AMBIENT-IT

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

Updated on 26/09/2024

Sign up

Apache Sedona training

2 days (14 hours)

Presentation

Do you work with geospatial data and want to integrate it into your Big Data pipelines? Our Apache Sedona training course will give you complete mastery of geospatial computation at massive scale by harnessing the power of Apache Spark.

Apache Sedona, an open source framework, facilitates the analysis of large quantities of geospatial data with advanced features such as spatial indexing, geospatial joins and spatial data visualization.

You'll discover how Sedona can be integrated with Spark to run geospatial SQL queries and analyze massive datasets.

This training course will enable you to understand the fundamental concepts of geospatial data, configure Apache Sedona, and effectively manipulate this data to gain insights from interactive visualizations.

As with all our courses, Apache Sedona training will be presented with its latest features (at the time of writing).

Objectives

- Understanding the basics of Apache Sedona and its integration with Apache Spark
- Learn to manipulate and index geospatial data
- Write and optimize geospatial queries with Spark SQL
- Master large-scale spatial data visualization and analysis techniques
- Use optimization tools for distributed geospatial query processing

Target audience

- Data engineers
- Data analysts
- Big Data Architects
- Developers

Prerequisites

- Basic knowledge of Apache Spark
- Understanding Big Data concepts
- Geospatial knowledge (spatial data types) a plus but not essential

APACHE SEDONA TRAINING PROGRAM

INTRODUCTION TO APACHE SEDONA

- Introduction to Apache Sedona and its role as a geodatabase extension for Apache Spark
- Discussion on the importance of geospatial data analysis
- Overview of Sedona components: SedonaCore, SedonaSQL, SedonaViz
- Exploring use cases and practical applications

INSTALLATION AND CONFIGURATION

- Hardware and software requirements for installing Apache Sedona
- Detailed procedure for installing Apache Sedona and its integration with Apache Spark
- Initial configuration and installation check
- Solving common installation problems

HANDLING GEOSPATIAL DATA

- Introduction to supported geospatial data formats: WKT, WKB, GeoJSON, Shapefiles
- Techniques for loading and managing spatial data in Apache Sedona
- Practical exercises on manipulating basic geospatial data

SPATIAL INDEXING AND GEOSPATIAL QUERIES

- Understanding the different types of spatial indexing: Grid, Quadtree, R-tree
- Importance of indexing for spatial query optimization
- Create and execute complex geospatial queries
- Use geospatial joins to enrich analyses

INTEGRATION WITH SPARK SQL

- Using Spark SQL to run queries on geospatial data
- Write geospatial-specific SQL queries in the Apache Sedona environment
- Practical exercises for writing and optimizing geospatial SQL queries

VISUALIZATION AND ANALYSIS OF GEOSPATIAL DATA

- Introducing SedonaViz for spatial data visualization
- Techniques for creating interactive maps and visualizing large amounts of spatial data
- Analysis of case studies and use of visualizations to gain insights from data. geospatial data
- Discussion of best practices for effective geospatial data analysis and visualization

Companies concerned

This training 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 security difficulties within the company (intra-company or virtual classroom) which could be 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.