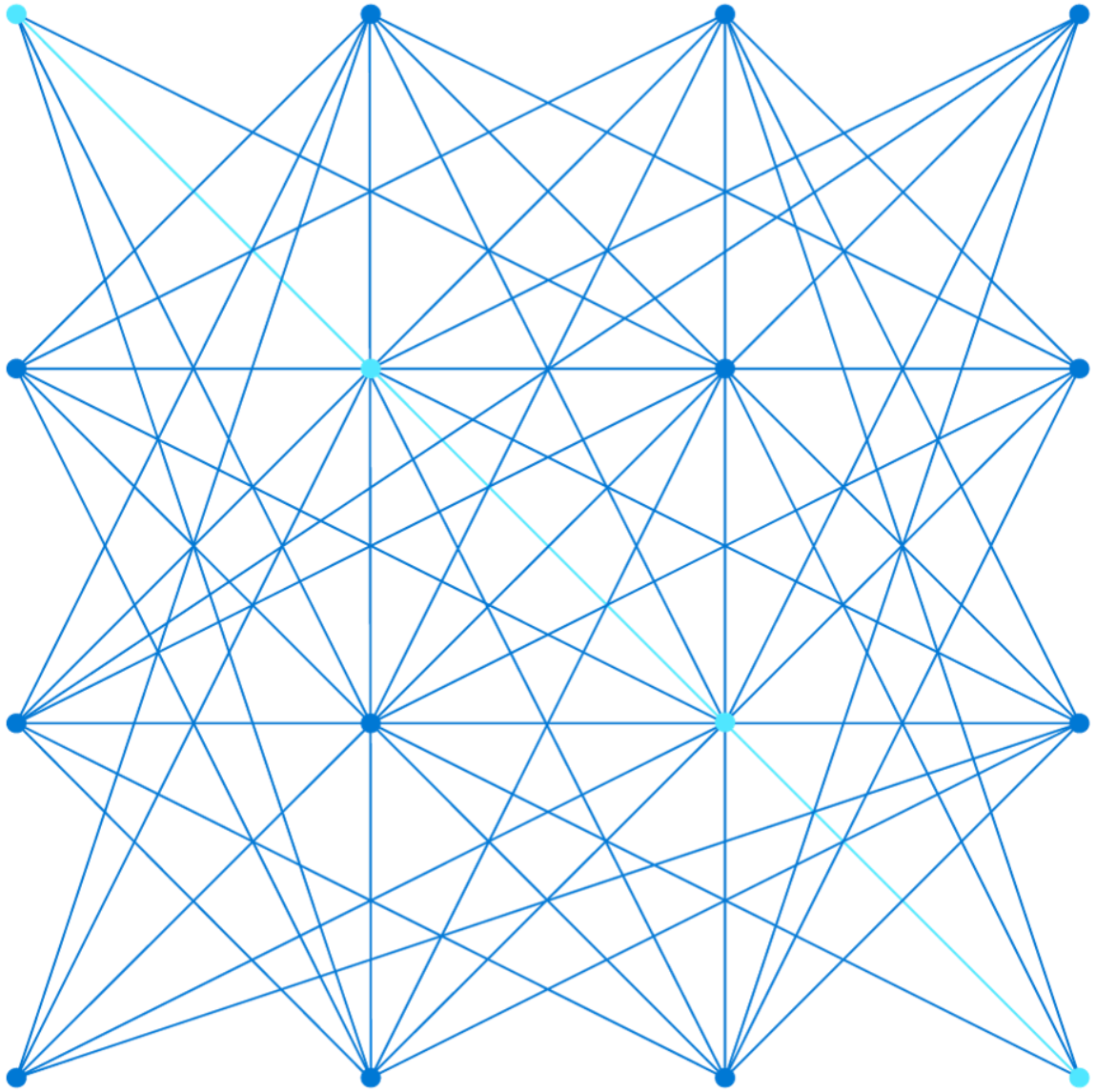
Microsoft Azure

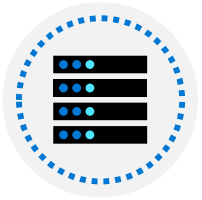
DP-203T00: Support Hybrid Transactional Analytical Processing (HTAP) with Azure Synapse Link



Agenda

Lesson 01 – Design hybrid transactional and analytical processing using Azure Synapse Analytics

Lesson 02 – Configure Azure Synapse Link with Azure Cosmos DB

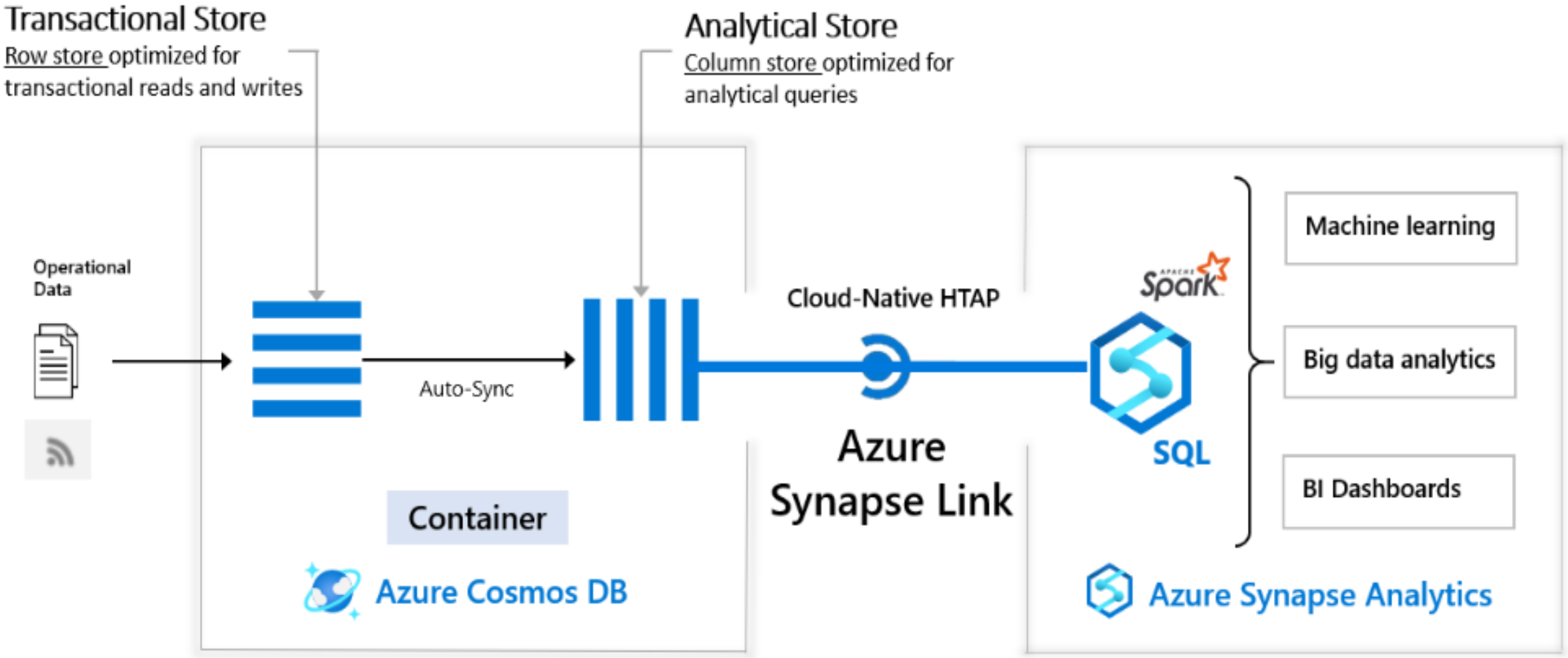


Lesson 03 – Query Azure Cosmos DB with Apache Spark for Azure Synapse Analytics

Lesson 04 – Query Azure Cosmos DB with SQL Serverless for Azure Synapse Analytics

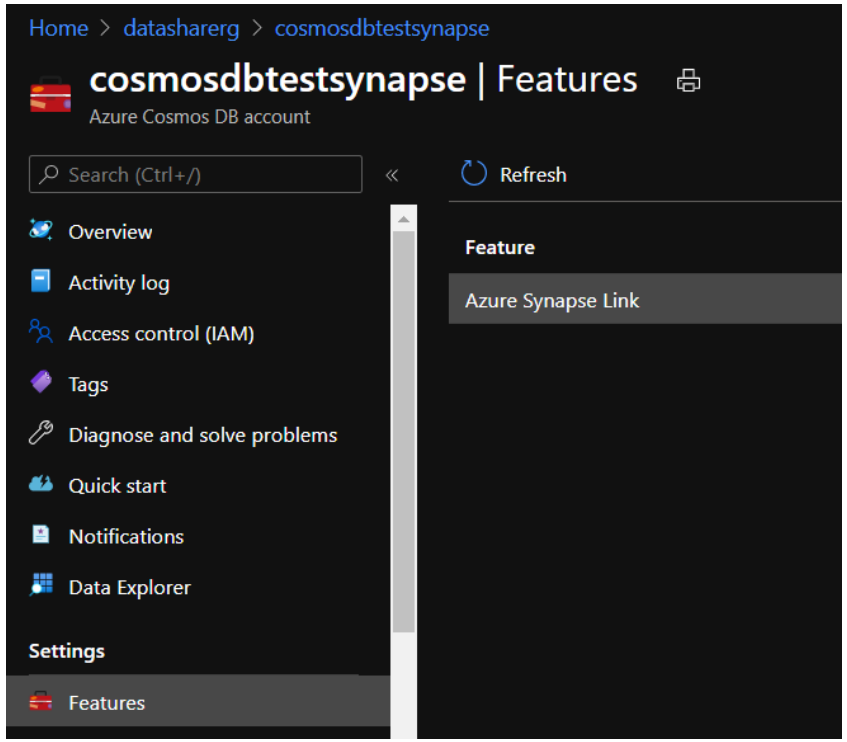
Lesson 01: Design hybrid transactional and analytical processing using Azure Synapse Analytics

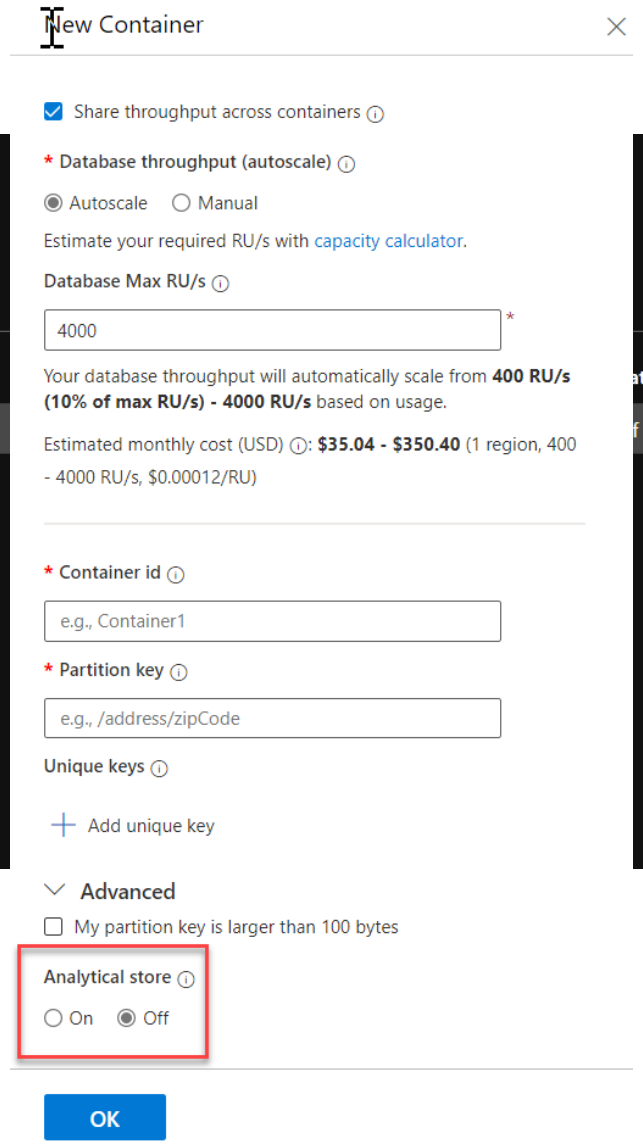
Design hybrid transactional and analytical processing using Azure Synapse Analytics

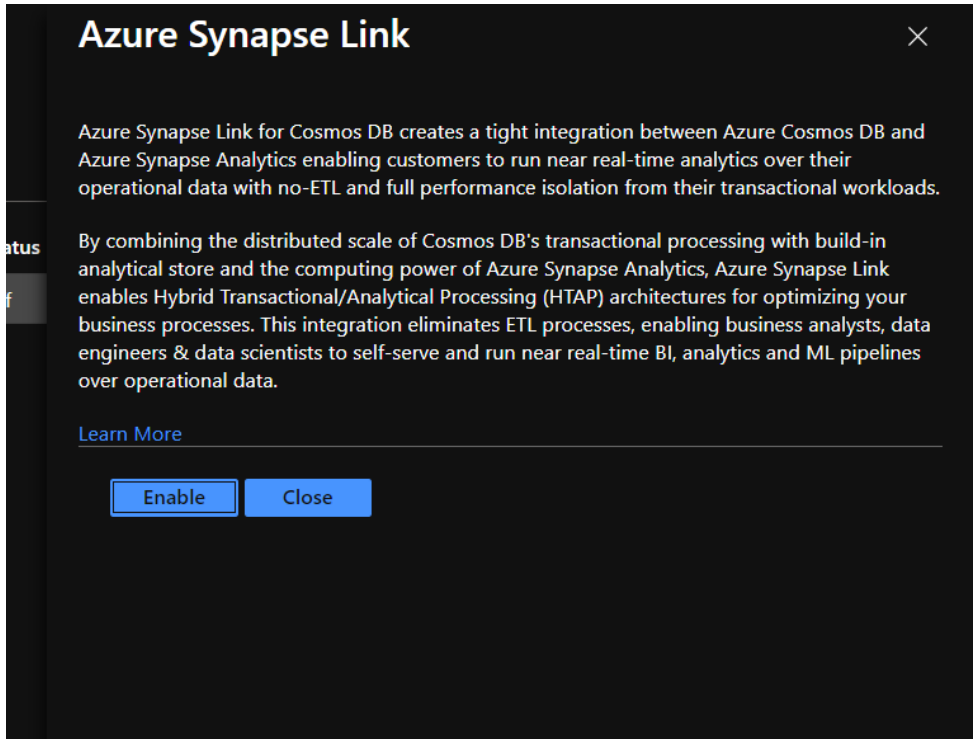


Lesson 02: Configure Azure Synapse Link with Azure Cosmos DB

Configure Azure Synapse Link with Azure Cosmos DB



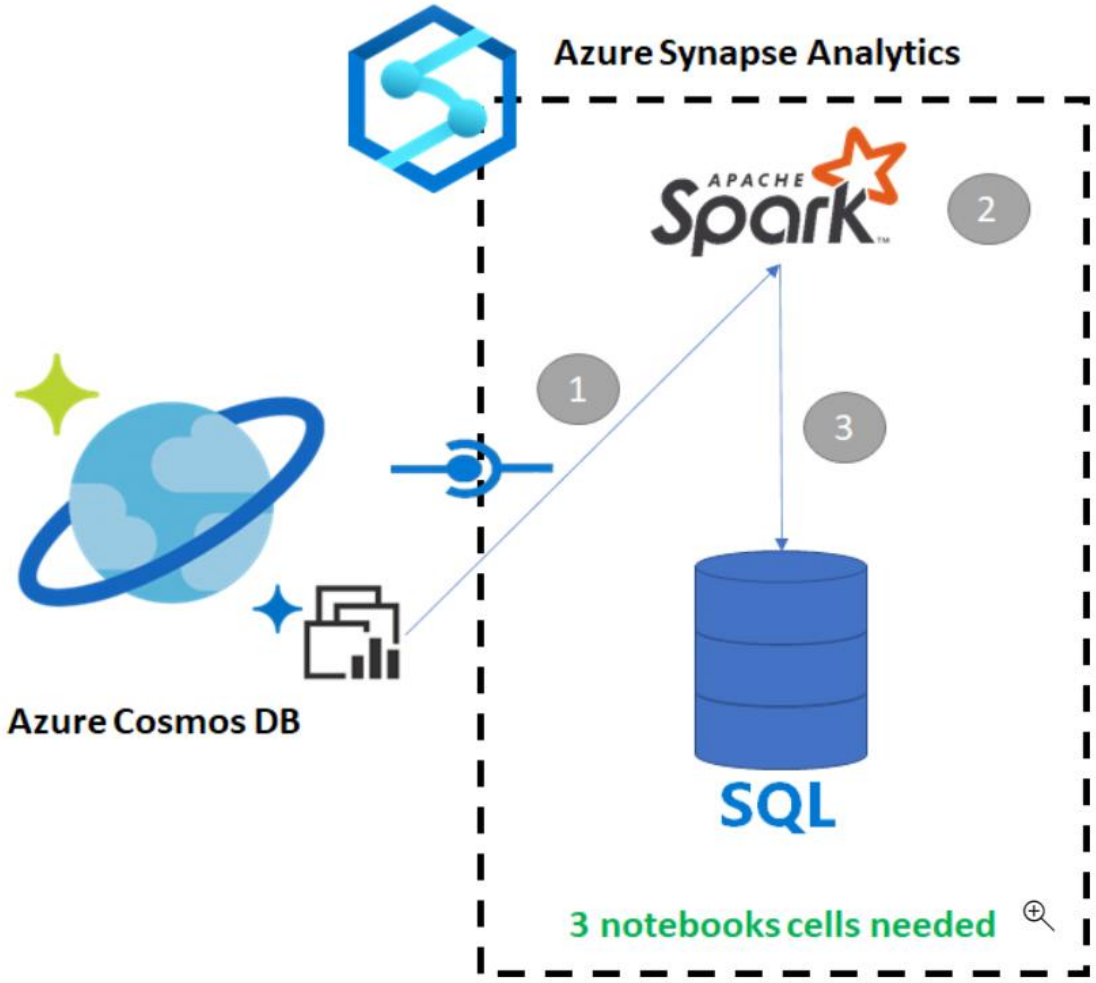




Lesson 03: Query Azure Cosmos DB with Apache Spark for Azure Synapse Analytics

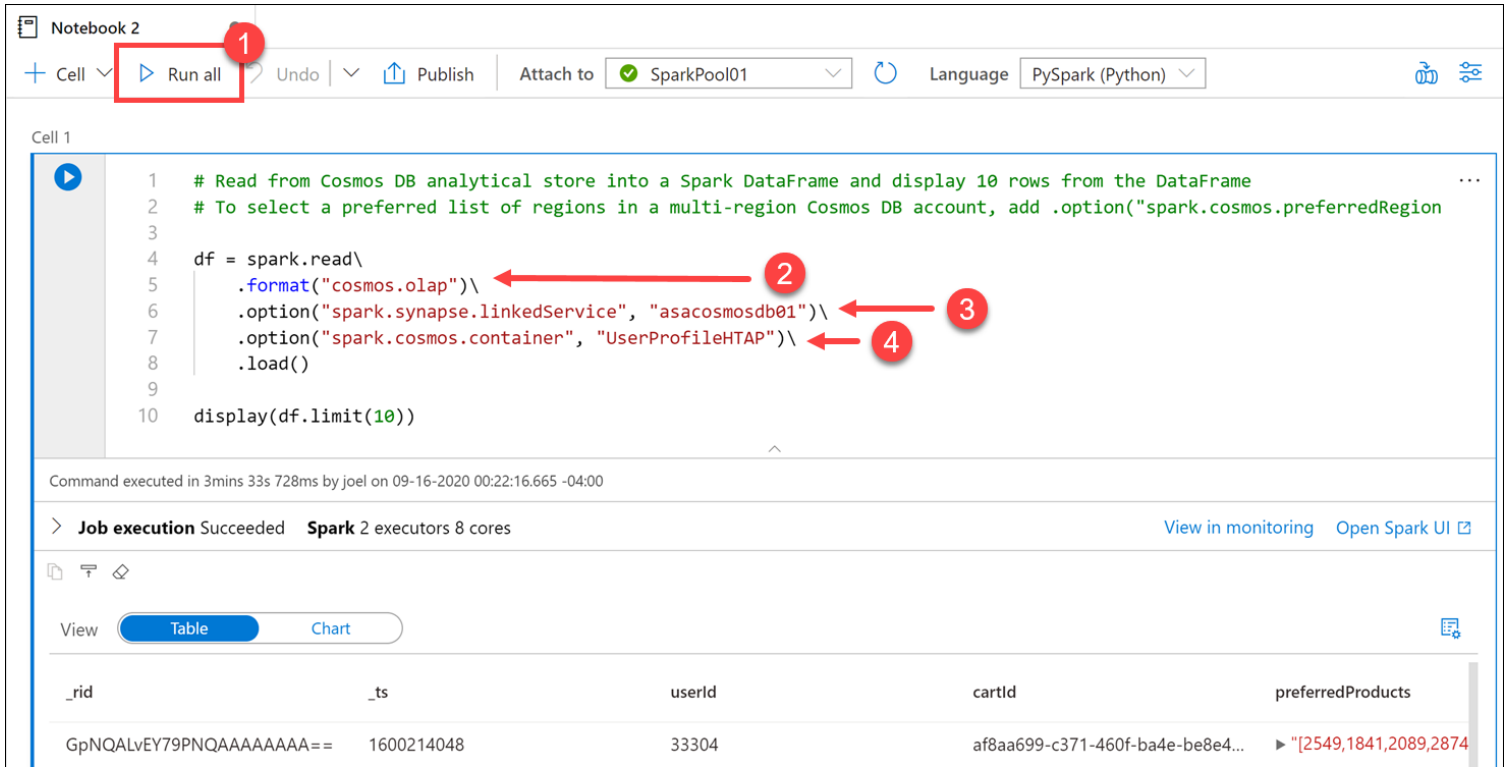
Query Azure Cosmos DB with Apache Spark for Azure Synapse Analytics

- Step 1: Load the data in Spark
- Step 2: Create a base DataFrame
- Step 3: Flatten JSON data
- Step 4: Create the final DataFrame



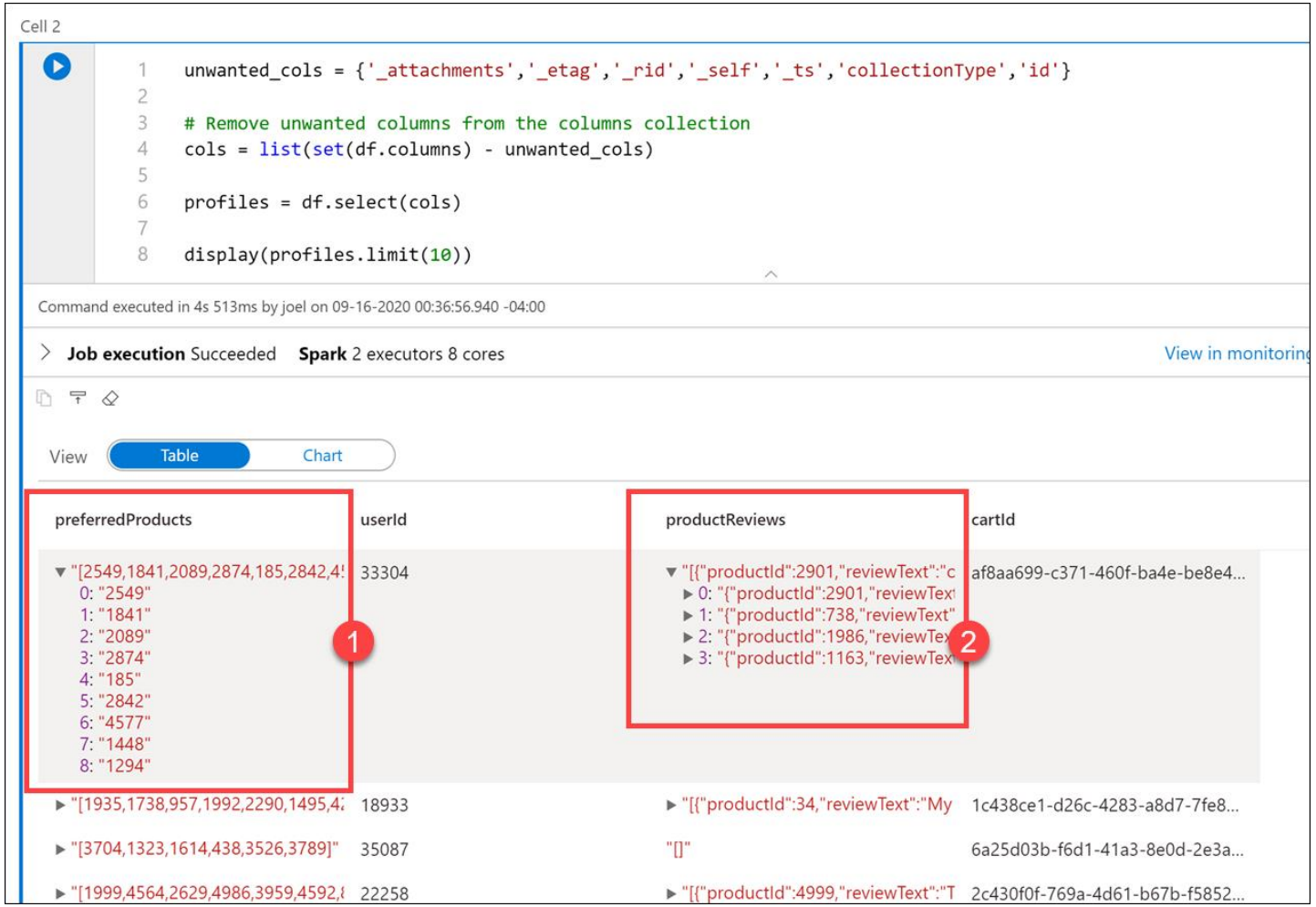
Query Azure Cosmos DB with Apache Spark for Azure Synapse Analytics (continued #1)

Step 1:
Load data to DataFrame



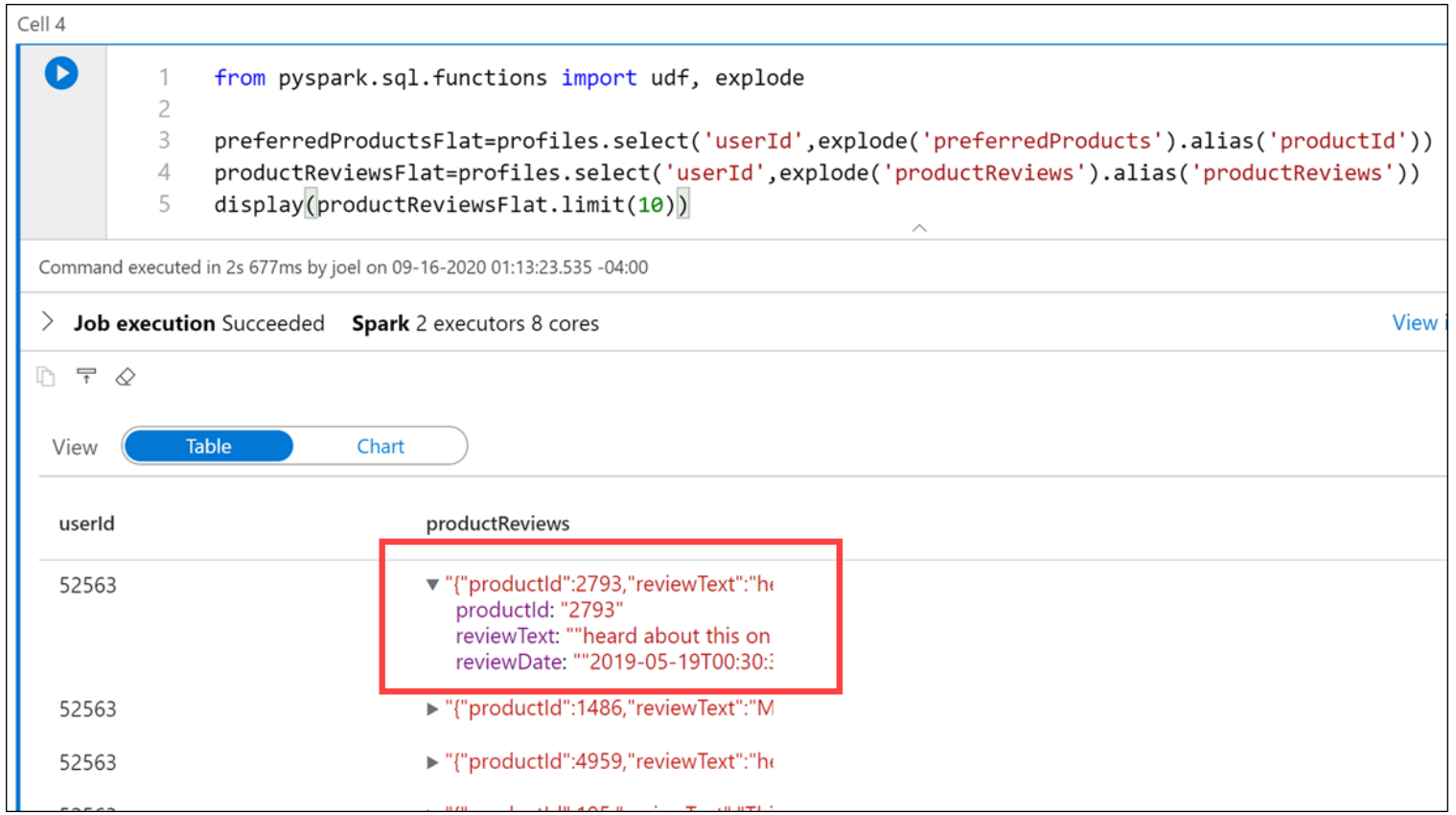
Query Azure Cosmos DB with Apache Spark for Azure Synapse Analytics (continued #2)

Step 2:
Create a base DataFrame



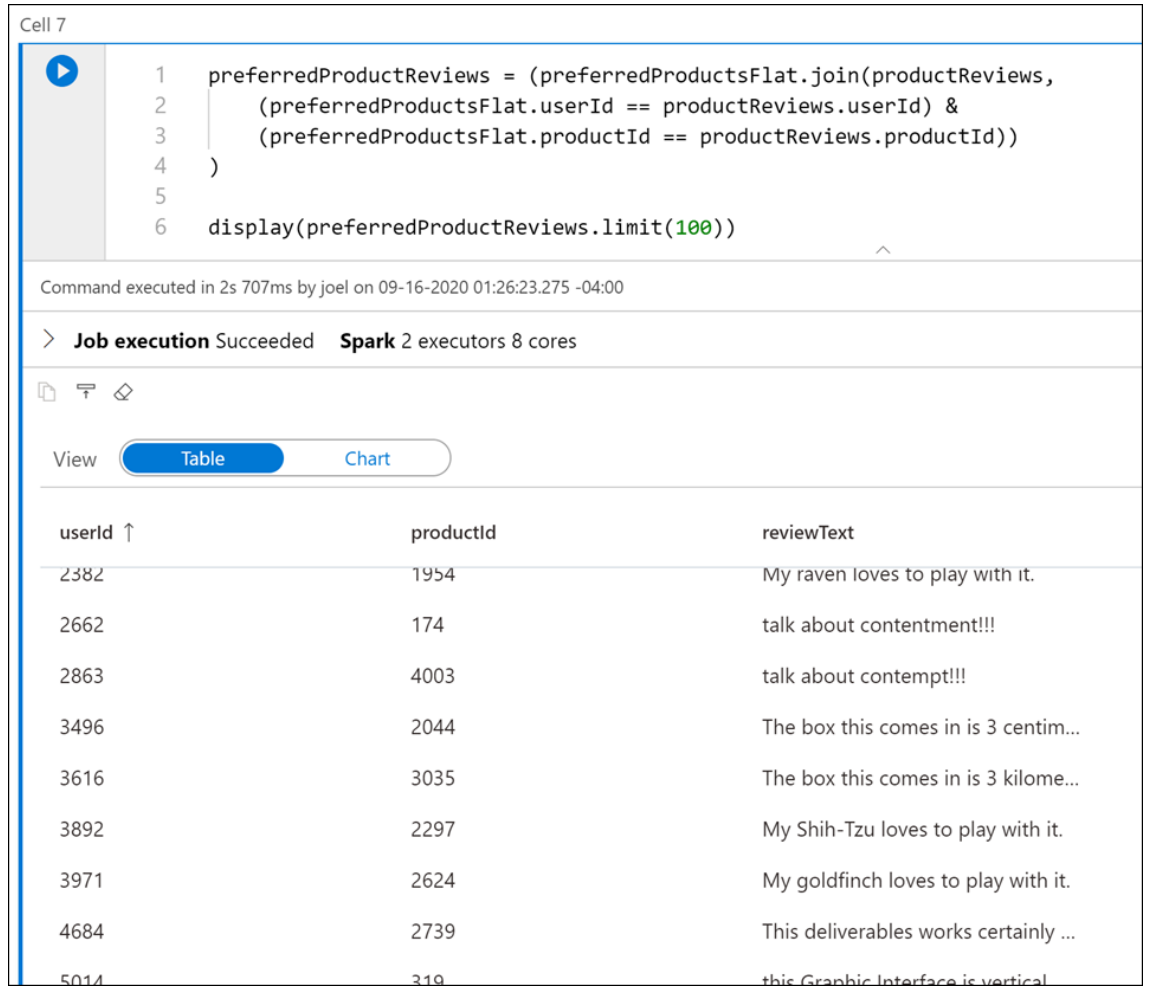
Query Azure Cosmos DB with Apache Spark for Azure Synapse Analytics (continued #3)

Step 3:
Flattening JSON data



Query Azure Cosmos DB with Apache Spark for Azure Synapse Analytics (continued #4)

Step 4:
Creating the final DataFrame



Lesson 04: Query Azure Cosmos DB with SQL Serverless for Azure Synapse Analytics

Query Azure Cosmos DB with SQL Serverless for Azure Synapse Analytics

Step 1:
Create a View

```
15
16  CREATE VIEW UserProfileHTAP                                    1
17  AS
18  SELECT
19      *
20  FROM OPENROWSET(                                               2
21      'CosmosDB',
22      N'account=asacosmosdbinaday84;database=CustomerProfile;key=1kxCTXbqWQ8wf0ojuzVQjCbFuAsoV6rlMR7KyD
23      UserProfileHTAP
24  )
25  WITH (
26      userId bigint,
27      cartId varchar(50),                                        3
28      preferredProducts varchar(max),
29      productReviews varchar(max)
30  ) AS profiles
31  CROSS APPLY OPENJSON (productReviews)
32  WITH (
33      productId bigint,                                          4
34      reviewText varchar(1000)
35  ) AS reviews
36  GO
```

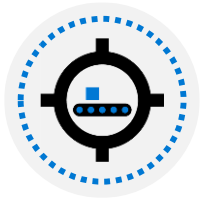
Results | Messages

11:46:22 AM Started executing query at Line 1 — 5

(Changed database context to 'Profiles'. Changed database context to 'master'.)
Total execution time: 00:00:04.304

00:00:04 Query executed successfully.

Review questions

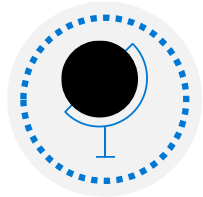


Q01 – Where do you enable Azure Synapse Link for Azure Cosmos DB?

A01 – In Azure Cosmos DB

Q02 – How can you manage the lifecycle of data and define how long it will be retained for in an analytical store?

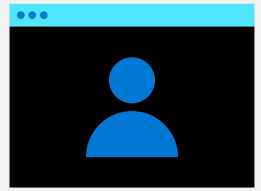
A02 – Configure the default Time to Live (TTL) property for records stored.



Q03 – What is the name of the application architecture that enables near real-time querying to provide insights?

A03 – HTAP

Lab: Support Hybrid Transactional Analytical Processing (HTAP) with Azure Synapse Link



http://fmdk.io/dp2039

Lab overview

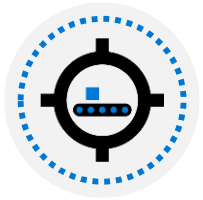
This lab teaches you how Azure Synapse Link enables seamless connectivity of an Azure Cosmos DB account to an Azure Synapse workspace. You will understand how to enable and configure Synapse link, then how to query the Azure Cosmos DB analytical store using Apache Spark and SQL Serverless.

Lab objectives

After completing this lab, you will be able to:

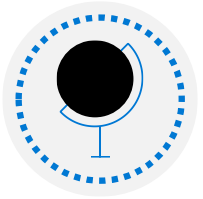
- Configure Azure Synapse Link with Azure Cosmos DB
- Query Azure Cosmos DB with Apache Spark for Synapse Analytics
- Query Azure Cosmos DB with serverless SQL pool for Azure Synapse Analytics

Lab review



Q01 – What SparkSQL method reads data from the analytical store?

Q02 – What function provides a rowset view over a JSON document?



Q03 – Which PySpark function helps flatten columns for better readability and ease of querying?

Module summary

In this module, you have learned about:

- HTAP
- SQL Serverless
- Flattening JSON Files
- Azure Synapse Link
- Azure Cosmos DB

Next steps

After the course, consider visiting the website that explores a [HTAP with Cosmos DB and Synapse Link] pattern and the associated documentation that goes into more depth about this architecture.

Microsoft Azure

Course feedback (100% anonymous)
http://fmdk.io/dp203feedback