ETL Talend Fundamentals 1

Cyrus Lentin

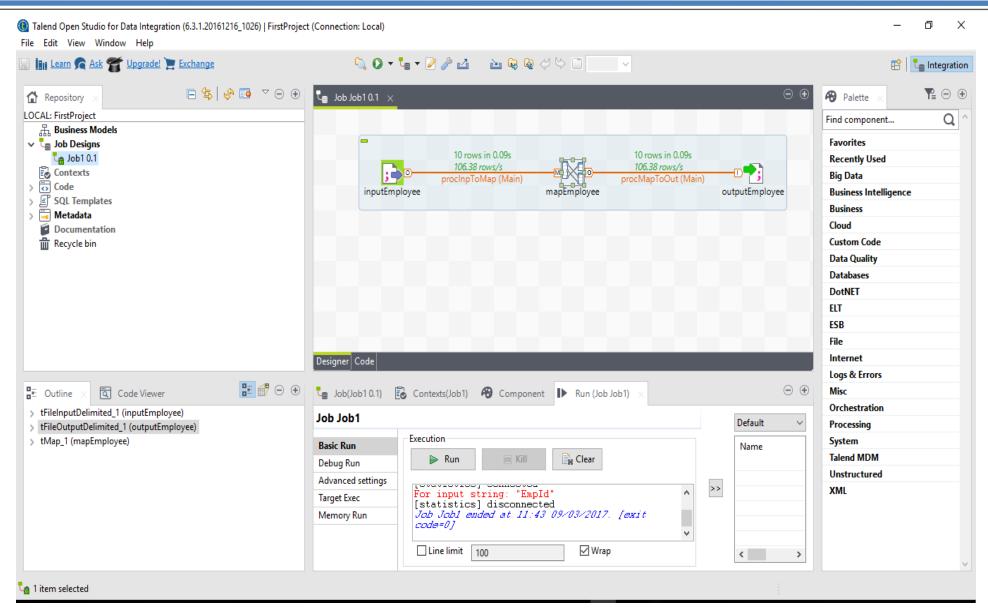
Talend

- Visual Programming IDE For ETL Applications
- Java Code Generator
- 600+ Connectors For Open And Proprietary Data Systems
- Easily Embeddable In Custom Applications
- Cross-platform
- Central Metadata Repository
- Available In Both Open Source And Premium Flavors

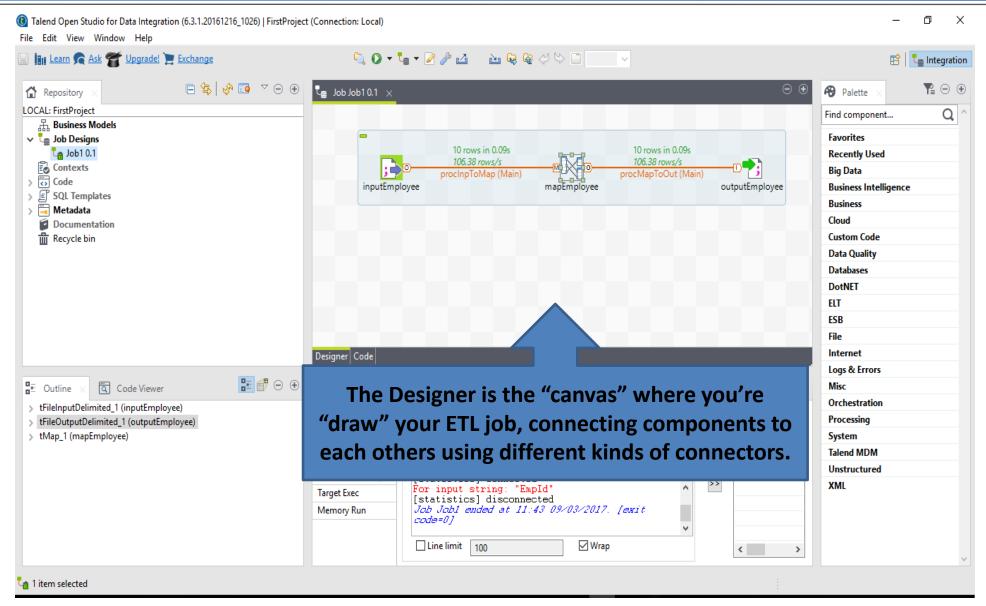
Talend Open Studio

- IT's The Open Source, Free To Use, Community-supported Version Of Talend For Data Integration
- Often Abbreviated In "TOS", To Differ From The Premium Version ("TIS")
- Features-lite, But Still Completely Usable
- Same Set Of Connectors And Components Of The Premium Version
- It Misses Team Working And Enterprise Capabilities Like SVN, Scheduling, Process Orchestrations
 And Monitoring Console

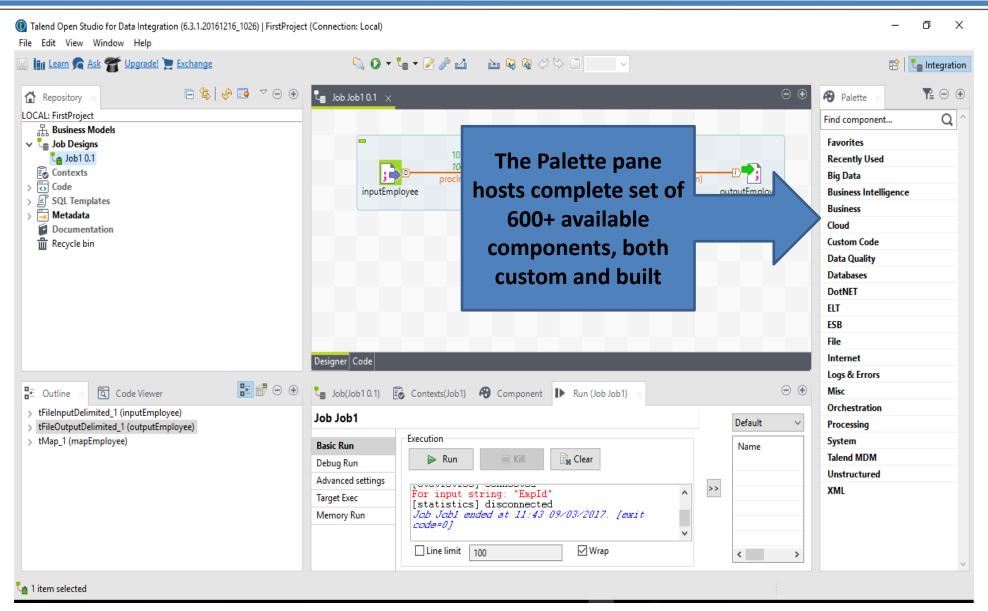
TOS – First Look



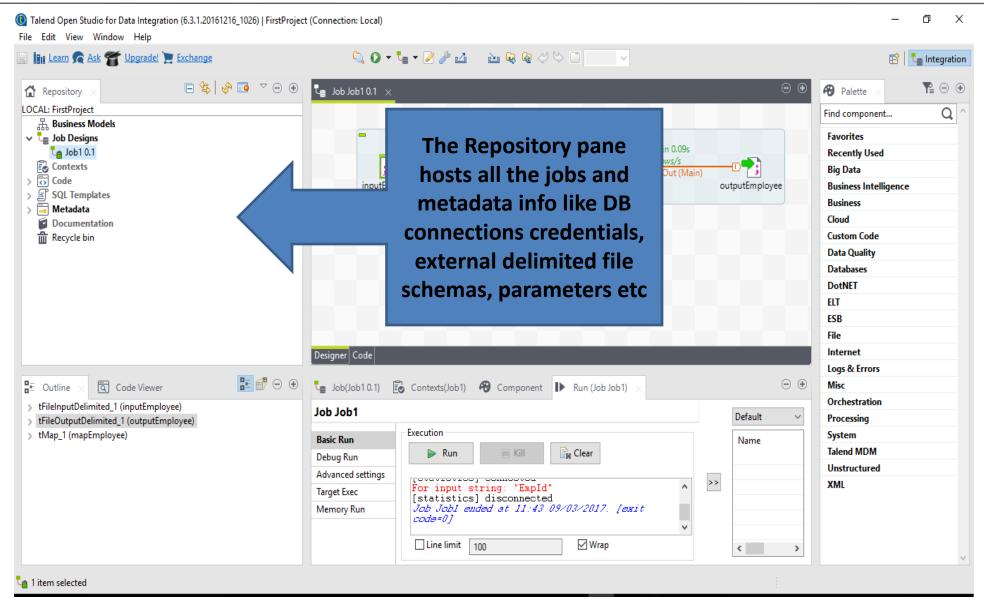
Designer



Palette

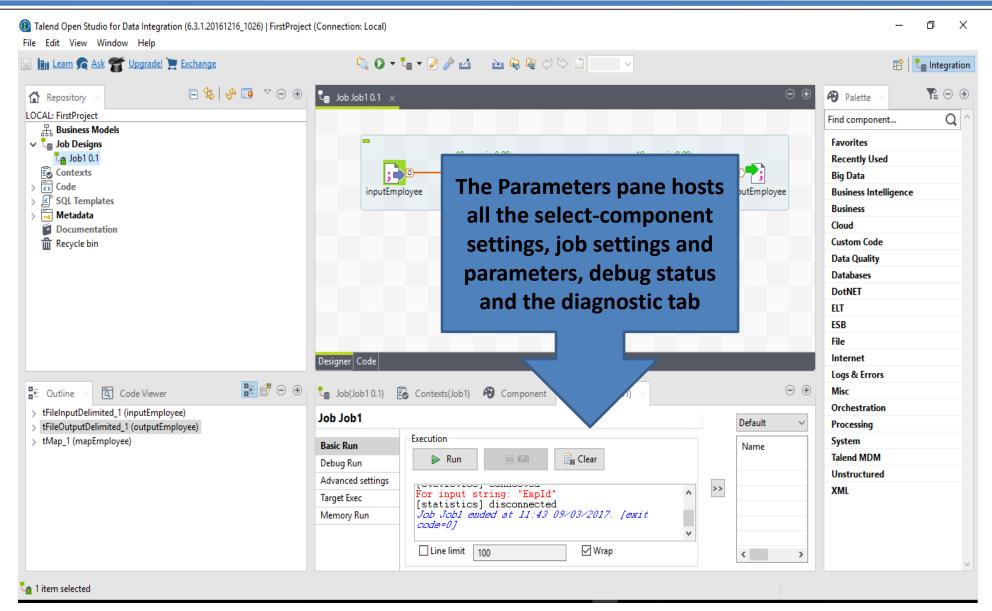


Repository



ETL – Cyrus Lentin 6

Parameters



Workspace

- A Workspace Is A Container Of Projects Which Shares The Same TOS Version And The Same Components Palette
- You Can Choose Which One To Use When The Program Starts
- In TOS, It's A Folder In The Local Drive

Business Diagrams

- It stores diagrams used to conveniently describe business models and to embed them with ETL
- It offers a small set of drawing capabilities in UML-fashion
- It is not widely used, but it's proven to be useful to quickly sketch-up transformation goals
- It is also useful for auto documenting ETL

Repository

- Repository Provides TOS With Preliminary Definition & Description Of Jobs & Metadata
- There Are 8 Types Of Repositories
- Important Repositories Are:
 - Business Models
 - Jobs
 - Contexts
 - Code
 - Metadata
 - SQL Templates

Projects

- A Project Is A Set Of Jobs And Involved Metadata
- It's Defined Under A Subfolder Into The Workspace
- All TOS Preferences Are Project-based
- In Other Words, Different Projects In The Same Workspace Can Have Different Settings
- Internally, It's A Mix Of XML, .Items And .Properties Files

Jobs

- It Is The Core Of TOS Repository
- The Jobs Graphically Describes The Jobs
- All Components Used In The Job Are Described Here
 Components Connectors, Signals, Parameters, Colors And Presentation
- Jobs Should Be Organized Properly For Better Clarity

Thank you!

Contact:

Cyrus Lentin cyrus@lentins.co.in +91-98200-94236