Turbocharge Power BI using Azure Synapse Analytics

Azure Synapse Analytics
Data Toboggan
Saturday June 12th

Andy Cutler

Independent Consultant & Contractor

Azure Data Platform & Power BI

www.datahai.co.uk

https://twitter.com/MrAndyCutler

https://www.linkedin.com/in/andycutler/

SCAN ME

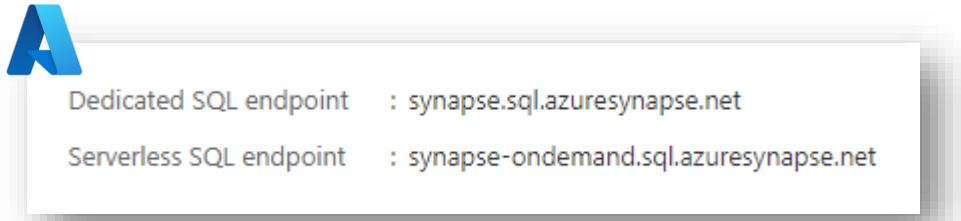
datahai

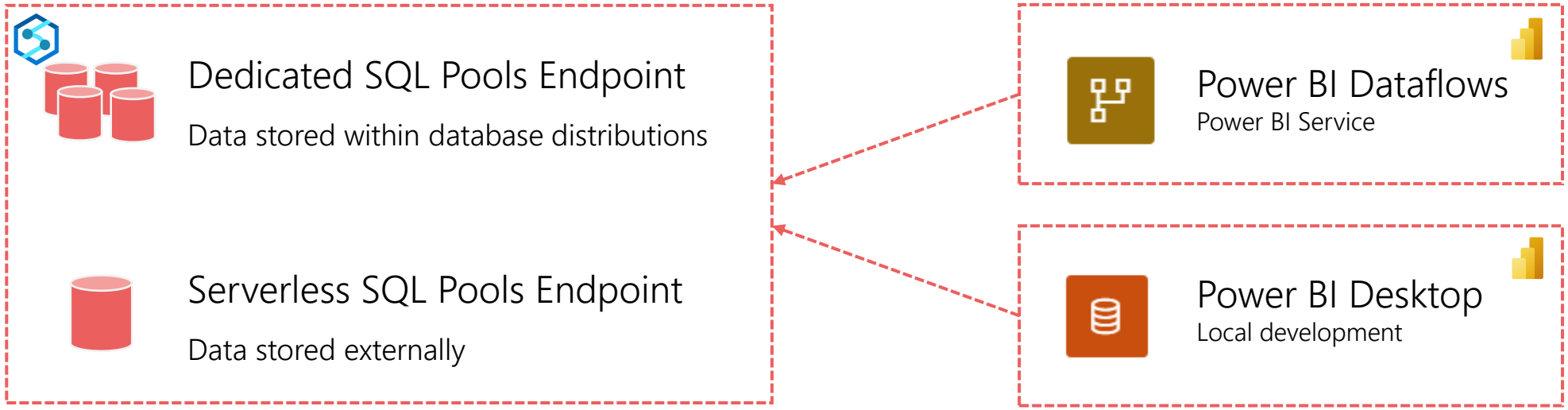
Session Overview

- Connecting to SQL Pools
- Using Serverless SQL Pools to load external data
- Creating Materialized Views in Dedicated SQL Pools
- Further reading

Connecting to SQL Pools

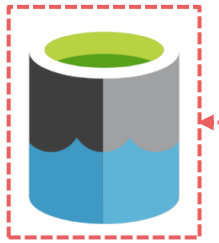
Both Dedicated SQL Pools and Serverless SQL Pools expose a SQL endpoint which external tools can connect to.





Using Serverless SQL Pools to load external data

We can connect to CSV, JSON, & Parquet file data in Azure Storage using Serverless SQL Pools and run SQL queries on this data.



Azure Blob Storage
and Data Lake Gen1/2

```
CREATE VIEW LDW.vwSalesOrdersAggregates
AS
SELECT SalesOrderID,
       OrderDate,
       CustomerID,
       SUM(OrderTotal) AS TotalOrderAmount
FROM
OPENROWSET
(
    BULK 'sourcedatapartitionsalesorder/*/*.csv',
    DATA_SOURCE = 'ExternalDataSourceDataLake',
    FORMAT = 'CSV',
    PARSER_VERSION = '2.0',
    HEADER_ROW = TRUE,
    FIELDTERMINATOR ='|'
) AS fct
GROUP BY SalesOrderID,
       OrderDate,
       CustomerID
```

DirectQuery
or Import

Power BI
Desktop/Dataflows

Connect to External Tables and Views and run SQL via Power BI. Power BI can use Query Folding to push SQL to Serverless SQL Pools

Scenarios include aggregating large source data for loading into Power BI.

Creating Materialized Views in Dedicated SQL Pools

We can create a Materialized View in Dedicated SQL Pools to pre-calculate and store dataset ready for use in Power BI.

Data stored within Dedicated SQL Pools databases

```
CREATE MATERIALIZED VIEW vwCustomerSalesAnalysis
WITH (DISTRIBUTION=HASH(customer id))
AS
SELECT DC.customer_id AS customer_id
       ,DC.first_name AS customer_first_name
       ,DC.last_name AS customer_last_name
       ,DD.[year] AS SalesYear
       ,sum(isnull(ss_ext_list_price-ss_ext_wholesale_cost
                   -ss_ext_discount_amt+ss_ext_sales_price, 0)/2) year_total
         , count_big(*) AS Sales
FROM DW.DimCustomer DC
INNER JOIN DW.FactSales FS ON FS.customer_id = DC.customer_id
INNER JOIN DW.DimDate DD ON DD.DateKey = FS.SalesDateKey
GROUP BY DC.customer_id
         ,DC.first_name
         ,DC.last_name
         ,DD.[year]
```

DirectQuery or Import

Power BI Desktop/Dataflows

Scenarios include complex logic calculation with data stored in Dedicated SQL Pools.

Further Reading

Microsoft | Docs

- https://docs.microsoft.com/en-us/azure/synapse-analytics/sql-data-warehouse/sql-data-warehouse-overview-what-is
- https://docs.microsoft.com/en-us/azure/synapse-analytics/sql/on-demand-workspace-overview

Serverless SQL Blog

Azure Synapse Analytics Serverless SQL Pools

- https://www.serverlesssql.com/learning/harnessing-azure-synapse-analytics-sql-serverless-in-power-bi-dataflows/
- https://www.serverlesssql.com/getting-started/getting-started-with-azure-synapse-analytics-sql-serverless/