

HDP Developer: Custom YARN Applications

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

Duration: **2 Days**

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.

Prerequisites

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

Who Should Attend

Architects, software developers, analysts and data scientists who need to apply data science and machine learning on Hadoop.

What You Will Learn

At the completion of the course students will be able to:

- 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

Outline

- 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