

Rolf Tesmer

Microsoft Australia

Azure Cloud Solution Architect | Data | Analytics | AI

Linked In: <https://www.linkedin.com/in/rolftesmer/>

Blog: <https://mrfoxsql.wordpress.com/>



Building Streaming Data Pipelines Using Azure Cloud Services



My Assumptions for Today

1. You know what the **cloud** and what **Azure** is
2. You have **awareness** of the **main Azure services**
3. You have used the **Azure Portal** before
4. You have **knowledge** of **database services**
5. You will **rate** my session **10 / 10**



Basic Program for Today

- 10 Introduction
- 20 What exactly is the data platform nowadays?
- 30 Data pipeline services in Azure
- 40 Demonstration: Lets see some data pipelines!
- 50 Where to next? Wrap up and summary
- 60 goto 10

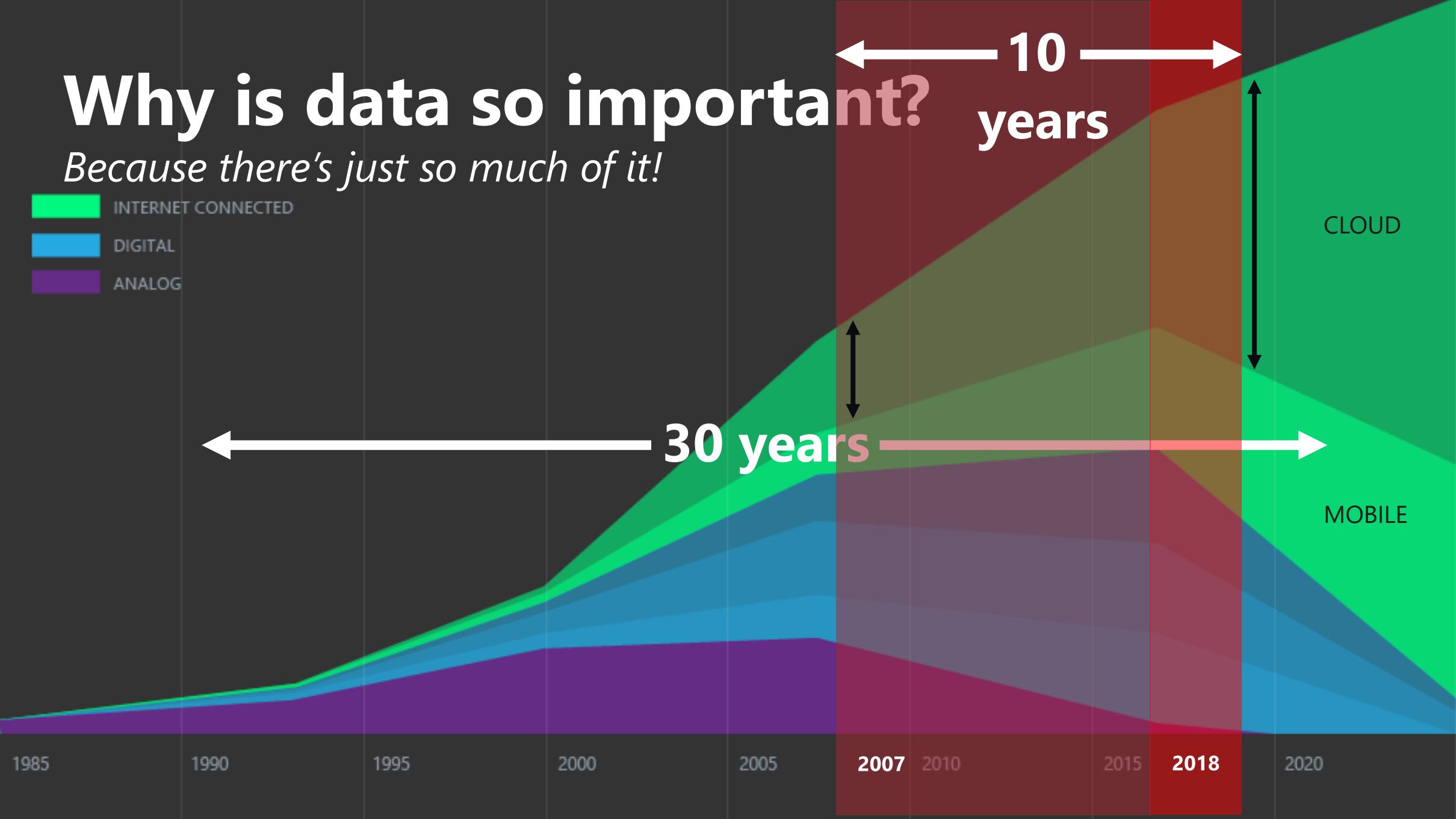




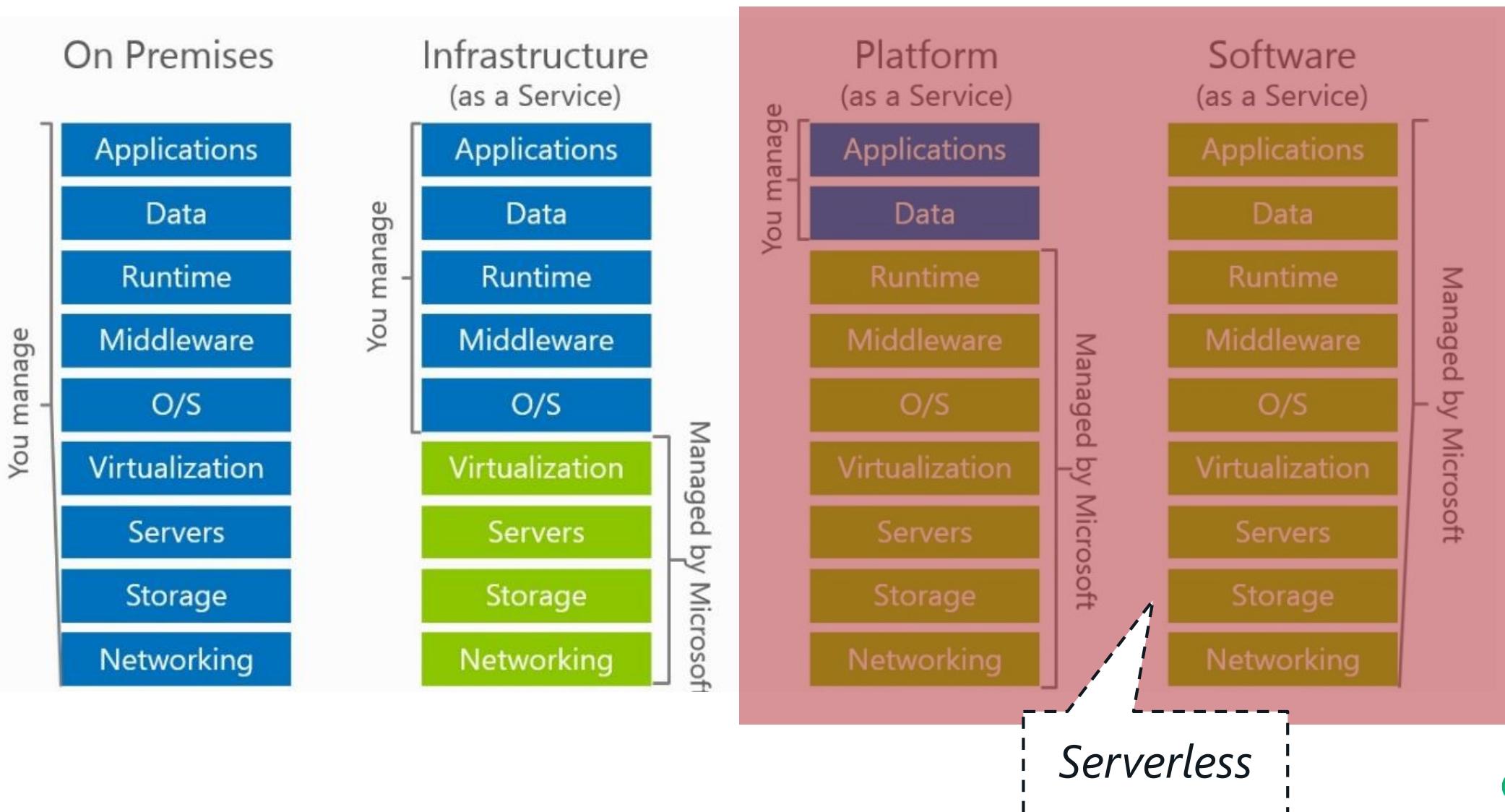
Introduction

Why is data so important?

Because there's just so much of it!

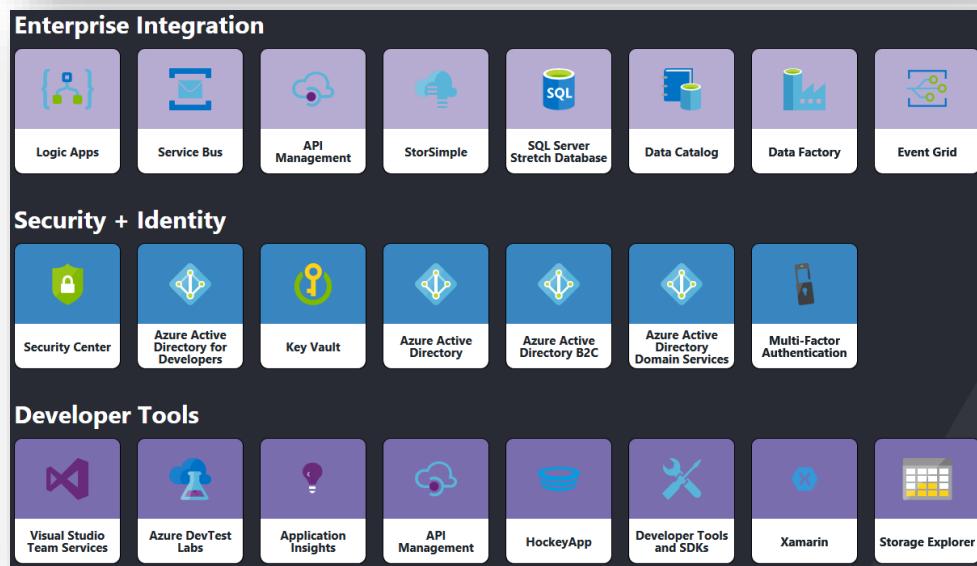
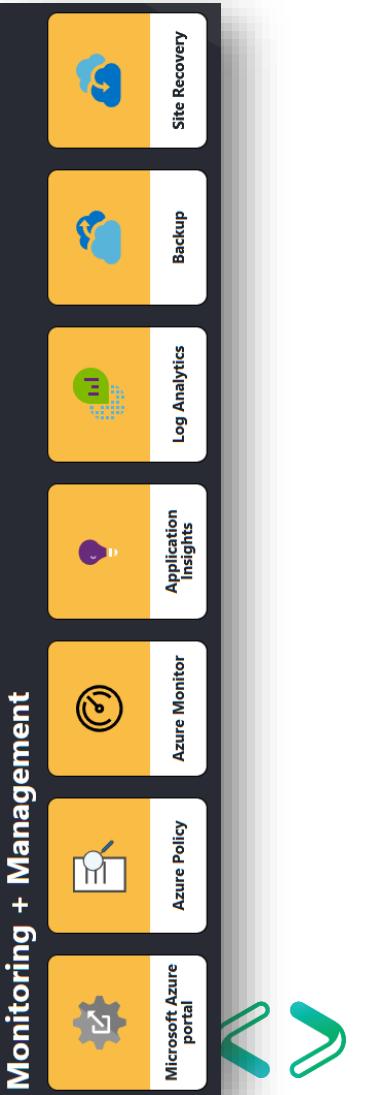
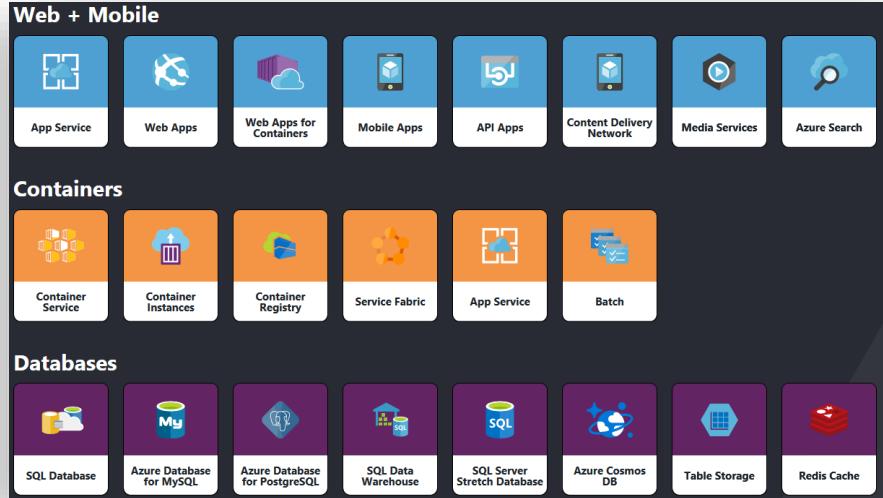


On-Prem vs IaaS vs PaaS vs SaaS – Which One?



Azure Services – Which One?

<http://azureinteractives.azurewebsites.net/Azure101Cards/default.html>



And so what exactly is a “data pipeline” anyway?

- Different definitions **depending on which vendor you talk to**
- **Microsoft** have **no formal definition**, but some I like...

*“**pipelines** are formed from multiple individual ‘fit for purpose’ services aligned in sequences that perform a set of specific targeted actions on data that is typically in transit.”*

Source: (Rolf Tesmer) ☺

*“a **pipeline** is a set of data processing elements connected in series, where the output of one element is the input of the next one. The elements of a **pipeline** are often executed in parallel or in time-sliced fashion”*

Source: (Wikipedia)

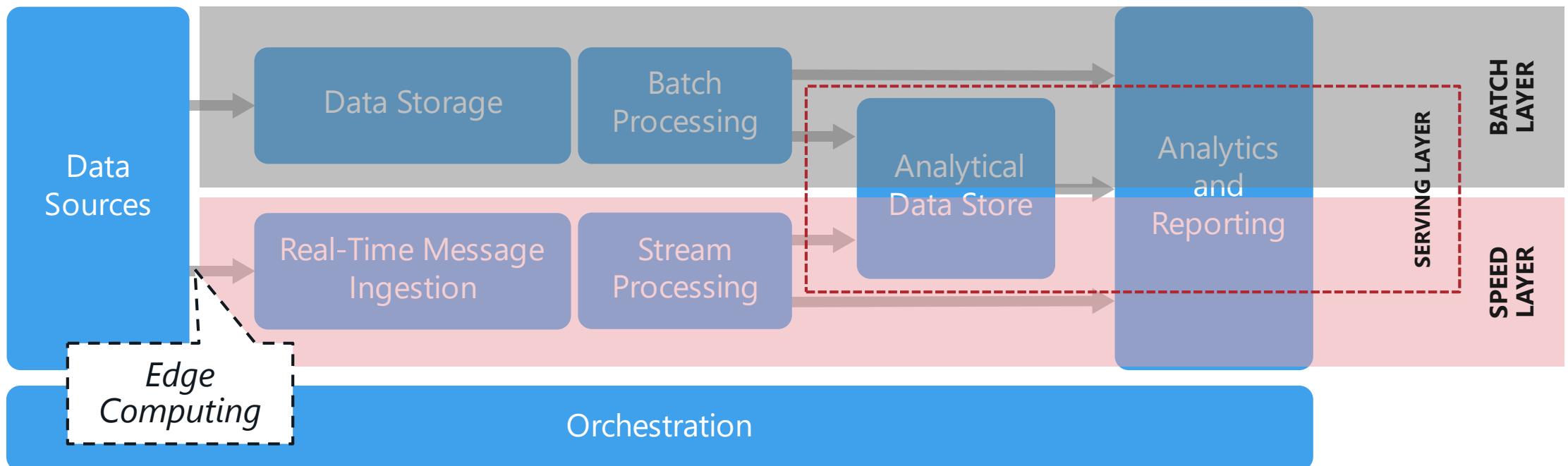
*“a data **pipeline** is the software that consolidates data from multiple sources and makes it available to be used strategically”*

Source: (Unknown)



What is the LAMBDA architecture?

*"The Objective of **Lambda Architecture** is to leverage the combined power of both **batch** & **real-time** processing to address the business scenarios where it requires both **historic view of the data** as well as getting insight into the **data in real-time** as business happens."*



<https://azure.microsoft.com/en-au/blog/the-emerging-big-data-architectural-pattern/>

<https://docs.microsoft.com/en-us/azure/architecture/guide/architecture-styles/big-data>

<https://azure.microsoft.com/en-us/services/iot-edge/>



Where did this come from, and why do we care?

1. Customers are on a **multi-year transformational journey**
2. Many **data sources** are **not static** or **at rest**
3. Solutions **cannot wait** for data to be landed before using it
4. building pipelines...
 - **Historically** → Complex, costly capital investment, time consuming
 - **Today** → Fast, simple, “fit for purpose” services, same **data platform**

As modern day Data Professionals we have to deal with it





What *exactly* is the
Data Platform
Nowadays?

What was the data platform?

Up till ~5 years ago it was typically a **relational platform**

...and... **included relational-like services** (OLTP, OLAP, DW, ETL, MDM, +)

...and... often **on-prem**, or in a hosted DC

...and... rarely hosted in external **public cloud** providers

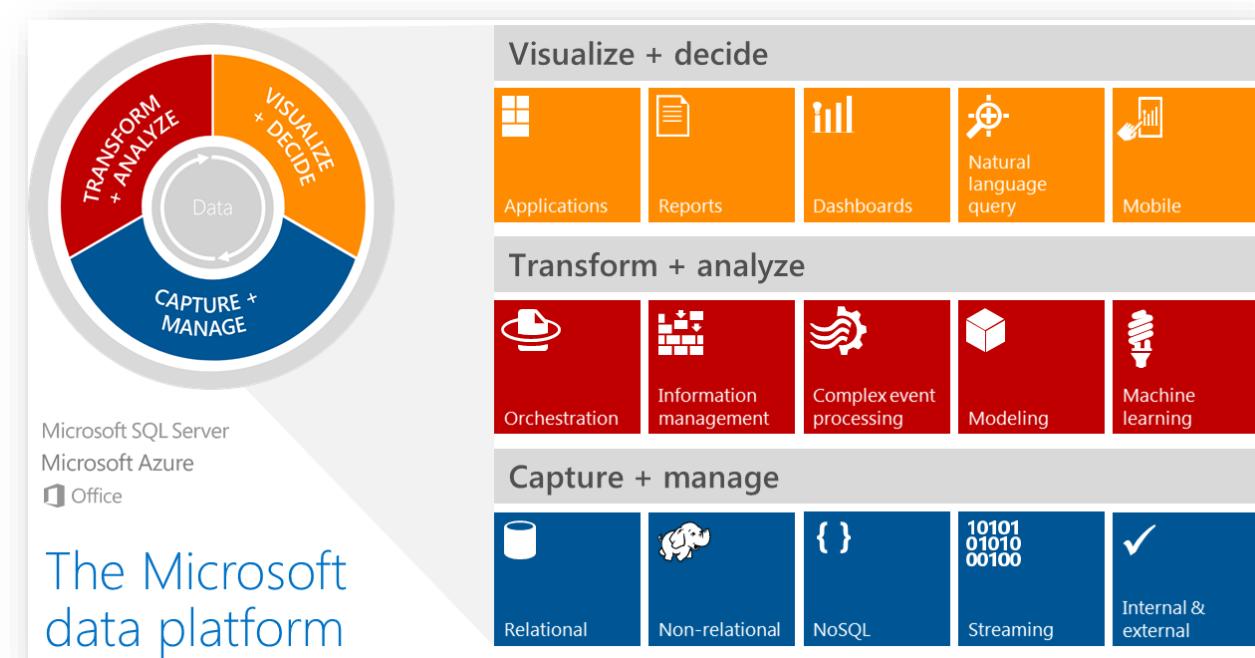
Occasionally included **special projects** (ie Big Data, NoSQL, IoT)



What is the data platform now?

- *A mix of... different data types, shapes, rates & sizes*
- *A mix of... on-prem & public cloud*
- *A mix of... deployment models (IaaS, PaaS, SaaS)*
- *A mix of... specific “fit for purpose” individual data services*

- These services are across a range of uses including;
 1. *Ingestion*
 2. *Transformation*
 3. *Storage*
 4. *Analytics*
 5. *Visualisation*

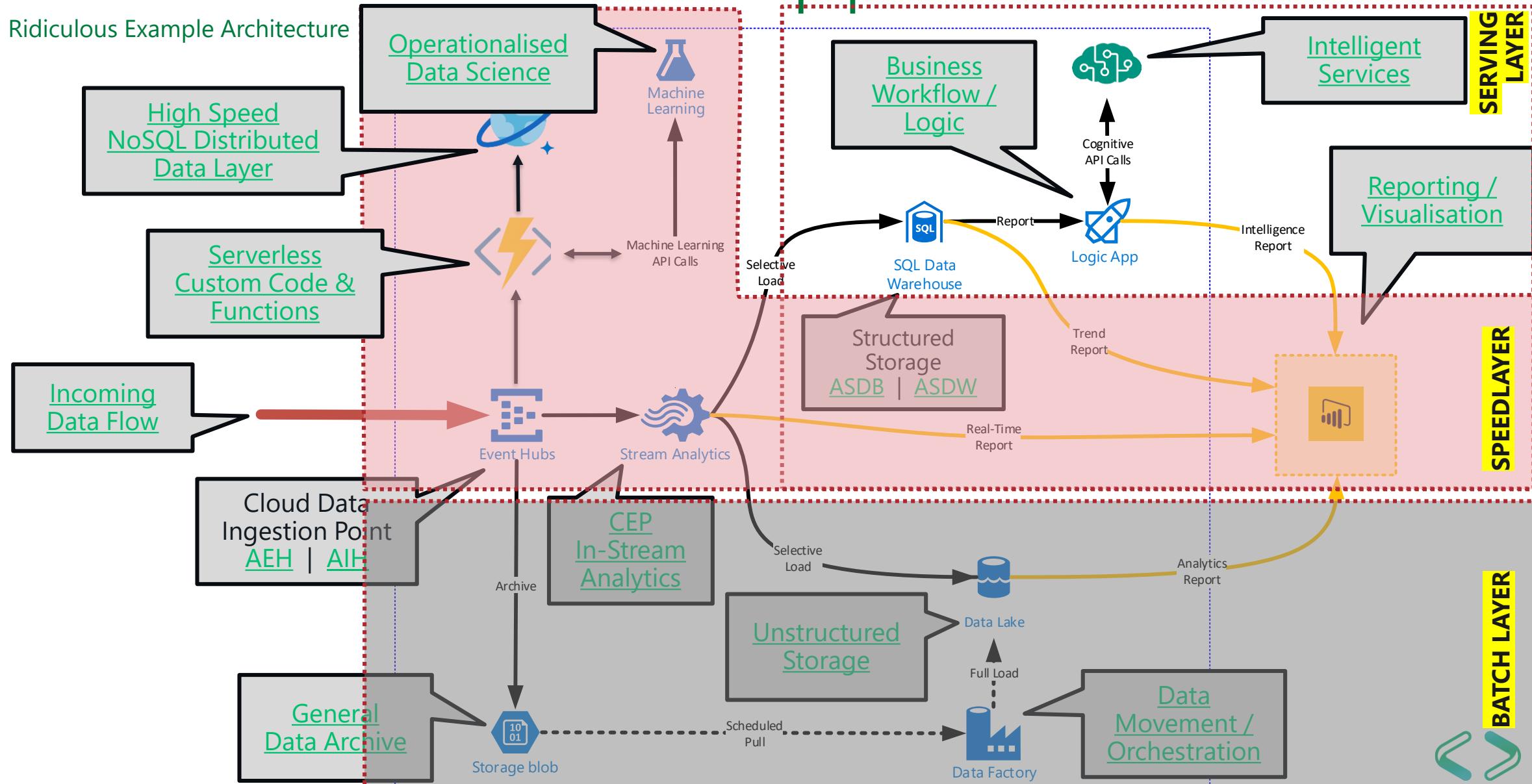




Data Pipeline Services in Azure

What are some of the Azure pipeline services?

Ridiculous Example Architecture

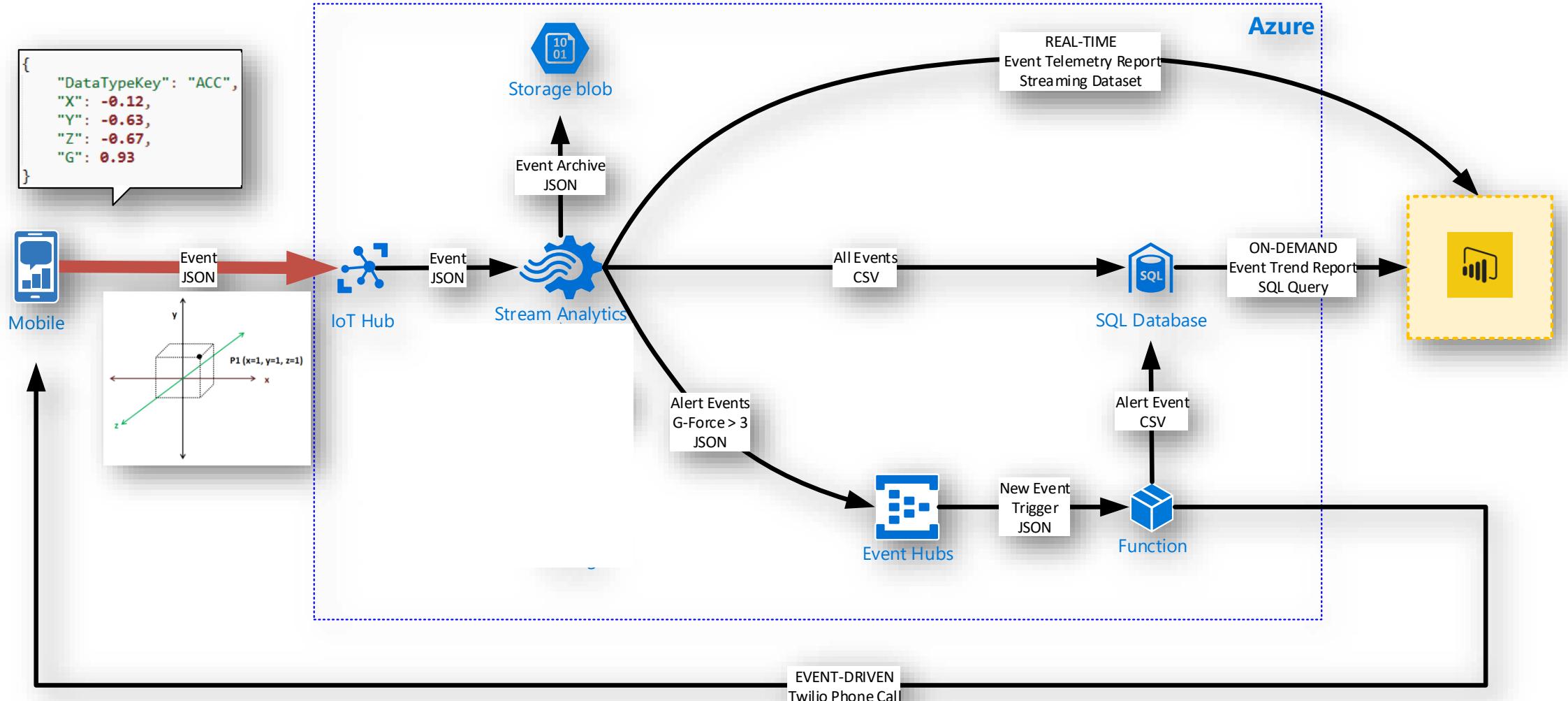




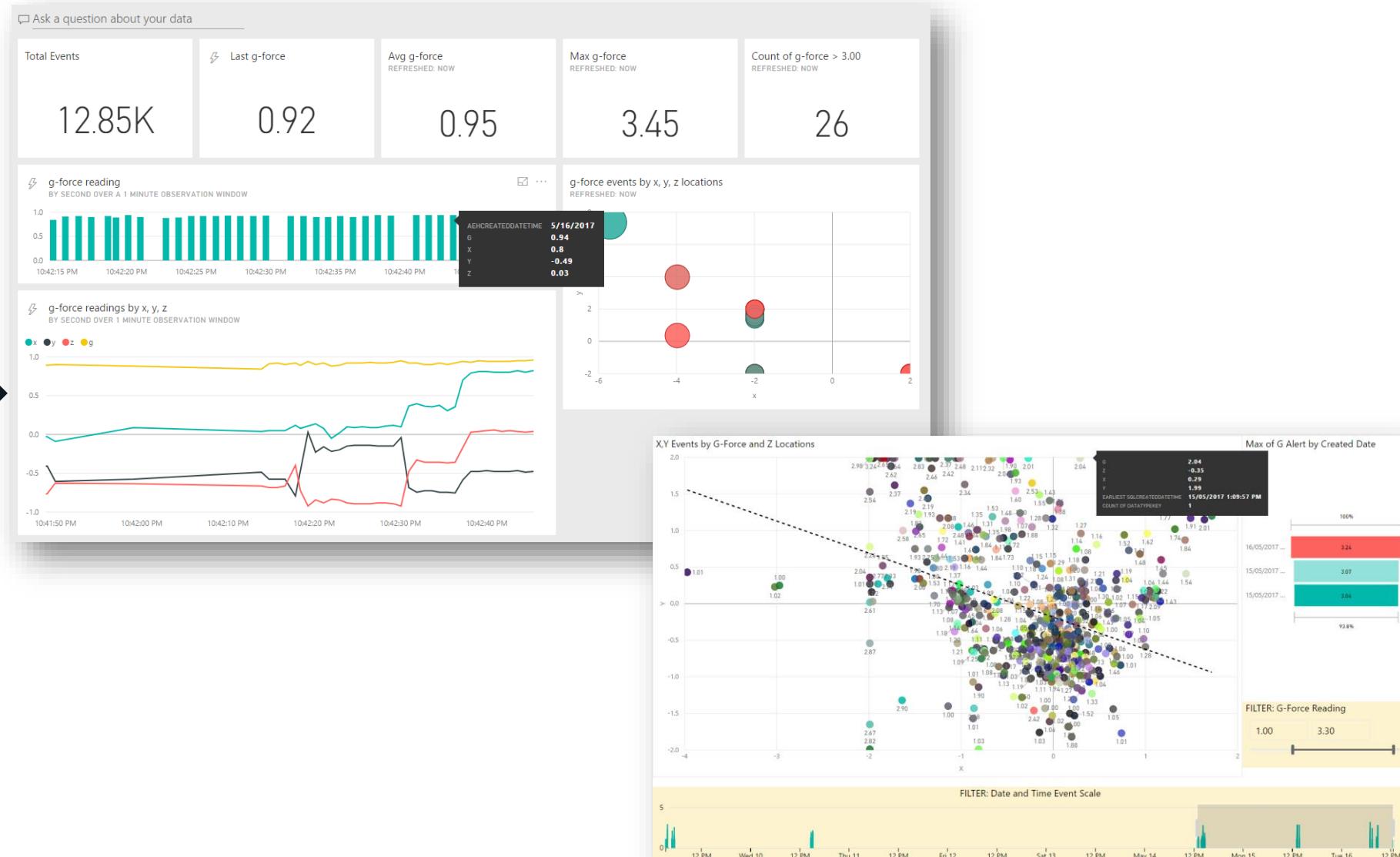
Lets see some Azure pipelines!

Demos / Examples:

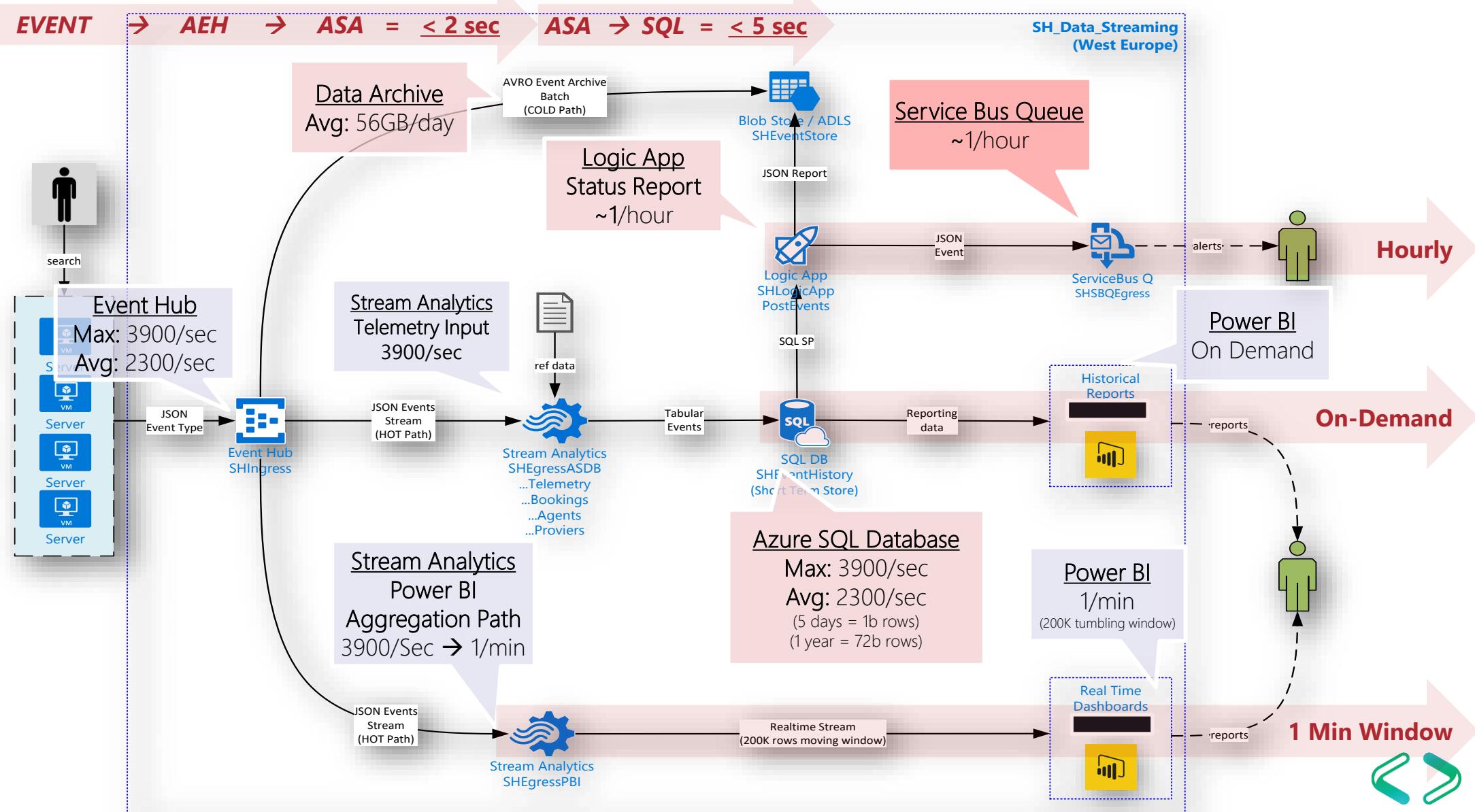
Demonstration → Mobile G-Force Solution - !



Demonstration → Mobile G-Force Solution - !



Other Examples → High Scale Web Search Telemetry



Web Search Telemetry – Total Events (By Day)

Plot

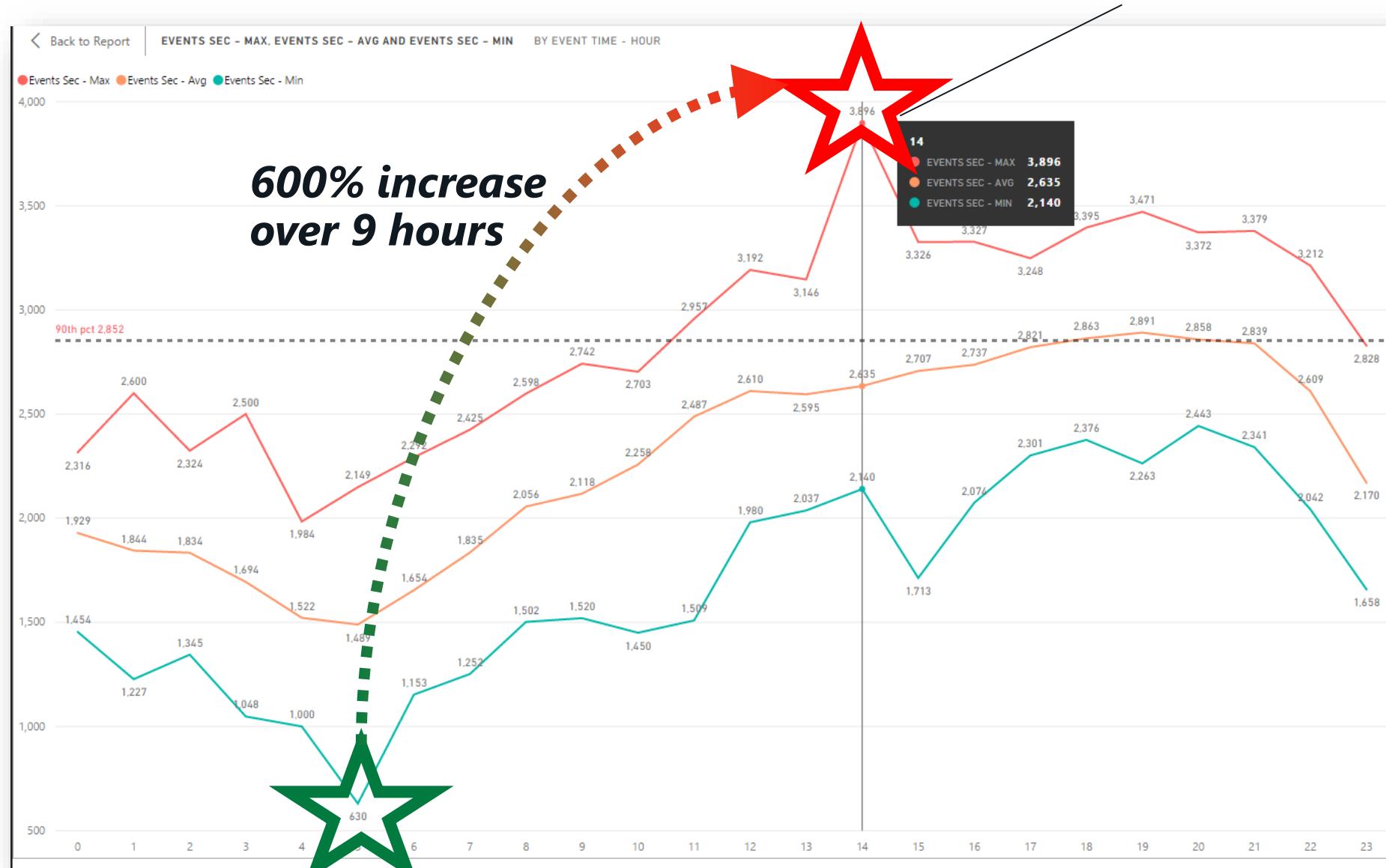


INCOMING MESSAGES
1.41 B

AVG Workload → 1,410,000,000 / week
= 201,000,000 / day
= 8,392,000 / hour
= 139,000 / min
= **2,330 / sec**



Web Search Telemetry – Events/Sec (By Hour)



When is scale an issue?

- What do you mean by "scale"?

IoT Device – Streaming Telemetry Workload:

29,000 / sec → 2,505,600,000 / day → 914,544,000,000 / year

- Lambda principals still apply!!
 - **Ingestion** → handling the "peak" rate without latency/delay/error
 - **Processing/Speed** → need data granularity, or are aggregate windows OK
 - **Storage/Batch** → need adhoc on-demand data engineering, or recurring
 - **Serving** → what granularity is important, what decisions will be made
- **Question:** can you pre-process at "**the edge**"?





So where to from here?
Wrap up and summary

What's next for the data platform? *...and what does this mean for us Data Professionals?*

1. **On-prem** hosted/deployed data solutions are **diminishing**
2. **Public cloud** data ecosystem is **mature** and **expanding**
3. **IaaS is popular, PaaS is King** (ie *Serverless world is the future*)
4. **Customer “expectation”...**
...This is the “Domain of the Data Professional”

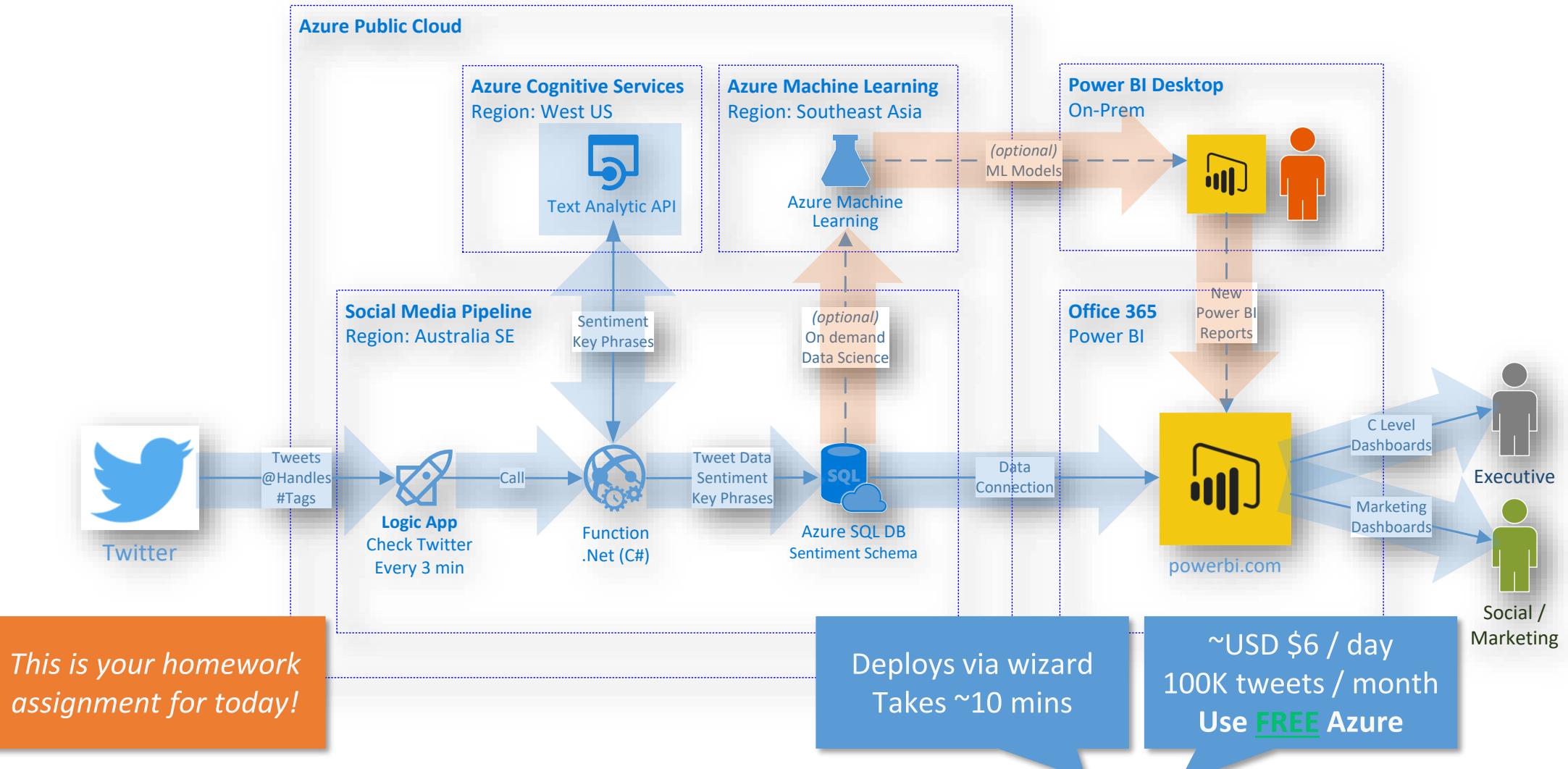


Where can I try this out – or learn more?

- **Cortana Intelligence Gallery Pre-Built Solutions (one-click deploy)**
 - **Vehicle Telemetry**
<https://gallery.cortanaintelligence.com/Solution/Telemetry-Analytics>
 - **Personalised Offers**
<https://gallery.cortanaintelligence.com/Solution/Personalized-Offers-2>
 - **Energy Demand Forecasting**
<https://gallery.cortanaintelligence.com/Solution/Demand-Forecasting-3>
- **EdX Self-Paced Courses (3-4 hrs/week for ~4 weeks)**
 - **Developing IoT Solutions with Azure IoT**
<https://www.edx.org/course/developing-iot-solutions-azure-iot-microsoft-dev225x>
 - **Processing Real-Time Data Streams in Azure**
<https://www.edx.org/course/processing-real-time-data-streams-azure-microsoft-dat223-2x-0>
 - **Orchestrating Big Data with Azure Data Factory**
<https://www.edx.org/course/orchestrating-big-data-azure-data-microsoft-dat223-3x-0>



Your Homework → Twitter Social Media Analytics



<https://powerbi.microsoft.com/en-us/solution-templates/brand-management-twitter/>





[End of Presentation]



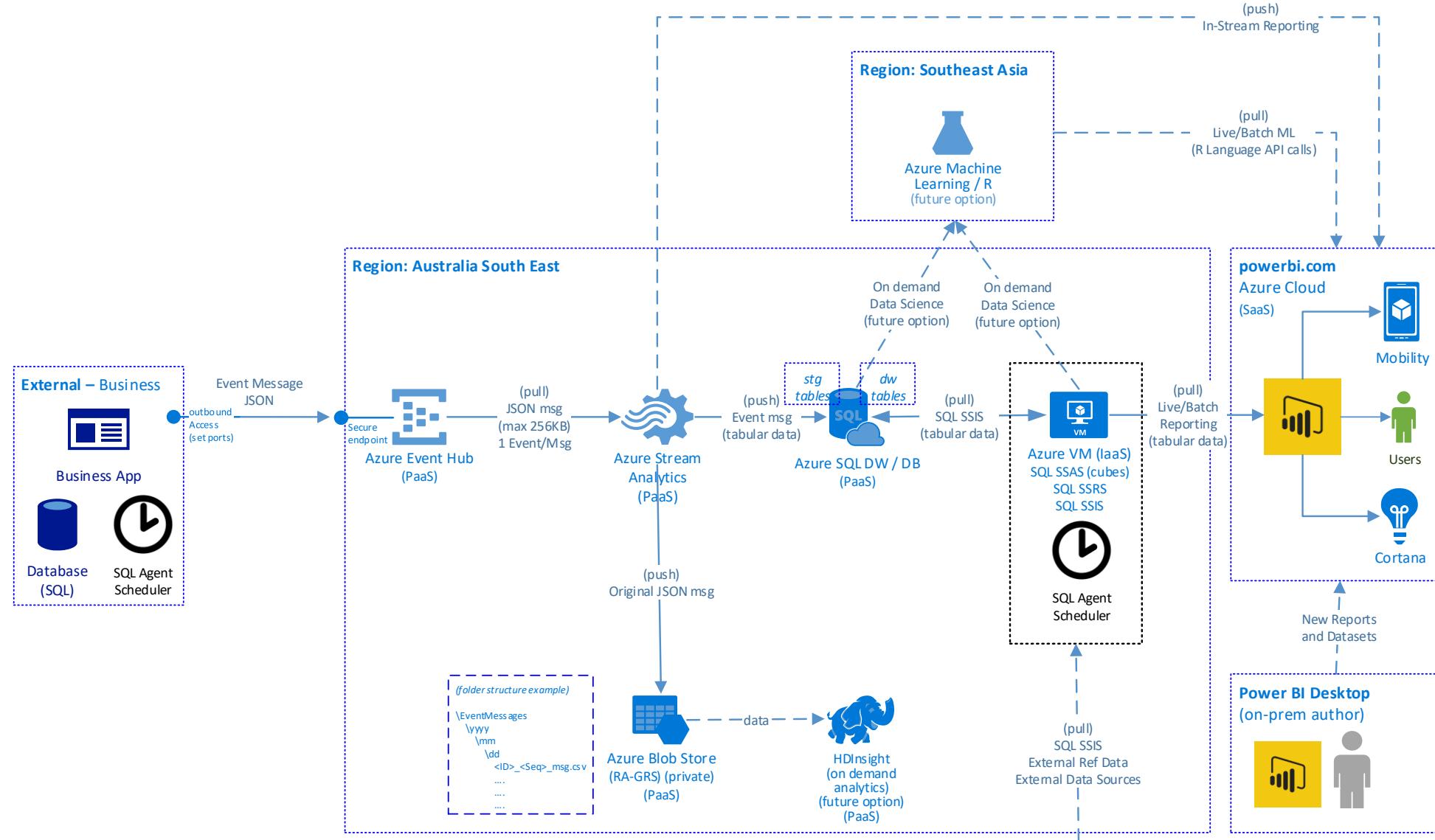
Appendix

Appendix and References

- Online Azure Interactive Services Diagram - <http://azureplatform.azurewebsites.net/en-us/>
- Azure Time Series Insights Service - <https://azure.microsoft.com/en-au/blog/announcing-azure-time-series-insights/>
- Service Bus Explorer – Github - <https://code.msdn.microsoft.com/windowsapps/Service-Bus-Explorer-f2abca5a>
- Cortana Intelligence Gallery - <https://gallery.cortanaintelligence.com/>
- Predictive Maintenance - <https://docs.microsoft.com/en-us/azure/machine-learning/cortana-analytics-playbook-predictive-maintenance>
- Anomaly Detection - <https://docs.microsoft.com/en-us/azure/machine-learning/machine-learning-apps-anomaly-detection-api>
- EdX - Developing IoT Solutions with Azure - <https://www.edx.org/course/developing-iot-solutions-azure-iot-microsoft-dev225x>
- EdX - Processing Real-Time Data in Azure - <https://www.edx.org/course/processing-real-time-data-streams-azure-microsoft-dat223-2x-0>
- EdX - Orchestrating Big Data with ADF - <https://www.edx.org/course/orchestrating-big-data-azure-data-microsoft-dat223-3x-0>
- Microsoft Lambda Architecture - <https://social.technet.microsoft.com/wiki/contents/articles/33626.lambda-architecture-implementation-using-microsoft-azure.aspx>
- Microsoft Lambda Reference Architecture - <https://azure.microsoft.com/en-au/updates/microsoft-azure-iot-reference-architecture-available/>
- Wiki Lambda Architecture - https://en.wikipedia.org/wiki/Lambda_architecture
- Azure Stream Analytics Query Language - <https://msdn.microsoft.com/en-us/library/azure/dn834998.aspx>
- Azure Stream Analytics Query Windowing Functions - <https://msdn.microsoft.com/en-us/library/azure/dn835019.aspx>
- Azure Stream Analytics Query Patterns - <https://docs.microsoft.com/en-us/azure/stream-analytics/stream-analytics-stream-analytics-query-patterns>
- Azure Storage Explorer - <http://storageexplorer.com/>
- Blog – Making Phone Calls from Azure - <https://mrfoxsql.wordpress.com/2017/05/31/making-phone-calls-from-azure-event-hub-messages/>
- Blog – What is the Data Platform Today - <https://mrfoxsql.wordpress.com/2017/04/19/what-exactly-is-the-data-platform-nowadays/>



Other Examples → Business Incident Management



Where can I find even more examples of this stuff?

Cortana Intelligence Gallery

Browse all Industries Solutions Experiments Machine Learning APIs Custom Modules Learning More

Refine by

CATEGORIES

- Solution
- Experiment
- Machine Learning API
- Competition
- Tutorial
- Collection
- Notebook
- Custom Module
- Classroom Training
- Video Training

SHOW

- Microsoft content only

TAGS

- R
- Python
- Classification
- Microsoft R Server
- Linear Regression

Show all

AI ALGORITHMS USED

Results

You've selected: Solution

Sort by: New and Noteworthy

Image	Name	Description	Views	Created	Provider
	Personalized Offers	In today's highly competitive and connected environment, modern businesses can no longer survive with generic, static online content...	1066	11 days ago	Microsoft
	Demand Forecasting and P...	Pricing is recognized as a pivotal determinant of success in many industries and can be one of the most challenging tasks. Companies o...	1772	12 days ago	Microsoft
	Campaign Optimization wit...	This solution demonstrates how to build and deploy a machine learning model with SQL Server 2016 with R Services to recommend a...	6753	21 days ago	Microsoft
	Campaign Optimization wit...	This solution demonstrates how to build and deploy a machine learning model with Microsoft R Server on Azure HDInsight Spark clu...	257	36 days ago	Microsoft
	Quality Assurance	Quality assurance systems allow businesses to prevent defects throughout their processes of delivering goods or services to customers...	1116	11 days ago	Microsoft

[https://gallery.cortanaintelligence.com/browse?categories=\[\"10\"\]&orderby=freshness desc](https://gallery.cortanaintelligence.com/browse?categories=[\)

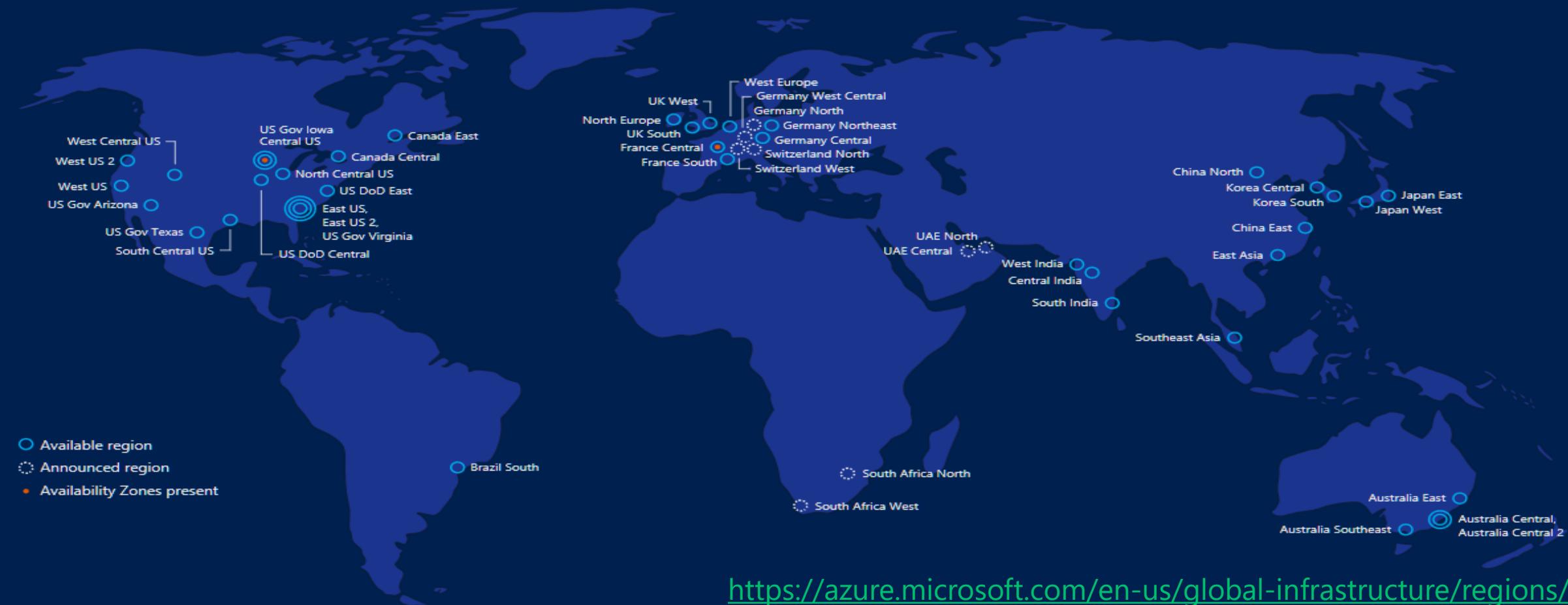


Azure regions

Azure has more global regions than any other cloud provider—offering the scale needed to bring applications closer to users around the world, preserving data residency, and offering comprehensive compliance and resiliency options for customers.

50 regions worldwide

140 available in 140 countries



Microsoft Azure Data Services

IaaS

SQL Server, Oracle,
Mongo, MySQL, +
in Azure VM



PaaS

Azure Relational
Databases
SQL, PostgreSQL, MySQL



Azure NoSQL
Databases
CosmosDB



Azure Table
Storage



Azure Blob
Storage



fully featured RDBMS

transactional processing

rich query

managed as a service

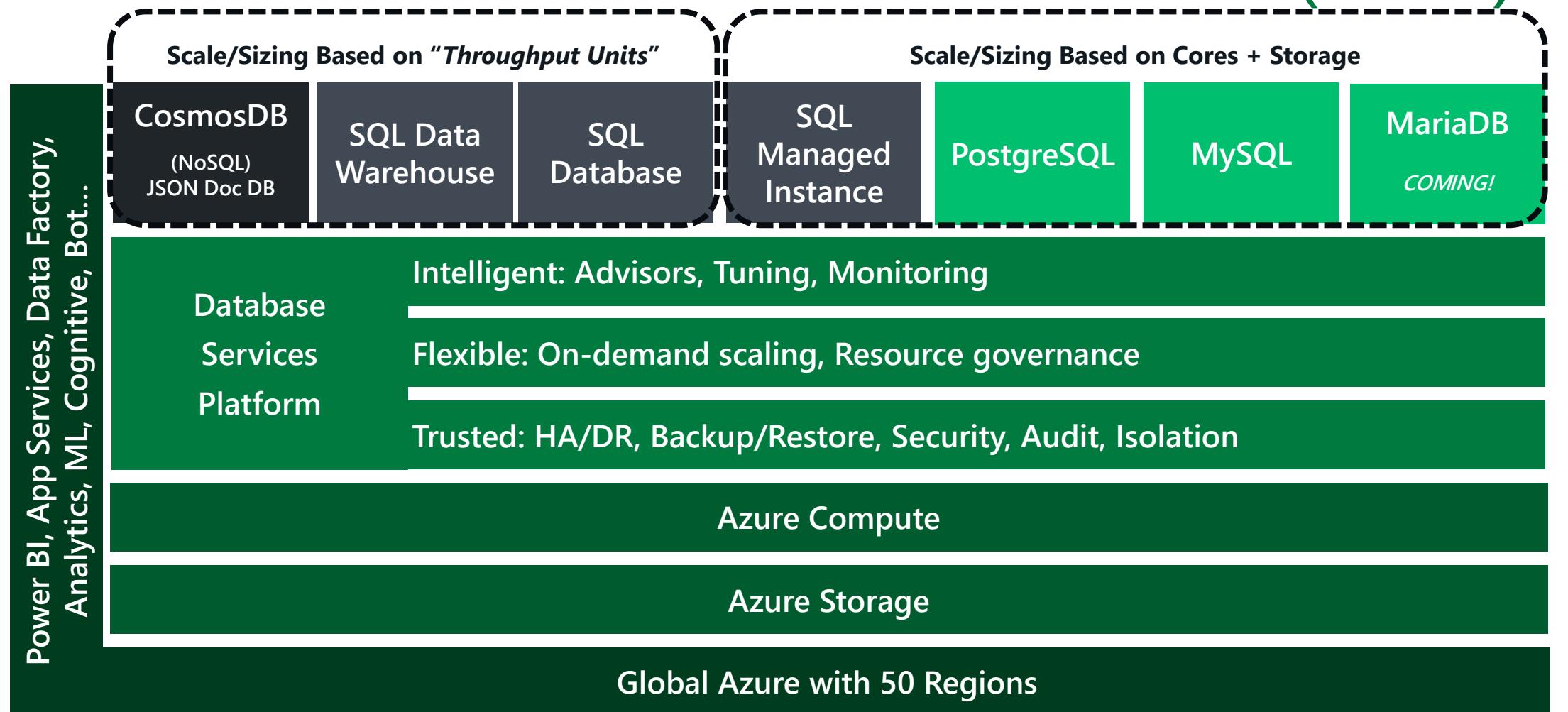
elastic scale

schema-free data model

Internet accessible http/rest

arbitrary data formats

Azure Relational Database Platform (PaaS)



<https://docs.microsoft.com/en-us/azure/architecture/guide/technology-choices/data-store-overview>

<https://docs.microsoft.com/en-us/azure/architecture/data-guide/technology-choices/analytical-data-stores>



Get Azure...?

Your Enterprise Agreement (EA)

Various options – Currently being setup and configured for MLC
Would be linked to your corporate identity/login/account

Azure 30 day free account up to \$260 (time boxed to 30 days)

<https://azure.microsoft.com/en-au/free/>

Would be linked to your personal identity/login/account

MSDN

There are free monthly Azure credits within MSDN subscriptions. Rolls over month to month
<https://azure.microsoft.com/en-us/pricing/member-offers/msdn-benefits-details/>

Would be linked to your corporate identity/login/account

Research Programs and Grants

Free credits available for specific research programs = <https://www.microsoft.com/en-us/research/academic-program/microsoft-azure-for-research/>

You can apply for a Microsoft Azure for Research Grant = <http://research.microsoft.com/en-us/projects/azure/default.aspx>



Learn Azure...? Free Online-Training

edX – Free online courses on Microsoft Azure

45+ Free Azure Courses = https://www.edx.org/course?search_query=azure

Introduction to Azure - <https://www.edx.org/course/introduction-microsoft-azure-microsoft-azure201x>

Architecting Azure Solutions - <https://www.edx.org/course/architecting-microsoft-azure-solutions-microsoft-dev205bx-3>

Developing Azure Solutions - <https://www.edx.org/course/developing-microsoft-azure-solutions-microsoft-dev233-1>

Developing Apps and Bots - <https://www.edx.org/course/developing-intelligent-apps-bots-microsoft-dat211x-1>

Deliver a DW in the Cloud - <https://www.edx.org/course/delivering-data-warehouse-cloud-microsoft-dat220x-0>

Delivery Big Data Solutions with Machine Learning - <https://www.edx.org/course/developing-big-data-solutions-azure-microsoft-dat228x>

Provision SQL Databases in Azure - <https://www.edx.org/course/provisioning-databases-azure-sql-server-microsoft-dat219x-0>



Learn Azure...? Free 1x day In-Person Training

<https://azure.microsoft.com/en-us/community/events/?Country=Australia&query=azure+discovery+day>



Reserve your spot. Accelerate your learning.

Microsoft Azure is a growing collection of integrated cloud services – analytics, computing, database, storage, mobile, networking, the web – to help you move faster, work smarter and achieve more.

Its potential is unlimited if you know how to unlock it. Which is why we offer Azure Discovery Days; free sessions led by Microsoft Azure experts and built to help you extract extra value out of Azure.

Think big, then delve deeper

Each Discovery Day offers an Azure overview, plus primers on security, billing, identity, storage, networking and compute. We'll also cover these core topics.

- Administer Azure via the Azure Portal & PowerShell
- Create & configure VMs in Azure to connect securely to On-Premises
- Monitoring & Recovery Services
- Monitoring and operational Insights
- Automation and Azure Quick Start Templates
- MyApps portal & Self Service capabilities
- Multi Factor Authentication
- Protecting, encrypting, and tracking documents

Pick your city. Start your journey.



Learn Azure...? Patterns, Blogs and Feedback

Need Azure Patterns and Guidance? Check out the Azure Architecture Centre - !

- <https://docs.microsoft.com/en-au/azure/architecture/>
 - Reference Architectures - <https://docs.microsoft.com/en-us/azure/architecture/reference-architectures/>
 - Application Architectures - <https://docs.microsoft.com/en-us/azure/architecture/guide/>
 - Azure Design Patterns - <https://docs.microsoft.com/en-au/azure/architecture/patterns/>
 - Azure Service Roadmap - <https://azure.microsoft.com/en-us/roadmap/>
 - Data Architecture Guide - <https://docs.microsoft.com/en-us/azure/architecture/data-guide/>

Need Some Azure Updates? Subscribe to the Global Azure Blog & Update Feed - !

- Blog - <https://azure.microsoft.com/en-us/blog/>
- **Updates** - <https://azure.microsoft.com/en-us/updates/>

You Have a Cool New Azure Idea? Submit it to Azure Ideas & Feedback - !

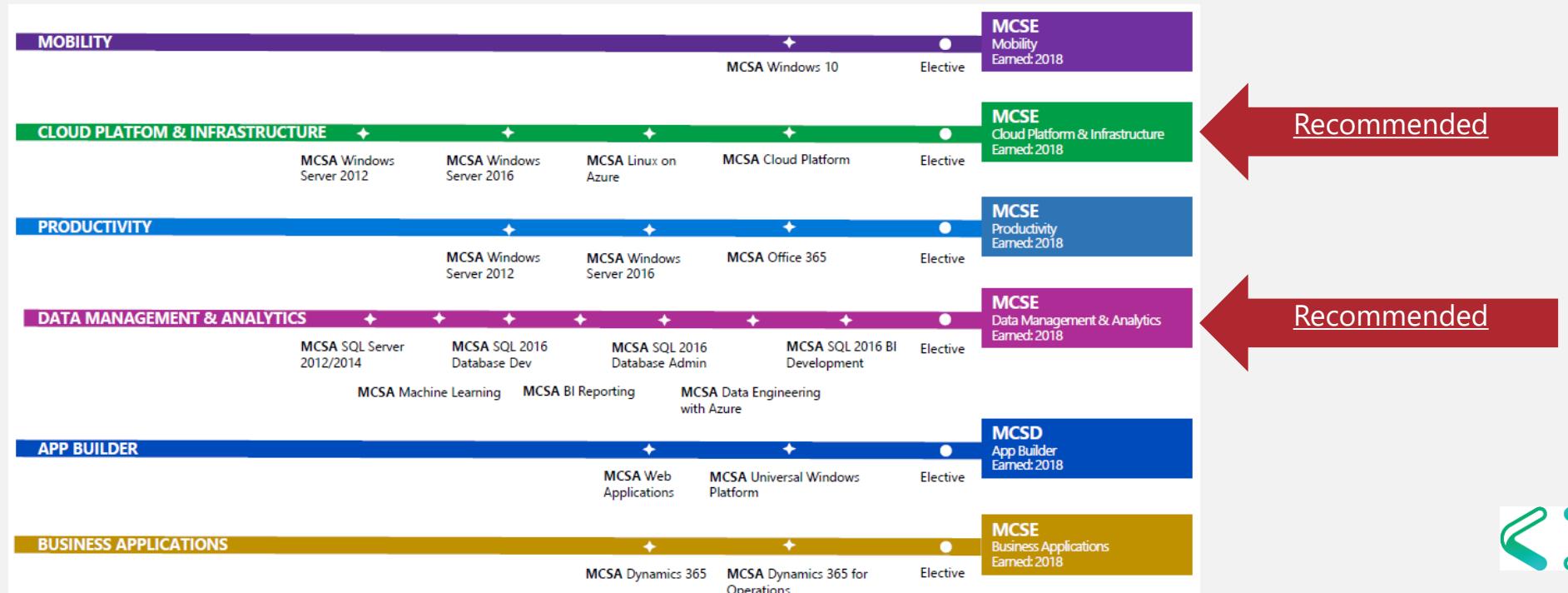
- <https://feedback.azure.com/forums/34192--general-feedback>



Certify in Azure...?

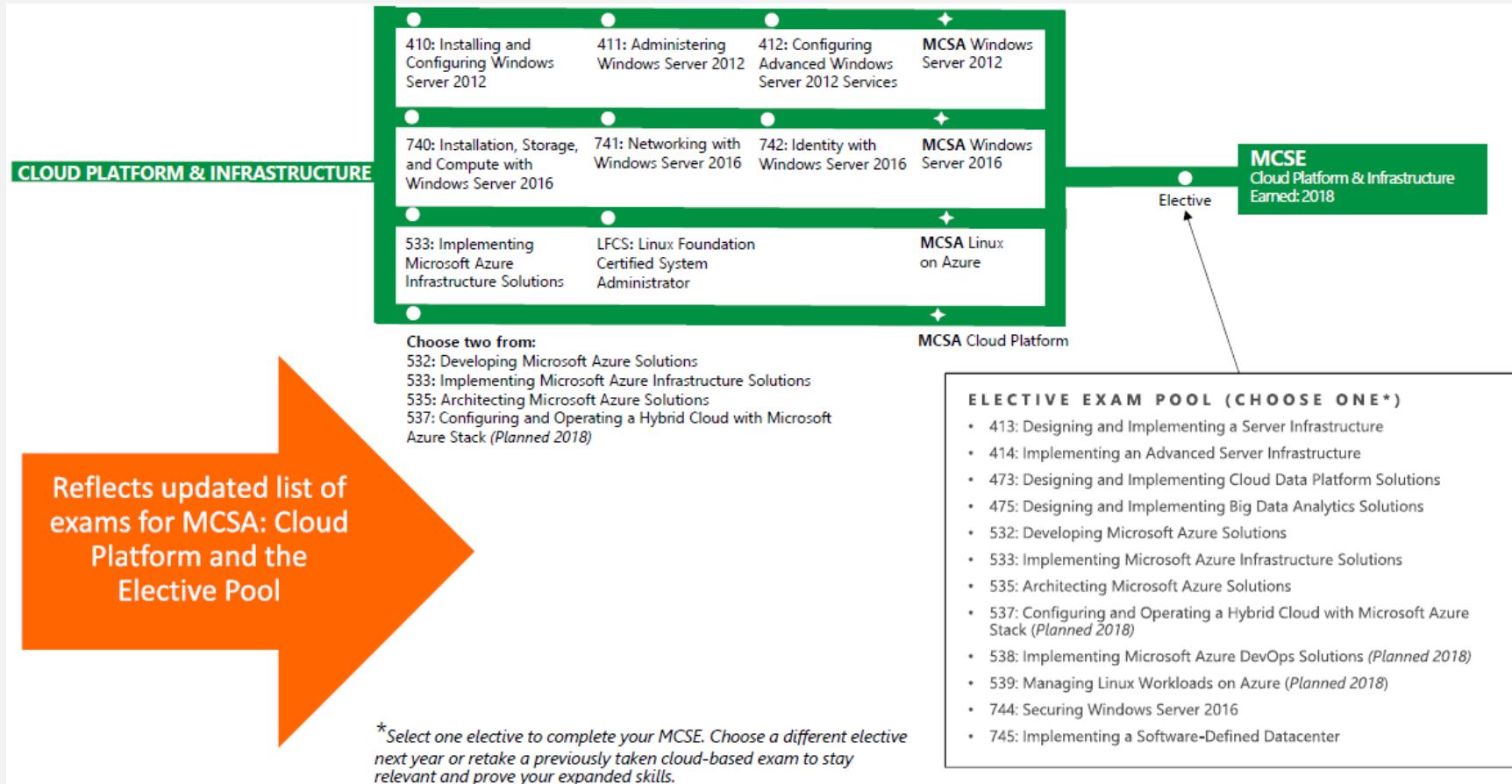
Azure Certifications

- Overview - <https://www.microsoft.com/en-au/learning/certification-overview.aspx>
- **Detailed Guide** - <https://download.microsoft.com/download/2/B/F/2BFA3E78-546B-4C71-9EC3-2CB7751444BF/MCP Cert Paths 01 01 18.pdf>



Certify in Azure...?

MCSE Cloud Platform & Infrastructure



Certify in Azure...?

MCSE Data Management & Analytics

