Talend - Course Curriculum

1. Role of Open Source ETL Technologies in Big Data

- Overview on: TOS (Talend Open Studio) for Data Integration
- ETL concepts
- Data warehousing concepts

2. Talend

- Why Talend?
- Features
- Advantages
- Talend Installation/System Requirements
- GUI layout (designer)

Uplatz
training.uplatz.com
Training Provider for IT and Certification Courses

- Understanding it's Basic Features

- Comparison with other market leader tools in ETL domain

- Important areas in Talend Architecture: Project

- Workspace

- Job

- Metadata

- Propagation

- Linking components

3. Talend: Read & Write various Types of Source/Target System

- Data Source Connection

- File as Source

- Create meta data

- Database as source

- Create metadata

- Using MySQL database (create tables, Insert, Update Data from Talend)

- Read and write into excel files, into multiple tabs

- View data

- How to capture log and navigate around basic errors

- Role of tLogrow and how it makes developers life easy

4. Talend: How to Transform Your Business: Basic

- Using Advanced components like: tMap, tJoin, tFilter, tSortRow, tAggregateRow, tReplicate, tSplit, Lookup, tRowGenerator

5. Talend: How to Transform Your Business: Advanced 1

- Trigger (types) and Row Types

- Context Variables (parameterization)

- Functions (basic to advanced functions to transform business rules such as string, date, mathematical etc.)

- Accessing job level / component level information within the job

6. Talend: How to Transform Your Business: Advanced 2

- Type Casting (convert data types among source-target platforms)

- Looping components (like tLoop, tFor)

- tFileList

- tRunJob

- How to schedule and run talend DI jobs externally (not in GUI)

7. Working with Hierarchical File Structures

- Read and Write an XML file, configure the schema and XPath expression to parse an XML file

- Read and Write a JSON file, configure the schema and JSONPath expression to parse a JSON file

- Read and write delimited, fixed width files.

8. Context Variables and Global Variables

- Create context/global variables

- Use context/global variables in the configuration of Talend components

- Load context variables from a flow

9. Best practices

- Working with databases and implementing data warehousing concepts

- Working with files (excel, delimited, JSON, XML etc.)

10. Orchestration and Controlling Execution Flow

- Files - Use components to list, archive, and delete files from a directory

- Database – Controlling Commit and Rollback

- COMMIT at end of job/ every x number of rows
- Rollback on error

11. Shared DB connection across jobs and subjobs

- Use triggers to connect components and subJobs
- Orchestrate several jobs in master jobs.
- Handling Errors
  - Kill a Job on a component error
  - Implement a specific Job execution path on a component error
  - Configure the log level in the console

statement of accounts