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

What is Azure Data Lake?

Storage & Compute

Why use Data Lake?

The Modern Data Warehouse

How can we work with Data Lake?

Development & Management

Agenda

What is Azure Data Lake?

Storage & Compute

Why use Data Lake?

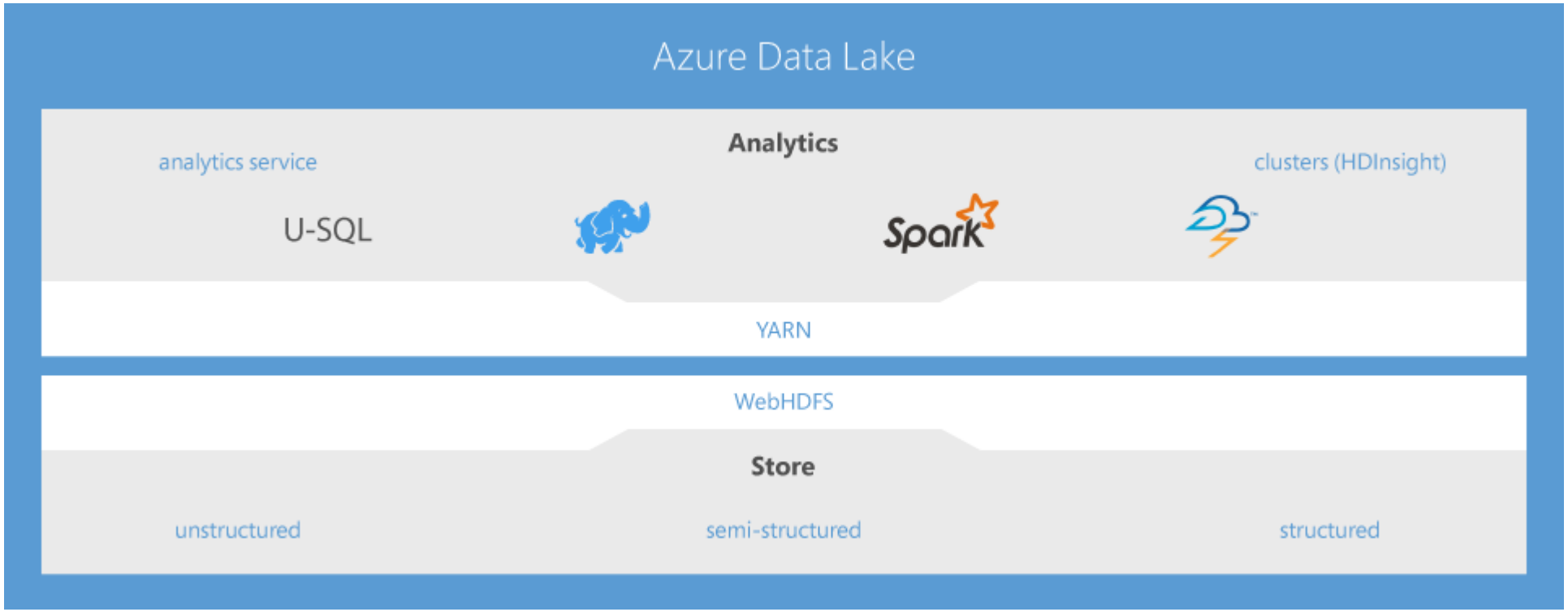
The Modern Data Warehouse

How can we work with Data Lake?

Development & Management

What is Azure Data Lake?

The Microsoft version:

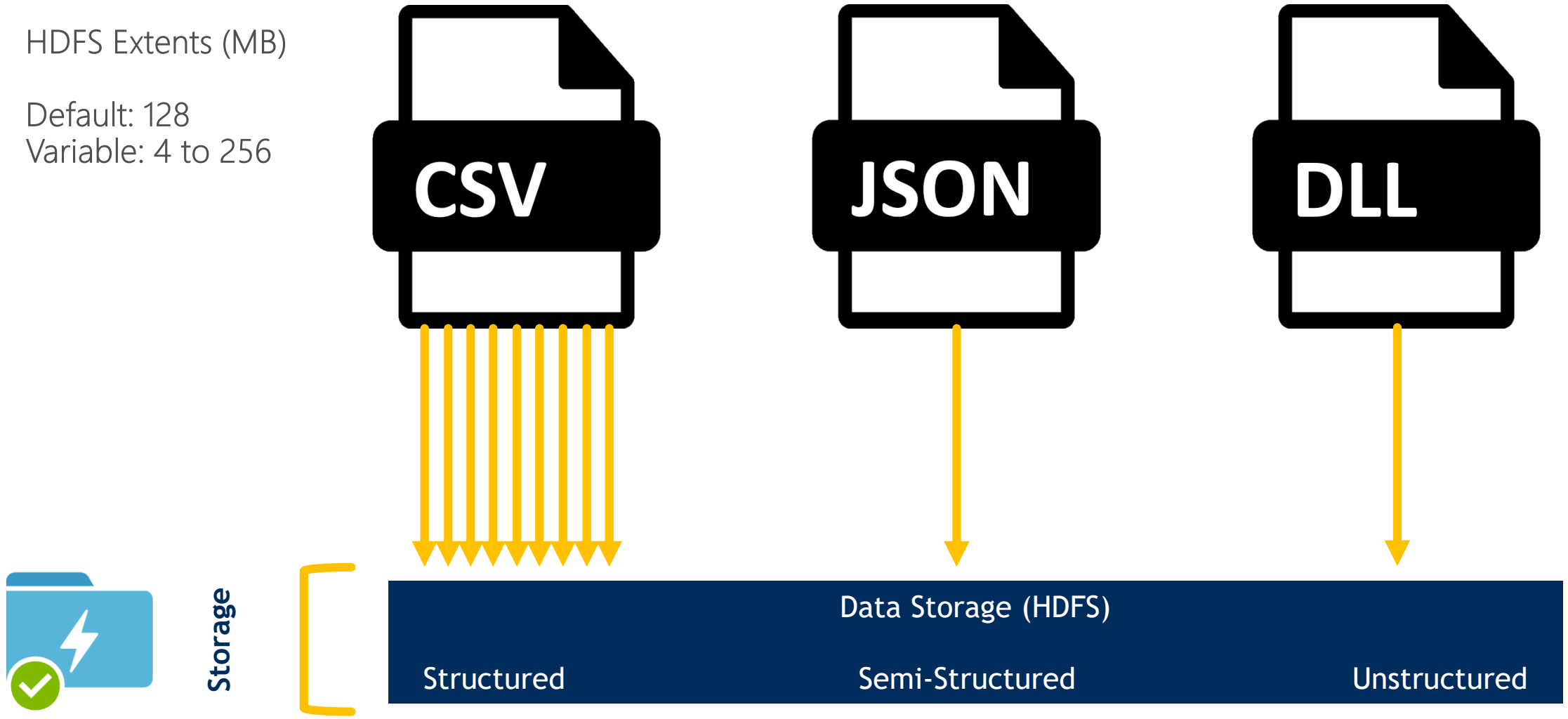


https://azure.microsoft.com/en-gb/solutions/data-lake/

What is Azure Data Lake?

HDFS Extents (MB)

Default: 128
Variable: 4 to 256



https://docs.microsoft.com/en-us/azure/data-lake-store/data-lake-store-performance-tuning-guidance

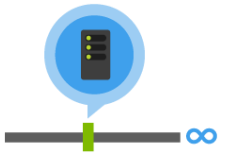
What is Azure Data Lake?

Analytics

U-SQL

Processing Engine

Node Node Node Node Node Node Node Node Node Node Node Node Node Node Node Node



Analytics Units (AU's) – Scale Out Compute

Storage

Data Storage (HDFS)

Structured Semi-Structured Unstructured

https://blogs.msdn.microsoft.com/azuredatalake/2016/10/12/understanding-adl-analytics-unit/

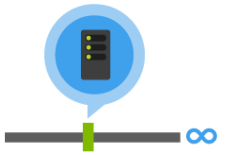
What is Azure Data Lake?

Analytics

U-SQL

Processing Engine

Node Node Node Node Node Node Node Node Node Node Node Node Node Node Node Node



Analytics Units (AU's) – Scale Out Compute

Storage

Data Storage (HDFS)

Structured Semi-Structured Unstructured

https://blogs.msdn.microsoft.com/azuredatalake/2016/10/12/understanding-adl-analytics-unit/

Azure Data Lake vs Other Data Services

Azure SQL Database (SQLDB)

T-SQL

Compute Node

Database Storage Engine

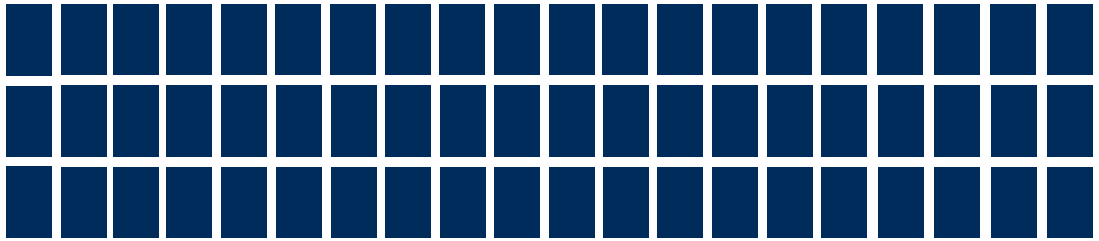
Database Transaction Units (DTU's) – Scale Up Compute

Azure SQL Data Warehouse (SQLDW)

D-SQL

Control Node

Compute Node | Compute Node | Compute Node | Compute Node



Database Warehouse Units (DWU's) – Scale Out Compute

Agenda

What is Azure Data Lake?

Storage & Compute

Why use Data Lake?

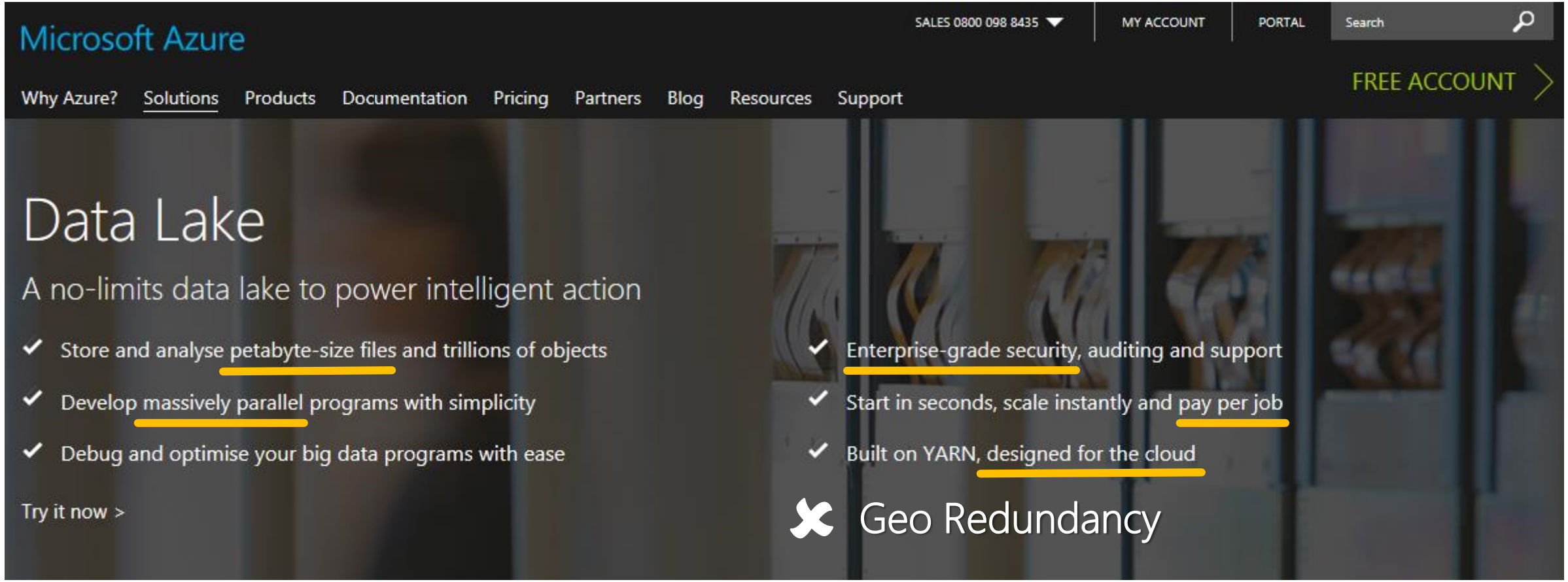
The Modern Data Warehouse

How can we work with Data Lake?

Development & Management

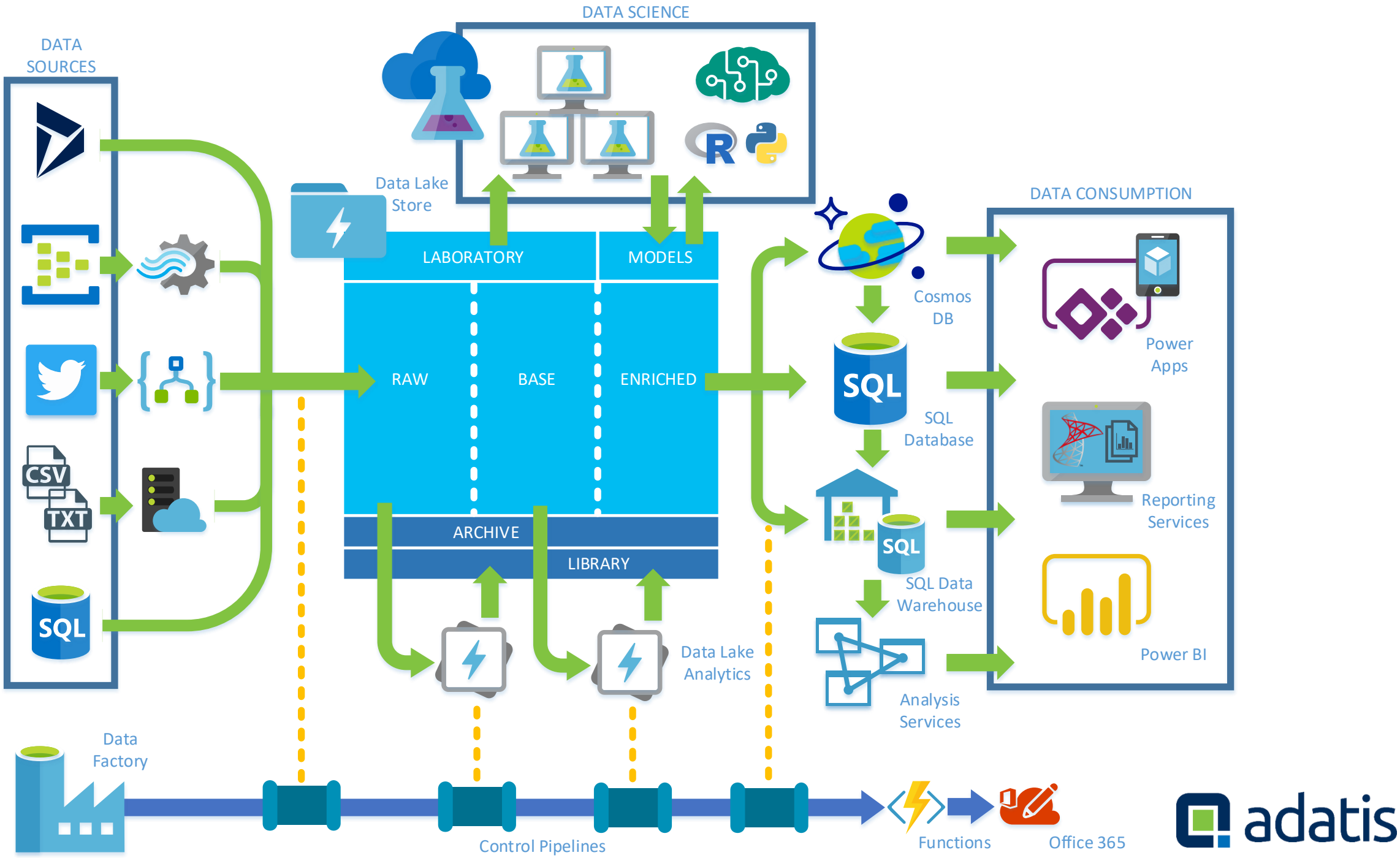
Why use Azure Data Lake?

The Microsoft version:



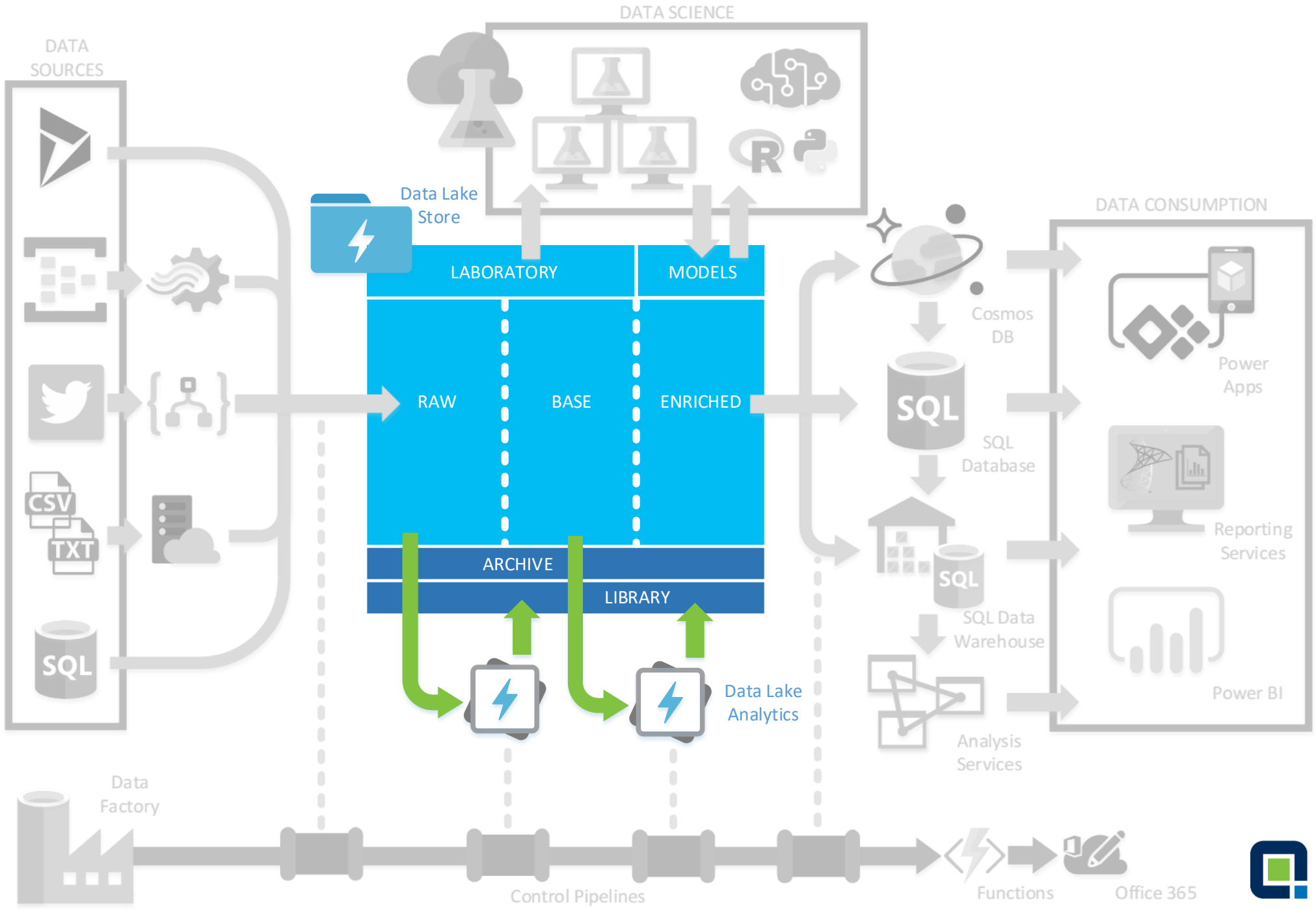
https://azure.microsoft.com/en-gb/solutions/data-lake/

The Modern Data Warehouse



adatis

The Modern Data Warehouse



adatis

Agenda

What is Azure Data Lake?

Storage & Compute

Why use Data Lake?

The Modern Data Warehouse

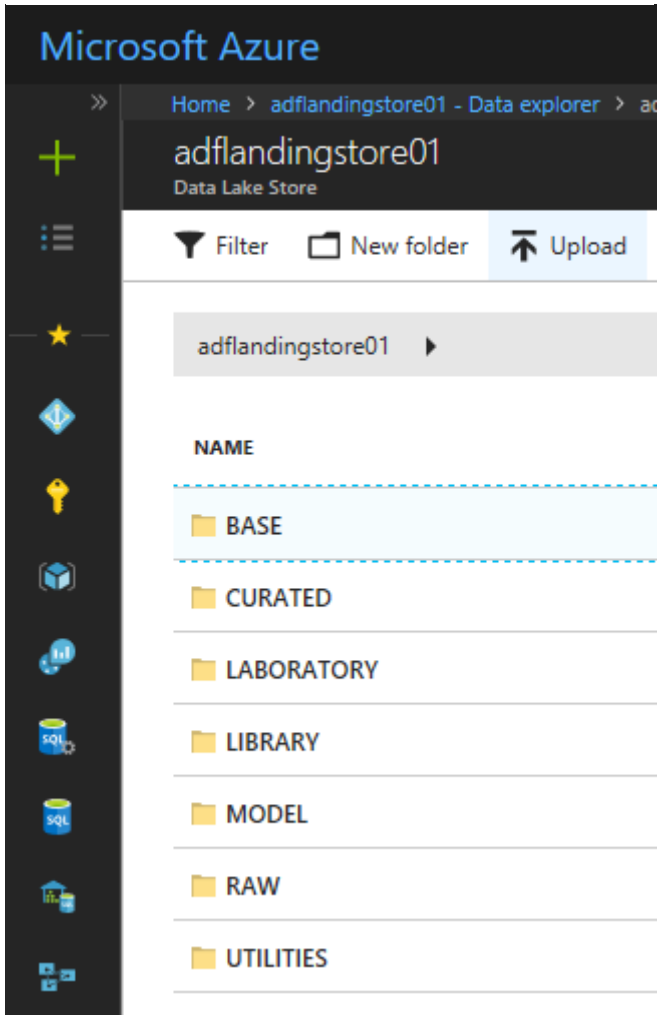
How can we work with Data Lake?

Development & Management

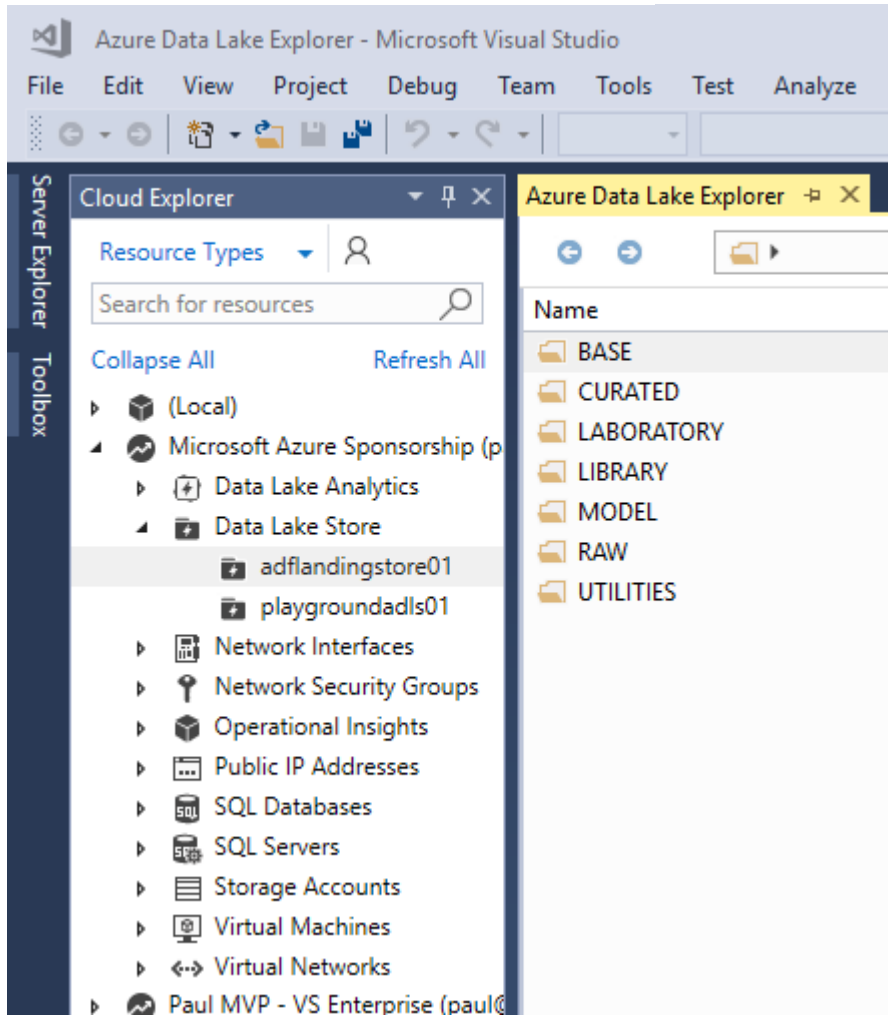
Working with Azure Data Lake Storage

Manual File Uploads

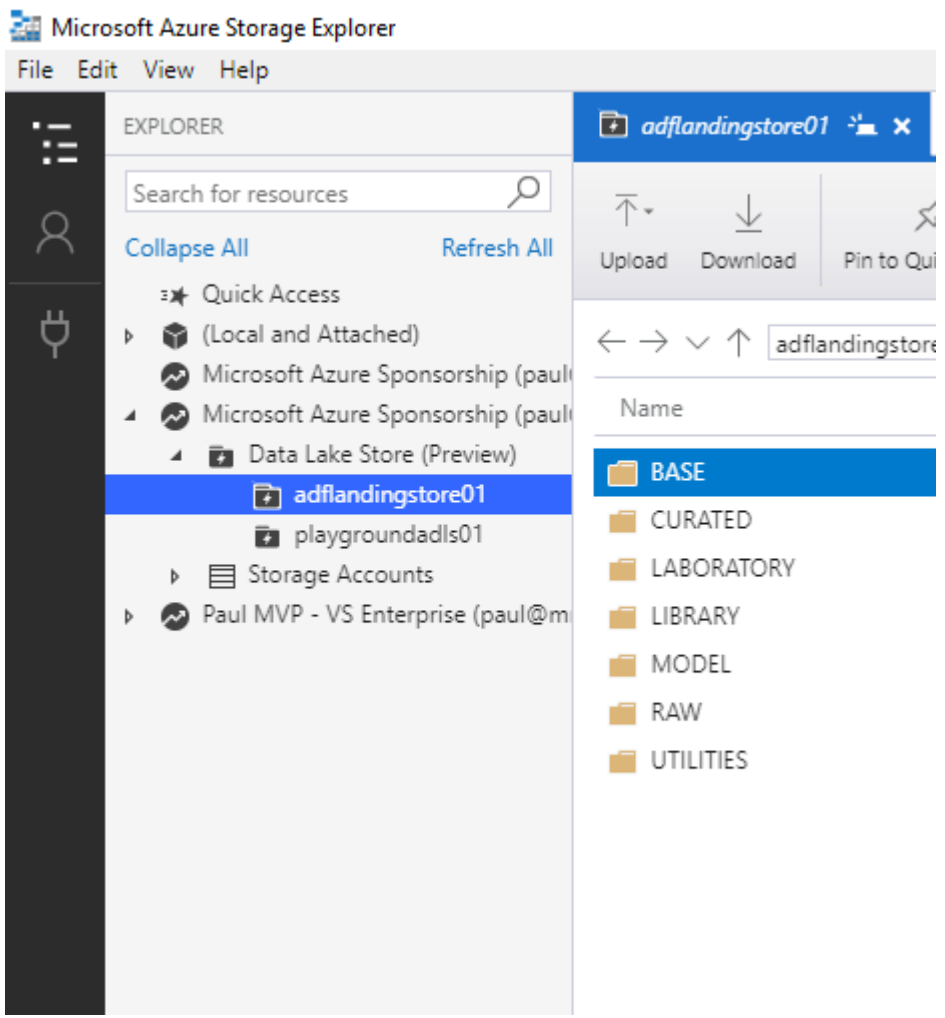
Azure Portal



Visual Studio Cloud Explorer



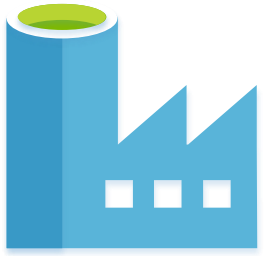
Azure Storage Explorer



Working with Azure Data Lake Storage

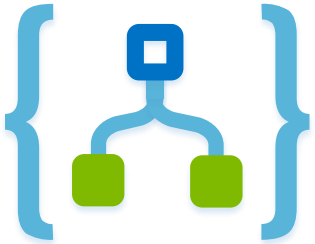
Automatic File Uploads

Data Factory



Stream Analytics

Logic Apps



SSIS

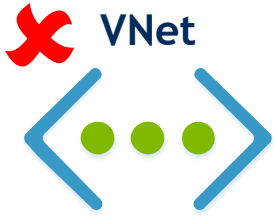
.Net SDK

PowerShell

Python

REST API

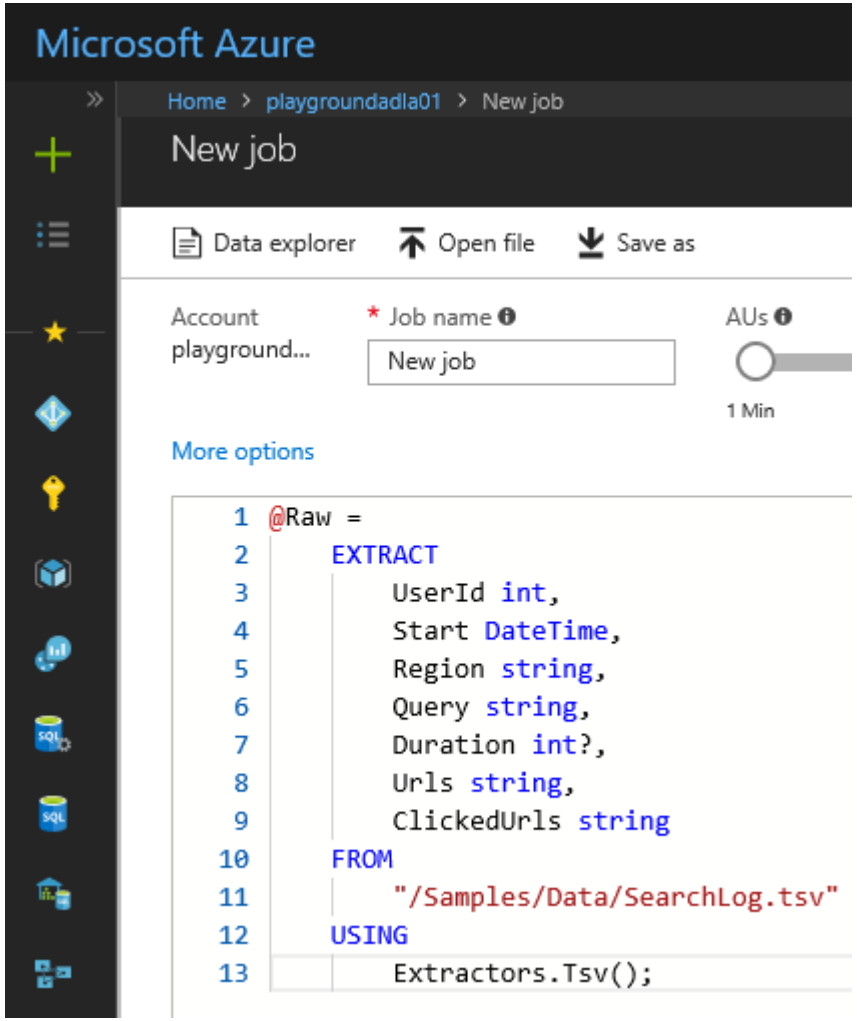
VNet



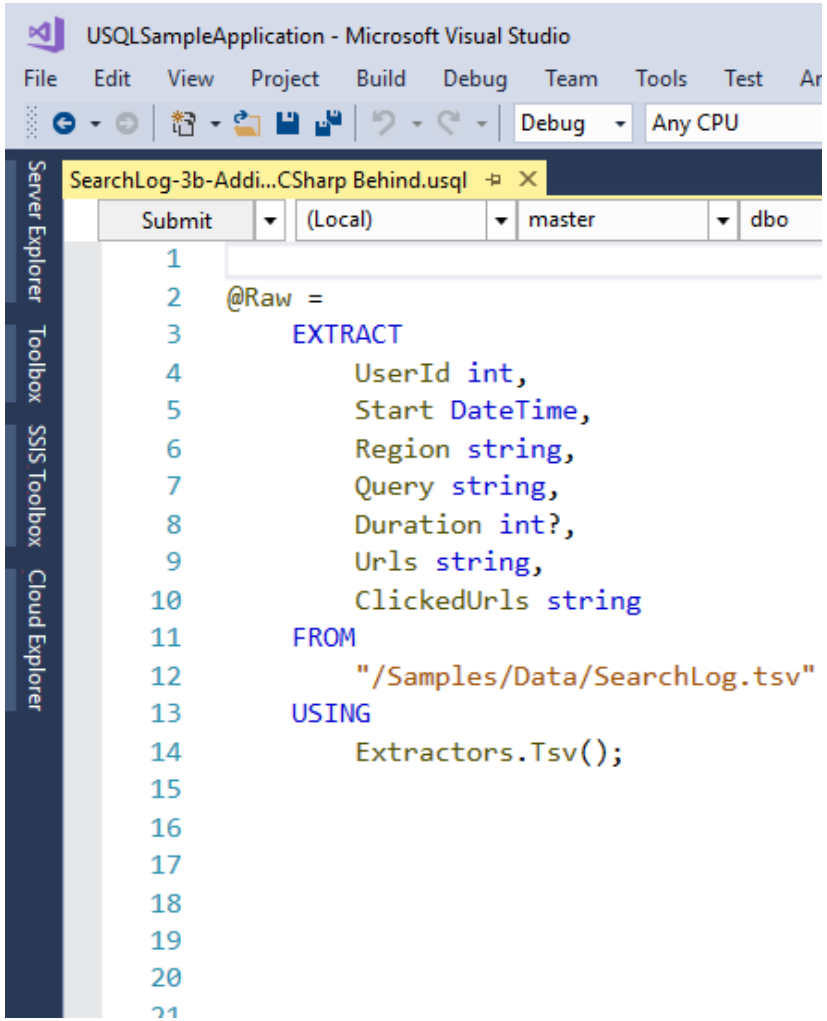
Working with Azure Data Lake Analytics

Manual U-SQL Job Execution

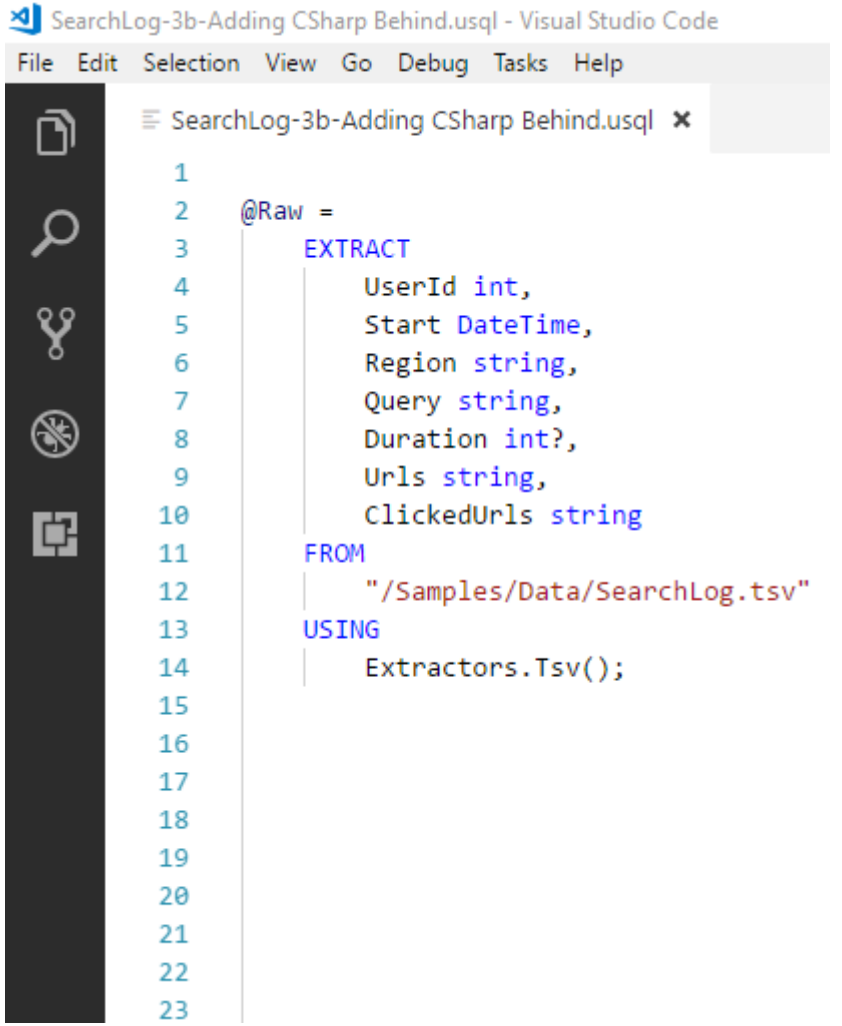
Azure Portal



Visual Studio Project

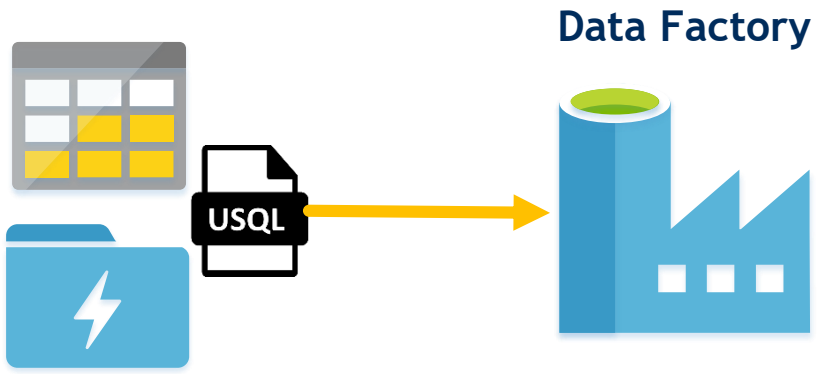


Visual Studio Code



Working with Azure Data Lake Analytics

Automatic U-SQL Job Execution

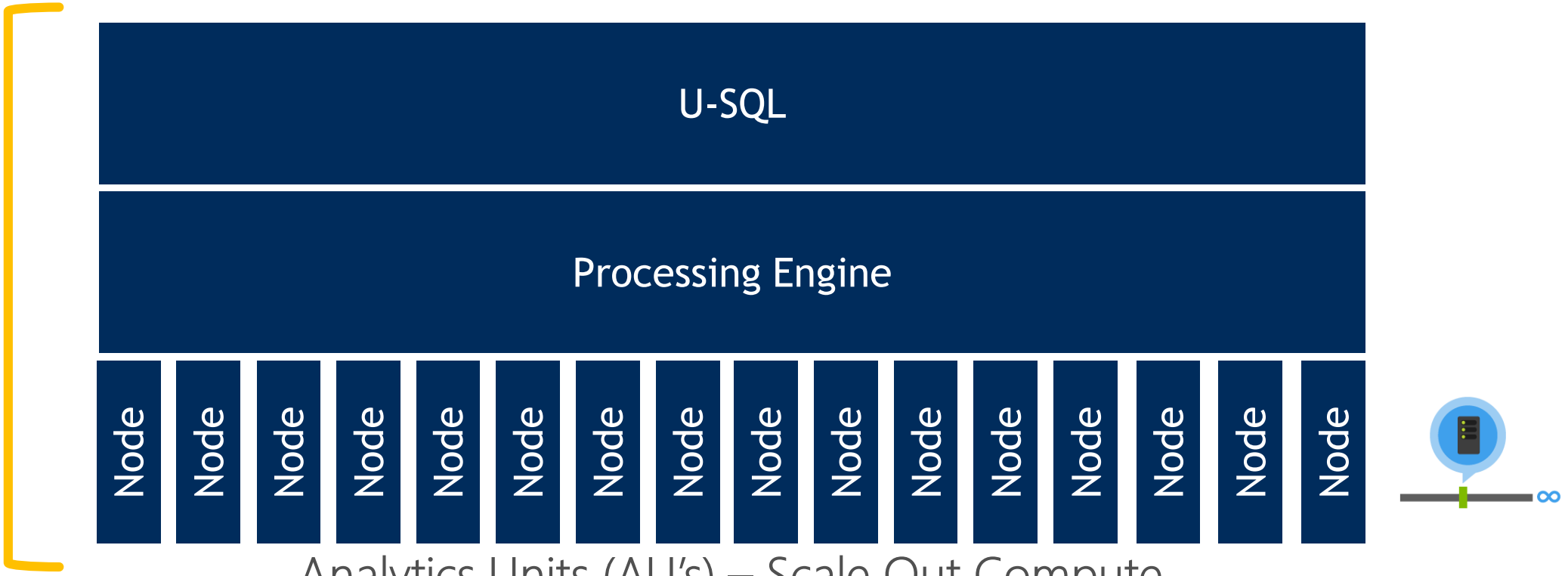


.Net SDK PowerShell Azure CLI Python Java Node.js

Working with Azure Data Lake Analytics

Job Execution

Analytics



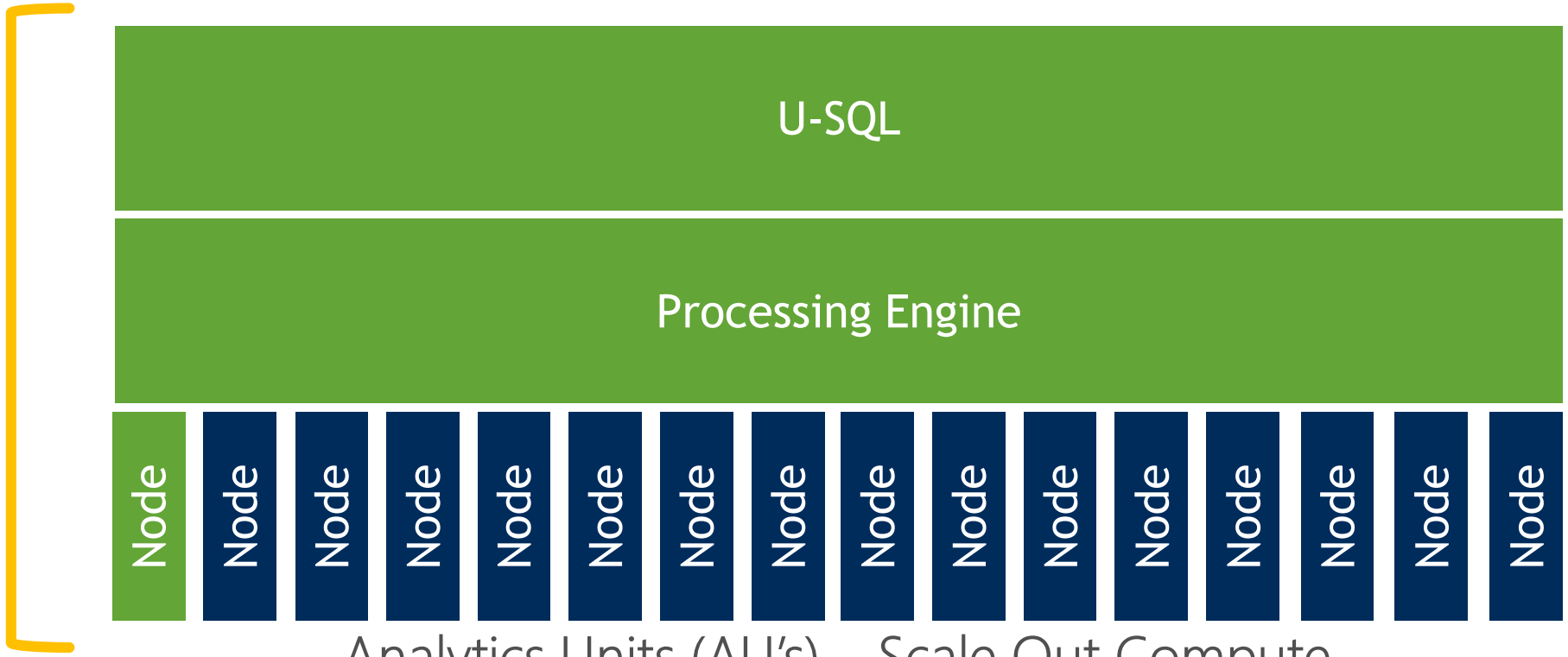
Analytics Units (AU's) – Scale Out Compute

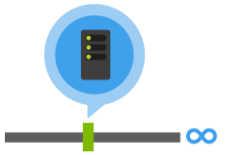
AU/hours

Working with Azure Data Lake Analytics

Job Execution

Analytics





Analytics Units (AU's) – Scale Out Compute

1 x AU/hour = £1.49 *

* Price Checked April 2018

Working with Azure Data Lake Analytics

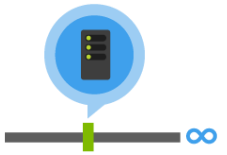
Job Execution

Analytics

U-SQL

Processing Engine

Node Node Node Node Node Node Node Node Node Node Node Node Node Node Node Node



Analytics Units (AU's) – Scale Out Compute

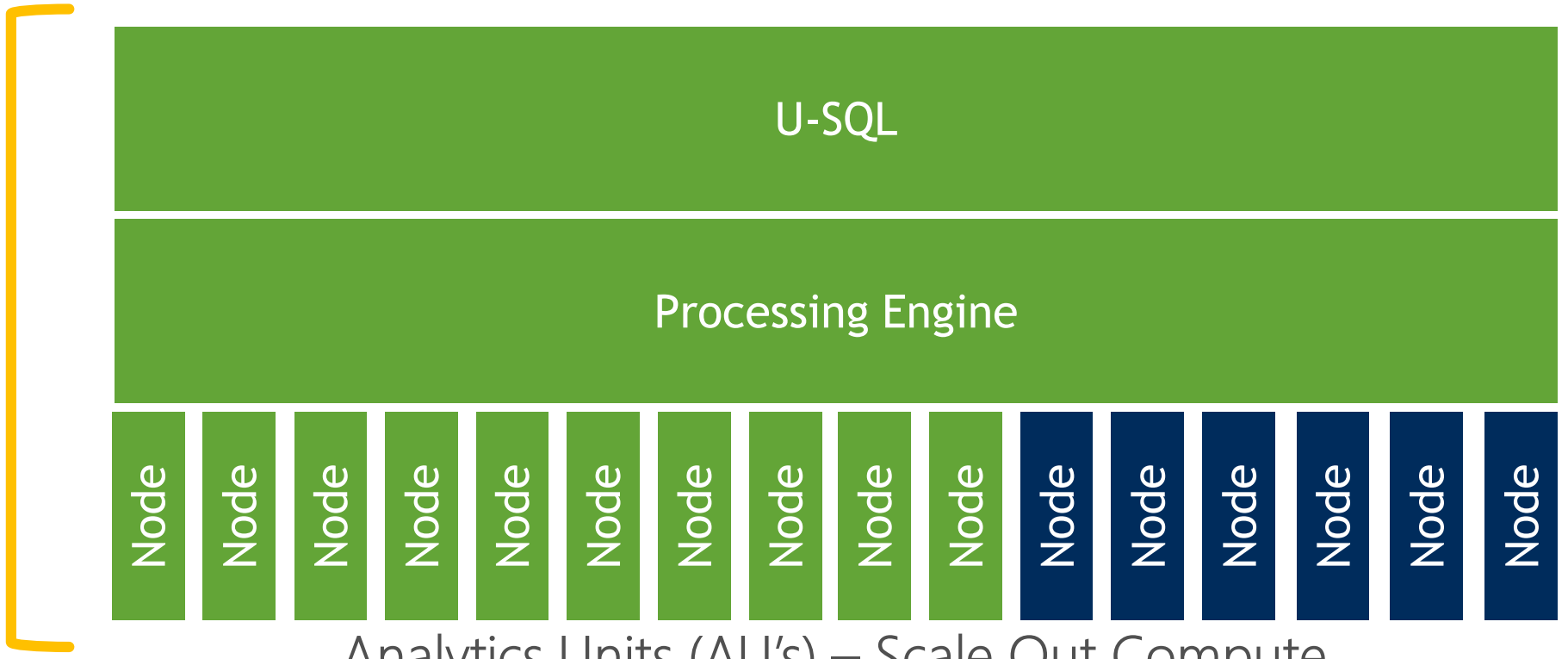
2 x AU/hour = £2.98 *

* Price Checked April 2018

Working with Azure Data Lake Analytics

Job Execution

Analytics



Analytics Units (AU's) – Scale Out Compute

10 x AU/hour = £14.90 *

* Price Checked April 2018

Working with Azure Data Lake Analytics

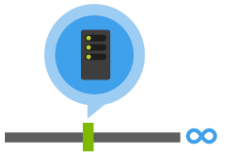
Job Execution

Analytics

U-SQL

Processing Engine

Node Node Node Node Node Node Node Node Node Node Node Node Node Node Node



Analytics Units (AU's) – Scale Out Compute

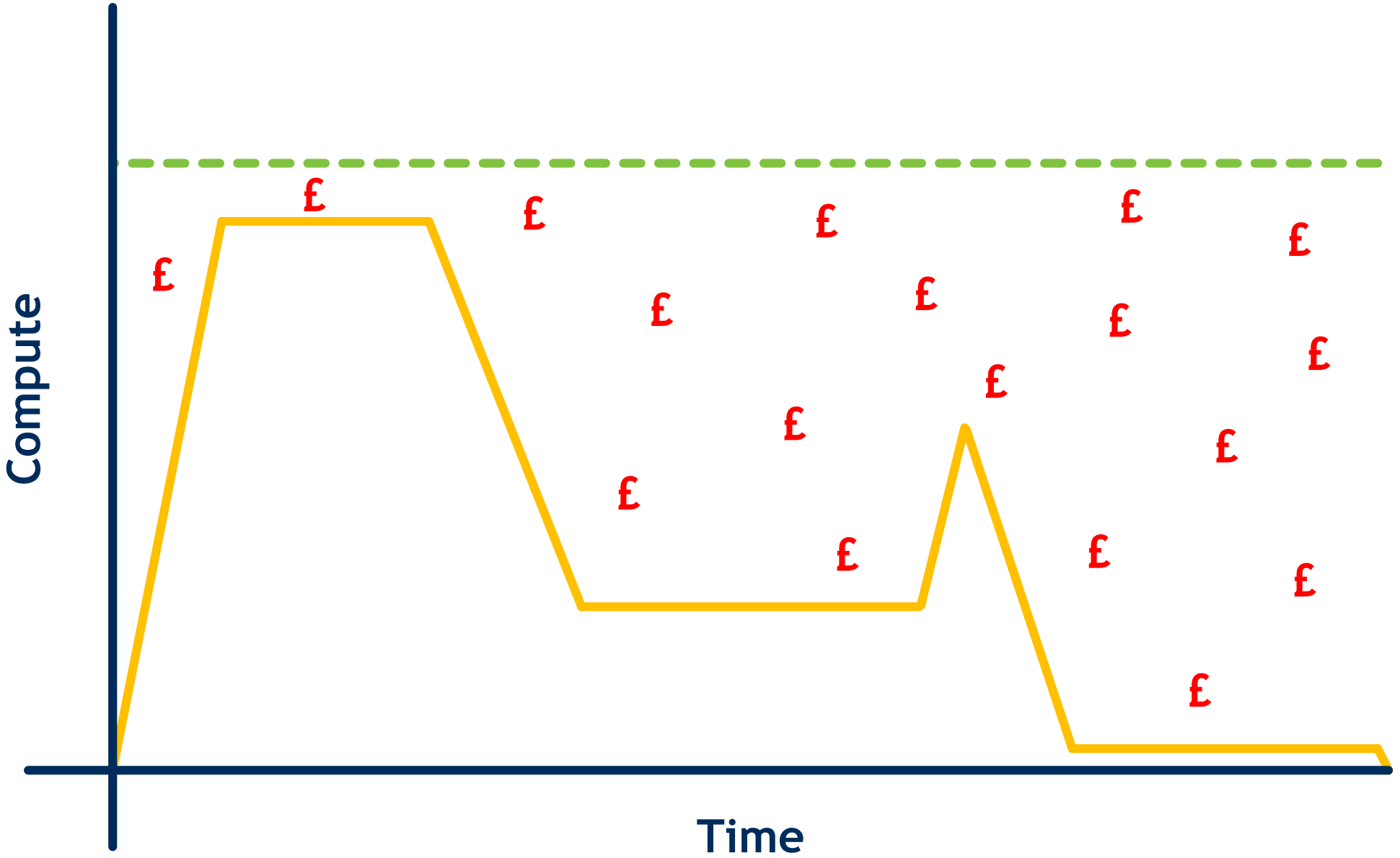
1 minute to complete

1 x AU/hour = ~~£1.49~~ *

£0.02

* Price Checked April 2018

Working with Azure Data Lake Analytics



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