

Supported Deployment Scenarios for Cloudera

Contents:

- *Supported Cloudera Distributions*
- *Supported Deployments*
 - *Deployment System*
 - *Running Environment*
 - *Platform Security*
 - *High Availability*
 - *Metadata Publishing*
 - *Supported File Formats*
- *Connectivity*
 - *Hadoop Connectivity*
 - *External Connectivity*
- *Notes*

Supported Cloudera Distributions

NOTE: By default, Cloudera may be installed with Java JDK 1.7 or earlier. If so, you must upgrade each node in the cluster to Java JDK 1.8. For more information, see https://www.cloudera.com/documentation/enterprise/latest/topics/cdh_ig_jdk_installation.html.

For this release, the Trifacta® platform supports the following Cloudera versions.

NOTE: Cloudera 6.x and later requires use of native Hadoop libraries from the cluster.

- Cloudera 6.3.x (recommended)
- Cloudera 6.2.x
- Cloudera 6.1.x
- Cloudera 5.16.x

NOTE: Cloudera 5.14.x and 5.15.x are no longer supported. For best results, please upgrade your Hadoop distribution.

- Cloudera Data Platform 7.1

Notes:

- **Update Date:** August 25, 2020
- The Trifacta platform supports all variants of patch or point releases (X.Y.* and X.Y.*.* releases) through the Hadoop vendor's backwards compatibility policy.
- For individual versions of Hadoop components (such as HDFS, Spark, and Hive), the Trifacta platform supports the component version that is bundled with the vendor's package for the supported Hadoop distribution.
- For more information on how to set up your Hadoop distribution, please consult the documentation provided with your distribution or contact your distribution vendor.

Supported Deployments

NOTE: Unless otherwise noted, all items listed below are supported across all Hadoop distribution versions listed above. Unlisted items are not supported. Please contact *Trifacta Support* or your sales representative for items not listed here.

Deployment System

NOTE: The Trifacta platform software must be installed on a gateway node of the Cloudera cluster. For more information, see *System Requirements* in the Planning Guide.

Item	Description
Physical On Premise Machines	Supported.
VMWare / VXServer	Supported.

NOTE: Deployment to an Amazon EC2 is supported. See *Supported Deployment Scenarios for AWS*.

Running Environment

Item	Description
Spark	Supported.
Trifacta Photon	Supported.

Platform Security

Item	Description
HDFS File Permissions	Supported.
HDFS Privileges	Supported through Sentry.
Hive Privileges	Supported through Sentry.
Kerberos-Enabled Hadoop Cluster	Supported.
Secure User Impersonation	Supported.

High Availability

Item	Description
Name Node, Resource Manager, HttpFS	Supported.

Metadata Publishing

Item	Description
Cloudera Navigator	Not supported.

Hive Publishing	Supported.
Redshift Publishing	Supported.

Supported File Formats

See *Supported File Formats* in the User Guide.

Connectivity

Hadoop Connectivity

The Trifacta platform supports connectivity for execution to the following Hadoop environments for this vendor's distributions. Connectivity exceptions are listed below:

Running Environment	HDFS Reader	HDFS Writer	Hive Reader w/ HiveServer2
Spark	Supported.	Supported.	Supported.

Profiling Environment	HDFS Reader	HDFS Writer
Profiling on Spark	Supported.	Supported.

External Connectivity

Storage Platform	HDFS Reader	HDFS Writer
S3	Supported.	Supported.

Storage Platform	Amazon S3 Reader	Amazon S3 Writer
Spark Profiling	Supported.	Supported.

Notes

- none.