+33 (0)1 48 10 09 46
formation@ambient-it.net
140 Avenue Jean Lolive 93500 Pantin - France

Updated 07/27/2023

Sign up

Industry 4.0 Training: Real Time Financial & IoT Data Processing

3 days (21 hours)

Presentation

We've just entered the era of Industry 4.0. Thanks to the contributions of new-generation Big Data technologies such as Apache Kudu, Impala, Spark, and Flink, new approaches have emerged, including Real Time Financial & IoT Data Processing.

The course takes a holistic approach and is intended as a practical guide covering the most important aspects of Big Data for modern businesses.

You'll be able to grasp and solve the majority of your architecture, performance and storage problems arising from your company's day-to-day IT issues.

We'll be using next-generation tools and applications such as :

- For storing large amounts of data
- Data warehouse optimization
- Real-time and batch data ingestion and processing
- Real-time data visualization
- Big data governance
- In-memory Big Data computing

Objectives

- Install Apache Kudu, Impala, Spark, Flink to modernize enterprise data warehousing and business intelligence environments, with concrete, easy-to-follow examples and practical tips.
- Integration of HBase, Solr, Oracle, SQL Server, MySQL, Flume, Kafka, HDFS and Amazon S3 with Apache Kudu, Impala and Spark

- Understand enterprise Big Data topics such as Big Data governance, metadata management, data lineage, impact analysis and policy enforcement, and how to use Cloudera Navigator to perform common data governance tasks.
- Implement big data use cases such as data warehousing optimization, the Internet of Things, real-time data ingestion and analysis, complex event processing and scalable predictive modeling.
- Be able to deploy solutions through various case studies: massive data carried out in the real world with innovative companies such as Navistar, Cerner, British Telecom, Shopzilla, Thomson Reuters and Mastercard.

Target audience

- Developer, Lead Developer, Data Engineer

Prerequisites

- Good programming skills in a recent language
- Know the basics of a Linux system
- Senior Big Data Engineer

Further information

- We also recommend the following training courses:
  - IBM Bluemix & Watson
  - Machine Learning with Spark
  - Deep Learning with TensorFlow

Apache Flink, Kudu, Impala & Spark training program

- New-generation Big Data
- Introduction to Kudu
- Introduction to Impala
- High-performance data analysis with Impala
- Introduction to Spark and Flink
- High-performance data processing with Spark, Flink and Kudu
- Batch and real-time data ingestion and processing
- Large data warehousing
- Big Data visualization
- Big Data Computing
- Governance and management
- Big Data in the cloud
- Big Data case studies

Companies concerned

This course is aimed at both individuals and companies, large or small, wishing to train their teams in a new advanced computer technology, or to acquire specific business knowledge or modern methods.

Teaching methods

Practical course: 60% Practical, 40% Theory. Training material distributed in digital format to all participants.

Organization

The course alternates theoretical input from the trainer, supported by examples, with brainstorming sessions and group work.

Validation

At the end of the session, a multiple-choice questionnaire verifies the correct acquisition of skills.

Sanction

A certificate will be issued to each trainee who completes the course.