

Zeppelin Installation for Analytics

Date of publish: 2017-11-06



Legal Notice

© Cloudera Inc. 2019. All rights reserved.

The documentation is and contains Cloudera proprietary information protected by copyright and other intellectual property rights. No license under copyright or any other intellectual property right is granted herein.

Copyright information for Cloudera software may be found within the documentation accompanying each component in a particular release.

Cloudera software includes software from various open source or other third party projects, and may be released under the Apache Software License 2.0 ("ASLv2"), the Affero General Public License version 3 (AGPLv3), or other license terms. Other software included may be released under the terms of alternative open source licenses. Please review the license and notice files accompanying the software for additional licensing information.

Please visit the Cloudera software product page for more information on Cloudera software. For more information on Cloudera support services, please visit either the Support or Sales page. Feel free to contact us directly to discuss your specific needs.

Cloudera reserves the right to change any products at any time, and without notice. Cloudera assumes no responsibility nor liability arising from the use of products, except as expressly agreed to in writing by Cloudera.

Cloudera, Cloudera Altus, HUE, Impala, Cloudera Impala, and other Cloudera marks are registered or unregistered trademarks in the United States and other countries. All other trademarks are the property of their respective owners.

Disclaimer: EXCEPT AS EXPRESSLY PROVIDED IN A WRITTEN AGREEMENT WITH CLOUDERA, CLOUDERA DOES NOT MAKE NOR GIVE ANY REPRESENTATION, WARRANTY, NOR COVENANT OF ANY KIND, WHETHER EXPRESS OR IMPLIED, IN CONNECTION WITH CLOUDERA TECHNOLOGY OR RELATED SUPPORT PROVIDED IN CONNECTION THEREWITH. CLOUDERA DOES NOT WARRANT THAT CLOUDERA PRODUCTS NOR SOFTWARE WILL OPERATE UNINTERRUPTED NOR THAT IT WILL BE FREE FROM DEFECTS NOR ERRORS, THAT IT WILL PROTECT YOUR DATA FROM LOSS, CORRUPTION NOR UNAVAILABILITY, NOR THAT IT WILL MEET ALL OF CUSTOMER'S BUSINESS REQUIREMENTS. WITHOUT LIMITING THE FOREGOING, AND TO THE MAXIMUM EXTENT PERMITTED BY APPLICABLE LAW, CLOUDERA EXPRESSLY DISCLAIMS ANY AND ALL IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO IMPLIED WARRANTIES OF MERCHANTABILITY, QUALITY, NON-INFRINGEMENT, TITLE, AND FITNESS FOR A PARTICULAR PURPOSE AND ANY REPRESENTATION, WARRANTY, OR COVENANT BASED ON COURSE OF DEALING OR USAGE IN TRADE.

Contents

- Introduction to Using Zeppelin With CCP.....4**
- Installing Zeppelin..... 4**
 - Install Apache Zeppelin Using Ambari.....4
 - Import the Apache Zeppelin Notebook Manually.....5

Introduction to Using Zeppelin With CCP

The Zeppelin dashboard is intended for use by Security Operations Center (SOC) analysts and investigators.

Like the Metron dashboard, the Zeppelin dashboard can be used to view and analyze the enriched telemetry data provided by CCP. However Zeppelin can be used by a data scientist to create runbooks for recreatable investigations. These runbooks can be static, which require no input, or dynamic, which require you to enter or choose information.

Installing Zeppelin

You can install Zeppelin either using Ambari or manually.

Install Apache Zeppelin Using Ambari

How to install Apache Zeppelin on an Ambari-managed cluster.

Before you begin

Install Zeppelin on a node where Spark clients are already installed and running. This typically means that Zeppelin will be installed on a gateway or edge node.

Zeppelin requires the following software versions:

- HDP 3.0 or later.
- Apache Spark 2.0.
- Java 8 on the node where Zeppelin is installed.

The optional Livy server provides security features and user impersonation support for Zeppelin users. Livy is installed as part of Spark.

- After installing Spark, Livy, and Zeppelin, refer to "Configuring Zeppelin" in this guide for post-installation steps.

Install Zeppelin Using Ambari

The Ambari installation wizard sets default values for Zeppelin configuration settings. Initially, you should accept the default settings. Later, when you are more familiar with Zeppelin, consider customizing the Zeppelin configuration settings.

To install Zeppelin using Ambari, add the Zeppelin service:

1. Click the ellipsis (...) symbol next to Services on the Ambari dashboard, then click Add Service.
2. On the Add Service Wizard under Choose Services, select Zeppelin Notebook, then click Next.
3. On the Assign Masters page, review the node assignment for Zeppelin Notebook, then click Next.
4. On the Customize Services page, review the default values, then click Next.
5. If Kerberos is enabled on the cluster, review the principal and keytab settings on the Configure Identities page, modify the settings if desired, then click Next.
6. Review the configuration on the Review page, then click Deploy to begin the installation.
7. The Install, Start, and Test page displays the installation status.
8. When the progress bar reaches 100% and a "Success" message appears, click Next.
9. On the Summary page, click Complete to finish installing Zeppelin.

To validate the Zeppelin installation, open the Zeppelin Web UI in a browser window. Use the port number configured for Zeppelin (9995 by default); for example:

```
http://<zeppelin-host>:9995
```

You can also open the Zeppelin Web UI by selecting Zeppelin Notebook > Zeppelin UI on the Ambari dashboard.

To check the Zeppelin version number, type the following command on the command line:

```
/usr/hdp/current/zeppelin-server/bin/zeppelin-daemon.sh --version
```

Zeppelin stores configuration settings in the `/etc/zeppelin/conf` directory. Note, however, that if your cluster is managed by Ambari you should not modify configuration settings directly. Instead, use the Ambari web UI.

Zeppelin stores log files in `/var/log/zeppelin` on the node where Zeppelin is installed.

Import the Apache Zeppelin Notebook Manually

As an alternative to using Ambari to install Apache Zeppelin you can manually install the tool.

Procedure

1. Use ssh to navigate to the host where you want to install Zeppelin.

```
ssh $METRON_HOME
```

2. Use the following command to import the `$METRON_HOME/config/zeppelin/metron-yaf-telemetry.json` file onto your Zeppelin host.

```
curl -s -XPOST https://github.com/apache/incubator-metron/blob/master/metron-platform/metron-indexing/src/main/config/zeppelin/metron-yaf-telemetry.json/api/notebook/import -d @"$METRON_HOME/config/zeppelin/metron-yaf-telemetry.json"
```

3. Navigate to `http://$ZEPPELIN_HOST:9995`.