

Updated on 16/10/2023

Sign up

Talend DI Advanced training

2 days (14 hours)

Presentation

Talend is the leading commercial provider of open source data integration software.

Talend is an open source data integration platform. It provides a variety of software and services for data integration, data management, enterprise application integration, data quality, cloud storage and Big Data.

In this course, you'll learn all about Talend, from the fundamentals to advanced data integration and big data, with examples.

We'll teach you the skills you need to master for advanced use of Talend Data Integration. This includes joblets, routines, performance optimization and data manipulation.

Like all our training courses, this one will introduce you to the latest stable release and its new features: [Talend 8](#).

Objectives

- Understanding and using Talend routines effectively
- Running and debugging a remote Job from Talend Studio
- Knowledge of parallel execution methods and data storage
- Manage and optimize data handling

Target audience

- Data engineers
- Developers

- System administrators
- Database administrators

Prerequisites

- Knowledge of Java and SQL
- Knowledge of databases
- Have completed our [Talend DI](#) training course or master the entire program

Installation requirements

- Access to our Azure machines via Remote Desktop (RDP)
- A few days before the session, you will need to create a 14-day trial account (typically from your Gmail), in order to be able to carry out the practical work in the same Talend Cloud environment.

Talend DI Advanced training program

Introduction

- Talend Data Integration basics
- ELT vs ETL
- Starting Talend Studio
- Git in Talend Studio
- Connecting to a remote repository

Routines

- Routines for thongs and characters
- Time and date routines
- Arithmetic operation routines

Remote Job execution

- Running a Job from Talend Studio on a remote JobServer
- Best practices
- Debugging

Parallel execution

- Parallel execution methods in a Job

- JVM parameters

Data storage

- Job design using memorization logic
- Using buffer and hash components
- Joblet benefits and use cases
- Creating joblets
- Using Joblets

Database administration

- Connections to databases and other sources
 - MongoDB
 - BigQuery
 - Snowflake
 - FTP
 - Email
- Commit and rollback
- Insert and Update
- Using bulk components
- Stored procedures
- Optimizing performance

Data handling

- File and copy management
- Using XML
- Using JSON

Putting skills and webservices into practice

- Fast reading of large XML files
- Currency conversion with a REST service
- Data testing
- Discovering WebServices components

Further information

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.