Apache Ambari

-------------------------------------------------------------------------------------------------------------------------------

Apache Ambari is defined as a software project which is deployed on top of the [Hadoop cluster](https://hadoop.apache.org/docs/current/hadoop-project-dist/hadoop-common/ClusterSetup.html). It is responsible for keeping track of running applications and their status. Moreover, it is a web-based management tool that manages, monitors, and provisions the health of Hadoop clusters.

It is very flexible and scalable user-interface, which permits a range of tools, for example**,** [**Pig**](https://www.tutorialandexample.com/apache-pig-tutorial/)**, MapReduce,** [**Hive**](https://www.tutorialandexample.com/hive-tutorial/), etc.

Following are some profiles for which Apache Ambari should be used:

* Hadoop administrators.
* Database professionals.
* Mainframe and Hadoop testing professionals.
* DevOps Professionals.

**Why Use Ambari?**

* It is being used for targeting the management of Hadoop for the developers and administrators.
* It uses web-based APIs for Hadoop Management.
* It provides career opportunities to the students.

### **Ambari Architecture**

The architecture of Ambari is defined as a framework which shows the interaction between different components and APIs to make a connection with it.

It works to automate the operations in the Hadoop cluster, and it also provides a secure interface for the users.

Ambari Architecture has two major components :

* Ambari Server
* Ambari Agent

**Ambari Server**

It provides communication between the agents and installed on each node of the cluster.

**Ambari Agent**

The participants provide the health status of every node along with various operational metrics are referred as **Ambari agents.**

### **Getting Started with Ambari**

The latest version is a 64-bit version of the following Operating Systems:

* RHEL 7.4, 7.3, 7.2
* CentOS 7.4, 7.3, 7.2
* OEL (Oracle Enterprise Linux) 7.4, 7.3, 7.2
* Amazon Linux 2
* SLES 12 SP3, 12 SP2
* Ubuntu 14 and 16
* Debian 9

**Recent Improvements:**

* Usability (provide a framework across Hortonworks products)
* Management level (handle 5,000 node clusters)
* Secure level Configuration (used for DPS services)

**Recovery in Ambari**

Recovery in Ambari can be achieved in two ways, which are given as follows:

1. **Based on actions**

Restart master is responsible for seeking pending actions, and then rescheduled them.  Besides, it defines the state of an individual process.

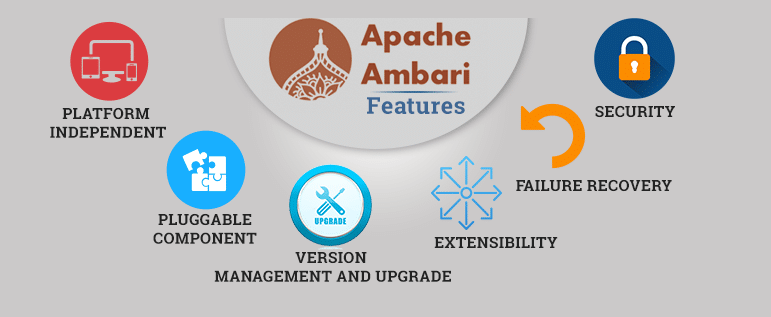
The actions should be idempotent (special consideration is taken) and those actions which doesn not show as complete, the master restarts them.

1. **Based on the desired state**

When a system restarts, the master converts clusters into the active state as per the desired state.

### **Features of Apache Ambari**

Features of Ambari are as follows:



**Platform independent:** It can be run on a different platforms without changing its configuration, i.e., Windows, Mac, Ubuntu, SLES, RHEL, etc.

**Pluggable component:** Pluggable provide an efficient way to customize current Ambari application by encapsulating specific tools and technologies.

**Version management and upgrade:** Ambari is capable of maintaining its versions. We don’t require any external tool like Git.

**Extensibility:** In Ambari, the functionality of its applications can be enhanced by using various components.

**Failure recovery:** If a system is accidentally closed due to power failure or any other internal or external factor, the data at that time can be recovered.

**Security:** It provides robust protection of an application.

### **Advantages of Apache Ambari**

Advantages of Ambari are as follows:

1. **Installation, Configuration, and Management is a way.**

Client components and Master-slaves are used for handling and configuring services.

1. **Centralized Security and Application**

Ambari handles the cluster securities and the administrative authorities by preventing the data from unauthorized persons or eavesdroppers. Furthermore, it provides an automated setup of advanced security enable like Kerberos & Ranger.

1. **Open-source**

Apache Ambari is an open source platform where a user can easily access or improve the source code.

1. **Extensible**

We can extend the functionality by using various components.

### **Apache repositories**

Apache generally uses four repositories.

* Ambari
* HDP
* HDP UTILS
* EPEL

**Ambari:** It is responsible for hosting the Ambari server and Ambari agents.

**HDP:** It is used for hosting the Hadoop Stack Packages, i.e., **Pig, Hive,** [**HBase**](https://www.tutorialandexample.com/hbase-tutorial/), **Oozie**, etc.

**HDP-UTILS:** It is defined as a utility package for enhancing the capabilities provided by the operating system.

**EPEL (Extra Packages for Enterprise Linux):** It provides a bundle of additional packages used in Enterprise Linux.

### **How Ambari Uses Repositories**

First of all, Ambari Repository tells its server about which Ambari Utility Repository is to be used, then the server is to be connected to those hosts in which HDP repository is being used.

Finally, the server tells all hosts in the cluster about what HDP repository is used.

**Ambari User Views**

A “view” is defined as an enhancement in Ambari that interacts with third parties to plug in new resource types along with the APIs.

Most importantly, the Ambari User Views contribution work actively in the community. It also helps to provide capabilities that deal with the operational tasks of application development and also manages the workload.

**Tez:** It is used for optimizing your cluster resource usage by using the views.

**Hive:** Its main task is to write SQL Queries into the cluster. Also, it shows the history of all SQL Queries.

**Pig:** It is similar to Hive View and used for writing and running Pig Script.

**Files:** It allows the user to manage, browse, and upload files and folders in **HDFS**.

### **Troubleshooting Ambari**

Troubleshooting Ambari generally helps to resolve the general problems faced while using Ambari.

So, let’s begin Troubleshooting  Ambari.

**Problem 1**

Ambari can show an error named “Server Fails to Start: No Driver.”

Check /var/log/ambari-server/ambari-server.log for :

ExceptionDescription:Configurationerror.

Class[oracle.jdbc.driver.OracleDriver] not found.

“Unable to found the Oracle JDBC .jar file.”

**Solution:** You can remove an error by only re-run **ambari-server setup,** if you are sure about that file that it is saved in the appropriate directory on the server.

**Problem 2**

Ambari can show an error named: “No Connection.”

Check /var/log/ambari-server/ambari-server.log for :

The Network Adapter could not establish the connection

Error Code: 17002

Database and the Ambari Server are not connected.

**Solution:** It should be confirmed that the database is correctly configured by reading /etc/Ambari-server/conf/ambari.properties and also encounter that the host is reachable to the Ambari Server.

**Problem 3**

Ambari can show an error named: Bad Username.

Check /var/log/ambari-server/ambari-server.log for :

Internal Exception: java.sql.SQLException: ORA¬01017:

invalid username/password; logon denied

If a username /password is invalid, then it shows an error.

**Solution:** To remove this error, you should make sure that the user account has the authority to access it.

**Problem 4**

Ambari can show an error named: “No Schema.”

Check /var/log/ambari-server/ambari-server.log for :

Internal Exception: java.sql.SQLSyntaxErrorException: ORA¬00942:

table or view does not exist

Means that the schema has not been loaded.

**Solution:** Ensure that you should have loaded the database schema.

### **What is Ambari Security?**

Security is a vital task which helps to secure user data and identifies the authority of a user by preventing the data from unauthorized person. It has many advanced security options for secure authentication, which are given below:

* Configuring Ambari and Hadoop for Kerberos.
* For LDAP or active directory authentication.
* Configuring Ambari for non-root.
* Optional: Encrypt database and LDAP passwords.
* Set up SSL for Ambari: Optional.

**Ports used by Ambari:**

**Default ports:**

* **8080**– Ambari web interface.
* **8440** – for establishing a connection between Ambari agents and Ambari server.
* **8441** – used for registration and providing the interface to Ambari agents and Ambari server.

In case if it does not connect to the Ambari server, there are some essential checkpoints that you can perform:

* If a firewall occurs between the Ambari host and server, check only port 8440 and 8441.
* Check the IP cables.
* Disable SELinux on both server and clients. Also, check if any new host has connected to the server.
* Check the logs which are available at location /var/lob/ambari-agent/ambari-agent.log to see the error messages.

----------------------------------------------------------------------------------------------------------