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

Updated on 27/11/2024

Sign up

Samza training

3 days (21 hours)

Presentation

Samza is a scalable data processing engine that enables you to process and analyze your data in real time. Our Samza training course will give you an in-depth understanding and simplify your data flow.

In this hands-on training course, you'll learn how to manage, set up, administer and configure Samza.

You'll discover how to use Samza to manipulate data pipelines, master its concepts, use its main components, its basic architecture and call its APIs.

At the end of this course, you'll be able to configure, optimize and automate your data flows. You'll master API configuration and deployment, performance optimization, latency management and scalability.

As with all our courses, Samza will be presented with its latest version: Samza 1.8.0.

Objectives

- Understanding Samza's architecture
- Understanding how data flows work
- Using and understanding tools and APIs
- **Securing** the various flows
- Understand how Samza works and integrate it successfully.

Target audience

- Data Scientists
- Developers
- Architects
- Data analysts
- Data engineers

Prerequisites

- Good understanding of databases and structured file manipulation
- Experience with a programming language such as Python or Java
- Basic knowledge of massive data processing (Hadoop, Spark, etc.).
- Familiarity with columnar storage and compression concepts

Samza training program

INTRODUCTION TO SAMZA

- Introducing **Samza**
- Advantages of using Samza to process large amounts of data
- Typical use cases
- Introduction to Samza

PRESENTATION, BASIC CONCEPTS

- Introduction to Samza
- Flow and partition
- **Flow requests**
- Status processing
- Processing time and event time

ARCHITECTURE

- Threading model and commands
- Status management
- State fault tolerance
- Distributed execution
  - Tasks
  - container
  - Coordinator
- Different host affinities

SAMZA CONFIGURATION

- Samza application configuration

- Advanced application configuration
- Control points, systems and flows
- Consumption and production configurations on Amazon Kinesis

API PRESENTATION

- Introduction of Samza APIs
- Key concepts and Samza application
- Flux and table Descriptors
- Define flow processing logic

HIGH AND LOW LEVEL API

- Difference, API definition and code examples
- Key concepts and operators
- Data serialization and application
- Samza TABLE API

SAMZA DEPLOYMENT

- Deployment options
- How it works on YARN
- Samza library integration and execution

THE DIFFERENT SAMZA CONNECTORS

- Connector presentation
- Apache Kafka
- Azure Event Hubs
- AWS Kinesis

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.

Positioning on entry to training

Positioning at the start of training complies with Qualiopi quality criteria. As soon as registration is finalized, the learner receives a self-assessment questionnaire which enables us to assess his or her estimated level of proficiency in different types of technology, as well as his or her expectations and personal objectives for the training to come, within the limits imposed by the selected format. This questionnaire also enables us to anticipate any connection or internal security difficulties within the company (intra-company or virtual classroom) that might be encountered.

problematic for the follow-up and smooth running of the training session.

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.