

Data Integration - Advanced

Duration	2 days (16 hours)
Target audience	Anyone who wants to use Talend Data Integration to perform data integration and management tasks: software developers and development managers
Prerequisites	Completion of Talend Data Integration Basics and knowledge of computing, including familiarity with Java or another programming language, SQL, and other general database concepts
Course objectives	<p>After completing this course, you will be able to:</p> <ul style="list-style-type: none"> • Start and connect the Talend Studio to a remote repository • Use SVN branches in the Studio • Run a Job in the Studio on a remote Job server • Monitor host CPU and JVM memory in real time during Job execution • Use debugging features in the Studio • Configure a Talend project to capture statistics and logs, and monitor them from the Activity Monitoring Console (AMC) • Implement several methods of parallel execution in a Talend Job • Create Joblets • Create a unit test from a working Job • Configure a database to monitor and log changes in a separate change data capture (CDC) database, and use it to perform incremental updates between the source and target • Understand SDLC Process and Tools <ul style="list-style-type: none"> ◦ Source Code Management ◦ Binaries Management ◦ Automatic Promotion and Deployment
Course agenda	<p>Connecting to a remote repository</p> <ul style="list-style-type: none"> • Creating a remote connection <p>SVN in the Studio</p> <ul style="list-style-type: none"> • Copying a Job to a branch • Comparing Jobs • Resetting a branch <p>Remote Job execution</p> <ul style="list-style-type: none"> • Creating and running a Job remotely <p>Activity Monitoring Console (AMC)</p> <ul style="list-style-type: none"> • Configuring statistics and logging • Using the Activity Monitoring Console (AMC)

Parallel execution

- Writing large files
- Writing to databases
- Parallelizing automatically
- Partitioning

Joblets

- Creating a Joblet from an existing Job
- Creating a Joblet from scratch
- Triggering Joblets

Unit test

- Creating a unit test

Change data capture

- Examining databases
- Configuring the CDC database
- Monitoring changes
- Updating the warehouse
- Resetting the databases

Version Control

- Managing version with SVN/Git
- Branching and Tagging

Software Development Life Cycle (SDLC)

- Binaries Management.
- Publisher
- Promoting Projects
- Deployment