

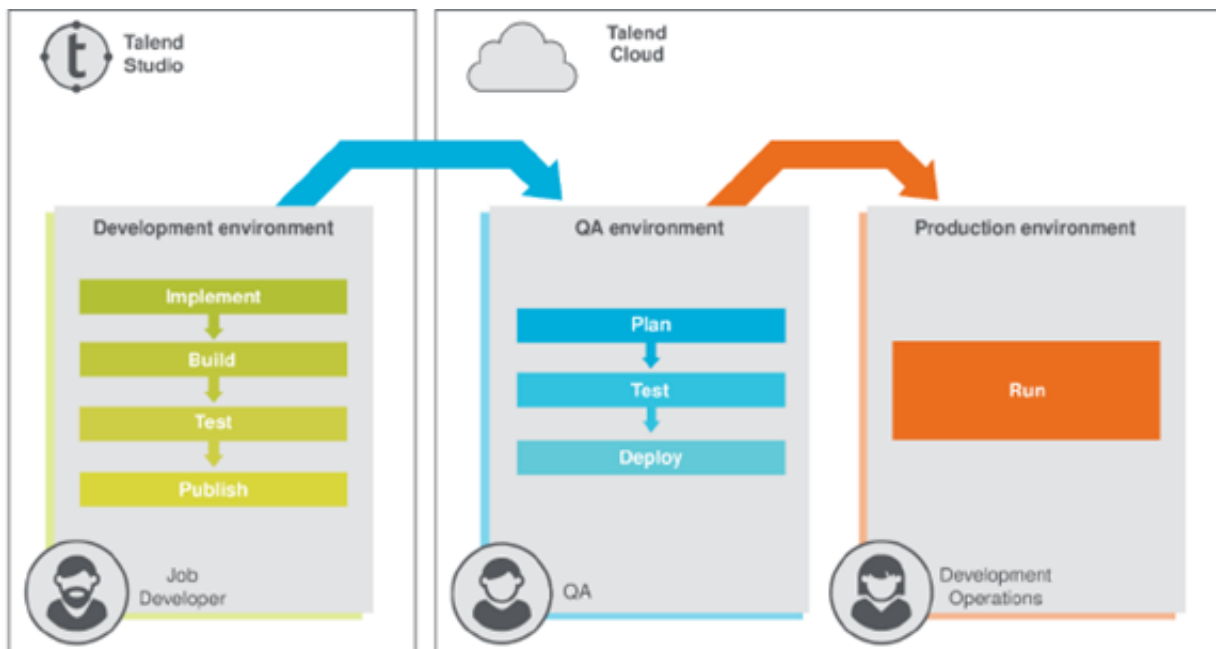
# Talend Development Basics

Last updated by | Archana Balachandran | May 27, 2020 at 5:09 PM EDT

## Contents

- Development Process
- Data Sources
- Contexts, Context Variables, and Context Groups
- Talend-Snowflake Interaction
  - Staging Data in Snowflake
- Labs
  1. How to Create Connections and Context Variables
  2. How to Create a Staging Job

## Development Process

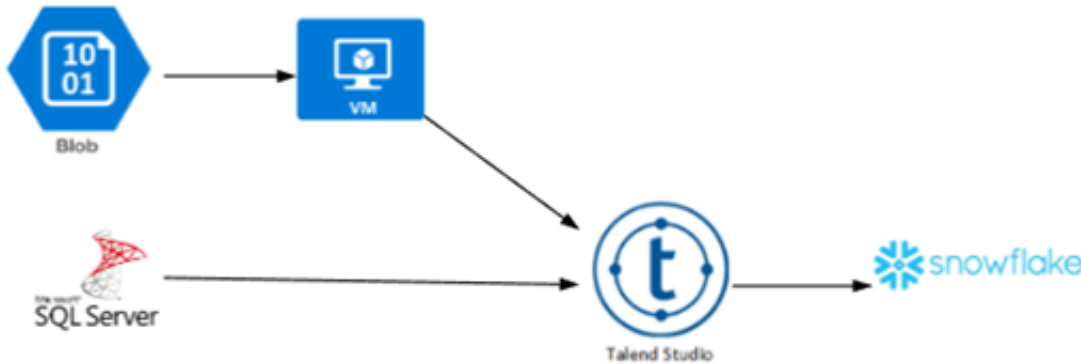


Jobs designed in Talend Studio by Job Developers are published to Talend Cloud QA environment as artifacts.

QA team members then create tasks using these artifacts and run a series of tests.

The artifacts then pass through UAT environment and finally goes into production where its scheduled to run daily.

## Data Sources



Talend loads data from the following sources:

SQL Server: DCOD, COV-SQL, OMA-EDW, ACC-SQL, COV-APP-SQL, COV-EDW

Azure Blob Storage: IS4East, IS4West, Conformance, Midwest, Xref, GoalWorkbook, Sharepoint

## Contexts, Context Variables, and Context Groups

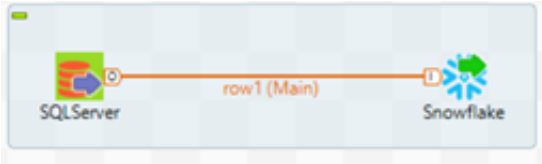
- **Context variable:** A variable which can be set either at compile time or runtime. It can be changed and allows variables which would otherwise be hardcoded to be more dynamic.
- **Context:** The environment or category of the value held by the context variable. Most of the time Contexts are DEV, TEST, PROD, UAT, etc. This allows you to set up one context variable and assign a different value per environment.
- **Context Group:** A group of context variables which are packaged together for ease of use. Context Groups can be dragged and dropped into jobs so that you do not have to set up the same context variables in different jobs. They can also be updated (added to) in one location and then the changes can be distributed to the jobs that use those Context Groups.

Reference: <https://www.talend.com/blog/2019/02/12/best-practices-for-using-context-variables-with-talend-part-1/> [🔗](#)

## Talend-Snowflake Interaction

### Staging Data in Snowflake

How data is uploaded to Snowflake via Talend Below is a simple Talend job design to load data into Snowflake from SQL server



Talend creates temporary files behind the scenes

snowflake > stage > CLAIMFINANCIALALLOCATIONBYEXPENSE > INSERT > 20200225_093930_716_HQTA92_1			
Name	Date modified	Type	Size
stream_20200225_093930_7160.gz	2/25/2020 9:39 AM	GZ File	1 KB

✓	0192d0a6-...	commit	DEV_TALEND	DEV_ETL_XS...			1416050983...	9:26:28 AM	9:26:28 AM
✓ Step 5	0192d0a6-...	RM '@~/DBO_ALLEGATION\INSERT\20200311_132554_169_474XJG_1'	DEV_TALEND	DEV_ETL_XS...			1416050983...	9:26:28 AM	9:26:28 AM
✓ Step 4	0192d0a6-...	INSERT INTO "DEV_STAGING"."CLINICAL_CODING"."DBO_ALLEGATION"("ALLEGATION_CODE";"ALLEGATION_DESCRIPTION";"STA...	DEV_TALEND	DEV_ETL_XS...	1	X-Small	1416050983...	9:26:27 AM	9:26:28 AM
✓ Step 3	0192d0a6-...	COPY INTO "DBO_ALLEGATION_20200311_132554_169_1" FROM '@~/DBO_ALLEGATION\INSERT\20200311_132554_169_474XJG_...	DEV_TALEND	DEV_ETL_XS...	1	X-Small	1416050983...	9:26:25 AM	9:26:27 AM
✓ Step 2	0192d0a6-...	CREATE TEMPORARY TABLE "DBO_ALLEGATION_20200311_132554_169_1" AS SELECT "ALLEGATION_CODE";"ALLEGATION_DES...	DEV_TALEND	DEV_ETL_XS...	1	X-Small	1416050983...	9:26:24 AM	9:26:25 AM
✓ Step 1	0192d0a6-...	PUT file://C:\Users\labalechandrani\AppData\Local\Temp\snowflake\stage\DBO_ALLEGATION\INSERT\20200311_132554_...	DEV_TALEND	DEV_ETL_XS...			1416050983...	9:26:24 AM	9:26:24 AM
✓	0192d0a6-...	show /*_JDBC:DatabaseMetaData.getColumns() */ columns in database "DEV_STAGING"	DEV_TALEND	DEV_ETL_XS...			1416050983...	9:26:00 AM	9:26:23 AM

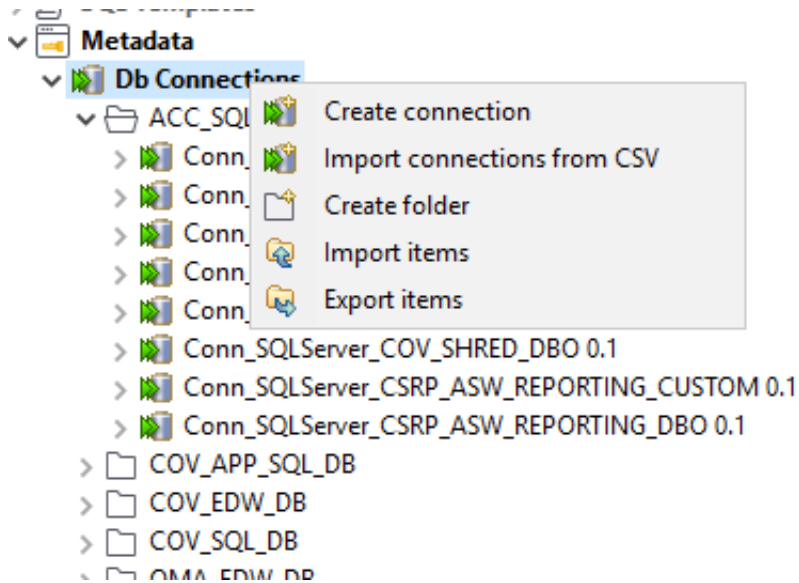
Here's the series of events happening within snowflake:

1. The temporary file created by Talend is then uploaded into a snowflake internal stage area
2. A temporary table with the same structure as the target table is created in Snowflake
3. Data is then loaded into the temporary table.
4. Data from Temporary table is loaded into the target table.
5. The temporary table is then deleted.

## Labs

### 1. How to Create Connections and Context Variables

The lifecycle of a context variable begins when you create a new connection under Metadata and export it as a context. You can create new connections under Metadata section in your repository view, by right clicking the connection type and selecting 'Create Connection'.



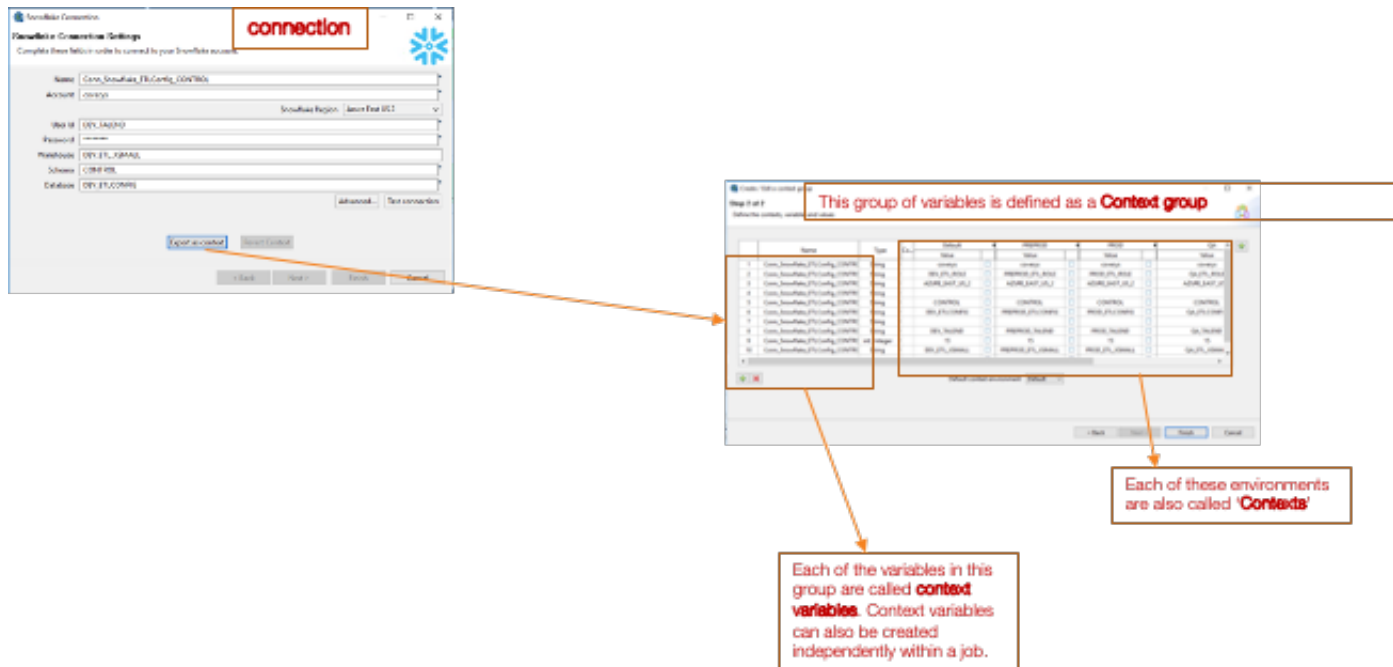
*Naming Standard for connections: Conn<ConnectionType><Database\_name><Schema\_name>\_*

e.g. Conn\_SQLServer\_COV BILLING\_DBO

Once the connection details are entered in the Connection setup window, select 'Test Connection' to validate the connection parameters you entered. Once the test is successful, select 'Export as context' and 'Create a new Context' and complete the steps in the wizard.

*Naming Standard for Context Groups: ConGrp<ConnectionType><Database\_name><Schema\_name>\_*

e.g. ConGrp\_SQLServer\_COV BILLING\_DBO



References: <https://help.talend.com/reader/93oICfmQi615MRwYBjy30g/S6QBejHCyak6a6H9bWWePA>

## 2. How to Create a Staging Job

1. Open a new Job Design.
2. Insert the required input and output components from Palette.
3. For each component, configure the connection and query details in 'Component' tab. Ensure the correct repository connection is used, and also ensure that their corresponding context group is included in the 'Context' tab.
4. Run the job and ensure that the job performs as expected.
5. Apply the supporting control components (PreJob, PostJob, LogProcessing Listener) along with the Job Configuration context group.

Reference: <https://help.talend.com/reader/93olCfmQi615MRwYBjy30g/y9Orb7R3cKh9D52KZsevlw> 