



The Power of Hadoop, Built for the Cloud

Flexible

Out-of-the-box integrations with AWS, Azure, and GCP and an open framework eliminates cloud environment lock-in and ensures seamless deployments

Simple

Single, unified interface to manage multiple clusters, across multiple environments

Transient cluster support

Amazon S3 object store integration for several CDH components including Impala, Hive, and Spark

Reliable

Consistent experience across environments, complete with workload templates and cluster cloning for repeatable efficiency

Production-Ready

Pre-configured for immediate production use and includes native security and high availability at scale

Integrated

Part of the Cloudera Enterprise platform featuring the same performance, security, governance, data management, and administration necessary for all production Hadoop deployments

Cloudera Director

The fastest, easiest way to run Hadoop across cloud environments

Public cloud adoption is surging among enterprise organizations. At the same time, cloud-borne data – social, clickstream, sensor data, Internet of things, etc... – continue to grow exponentially. These converging market dynamics, paired with the advantages cloud offers in terms of flexibility, elasticity, and agility, makes public cloud infrastructure an environment that's well suited for modern data analytics.

For an increasing number of organizations, Apache Hadoop is the platform of choice for deriving valuable business insights and action from large volumes of structured and unstructured data. Cloudera Enterprise is the fastest, easiest, and most secure data management and analytics platform built on Apache Hadoop and the latest open source technologies.

Cloudera Director makes it fast and easy to deploy and manage enterprise-grade Hadoop in multiple cloud environments. Key features of Cloudera Director include:

Multicloud support

- Supported out-of-the-box deployment of CDH clusters onto Amazon Web Services (AWS), Google Cloud Platform (GCP), and Microsoft Azure
- Open Cloud Connector framework with open-sourced cloud provider plugins and ability to extend support to other cloud environments

Cluster lifecycle support

- Support for elastic compute and object storage
- Ability to bootstrap with kerberos and high availability enabled
- Out-of-box cluster configuration templates and ability to customize templates for personalized cluster configurations
- Support for large-scale Hadoop clusters

Cluster Management

- Deployment and management across multiple regions from a single Director instance
- Support for AWS spot instances and GCP preemptible instances to minimize cloud infrastructure cost
- Automated billing and metering based on consumption pricing model
- Tightly coupled with Cloudera Manager for robust monitoring and troubleshooting
- High availability and Kerberos configurable

Cloud Object Store support

Many of the most common Hadoop-based workloads in the cloud are short-lived jobs like extract/transform/load (ETL), data modeling, dev/test, and ad hoc business intelligence and reporting. Processes like these may be deployed on transient clusters where a cloud object store acts as the primary data source.

Cloudera is focused on delivering tight integration with core Hadoop processing and analytic components in CDH and cloud-native object stores like Amazon S3.

Object store support

CDH Version	Components supporting Amazon S3
5.5	Apache Hive on YARN, Apache Spark,
5.7	Hive on YARN, Spark, Hive-on-Spark
5.8	Hive on YARN, Spark, Hive-on-Spark, Apache Impala (incubating)

Technical Specifications

Cloudera's enterprise data hub offering, which includes production-grade data management, security, and governance, is supported on public cloud infrastructure using local HDFS with backup to Amazon S3. Check out the Getting Started guides for more detail.

Amazon Web Services Google Cloud Platform Microsoft Azure

	Cloud	Cloudera Director	Cloudera Manager & CDH (with Director)
Operating System (64-bit only)	AWS	RHEL & CentOS 6.5, 6.7, 7.1, and 7.2	RHEL & CentOS 6.5, 6.7, 7.1, and 7.2
	Azure	CentOS 6.6, 6.7	CentOS 6.6, 6.7
	GCP	RHEL & CentOS 6.7, 7.1, 7.2	RHEL & CentOS 6.7, 7.1, 7.2
Default databases		Embedded H2 database	Embedded PostgreSQL database
Supported Databases		MySQL 5.5, 5.6	MySQL 5.5, 5.6
		MariaDB 5.5	MariaDB 5.5

Download Cloudera Director

Cloudera Director is available as a free download for use with CDH.

[Download Cloudera Director
cloudera.com/director](https://cloudera.com/director)

[Try it with AWS Quickstart
aws.amazon.com/quickstart/](https://aws.amazon.com/quickstart/)