

Cloudera Administrator Training for Apache Hadoop

Learn via: **Classroom / Virtual Classroom / Online**

Duration: **4 Gün**

<https://bilginc.com/tr/egitim/cloudera-administrator-training-for-apache-hadoop-1055-egitimi/>

Overview

Cloudera University'nin bu dört günlük Apache Hadoop yönetici eğitimi, katılımcılara Cloudera Manager kullanarak Hadoop kümesini çalıştırabilmek ve bakımını yapabilmek için gereken tüm adımları kapsamlı bir şekilde anlatmaktadır. Kurulum ve yapılandırmadan yük dengeleme ve ayara dek bu Cloudera eğitimi, Hadoop yöneticilerinin gerçek dünyada karşı karşıya kaldığı sorunlar için en iyi hazırlık imkanını sunmaktadır.

Prerequisites

Herhangi bir ön koşul yoktur.

Who Should Attend

Bu eğitim en çok temel Linux deneyimine sahip olan sistem yöneticileri ile BT yöneticileri için uygundur. Apache Hadoop hakkında önceden bilgi sahibi olunması zorunlu değildir.

What You Will Learn

Bir eğitmenin yönettiği ve tartışma ve etkileşim ağırlıklı uygulamalı alıştırma ile katılımcılar Hadoop ortamında gezinirken aşağıdaki konularda da bilgi sahibi olacaklardır:

- Toplu günlük kaydı tutma, yapılandırma yönetimi, kaynak yönetimi, raporlar, uyarılar ve servis yönetimi gibi kümelerinizi daha kolay yönetmenizi sağlayan Cloudera Manager özellikleri.
- YARN, MapReduce, Spark ve HDFS'nin özellikleri
- Kümeniz için doğru donanım ve altyapıyı belirleme
- Veri merkezinizle entegre olacak doğru küme yapılandırması ve kurulumu
- Flume kullanarak dinamik oluşturulmuş dosyalardan ve Sqoop kullanarak ilişkisel veritabanı yönetim sisteminden verilerin kümeye nasıl yükleneceği
- Bir kümenin birden çok kullanıcıya hizmet seviyesi anlaşmaları sunmak amacıyla FairScheduler'in yapılandırılması
- Apache Hadoop'un hazırlanmasına ve üretim ortamında bakımının yapılmasına yönelik en iyi uygulamalar
- Sorun tespit ve giderme, ayar yapma ve Hadoop sorunlarının çözülmesi

Outline

Introduction

The Case for Apache Hadoop

- Why Hadoop?
- Fundamental Concepts
- Core Hadoop Components

Hadoop Cluster Installation

- Rationale for a Cluster Management Solution
- Cloudera Manager Features
- Cloudera Manager Installation
- Hadoop (CDH) Installation

The Hadoop Distributed File System (HDFS)

- HDFS Features

- Writing and Reading Files
- NameNode Memory Considerations
- Overview of HDFS Security
- Web UIs for HDFS
- Using the Hadoop File Shell

MapReduce and Spark on YARN

- The Role of Computational Frameworks
- YARN: The Cluster Resource Manager
- MapReduce Concepts
- Apache Spark Concepts
- Running Computational Frameworks on YARN
- Exploring YARN Applications Through the Web UIs, and the Shell
- YARN Application Logs

Hadoop Configuration and Daemon Logs

- Cloudera Manager Constructs for Managing Configurations
- Locating Configurations and Applying Configuration Changes
- Managing Role Instances and Adding Services
- Configuring the HDFS Service
- Configuring Hadoop Daemon Logs
- Configuring the YARN Service

Getting Data Into HDFS

- Ingesting Data From External Sources With Flume
- Ingesting Data From Relational Databases

With Sqoop

- REST Interfaces
- Best Practices for Importing Data

Planning Your Hadoop Cluster

- General Planning Considerations
- Choosing the Right Hardware
- Virtualization Options*
- Network Considerations
- Configuring Nodes

Installing and Configuring Hive, Impala, and Pig

- Hive
- Impala
- Pig

Hadoop Clients Including Hue

- What Are Hadoop Clients?
- Installing and Configuring Hadoop Clients
- Installing and Configuring Hue
- Hue Authentication and Authorization

Advanced Cluster Configuration

- Advanced Configuration Parameters
- Configuring Hadoop Ports
- Configuring HDFS for Rack Awareness
- Configuring HDFS High Availability

Hadoop Security

- Why Hadoop Security Is Important
- Hadoop's Security System Concepts
- What Kerberos Is and how it Works
- Securing a Hadoop Cluster With Kerberos
- Other Security Concepts

Managing Resources

- Configuring cgroups with Static Service Pools
- The Fair Scheduler
- Configuring Dynamic Resource Pools

- YARN Memory and CPU Settings
- Impala Query Scheduling

Cluster Maintenance

- Checking HDFS Status
- Copying Data Between Clusters
- Adding and Removing Cluster Nodes
- Rebalancing the Cluster
- Directory Snapshots
- Cluster Upgrading

Cluster Monitoring and Troubleshooting

- Cloudera Manager Monitoring Features
- Monitoring Hadoop Clusters
- Troubleshooting Hadoop Clusters
- Common Misconfigurations

Conclusion