ODS – Operational Data Store.

Replicate data to Azure SQL Database using ODS.

**Contributors**

XXXX

**Date**

Jan 2023

**Version**

1.0

Table of Contents

[Overview 3](#_Toc132274364)

[Prerequisites 4](#_Toc132274365)

[Architecture 5](#_Toc132274366)

[Type of Columns 6](#_Toc132274367)

[Process Workflow 7](#_Toc132274368)

[Solution Design 9](#_Toc132274369)

[Best Practices when using ODS framework. 13](#_Toc132274370)

[Deployment Strategy 14](#_Toc132274371)

[Create Data Factory 14](#_Toc132274372)

[Import ARM Template 14](#_Toc132274373)

[Post deployment validation 16](#_Toc132274374)

[Pipeline Execution 17](#_Toc132274375)

[Known Issues 18](#_Toc132274376)

[Appendix 19](#_Toc132274377)

[Acronyms 19](#_Toc132274378)

[Reference Documents 19](#_Toc132274379)

**Revision and Signoff Sheet**

Change Record

|  |  |  |  |
| --- | --- | --- | --- |
| **Date** | **Commented by** | **Version** | **Change reference** |
| 01/20/2023 | Chetan Hiran | 1.0 | Draft |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

**Approvers**

|  |  |  |  |
| --- | --- | --- | --- |
| **Name** | **Version approved** | **Position** | **Date** |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

# Overview

The ODS is a framework developed using Azure Service to replicate data from Microsoft Dataverse to an Azure SQL Database store in a user owned Azure Subscription. The supported target destination is Azure SQL Database, but the framework can be easily extended to support SQL Server on Azure virtual machines or on-Premises SQL server. The Data Export Service intelligently synchronizes the entire data initially and thereafter synchronizes on a continuous basis as changes occur (delta changes) in the system.

# Prerequisites

Azure SQL database Service

A SQL database and a service account with the Database permissions.

| **Permission type code** | **Permission name** |
| --- | --- |
| CRTB | CREATE TABLE |
| CRTY | CREATE TYPE |
| CRVW | CREATE VIEW |
| CRPR | CREATE PROCEDURE |

Schema permissions required.

| **Permission type code** | **Permission name** |
| --- | --- |
| AL | ALTER |
| IN | INSERT |
| DL | DELETE |
| SL | SELECT |
| UP | UPDATE |
| EX | EXECUTE |
| RF | REFERENCES |

**Azure Data Factory**

**Azure Key Vault service**

Key Vault Secrets User Role for Data Factory.

**Azure Blob Storage**

Storage Blob Data Contributor role for Data factory.

**Microsoft Dataverse**

Service account user with System Administrator security role.

# Architecture



## Type of Columns

The following table includes the corresponding sink type.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Power Apps data type** | **Solution Explorer type** | **API type** | **Sink Data Type** | **Additional Sink Columns** |
| Big Integer | Time Stamp | BigIntType | bigint |  |
| Yes/No | Two Options | BooleanType | bit | <<attributename>>name : String |
| Customer | Customer | CustomerType | uniqueidentifier | <<attributename>>name : String  <<attributename>>Type : String |
| Date and Time | Date and Time Date and Time Format | DateTimeType | datetime |  |
| Date Only | Date and Time Date Only Format | DateTimeType | datetime |  |
| Decimal Number | Decimal Number | DecimalType | decimal |  |
| Floating Point Number | Floating Point Number | DoubleType | decimal |  |
| File | File | FileType | uniqueidentifier |  |
| Image | Image | ImageType | varchar(MAX) |  |
| Duration | Whole Number Duration Format | IntegerType | int |  |
| Language | Whole Number Language Format | IntegerType | int |  |
| Timezone | Whole Number Time Zone Format | IntegerType | int |  |
| Whole Number | Whole Number None Format | IntegerType | int |  |
| Lookup | Lookup | LookupType | uniqueidentifier | <<attributename>>name : String |
| Multiline Text | Multiple Lines of Text | MemoType | nvarchar(MAX) |  |
| Currency | Currency | MoneyType | decimal |  |
| Choices | MultiSelect Field | MultiSelectPicklistType | nvarchar(MAX) |  |
| Owner | Owner | OwnerType | uniqueidentifier | <<attributename>>name : String  <<attributename>>Type : String |
| Choice | Option Set | PicklistType | int | <<attributename>>name : String |
| Status | Status | StateType | int | <<attributename>>name : String |
| Status Reason | Status Reason | StatusType | int | <<attributename>>name : String |
| Email | Single Line of Text Email Format | StringType | nvarchar |  |
| Phone | Single Line of Text Phone Format | StringType | nvarchar |  |
| Text | Single Line of Text Text Format | StringType | nvarchar |  |
| Text Area | Single Line of Text Text Area Format | StringType | nvarchar |  |
| Ticker Symbol | Single Line of Text Ticker Symbol Format | StringType | nvarchar |  |
| URL | Single Line of Text URL Format | StringType | nvarchar |  |
| Unique Identifier | Unique Identifier Primary Key | UniqueidentifierType | uniqueidentifier |  |

## Process Workflow

Data Sync Pipeline



Schema Sync Pipeline



## Solution Design

Database Model

Tables

* ODS.EntitySync

|  |  |  |
| --- | --- | --- |
| **COLUMN\_NAME** | **DATA\_TYPE** | **IS\_NULLABLE** |
| ID | int | NO |
| DisplayName | nvarchar(500) | NO |
| EntityName | varchar(500) | NO |
| RelativeURL | varchar(500) | NO |
| DeltaToken | varchar(500) | YES |
| LastSyncTimeStamp | datetime | YES |
| AttributeSchemaURL | varchar(500) | YES |
| Mapping | varchar(MAX) | YES |
| FirstInsertMapping | varchar(MAX) | YES |
| SyncReady | bit | YES |
| ChangeTrackingEnabled | bit | YES |
| IsNNRelationShip | bit | YES |
| SyncFrequency | varchar(100) | YES |
| PrimaryIdAttribute | nvarchar(500) | YES |
| IsBPFEntity | bit | YES |
| IsActivity | bit | YES |
| IsSystemVersioningEnabled | bit | NO |

* ODS. EntityAttributeSchemaMapping

|  |  |  |
| --- | --- | --- |
| **COLUMN\_NAME** | **DATA\_TYPE** | **IS\_NULLABLE** |
| SourceSchema | varchar(500) | YES |
| DestinationSchema | varchar(500) | YES |
| PrimaryColumnMapping | varchar(MAX) | YES |
| SecondaryColumnMapping | varchar(MAX) | YES |
| AttributeTypeName | nvarchar(500) | YES |
| SecondaryColumnMappingJson | varchar(MAX) | YES |

* ODS.EntityAttributeSchema

|  |  |  |
| --- | --- | --- |
| **COLUMN\_NAME** | **DATA\_TYPE** | **IS\_NULLABLE** |
| ID | int | NO |
| AttributeOf | varchar(500) | YES |
| AttributeType | varchar(500) | YES |
| EntityLogicalName | varchar(500) | YES |
| IsPrimaryId | bit | YES |
| IsValidODataAttribute | bit | YES |
| IsPrimaryName | bit | YES |
| AttributeLogicalName | varchar(500) | YES |
| AttributeSchemaName | varchar(500) | YES |
| MaxLength | int | YES |
| DatabaseLength | int | YES |
| Precision | int | YES |
| IsRetrievable | bit | YES |
| IsValid | bit | YES |
| AttributeTypeName | nvarchar(500) | YES |

* ODS. ExecutionLog

|  |  |  |
| --- | --- | --- |
| **COLUMN\_NAME** | **DATA\_TYPE** | **IS\_NULLABLE** |
| ID | int | NO |
| ExecutionId | uniqueidentifier | YES |
| PackageName | varchar(500) | YES |
| RunStatus | varchar(100) | YES |
| StartDate | datetime | YES |
| EndDate | datetime | YES |
| Message | nvarchar(500) | YES |

* ODS. ExecutionLogDetail

|  |  |  |
| --- | --- | --- |
| **COLUMN\_NAME** | **DATA\_TYPE** | **IS\_NULLABLE** |
| ID | int | NO |
| ExecutionId | uniqueidentifier | YES |
| PipelineRunID | uniqueidentifier | YES |
| StepName | varchar(500) | YES |
| StartDate | datetime | YES |
| Message | nvarchar(500) | YES |
| LogType | varchar(100) | YES |
| EntityName | varchar(500) | YES |

* ODS. EntityDataSyncLog

|  |  |  |
| --- | --- | --- |
| **COLUMN\_NAME** | **DATA\_TYPE** | **IS\_NULLABLE** |
| ID | int | NO |
| EntityName | nvarchar(500) | NO |
| ExecutionId | uniqueidentifier | NO |
| PipelineRunID | uniqueidentifier | NO |
| InsertCount | int | YES |
| UpdateCount | int | YES |
| DeleteCount | int | YES |
| SyncDate | datetime | NO |
| SkippedRows | int | NO |
| LogFilePath | nvarchar(500) | YES |

* ODS. EntityDeltaTokenLog

|  |  |  |
| --- | --- | --- |
| **COLUMN\_NAME** | **DATA\_TYPE** | **IS\_NULLABLE** |
| ID | bigint | NO |
| ExecutionId | uniqueidentifier | NO |
| PipelineRunId | uniqueidentifier | NO |
| EntityName | nvarchar(500) | NO |
| DeltaToken | varchar(500) | NO |
| LogDate | datetime | NO |

* ODS. EntitySchemaSyncLog

|  |  |  |
| --- | --- | --- |
| **COLUMN\_NAME** | **DATA\_TYPE** | **IS\_NULLABLE** |
| Id | bigint | NO |
| ExecutionId | uniqueidentifier | YES |
| SchemaChangeQuery | nvarchar(MAX) | YES |
| SyncDate | datetime | YES |
| TableName | nvarchar(500) | YES |

Custom Data Types

* ODS.EntityAttributeSchemaType: User defined Table Type
* ODS.EntitySyncType: User defined Table Type

Procedures

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Schema** | **ObjectName** | **ParameterName** | **ParameterDataType** | **Description** |
| ODS | usp\_CreateStagingTable | @sourceEntityName | nvarchar(500) | Procedure to create staging table |
| @PrimaryIdAttribute | nvarchar(500) |
| ODS | usp\_ExecutionLogDetail | @ExecutionId | uniqueidentifier(8) | Procedure to log execution details |
| @PipelineRunID | uniqueidentifier(8) |
| @StepName | varchar(250) |
| @Message | nvarchar(500) |
| @LogType | varchar(50) |
| @entityName | varchar(250) |
| ODS | usp\_InsertEntityDataSyncLog | @EntityName | nvarchar(500) | Procedure to details amount of data written/created in sink |
| @ExecutionId | uniqueidentifier(8) |
| @PipelineRunID | uniqueidentifier(8) |
| @InsertCount | int |
| @UpdateCount | int |
| @DeleteCount | int |
| @SkippedRows | int |
| @LogFilePath | nvarchar(500) |
| ODS | usp\_InsertEntityDeltaTokenLog | @ExecutionId | uniqueidentifier(8) | Procedure to log delta token logs. |
| @DeltaToken | varchar(250) |
| @PipelineRunId | uniqueidentifier(8) |
| @EntityName | nvarchar(500) |
| ODS | usp\_InsertEntitySchemaSyncLog | @ExecutionId | uniqueidentifier(8) | Procedure to capture schema changes on sink tables. |
| @SchemaChangeQuery | nvarchar(0) |
| @tableName | nvarchar(500) |
| ODS | usp\_InsertOrUpdateEntityTable | @sourceTableName | nvarchar(500) | Procedure to write data to sink |
| @primaryColumn | nvarchar(500) |
| @ExecutionId | nvarchar(50) |
| ODS | usp\_LoadEntityAttributeSchema | @passing | EntityAttributeSchemaType | Procedure to load master data for AttributeSchemaMapping |
| @sourceTableName | nvarchar(100) |
| ODS | usp\_LoadEntitySyncData | @Sourcetbl | EntitySyncType(0) | Procedure to load master data for EntitySync Table |
| ODS | usp\_LogEndExecution | @ExecutionId | uniqueidentifier(8) | Procedure to log pipeline completion status. |
| @RunStatus | varchar(50) |
| @Message | nvarchar(500) |
| ODS | usp\_LogStartExecution | @ExecutionId | uniqueidentifier(8) | Procedure to log pipeline execution start. |
| @PackageName | varchar(250) |
| @RunStatus | varchar(50) |
| ODS | usp\_SyncEntitySchema | @sourceEntityName | nvarchar(500) | Procedure to sync sink table schema |
| @PrimaryIdAttribute | nvarchar(500) |
| @ExecutionId | uniqueidentifier(8) |
| @StagingDBSchema | varchar(50) |
| ODS | usp\_TruncateStagingTable | @sourceEntityName | varchar(250) | Procedure to truncate staging tables. |
| @databaseSchema | varchar(50) |
| ODS | usp\_UpdateAttributeMapping\_WithJson | @sourceTableName | nvarchar(500) | Procedure to update Attribute Mapping |
| @isActivity | bit |
| @primaryColumnName | nvarchar(100) |
| ODS | usp\_UpdateDeltaToken | @sourceTableName | nvarchar(500) | Procedure to update delta tokens |
| @sourceSchema | nvarchar(100) |
| @pipelineRunId | nvarchar(100) |
| ODS | usp\_ValidateAttributes | @sourceTableName | varchar(250) | Procedure to validate attributes. |

## Best Practices when using ODS framework.

* To avoid synchronization errors due to resource throttling, we recommend that you have an Azure SQL Database Premium P1 or better plan
* Enable Change tracking on Dataverse Entities. More Information: [Enable Change Tracking in Power Platform.](https://learn.microsoft.com/en-us/power-platform/admin/enable-change-tracking-control-data-synchronization)
* Periodically update database statistics on tables and indexed views in the SQL database. More information: [Update Statistics](https://learn.microsoft.com/en-us/sql/relational-databases/statistics/update-statistics?view=sql-server-ver16)
* Monitor the SQL database's utilization. More information: [Perf monitoring](https://learn.microsoft.com/en-us/azure/azure-sql/database/monitor-tune-overview?view=azuresql)
* Monitor Data Factory Pipeline execution. More Information: [Monitor using Azure Monitor](https://learn.microsoft.com/en-us/azure/data-factory/monitor-using-azure-monitor)

# Deployment Strategy

## Create Data Factory

Follow the steps to create the data factory.

<https://learn.microsoft.com/en-us/azure/data-factory/quickstart-create-data-factory>

## Import ARM Template

1. Navigate to Manage -> ARM Template -> Import ARM Template.

A screenshot of a computer

Description automatically generated

1. Select *“Build your own template in the editor.”*

A screenshot of a computer

Description automatically generated with medium confidence

1. Browse to “*ARMTemplateForFactory.json*” file.

A screenshot of a computer

Description automatically generated with medium confidence

1. Click Save
2. Enter the configuration details.

A screenshot of a computer

Description automatically generated with medium confidence

1. Click Review + Create.

This will deploy the pipeline to the Data factory.

## Post deployment validation

1. Open Data Factory Studio
2. Navigate to Manage -> Linked Services.

A screenshot of a computer

Description automatically generated

1. Test all the connections.

A screenshot of a computer

Description automatically generated with medium confidence

# Pipeline Execution

1. Navigate to Author -> 01 Master Data Load -> Add Trigger -> Trigger Now

A screenshot of a computer

Description automatically generated

The pipeline will load data on ODS.EntitySync table.

1. Update SyncReady Flag in ODS.EntitySync table using the below sample script.

UPDATE ODS.EntitySync

SET SyncReady = 1

where EntityName = 'account'

***Note:*** Only the entities flagged as SyncReady = 1 will be set for data sync.

1. Navigate to “02 – Entity Schema Sync Master” -> Add Trigger -> Trigger Now

A screenshot of a computer

Description automatically generated

***Note:*** This pipeline will create the required tables in the database. Execute this pipeline manually after each solution deployment to environment. This pipeline helps to sync the schema between D365 and the replicated database.

1. After completion of Schema Sync pipeline execute the “03 – Data Sync Master” pipeline.

***Note:*** This pipeline will sync the replicate data from D365 to Azure SQL db. It is recommended to create triggers to sync the data at periodic intervals.

# Known Issues

There are no known bugs when using the framework.

# Appendix

## Acronyms

## Reference Documents