HDP Developer: Custom Apache YARN Applications

Overview

This course is designed for developers who want to create custom YARN applications for Apache Hadoop. It will include: the YARN architecture, YARN development steps, writing a YARN client and ApplicationMaster, and launching Containers. The course uses Eclipse and Gradle connected remotely to a 7-node HDP cluster running in a virtual machine.

Course Objectives

• Describe the YARN architecture
• Describe the YARN application lifecycle
• Write a YARN client application
• Run a YARN application on a Hadoop cluster
• Monitor the status of a running YARN application
• View the aggregated logs of a YARN application
• Configure a ContainerLaunchContext
• Use a LocalResource to share application files across a cluster
• Write a YARN ApplicationMaster
• Describe the differences between synchronous and asynchronous ApplicationMasters
• Allocate Containers in a cluster
• Launch Containers on NodeManagers
• Write a custom Container to perform specific business logic
• Explain the job schedulers of the ResourceManager
• Define queues for the Capacity Scheduler

Hands-On Labs

• Run a YARN Application
• Setup a YARN Development Environment
• Write a YARN Client
• Submit an ApplicationMaster
• Write an ApplicationMaster
• Requesting Containers
• Running Containers
• Writing Custom Containers

Duration

2 days

Target Audience

Java software engineers who need to develop YARN applications on Hadoop by writing YARN clients and ApplicationMasters.

Prerequisites

Students should be experienced Java developers who have attended *HDP Developer: Java* **OR** *HDP Developer: Pig and Hive* **OR** are experienced with Hadoop and MapReduce development.

Format

50% Lecture/Discussion
50% Hands-on Labs

Certification

Hortonworks offers a comprehensive certification program that identifies you as an expert in Apache Hadoop. Visit *hortonworks.com/training/certification* for more information.

Hortonworks University

Hortonworks University is your expert source for Apache Hadoop training and certification. Public and private on-site courses are available for developers, administrators, data analysts and other IT professionals involved in implementing big data solutions. Classes combine presentation material with industry-leading hands-on labs that fully prepare students for real-world Hadoop scenarios.