

BP TD - Bollate - Raw Data

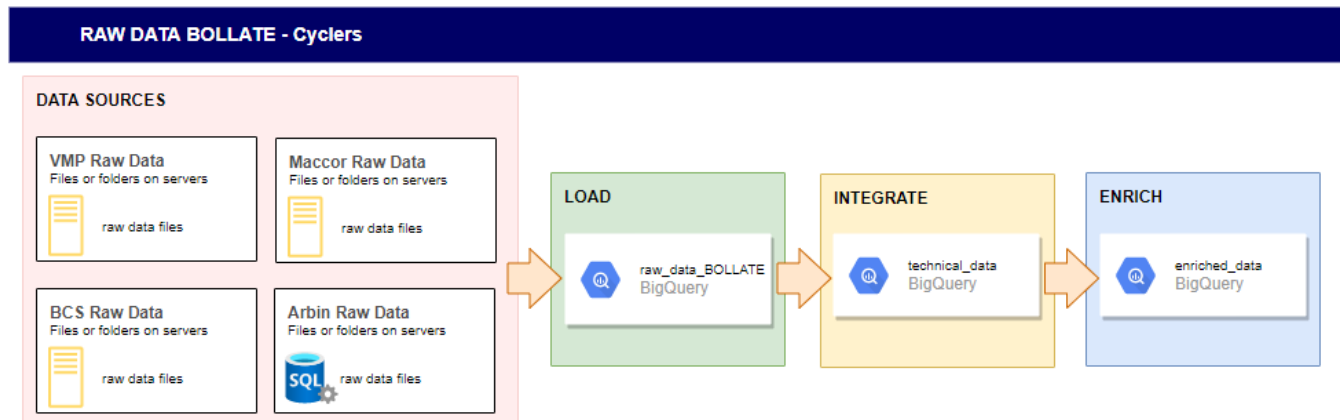
Technical documentation for the Bollate raw data load and integration.

Summary

- [Sum-up](#)
- [Macro Schema](#)
- [Data sources](#)
- [Load and Integration](#)
 - [Schema](#)
 - [Integration and Computing](#)
- [Talend Project](#)
 - [Rnl_Battery_Bollate](#)
 - [Talend jobs](#)

Sum-up

Macro Schema



Schemas

All schemas can be found on the following GDrive link:

#133 Battery Data management platform >> 06 - Documentation >> [Architecture](#) >> [Battery Data Scheme and Diagrams](#)

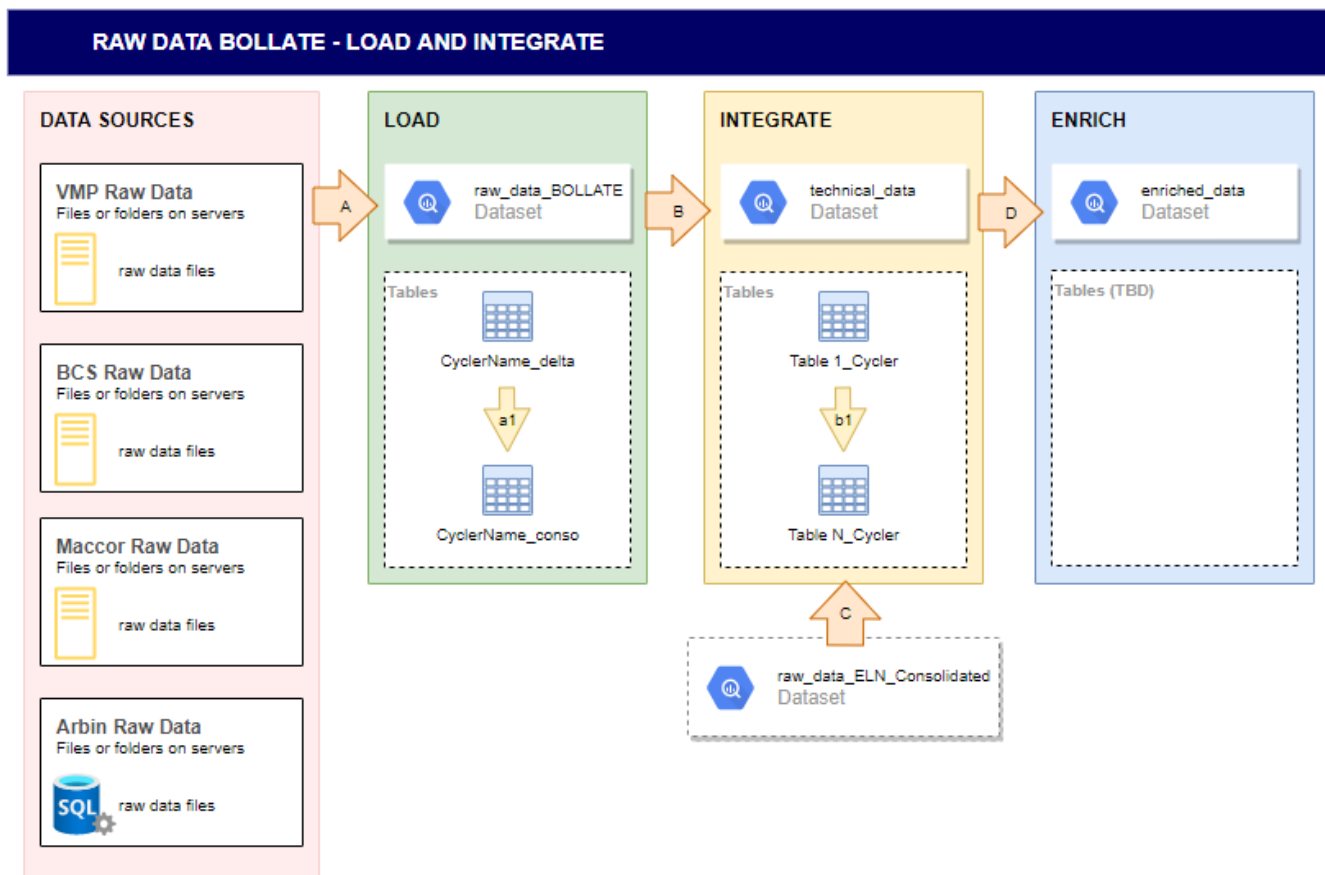
Data sources

Raw data files have different structures and are extracted from different the servers:

Cyclers	Description	Data Status	Flow Status
Maccor	CSV Raw files coming from Maccor machine. The source server is located in Italy	HOT	ACTIVE
Biologic BCS	CSV Raw files coming from Biologic BCS machine. The source server is located in Italy	HOT	ACTIVE
Biologic VMP	CSV Raw files coming from Biologic VMP machine. The source server is located in Italy	HOT	ACTIVE
Arbin	Data coming from SQL Server database? for Arbin machine	HOT	ON HOLD

Load and Integration

Schema



N	Description
A	A Talend job is responsible for getting all files from the sources and bringing them to the Talend server to finally load the raw data to BigQuery.
a1	All new raw data is loaded in a "delta" table. A SQL script is responsible to consolidate the new data to the already existing. Every cyclers has its own tables.
B	Once the data is consolidated, it's taken to another dataset where data is transformed, enriched and integrated on the same table with other Bollate's cycles.
b1	Data is transformed, enriched and integrated passing through many tables. There is one or many SQL scripts to do this task.
C	Data coming from ELN is integrated at this point. Check ELN page for more details.
D	Data is pushed to the enriched_data dataset where it can be accessed by the users.

SQL Scripts

All SQL scripts can be found:

- Project Saved Queries
- Git?
- GCP Cloud Storage

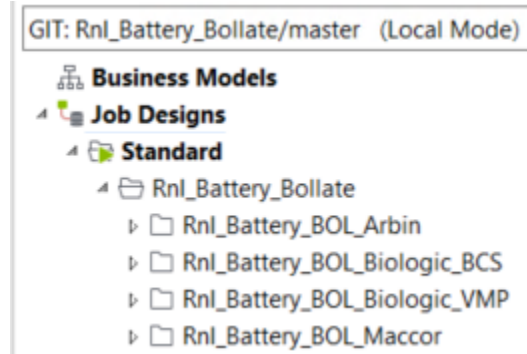
Integration and Computing

Data mapping to have the same table structure for all the cyclers files:

Talend Project

Rnl_Battery_Bollate

This project holds all jobs necessary to load all Battery Bollate raw data



Talend jobs