
Data-as-a-Service (DaaS)

CALYX™

Contents

Data-as-a-Service (DaaS) Overview.....	1
DaaS and Calyx RIM Interaction.....	2
DaaS Access.....	3
Pipelines and Schemas.....	4
Product Support and Client Responsibilities.....	5
Index.....	7

Data-as-a-Service (DaaS) Overview

Use DaaS to operate a database in the cloud for data provisioning, licences support, and maintenance.

Data-as-a-Service

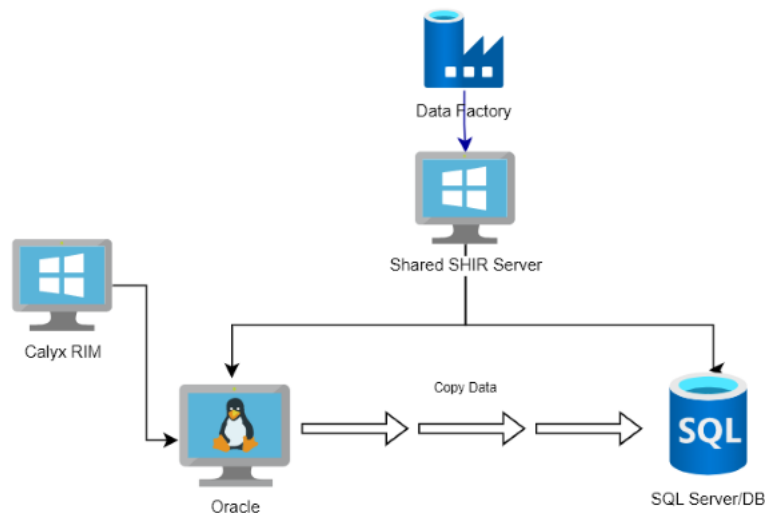
Data-as-a-Service (DaaS) is a cloud computing service that gives access to a database without the setup of a physical hardware, the database configuration and the required software installation. DaaS is used to access and interact with data.

Benefits of DaaS include:

- Minimum set up so that you can start working with database immediately.
- Support and maintenance of the base-line infrastructure.
- Faster performance with the software and timely updates of tools.

DaaS and Calyx RIM Interaction

Use the information about the high-level architecture of the DaaS solution to access and interact with data. Calyx RIM uses an Oracle database for data storage. DaaS provides the ability to query the data, process datasets or generate reports.



Data Factory is used to transfer the data from the source to the destination point. In the example above, Oracle database with the Calyx RIM data is the source point and Azure SQL database is the destination point.

The self-hosted integration server is required to set up a connection between the source and destination points. By default, Data factory is not connected to the source or destination points. Data Factory pipelines are run on self-hosted integration server which provide the data transfer from the Oracle database to Azure SQL database.

SQL server and SQL database are deployed to the environment and the required information is given to the client to access the database. SQL database is installed on the database server and the credentials are shared with users for access to the database.

DaaS Access

Access DaaS using the credentials provided by Calyx.

To allow user access to database, user IP address is granted access in the database server. VPN is not required to access data. After the IPs are added to the database server, Calyx will provide the SQL database and server information that includes:

- Server Name
- Database Name
- User ID (by default, it is daas_reader)
- Password

To reset a password, user should raise a ticket to Service Now.

Pipelines and Schemas

Calyx RIM uses different types of pipelines and schemas for DaaS architecture.

Pipeline

A database pipeline enables flow of data from source to destination. There are two types of pipelines:

- Full Load: In the Full Load pipeline, an existing dataset is replaced by the newly updated dataset in the next data loading. All datasets will have an initial full load.
- Incremental Load: In the Incremental Load pipeline, the data coming from the source is compared with the existing data in the destination. Only data that has been changed from the source will be replaced.

Schemas

The database schema structure used by Calyx RIM includes:

- Audit Schema: The Audit schema include an overview of the fields used, the data and the source of the data. All the **audit** tables from the Calyx RIM Oracle database are copied to the **s_audit** table of SQL database using Azure Data Factory.
- Reporting Schema: The Reporting schema defines the tables, the fields in each table, and the relationships between fields and tables. All the **MGR, ODS** tables from the Calyx RIM Oracle database are copied to **s_reporting** of SQL database using Azure Data Factory.

Work with Data

Before copying, replacing or further data processing, check if the data is current by selecting the **last_refresh_date** from **dbcommon.df_run_parameters** where **run_key=<run_key>**. **Run_key** is a unique string key used to get pipeline settings from **df_run_parameters**. **Run_key** defines the **Table_owner**, **Include_list** and **Exclude list**.

A table is refreshed every five minutes.

Product Support and Client Responsibilities

Use the provided service account to interface systems to the database.

Product Support

By enabling permission to client IP address, data can be accessed by client systems through service account information that includes credentials, SQL database and server information. VPN is not required for accessing data.

You can use the client system to work with data without any constraints except for egress which has additional cost implications.

The service account provided is used to enable interaction between the client systems and the database and supports the following:

- Ingest process: the use of a script or system based process to query the data required to ingest into the environment.
- Simple reporting process: interface a system (such as a reporting platform) to the database as a data source:
 - Depending on your requirement, you may choose for the performance between Azure SQL and Internet model you have.
 - The standard principle for compute intensive operations is that data must be local to the compute resources. Therefore, clients should ingest data into their own environment such that they can then operate on it locally.
 - Calyx can increase the sizing of the database both vertically and horizontally, but this is at additional cost and mediated by Pro Services architecture team.
- Calyx will not provision access to individual user accounts to run queries against the database.
 - Calyx will not manage user accounts / starters / movers / leavers process.
 - If a client wants to present data to their users with granular controls around access, they may do so from their own solutions / environment.

Client Responsibilities

Clients are responsible for:

- Connecting to their own data. Any SQL client may be used while accessing the SQL Reporting database.
- Data replication into their own environments.
- Use of reporting software. Any reporting software that supports SQL may be used.

Clients may encounter the following performance issues while accessing data:

- Accessing data over the Internet may increase performance issues which are not controlled by Calyx. These issues may increase when loading large amounts of data which tend to be slow and time-consuming.
- Accessing data over the Internet may require an increased amount of time. This can cause performance issues and are not controlled by Calyx other than potential metrics in line with different latencies.

Clients will have real-time insight into performance matrix. Performance data is available in the portal for the SQL database and alerting (this is CPU, memory type utilization positions).

Client owns the source system interfacing (and the performance of any queries) with Azure SQL and MS own the performance of Azure SQL as a service position, and networking is dependent on latency.

Index

A

Architecture 1–5