

Administrator Training: CDP Private Cloud Base

Price
\$3,195.00

Duration
4 Days

Delivery Methods
VILT, Private Group

Cloudera's four-day administrator training course for CDP Private Cloud Base provides participants with a comprehensive understanding of all the steps necessary to operate and maintain on-premises clusters using Cloudera Manager. From installation and configuration through load balancing and tuning, this Cloudera training course is the best preparation for the real-world challenges faced by administrators who run CDP Private Cloud Base.

Who Should Attend

This course is best suited to systems administrators who have at least basic Linux experience. Prior knowledge of CDP, nor earlier platforms such as Cloudera's CDH or Hortonworks HDP, is not required.

Course Objectives

- Install Cloudera Manager
- Use Cloudera Manager to install a CDP Private Cloud Base cluster
- Configure and monitor the cluster using Cloudera Manager
- Understand, evaluate, and select the most appropriate data storage option
- Optimize cluster performance
- Perform routine cluster maintenance tasks
- Detect, troubleshoot, and repair problems with the cluster

Agenda

1 - CLUDERA DATA PLATFORM



Upcoming Class Dates and Times

[Click Here to View Course Online and Enroll](#)

Jun 4, 5, 6, 7

\$3,195.00

8:00 AM - 4:00 PM

Virtual: Online - CST

- Industry Trends for Big Data
- The Challenge to Become Data-Driven
- The Enterprise Data Cloud
- CDP Overview
- CDP Form Factors

2 - CDP PRIVATE CLOUD BASE INSTALLATION

- Installation Overview
- Cloudera Manager Installation
- CDP Runtime Overview
- Cloudera Manager Introduction

3 - CLUSTER CONFIGURATION

- Overview
- Configuration Settings
- Modifying Service Configurations
- Configuration Files
- Managing Role Instances
- Adding New Services
- Adding and Removing Hosts

4 - DATA STORAGE

- Overview
- HDFS Topology and Roles
- HDFS Performance and Fault Tolerance
- HDFS and Hadoop Security Overview
- Working with HDFS
- HBase Overview
- Kudu Overview
- Cloud Storage Overview

5 - DATA INGEST

- Data Ingest Overview
- File Formats
- Ingesting Data using File Transfer or REST Interfaces
- Importing Data from Relational Databases with Apache Sqoop
- Ingesting Data Using NiFi
- Best Practices for Importing Data

6 - DATA FLOW

- Overview of Cloudera Flow Management and NiFi
- NiFi Architecture
- Cloudera Edge Flow Management and MiNiFi
- Controller Services
- Apache Kafka Overview
- Apache Kafka Cluster Architecture
- Apache Kafka Command Line Tools

7 - DATA ACCESS AND DISCOVERY

- Apache Hive
- Apache Impala
- Apache Impala Tuning
- Search Overview
- Hue Overview
- Managing and Configuring Hue
- Hue Authentication and Authorization
- CDSW Overview

8 - DATA COMPUTE

- YARN Overview
- Running Applications on YARN
- Viewing YARN Applications
- YARN Application Logs
- MapReduce Applications
- YARN Memory and CPU Settings
- Tez Overview
- Hive on Tez
- ACID for Hive
- Spark Overview
- How Spark Applications Run on YARN
- Monitoring Spark Applications
- Phoenix Overview

9 - MANAGING RESOURCES

- Configuring cgroups with CPU Scheduling
- The Capacity Scheduler
- Managing Queues
- Impala Query Scheduling

10 - PLANNING YOUR CLUSTER

- General Planning Considerations
- Choosing the Right Hardware
- Network Considerations

- CDP Private Cloud Considerations
- Configuring Nodes
-

11 - ADVANCED CLUSTER CONFIGURATION

- Configuring Service Ports
- Tuning HDFS and MapReduce
- Managing Cluster Growth
- Erasure Coding
- Enabling HDFS High Availability

12 - CLUSTER MAINTENANCE

- Checking HDFS Status
- Copying Data Between Clusters
- Rebalancing Data in HDFS
- HDFS Directory Snapshots
- Host Maintenance
- Upgrading a Cluster

13 - CLUSTER MONITORING

- Cloudera Manager Monitoring Features
- Health Tests
- Events and Alerts
- Charts and Reports
- Monitoring Recommendations

14 - CLUSTER TROUBLESHOOTING

- Overview
- Troubleshooting Tools
- Misconfiguration Examples

15 - SECURITY

- Data Governance with SDX
- Hadoop Security Concepts
- Hadoop Authentication Using Kerberos
- Hadoop Authorization
- Hadoop Encryption
- Securing a Hadoop Cluster
- Apache Ranger
- Apache Atlas
- Backup and Recovery

16 - PRIVATE CLOUD / PUBLIC CLOUD

- CDP Overview
- Private Cloud Capabilities
- Public Cloud Capabilities
- What is Kubernetes?
- WXM Overview
- Auto-scaling