

ADMIN-336: Running Cloudera Public Cloud

Learn to administer and operate CDP Public Cloud efficiently

Course Overview

Course Type

Instructor-led training course

Level

Advanced

Duration

4 days

Platform

CDP Public Cloud

Topics Covered

- CDP Architecture
- CDP CLI
- CDP Data Hubs
- CDP Data Services
- CDP SDX
- Observability

About This Training

CDP Public Cloud Administrator Training provides participants with a comprehensive understanding of all the steps required to configure, operate, and maintain CDP Public Cloud instances. This four (4) day instructor-led course covers everything from setup to configuring various data services to execute workloads on the cloud on all major cloud providers using Cloudera Management Console. It also covers various configuration options using the web interface and automation scenarios using Ansible. On the optimization side, it covers load balancing and tuning CDP PC instances. This Cloudera training course is the best preparation for the real-world challenges faced by administrators running CDP Public Cloud.

What You Will Learn

Through instructor-led discussion and interactive, hands-on exercises, you will learn how to:

- Evaluate and select the appropriate deployment option
- Setup CDP Public Cloud using Cloudera Management Console
- Setup and configure various data services
- Configure and monitor instances using Cloudera Manager
- Optimize cluster performance and security
- Detect, troubleshoot, and repair problems with the cluster
- Auto scale Data Hub clusters and Data Services

What to Expect

This course is best suited to cloud systems administrators and operators who have at least basic Linux and AWS/Azure/GCP experience. Prior knowledge of CDP, nor earlier platforms such as Cloudera's CDH or Hortonworks HDP, is not required but will be helpful.

Preparation

Students are highly recommended to go through the free OnDemand courses to make the best of the instructor-led classroom learning experience:

- Introducing AWS for CDP Public Cloud (FREE!)
- Cloudera Essentials for CDP (FREE!)
- Introducing - CDP Public Cloud Administration (FREE!)
- Quickstart: Azure for CDP (FREE!)
- Quickstart: AWS for CDP (FREE!)

ADMIN-336: Running Cloudera Public Cloud

Training Outline (Page 2 of 4)

Installation Overview (Quick Start)

- Cloudera Management Console
- CDP Credentials
- CDP Control Plane Regions
- Register a CDP environment
- Cloudera Data Platform
- Industry Trends for Big Data
- The Challenge to Become Data-Driven
- The Enterprise Data Cloud
- CDP Overview
- CDP Form Factors

CDP Architecture

- Overview
- Key Concepts & Components
- CDP Runtime Overview
- Minimum Hardware
- Outbound Connections

Control Plane Overview

- Accessing and Managing an Environment
- Data Management Overview
- Management Console
- Dashboard
- Environments
- Data Lakes
- User Management
- Classic Clusters
- Data Hubs
- Data Catalog
- Replication Manager
- Observability

CDP CLI (Command Line Interface)

- CDP CLI Command Line Interface
- Installing CDP CLI / CLI Client Setup
- CLI Modules
- Generating an API access key / Configuring CDP client
- Logging into the CDP CLI/SDK
- Configuring CLI autocomplete / CLI reference / Accessing CLI help
- CDP API overview / CDP SDK for Java overview / CDP curl overview

Managing CDP Access

- Management Console
- User Management
- Create Machine User
- User Permissions
- Sync Users
- Configure Groups
- Identity Providers
- Roles and Resource Roles
- Global Settings
- Audit Data Storage Credential

Data Hubs Overview

- Data Hubs
- Planning / Creating your Data Hub Cluster
- General Planning Considerations
- Configuring Nodes
- Managing Data Hub
- Choosing the Right Hardware
- Advanced Cluster Configuration
- Data Hub Types
- DataFlow
- Data Engineering
- Troubleshooting

ADMIN-336: Running Cloudera Public Cloud

Training Outline (Page 3 of 4)

Managing Data Hubs

- Best Practices on Data Hubs
- Sizing Data Hubs
- Cloudera Manager
- Data Hub Services
- Autoscaling/Data Hub Info
- Checking Cluster Health Status / Events and Alerts
- Host Maintenance
- Upgrading a Data Hub Cluster
- Monitoring / Monitoring Features

Data Services Overview

- Data Services Overview
- Data Services
- Planning Your Data Service Cluster
- Choosing the Right Hardware / Network Considerations
- Creating Data Services
- DataFlow
- Data Engineering
- Data Warehouse
- Operational Database
- Machine Learning
- Troubleshooting

DataFlow

- DataFlow Service Overview
- Data Ingest Overview
- Ingesting Data using File Transfer or REST Interfaces
- Ingesting Data Using NiFi
- Autoscaling

Data Engineering

- Data Engineering Service Overview
- Apache Spark/Flink/Kafka streams Overview
- Autoscaling

Data Warehouse

- Data Warehouse Service Overview
- Adding and Managing a Database Catalog
- Adding and Tuning a Virtual Warehouse
- Querying a Data Warehouse
- Data Visualization
- Monitoring & Troubleshooting

Operational Database

- Operational Database Service Overview
- Apache HBase/Search Overview
- Autoscaling

Machine Learning

- Machine Learning Service Overview
- CML Engines
- Requirements for CML Workspaces
- Provisioning a CML Workspace
- CML Auto-Scaling
- Monitoring

ADMIN-336: Running Cloudera Public Cloud

Training Outline (Page 4 of 4)

Monitoring and Management

- Monitoring and Management in CDP Public Cloud
- Data Lake Cluster Monitoring and CDP Auditing
- Getting Started with Monitoring in CDP
- Monitoring with Cloudera Manager: Health Tests and Dashboards
- Monitoring Clusters, Services, Hosts, Roles, and Activities
- Troubleshooting Cluster Configuration and Operation

Data Management

- SDX - Security and Governance
- Security Concepts
- Access Cloud Storage
- Data Lake Security: SDX
- Apache Ranger
- CDP Authorization / Authentication
- Data Governance
- Apache Atlas
- Data Catalog

Observability

- Overview
- Support
- Observability deployment architecture
- Monitoring capabilities
- Working with alerts, costs, and reports