Building an Azure Business Intelligence Solution End to End

Hands On Workshop

Paul Andrew | Senior Consultant

Terry McCann | Principal Consultant

Simon Whiteley | Cloud Architect

https://github.com/Adatis

ModernDataWarehouseWorkshop

adatis

Agenda for the Day

Module 1 Microsoft Azure	Module 2 Storage Uploading Data Data Lake	Module 3 Real-time Data Streaming Power BI	Module 4 U-SQL - Data Transformation Basics
Module 5 USQL - Advanced Analytics Cognitive Services	Module 6 Data Factory Orchestration Dynamic Pipelines	Module 7 Data Presentation & Consumption Power BI Models	Module 8 Other Services Q&A

Session Agenda

Real-time Data

Lambda
History
Problem/Solution

Endpoints

IoT Hub
Event Hub

Streaming

Stream Analytics
Query Windows

Session Agenda

Real-time Data

Lambda
History
Problem/Solution

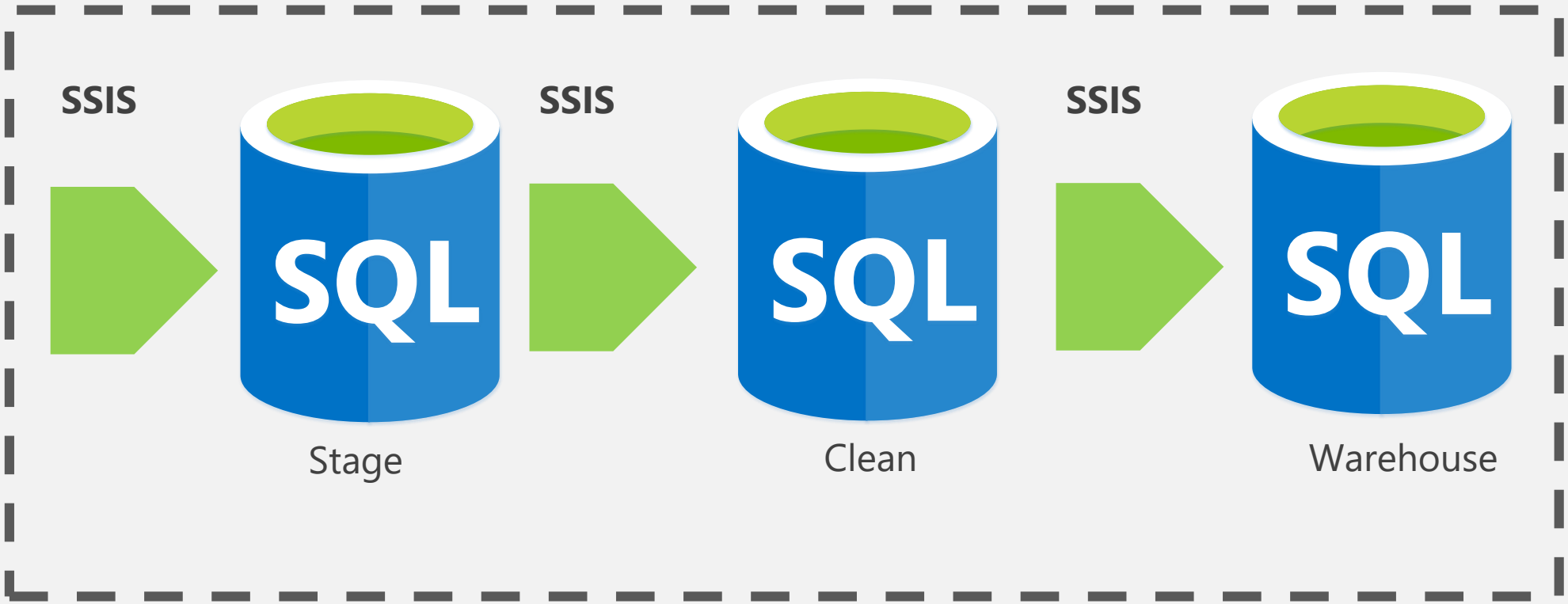
Endpoints

IoT Hub
Event Hub

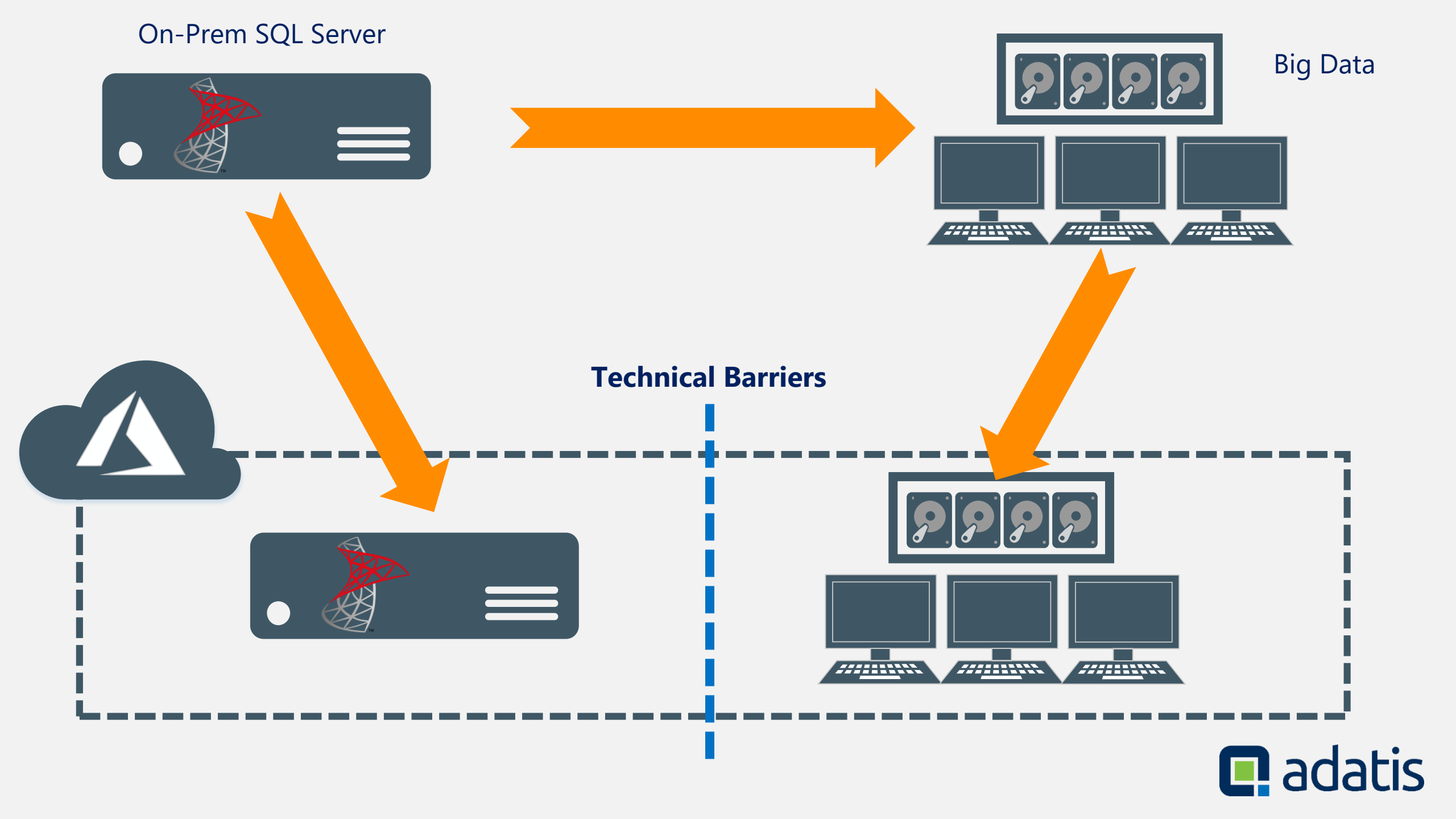
Streaming

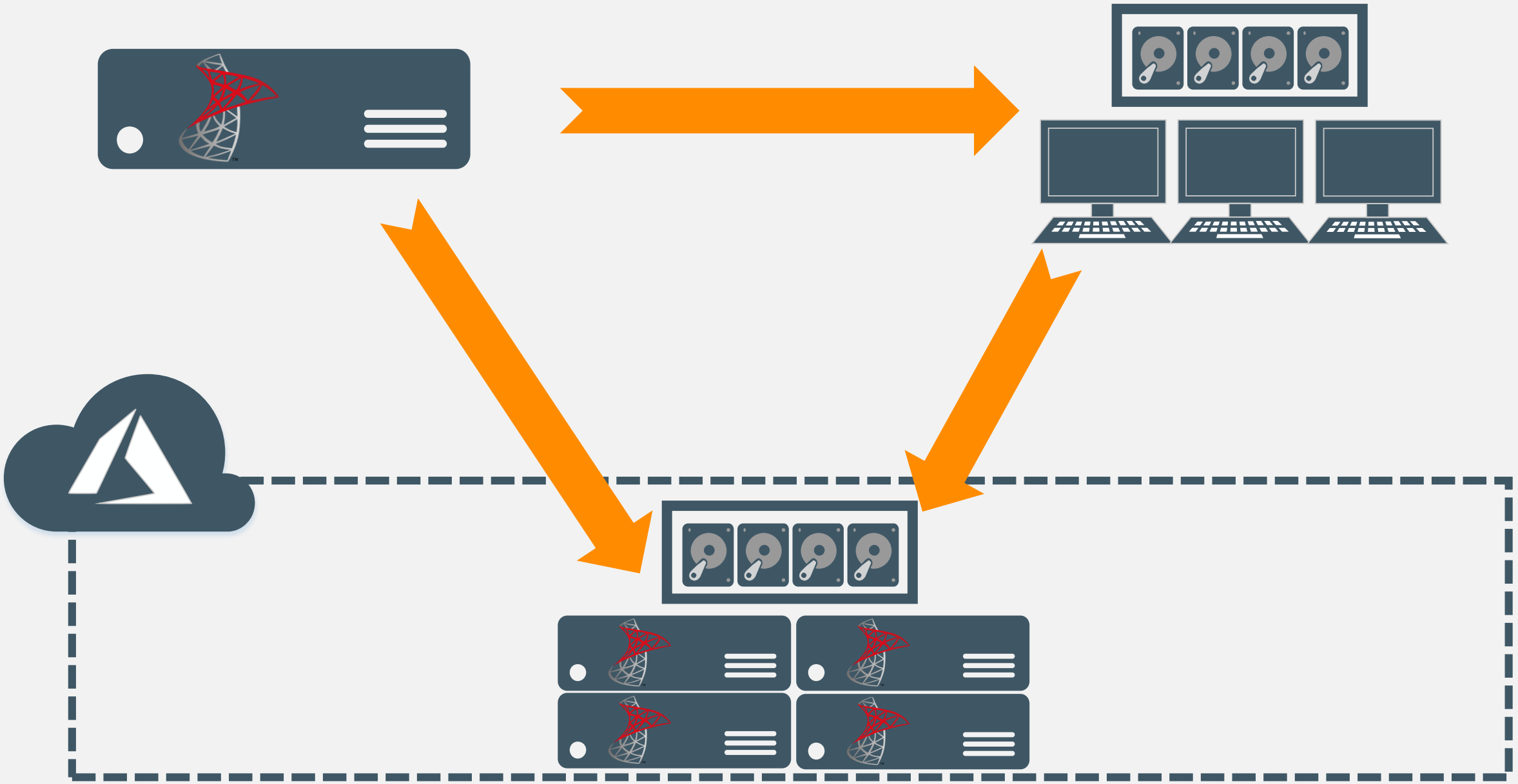
Stream Analytics
Query Windows

One-Box SQL BI Architecture



adatis


On-Prem SQL Server
Big Data
Technical Barriers
adatis



Modern Data Warehouse

adatis

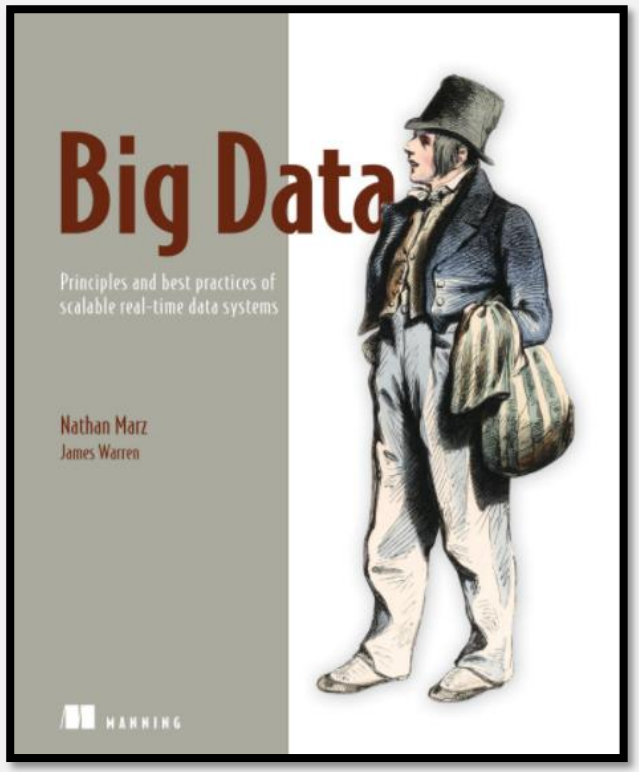
My Life Goal:

Never to manage another Server

adatis

Lambda Architecture

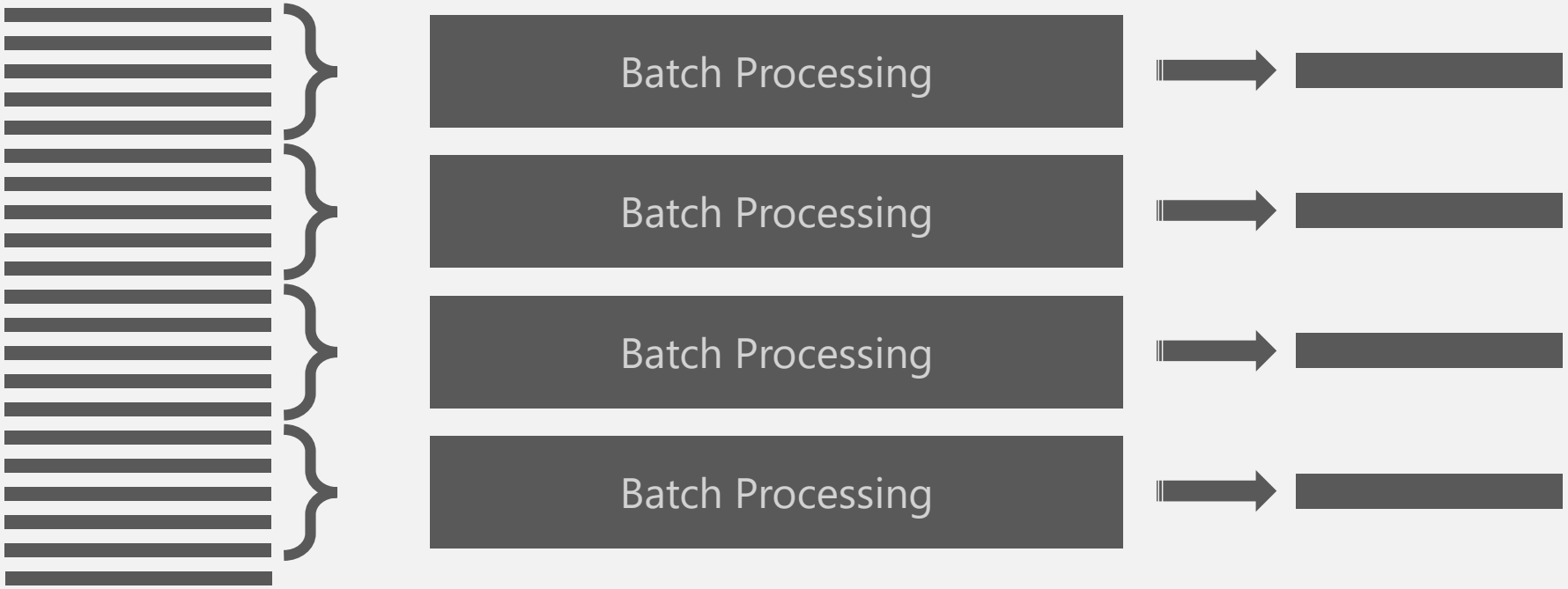
Use Batch and Stream technologies together to balance latency, throughput and fault-tolerance



Nathan Marz &
James Warren

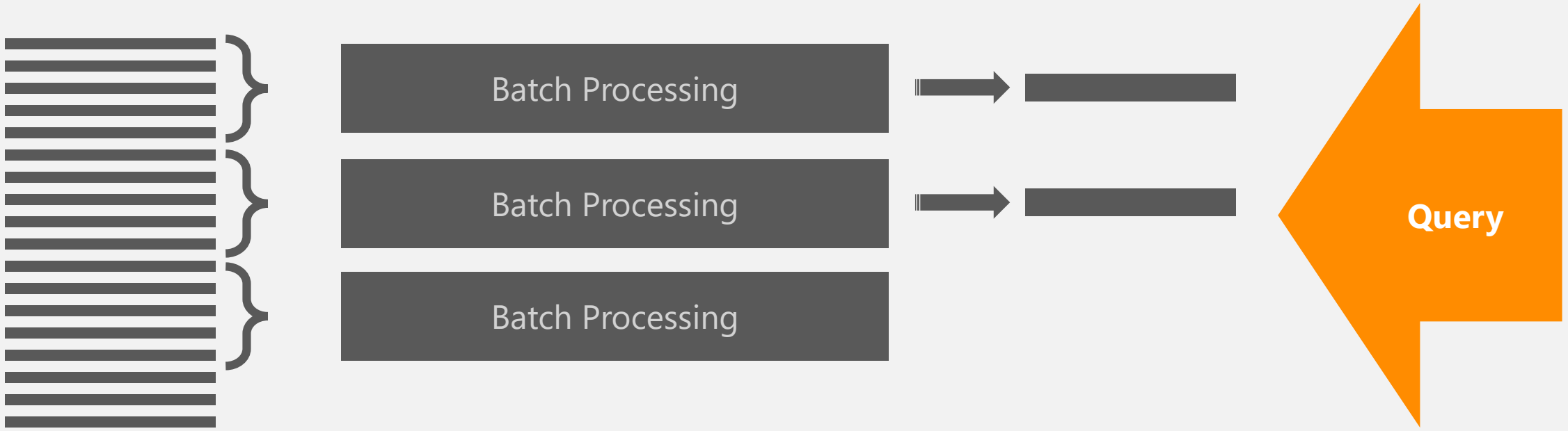
adatis

The Problem…

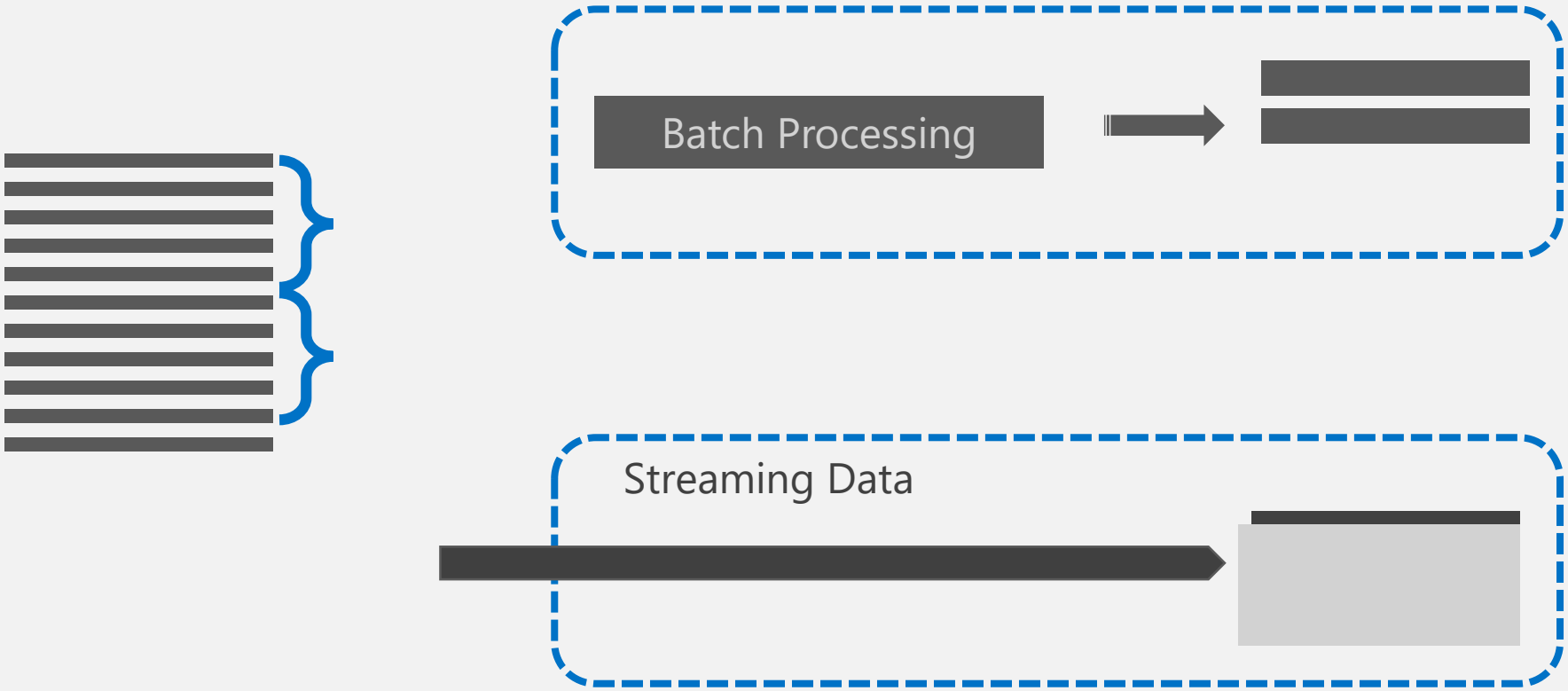


adatis

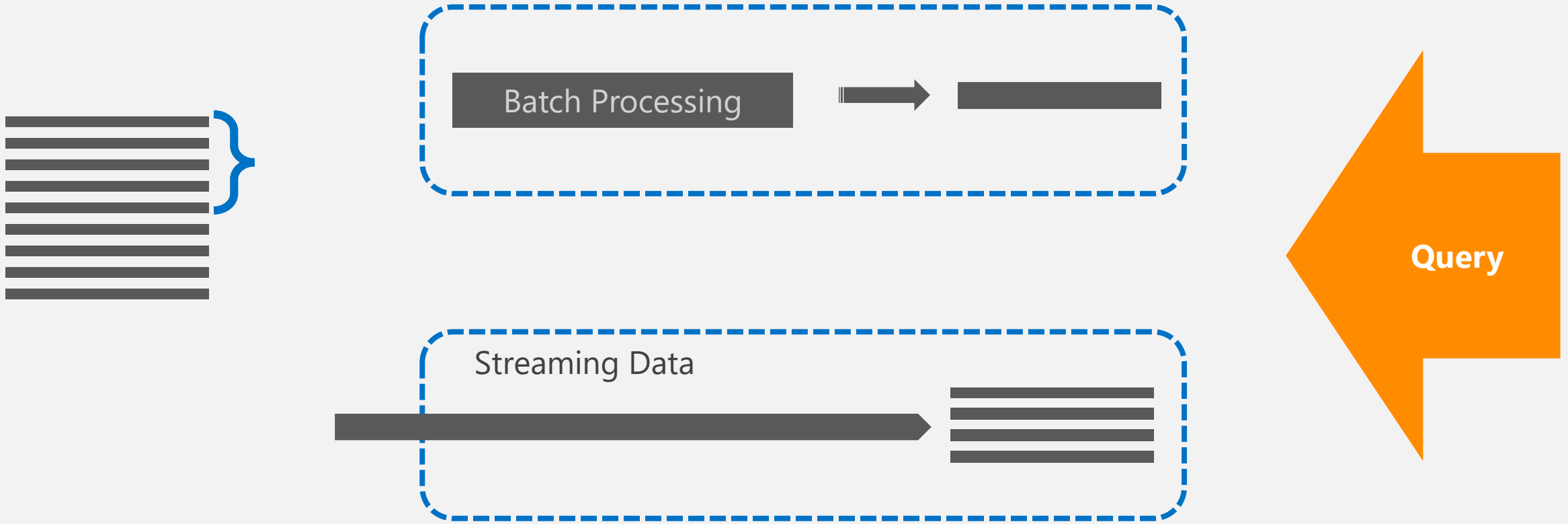
The Problem…



adatis

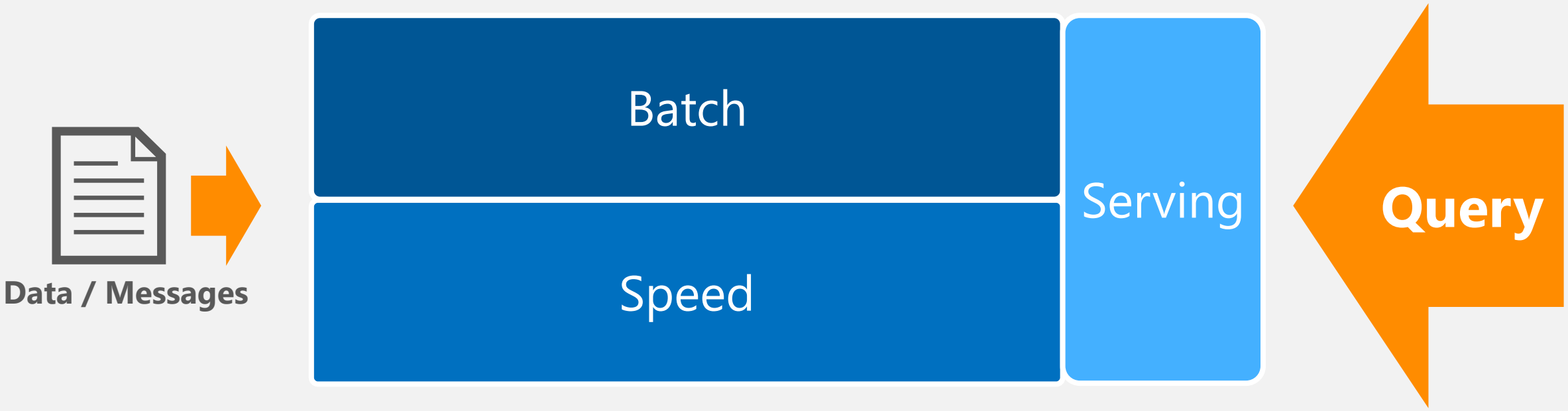


adatis



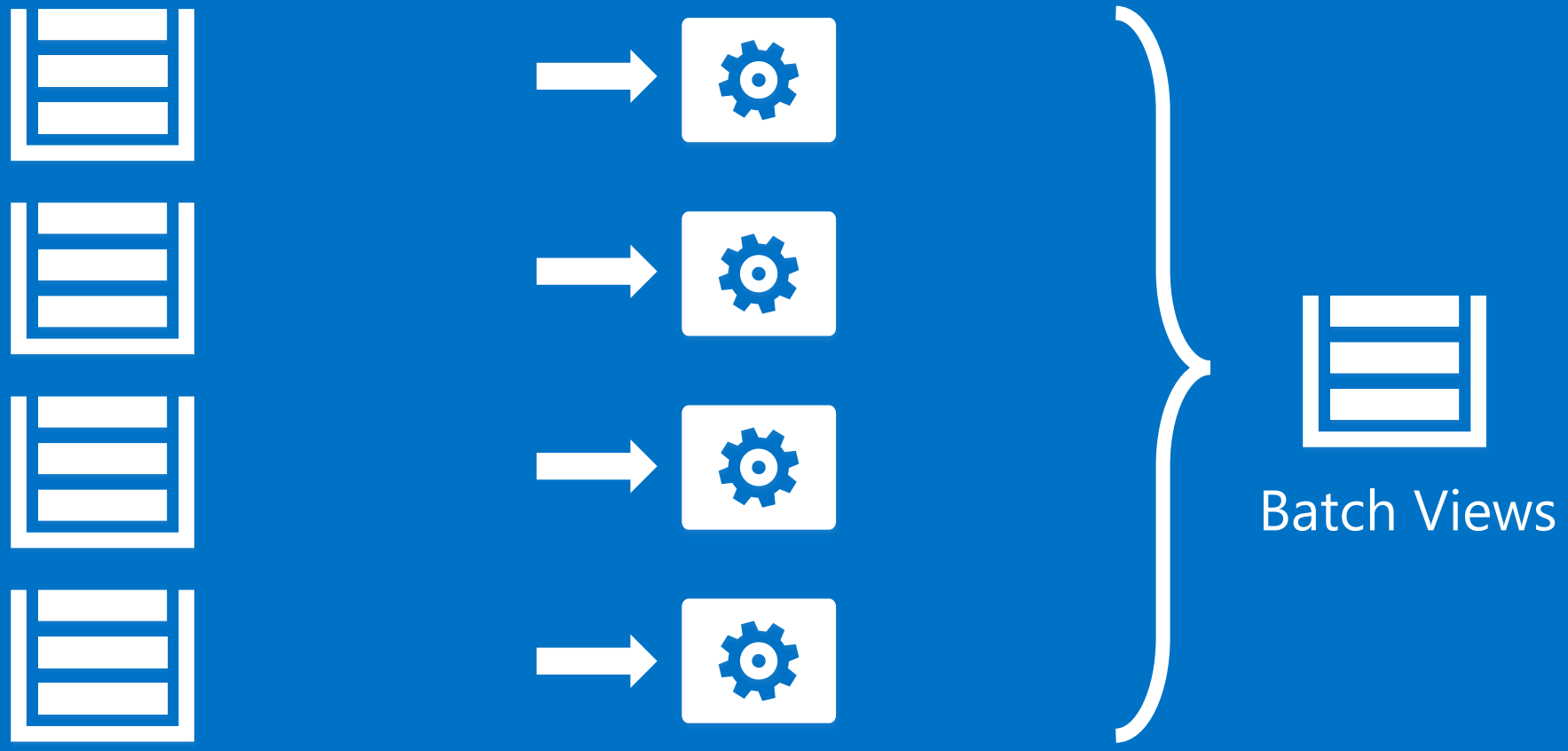
adatis

The Lambda Architecture

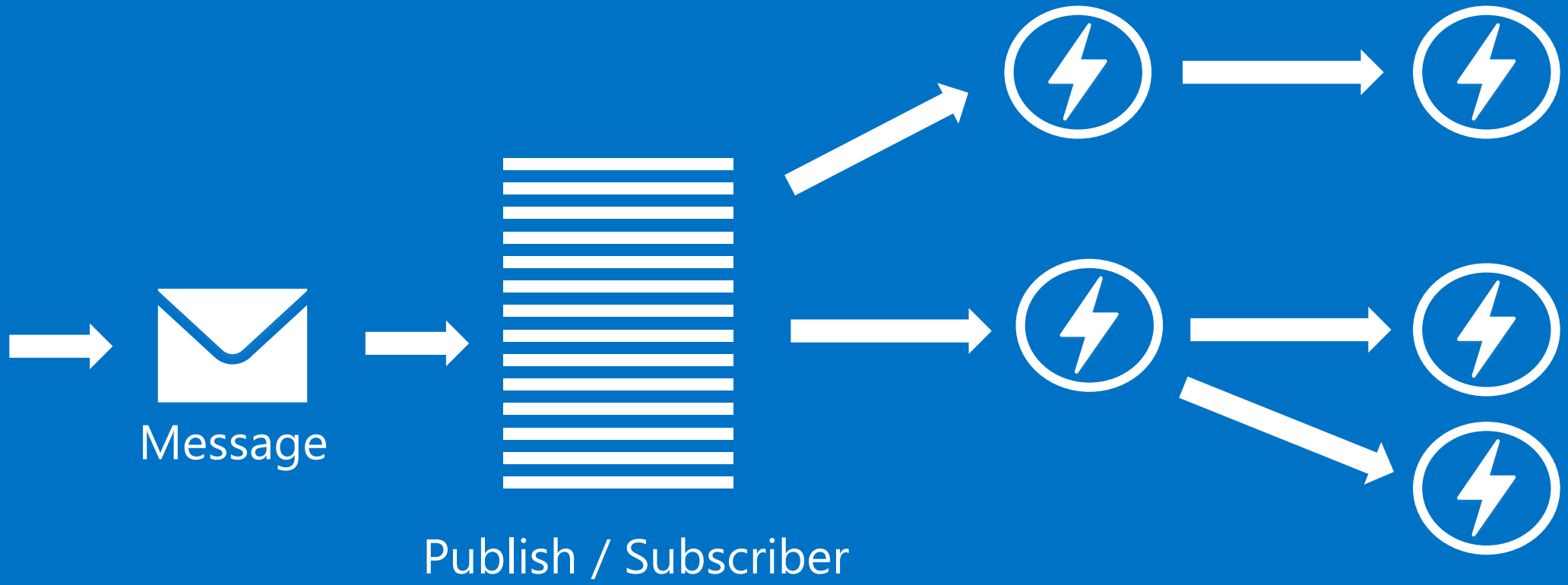


adatis

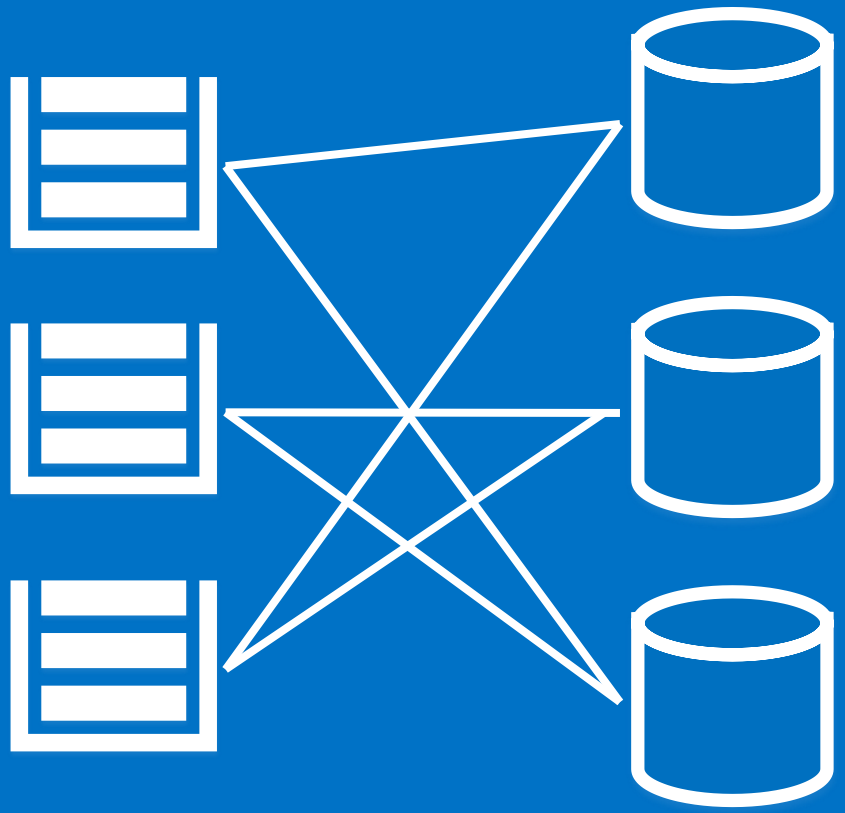
Batch Layer


Batch Views
Distributed Storage
(HDFS)
Parallel Processing
(MapReduce)

Speed Layer



Serving Layer



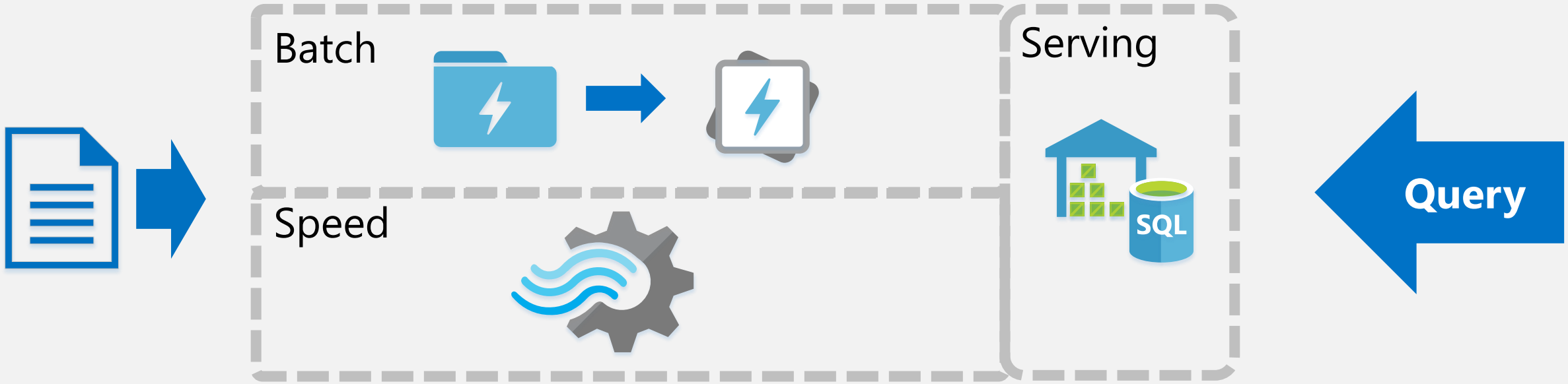
Batch Views

Sharded
Database Layer

The Marz Lambda Architecture

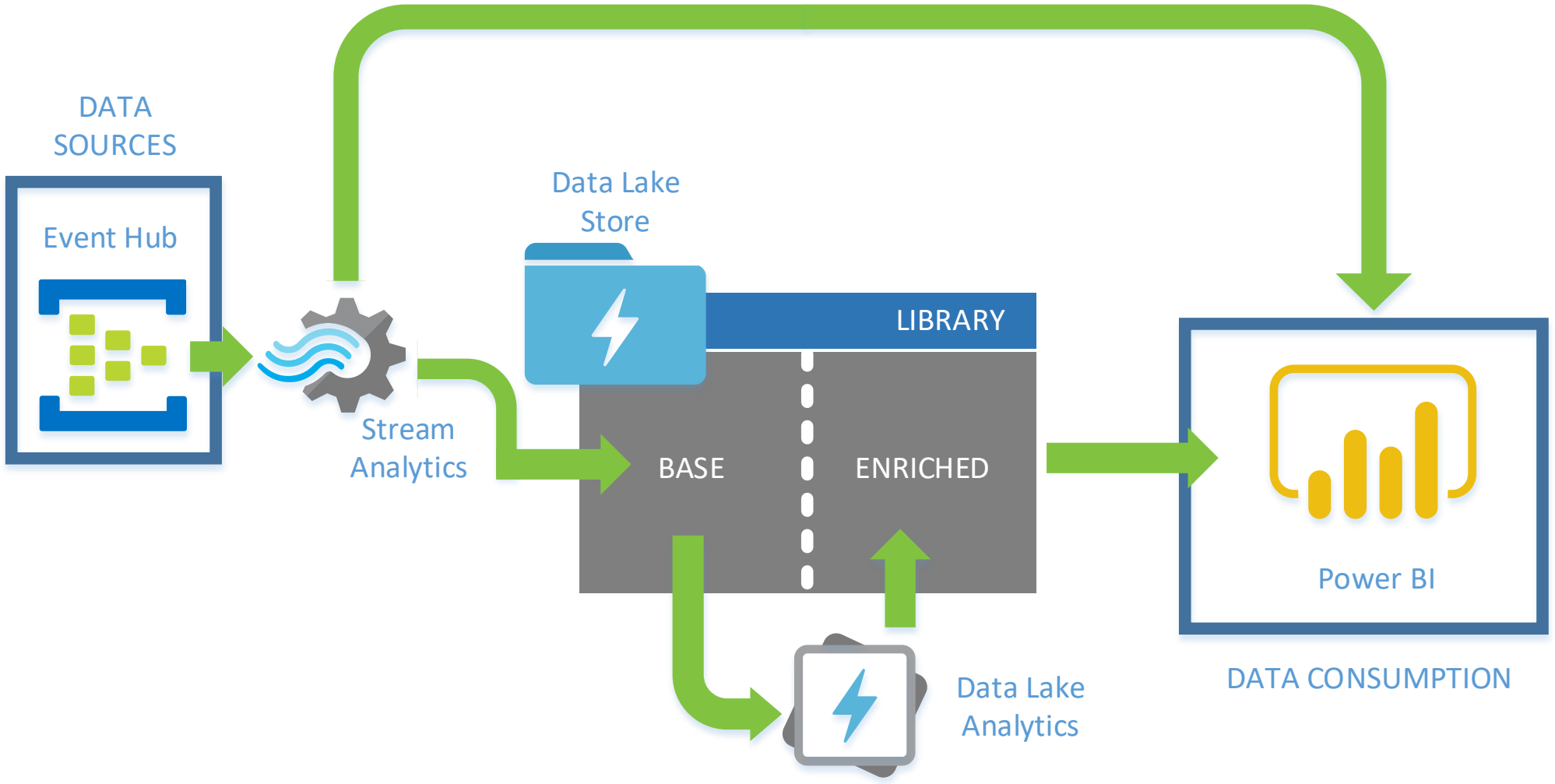
adatis

Applying Lambda to Azure



adatis

Applying Lambda to our Architecture



Session Agenda

Real-time Data

Lambda
History
Problem/Solution

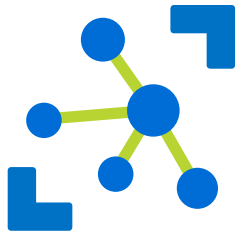
Endpoints

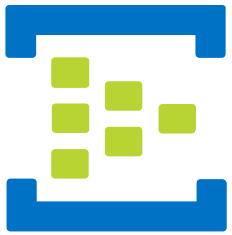
IoT Hub
Event Hub

Streaming

Stream Analytics
Query Windows

IoT Hub vs Event Hub





Feature	IoT Hub	Event Hub
Message Direction	2 Way	1 Way
Protocol Support	MQTT, AMQP, HTTP	AMQP, HTTP
Scaling	Configured	Automatic
Message Routing	Yes	No
Security	Device Level	Hub Level
Device State Support	Yes	No
Message Capturing	No	Yes
Multiple Hubs	No	Yes
Tiers	F1/S1/S2/S3	Basic/Standard

https://docs.microsoft.com/en-us/azure/iot-hub/iot-hub-compare-event-hubs

Session Agenda

Real-time Data

Lambda
History
Problem/Solution

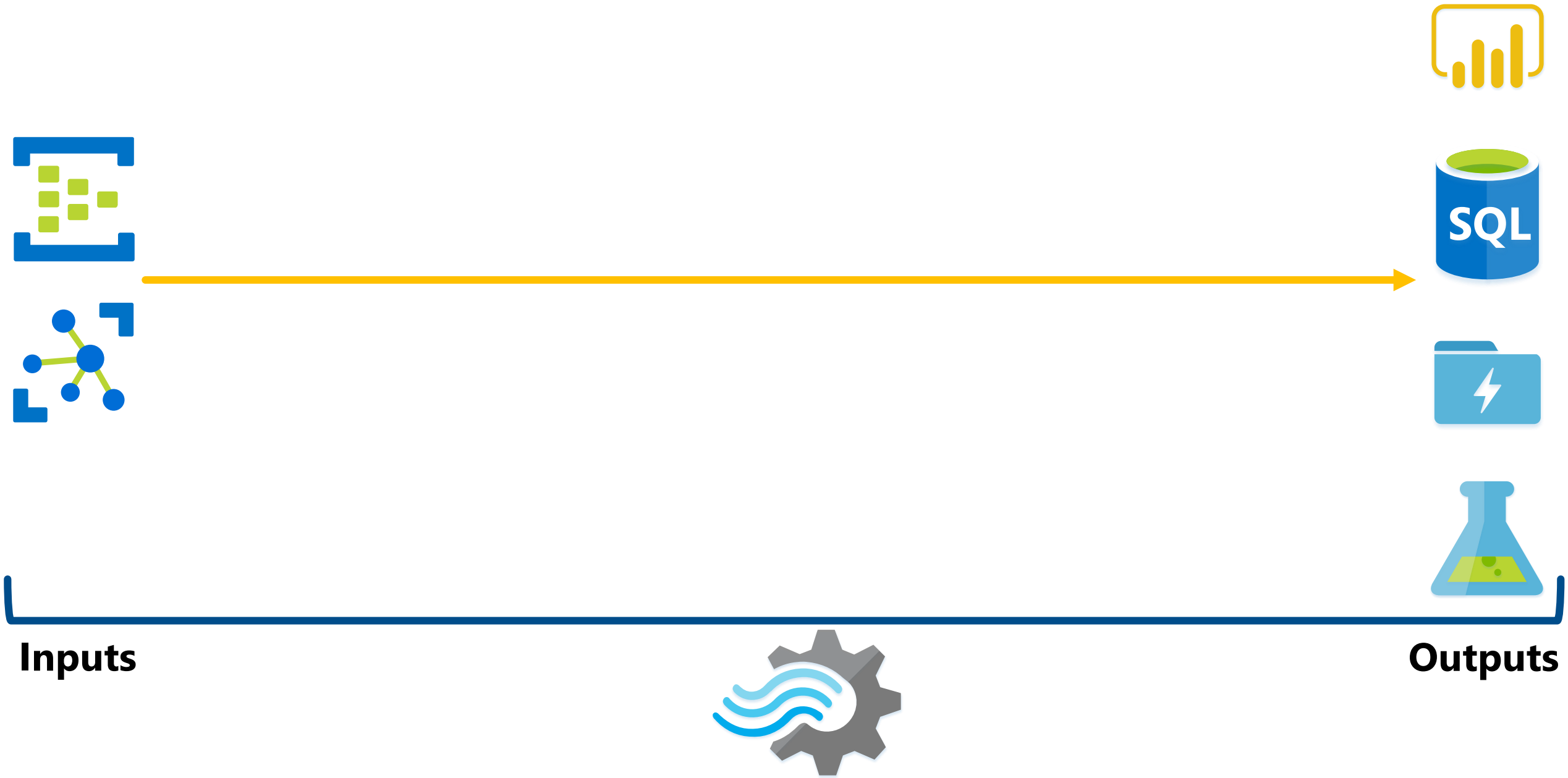
Endpoints

IoT Hub
Event Hub
Service Bus

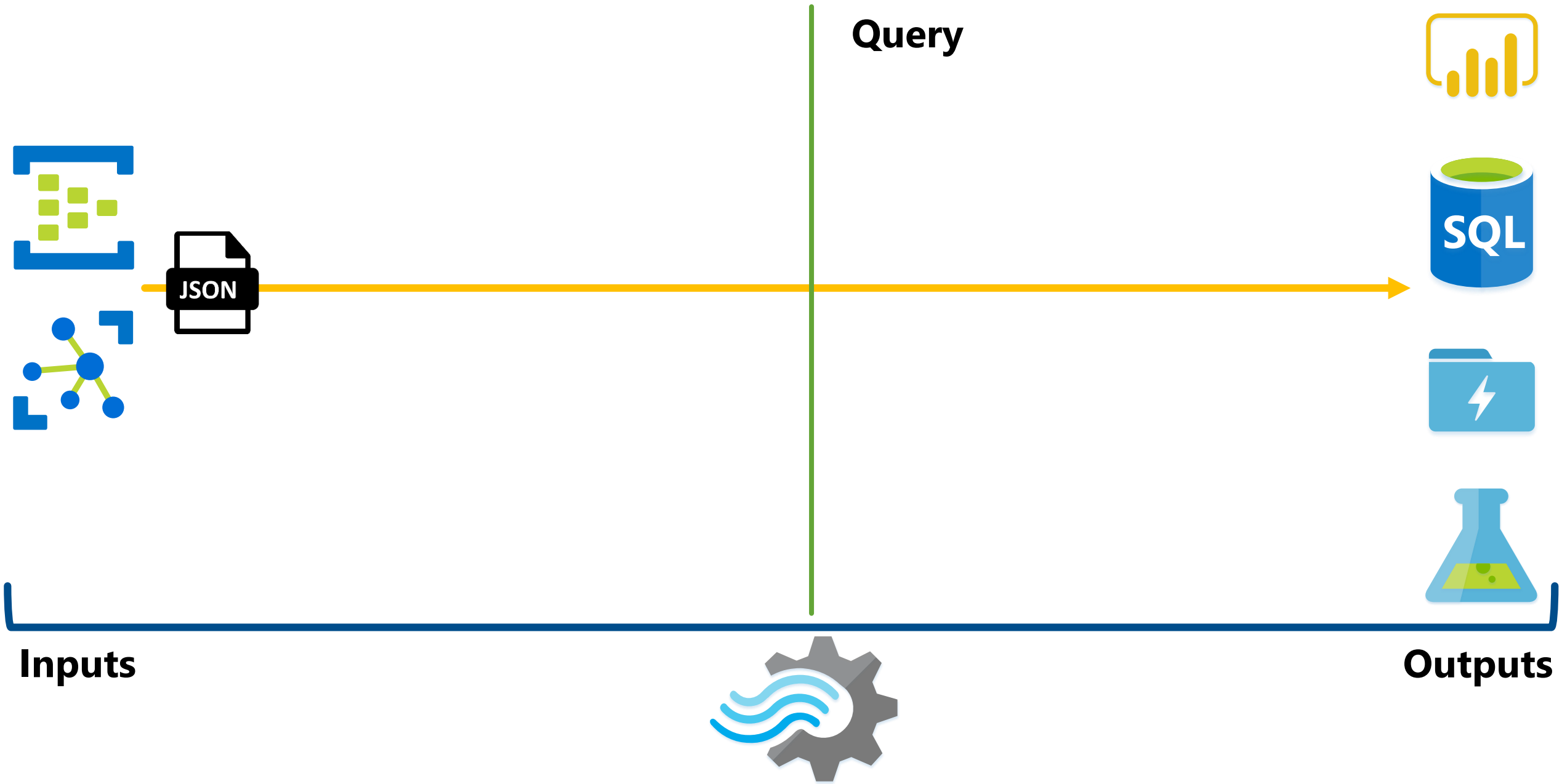
Streaming

Stream Analytics
Query Windows

Stream Analytics



Stream Analytics



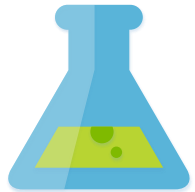
Stream Analytics

```
SELECT


    SUM(CAST(eh.UnitPrice AS float)) AS UnitPrice,
    SUM(CAST(eh.LineTotal AS float)) AS LineTotal,
    SUM(CAST(eh.OrderQty AS float)) AS OrderQty,
    COUNT(*) AS RecordCount
INTO
    [powerbi]
FROM
    [eventhub] AS eh


GROUP BY
    eh.EventEnqueuedUtcTime,

    SlidingWindow(second, 30)
```



Inputs

Outputs

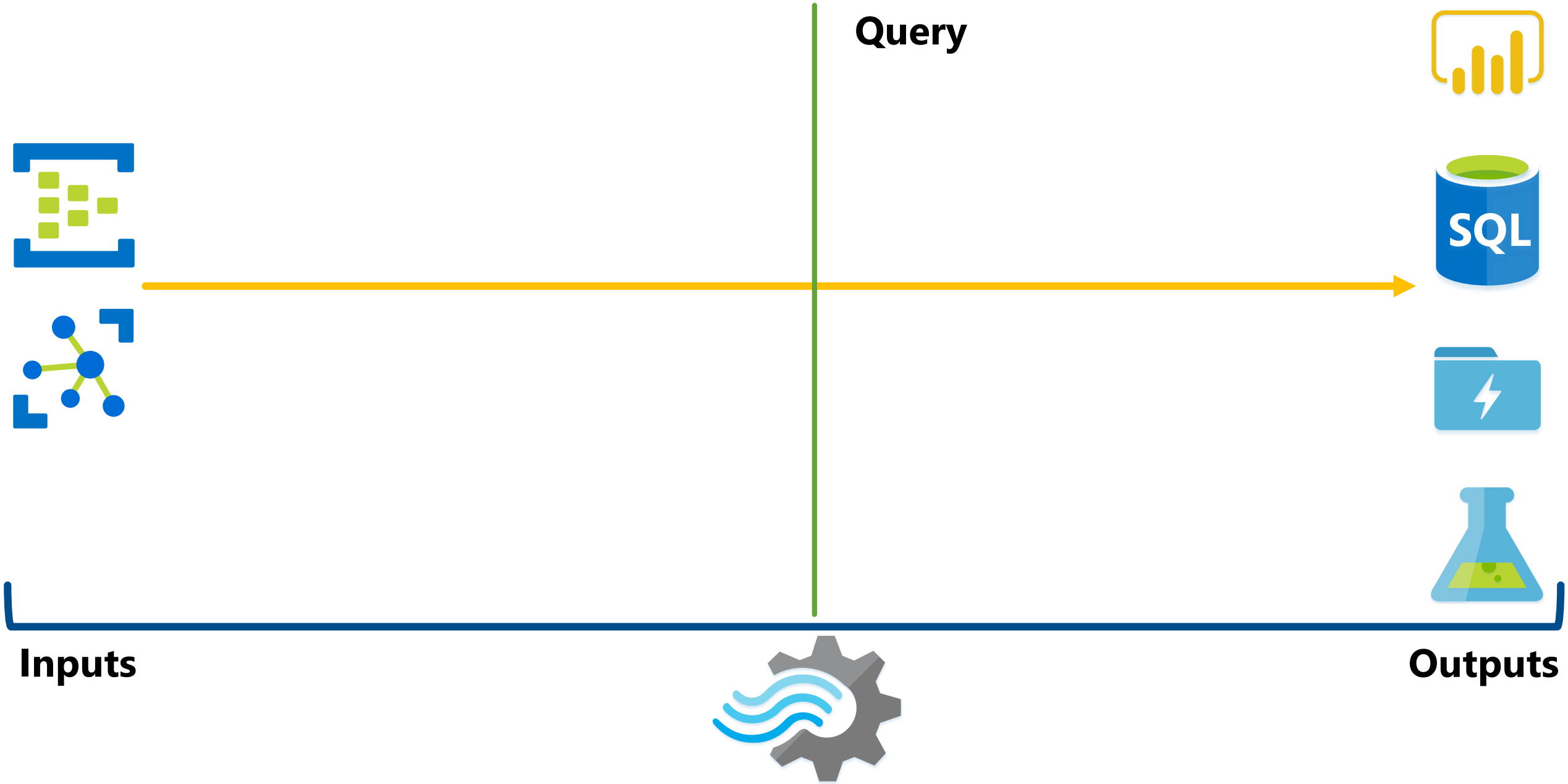
Stream Analytics


```
SELECT
    eh.EventEnqueuedUtcTime,
    prd.Name AS ProductName,
    SUM(CAST(eh.UnitPrice AS float)) AS UnitPrice,
    SUM(CAST(eh.LineTotal AS float)) AS LineTotal,
    SUM(CAST(eh.OrderQty AS float)) AS OrderQty,
    COUNT(*) AS RecordCount
INTO
    [powerbi]
FROM
    [eventhub] AS eh
    INNER JOIN [Products] AS prd
        ON eh.[ProductId] = prd.[ProductId]
GROUP BY
    eh.EventEnqueuedUtcTime,
    prd.Name,
    SlidingWindow(second, 30)
```

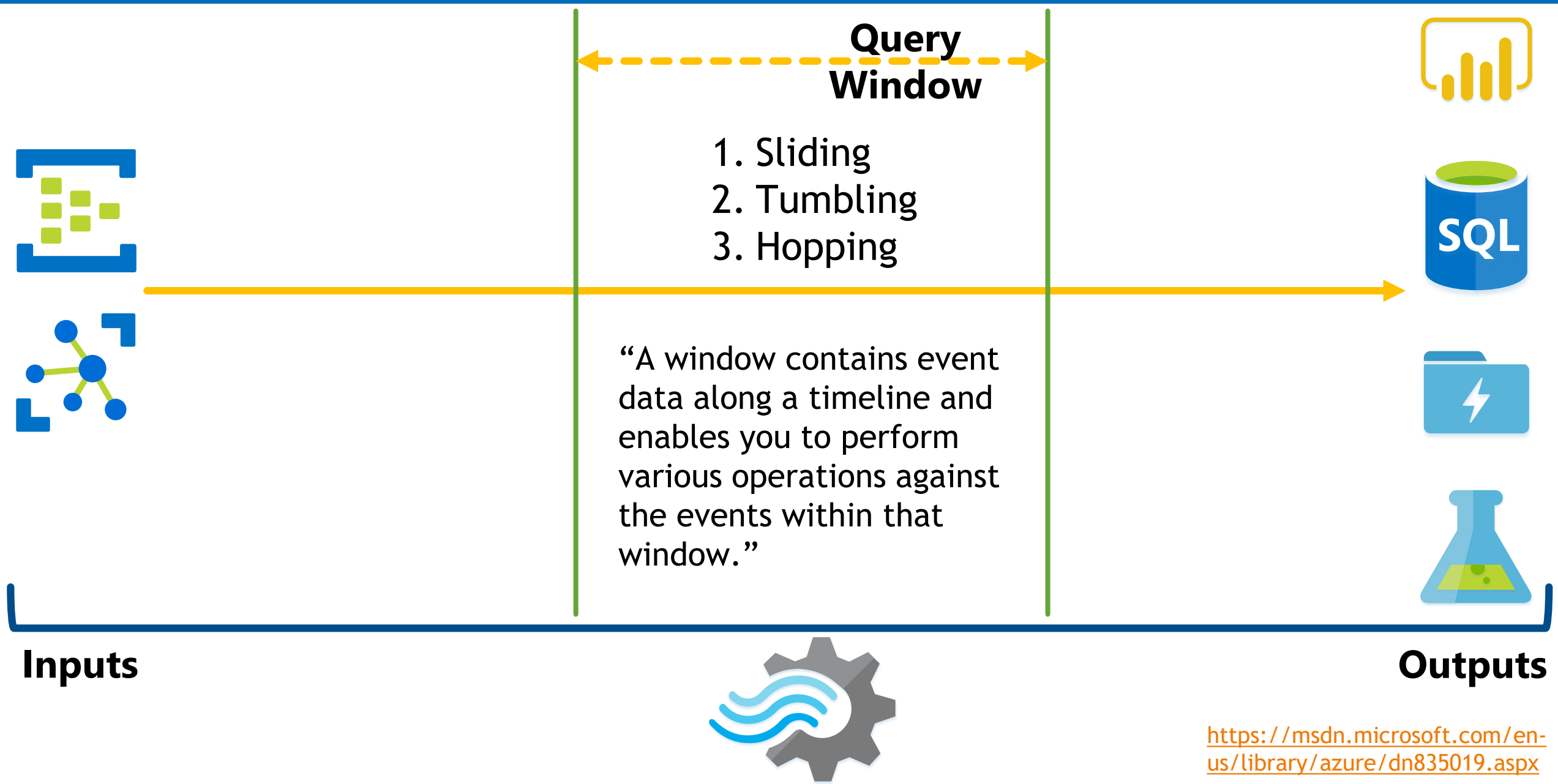
Inputs

Outputs

Stream Analytics



Stream Analytics

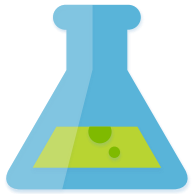


https://msdn.microsoft.com/en-us/library/azure/dn835019.aspx

Stream Analytics

Sliding Window

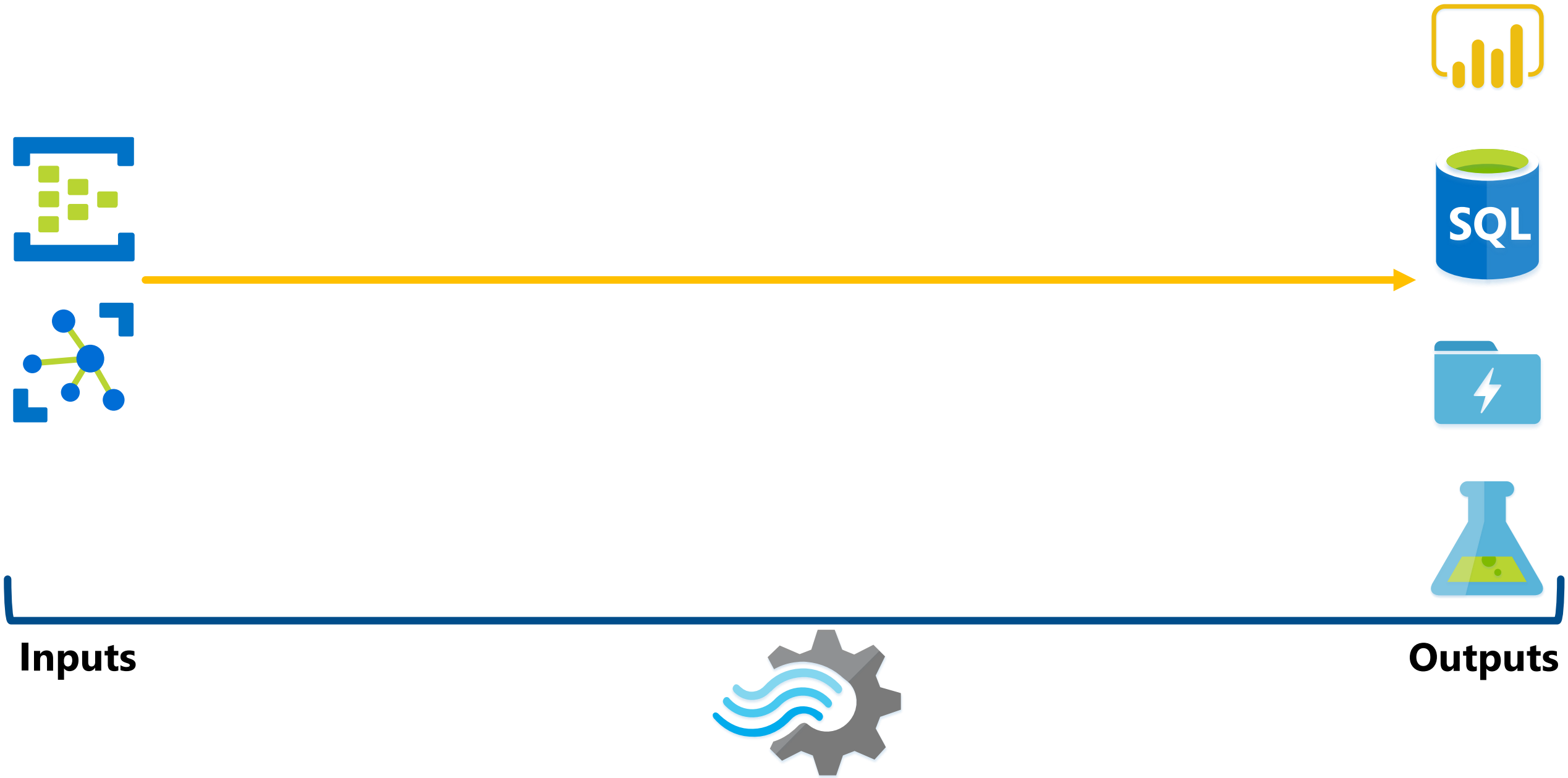
```
SELECT
    MIN([value]) AS 'Min',
    MAX([value]) AS 'Max'
INTO
    SQLDBAvg
FROM
    IoTHub
TIMESTAMP BY
    timecreated
GROUP BY
    SlidingWindow(second, 5))
```



Inputs

Outputs

Stream Analytics



Stream Analytics

Tumbling Window

Tumbling Window

Tumbling Window

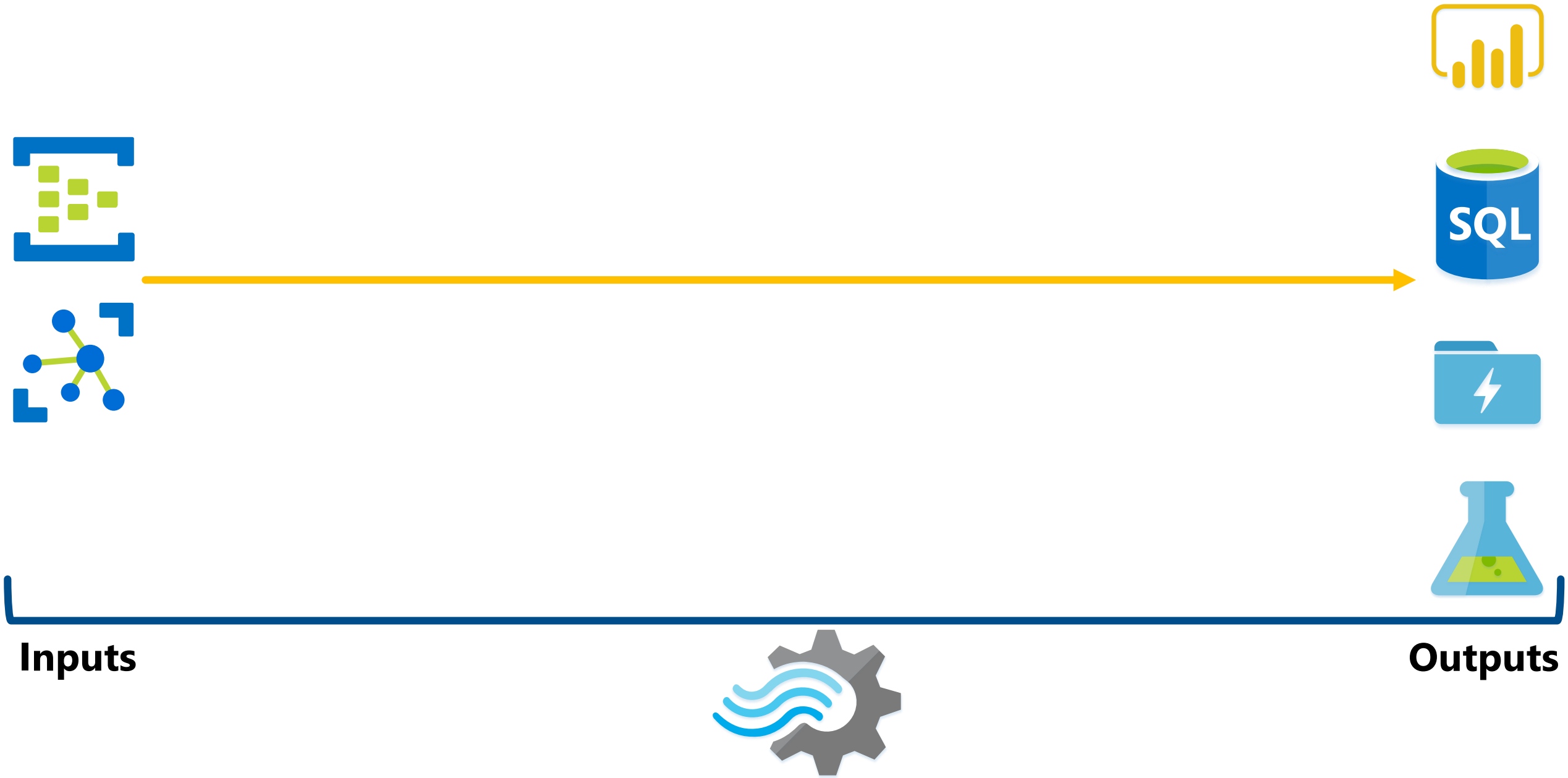
Tumbling Window

```
SELECT
    MIN([value]) AS 'Min',
    MAX([value]) AS 'Max'
INTO
    SQLDBAvg
FROM
    IoTHub
TIMESTAMP BY
    timecreated
GROUP BY
    TumblingWindow(Duration(MINUTE, 15), Offset(millisecond, -1))
```

Inputs

Outputs

Stream Analytics



Stream Analytics

Hopping Window

Hopping Window

Hopping Window

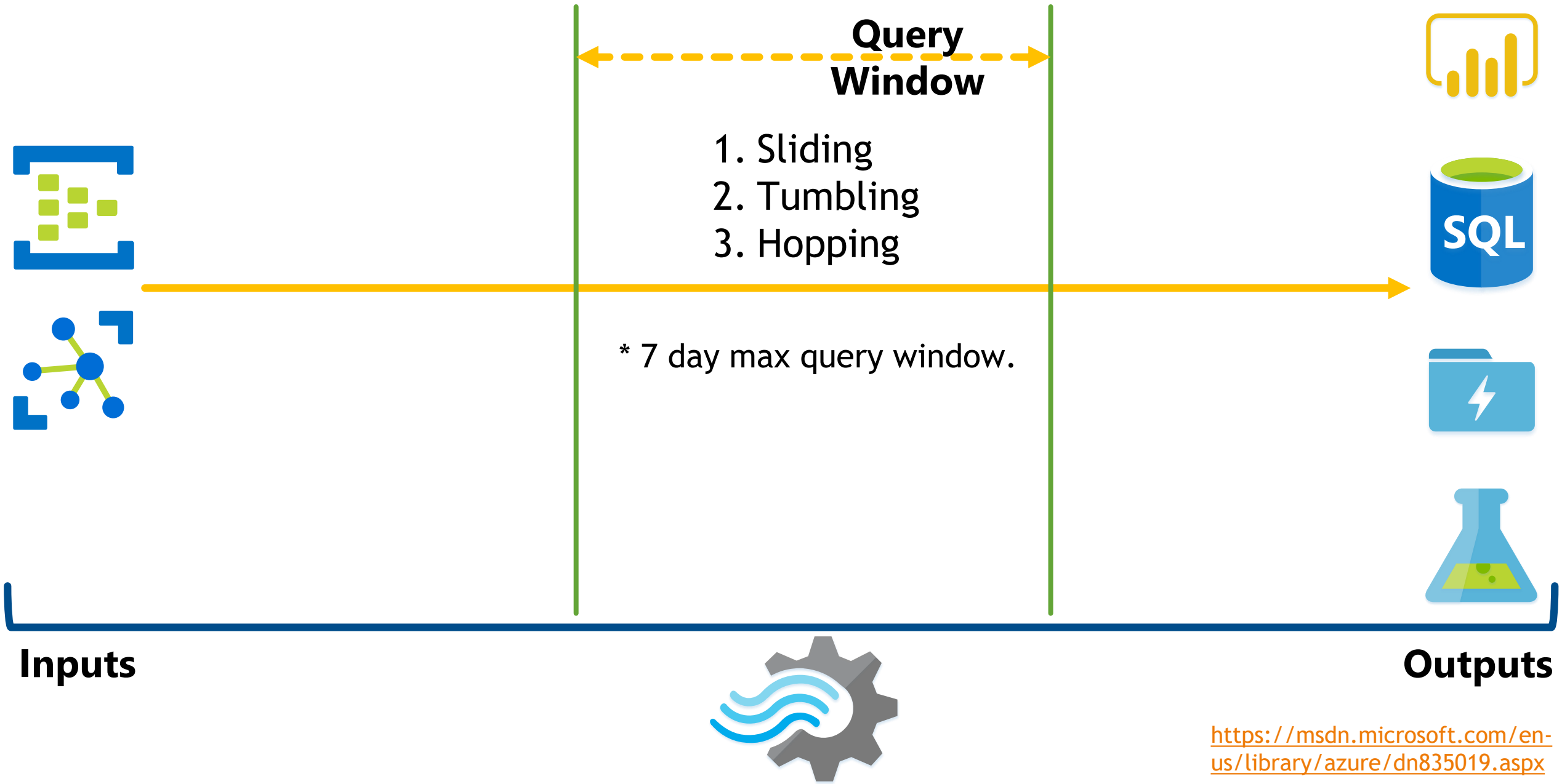
Hopping Window

```
SELECT
    MIN([value]) AS 'Min',
    MAX([value]) AS 'Max'
INTO
    SQLDBAvg
FROM
    IoTHub
TIMESTAMP BY
    timecreated
GROUP BY
    HoppingWindow(Duration(MINUTE, 15), Hop(MINUTE, 5), Offset(millisecond, -1))
```

Inputs

Outputs

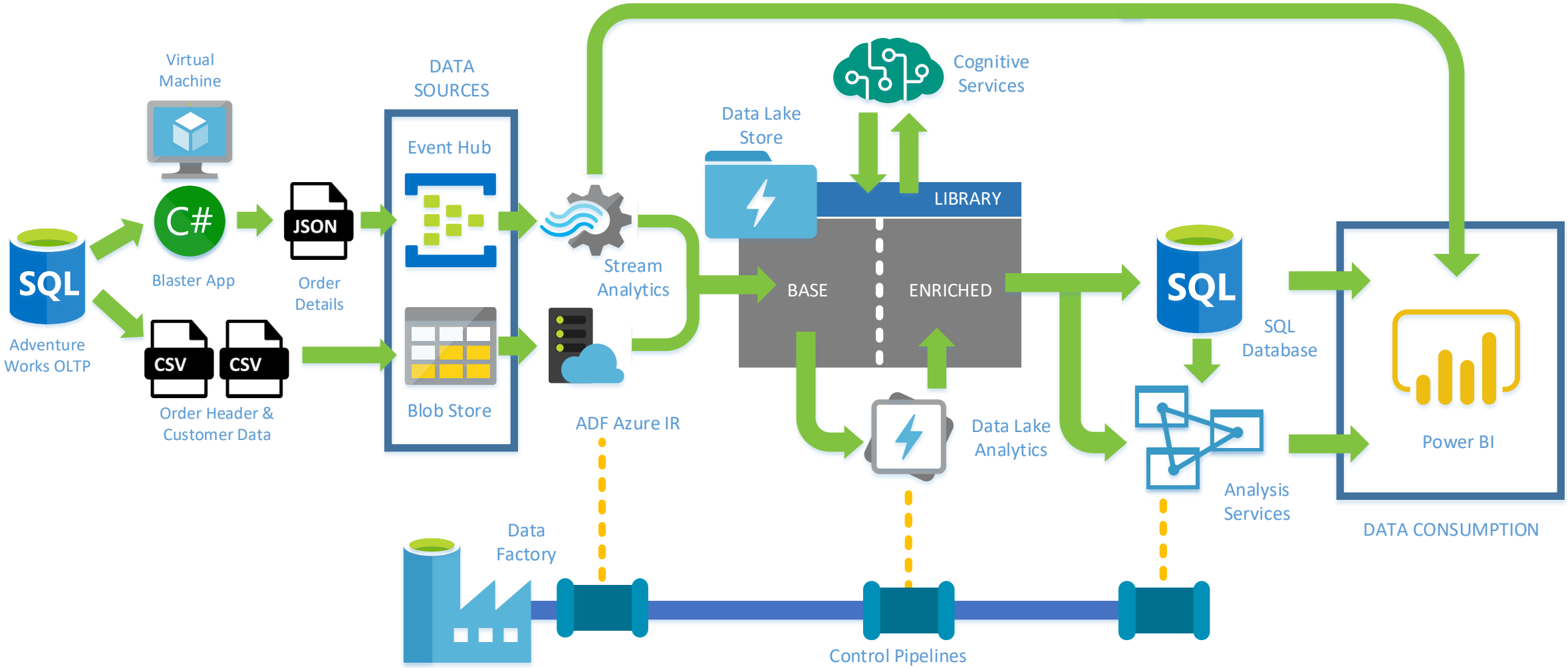
Stream Analytics



https://msdn.microsoft.com/en-us/library/azure/dn835019.aspx


adatis

Our Architecture



Agenda for the Day

Module 1 Microsoft Azure	Module 2 Storage Uploading Data Data Lake	Module 3 Real-time Data Streaming Power BI	Module 4 U-SQL - Data Transformation Basics
Module 5 USQL - Advanced Analytics Cognitive Services	Module 6 Data Factory Orchestration Dynamic Pipelines	Module 7 Data Presentation & Consumption Power BI Models	Module 8 Other Services Q&A