ADMIN-335

Administering Data Services on premises

Course Overview

Course Type
Instructor-Led Training

Level
Advanced

Duration
4 days

Platform
Cloudera on premises

Topics Covered
- Architecture for Cloudera Data Services
- Docker Images and Containers
- Docker Private Registry
- Setup and install of Embedded Containerized Services
- Configure Management Console and Environments
- Theory of Kubernetes
- Install, Configure, and Administer Cloudera Data Warehouse
- Install, Configure, and Administer Cloudera Data Engineering
- Install, Configure, and Administer Cloudera AI

About This Training

This four-day course teaches the architecture, deployment, and configuration of Cloudera Data Services on Embedded Containerized Services (ECS). Cloudera Data Services provide a state-of-the-art, low-code platform that unifies the entire data lifecycle, reducing development costs and accelerating the development and deployment of use cases.

The course starts by covering best practices for managing Docker images and containers. Students will then build a Docker private registry. This Docker private registry will be used to deploy a Data Services cluster on ECS. Students will install, configure, and validate Cloudera Data Engineering, Cloudera Data Warehouse, and Cloudera Machine Learning. Through hands-on exercises, students will gain experience with Kubernetes, install a Private Cloud Embedded Container Service (ECS), and deploy Cloudera Data Services. Additionally, the course covers networking and hardware requirements and explains how Kubernetes pods dynamically scale to support Cloudera Data Services.

Who Should Take This Course?

This immersive course is designed for Cloudera Administrators transitioning to managing Cloudera Data Services on premises. Students should have at least 3 to 5 years of system administration experience. Students must have proficiency in the Linux Command Line Interface and knowledge of Identity Management, including Transport Layer Security and Kerberos. Familiarity with SQL select statements is recommended. Prior experience with Cloudera products is required. Students need reliable internet access to connect to the Amazon Web Services environment used in this course.

Recommended Prerequisite Courses

- ADMIN-230: Administering Cloudera on premises
- ADMIN-332: Securing Cloudera on premises

ADMIN-335

Administering Data Services on premises

Big Data to Big Compute

- The importance of microservices
- Overview of containerized applications
- Design Principles for Cloudera Data Services

Docker Images and Containers

- Docker CLI for images and containers
- Building a containerized application

Architecture for Cloudera Platform

- Architecture for Data Platform, Data Lake, and Data Services
- Architecture for Embedded Container Service

Docker Private Registry

- Build a Docker private registry
- Load Cloudera image packages

Clusters for Embedded Containerized Services

- Network requirements
- Hardware requirements
- Cluster recommendations

Shared Data Experience

- The Importance of Ranger, Atlas, and Hive Metastore
- Create Ranger policies on Hive databases in support of Cloudera Data Warehouse

Set Up Embedded Container Services

- Managing an air gapped repo
- Securing a Metastore database
- Creating wildcard DNS and wildcard certificates

Install Embedded Containerized Services

- Install Embedded Containerized Services
- Manage ECS Docker registry

Setup Management Console

- Configure secure LDAP
- Configure an administrator

Setup Environments

- Add users and groups
- Create environments
- Assign users to environments

Operate Management Console

- Manage alerts
- Create support bundles

Theory for Kubernetes

- Theory of Kubernetes clusters
- Theory of Kubernetes namespaces
- Theory of Kubernetes deployments
- Theory of Embedded Containerized Services

Manage Kubernetes

- Use Kubernetes web UI
- Use kubectl CLI
- Use K9 terminal interface

Manage Storage for ECS

- Architecture for Storage
- Architecture for Longhorn

ADMIN-335

Administering Data Services on premises

Run Cloudera Data Warehouse

- Create virtual warehouses for Hive and Impala
- Administrator virtual warehouses

Run Cloudera Data Engineering

- Setup principals and wildcard certificates
- Create virtual clusters for Spark
- Administrator virtual clusters

Run Cloudera Machine Learning

- Create virtual clusters for Spark
- Administrator virtual clusters
- Administrator workspaces

Administration for Embedded Containerized Services

- Stop and start Cloudera clusters
- Uninstall ECS

Cloudera, Inc. 5470 Great America Parkway, Santa Clara, CA 95054