DATA
& AI

Microsoft Azure Solutions Business

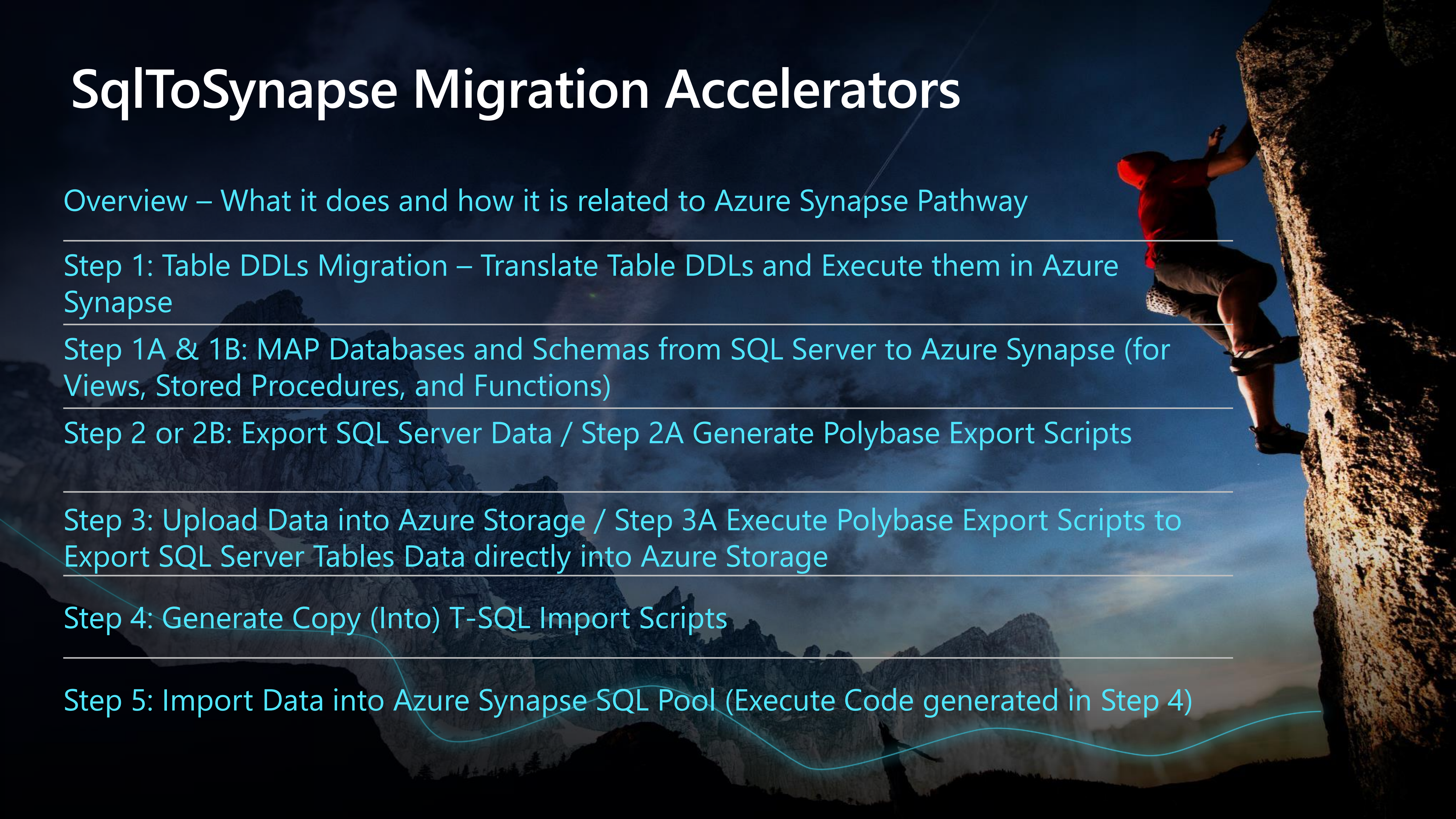
AAAP
Analytics & AI Accelerate Program

SQL Server to Azure Synapse Migration Accelerators
(Migration Process, Modules, Scripts)

Gaiye "Gail" Zhou
February 2022

Disclaimer

Theis document was developed in consultation and collaboration with Microsoft Corporation technical architects. Because Microsoft must respond to changing market conditions, this document should not be interpreted as an invitation to contract or a commitment on the part of Microsoft. Microsoft has provided high-level guidance in this document with the understanding that MICROSOFT MAKES NO WARRANTIES, EXPRESS OR IMPLIED, WITH RESPECT TO THE INFORMATION CONTAINED HEREIN. This document is provided “as-is”. Information and views expressed in this document, including URL and other Internet Web site references, may change without notice. Some examples depicted herein are provided for illustration only and are fictitious. No real association or connection is intended or should be inferred. This document does not provide you with any legal rights to any intellectual property in any Microsoft product. You may copy and use this document for your internal, reference purposes.



SqlToSynapse Migration Accelerators

Overview – What it does and how it is related to Azure Synapse Pathway

Step 1: Table DDLs Migration – Translate Table DDLs and Execute them in Azure Synapse

Step 1A & 1B: MAP Databases and Schemas from SQL Server to Azure Synapse (for Views, Stored Procedures, and Functions)

Step 2 or 2B: Export SQL Server Data / Step 2A Generate Polybase Export Scripts

Step 3: Upload Data into Azure Storage / Step 3A Execute Polybase Export Scripts to Export SQL Server Tables Data directly into Azure Storage

Step 4: Generate Copy (Into) T-SQL Import Scripts

Step 5: Import Data into Azure Synapse SQL Pool (Execute Code generated in Step 4)

SqlToSynapse Migration Accelerator (PowerShell Scripts)

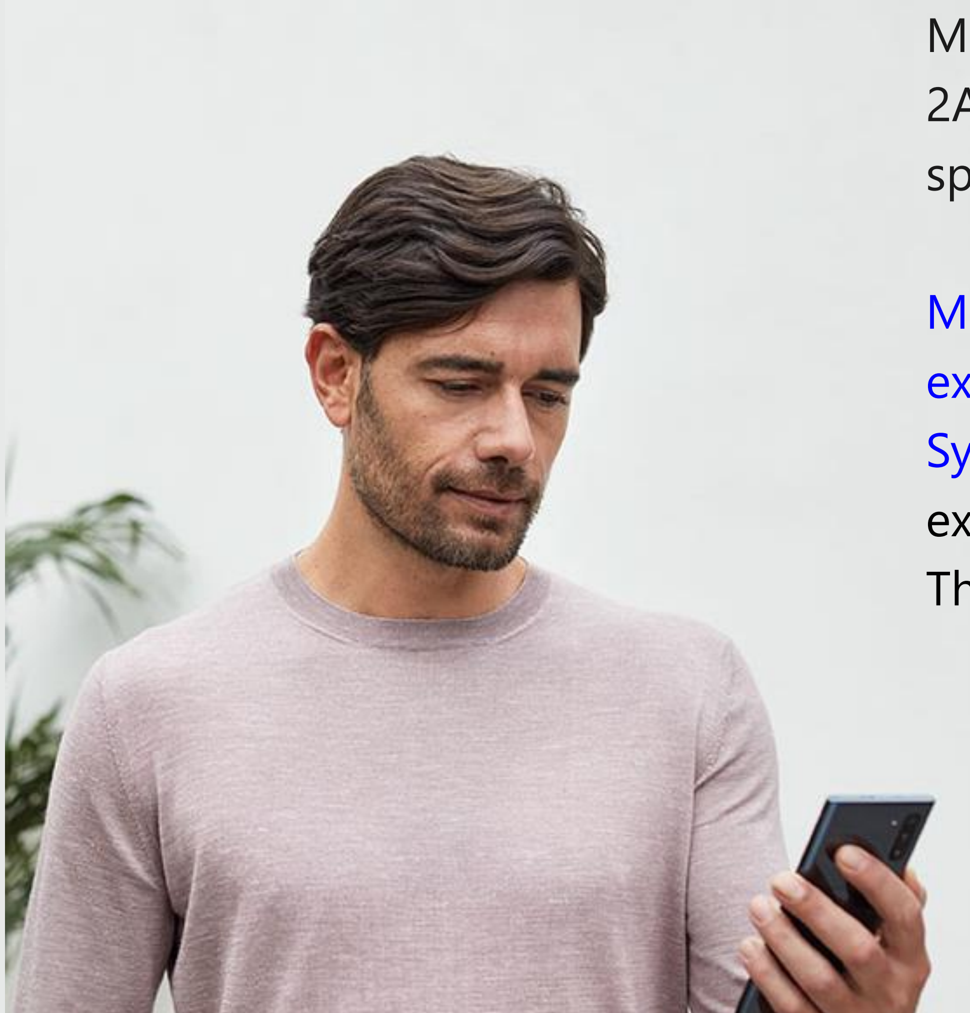
What does the Scripts do?

- ✓ Translate SQL Server Table DDLs into Azure Synapse DDLs (Create Table Statements) and execute Table DDLs in Azure Synapse
- ✓ Export SQL Server tables to .csv or .parquet files
- ✓ Upload .csv or .parquet files into Azure Storage
- ✓ Extract SQL Server Code (Views, Stored Procedures and Functions) and perform database and schema mapping
- ✓ Generate Polybase Export T-SQL Scripts.
- ✓ Execute Polybase Export T-SQL Scripts – Export data directly into Azure Storage from SQL Server
- ✓ Generate T-SQL Copy Into Import Scripts.
- ✓ Execute T-SQL Copy Into Scripts - Import Data into Azure Synapse from Azure Storage.

Where are the scripts and documentation?

AzureSynapseScriptsAndAccelerators/Migration/SQLServer at main · microsoft/AzureSynapseScriptsAndAccelerators (github.com)

Reusability



Module 1, 1A, and 1B are designed for code translation; Module 2, 2A, and 2B are for SQL Server data export. Only these modules are specific to SQL Server.

Module 3, 4, 5 are reusable for other types of migrations, for example, Netezza or Teradata or Exadata or Oracle to Azure Synapse migrations. After the code is translated, and data is exported out of source systems, the rest of the tasks are the same. Therefore module 3-5 can be utilized for any of those migrations.

What Do I need to Use the Utilities?

Choose One of the Environments:

1. Windows PowerShell ISE (preferred)
2. Visual Studio Code (with PowerShell Extension Installed)

PowerShell Modules required:

(1) Install-Module (Import-module) -name SqlServer
(2) Install-Module (Import-module) -name ImportExcel
(3) Install-Module (Import-module) -Name Az –AllowClobber

PowerShell Permissions (Permissions may be denied if the Scripts are from GitHub or Email):
Use one of the options to set Powershell permissions (examples):

```
Set-ExecutionPolicy Unrestricted -Scope CurrentUser
Unblock-File -Path C:\migratemaster\modules\1_TranslateTableDDLs\TranslateTables.ps1
```

Download and Install AzCopy (Only if you will be using BCP Export Method). This task can be skipped if you will be using other methods to upload data into Azure Storge such as Azure Data Box Gateway or Azure Data Explore).

Copy or move data to Azure Storage by using AzCopy v10 | Microsoft Docs

Local .csv/.parquet files or Polybase Export Path

# Step	Export Tables into local .csv or .parquet format	Polybase Export	Results
1 – Task 1	**Translate SQL Tables** (Module 1_TranslateTableDDLs)	← Same	Azure Synapse Table DDLs generated
1 – Task 2	**Execute Synapse Tables** DDLs (or DMLs) (Module 5_RunSqlFilesInfolder)	← Same	Tables Created in Azure Synapse
1 – Task 3	Extract SQL Server Code and Map Databases and Schemas (Module 1A_ExtractCodeDDLs, and Module 1B_MapDatabasesAndSchemas)	← Same	Code (Views, Stored Procedures, Functions) have the desired Azure Synapse DB name and Schema Name
1 – Task 4	Perform Additional Code Translation (Manually or using Azure Synapse Pathway) and then Execute Code using (Module 5_RunSqlFilesInfolder)	← Same	Code is migrated into Azure Synapse
2 or 2B	**Export SQL Server Table data into .csv files** (Module 2_ExportSourceDataWithBCP) **Export SQL Server Table data into .parquet files** (Module 2B_ExportSourceDataToParquet)		Data is exported to local files (.csv or .parquet)
2A – Generate Polybase Code		**Generate Polybase Export Scripts** (Module 2A_GeneratePolybaseExportScripts)	Polybase Export T-SQL Scripts generated
3	**Upload exported Data into Azure Storage** (Module 3_LoadDataIntoAzureStorage)		Data loaded to Azure Storage
3A – Run Polybase Code		**Execute Polybase Export Scripts** **(**Module 5_RunSqlFilesInfolder)	Data exported directly to Azure Storage from SQL Server
4	**Generate COPY Into Scripts** (Module 4_GenerateCopyIntoScripts)	Same as BCP	Copy Into T-SQL Scripts generated
5	**Execute COPY Into Scripts** (Module 5_RunSqlFilesInFolder)	Same as BCP	Data imported to Synapse

PowerShell Modules for Exporting into Local files

There are five modules that contain PowerShell Scripts and T-SQL Scripts designed to accomplish key task(s) if you will use BCP Export method.

The modules are summarized as below:

1_TranslateTableDDLs: Translate SQL objects (DDLs) from source system format to Azure Synapse format. The output is stored as .sql files in specified output file folder (configurable).).

1A_ExtractCodeDDLs: Script out SQL Objects (Views, Functions, Stored Procedures, and Triggers) and write each object into a separate .sql file.

1B_MapDatabasesAndScheams: Map SQL Server Database & Schema into Synapse Database and Schema. In addition, unsupported data types in the code are mostly discovered and counted.

2_ExportSourceDataWithBCP: Export SQL Server Tables into data files stored in predefined structure and format (.csv or .txt)

2B_ExportSourceDataWithBCP: Export SQL Server Tables into data files stored in parquet format (.parquet)

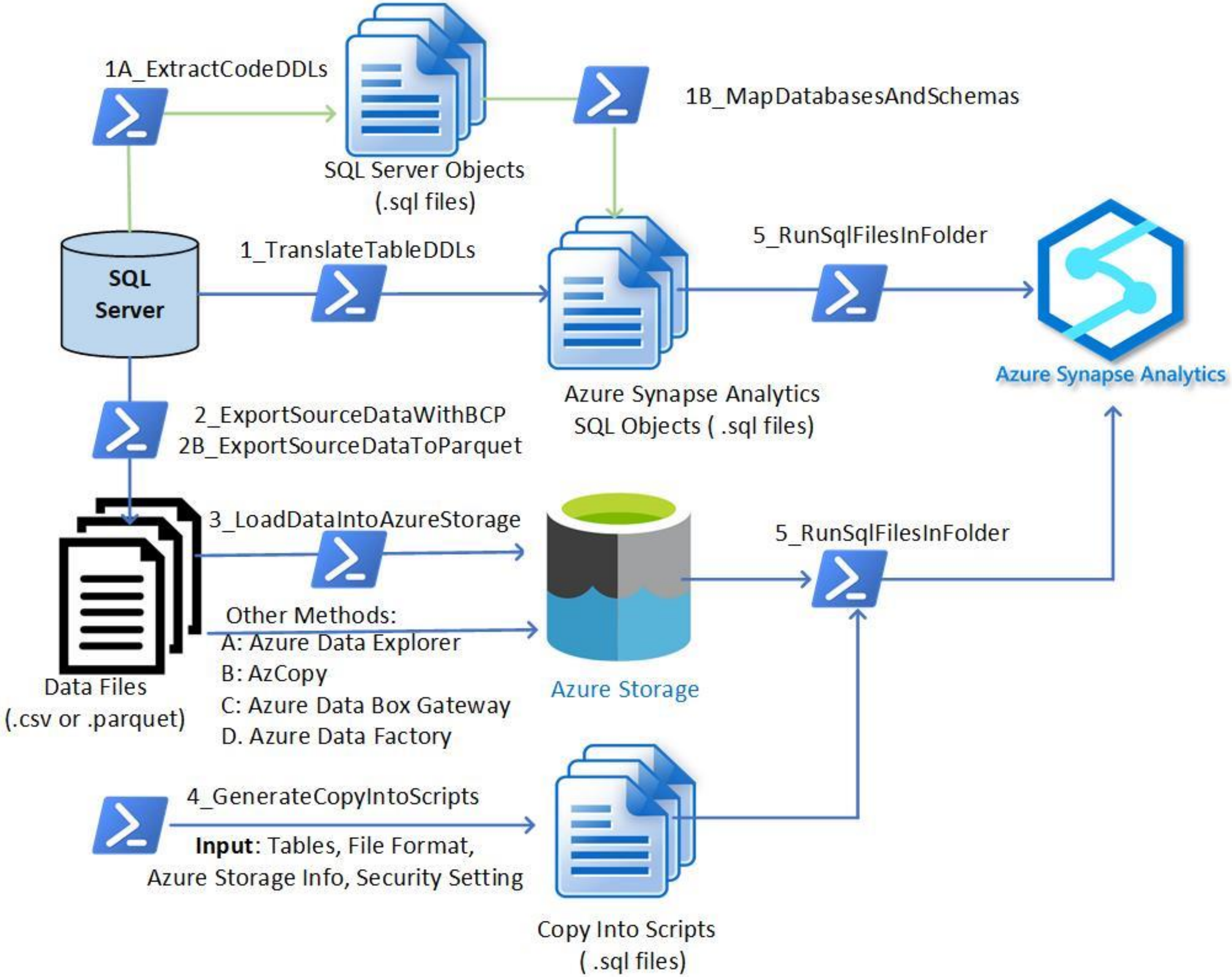
3_LoadDataIntoAzureStorage: Load exported data files into specified container in Azure Storage.

4_GenerateCopyIntoScripts: Generate "COPY Into" T-SQL Scripts that will move data from Azure Storage into Azure Synapse SQL Pool tables, once executed.

5_RunSqlFilesInFolder: Run all T-SQL Scripts defined in .sql files stored in a specified file folder. The T-SQL Scripts can be DDL, DML, Data Movement Scripts (such as Copy Into scripts or Polybase Export Scripts), or any other scripts such as create/update statistics or indexes. In fact, this module is designed to run all T-SQL scripts in a folder (against Azure Synapse or SQL Server).

Note: **5_RunSqlFilesInFolder is reused twice in the process:** (1) Run T-SQL Scripts generated from 1_TranslateMetadata, and (2) Run T-SQL Scripts generated from 4_GenerateCopyIntoScripts. You can use module 5_RunSqlFileInFolder to execute code manually translated, mapped by module 1B or translated by Azure Synapse Pathway.

Export SQL Server data into local .csv or .parquet files – All Versions of SQL Servers



PowerShell Modules to Use for Polybase Export

There are four modules that contain PowerShell Scripts and T-SQL Scripts designed to accomplish key task(s) if you will use Polybase Export method.

The modules are summarized as below:

1_TranslateTableDDLs: Translate SQL objects (DDLs) from source system format to Azure Synapse format. The output is stored as .sql files in specified file folder (configurable).

1A_ExtractCodeDDLs: Script out SQL Objects (Views, Functions, Stored Procedures, and Triggers) and write each object into a separate .sql file.

1B_MapDatabasesAndScheams: Map SQL Server Database & Schema into Synapse Database and Schema. In addition, unsupported data types in the code are mostly discovered and counted.

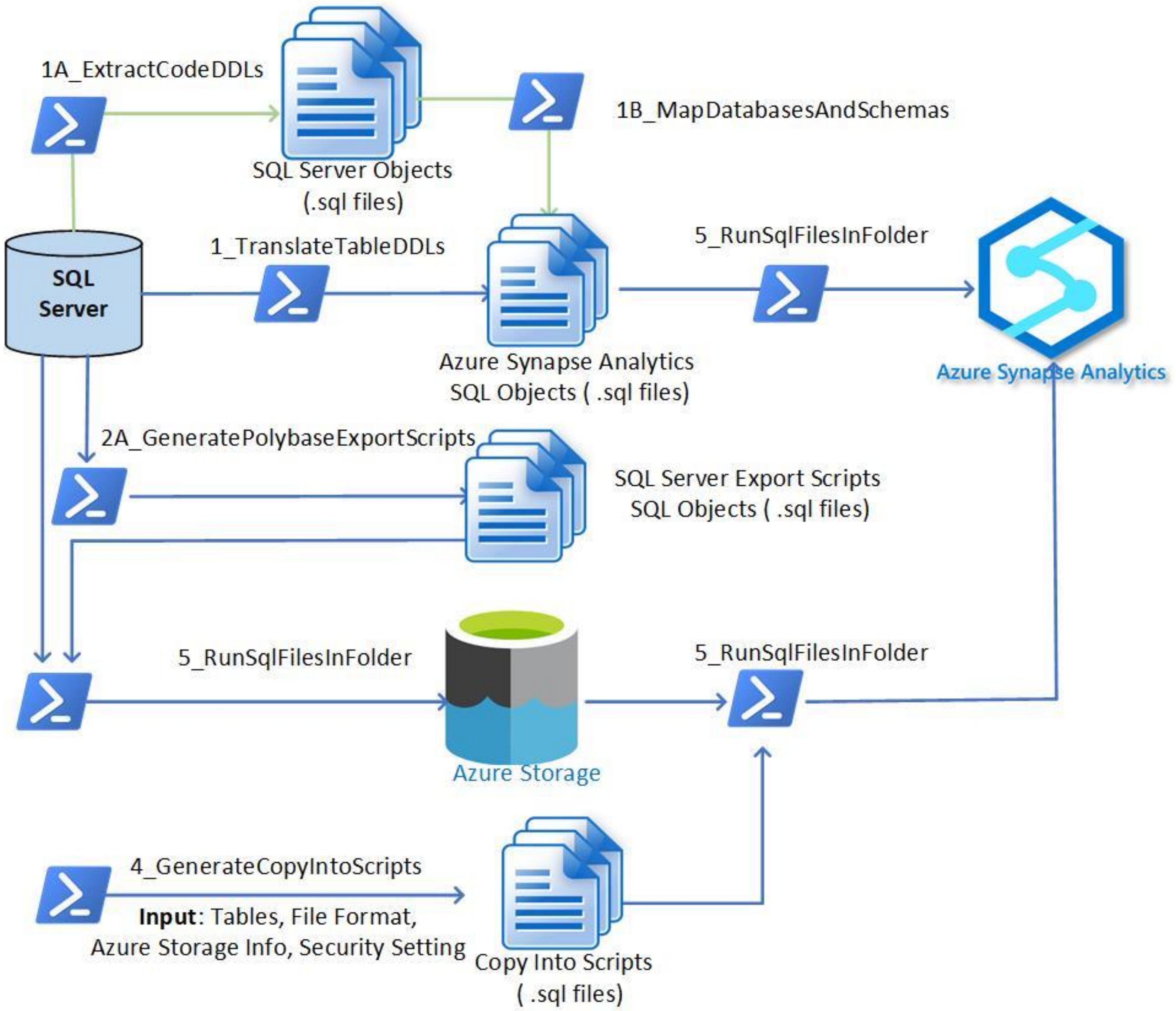
2A_GeneratePolybaseExportScripts: Generate Polybase Export T-SQL Script for each table. Polybase export set up examples are provided in subfolder "Utilities" inside this module.

4_GenerateCopyIntoScripts: Generate "COPY Into" T-SQL Scripts that will move data from Azure Storage into Azure Synapse SQL Pool tables, once executed.

5_RunSqlFilesInFolder: Run all T-SQL Scripts defined in .sql files stored in a specified file folder. The T-SQL Scripts can be DDL, DML, Data Movement Scripts (such as Copy Into scripts or Polybase Export Scripts), or any other scripts such as create/update statistics or indexes. In fact, this module is designed to run all T-SQL scripts in a folder (against Azure Synapse or SQL Server).

Note: **5_RunSqlFilesInFolder is reused three times in the process:** (1) Run T-SQL Scripts generated from 1_TranslateMetadata, (2) Run T-SQL Scripts generated from 2A_GeneratePolybaseExportScripts, and (3) Run T-SQL Scripts generated from 4_GenerateCopyIntoScripts. You can use module 5_RunSqlFileInFolder to execute code manually translated, mapped by module 1B or translated by Azure Synapse Pathway.

SQL Server to Synapse Migration Tasks Carried out by PowerShell Scripts (Modules) – Polybase Export – SQL Servers 2016+

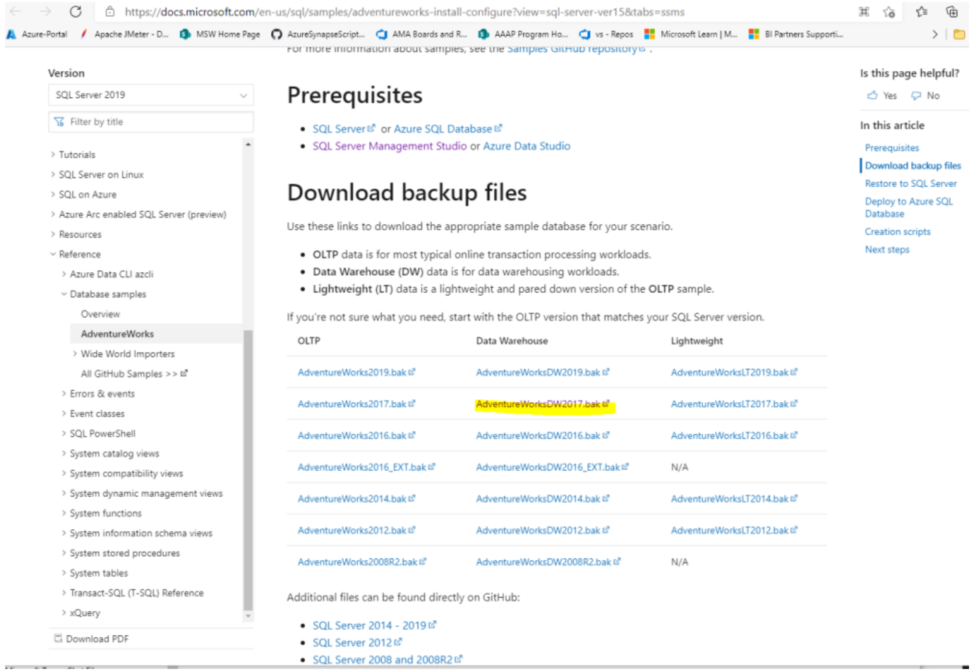


To Test the Scripts with an actual SQL Server & Azure Synapse

(1) Set up your Azure Synapse Workspace with Azure SQL Pool Created

(2) Download **AdventureWorksDW2017.bak** from the link below and restore it to your SQL Server.

AdventureWorks sample databases - SQL Server | Microsoft Docs



For more information about samples, see the Samples GitHub repository.

Prerequisites

- SQL Server or Azure SQL Database
- SQL Server Management Studio or Azure Data Studio

Download backup files

Use these links to download the appropriate sample database for your scenario.

- **OLTP** data is for most typical online transaction processing workloads.
- **Data Warehouse (DW)** data is for data warehousing workloads.
- **Lightweight (LT)** data is a lightweight and pared down version of the **OLTP** sample.

If you're not sure what you need, start with the OLTP version that matches your SQL Server version.

OLTP	Data Warehouse	Lightweight
AdventureWorks2019.bak	AdventureWorksDW2019.bak	AdventureWorksLT2019.bak
AdventureWorks2017.bak	AdventureWorksDW2017.bak	AdventureWorksLT2017.bak
AdventureWorks2016.bak	AdventureWorksDW2016.bak	AdventureWorksLT2016.bak
AdventureWorks2016_EXT.bak	AdventureWorksDW2016_EXT.bak	N/A
AdventureWorks2014.bak	AdventureWorksDW2014.bak	AdventureWorksLT2014.bak
AdventureWorks2012.bak	AdventureWorksDW2012.bak	AdventureWorksLT2012.bak
AdventureWorks2008R2.bak	AdventureWorksDW2008R2.bak	N/A

Additional files can be found directly on GitHub:

- SQL Server 2014 - 2019
- SQL Server 2012
- SQL Server 2008 and 2008R2

How to Restore SQL Server Database to .bak file

Restore a Database Backup Using SSMS - SQL Server | Microsoft Docs

Download SQL Server 2017 if you do not have a SQL Server already

https://www.microsoft.com/en-us/sql-server/sql-server-2017

Download SSMS if you don’t have it already

Download SQL Server Management Studio (SSMS) - SQL Server Management Studio (SSMS) | Microsoft Docs

5-Step Migration Process with BCP Export

- You can use this method for all versions of SQL Servers

If you have SQL Server 2016 or later, you have additional option to export SQL server data directly into Azure Storage. See details later in this presentation.

(Polybase is available only for SQL Server 2016 or later)

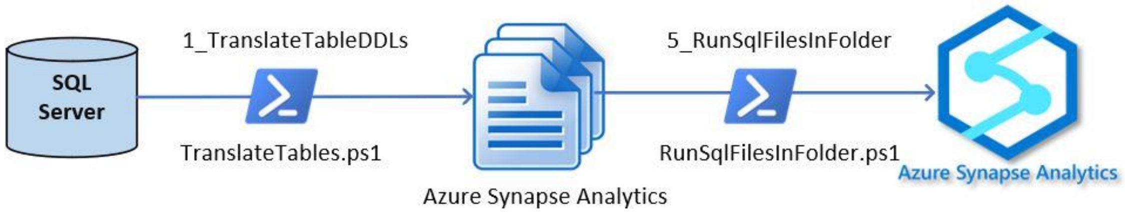
Export SQL Server Tables into Local .csv or .parquet

# Step	BCP Export	Results
1 – Task 1	**Translate SQL Tables** (Module 1_TranslateTableDDLs)	Azure Synapse Table DDLs generated
1 – Task 2	**Execute Synapse Tables** DDLs (Module 5_RunSqlFilesInfolder)	Tables Created in Azure Synapse.
1 – Task 3	**Map Databases and Schemas** (Module 1A_ExtractCodeDDLs, and Module 1B_MapDatabasesAndSchemas)	Code (Views, Stored Procedures, Functions) have the desired Azure Synapse DB name and Schema Name
1 – Task 4	**Perform Additional Code Translation** (Manually or using Azure Synapse Pathway) and then Execute Code using (Module 5_RunSqlFilesInfolder)	Code is migrated into Azure Synapse
2 – Export to .csv 2B – Export to .parquet	**Export SQL Server Table data into .csv files** (Module 2_ExportSourceDataWithBCP) **Export SQL Server Table data into .parquet files** (Module 2B_ExportSourceDataToParquet)	Data is exported to local files (.csv or .parquet)
3 – Upload	**Upload exported Data into Azure Storage** (3_LoadDataIntoAzureStorage)	Data loaded to Azure Storage
4	**Generate COPY Into Scripts** (Module 4_GenerateCopyIntoScripts)	Copy Into T-SQL Scripts generated
5	**Execute COPY Into Scripts** (Module 5_RunSqlFilesInFolder)	Data imported to Synapse

Step 1

(Same for BCP and Polybase Export)

Step 1: Code (DDLs) Migration – Translate SQL Server Tables and create them in Azure Synapse



SQL Server

1_TranslateTableDDLs

TranslateTables.ps1

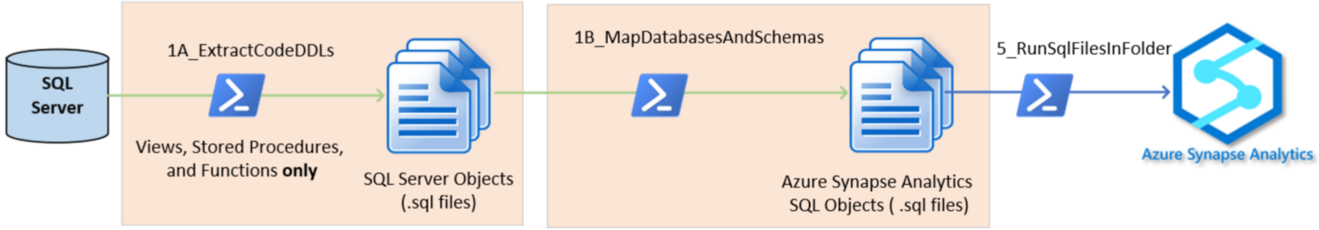
Azure Synapse Analytics

5_RunSqlFilesInFolder

RunSqlFilesInFolder.ps1

Azure Synapse Analytics

Step 1AB



SQL Server

1A_ExtractCodeDDLs

Views, Stored Procedures, and Functions **only**

SQL Server Objects (.sql files)

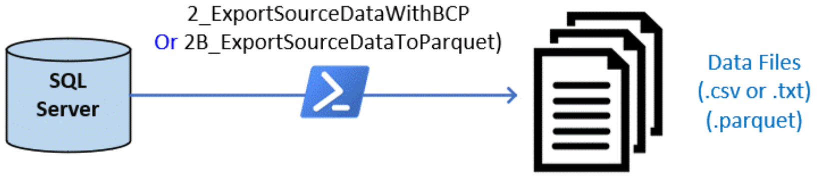
1B_MapDatabasesAndSchemas

Azure Synapse Analytics SQL Objects (.sql files)

5_RunSqlFilesInFolder

Azure Synapse Analytics

Step 2 or 2B



SQL Server

2_ExportSourceDataWithBCP

Or 2B_ExportSourceDataToParquet)

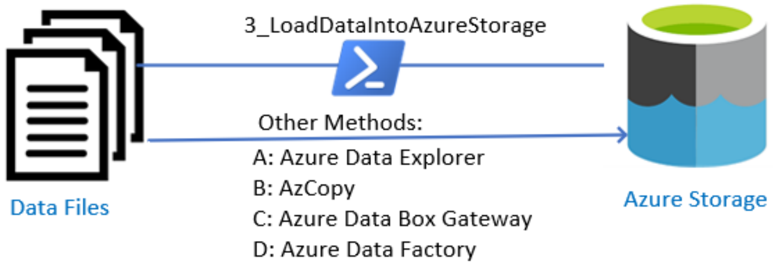
Data Files

(.csv or .txt)

(.parquet)

Step 3

Step 3: Upload Exported Data Files into Azure Storage (Azure Data Lake Store or Blob Storage)



3_LoadDataIntoAzureStorage

Other Methods:

A: Azure Data Explorer

B: AzCopy

C: Azure Data Box Gateway

D: Azure Data Factory

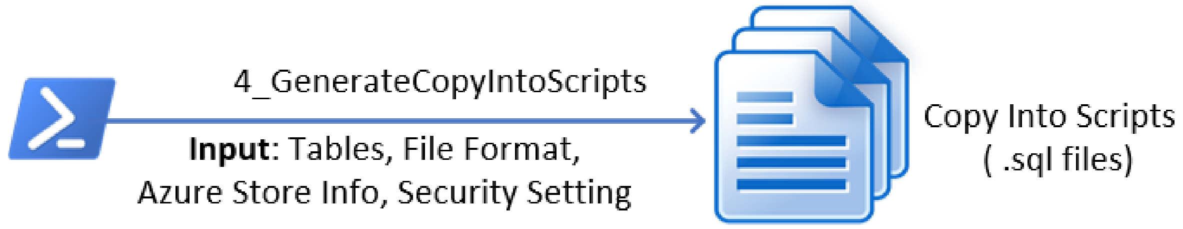
Data Files

Azure Storage

Note: Step 4 and Step 5 are the same for BCP or Polybase Export

Step 4

Step 4: Generate COPY Into T-SQL Script for Tables (to run in Azure Synapse Dedicated SQL POOL)



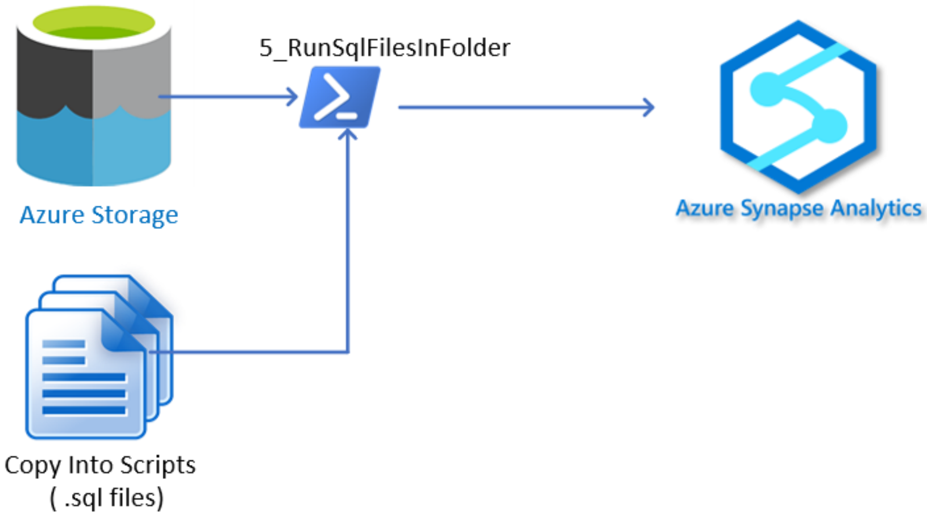
4_GenerateCopyIntoScripts

Input: Tables, File Format,
Azure Store Info, Security Setting

Copy Into Scripts
(.sql files)

Step 5

Step 5: Import Data From Azure Storage into Azure Synapse Dedicated SQL POOL)



5_RunSqlFilesInFolder

Azure Storage

Azure Synapse Analytics

Copy Into Scripts
(.sql files)

5-Step Migration Process with Polybase Export

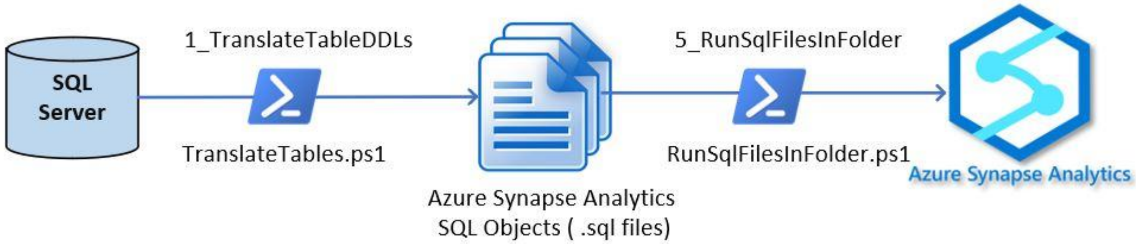
(Polybase is available only for SQL Server 2016 or later)

Polybase Export Path

# Step	Polybase Export	Results
1 – Task 1	**Translate SQL Tables** (Module 1_TranslateTableDDLs)	Azure Synapse Table DDLs generated
1 – Task 2	**Execute Synapse Tables** DDLs (Module 5_RunSqlFilesInfolder)	Tables Created in Azure Synapse.
1 – Task 3	**Map Databases and Schemas** (Module 1A_ExtractCodeDDLs, and Module 1B_MapDatabasesAndSchemas)	Code (Views, Stored Procedures, Functions) have the desired Azure Synapse DB name and Schema Name
1 – Task 4	**Perform Additional Code Translation** (Manually or using Azure Synapse Pathway) and then Execute Code using (Module 5_RunSqlFilesInfolder)	Code is migrated into Azure Synapse
2A – Generate Polybase Code	**Generate Polybase Export Scripts** (Module 2A_GeneratePolybaseExportScripts)	Polybase Export T-SQL Scripts generated
3A – Run Polybase Code	**Execute Polybase Export Scripts (**Module 5_RunSqlFilesInfolder)	Data exported directly to Azure Storage from SQL Server
4	**Generate COPY Into Scripts** (Module 4_GenerateCopyIntoScripts)	Copy Into T-SQL Scripts generated
5	**Execute COPY Into Scripts** (Module 5_RunSqlFilesInFolder)	Data imported to Synapse

Step 1: Code (DDLs) Migration – Translate SQL Server Tables and create them in Azure Synapse

Step 1
(Same for BCP
and Polybase
Export)



SQL Server

1_TranslateTableDDLs

TranslateTables.ps1

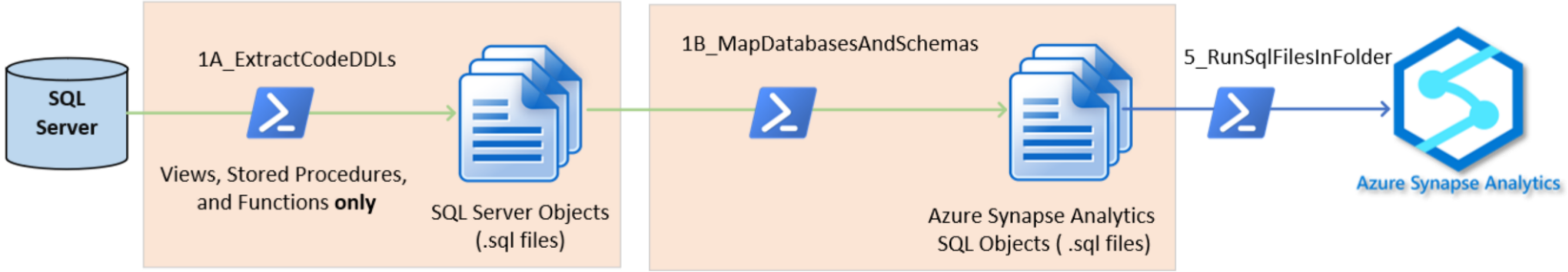
Azure Synapse Analytics
SQL Objects (.sql files)

5_RunSqlFilesInFolder

RunSqlFilesInFolder.ps1

Azure Synapse Analytics

Step 1AB



SQL Server

1A_ExtractCodeDDLs

Views, Stored Procedures,
and Functions **only**

SQL Server Objects
(.sql files)

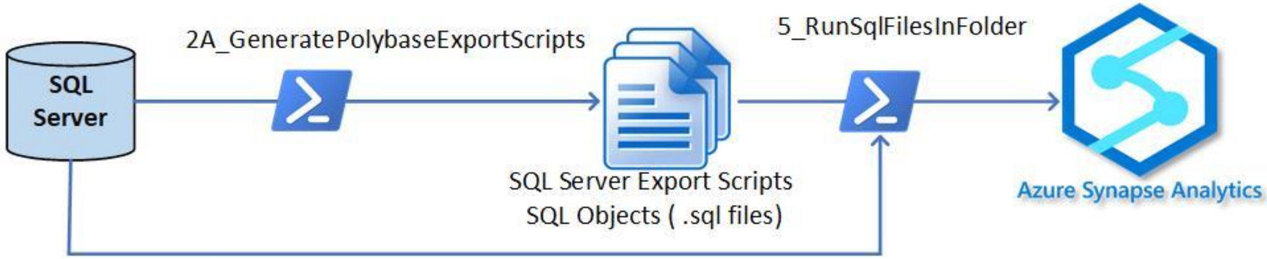
1B_MapDatabasesAndSchemas

Azure Synapse Analytics
SQL Objects (.sql files)

5_RunSqlFilesInFolder

Azure Synapse Analytics

Step 2A & 3A



SQL Server

2A_GeneratePolybaseExportScripts

SQL Server Export Scripts
SQL Objects (.sql files)

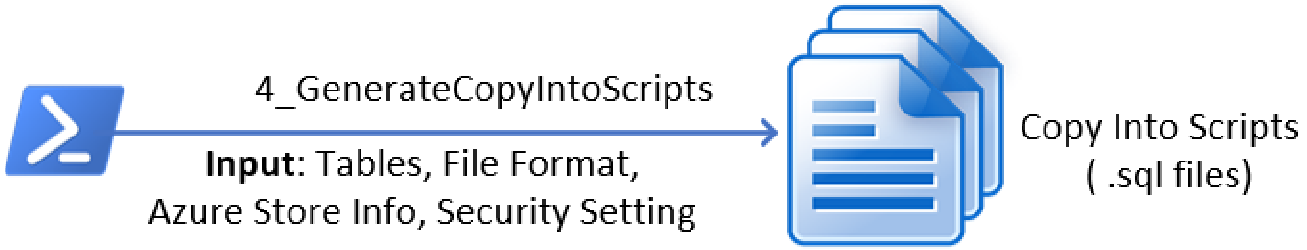
5_RunSqlFilesInFolder

Azure Synapse Analytics

Note: Step 4 and Step 5 are the same for BCP or Polybase Export

Step 4

Step 4: Generate COPY Into T-SQL Script for Tables (to run in Azure Synapse Dedicated SQL POOL)



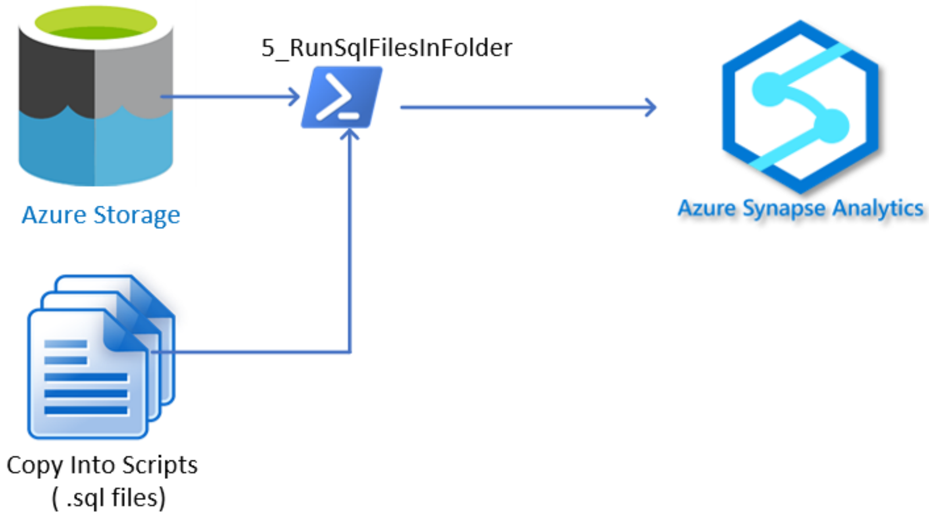
4_GenerateCopyIntoScripts

Input: Tables, File Format,
Azure Store Info, Security Setting

Copy Into Scripts
(.sql files)

Step 5

Step 5: Import Data From Azure Storage into Azure Synapse Dedicated SQL POOL)



Azure Storage

5_RunSqlFilesInFolder

Azure Synapse Analytics

Copy Into Scripts
(.sql files)

DATA
& AI

Microsoft Azure Solutions Business

Thank you