Explore data with

Azure Data Explorer

A short story of database engine developed by Microsoft for Microsoft….and how it became customer facing project

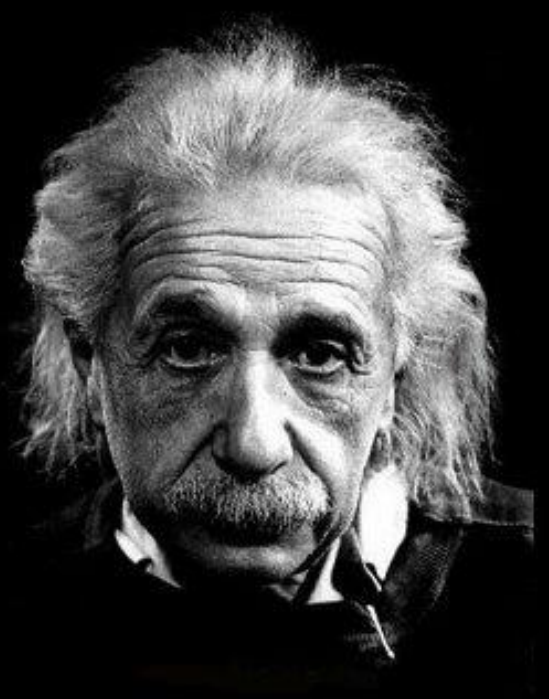
Bartłomiej Graczyk
Lead Cloud Solution Architect, Data & Analytics
bagra@microsoft.com

@GraczykBartek

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

"If you can't explain it simply, you don't understand it well enough."

Albert Einstein



Azure Data Explorer in a sentence

Any append-only stream of records

High volume
High velocity
High variance
(structured, semi-structured, free-text)

Relational query model: Filter, aggregate, join, calculated columns, …

Fully-managed

PaaS, Vanilla, Database

A big data analytics cloud platform optimized for interactive, ad-hoc queries

Purposely built

Rapid iterations to explore the data

Necessity is the mother of invention….

Telemetry

50B
connected devices
by 2030

175PB
total amount of data
by 2025

Telemetry Data in a Modern Analytics Environment

- Integrates data silos
- Enables real-time insights
- Monitors the health of systems, apps and devices
- Can quickly identify trends and anomalies
- Enables root cause analysis
- Provides scale and performance at an economical price

….let’s see it now… [DEMO]

Image source: https://tiffanyleestudios.com/product/hipster-giclee-wall-art-print-seeing-is-believing/

Let's look under the hood...

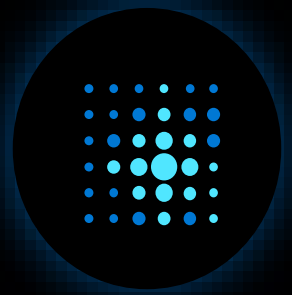
Image source: https://tiffanyleestudios.com/product/hipster-giclee-wall-art-print-seeing-is-believing/

Azure Data Explorer Key Pillars

Advanced Storage Engine

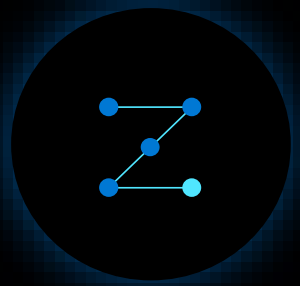
Also available in Azure Synapse Analytics

Optimized



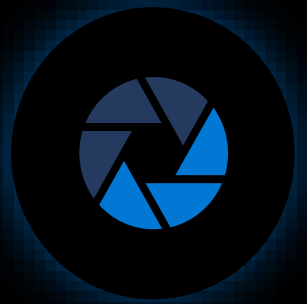
Time stamped data: logs, time series, telemetry

Native Support



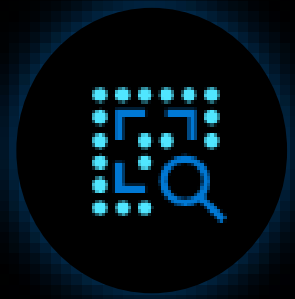
Free text, structured, and semi-structured

Low Latency



Near real-time data analytics at scale

Ad-Hoc Queries



Human friendly query language

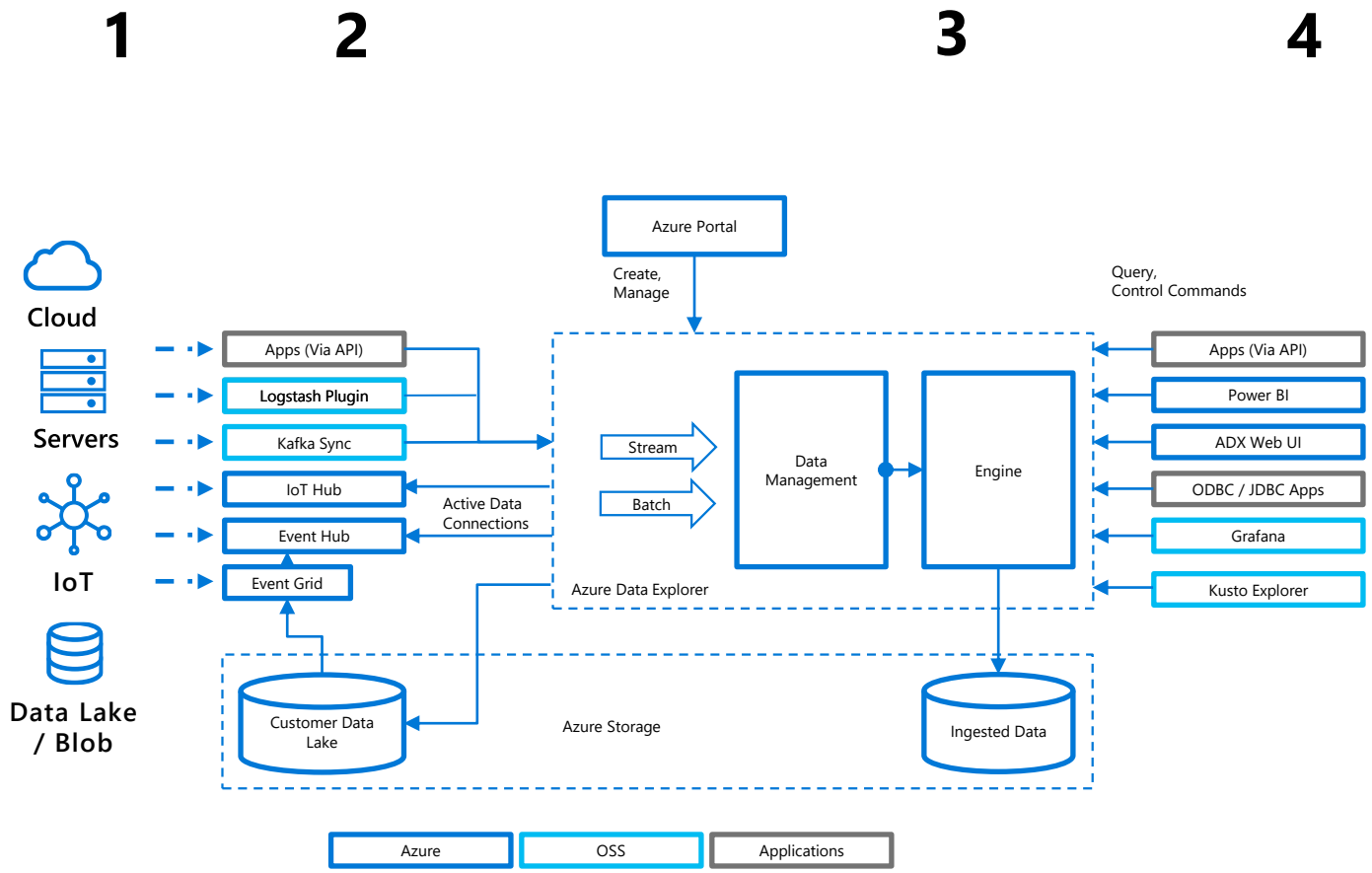
Azure Data Explorer overview

1. **Capability for many data types, formats, and sources**
   Structured (numbers), semi-structured (JSON\XML), and free text

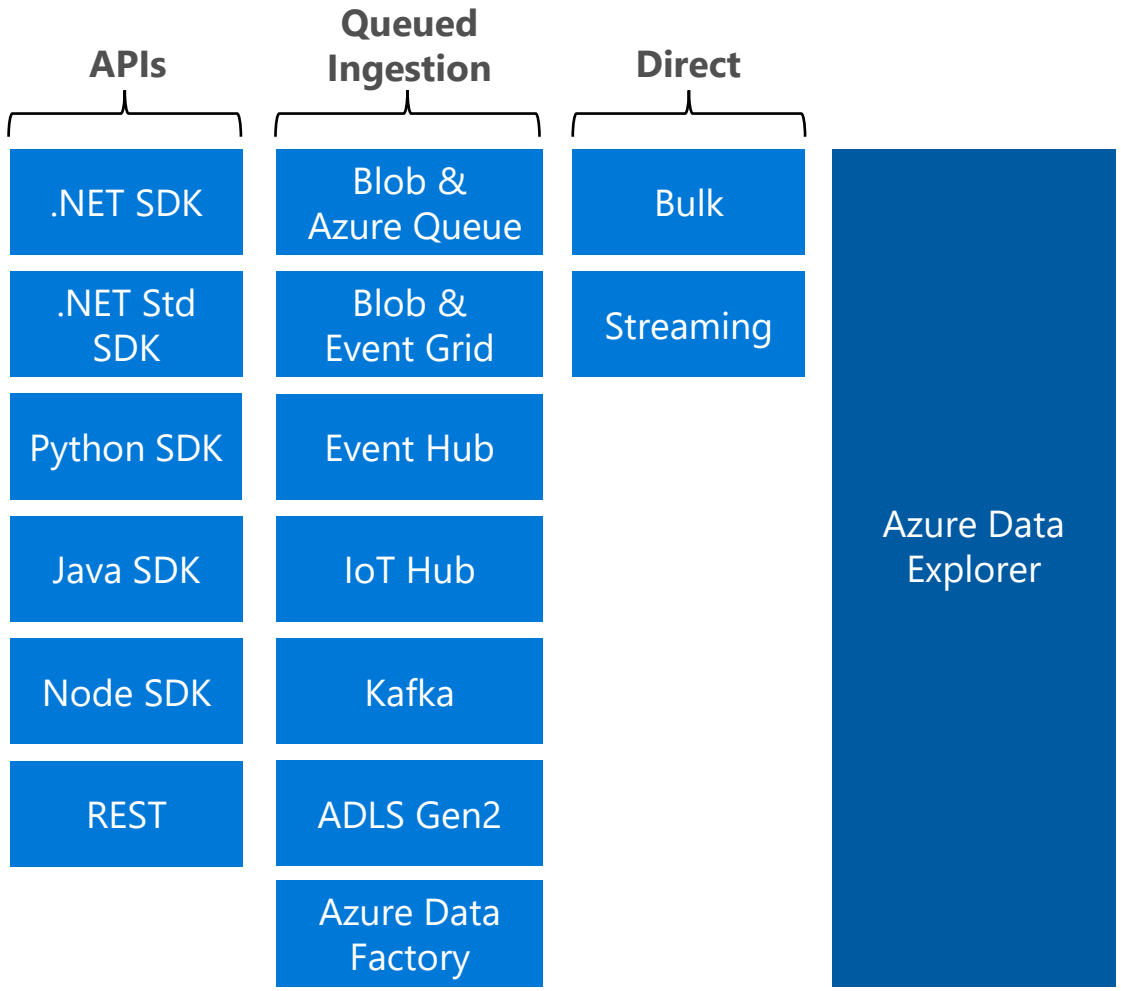
2. **Batch or streaming ingestion**
   Use managed ingestion pipeline or queue a request for pull ingestion

3. **Compute and storage isolation**
   - Independent scale out / scale in
   - Persistent data in Azure Blob Storage
   - Caching for low-latency on compute

4. **Multiple options to support data consumption**
   Use out-of-the box tools such as Power BI and connectors or use APIs/SDKs for custom solution



How to get the data in and how to access the data



How to get the data in and how to access the data
Fast ingestion

Optimized for streaming data

Easy input from multiple data sources

Multiple data sources

- Managed ingestion (e.g. Event Hub, IoT Hub) or programmatic ingestion (e.g. connectors, SDKs)

Versatile ingestion

- Use batch or streaming ingestion

Easy input from multiple formats

- Tabular formats: CSV, TSV, PSV, SCSV
- JSON (line-separated, multiline), Avro
- ZIP and GZIP compression (for Batch)

Instant integration with simple transforms

- Reshape the data with update policies (Database Ingest Triggers)

Managed services


Azure Event Grid


Azure Event Hub


Azure IoT Hub

Azure Blob

Azure Data Lake

Connectors/Plugins

SDKs and APIs

Intuitive querying

Designed for data exploration

Simple and powerful

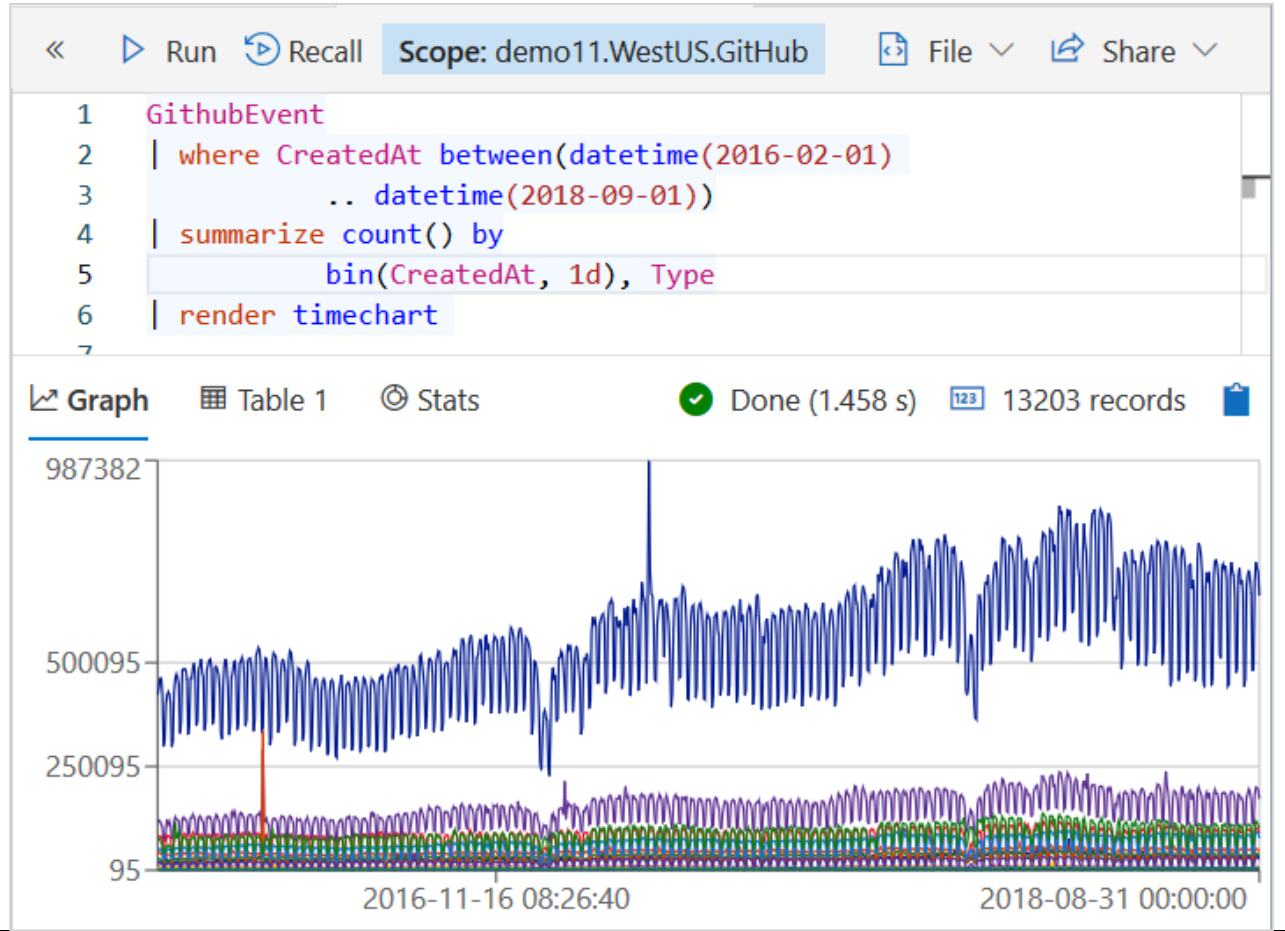
- Rich rational query language (filter, aggregate, join, calculated columns, and more)
- Built-in full-text search, time series, user analytics, and machine learning operators
- Out-of-the box visualization (render)
- Easy-to-use syntax + Microsoft IntelliSense
- Highly recognizable hierarchical schema entities

Comprehensive

- Built for querying over structured, semi-structured and unstructured data simultaneously

Extensible

- In-line Python
- SQL



Simple provisioning

Fully managed for efficiency

Easy provisioning

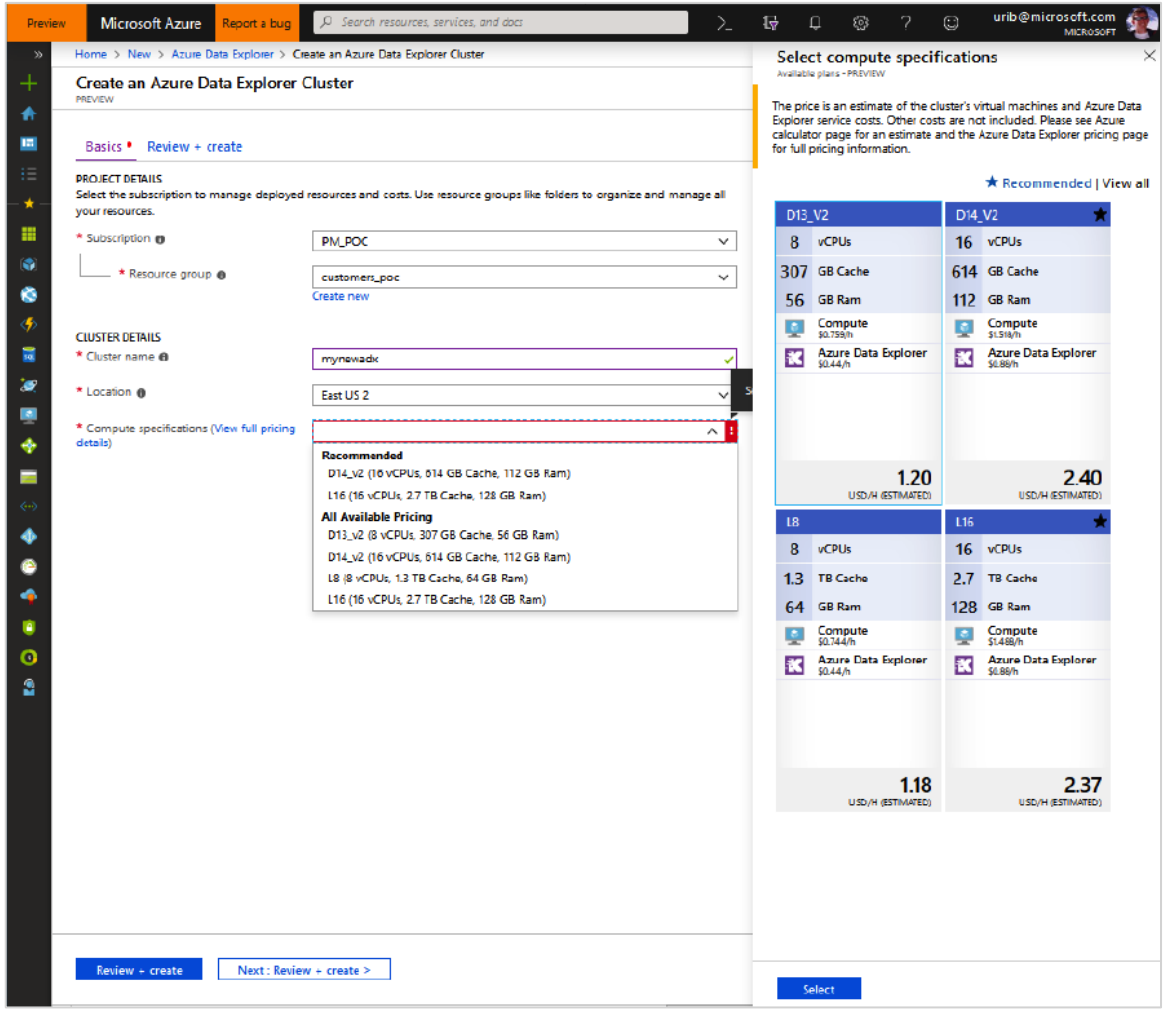
- No infrastructure to manage: Azure PaaS
- Use Azure Portal, APIs, or PowerShell to provision
- Storage Optimize/Compute Optimize SKUs
- Flexible data caching and retention options at database and table level

Rapid elasticity

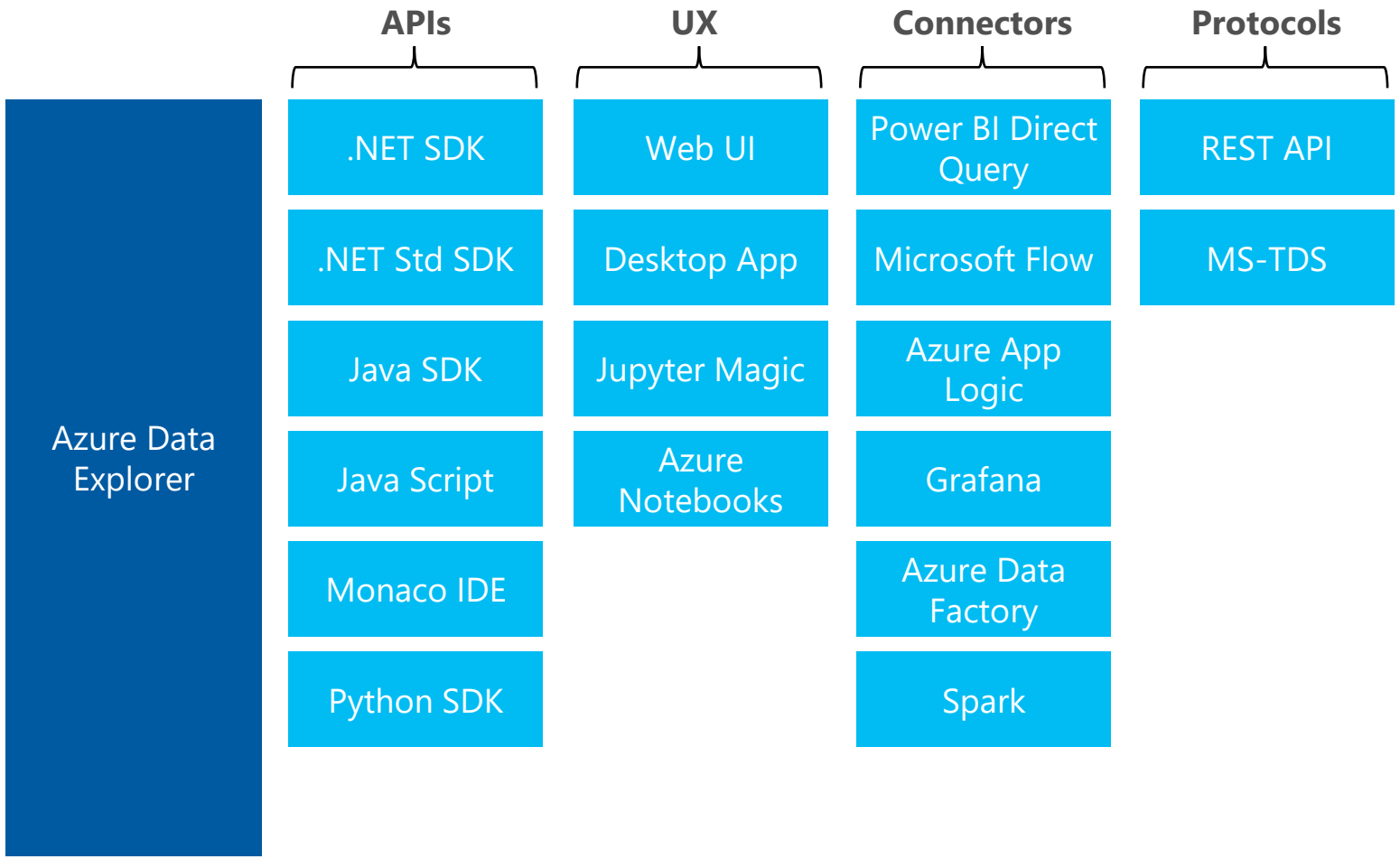
- Buy only what you need
- Scale out/in manually or use autoscale
- Dedicated resources

Maintenance-free

- All columns are compressed and indexed during ingestion
- No index maintenance required

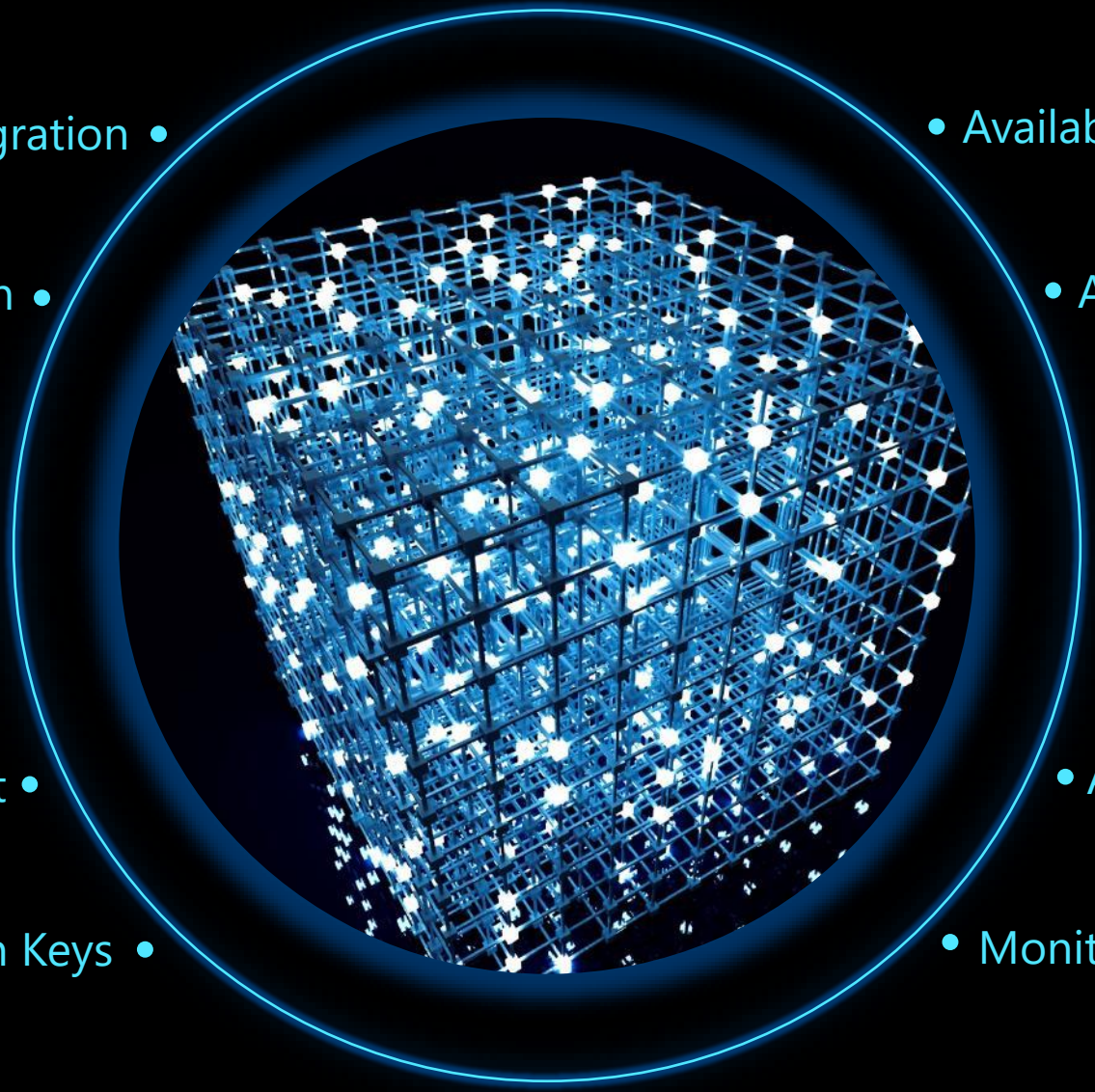


How to get the data in and how to access the data



Enterprise Ready – Mission Critical

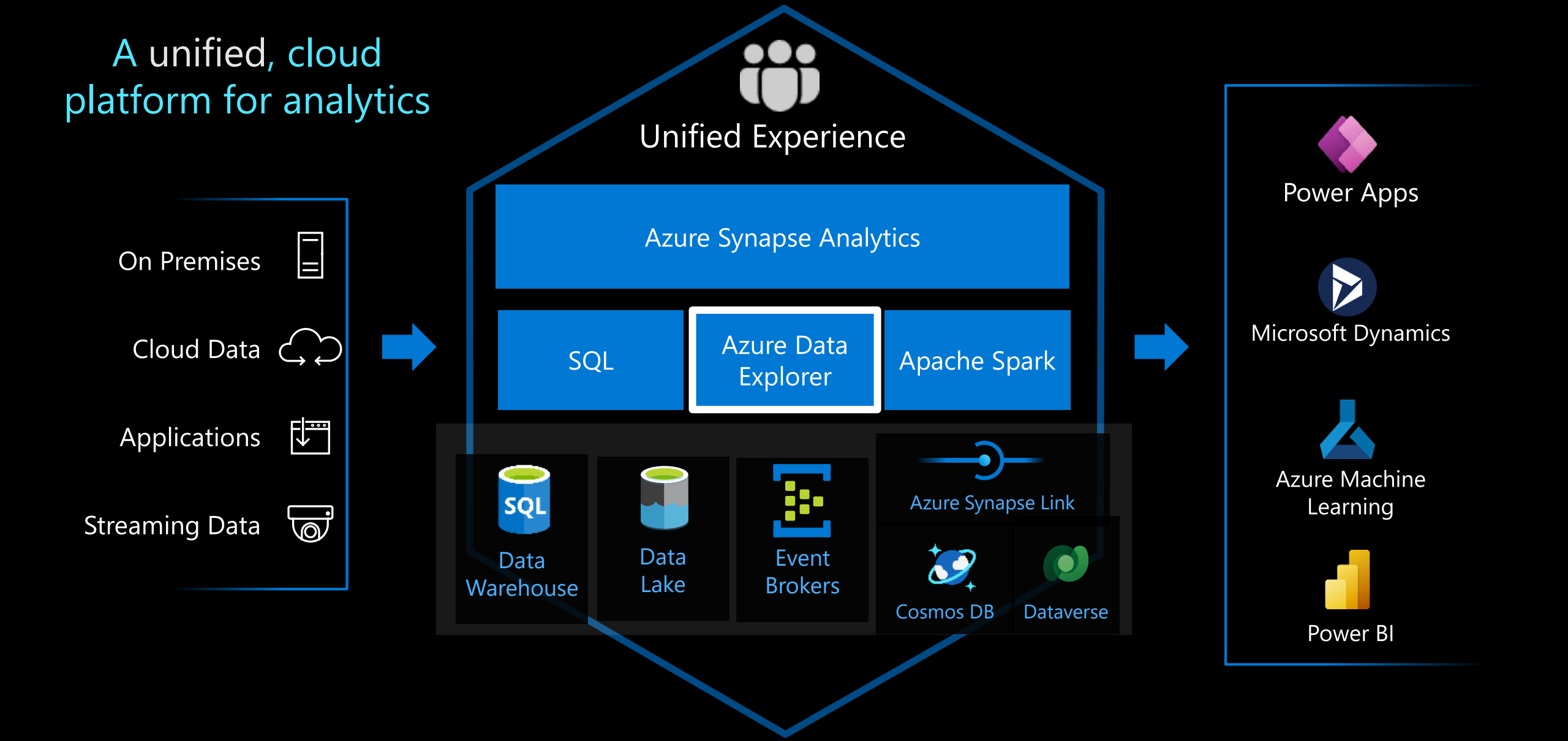
- Azure Active Directory Integration
- Role based authorization
- Network Perimeter Security
- Encryption with customer managed keys/"BYOK"
- Build Azure Policy Support
- Bring Your Own Keys



- Availability Zones
- Auto Scale-Up/In
- Globally available
- CI/CD Integration
- Automated provisioning
- Monitoring

Part of the Azure Analytics Solution

A unified, cloud platform for analytics

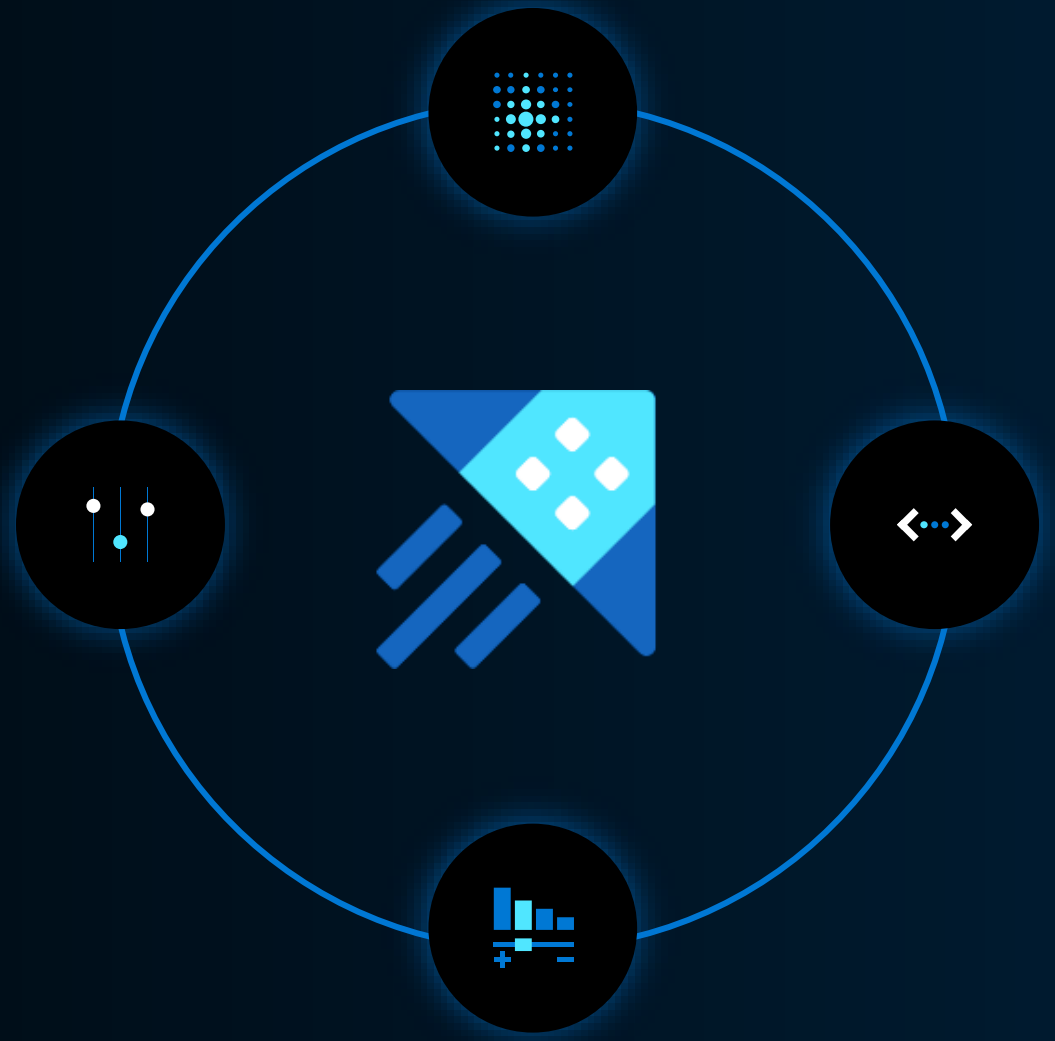


Microsoft

Benchmark: Azure Data Explorer

Conducted by GigaOm, a leading independent analyst.

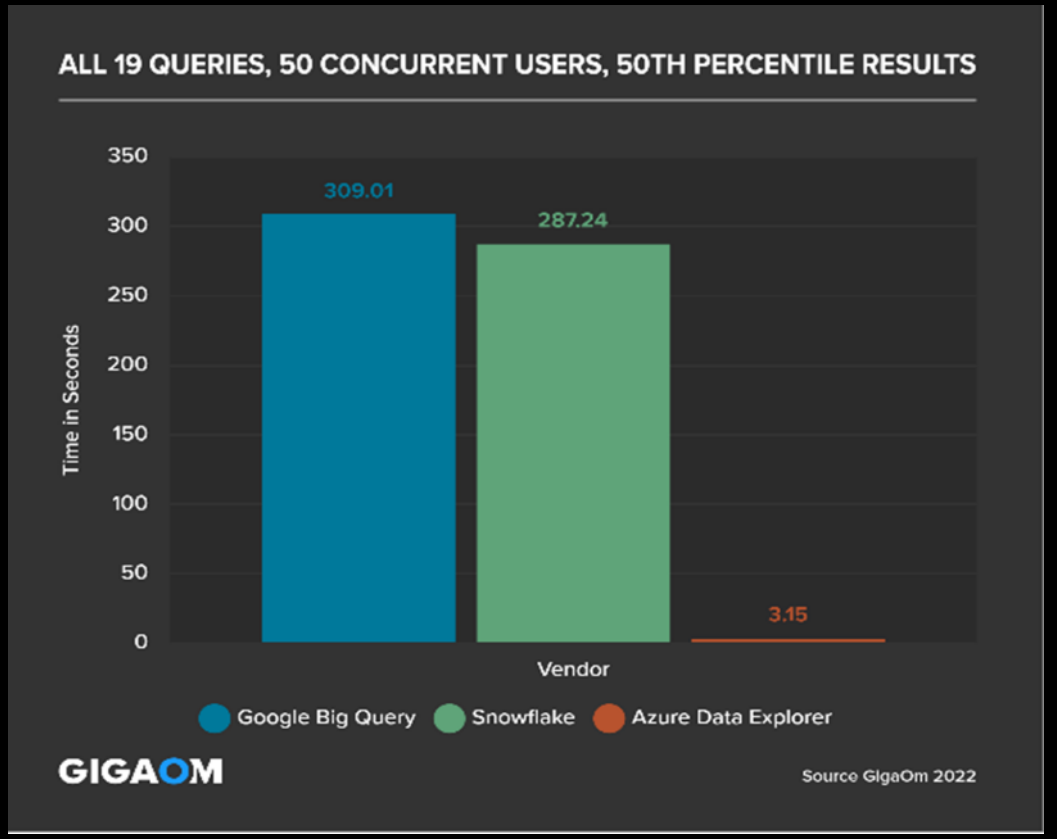
GigaOm Log Analytical Field Test: Workload simulates activities in complex, real-world environments.



Azure Data Explorer Shines!

Results

- ADX outperformed Google BigQuery and Snowflake on all 19 tests featuring a single user.
- ADX won in 18 of 19 tests featuring 50 concurrent users.
- Average execution time on BigQuery and Snowflake was up to 15x slower.
- BigQuery and Snowflake had eight queries that did not complete within the two-minute timeout.



"The Azure Data Explore query language, KQL, was by far the simplest and easiest syntax to write"

Full Report: https://aka.ms/adx.benchmark

….let’s continue the journey… [DEMO]

Microsoft

Welcome to the
Kusto Detective Agency

https://detective.kusto.io/

Thank You

Product

- Product Page: http://aka.ms/AzureDataExplorer
- Docs: https://aka.ms/adx.docs
- Cost Estimator: http://aka.ms/adx.cost
- Free online Courses:
  KQL from Scratch, Azure data exploring, How to start with Azure Data Explorer (blog), Advanced KQL (blog)
- Lab: https://aka.ms/adx.lab
- Whitepaper: https://azure.microsoft.com/en-us/resources/azure-data-explorer/en-us/
- 101 blog: https://azure.microsoft.com/en-us/blog/azure-data-explorer-technology-101/
- Reference Architectures - https://aka.ms/adx.architectures

Social and Community

- Twitter: @AzDataExplorer
- Tech Community Blog: https://aka.ms/adx.blog
- Tech Community Forum: http://aka.ms/adx.techcommunity
- Stack overflow: https://aka.ms/adx.sof
- YouTube Channel: https://aka.ms/adx.youtube