XCELERATE TRAINING INSTITUTE

your learning & development partner

INFORMATION TECHNOLOGY

Apache Spark

DURATION	LANGUAGE	FORMAT
5 days	English	Classroom / Online

Upcoming dates

Dates	Venue	Format	Fee
22 – 26 ?????? 2027	Live online	Online (live)	€1?200
2 – 6 ????? 2027	Live online	Online (live)	€1?200
6 – 10 ?????? 2027	Bahrain	Classroom	€4?800

Register or ask a question

Online: staging.x8ti.com/course/apache-spark/

+491631686678 · info@x8ti.com

Apache Spark

Introduction

This course by Xcelerate Training Institutes Apache Spark training program equips learners with the knowledge to understand and leverage Spark's in-memory processing capabilities for significantly faster data analysis compared to Hadoop Map Reduce. Participants will gain proficiency in Scale programming and explore various Spark APIs including Spark Streaming, Spark SQL, Spark RDD, Spark MLlib, and Spark Graph X. This course is essential for aspiring Big Data developers.

In today's data-driven world, extracting meaningful insights from vast datasets is crucial. While multiple big data processing tools exist, Spark stands out due to its ability to handle both batch and streaming data, making it an ideal choice for rapid big data analytics.

Learning Objectives

Upon completion, participants will:

- Master Scale programming and its application in Spark
- Install and operate Spark on Spark Shell
- Grasp the concept of Spark RDD
- Develop Spark applications on YARN (Hadoop)
- Utilize Spark Streaming API
- Implement machine learning models using Spark MLlib
- Analyze Hive and Spark SQL architecture
- Optimize performance using Broadcast variables and Accumulators
- Complete a hands-on project

Training Methodology

A comprehensive Apache Spark training program should incorporate a blend of theoretical concepts and hands-on exercises. Begin with foundational Spark concepts like RDDs, DataFrames, and Spark SQL. Advance to more complex topics like Spark Streaming, MLlib, and GraphX. Provide ample opportunities for participants to work on real-world datasets and projects, applying their knowledge to solve practical problems. Consider using a mix of lectures, demonstrations, and group activities to foster a collaborative and engaging learning environment.

Benefits for Your Organisation

Its in-memory processing capabilities enable extremely fast data processing and analysis, making it ideal for real-time applications and large-scale data sets. Spark's unified platform supports a wide range of data processing workloads, including batch processing, streaming, and machine learning, reducing the need for multiple tools and simplifying data management. Additionally, Spark's fault tolerance and scalability ensure high availability and the ability to handle growing data volumes. These benefits contribute to increased efficiency, improved decision-making, and overall organizational success.

Benefits for You

Apache Spark is a powerful, open-source data processing engine that offers numerous benefits. Its in-memory computing capability significantly accelerates data processing tasks, enabling rapid analysis and real-time applications. Spark's unified platform supports a wide range of data processing workloads, including batch processing, streaming, machine learning, and graph processing. Additionally, Spark's fault tolerance ensures data reliability and minimizes downtime. Its integration with various data sources and frameworks, such as Hadoop and Kafka, simplifies data ingestion and management. Overall, Spark's speed, versatility, and reliability make it a valuable tool for organizations seeking to extract insights from large and complex datasets.

Who Should Attend

Data scientists, analysts, developers, solution architects, and anyone eager to acquire new technical skills can benefit from this Apache Spark certification training.

Course Outline

Spark Fundamentals

- Introduction to Spark: purpose and components
- Understanding Resilient Distributed Datasets (RDDs)
- Overview of Scale and Python
- Hands-on experience with Spark's Scale and Python shells

RDDs and Data Frames

- Creating and managing parallel collections and external datasets
- Mastering RDD operations
- Working with shared variables and key-value pairs

Spark Application Development

- Exploring Spark Context and its applications
- Initiating Spark projects using different programming languages
- Executing Spark examples
- Passing functions to Spark
- Building and running standalone Spark applications
- Submitting applications to clusters

Spark Libraries

- Comprehensive overview of Spark libraries
- Deep dive into Spark Core programming
- Understanding and utilizing Spark SQL
- Introduction to Spark Machine Learning

Advanced Spark Components

- Exploring Machine Learning algorithms
- Practical examples

- Introduction to Spark Streaming

Spark Configuration, Monitoring, and Optimization

- Understanding Spark cluster architecture
- Configuring Spark properties
- Environment variables and logging
- Monitoring Spark performance using web UIs
- Metrics
- External tools
- Optimizing Spark performance

Copyright and intellectual property
This course, "Apache Spark", is an original course by Xcelerate Training Institute. All materials, content and structure are the intellectual property of Xcelerate Training Institute. Reproduction, distribution or use of any part of this course without written consent is prohibited.

Registration Form

Apache Spark

Preferred dates: ..... Venue or online: .....

Duration: 5 days Language: .....

Please complete and email this form to us, or register online. We confirm your place and send the invoice within one working day.

Delegate information

Full name (Mr / Ms / Dr / Eng): .....

Position: ..... Telephone / mobile: .....

Official email: ..... Personal email: .....

Company information

Company name: .....

Address: .....

City / country: ..... VAT / tax number: .....

Person responsible for training & development

Full name (Mr / Ms / Dr / Eng): .....

Position: ..... Telephone / mobile: .....

Email: .....

Payment

☐ Please invoice me ☐ Please invoice my company

Signature: ..... Date: .....

Call	Email this form	Register online
+491631686678	info@x8ti.com	Open the course page