

Integrating Power BI with Azure Synapse Analytics

Thursday March 17th 2022

Andy Cutler

»» MICROSOFT ««
POWER BI - UK
USER GROUP

Microsoft
Partner

Silver Data Analytics
Silver Data Platform



Microsoft®
Most Valuable
Professional

Andy Cutler

Independent BI/DW Consultant

Azure Data Platform & Power BI



datahai.co.uk/blog

serverlesssql.com

twitter.com/MrAndyCutler

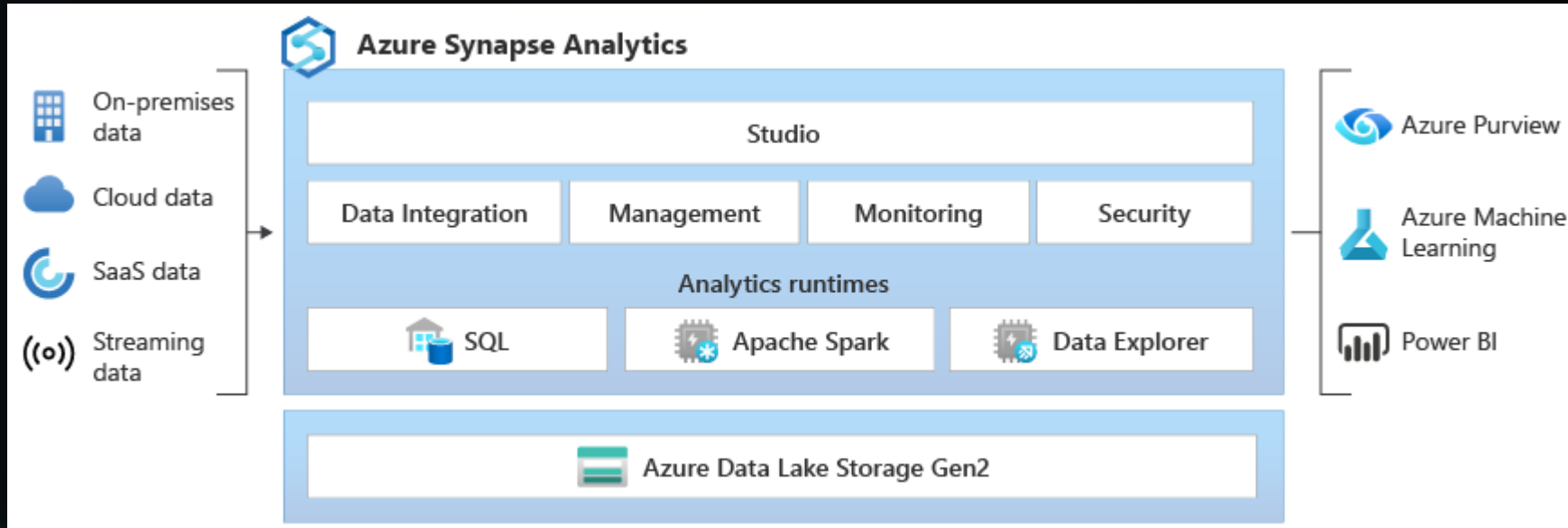
linkedin.com/in/andycutler/



SCAN ME

» MICROSOFT «
POWER BI - UK
USER GROUP

Azure Synapse Analytics



Dedicated SQL Pools to run SQL data warehousing workloads in a Massively Parallel Processing (MPP) system

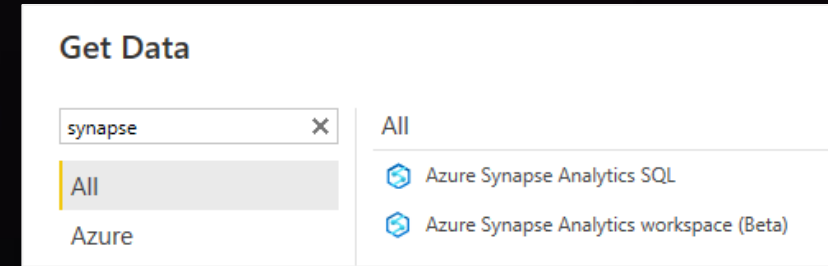
- Import Data into Dedicated SQL Pools Tables with appropriate indexing (ColumnStore, Clustered) and data distribution

Serverless SQL Pools to run SQL workloads over data in Azure Storage, Data Lake, CosmosDB, Dataverse

- Use T-SQL to run SELECT commands over data stored externally, without moving data

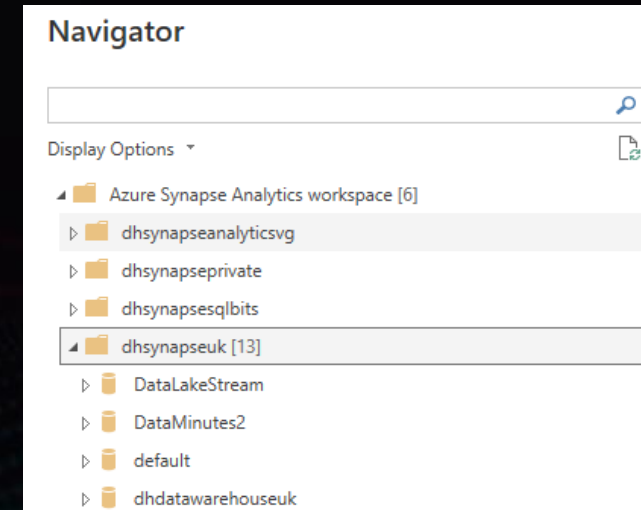
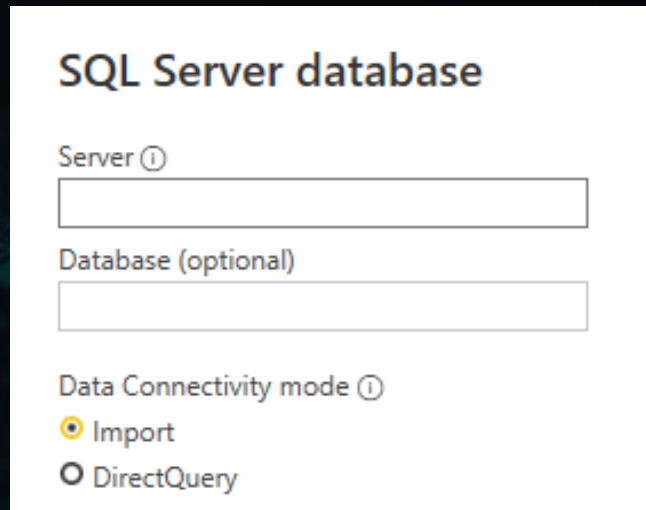
Connecting to Synapse Analytics

There are 2 connectors in Power BI Desktop



Azure Synapse Analytics SQL

Azure Synapse Analytics workspace (Beta)



DirectQuery & Import

Import only

Features



Synapse Analytics

Materialized Views

Result-Set Caching

Workload Classification



Power BI

Aggregations (Manual & Automatic)

Incremental Refresh

Hybrid Tables (Import & DirectQuery)



Dedicated SQL Pools

Query concurrency depends on:

- The DWU tier the pool is currently set to
- The Resource Class a user is assigned to when running a query

Maximum Number of Queries per DWU tier:

DWU100c



DWU200c



DWU500c



DWU1000c



DWU30000c



RAM (GB) 60

120

600

1,200

18,000



Serverless SQL Pools

Serverless SQL Pools are always on, no need to resume/pause/stop

Resources are allocated automatically, there are no scaling options (no DWU concept)

Best Practices

- Try and ensure Power BI, Synapse Analytics, and Azure Storage are in the same region
- Use Parquet file format where possible
- Assign the smallest data types possible
- File sizes evenly ranged from 100MB to 10GB each



Serverless SQL Pools

We can use a SQL metadata function **filepath()** to “partition prune” when querying data in Azure Storage



Data Lake



Serverless SQL Pools View



Power BI Query

```
\EventYear=2022\  
  \EventMonth=03\  
    \EventDate=2022-03-01  
    \EventDate=2022-03-02  
    \EventDate=2022-03-03
```

```
CREATE VIEW...  
AS  
SELECT filepath(1) AS EventYear,  
       filepath(2) AS EventMonth,  
       filepath(3) AS EventDate,  
       ...
```

```
SELECT...  
...WHERE EventDate = '2022-03-02'
```

EventDate is the 3rd level in the folder hierarchy



Demo

Dedicated SQL Pools – Materialized Views

Serverless SQL Pools – Partition Filtering

New Certification



Exam DP-500: Designing and Implementing Enterprise-Scale Analytics Solutions Using Microsoft Azure and Microsoft Power BI (beta)

Candidates for this exam should have advanced Power BI skills, including managing data repositories and data processing in the cloud and on-premises, along with using Power Query and Data Analysis Expressions (DAX). They should also be proficient in consuming data from Azure Synapse Analytics and should have experience querying relational databases, analyzing data by using Transact-SQL (T-SQL), and visualizing data.

📌 Important

This exam will be available in mid-April.

Part of the requirements for: [Microsoft Certified: Azure Enterprise Data Analyst Associate](#)

Related exams: none

Important: [See details](#)

[Go to Certification Dashboard](#)

Available ~mid-April

Covers:

- Power BI
- Azure Purview
- Azure Synapse Analytics
- Azure Data Lake Gen2

[DP-500: Designing and Implementing Enterprise-Scale Analytics Solutions Using Microsoft Azure and Microsoft Power BI \(beta\) - Learn | Microsoft Docs](#)

» MICROSOFT «
POWER BI - UK
USER GROUP

References

- [Memory and concurrency limits - Azure Synapse Analytics | Microsoft Docs](#)
- [Quickstart: Linking a Power BI workspace to a Synapse workspace - Azure Synapse Analytics | Microsoft Docs](#)
- [Performance tuning with materialized views - Azure Synapse Analytics | Microsoft Docs](#)
- [Performance tuning with result set caching - Azure Synapse Analytics | Microsoft Docs](#)
- [Quickstart: Create a workload classifier - T-SQL - Azure Synapse Analytics | Microsoft Docs](#)
- [Azure SQL DW Materialized Views \(part 1\) – Niko Neugebauer \(nikoport.com\)](#)
- [Quickstart: Create a workload classifier - Portal - Azure Synapse Analytics | Microsoft Docs](#)
- [COPY INTO \(Transact-SQL\) - \(Azure Synapse Analytics\) - SQL Server | Microsoft Docs](#)
- <https://docs.microsoft.com/en-us/power-bi/enterprise/aggregations-auto>
- <https://powerbi.microsoft.com/en-us/blog/announcing-public-preview-of-hybrid-tables-in-power-bi-premium/>
- [Monitor your dedicated SQL pool workload using DMVs - Azure Synapse Analytics | Microsoft Docs](#)
- [Copy and transform data in Azure Synapse Analytics - Azure Data Factory & Azure Synapse | Microsoft Docs](#)
- [Capacity limits for dedicated SQL pool - Azure Synapse Analytics | Microsoft Docs](#)
- [DBCC PDW_SHOWMATERIALIZEDVIEWOVERHEAD \(Transact-SQL\) - SQL Server | Microsoft Docs](#)
- [Monitor your dedicated SQL pool workload using DMVs - Azure Synapse Analytics | Microsoft Docs](#)
- [GitHub - JocaPC/qpi: Query Performance Insight - analyze performance of your SQL Server Database Engine](#)