



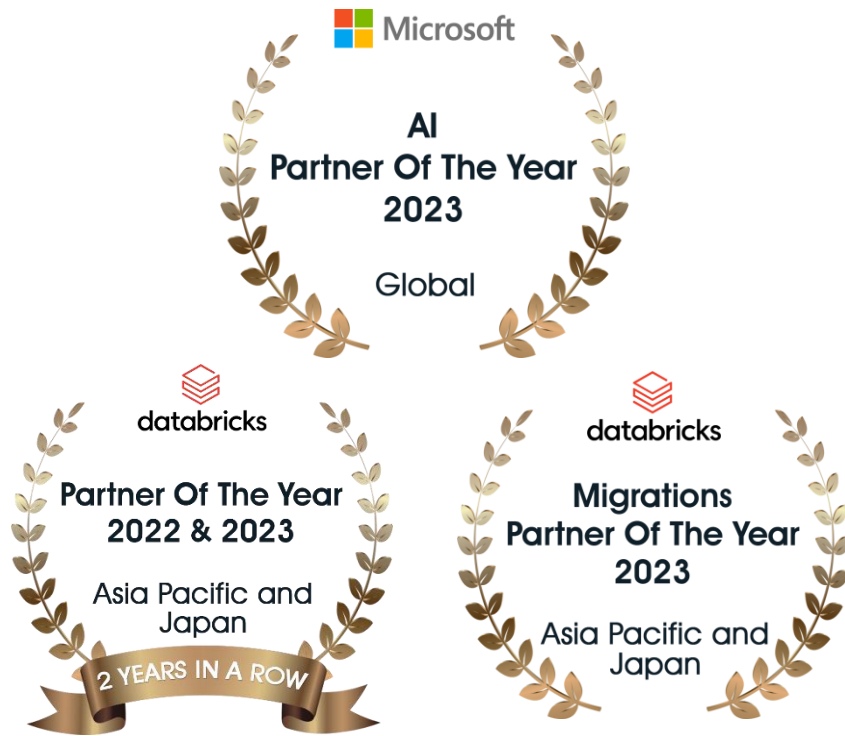
Celebal-SAP Data Ingestion Accelerator



Infrastructure (Azure)
Data & AI (Azure)
Digital & App Innovation (Azure)
Security
Biz Applications



Celebal Specialization and Strength



Partnerships



Advanced Specialization



- AI & Machine Learning
- Analytics
- Infra and Database Migration
- Kubernetes
- Cloud Security
- Low Code No Code
- Intelligent Automation



INDIA | USA | CANADA | APJ | MIDDLE EAST | AUS

2800+
Employees

800+
AI experienced
professionals

500+
AI Certifications



Industries We Serve



Manufacturing



Retail & CPG



Financial
Services



Energy &
Sustainability

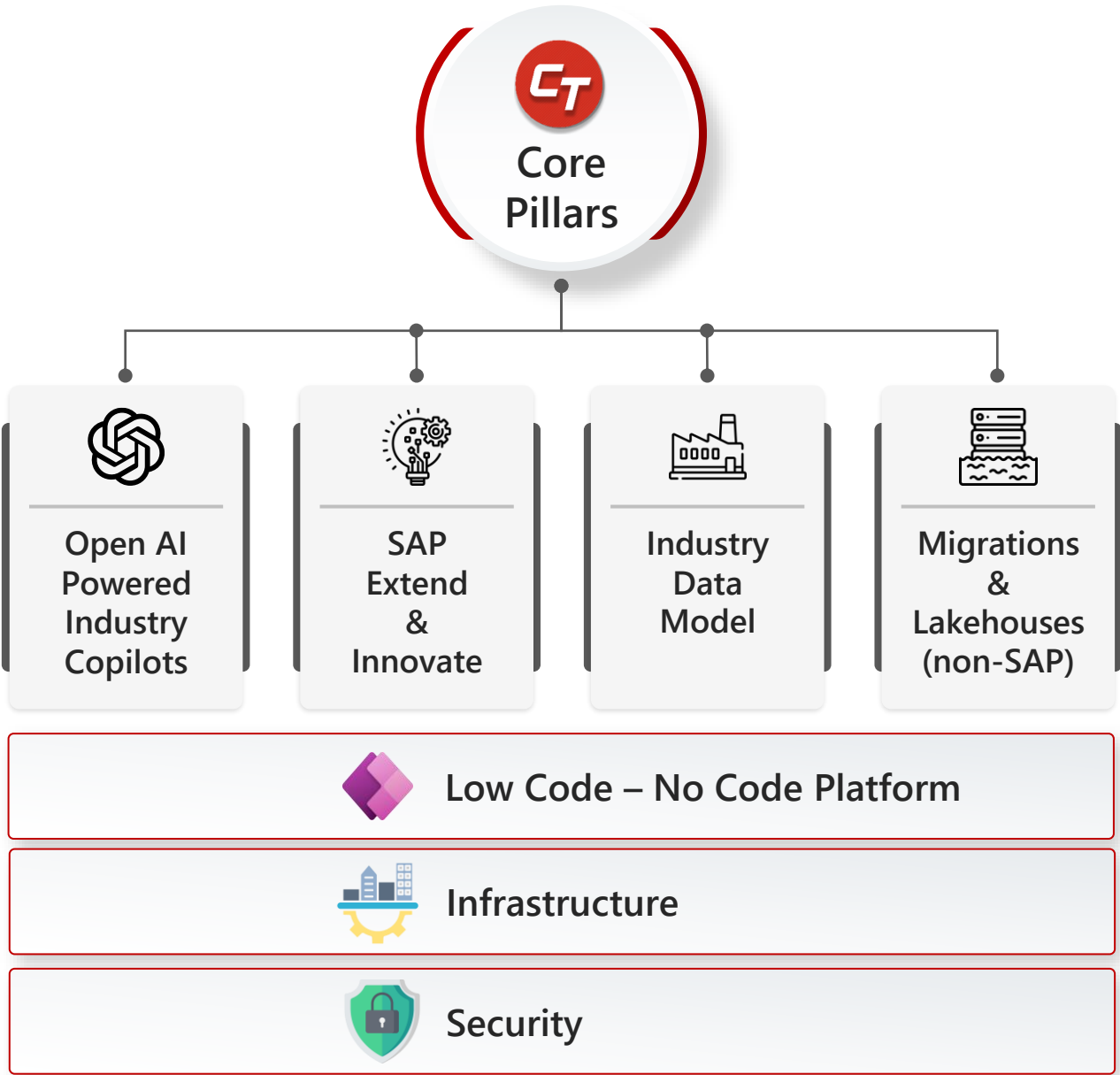


Healthcare &
Life Sciences



Media &
Entertainment

Celebal Core Pillars



Purpose: To build a framework that accelerates the Data Ingestion process and establishes a multi-hop architecture. Automation is currently limited to the Silver Layer. The primary goal of this framework is to address the following challenges:

Problems:

1. Manual Creation of ADF pipelines for full load and delta load
2. Involvement of multiple developers in the above process
3. Maintenance of full load and delta load pipelines
4. Lack of automation
5. Lack of standardization
6. Parallel execution of different workloads leads to deadlock issues
7. No user visibility into the ingestion process
8. Databricks notebooks for silver layer creation are manually developed for each base SAP table

The SAP Data Ingestion Accelerator is a tool that automates data ingestion from source to destination. It features a user-friendly interface that provides customers with enhanced transparency and control, allowing them to ingest specific tables from any module within their SAP system.

User-Friendly Design

- Leverage pipeline monitoring, debugging, and scheduling capabilities from ADF and Synapse Analytics.
- Seamlessly toggle between your source, destination, and ELT platform.
- Enable full load Ingestion functionality.
- Enable delta load ingestion functionality.
- Monitor the last delta/full load update date for ingested tables.
- Configure full load settings to handle large-volume tables (e.g., BSEG, MSEG, ACDOCA, etc.)
- Gain transparency into ingested tables and identify additional tables pending ingestion.

Automated and Unified Delta Load and Full Load Ingestion Process

- Unified data orchestration process.
- Unified transformation process for ingesting data from Bronze to Silver layer.
- Automatically generates RFC expressions based on delta conditions via the Delta Control Table.
- Pre-built and customizable control tables, tailored to project needs, for defining optimal distribution on Synapse tables, partitioning options and columns (ADF partitioning on the source side), and delta load conditions for header and item tables.

Automated Creation of Bronze and Silver Layer

- Automated ingestion of raw data into the Bronze Layer (Data Lake) following a predefined hierarchy.
- Automated creation of the Silver Layer by fetching metadata from the SAP Landscape (DD03L) for user-selected tables.
- Automated handling of schema drift from the source side.
- Automated upsert operations in Silver layer tables.
- No need to manually create data securable layers on DBX/Synapse for Bronze to Silver transitions.

Value for Customer

- **Reduced time consumption in the data ingestion process.**
- **Clean Databricks/Synapse environment enabled by unified scripts for data transformation and delta operations.**
- **Greater focus on data visualization.**
- **Transparency and improved ease of use.**

Source



Data Discovery

S/4 HANA

Table Connector

Search..



☐ Materials Management

☐ Human Resource

☐ Quality Management

☐ Plant Maintenance

☐ Production Planning

☐ Financial Accounting

☐ Sales and Distribution

Full Load

Delta Load

Destination



Added Tables

Databricks

Search..



PRESTG_TC_s4Hana

STG_TC_s4Hana

Silver_TC_s4Hana













Summary



Data Factory Pipeline

Synapse pipeline

Fetching Data

	Source	Destination	Pipeline Name	Action
	SAP_s4_HANA	synapse	PL_Service_TC_s4Hana_dev	
	SAP_s4_HANA	synapse	PL_Large_table_FullLoad_TC_s4Hana_dev	
	SAP_s4_HANA	synapse	pl_DependentDelta_TC_s4Hana_dev	
	SAP_s4_HANA	synapse	PL_Installation_TC_s4Hana_dev	
	SAP_s4_HANA	synapse	PL_FullLoad_TC_s4Hana_dev	
	SAP_s4_HANA	synapse	PL_DailyProcessComposite_TC_s4Hana_dev	



Source

Data Discovery

S/4 HANA

Table Connector

Table Connector

CDC With SLT

CDC w/o SLT

☐ Human Resource

☐ Quality Management

☐ Plant Maintenance

☐ Production Planning

☐ Financial Accounting

☐ Sales and Distribution

Full Load

Delta Load

Destination

Added Tables

One Lake

Search...

PRESTG_TC_s4Hana

STG_TC_s4Hana

Silver_TC_s4Hana

Summary

Data Factory Pipeline

Synapse pipeline

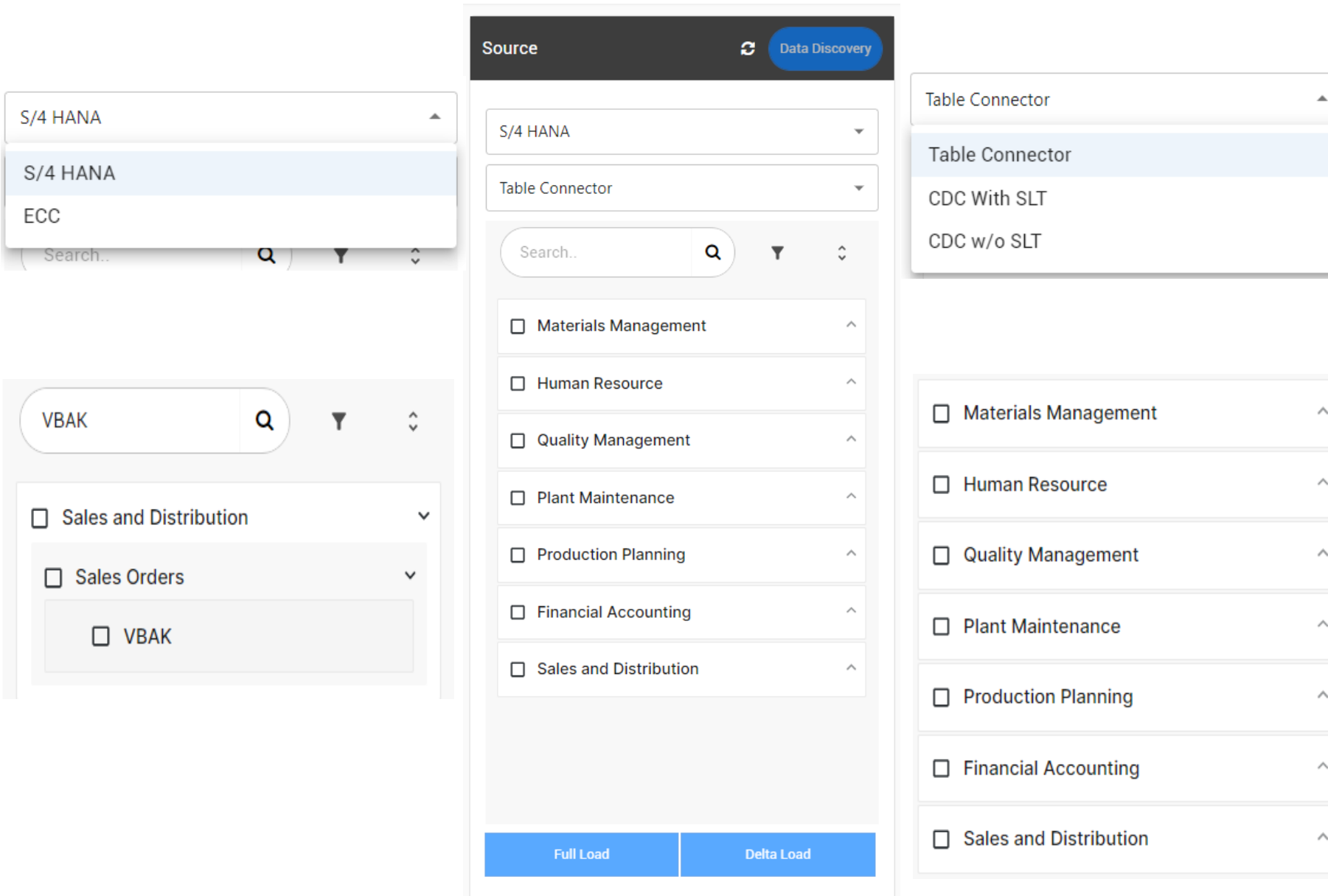
Fetching Data

	Source	Destination	Pipeline Name	Action
	SAP_s4_HANA	synapse	PL_Service_TC_s4Hana_dev	
	SAP_s4_HANA	synapse	PL_Large_table_FullLoad_TC_s4Hana_dev	
	SAP_s4_HANA	synapse	pL_DependentDelta_TC_s4Hana_dev	
	SAP_s4_HANA	synapse	PL_Installation_TC_s4Hana_dev	
	SAP_s4_HANA	synapse	PL_FullLoad_TC_s4Hana_dev	
	SAP_s4_HANA	synapse	PL_DailyProcessComposite_TC_s4Hana_dev	

« Previous

1

Next »



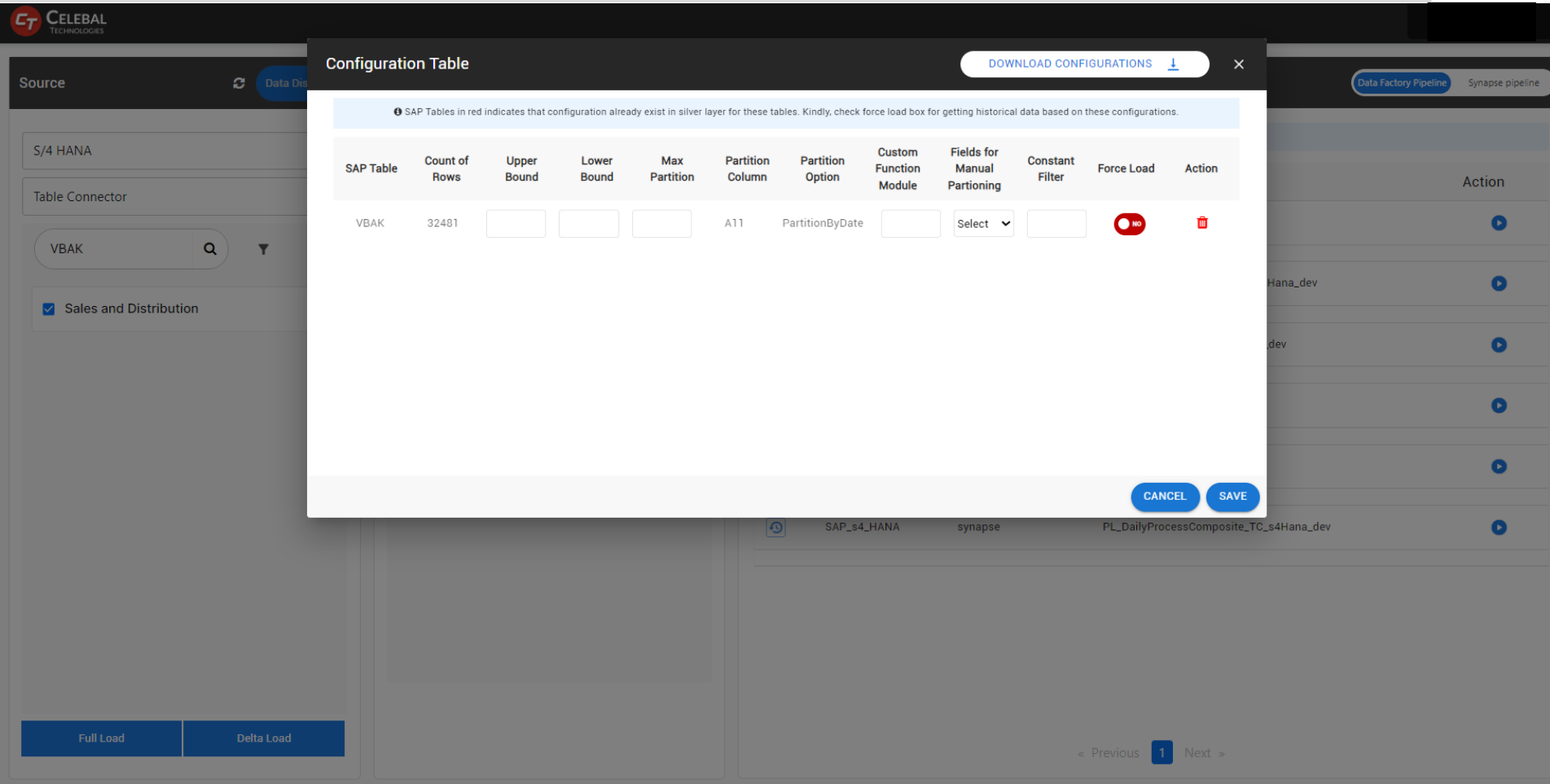
The screenshot displays the 'Source Overview' interface, which is used for configuring data sources. It features three main panels:

- Left Panel:** A search bar with 'VBAK' entered. Below it, a list of modules is shown, including 'Sales and Distribution' and 'Sales Orders', with 'VBAK' selected under 'Sales Orders'.
- Center Panel:** A 'Source' dropdown menu with 'S/4 HANA' selected. Below it, a 'Table Connector' dropdown menu is open, showing options like 'Table Connector', 'CDC With SLT', and 'CDC w/o SLT'. A search bar and filter icons are also present.
- Right Panel:** A list of modules with checkboxes, including 'Materials Management', 'Human Resource', 'Quality Management', 'Plant Maintenance', 'Production Planning', 'Financial Accounting', and 'Sales and Distribution'.

At the bottom of the center panel, there are two buttons: 'Full Load' and 'Delta Load'.

- **Source Switch:** Switch between S/4 Hana or SAP ECC as the source.
- **Connector Switch:** Switch between ADF native connectors to connect with your source.
- **Modules:** Modules contain sub-types, which are followed by base tables.

Full Load Configuration



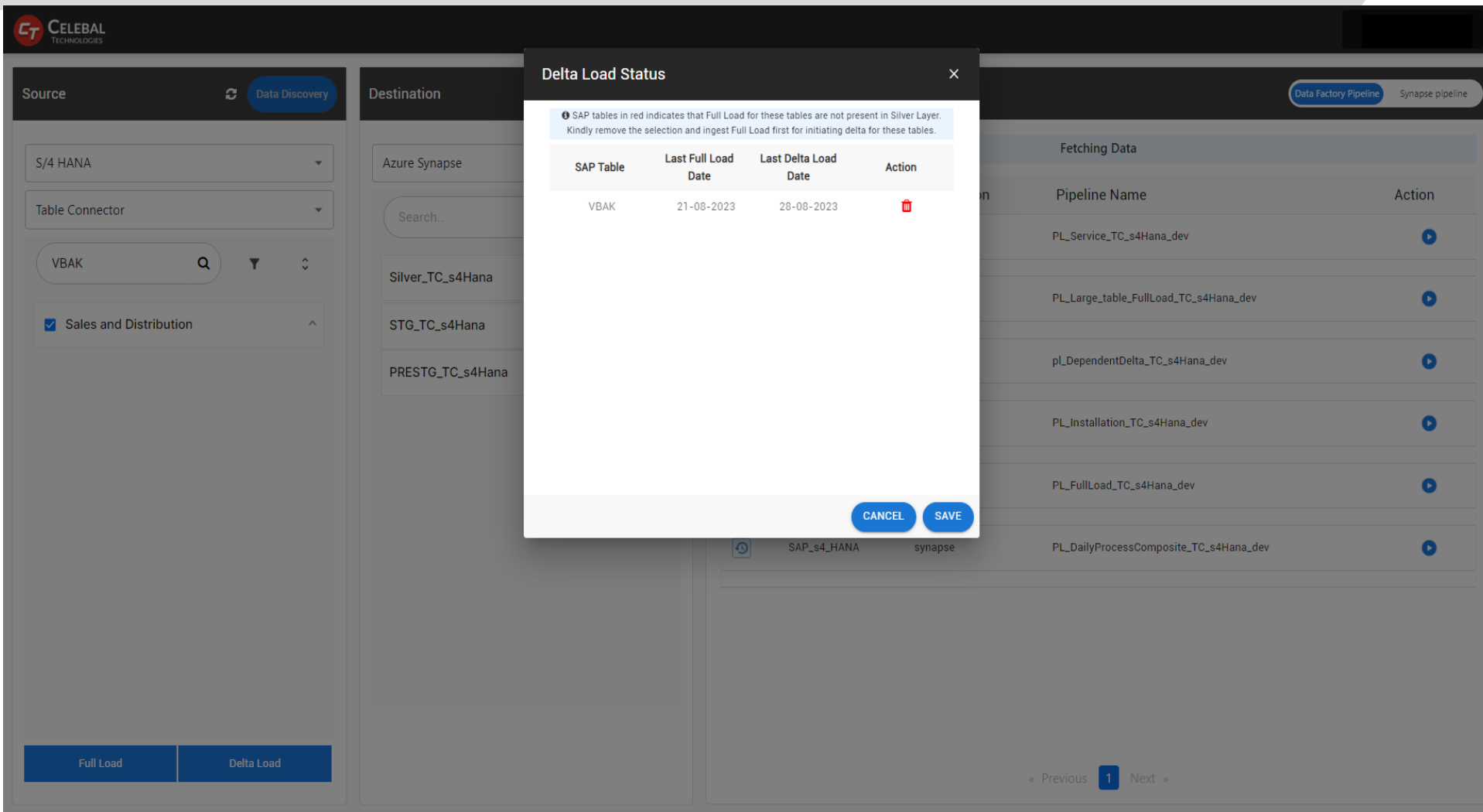
The screenshot shows the 'Configuration Table' dialog box in the CELEBAL interface. The dialog has a title bar with 'Configuration Table' and a 'DOWNLOAD CONFIGURATIONS' button. Below the title bar is a message: 'SAP Tables in red indicates that configuration already exist in silver layer for these tables. Kindly, check force load box for getting historical data based on these configurations.' The main table has the following columns: SAP Table, Count of Rows, Upper Bound, Lower Bound, Max Partition, Partition Column, Partition Option, Custom Function Module, Fields for Manual Partitioning, Constant Filter, Force Load, and Action. The table contains one row for 'VBAK' with a count of 32481. The 'Force Load' column has a red toggle switch. The 'Action' column has a red trash icon. At the bottom of the dialog are 'CANCEL' and 'SAVE' buttons. The background shows the 'Table Connector' section with 'S/4 HANA' and 'VBAK' selected, and a 'Full Load' button.

SAP Table	Count of Rows	Upper Bound	Lower Bound	Max Partition	Partition Column	Partition Option	Custom Function Module	Fields for Manual Partitioning	Constant Filter	Force Load	Action
VBAK	32481				A11	PartitionByDate		Select		<input checked="" type="checkbox"/>	


- **Upper Bound:** Upper limit of the partition column.
- **Lower Bound:** Lower limit of the partition column.
- **Custom Functional Module:** Pull data based on a user-defined custom function module.
- **Fields for Manual Partitioning:** Select fields for advanced partitioning.
- **Constant Filter:** Filter the dataset (e.g., BUKRS EQ '10')
- **Force Load:** Enables a check that warns the user if a similar configuration already exists in the Silver layer for the selected table. If the user still wants to override the data, they can toggle this option to "Yes."
- **Action:** Remove unwanted tables selected due to human error.
- **Download Configuration:** Export configurations to Excel.

Advanced partitioning: A method built into SAP DIA to handle large datasets with more than 10 million records. The user must identify one or more fields from the selected table that can divide the large dataset into approximately uniform row counts based on the distinct values in those fields. These fields are specified in the Fields for Manual Partitioning cell of the configuration table.

Delta load Status



The screenshot shows the CELEBAL interface with a 'Delta Load Status' dialog box open. The dialog box contains a table with the following data:

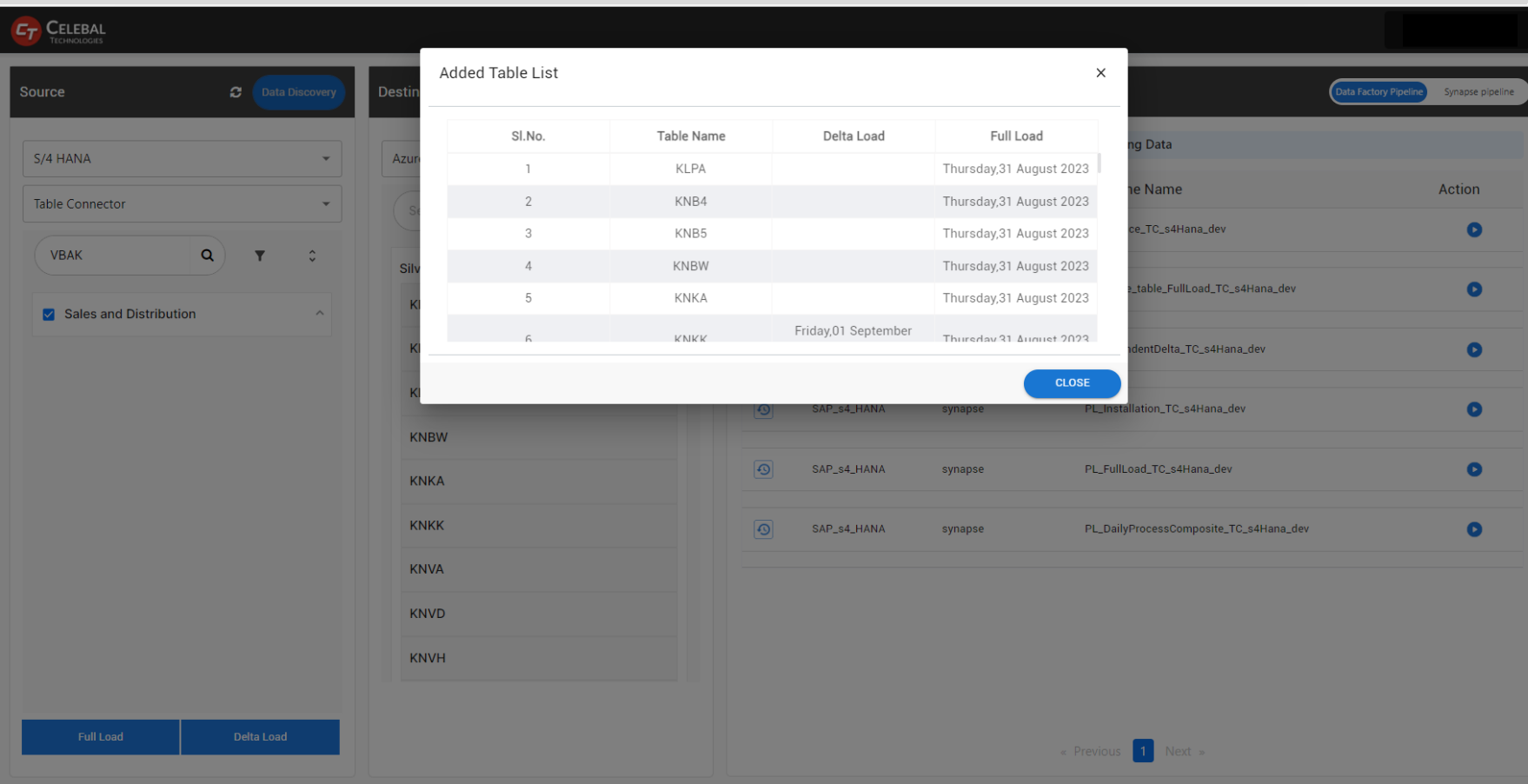
SAP Table	Last Full Load Date	Last Delta Load Date	Action
VBAK	21-08-2023	28-08-2023	

Below the table are 'CANCEL' and 'SAVE' buttons. A message at the top of the dialog states: 'SAP tables in red indicates that Full Load for these tables are not present in Silver Layer. Kindly remove the selection and ingest Full Load first for initiating delta for these tables.'

The background interface shows the 'Source' section with 'S/4 HANA' and 'Table Connector' selected, and the 'Destination' section with 'Azure Synapse' selected. A list of SAP tables is visible, including 'Silver_TC_s4Hana', 'STG_TC_s4Hana', and 'PRESTG_TC_s4Hana'. At the bottom, there are 'Full Load' and 'Delta Load' buttons.

- **Last Full Load Date:** The date on which the last full load was ingested for the selected table.
- **Last Delta Load Date:** The date on which the last delta load was ingested for the selected table.
- **Tables Highlighted in Red:** A feature in SAP DIA that highlights tables in red if they do not have a full load present in the Silver layer. To initiate a delta load, the user must first perform an initial/full load for that table in the Silver layer.
- **Action:** Remove unwanted tables that were selected due to human error.

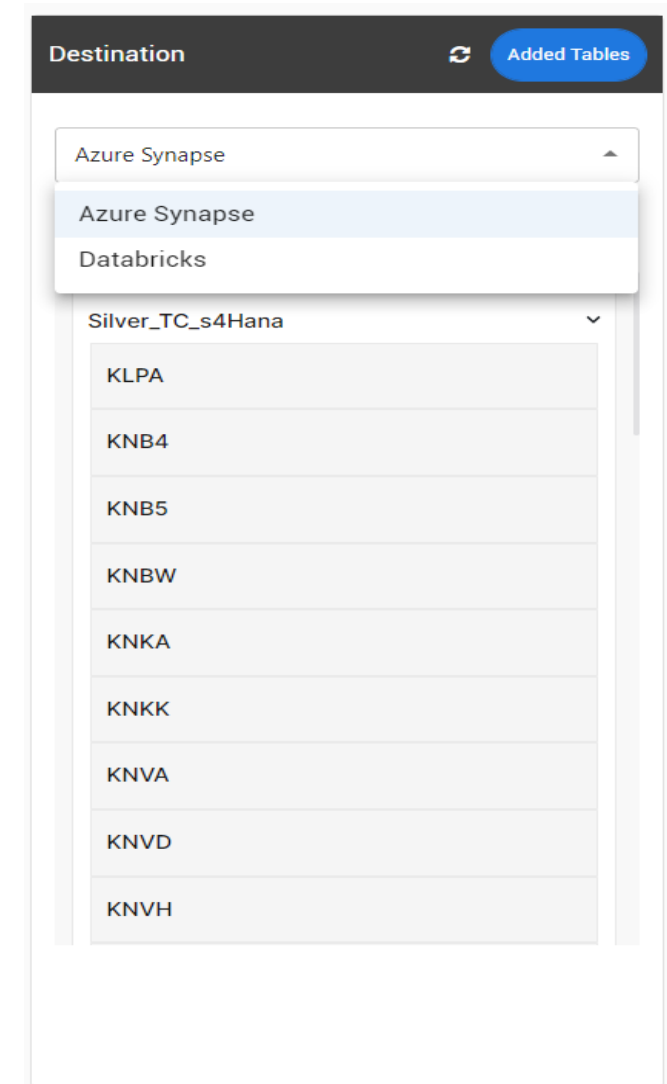
Destination Overview



The screenshot shows the CELEBAL interface with a modal titled "Added Table List" in the foreground. The modal contains a table with the following data:

Sl.No.	Table Name	Delta Load	Full Load
1	KLPA		Thursday,31 August 2023
2	KNB4		Thursday,31 August 2023
3	KNB5		Thursday,31 August 2023
4	KNBW		Thursday,31 August 2023
5	KNKA		Thursday,31 August 2023
6	KNKK	Friday,01 September	Thursday,31 August 2023

The background shows the "Destination Overview" page with a "Source" dropdown set to "S/4 HANA", a "Table Connector" dropdown, and a "Sales and Distribution" checkbox. The "Destination" dropdown is set to "Azure Synapse". The "Data Discovery" button is visible. The "Full Load" and "Delta Load" buttons are at the bottom.



The screenshot shows the "Destination Overview" page with the "Destination" dropdown set to "Azure Synapse". The "Added Tables" button is visible. The list of tables is displayed below:













- Azure Synapse
 - Azure Synapse
 - Databricks
- Silver_TC_s4Hana
 - KLPA
 - KNB4
 - KNB5
 - KNBW
 - KNKA
 - KNKK
 - KNVA
 - KNVD
 - KNVH

- The "Added Table List" provides a status check of all tables ingested into your silver layer, showing their last full load date and the most recent delta load update.
- Switch between your destinations to view the underlying list, which displays the schemas present in each destination along with the existing tables within each schema.

Summary Overview

Summary   Data Factory Pipeline Synapse pipeline

Fetching Data

Source	Destination	Pipeline Name	Action
 SAP_s4_HANA	synapse	PL_Service_TC_s4Hana_dev	 <div>Run Now Schedule Pipeline</div>
History			
Run Start ↑		Duration	Status
6/9/2023, 3:48:34 pm		00:02:02	Succeeded
 SAP_s4_HANA	synapse	PL_Large_table_FullLoad_TC_s4Hana_dev	
 SAP_s4_HANA	synapse	pl_DependentDelta_TC_s4Hana_dev	
 SAP_s4_HANA	synapse	PL_Installation_TC_s4Hana_dev	
 SAP_s4_HANA	synapse	PL_FullLoad_TC_s4Hana_dev	
 SAP_s4_HANA	synapse	PL_DailyProcessComposite_TC_s4Hana_dev	

« Previous 1 Next »

- **ELT Platform Toggle:** Switch between Data Factory and Synapse pipelines based on the availability of your existing ELT resources within your Azure subscription.
- **Action:** Trigger or schedule your data orchestration pipeline.
- SAP DIA also displays logs for pipelines that are in progress, cancelled, failed, or succeeded.
- Users can also filter historical logs using the filter button at the top left, after the Summary section.



CELEBAL
TECHNOLOGIES

Thank You