambari - my_first_c... - Mozilla Firefox
horton1.demo.lab:8080/#/main/hosts/horton1.demo.lab/summary
horton1.demo.lab
Summary
Configs
Alerts
Host Actions
Components
Grafana / Ambari Metrics
NameNode / HDFS
ResourceManager / YARN
ZooKeeper Server / ZooKeeper
DataNode / HDFS
Metrics Monitor / Ambari Metrics
NodeManager / YARN
Clients / HCat Client, HDFS Client, Hive Client, MapReduce2
Started
Installed
CPU Usage
Disk Usage
Last 1 hour

Unable to Delete Host

This host cannot be deleted since it has the following master components:

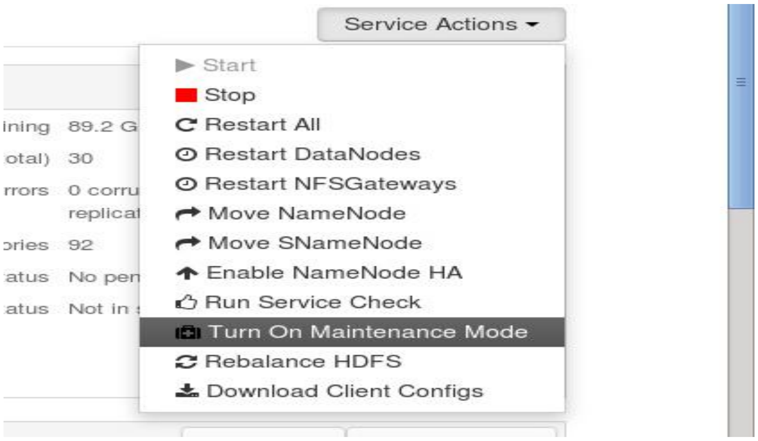
Grafana, NameNode, ResourceManager

To delete this host, you must first move all the master components listed above to another host.

OK

33. How to Turn On Maintenance Mode for a Service

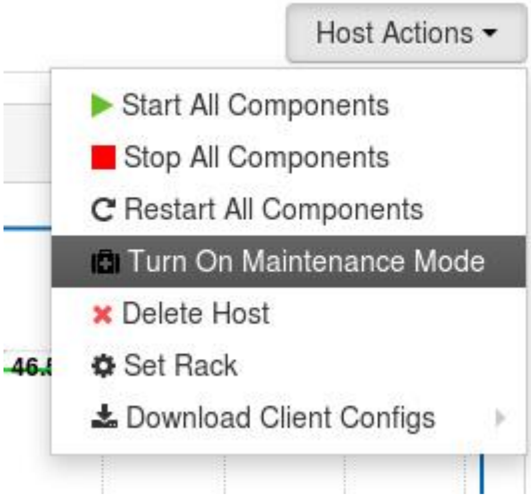
1. Using Services, select `HDFS`.
2. Select Service Actions, then choose `Turn On Maintenance Mode`.
3. Choose OK to confirm.



Notice, on Services Summary that Maintenance Mode turns on for the NameNode and SNameNode components.

34. How to Turn On Maintenance Mode for a Host

1. Using Hosts, select horton1.demo.lab
2. Select `Host Actions`, then choose `Turn On Maintenance Mode`.
3. Choose OK to confirm.



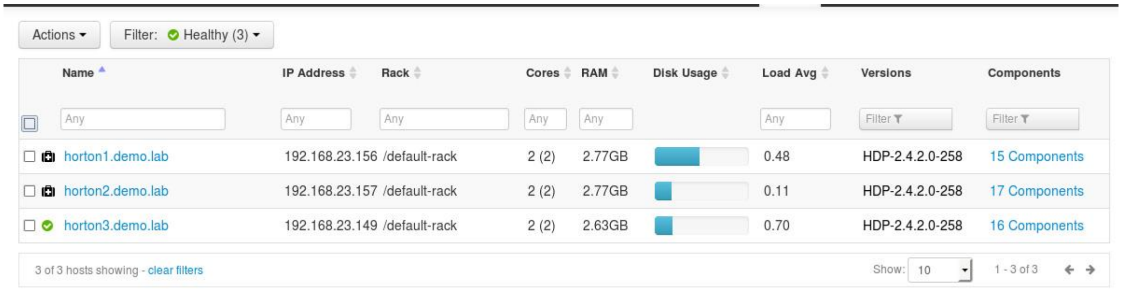
Notice on Components, that Maintenance Mode turns on for all components.

35. How to Turn On Maintenance Mode for a Host (alternative using filtering for hosts)

1. Using Hosts, select horton1.demo.lab.
2. In `Actions > Selected Hosts > Hosts` choose `Turn On Maintenance Mode`.
3. Choose `OK` to confirm.

Notice that Maintenance Mode turns on for host horton1.demo.lab.

Our list of Hosts now shows Maintenance Mode On for hosts horton1.demo.lab and horton2.demo.lab.



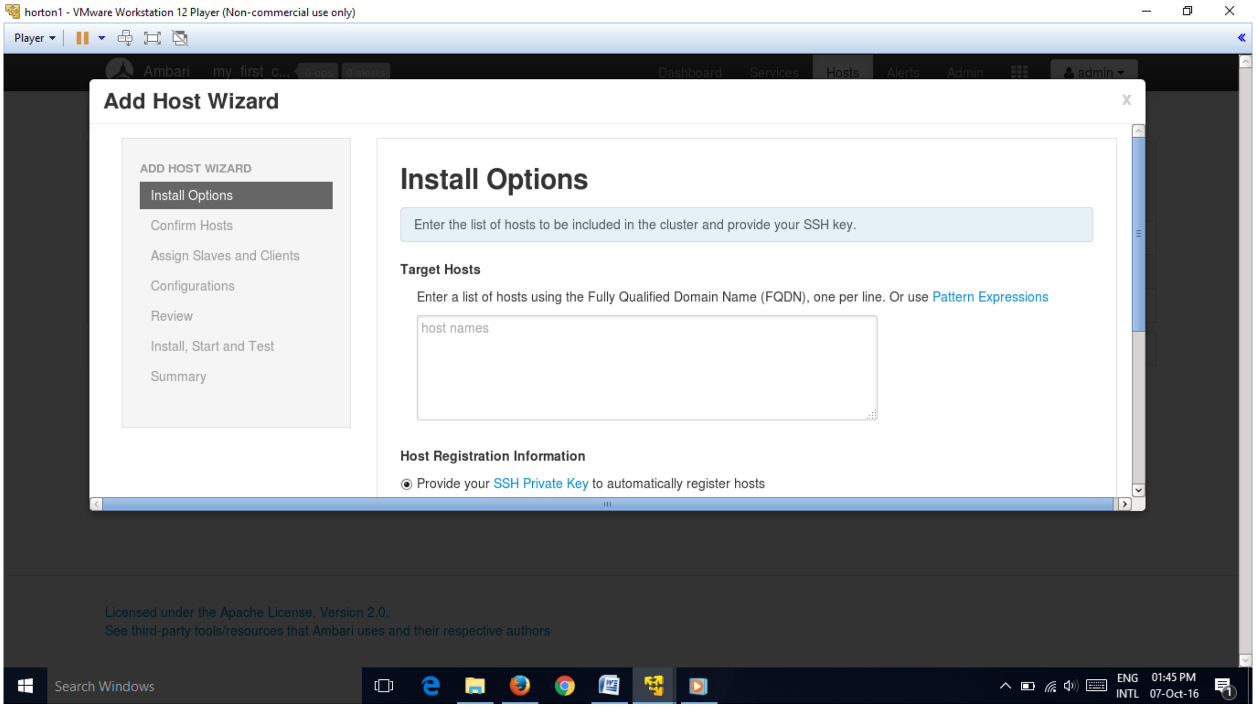
	Name	IP Address	Rack	Cores	RAM	Disk Usage	Load Avg	Versions	Components
□	Any	Any	Any	Any	Any		Any	Filter	Filter
□	horton1.demo.lab	192.168.23.156	/default-rack	2 (2)	2.77GB		0.48	HDP-2.4.2.0-258	15 Components
□	horton2.demo.lab	192.168.23.157	/default-rack	2 (2)	2.77GB		0.11	HDP-2.4.2.0-258	17 Components
□	horton3.demo.lab	192.168.23.149	/default-rack	2 (2)	2.63GB		0.70	HDP-2.4.2.0-258	16 Components

3 of 3 hosts showing - clear filters

Show: 10 1 - 3 of 3

36. Adding Hosts to a Cluster

To add new hosts to our cluster, browse to the Hosts page and select `Actions >+Add New Hosts`. The `Add Host Wizard` provides a sequence of prompts similar to those in the Ambari Install Wizard. Follow the prompts, providing information similar to that provided to define the first set of hosts in our cluster.



Add Host Wizard

ADD HOST WIZARD
- Install Options
- Confirm Hosts
- Assign Slaves and Clients
- Configurations
- Review
- Install, Start and Test
- Summary

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

host names

Host Registration Information

● Provide your SSH Private Key to automatically register hosts

Licensed under the Apache License, Version 2.0.
See third-party tools/resources that Ambari uses and their respective authors

37. Rack Awareness

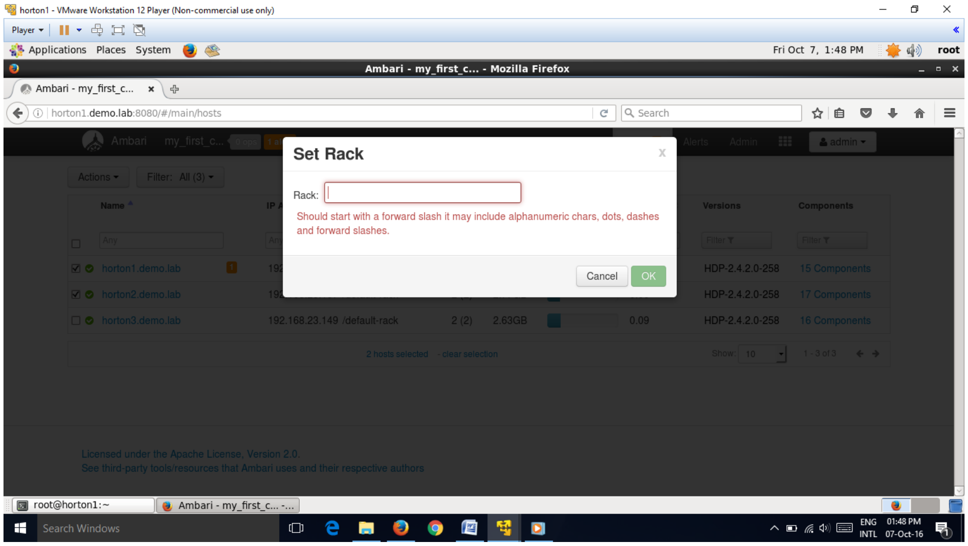
Ambari can manage Rack information for hosts. By setting the Rack ID, Ambari can display the hosts in heatmaps by Rack ID, as well users can filter & find hosts based on Rack ID on the Hosts page.

If HDFS is installed in our cluster, Ambari will pass this Rack ID information to HDFS via a topology script. Ambari generates a topology script at /etc/hadoop/conf/topology.py and sets the net.topology.script.file.name property in core-site automatically. This topology script reads a mappings file /etc/hadoop/conf/topology_mappings.data that Ambari automatically generates. When we make changes to Rack ID assignment in Ambari, this mappings file will be updated when we push out the HDFS configuration. HDFS uses this topology script to obtain Rack information about the DataNode hosts.

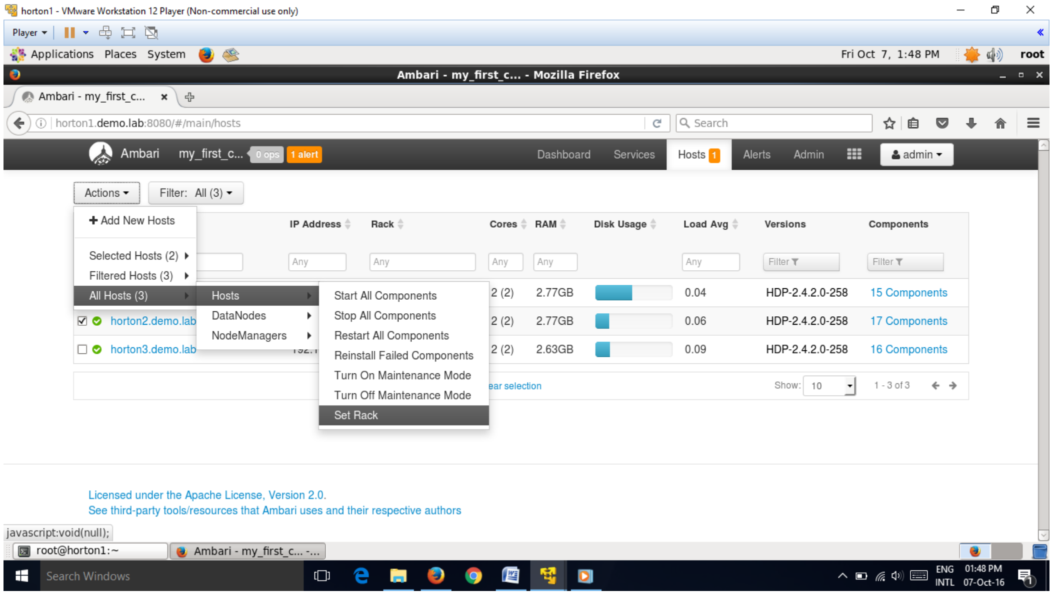
Setting Rack ID

There are two methods in Ambari Web for setting the Rack ID. We can set the Rack ID for hosts in bulk on the Hosts page using the Actions menu; and we can set the Rack ID on an individual host by viewing the Host page using the Host Actions menu.

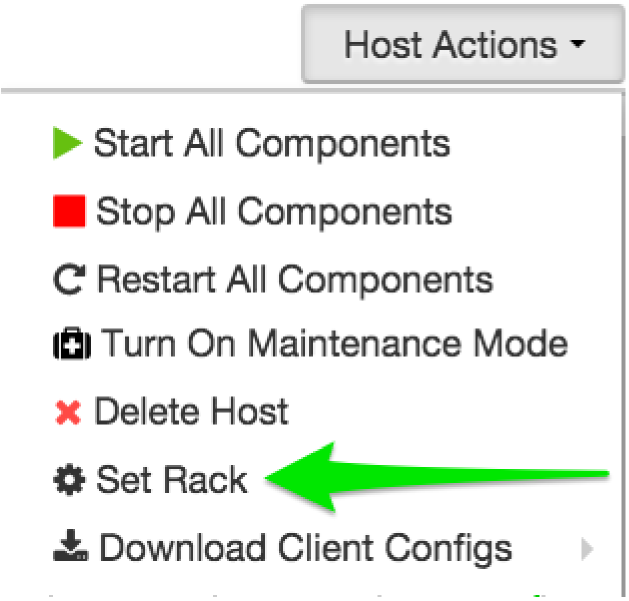
To set the Rack ID in bulk on the Hosts page, use the Actions menu and select Hosts > Set Rack (for All, Filtered or Selected hosts).


horton1 - VMware Workstation 12 Player (Non-commercial use only)
Ambari - my_first_c... - Mozilla Firefox
horton1.demo.lab:8080/#/main/hosts
Set Rack
Rack:
Cancel
OK

Should start with a forward slash it may include alphanumeric chars, dots, dashes and forward slashes.



To set the Rack ID on an individual host, browse to the Host page, use the Host Actions menu and select Set Rack.



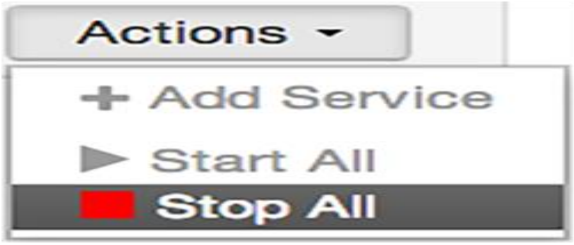
Using a Custom Topology Script

It is possible to not have Ambari manage the Rack information for hosts. Instead, we can use a custom topology script to provide rack information to HDFS and not use the Amabri-generated topology.py script. If we choose to manage Rack information on our own, we will need to **create our own topology script and manage distributing the script to all hosts**. Ambari will also not have any knowledge of host Rack information so heatmaps will not display by Rack in Ambari Web.

To manage Rack information on our own, in the `Services > HDFS > Configs`, modify the `net.topology.script.file.name` property. Set this property value to our own custom topology script (for example `/etc/hadoop/conf/topology.sh`). Distribute that topology script to our hosts and manage the Rack mapping information for our script outside of Ambari.

38. Starting and Stopping All Services

To start or stop all listed services at once, select `Actions`, then choose `Start All` or `Stop All`, as shown in the following example:

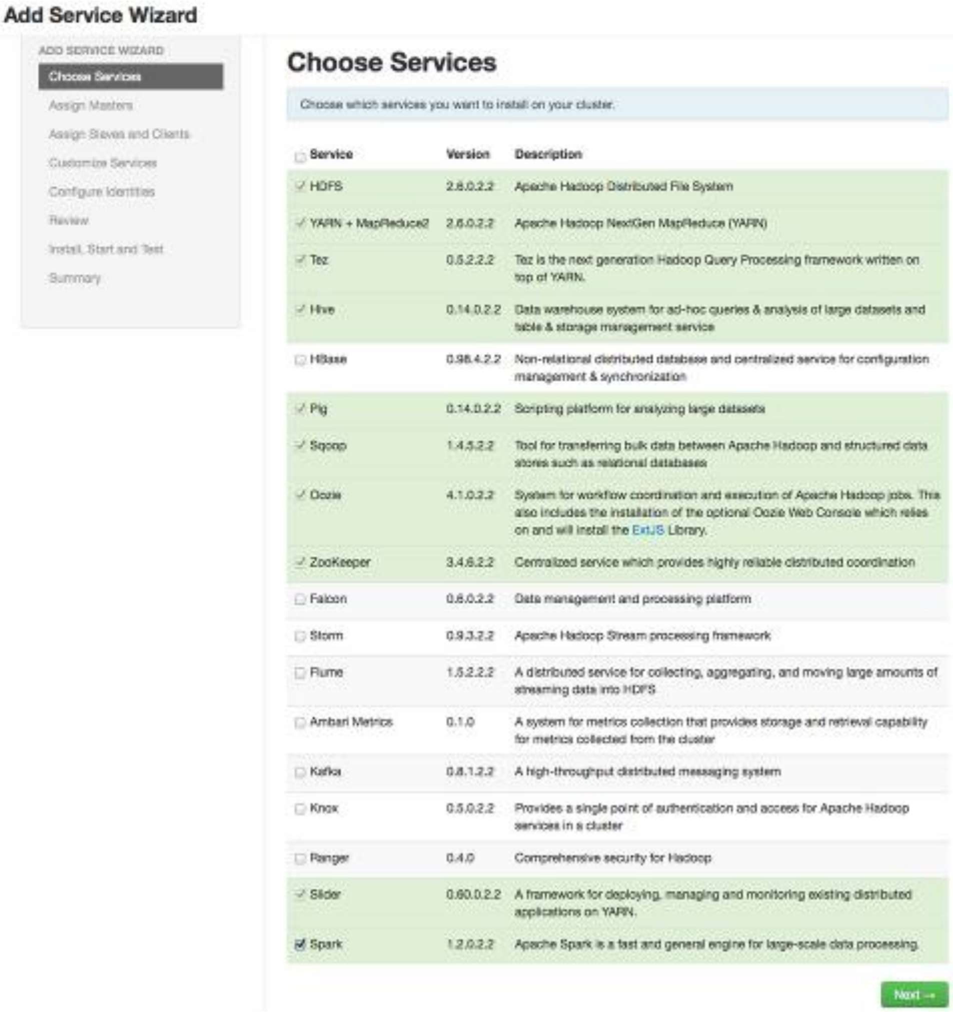


39. Adding a Service to our Hadoop cluster

This example shows the Falcon service selected for addition.

1. Choose `Services`.

   Choose an available service. Alternatively, choose all to add all available services to our cluster. Then, choose Next. The Add Service wizard displays installed services highlighted green and check-marked, not available for selection.



Add Service Wizard

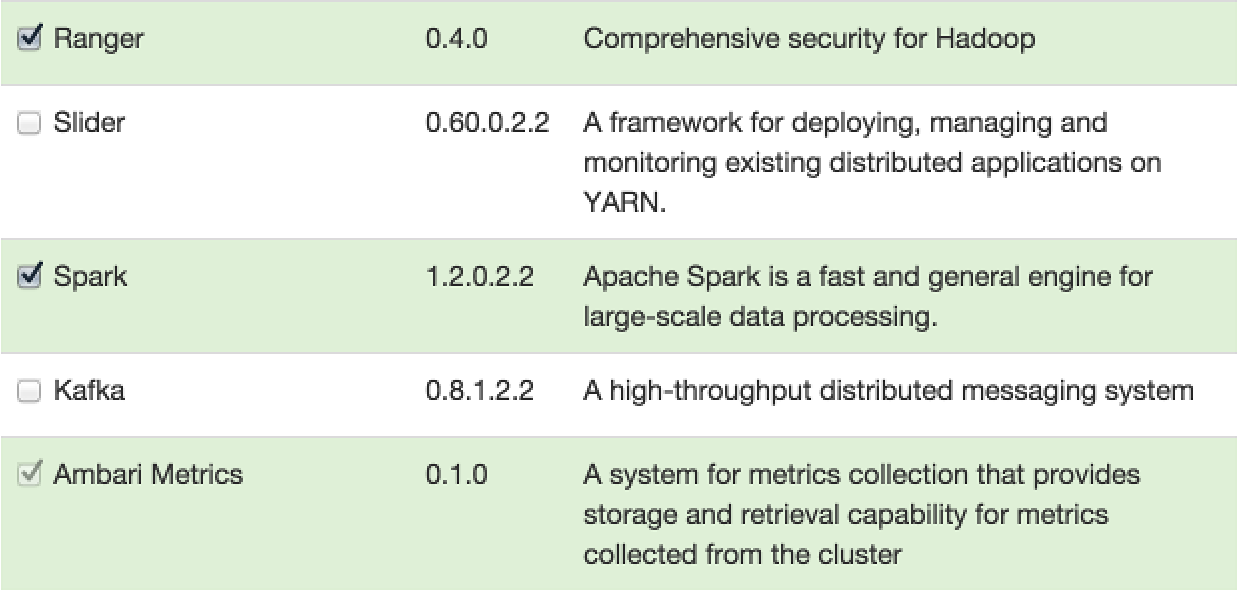
Choose Services

Choose which services you want to install on your cluster.

□ Service	Version	Description
✓ HDFS	2.6.0.2.2	Apache Hadoop Distributed File System
✓ YARN + MapReduce2	2.6.0.2.2	Apache Hadoop NextGen MapReduce (YARN)
✓ Tez	0.5.2.2.2	Tez is the next generation Hadoop Query Processing framework written on top of YARN.
✓ Hive	0.14.0.2.2	Data warehouse system for ad-hoc queries & analysis of large datasets and table & storage management service
□ HBase	0.98.4.2.2	Non-relational distributed database and centralized service for configuration management & synchronization
✓ Pig	0.14.0.2.2	Scripting platform for analyzing large datasets
✓ Sqoop	1.4.5.2.2	Tool for transferring bulk data between Apache Hadoop and structured data stores such as relational databases
✓ Oozie	4.1.0.2.2	System for workflow coordination and execution of Apache Hadoop jobs. This also includes the installation of the optional Oozie Web Console which relies on and will install the ExtJS Library.
✓ ZooKeeper	3.4.6.2.2	Centralized service which provides highly reliable distributed coordination
□ Falcon	0.6.0.2.2	Data management and processing platform
□ Storm	0.9.3.2.2	Apache Hadoop Stream processing framework
□ Flume	1.5.2.2.2	A distributed service for collecting, aggregating, and moving large amounts of streaming data into HDFS
□ Ambari Metrics	0.1.0	A system for metrics collection that provides storage and retrieval capability for metrics collected from the cluster
□ Kafka	0.8.1.2.2	A high-throughput distributed messaging system
□ Knox	0.5.0.2.2	Provides a single point of authentication and access for Apache Hadoop services in a cluster
□ Ranger	0.4.0	Comprehensive security for Hadoop
✓ Slider	0.60.0.2.2	A framework for deploying, managing and monitoring existing distributed applications on YARN.
☑ Spark	1.2.0.2.2	Apache Spark is a fast and general engine for large-scale data processing.

Note

Ambari 2.0 supports adding Ranger and Spark services, using the Add Services Wizard.



☑ Ranger	0.4.0	Comprehensive security for Hadoop
□ Slider	0.60.0.2.2	A framework for deploying, managing and monitoring existing distributed applications on YARN.
☑ Spark	1.2.0.2.2	Apache Spark is a fast and general engine for large-scale data processing.
□ Kafka	0.8.1.2.2	A high-throughput distributed messaging system
✓ Ambari Metrics	0.1.0	A system for metrics collection that provides storage and retrieval capability for metrics collected from the cluster

2. In `Assign Masters`, confirm the default host assignment. Alternatively, choose a different host machine to which master components for our selected service will be added. Then, choose Next.

   The Add Services Wizard indicates hosts on which the master components for a chosen service will be installed. A service chosen for addition shows a grey check mark.

   Using the drop-down, choose an alternate host name, if necessary.

   - A green label located on the host to which its master components will be added, or
   - An active drop-down list on which available host names appear.

3. In `Assign Slaves and Clients`, accept the default assignment of slave and client components to hosts. Then, choose Next.

   Alternatively, select hosts on which we want to install slave and client components. We must select at least one host for the slave of each service being added.

   Host Roles Required for Added Services

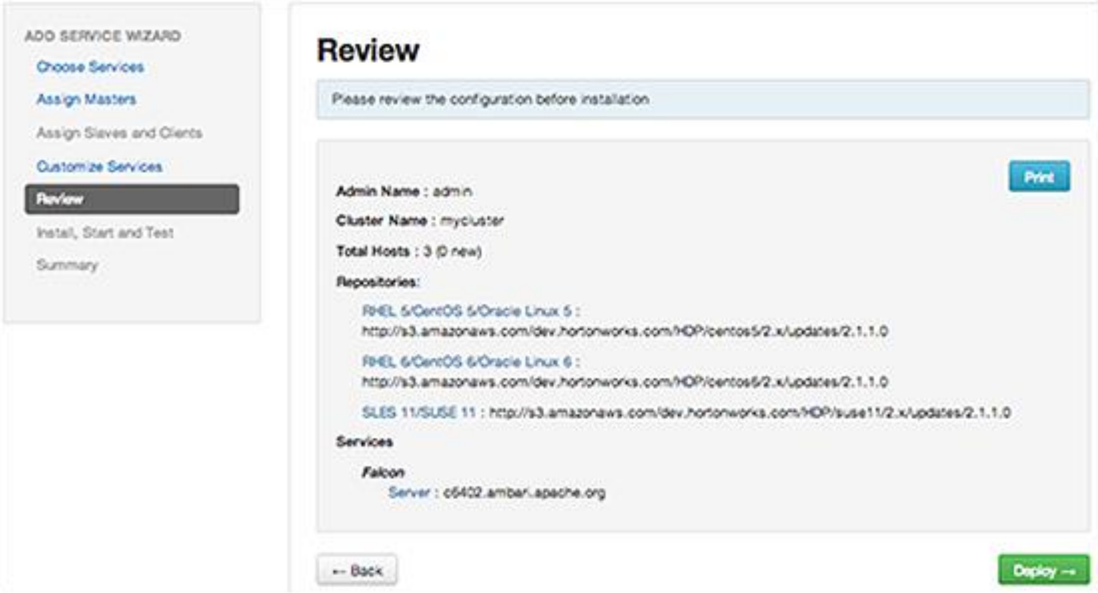
Service Added	Host Role Required
YARN	NodeManager
HBase	RegionServer

The Add Service Wizard skips and disables the Assign Slaves and Clients step for a service requiring no slave nor client assignment.

4. In `Customize Services`, accept the default configuration properties.

   Alternatively, edit the default values for configuration properties, if necessary. Choose Override to create a configuration group for this service. Then, choose Next.

5. In Review, make sure the configuration settings match our intentions. Then, choose Deploy.



Review

Please review the configuration before installation

Admin Name : admin

Cluster Name : mycluster

Total Hosts : 3 (0 new)

Repositories:

RHEL 5/CentOS 5/Oracle Linux 5 :
http://s3.amazonaws.com/dev.hortonworks.com/HDP/centos5/2.x/updates/2.1.1.0

RHEL 6/CentOS 6/Oracle Linux 6 :
http://s3.amazonaws.com/dev.hortonworks.com/HDP/centos6/2.x/updates/2.1.1.0

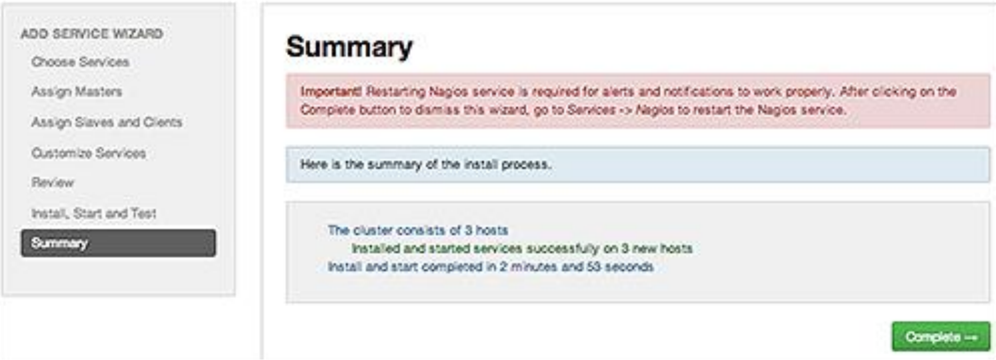
SLES 11/SUSE 11 : http://s3.amazonaws.com/dev.hortonworks.com/HDP/suse11/2.x/updates/2.1.1.0

Services

Falcon
Server : c6402.ambari.apache.org

6. Monitor the progress of installing, starting, and testing the service. When the service installs and starts successfully, choose Next.

7. Summary displays the results of installing the service. Choose Complete.



Summary

Important! Restarting Nagios service is required for alerts and notifications to work properly. After clicking on the Complete button to dismiss this wizard, go to Services -> Nagios to restart the Nagios service.

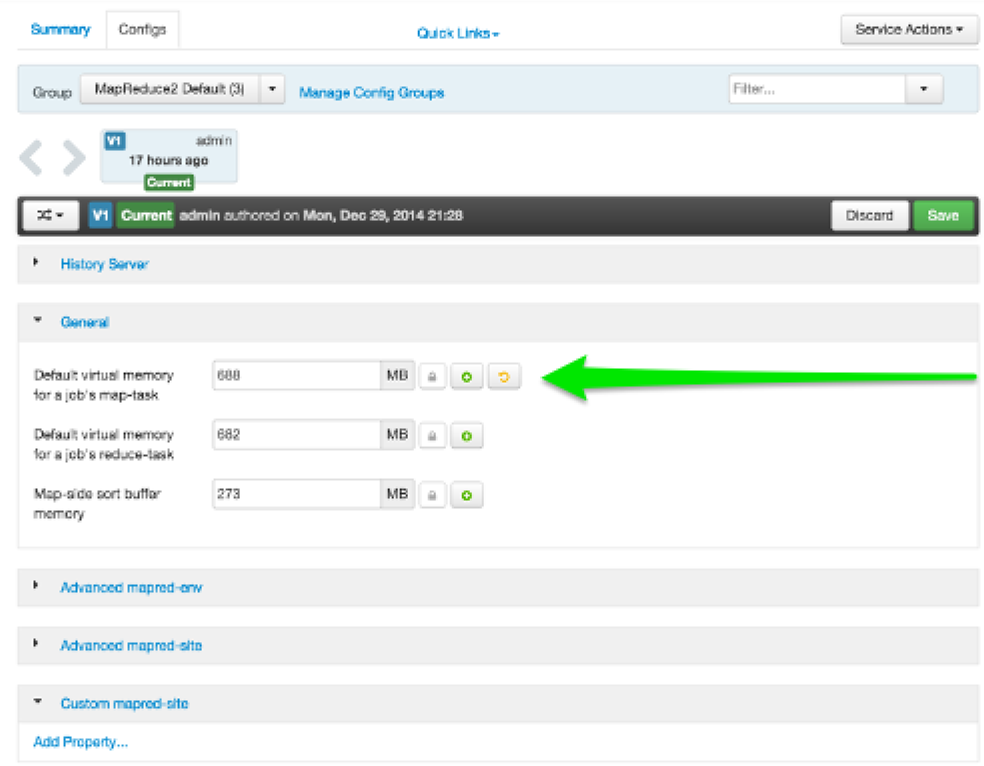
Here is the summary of the install process.

The cluster consists of 3 hosts
Installed and started services successfully on 3 new hosts
Install and start completed in 2 minutes and 53 seconds

8. Restart any other components having stale configurations.

40. Editing Service Config Properties

Select a service, then select `Configs` to view and update configuration properties for the selected service. For example, select MapReduce2, then select Configs. Expand a config category to view configurable service properties. For example, select General to configure Default virtual memory for a job's map task.



Group: MapReduce2 Default (3)

V1 admin 17 hours ago Current

V1 Current admin authored on Mon, Dec 29, 2014 21:28

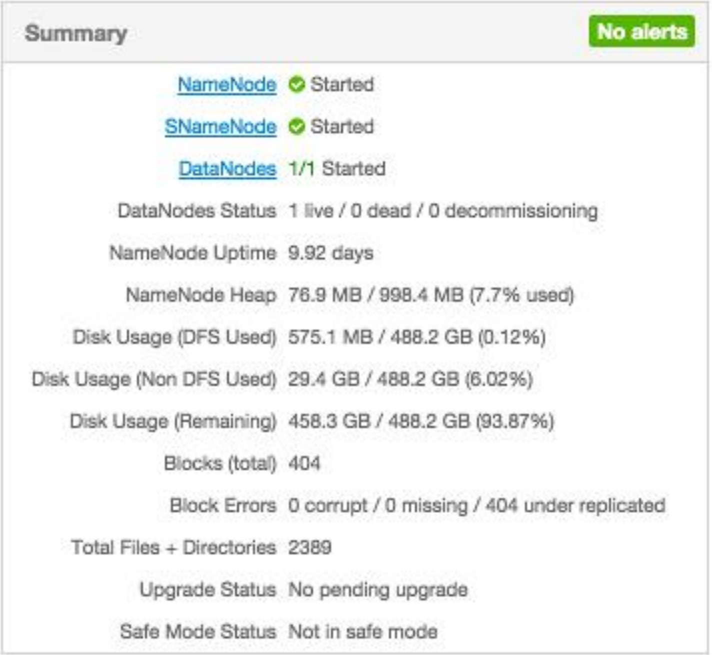
Default virtual memory for a job's map-task: 688 MB

Default virtual memory for a job's reduce-task: 682 MB

Map-side sort buffer memory: 273 MB

41. Viewing Service Summary and Alerts

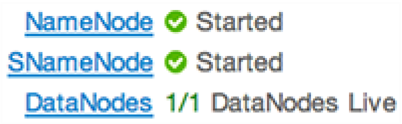
After we select a service, the `Summary` tab displays basic information about the selected service.



Summary — No alerts

NameNode	✓ Started
SNameNode	✓ Started
DataNodes	1/1 Started
DataNodes Status	1 live / 0 dead / 0 decommissioning
NameNode Uptime	9.92 days
NameNode Heap	76.9 MB / 998.4 MB (7.7% used)
Disk Usage (DFS Used)	575.1 MB / 488.2 GB (0.12%)
Disk Usage (Non DFS Used)	29.4 GB / 488.2 GB (6.02%)
Disk Usage (Remaining)	458.3 GB / 488.2 GB (93.87%)
Blocks (total)	404
Block Errors	0 corrupt / 0 missing / 404 under replicated
Total Files + Directories	2389
Upgrade Status	No pending upgrade
Safe Mode Status	Not in safe mode

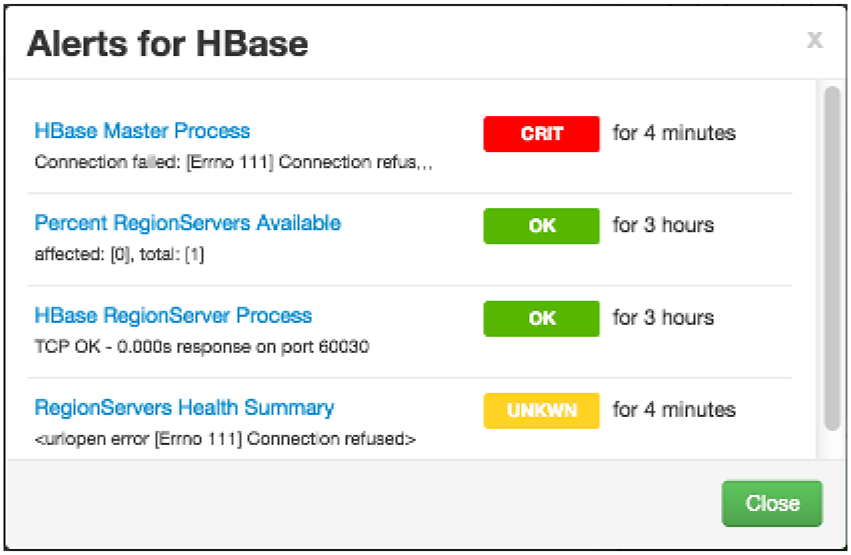
Select one of the `View Host` links, as shown in the following example, to view components and the host on which the selected service is running.



NameNode	✓ Started
SNameNode	✓ Started
DataNodes	1/1 DataNodes Live

42. Alerts and Health Checks

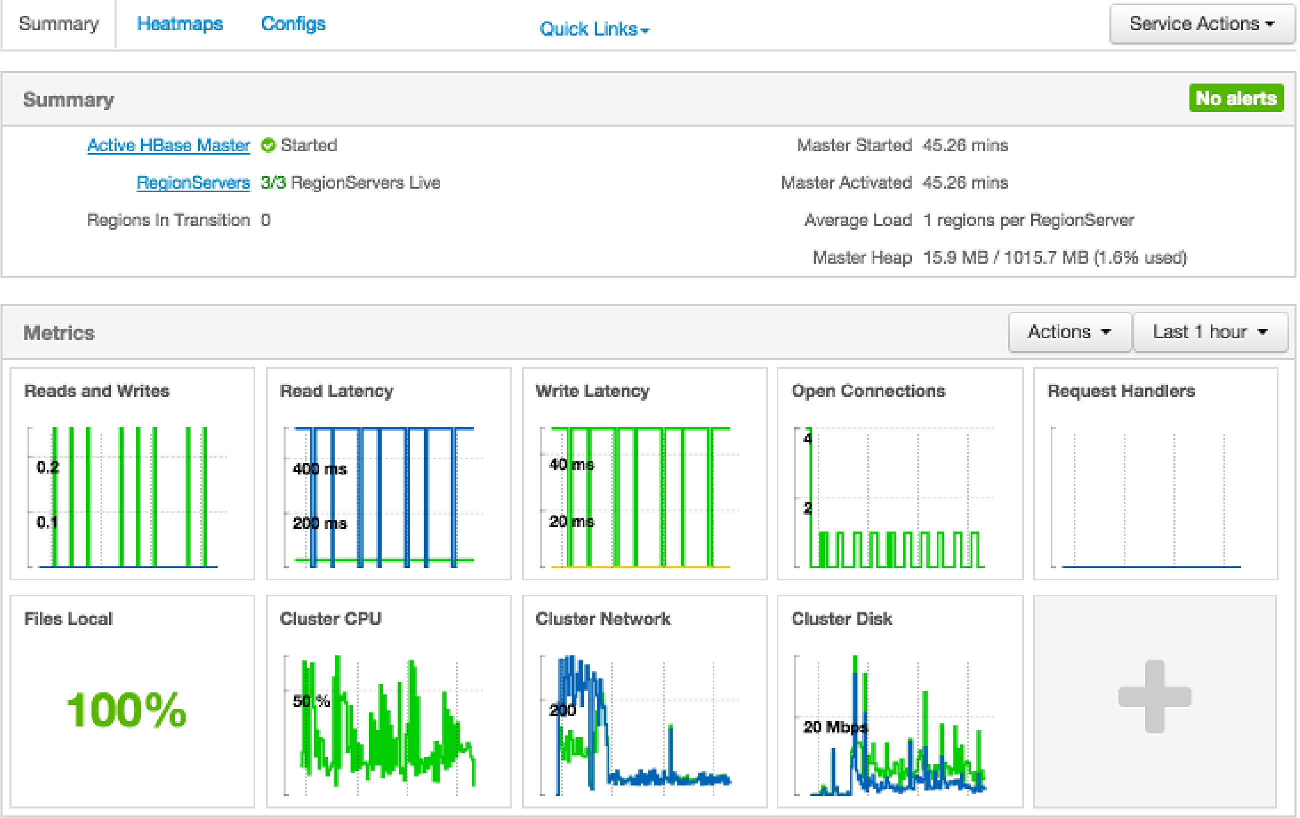
On each Service page, in the Summary area, click `Alerts` to see a list of all health checks and their status for the selected service. Critical alerts are shown first. Click the text title of each alert message in the list to see the alert definition. For example, On the HBase > Services, click Alerts. Then, in Alerts for HBase, click HBase Master Process.



HBase Master Process Connection failed: [Errno 111] Connection refus,,,	CRIT	for 4 minutes
Percent RegionServers Available affected: [0], total: [1]	OK	for 3 hours
HBase RegionServer Process TCP OK - 0.000s response on port 60030	OK	for 3 hours
RegionServers Health Summary <urlopen error [Errno 111] Connection refused>	UNKWN	for 4 minutes

43. Modifying the Service Dashboard

Depending on the Service, the Summary tab includes a Metrics section which is by default populated with important service metrics to monitor.



Summary

No alerts

Active HBase Master	Started	Master Started	45.26 mins
RegionServers	3/3 RegionServers Live	Master Activated	45.26 mins
Regions In Transition	0	Average Load	1 regions per RegionServer
		Master Heap	15.9 MB / 1015.7 MB (1.6% used)

This section of Metrics is customizable. We can add and remove widgets from the Dashboard as well as create new widgets. Widgets can be **private** only to we and our dashboard or **shared** in a Widget Browser library for other Ambari users to add/remove the widget from their Dashboard.

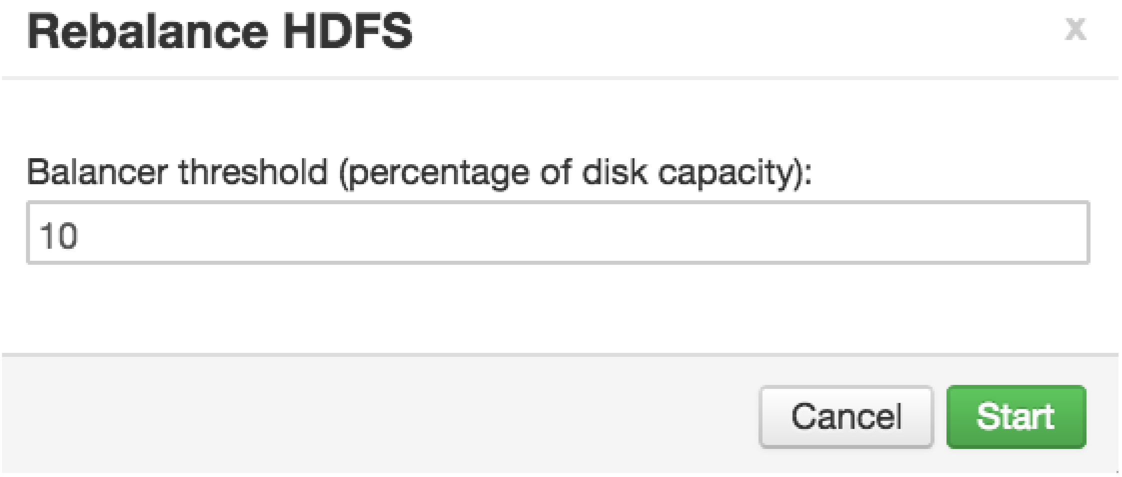
Important

We must have the Ambari Metrics service installed to be able to view, create, and customize the Service Dashboard. Only HDFS, Hive, HBase, and YARN have customizable service dashboards.

44. How to rebalance HDFS

This topic describes how we can initiate an HDFS rebalance from Ambari.

1. . In Ambari Web, browse to Services > HDFS > Summary.
2. Select Service Actions, then choose Rebalance HDFS.
3. Enter the Balance Threshold value as a percentage of disk capacity.



Rebalance HDFS

Balancer threshold (percentage of disk capacity): 10

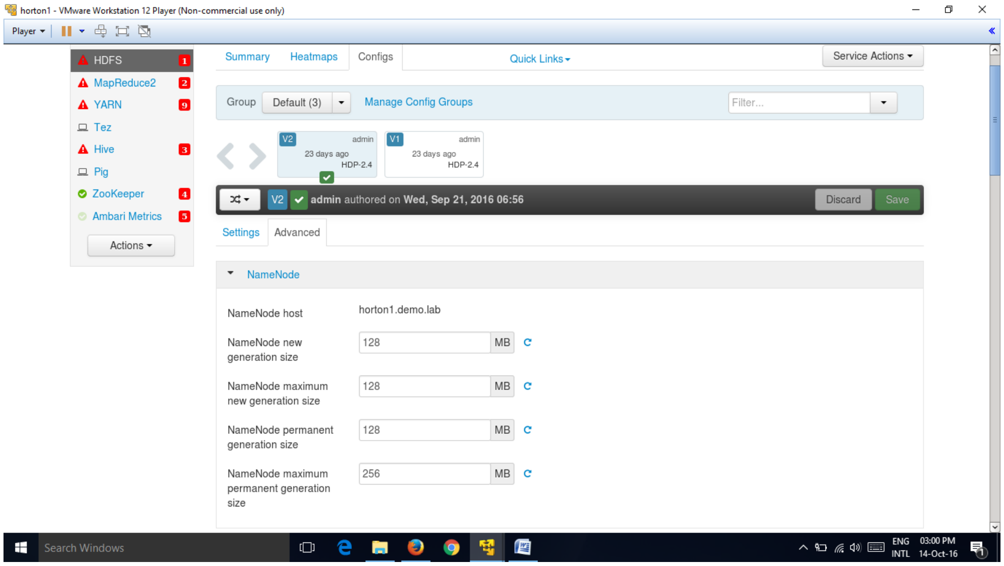
4. Click `Start` to begin the rebalance.
5. We can check rebalance progress or cancel a rebalance in process by opening the Background Operations dialog.

45. Updating Service Properties

1. Expand a configuration category.
2. Edit values for one or more properties that have the Override option.

   Edited values, also called stale configs, show an Undo option.

3. Choose Save.



V2 admin authored on Wed, Sep 21, 2016 06:56

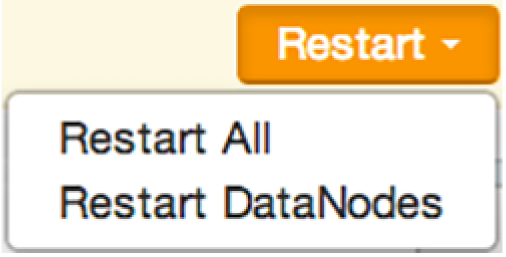
NameNode host	horton1.demo.lab
NameNode new generation size	128 MB
NameNode maximum new generation size	128 MB
NameNode permanent generation size	128 MB
NameNode maximum permanent generation size	256 MB

46. Restarting Components

After editing and saving a service configuration, Restart indicates components that we must restart.

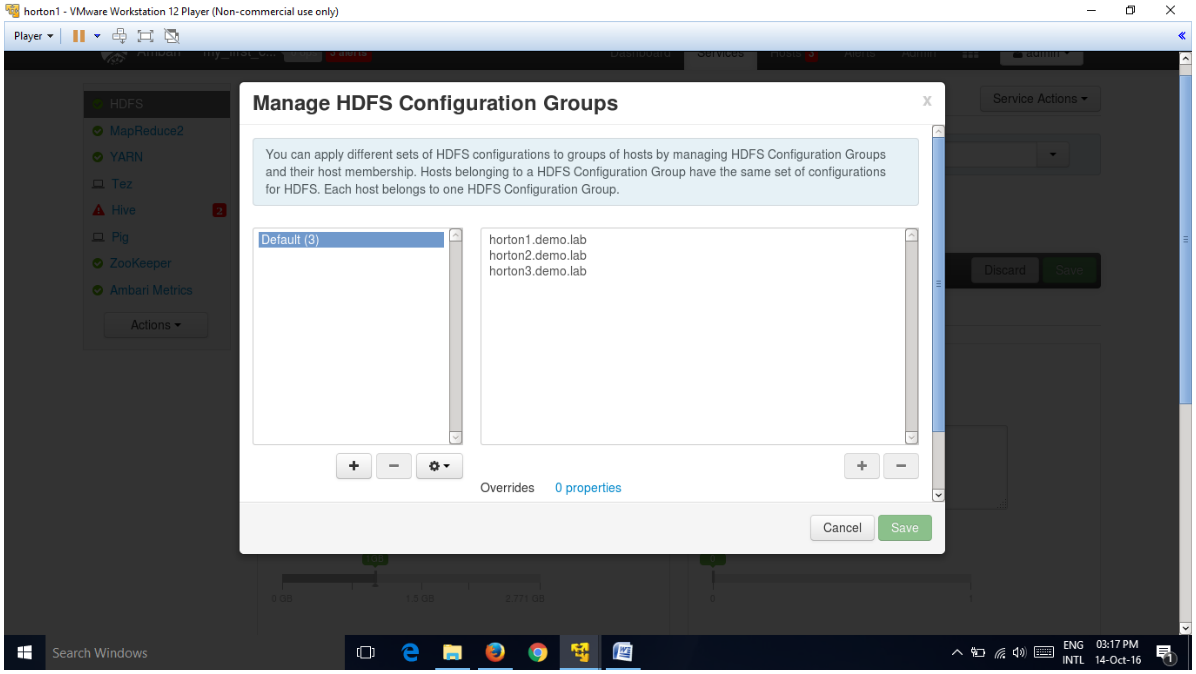
Select the Components or Hosts links to view details about components or hosts requiring a restart.

Then, choose an option appearing in Restart. For example, options to restart YARN components include:



47. Using Host Config Groups

Ambari initially assigns all hosts in our cluster to one, default configuration group for each service we install. For example, after deploying a three-node cluster with default configuration settings, each host belongs to one configuration group that has default configuration settings for the HDFS service. In Configs, select Manage Config Groups, to create new groups, re-assign hosts, and override default settings for host components we assign to each group.



Manage HDFS Configuration Groups

You can apply different sets of HDFS configurations to groups of hosts by managing HDFS Configuration Groups and their host membership. Hosts belonging to a HDFS Configuration Group have the same set of configurations for HDFS. Each host belongs to one HDFS Configuration Group.

Default (3)

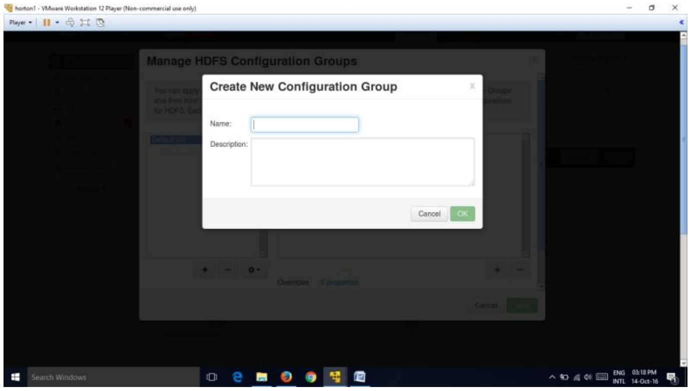
horton1.demo.lab
horton2.demo.lab
horton3.demo.lab

Overrides 0 properties

Cancel Save

To create a Configuration Group:

1. Choose `Add New Configuration Group`.
2. Name and describe the group, then choose Save.



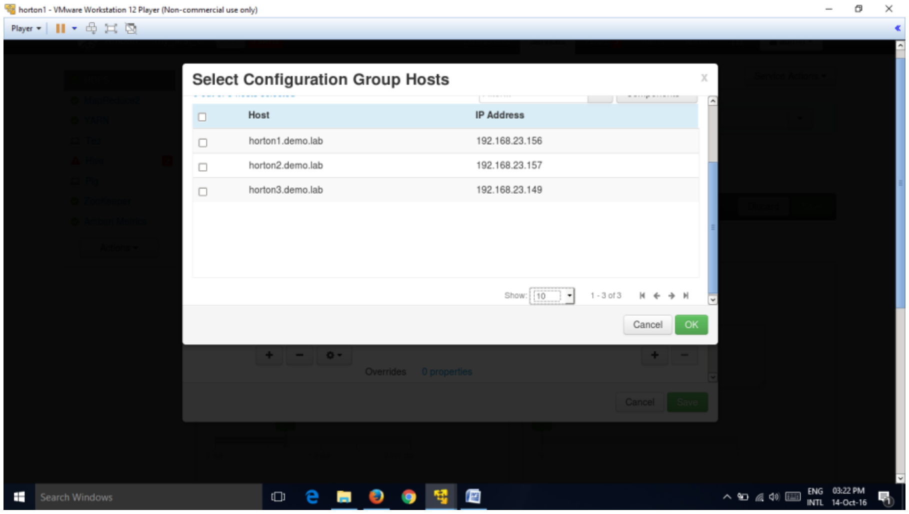
Create New Configuration Group

Name:

Description:

horton1 - VMware Workstation 12 Player (Non-commercial use only)
Manage HDFS Configuration Groups
Cancel
OK
Search Windows
03:18 PM
14-Oct-16

3. Select a Config Group, then choose Add Hosts to Config Group.
4. Select Components and choose from available Hosts to add hosts to the new group.

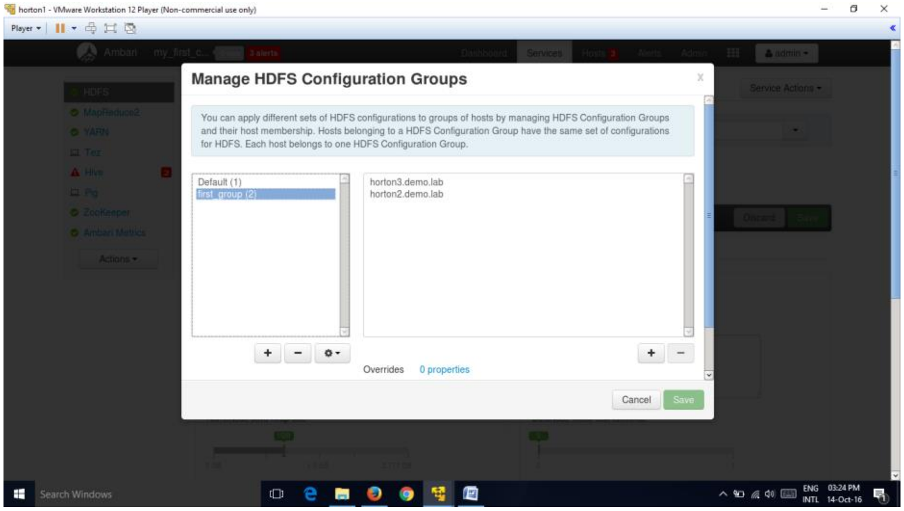


Select Configuration Group Hosts

□	Host	IP Address
□	horton1.demo.lab	192.168.23.156
□	horton2.demo.lab	192.168.23.157
□	horton3.demo.lab	192.168.23.149

Show: 10 1 - 3 of 3

Select Configuration Group Hosts enforces host membership in each group, based on installed components for the selected service.



Manage HDFS Configuration Groups

You can apply different sets of HDFS configurations to groups of hosts by managing HDFS Configuration Groups and their host membership. Hosts belonging to a HDFS Configuration Group have the same set of configurations for HDFS. Each host belongs to one HDFS Configuration Group.

Default (1)
first_group (2)

horton3.demo.lab
horton2.demo.lab

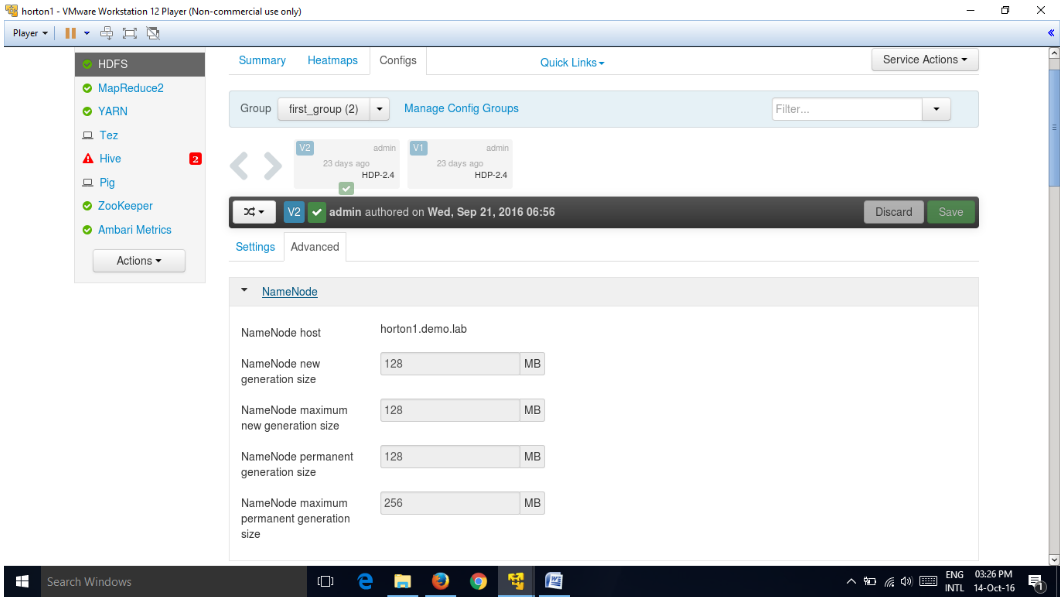
Overrides 0 properties

5. Choose OK.
6. In Manage Configuration Groups, choose Save.

To edit settings for a configuration group:

1. In Configs, choose a Group.
2. Select a Config Group, then expand components to expose settings that allow Override.
3. Provide a non-default value, then choose Override or Save.

   Configuration groups enforce configuration properties that allow override, based on installed components for the selected service and group.



Group first_group (2)

V2 ✓ admin authored on Wed, Sep 21, 2016 06:56

NameNode host: horton1.demo.lab

NameNode new generation size: 128 MB

NameNode maximum new generation size: 128 MB

NameNode permanent generation size: 128 MB

NameNode maximum permanent generation size: 256 MB

4. Override prompts we to choose one of the following options:

   a. Select an existing configuration group (to which the property value override provided in step 3 will apply), or
   b. Create a new configuration group (which will include default properties, plus the property override provided in step 3).
   c. Then, choose `OK`.

5. In Configs, choose Save.