# *First Part: Process Discovery*

Every workflow is supported by IT systems like SAP, Oracle or Salesforce, etc. Celonis detects digital footprints(CaseID, Activity, Timestamp) and reconstructs digital workflow traces. Using celonis we can visualize the actual process flow in real-time.

Celonis business intelligent business cloud(IBC) comprises the following range of applications:

1. Process Analytics
2. Action Engine
3. Transformation center
4. Intelligent Business Appstore
5. Event collection
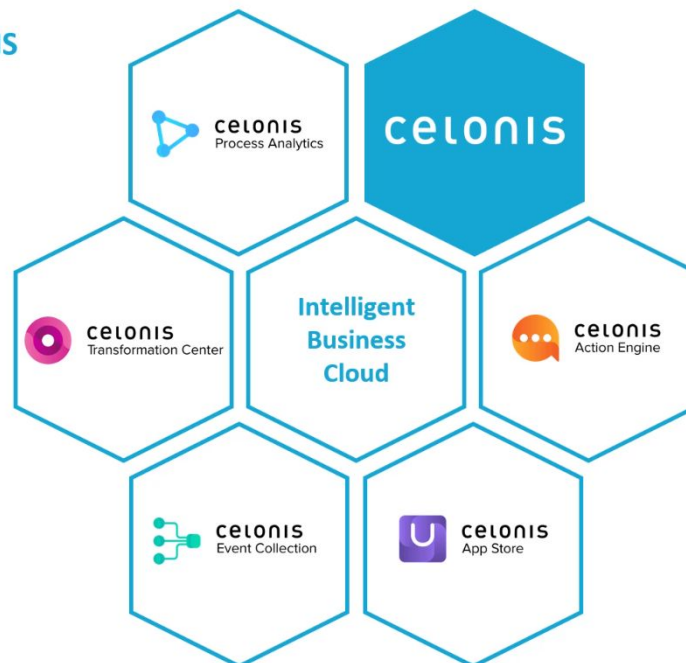
Process Analytics is the place where all the analyses are stored. Together with Process discovery and process-visualizations show you what's happening in your processes.

Action Engine helps you to involve everyone in your organization in the transformation journey. Action Engine operationalizes insights by proactively providing your business users with personalized recommendations for action. As an example, assume that you want to optimize the

cash discount realization in your Accounts Payable process. Action Engine can continuously analyze your open invoices, detect the ones with advancing discount expiry, notify the responsible business users and allow them to directly trigger the payment.

Transformation center gives you the tools to measure and monitor your progress toward achieving KPIs and business outcomes as you improve your processes over time.

Intelligent Business Appstore gathers the best practices and knowledge from celonis customers and partners which is bundled into apps. It provides you with configurations for all IBC Services to pursue your Business Transformation journey. The main benefits of the celonis app are to learn from best-in-class companies, deploy powerful solutions instantly with pre-built connectors and analyses, measure progress with transformation centers KPIs and power celonis action engine with pre-built skills.

Event Collection is a flexible development platform for integration and transformation of data from different sources into the intelligent Business cloud. Its main benefits are easy to set up, multi-functional where you can connect SAP and other systems, scalable and flexible infrastructure.

Understanding the Variant Explorer:

The variant explorer gives a good overview of the different variants within a process. A variant is an end-to-end path through the activities of a process. The initial path in the variant explorer shows the most common variant and within variant explorer, additional variants can be added.

Filter in components:

To filter the process you can create selections directly in your analysis. You can filter from any chart, table, explorer or selection components. When filtering, the entire analysis is instantly updated to this selection. It is important to delete all the filters before you start a new analysis.

Second Part: Analysis Building Basics

In the second part, it is important to have an idea or understanding of basic SAP table data such as P2P. For a great overview of SAP tables and their descriptions, we can always go to this link [Search SAP tables]. So, when building analyses from scratch it is crucial to know how to configure dimension, KPI and sorting. Dimension is which attributes to be displayed and KPI

such as case count, activity count, and throughput time calculation, etc. Sorting in the activity table is of integer data type and whenever two events of the same exact timestamp, the lower sorting number appears first in the process. We can add and configure tables and charts just by click on the add component button in the workspace analysis.

Third Part: Process connection

There are primarily four tasks for Data engineer
1. Connect
2. Extract
3. Transform
4. Set up a data model

The main structural element of Event collection is the Data Pool. Data Pools cluster data connections, data jobs, and schedules equipping you with everything you need to set up a data integration workflow. Below is the general structure for event collection:
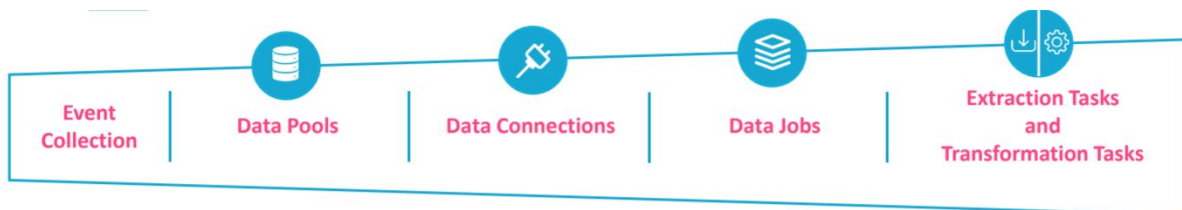
Connect to a database or source system:

Data connection enables you to connect to a specific source system(e.g SAP ECC, Salesforce) and to extract data from it.   The following data connection types are available: *1. Database(JDBC)   2. Oracle EBC 3. SalesForce 4. SAP Ariba 5.SAP Hybrid Cloud for Customers 6. SAP ECC 7. ServiceNow.* Connections are always secured via HTTPS.

You can upload excel or CSV files in the 'File Upload' section. The data types of the column will be derived based on the first rows, but mostly you need to specify it manually. The file upload limit is 1GB per file. Below is an example of setting up a data connection to an MS SQL database.

There are two basics implementation infrastructures scenario in the IBC:
1. On-premise system
2. Cloud system

The following steps must be conducted when setting up an on-premise connector:

1. Prepare the source system.
2. Whitelist firewall IPs and ports.
3. Set up an on-premise connector server.
4. Prepare the uplink connection in the IBC.
5. Install an on-premise connector on the connector server.
6. Configure connection to on-premise system.

Data Jobs -: Data jobs are used to extract data from different connected source systems and then transform the data to prepare it for process mining. You need to set up data jobs after you set up your data connection successfully.

You can create a new data job by clicking on the 'new data job' button your Data Pool while choosing the appropriate Data Connection and choose the desired data connection.
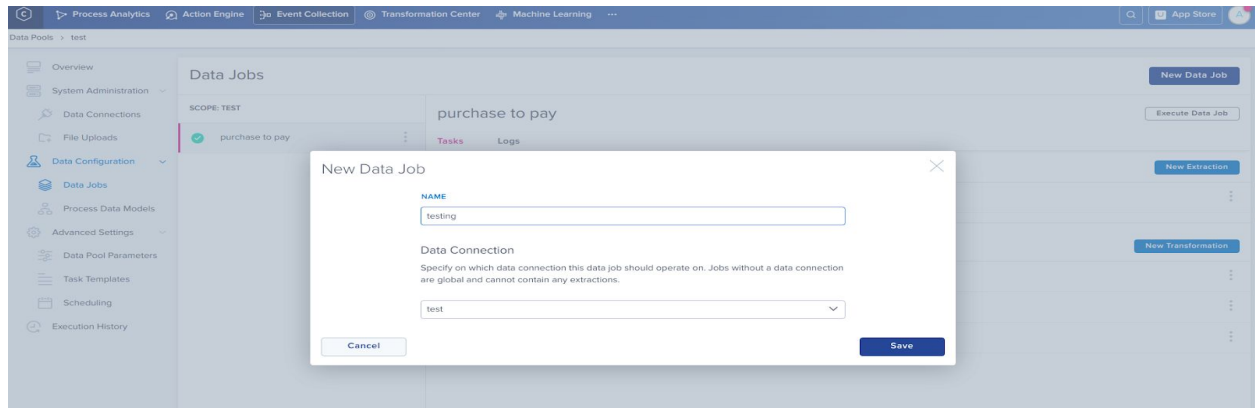
For **Data extraction**, we can add tables that are provided by the data connection. In the celonis intelligent business cloud, the SQL code has to be written in [Vertica SQL]**.**

Transformation -: Transformations are used to **create event logs** from the extracted data which are used by other cloud applications like Process Analytics. They are clean up, restructure and process data to be suitable as additions to the data model. Transformations are written in SQL code.

An example of Transformation in Celonis intelligent business cloud as below:
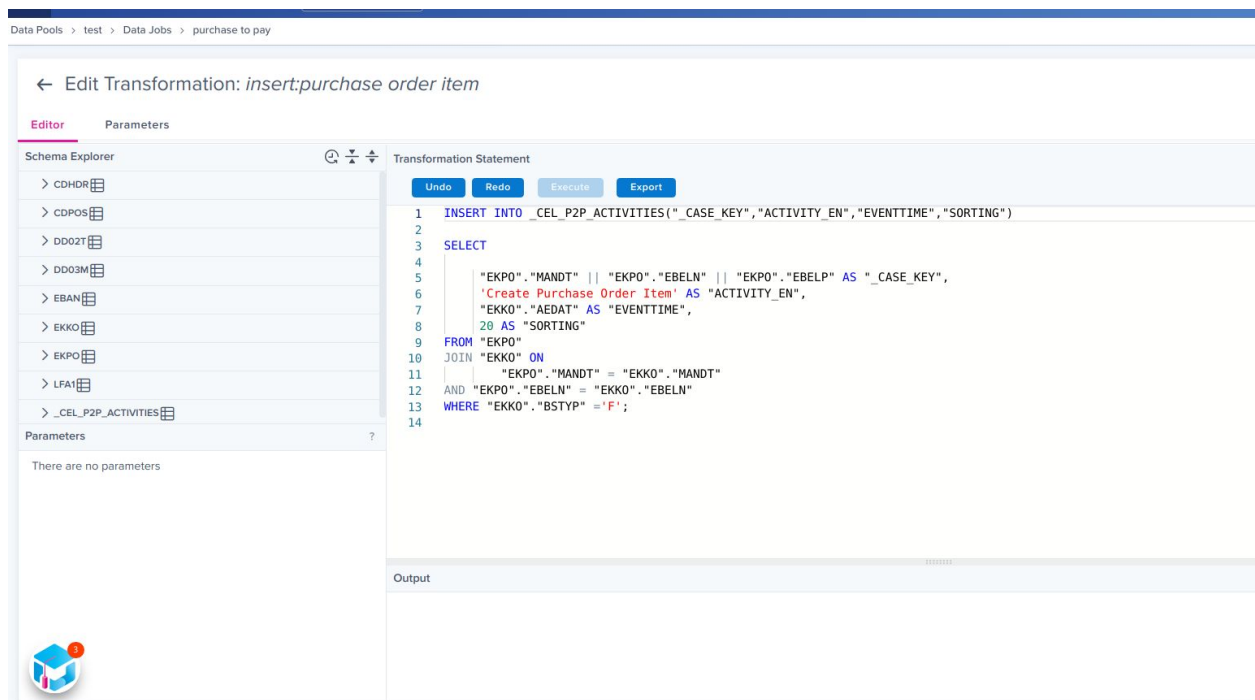
Set up process models -: In this step, we can set up one or multiple process data models and these can be used by other applications like process analytics or action engine. In order to create a process model, we have to load the new extracted and transformed data from the data pool into a data model. Once we set up and loaded our data model we can directly use it for our first analysis by navigating to the process analytics. Hence, in process analytic application we can create a new workspace and select our process model for analysis.

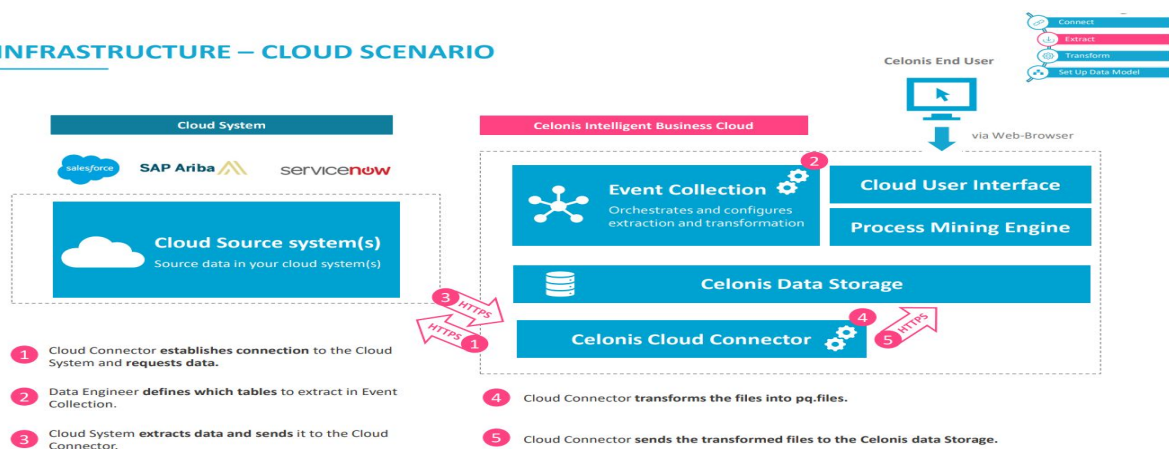## Extraction logic behind on-premise and cloud system:

1: On-premise
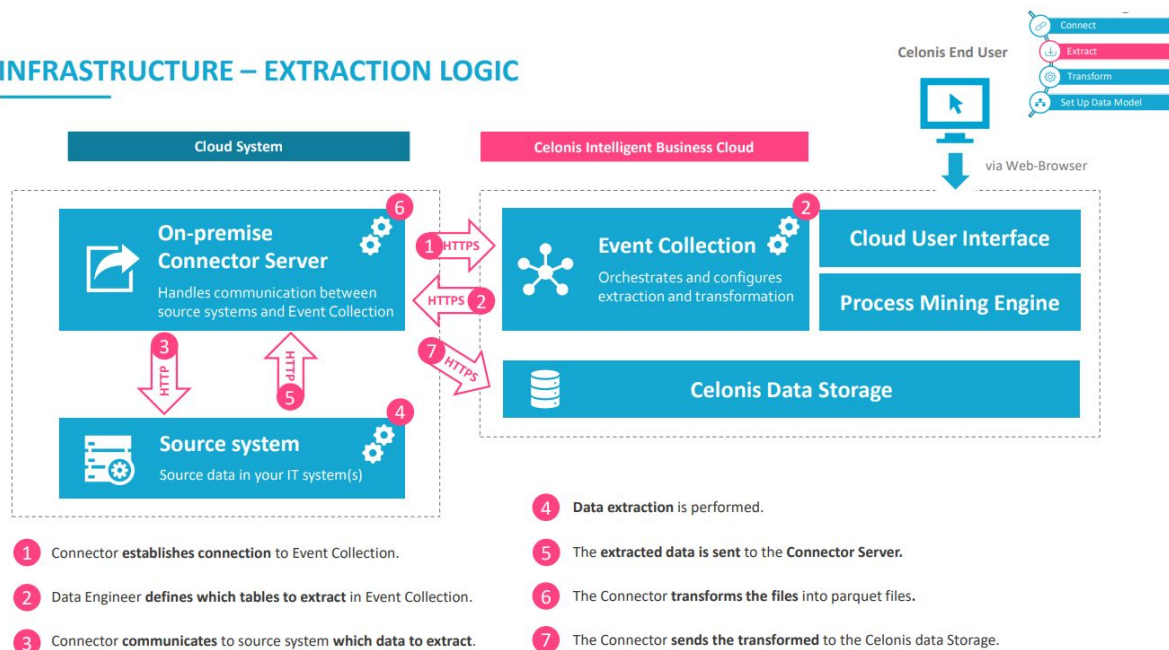
2: Cloud system

Name Mapping:

The name mapping allows you to enable translations for technical and language-specific terms. It is possible to use translation tables from databases, create an excel template and edit it on your machine or even define analysis specific translations.

The general methodology of connecting a new process in the Celonis Intelligent Business Cloud:
1. *Create a new Data Pool in Event Collection*
2. *Create data connections to your source systems*
3. *Create a Data Job and set up an Extraction Task to extract all relevant tables*
4. *Set up the Transformation Tasks*
   ▪ *Create an Activity Table*
   ▪ *Insert activities into the empty Activity Table*
      • *Specify object*
      • *Select information*
      • *Insert activity*
5. *Create a data model and extend it by applying the master data table*
6. *Applying name mapping*

Fourth Part: Process connection continue

CONNECTING MULTIPLE SYSTEMS

There are two different kinds of systems:

a) **Horizontal system:** In a horizontal system, initially we create a data job and set up connections for each source system. After that, create one extraction task as a template and reuse for all the other systems and the same applies for transformation. Next, create a global job(global transformation) and merge all the tables from the transformation. And in the end, we set one data model containing all the activity tables and raw data.

*A **task template** is a generalized task (extraction or transformation) which is useful if you want to reuse a task. A common use case is using the same extractions and transformations for multiple systems.*

b) **Vertical system:** In the vertical system, we cannot use the same extraction and transformation template as in the Horizontal system because in vertical system source systems store different activities. So, in the vertical system, we create separate data jobs for each source system and perform extraction and transformation tasks. Then create a global transformation and merge the activities and create one central data model.

Connecting multiple processes running in one system: creates one data connection and different data jobs for different tasks.

Data load Type :

There are two different load types available in Celonis namely: a) Full load b) Delta load

Full Load:

When you select to "reload all data" the tables in the existing database that have the same name as the ones in the extraction will be deleted and the new data is written to the database. However, the existing data is only deleted after the extraction was successful so you will not lose data unless the extraction worked.

Delta Load:

The purpose of delta loading is to reduce the amount of data needed to be transferred between the source system and the Celonis Event Collection to enable more frequent reloads. Delta loads, by contrast, are only supposed to load new entries (i.e. those cases that were created or updated in the days since the last load) these are much faster as they are not reloading old data, you already loaded last time. Duplicate values are deleted by means of the Primary Key.

The initial load is always a full load and after that, the scheduled loads are ordinarily delta loads.

Scheduling: Schedules are used to automatically execute jobs on a regular basis. Currently, they only allow sequential processing which means that all the jobs within one schedule are executed one after another and if one job fails, the remaining jobs will be canceled. There are three things to set for scheduling 1) Schedule setting either full load or delta load 2) Scheduling plan such as hourly, daily or weekly, etc 3) Scheduled data job: a list view of the jobs which allows you to add jobs to the schedule.

Important points to remember during data extraction and transformation :

1. Extract all the raw data
2. Create an activity table
3. Load the data
4. Create a new transformation for each activity
5. Create a data model and extend it by adding to the master data table
6. Configure name mapping and set the necessary case table
7. Link the tables via their foreign keys
8. Load data model again
9. Set schedule if it is required