

Durée : 4 jours soit 28h

SUMMARY

Learn how to harness the power of Apache Hadoop by building robust data processing applications that unlock the full potential of your data.

Cloudera, the leader in Apache Hadoop-based software and services, delivers a developer-focused training course that closely analyzes Hadoop's structure and provides hands-on exercises that teach you how to import data from existing sources; process data with a variety of techniques such as Java MapReduce programs and Hadoop Streaming jobs; and work with Apache Hive and Pig.

After the course, you can take the Cloudera Certified Developer exam to become the Cloudera certified expert in your organization.

AGENDA TOPICS

- The motivation for Hadoop
- Hadoop: Basic concepts
- Writing a MapReduce program
- The Hadoop ecosystem
- Integrating Hadoop into the workflow
- Common MapReduce algorithms
- Using Apache Hive and Pig
- Debugging MapReduce programs
- Advanced MapReduce programming
- Joining data sets in MapReduce jobs
- Graph manipulation in Hadoop
- The new Hadoop API
- Cloudera Certified Developer exam