

Advanced Big Data Analytics, Architecture, Management and Applications

Course Overview

This course covers the major architecture designs required to cater to different needs of the application, data center or deployment requirements. This course provides architectural designs and advanced hands-on training on topics covering Scaling of cluster to thousands of nodes and management, Lambda architecture for streaming analytics, Data Life Cycle management with HDFS tiered storage, and different approaches for Multi-tenant Hadoop cluster deployments with Openstack, UCSD Express, or with MapR volumes and Work-load Automation topics concerning the deployment of Big Data clusters. This course will cover the application and infrastructure architecture components for each use case with specific focus on Hadoop deployments. You will learn UCSD Express end to end automation for big data as well as best practices around Data Disaster Recovery and security.

Course Objectives

After completing this course you should be able to:

- Scale up to thousands of nodes
- Be familiar with Business Analytics including Splunk
- Install and configure Red Hat, SPLUNK
- Describe Edge Analytics (Social Media, IoT)
- Describe SQL on Hadoop
- Describe Data Life Cycle management for Tiered Storage of Hot, Warm and Cold data
- Install Hadoop as a Service (on OpenStack)/Multi-tenant Hadoop cluster
- Review End to End Automation with UCSD Express for Big Data
- Describe Hadoop as a Service on Bare Metal
- Describe Disaster Recovery
- Apply Security

Course Prerequisites

Attendees should meet the following prerequisites:

- Attended Cisco Big Data Analytics, Architecture and Management (ANDMB)
- Familiarity with Cisco UCS and Big Data Fundamentals

Recommended prerequisites:

- ANDMA - Advanced Big Data Analytics, Architecture, Management and Applications

Course Content

Module 1: Scaling to thousands of nodes - ACI

Module 2: Analytics Splunk

Module 3: Edge Analytics (Social Media, IoT)

Module 4: SQL on Hadoop

Module 5: Data Life Cycle management - Tiered Storage of Hot, Warm and Cold data

Module 6: Hadoop as a Service (on OpenStack)/Multi-tenant Hadoop cluster

Module 7: End to End Automation with UCS Director Express for Big Data

Module 8: Hadoop as a Service on Bare Metal

Module 9: Disaster Recovery

Module 10: Security

Labs:

- Lab 1: Introduction to Cisco ACI for Big Data Cisco Validated Designs
- Lab 2: Splunk
- Lab 3: Administering Hadoop on Cisco UCS
- Lab 4: SQL on Hadoop with Hive
- Lab 5: Cisco UCS Director Express for Big Data
- Lab 6: Apache Falcon Pipelines
- Lab 7: Security with Knox and Ranger