

Tencent Integration Platform

Getting Started

Product Documentation



Copyright Notice

©2013-2023 Tencent Cloud. All rights reserved.

Copyright in this document is exclusively owned by Tencent Cloud. You must not reproduce, modify, copy or distribute in any way, in whole or in part, the contents of this document without Tencent Cloud's the prior written consent.

Trademark Notice



All trademarks associated with Tencent Cloud and its services are owned by Tencent Cloud Computing (Beijing) Company Limited and its affiliated companies. Trademarks of third parties referred to in this document are owned by their respective proprietors.

Service Statement

This document is intended to provide users with general information about Tencent Cloud's products and services only and does not form part of Tencent Cloud's terms and conditions. Tencent Cloud's products or services are subject to change. Specific products and services and the standards applicable to them are exclusively provided for in Tencent Cloud's applicable terms and conditions.

Contents

Getting Started

 Data Integration Guide

Getting Started

Data Integration Guide

Last updated : 2023-08-04 09:49:56

Overview

Data integration integrates data from different sources and in different formats to provide your business with comprehensive data sharing and give you a unified and more valuable view of all your data. This helps you make better decisions based on complex and heterogeneous data sources in a fast and stable data environment.

Data integration helps you fully utilize data assets to make smart decisions and drive your business.

- iPaaS data integration allows you to connect to diverse heterogeneous data sources to meet your requirements for multiple data sources and versions.
- iPaaS supports scheduled, manual, and event-based trigger mechanisms as well as full and incremental data sync. This helps you integrate different types of data in different scenarios.
- iPaaS offers a rich set of data processing components that let you cleanse and process raw data, helping you guarantee the quality and value of your data.
- iPaaS supports sync of a large volume of data to help you migrate data to lakes and warehouses.

Use Cases

- Data lake development: Data integration can sync data from silo-like local platforms or SaaS services to a data lake so you can maximize the benefits of all your data.
- Data warehousing: Data integration can integrate data from different sources into a data warehouse for further computing and analysis.
- Digital and smart marketing: Data integration can move all your marketing data such as customer group characteristics or social networking and network analysis data to a certain place for analysis and subsequent operations.
- Industrial IoT: Data integration can integrate data from multiple IoT sources all into one place.
- Database replication: Data integration plays a key role in replicating data from a source database like Oracle, SQL Server, or MySQL to a cloud data warehouse.
- Application integration: Data integration integrates ERP systems, applications, SaaS services, and hybrid clouds through sync of the master data and business documents.

Samples

Database data sync

You can sync data from one or multiple tables from the source application database (MySQL in this example) to the target application database.

1. Use the Scheduler to sync the data at the scheduled time.
2. Connect the Database connector to the source application database. Configure the scope of data to be synced, including the row and column data. Similarly, connect the Database connector to the target application database and configure the target table for inserting the synced data.
3. After configuring the source and target, add the Mapper component between them to map and convert the source objects to the target objects, i.e., to save the obtained source data to the target database according to the mappings. This step can extract data from the source application database, map and convert it, and load it to the target application database at the scheduled time.

Directions

1. Configure the Scheduler. For more information, see Scheduler.
Click the **Trigger** section in the integration designer canvas and select the **Scheduler** component. On the pop-up configuration panel on the right, enter the expression for scheduled execution and configure the scheduling rule; for example, you can enter `0 0 12 * * ?` to execute the data integration task at 12:00 every day.
2. Configure the source application database. Set to extract data from the data source, transfer the extracted data to subsequent components for processing, and load the processed data into the target database. The specific steps are as follows:
 - a. Click **+** next to the **Scheduler** component in the canvas. Search for **Database** on the component selection panel, and click the **Database** connector. Click **Query** to add the connector to the canvas, and the configuration panel pops up on the right:
 - b. Configure the connection: Select **MySQL** as the data source type (select an option based on the actual use case), specify the database to be queried, and enter the database address, port number, username, and password.
 - c. Configure the query operation: The operation modes include **Simple mode - single table query**, **Simple mode - multi-table query**, and **SQL mode**.
 - Simple mode - single table query
 - Simple mode - multi-table query
 - SQL mode

After you select the single table query mode, the Database connector will read all tables in the specified database. You can select the data source table to be synced.

Directions

- i. Select **Simple mode - single table query**.
 - ii. Select a table.
 - iii. Select the fields to be queried.
 - iv. Configure the query conditions if needed.
 - v. In output configuration, select **RecordSet** as the output mode and **false** as the cache option and keep the default number of partitions, which is `1`.
3. Configure the target database and table for loading the extracted data.
- a. In the designer canvas, click **+** next to the source Database connector. On the component selection panel, search for and select the **Database** connector and then select **Batch insert** or **Batch merge**.
 - b. Configure on the Database batch insert panel as follows:
 - i. Configure the connection to the target database.
 - ii. Set the data set. If you select **Recordset** as the input mode of the previous component, enter `return msg.payload` here.
 - iii. Select a table in the target database for loading the data.
 - iv. Select the field from which to extract the data and map it to the target table field. Data of unmapped fields will be discarded.
 - v. Click **Schedule and run** to experience the effect.
4. Filter the data. If you need to filter the data extracted from the source, you can add the Filter component after the source Database connector for processing. The Filter component configuration panel consists of the following parts:
- a. Enter the data set to be filtered. If the data set is of the `Recordset` type, it does not need to be processed. The default value is `msg.payload`, and you can skip this step.
 - b. Configure the fields to be output after filtering. If you don't need the data of a field, you can deselect the field, and the component will filter it out.
 - c. Set the filters. Step b is to filter columns, while this step is to filter rows. You can configure fields and conditions such as `==`, `!=`, `>=`, `<=`, `>`, `<`, `StartWith`, `EndWith`, `Contain`, `In`, `NotIn`, `IsNull`, or `IsNotNull` to filter data; for example, you can filter out canceled orders.
5. Map the data. Source table and object structures are generally not simply in one-to-one correspondence to the target table and object structures. In this case, you can map and convert the data based on the correspondence between the extracted fields and fields to be loaded.
- a. Add the Mapper component between the source and target database connectors and configure the logic for mapping and conversion between the source and target object fields.

b. Drag and drop the input and output information (source and target fields) on the GUI to map them. If you need to process the source data, you can add processing logic for the relevant fields in the logic mapping section.