***Steps for Deploying Talend job into Snowflake using Airflow.***

1. **Setting up Snowflake Account:**

* Ensure that you have access to Snowflake and have necessary permissions to create databases, tables, and perform other operations.
* Obtain the necessary credentials (username, password, account name, etc.) to connect to Snowflake.

1. **Setting up Talend Job**

**Install Talend Open Studio:**

* Install Talend Studio on your local machine where you plan to develop your Talend job.

**Develop Talend Job:**

* Create and test your Talend job that interacts with Snowflake.
* Ensure your Talend job can be executed from the command line.
* Develop Talend job that performs the desired operations on data within Snowflake.

**Export/Build Talend Job:**

* Once your job is developed, export it from Talend Studio as a standalone Job.
* Make sure job-name\_run.sh file is present in the extract the same we used while we are writing DAG script of Airflow.

1. **Setting up Apache Airflow on AWS EC2:**

* Launch an EC2 instance on AWS with proper role and required packages to operate Talend, Python operator, snowflake, and airflow as follows.
* Install openjdk-8-jdk since Talend requires.
* Install unzip to extract build job-name.zip from Talend.
* Install virtual machine to configure Airflow to use an external metadata database (e.g., PostgreSQL, MySQL)
* Install Apache-airflow [Postgres]==2.6.0
* Set user Id and password for login to airflow using public Ip address of ec2.
* Install python3.
* Install Apache-airflow-providers-snowflake. ...etc.
* Create S3bucket give full accesses for the role you created while launching ec2.
* Create log folder in S3 bucket: where you capture all logs of the packages.
* Create code folder in S3 bucket: where you kept your DAG and job-name.zip.

1. **Configuring Airflow Connections:**

* Launch Airflow using Ip address of ec2.
* Configure Snowflake connection in Airflow's Connection settings with the Snowflake credentials obtained earlier.
* Configure other necessary connections for AWS (e.g., S3 staging)

1. **Creating an Airflow DAG:**

* Define an Airflow Directed Acyclic Graph (DAG) that includes tasks for executing your Talend job.
* Use the Bash Operator or Python Operator to execute your Talend job as a task within the DAG.
* Define any necessary dependencies between tasks.

1. **Start Airflow Schedular:** Ensure that the Airflow scheduler is running so that it can schedule and execute your DAGs according to the specified schedule.
2. **Monitor Execution:** Monitor the execution of your DAG in the Airflow UI. Check logs for any errors or issues that may arise during execution.
3. **Maintenance and Update:** Regularly maintain and update your Airflow environment, DAG scripts, and Talend jobs to ensure they continue to run smoothly and efficiently.