



# Microsoft AI

Transforming businesses with artificial intelligence

# Microsoft Cloud

Office 365



Dynamics

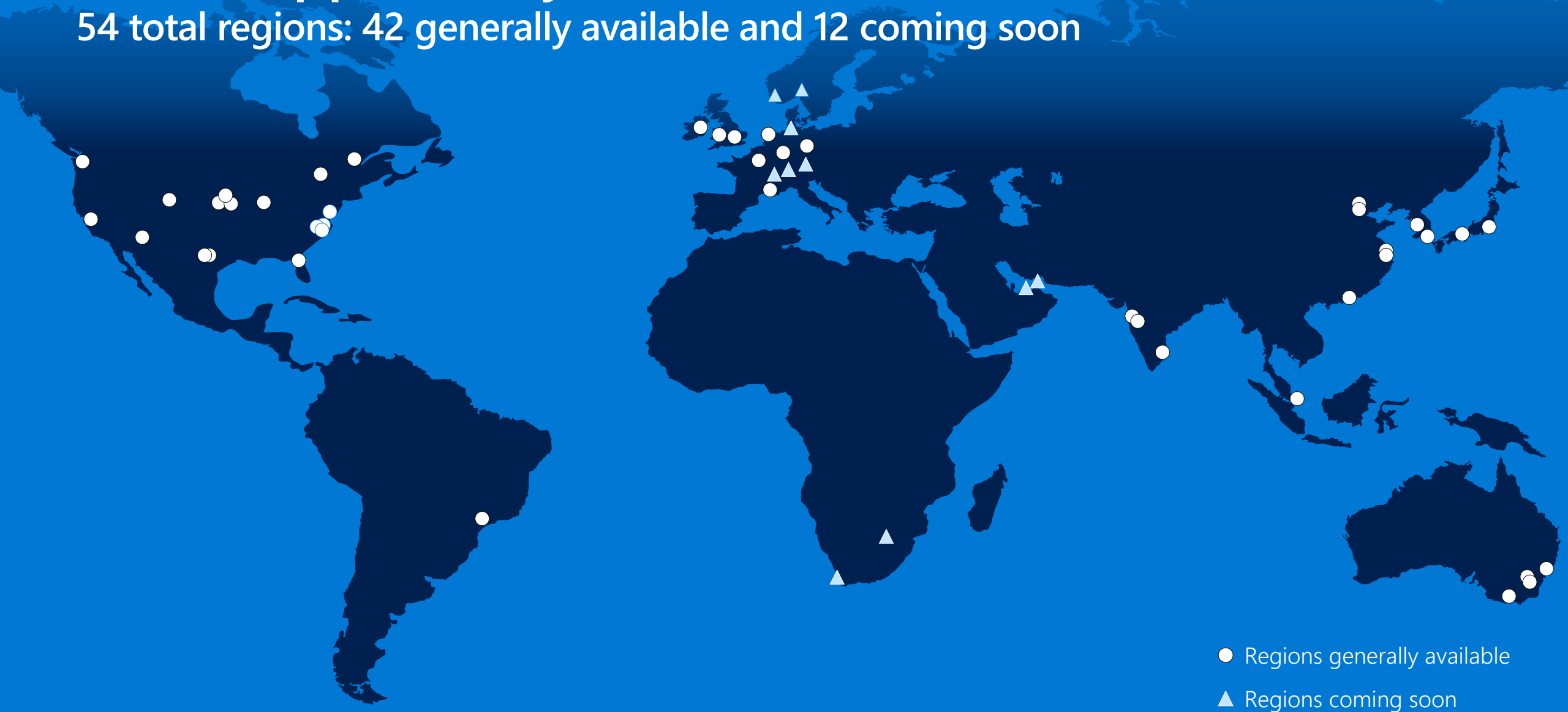


Azure



# Global opportunity with Azure infrastructure

54 total regions: 42 generally available and 12 coming soon







San Antonio



# Virginia (2017)















# Amsterdam



## Security & Management

-  Security Center
-  Portal
-  Azure Active Directory
-  Azure AD B2C
-  Multi-Factor Authentication
-  Automation
-  Scheduler
-  Key Vault
-  Store/Marketplace
-  VM Image Gallery & VM Depot






## Media & CDN

-  Media Services
-  Media Analytics
-  Content Delivery Network








## Integration

-  API Management
-  BizTalk Services
-  Logic Apps
-  Service Bus







## Compute Services

-  Container Service
-  VM Scale Sets
-  Batch
-  RemoteApp
-  Dev/Test Lab

## Application Platform


-  Web Apps
-  Mobile Apps
-  API Apps
-  Cloud Services
-  Service Fabric
-  Notification Hubs
-  Functions

## Developer Services

-  Visual Studio
-  Mobile Engagement
-  VS Team Services
-  Xamarin
-  Application Insights
-  HockeyApp

## Platform Services











## Data

-  SQL Database
-  SQL Data Warehouse
-  DocumentDB
-  SQL Server Stretch Database
-  Redis Cache
-  Storage Tables
-  Azure Search








## Intelligence

-  Cognitive Services
-  Bot Framework
-  Cortana

## Analytics & IoT

-  HDInsight
-  Machine Learning
-  Stream Analytics
-  Data Catalog
-  Data Lake Analytics Service
-  Data Lake Store
-  IoT Hub
-  Event Hubs
-  Data Factory
-  Power BI Embedded

## Hybrid Cloud

-  Azure AD Health Monitoring
-  AD Privileged Identity Management
-  Domain Services
-  Backup
-  Operational Analytics
-  Import/Export
-  Azure Site Recovery
-  StorSimple

## Infrastructure Services

## Compute

-  Virtual Machines
-  Containers

## Storage

-  Blob
-  Queues
-  Files
-  Disks

## Networking

-  Virtual Network
-  Load Balancer
-  DNS
-  Express Route
-  Traffic Manager
-  VPN Gateway
-  App Gateway

## Datacenter Infrastructure



# Build and innovate with choice and flexibility

## DevOps

**Nagios**



## Management



## Applications



## App frameworks & tools



**nodeJS**



## Databases & middleware



**cloudera**



## Infrastructure





# Microsoft AI Platform

## Azure AI Services

### PRE-BUILT AI

Cognitive Services

### CONVERSATIONAL AI

Bot Service



### CUSTOM AI

Azure Machine Learning

### CODING & MANAGEMENT TOOLS

VS Tools  
for AI

Azure ML  
Studio

Azure ML  
Workbench

Others (PyCharm, Jupyter Notebooks...)



### DEEP LEARNING FRAMEWORKS

3rd Party

Cognitive  
Toolkit

TensorFlow

Caffe

Others (Scikit-learn, MXNet, Keras,  
Chainer, Gluon...)

## Azure Infrastructure

### AI ON DATA

### AI COMPUTE

Cosmos  
DB

SQL  
DB

SQL  
DW

Data  
Lake

Spark

DSVM

Batch  
AI

AKS,  
ACI

IoT  
Edge

101010  
010101  
101010

CPU, FPGA, GPU

# AI infrastructure

## AI on data

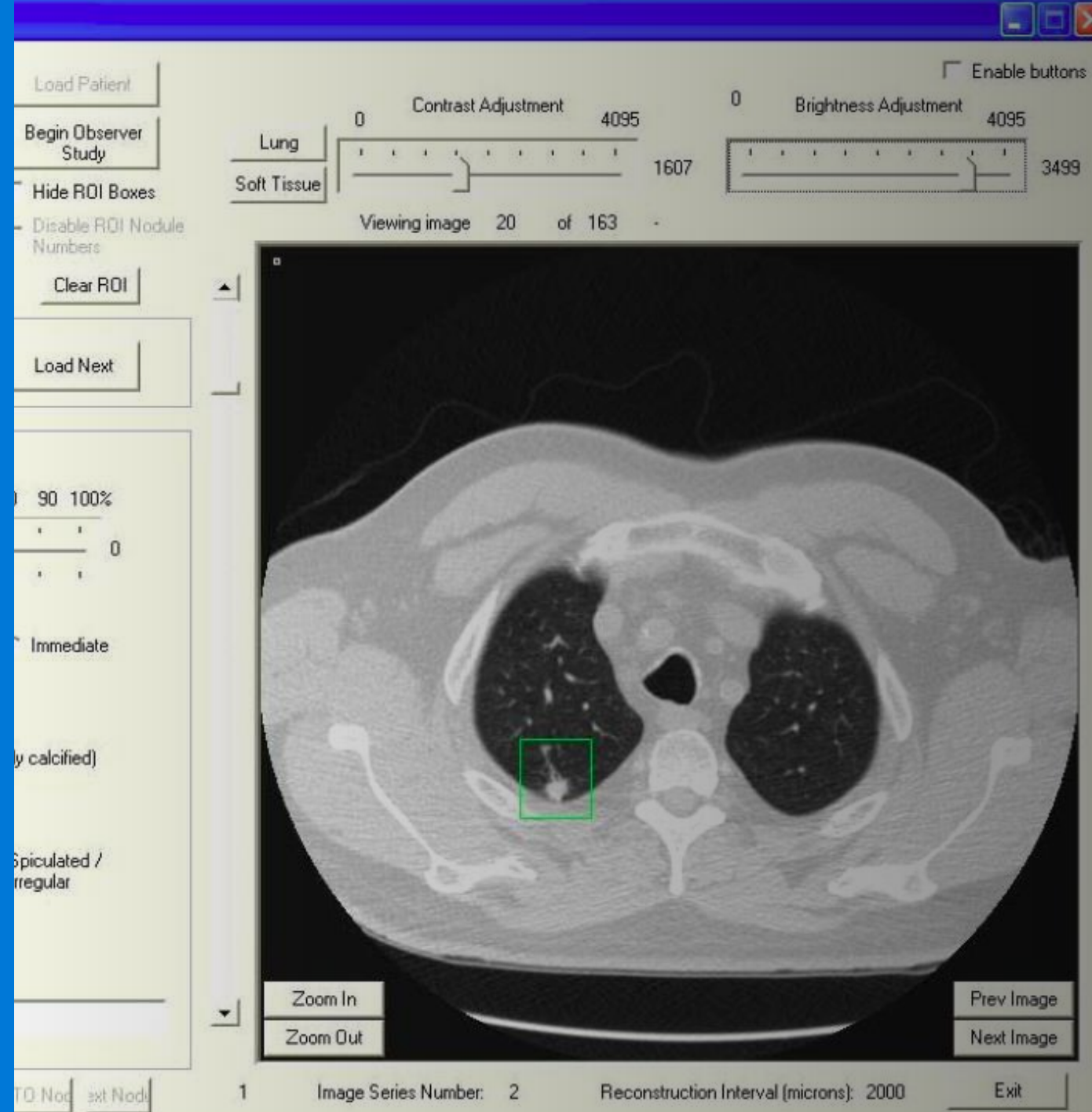
Built-in AI across your data.

## AI compute

Flexible compute infrastructure from virtually infinite cloud scale to the edge .

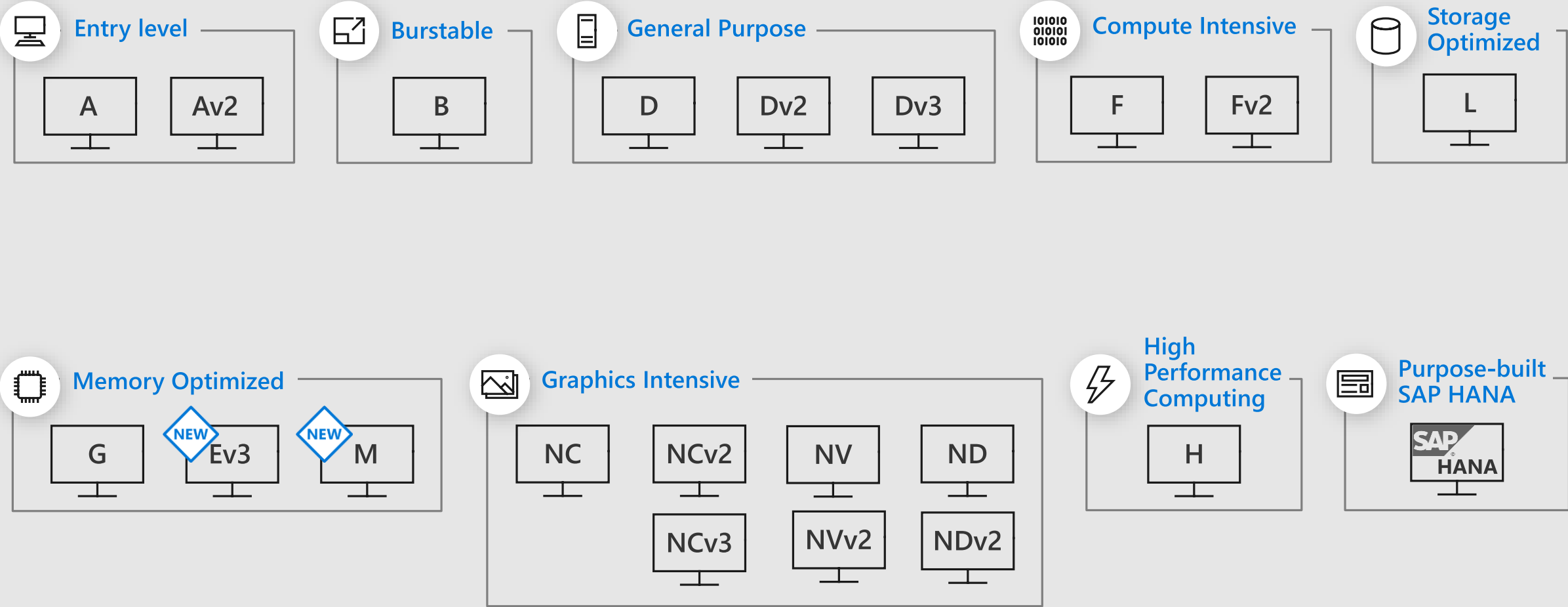
## AI core infrastructure

CPU, GPU, and FPGA.



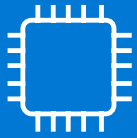
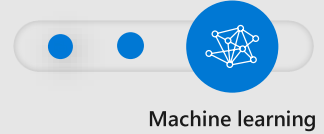


# Compute options for all types of apps



# Powerful infrastructure

Accelerate deep learning

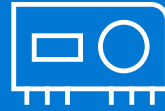


**CPUs**

---

General purpose  
machine learning

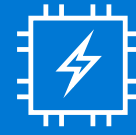
D, F, L, M, H Series



**GPUs**

---

Deep learning  
N Series



**FPGAs**

---

Specialized hardware  
accelerated deep learning

Project Brainwave



Optimized for flexibility



Optimized for performance



D e m o

(Azure VMs, Data Science VM)

# Recipe for Powerful AI

- Lots of experimentation
- Lots of GPUs
- Lots of training data

*Need to run Training at scale!*



# Infrastructure Can Get in Your Way

## Clusters

- Provision GPUs
- Install drivers and software
- Interactive use

## Scheduling

- Queue work
- Prioritize jobs
- Start MPI
- Monitor
- Handle failures

## Data

- Scale access to training data
- Output logs & models
- Secure & compliant

## Cost

- Scale up and down
- Share reserved instances
- Low priority

## Workflow

- Choose efficient hardware
- Tooling integration
- Laptop to cloud

## Azure Batch AI Service

- Managed Service
- Supports Role Based Access Control
- Run any toolkit (CNTK, Tensorflow, Caffee/Caffee2, Chainer, Keras, ...)
- Run experiments in Parallel
- Run in Containers or directly on VM
- Support various Shared File Systems
- Load based automatic scaling
- Only Storage and compute cost. Service is free



# D e m o

(Batch AI)

# Containers in Azure



## App Service

Deploy web apps or APIs using containers in a PaaS environment



## Service Fabric

Modernize .NET applications to microservices using Windows Server containers



## Kubernetes Service

Scale and orchestrate Linux containers using Kubernetes



## Container Instance

Elastically burst from your Azure Kubernetes Service (AKS) cluster



## Ecosystem

Bring your Partner solutions that run great on Azure



Azure Container Registry



Docker Hub

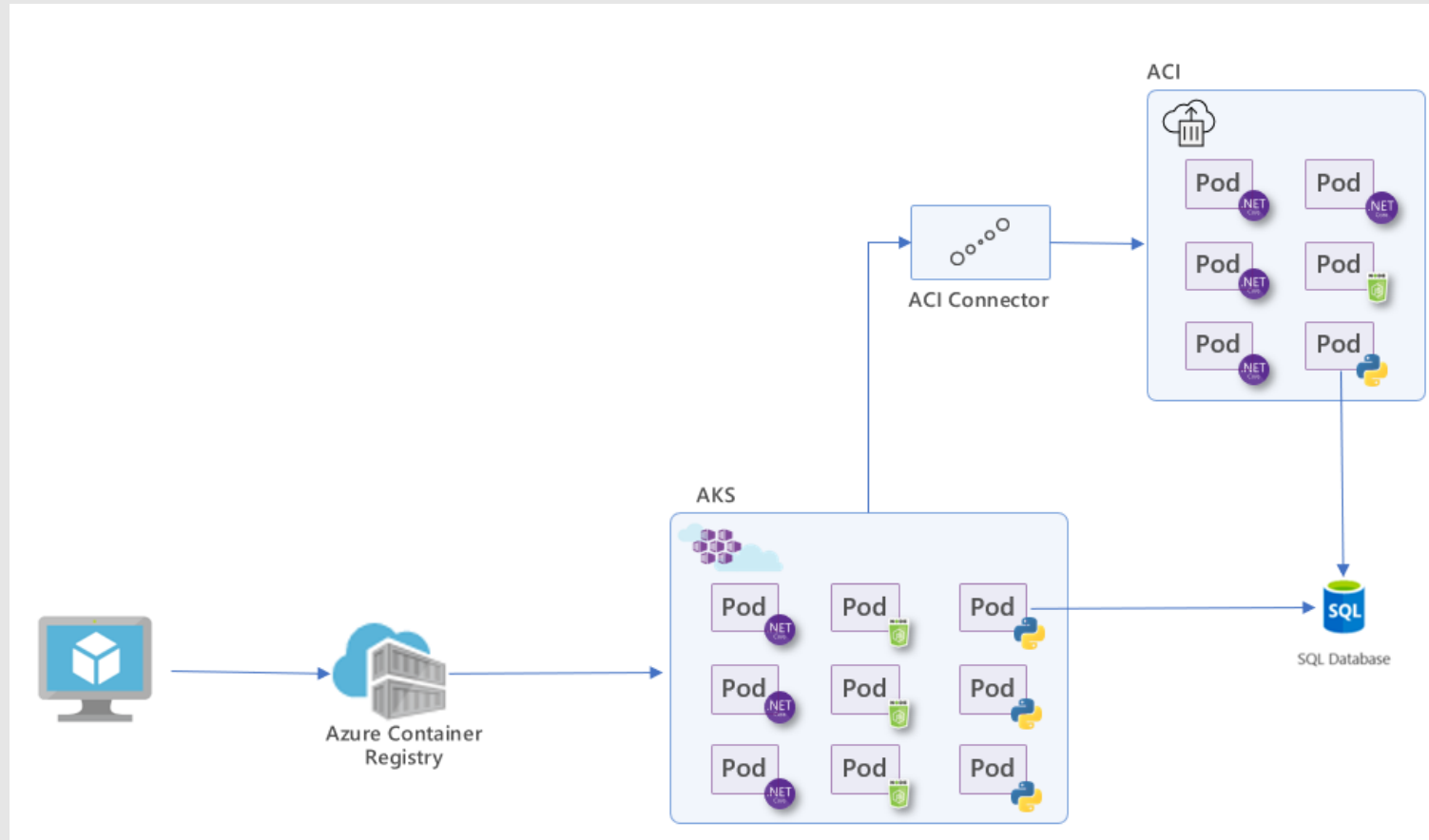
Choice of developer tools and clients

# ACI burst from AKS cluster

Run AKS with just enough capacity for your average workload and scale out in ACI without additional servers to manage. Meet traffic demand spikes by using ACI Connector to provision pods inside ACI that start in seconds.

- **Use the Horizontal Pod Autoscaler** in conjunction with the ACI Connector to easily autoscale applications without worrying about scaling infrastructure
- **Delegate a subnet for ACI** within the VNet already delegated to AKS. This enables customers to load balance services with pods backed by AKS and ACI
- **Connect apps to on-premises data.** With Hybrid Connections and VPN, it's easy to access data securely from on-premises data centers, anywhere in the world
- **Expose Azure Container Instances metrics and logs** together with the health of AKS clusters if customers have the ACI Connector enabled

An implementation of Virtual Kubelet.





# Azure ML Python SDK

[demo](#)

- Azure ML SDK supporting Python 2.7+ and 3.5+
- Runs in any Python IDE or Notebook, including Databricks
- Provision compute targets in Azure.
- Launch training jobs to local, remote VM, Azure Container Instance, Azure Batch AI cluster and HDI Spark cluster.
- Azure File-based shared data store easily accessible by all compute targets.
- Intelligent & efficient hyper-parameter sweep with policy-based early stopping.
- Query and visualize run history directly in Notebook interactive session.
- Register and deploy models

```
$ pip install azureml-sdk
```

```
import azureml.core
import azureml.train

from azureml.core.compute_target
    import BatchAITarget
# Create a configuration of a BatchAITarget.
# This defines a 16 node cluster of 4 GPU VMs
batchai_compute_target =
    BatchAITarget(project=cur_project,
                  node_count=16,
                  vm_size="Standard_NC24r")

from azureml.train.dnn import TensorFlow
# define a tensorflow job with 4 nodes
# (=16 GPUs) and 1 parameter server
resnet_estimator =
    TensorFlow(project=cur_project,
              entry_script='resnet50.py',
              node_count=4,
              ps_count=1)

# kick off the job (this will create the
# cluster for the job and take it down afterwards)
resnet_estimator.fit()
```

# Azure Storage

## Blobs

"Highly scalable, REST-based cloud object store"

Hot, Cool and Archive Tier

## Tables

"Massive auto scaling NoSQL store"

## Data Lake Store

"Highly scalable store supports HDFS and POSIX File System"

## Premium disks

SSD based, high IOPS, low latency  
Up to 64 TB, 80000 IOPs

## Standard disks

Magnetic disks, HDD based.  
Up to 500 IOPs/disk

## Files

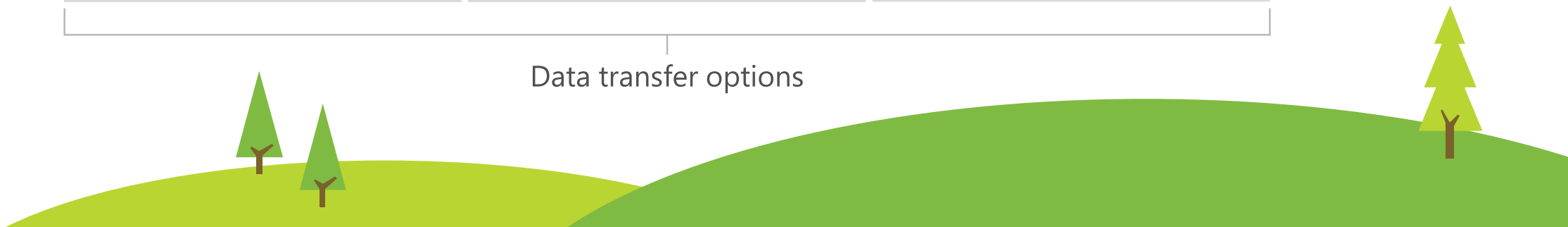
"SMB File Share"  
SMB 3.0  
Read/Write multiple VMs

Azure Import/Export

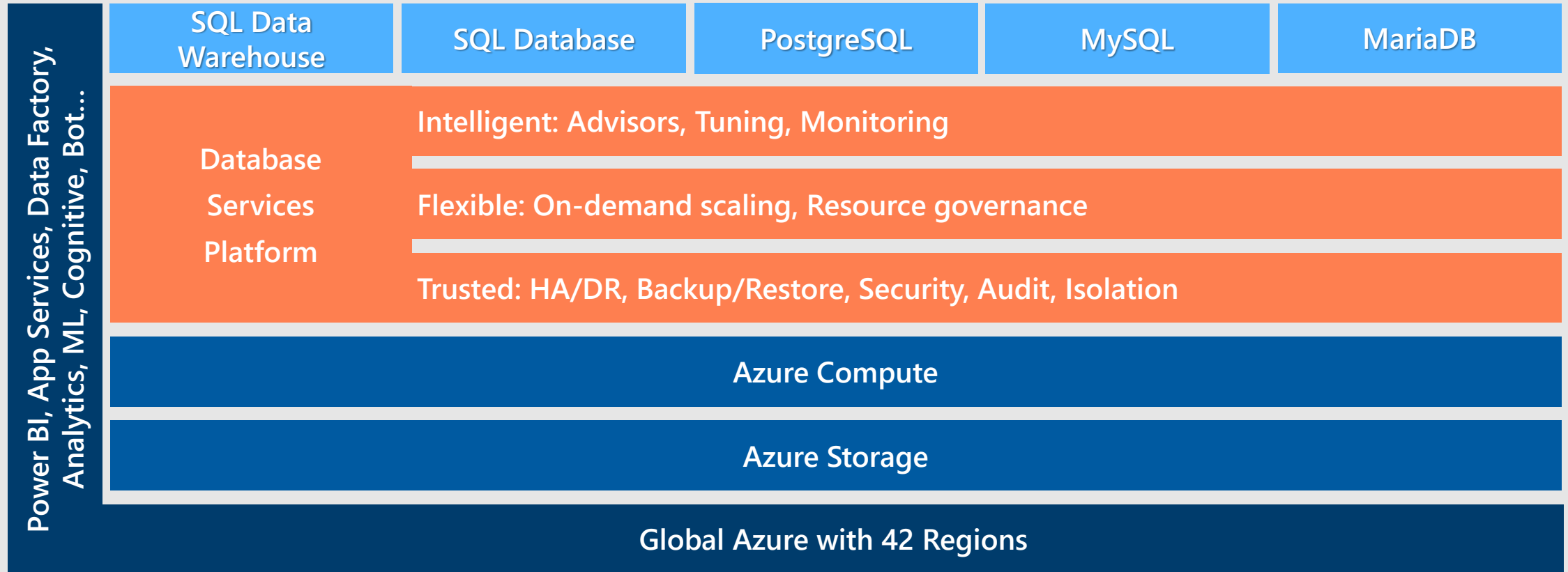
AZCopy

ExpressRoute

Data transfer options

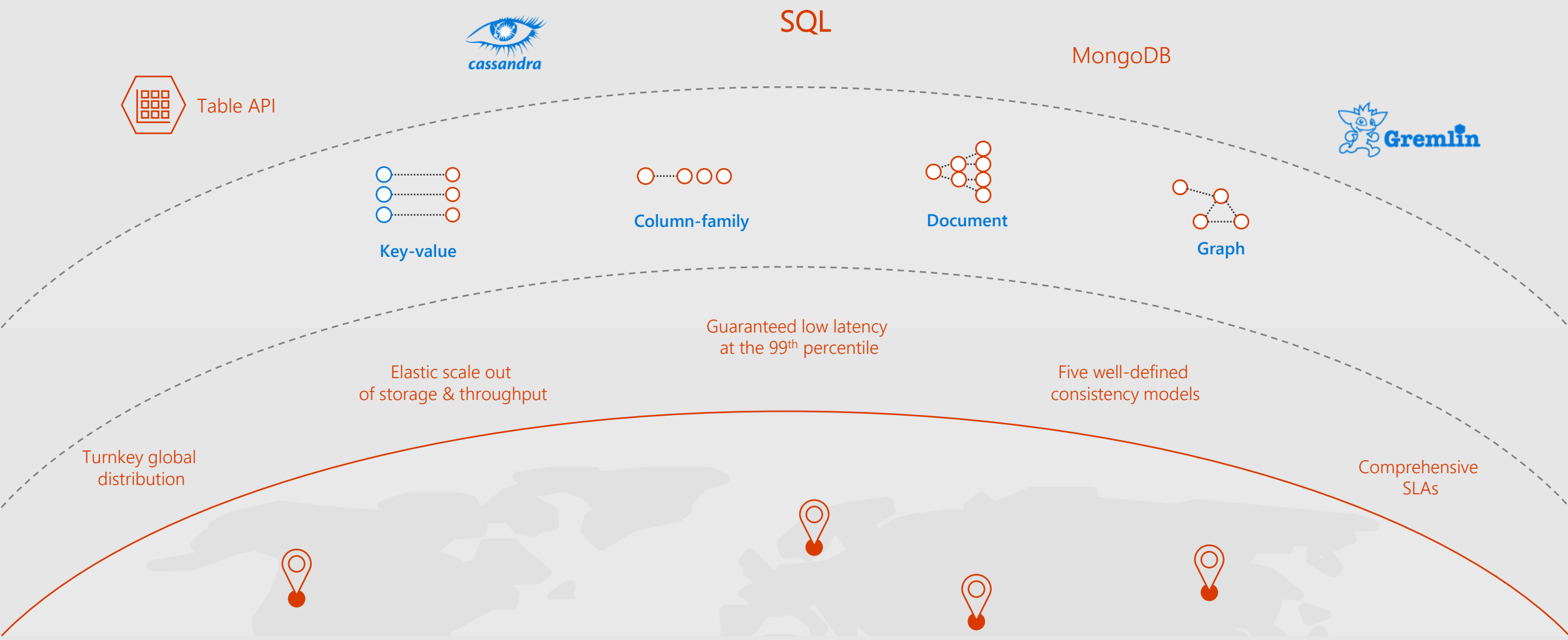


# Azure Relational Databases





# Cosmos DB (No SQL)



# Azure AI services

## Bot Service

Accelerated development for conversational AI.

## Cognitive Services

Vision, speech, language, knowledge and search pre-trained services customizable for any scenario.

## Azure Machine Learning

Experimentation and management services for creating AI models with productivity.



# Microsoft Cognitive Services

## Leverage API's in any language



### Vision

From faces to feelings, allow your apps to understand images and video



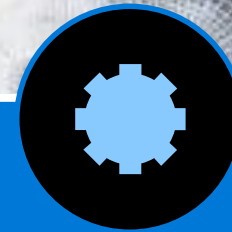
### Speech

Hear and speak to your users by filtering noise, identifying speakers, and understanding intent



### Language

Process text and learn how to recognize what users want



### Knowledge

Tap into rich knowledge amassed from the web, academia, or your own data



### Search

Access billions of web pages, images, videos, and news with the power of Bing APIs



# Microsoft Cognitive Services



## Vision

Computer Vision  
Content Moderator  
Emotion  
Face  
Video  
Video Indexer

Custom Vision  
Service



## Speech

Bing Speech  
Speaker  
Recognition

Custom Speech  
Service



## Language

Bing Spell Check  
Linguistic Analysis  
Text Analytics  
Translator Text  
& Speech  
Web Language  
Model

Language  
Understanding



## Knowledge

Academic  
Knowledge  
Entity Linking  
Knowledge  
Exploration  
Recommendations  
QnA Maker

Custom Decision  
Service



## Search

Bing  
Autosuggest  
Bing Image  
Search  
Bing News  
Search  
Bing Video  
Search  
Bing Web Search  
Bing Entity  
Search  
Bing Custom  
Search

# D e m o

(<http://jfkfiles2.azurewebsites.net/>)

(<https://www.how-old.net/>)

# Demo – The JFK Files

[JFKFiles](#)



Photo Credit: U.S. Embassy New Delhi



# Tools

## Visual Studio Tools for AI

Boost productivity with code-centric AI development and Azure integration.

## Azure Machine Learning Workbench

Full lifecycle support for AI and data wrangling productivity.

## Azure Machine Learning Studio

Drag and drop machine learning development for any skillset.

## Open deep learning framework support

Full support for Cognitive Toolkit, TensorFlow, Caffe and others.

Open standard for deep learning (ONNX).

Studio Gallery

## Binary Classification: Direct marketing

