Future Technologies in 2030 and beyond
A biased Perspective

Eric Pelletier
Technology Solution Professional

Lukasz Miroslaw, Ph.D.
HPC and Big Compute Switzerland

Microsoft
http://microsoft.com/hpc

Microsoft

eWIS-FutureTech, WMO, Geneva, March 2019

Predictions are (quite often) biased

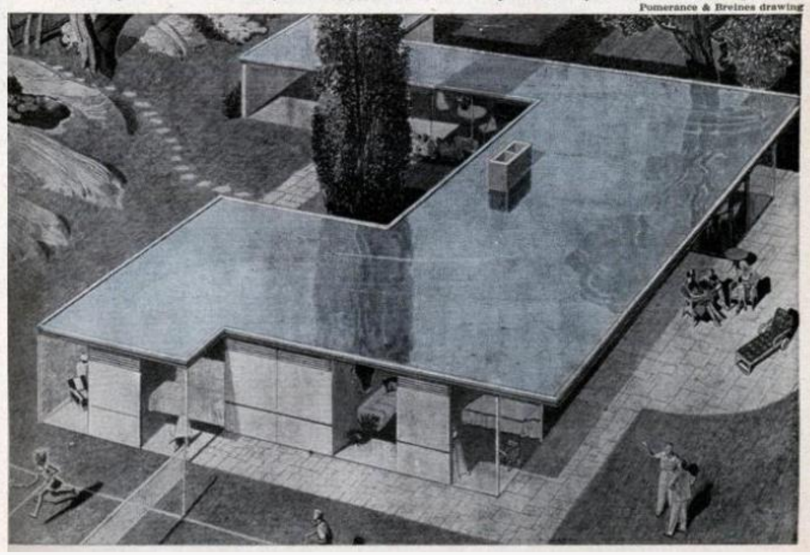
Because everything in her home is waterproof, the housewife of 2000 can do her daily cleaning with a hose

No more bouts with the razor for man of tomorrow. He'll whisk away whiskers with a chemical solution

Housewives in 50 years may wash dirty dishes—right down the drain! Cheap plastic would melt in hot water

Inches-deep lake on roof already cools Southern homes and may become important air-conditioning method

Pomerance & Breines drawing



(…) before a hurricane has a chance to gather much strength and speed, **oil is spread over the sea and ignited**. There is an updraft. Air from the surrounding region, which includes the developing hurricane, rushes in to fill the void. The rising air condenses so that some of the water in the whirling mass falls as rain.

Miracles You'll see in next 50 years.
Popular Mechanics, Feb. 1950

Technical Advances come (often) unexpected

ImageNet Classification with Deep Convolutional Neural Networks

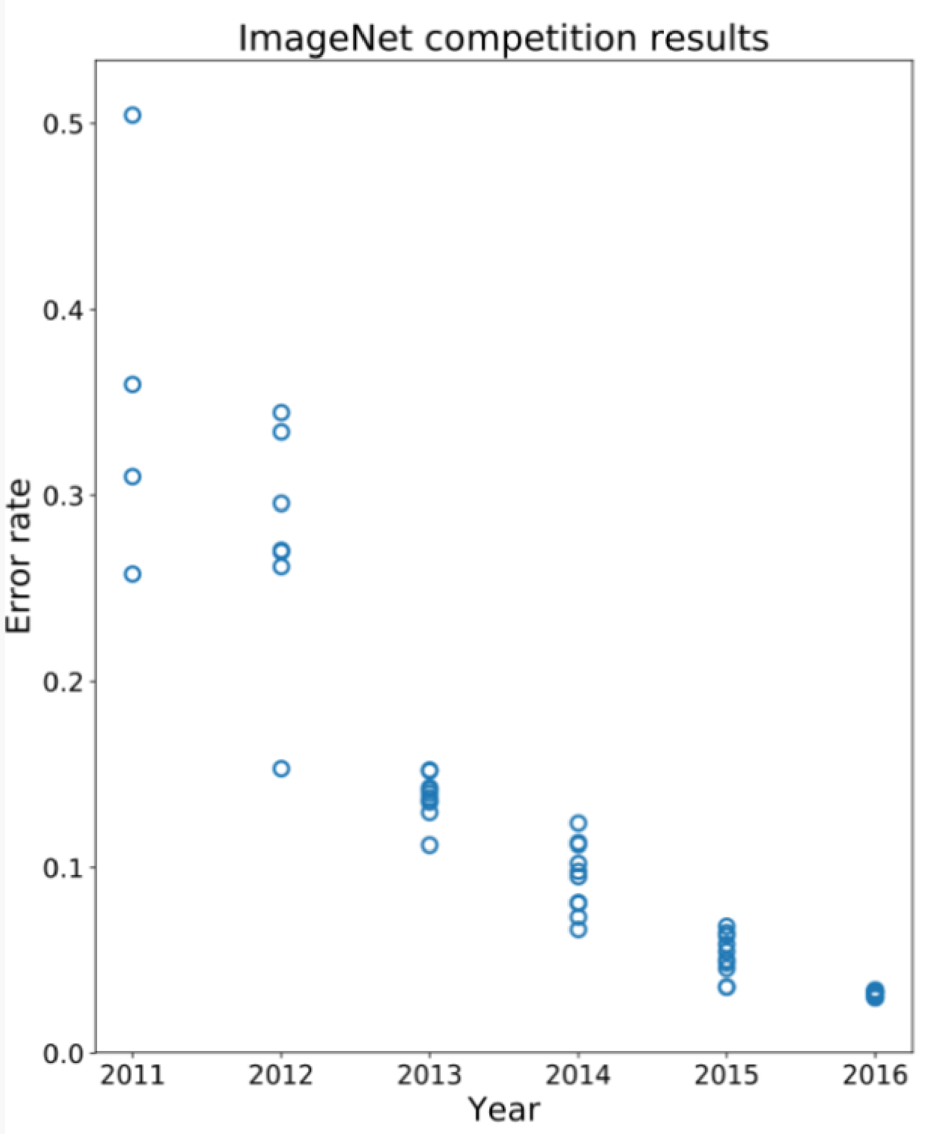
Alex Krizhevsky
University of Toronto
kriz@cs.utoronto.ca

Ilya Sutskever
University of Toronto
ilya@cs.utoronto.ca

Geoffrey E. Hinton
University of Toronto
hinton@cs.utoronto.ca

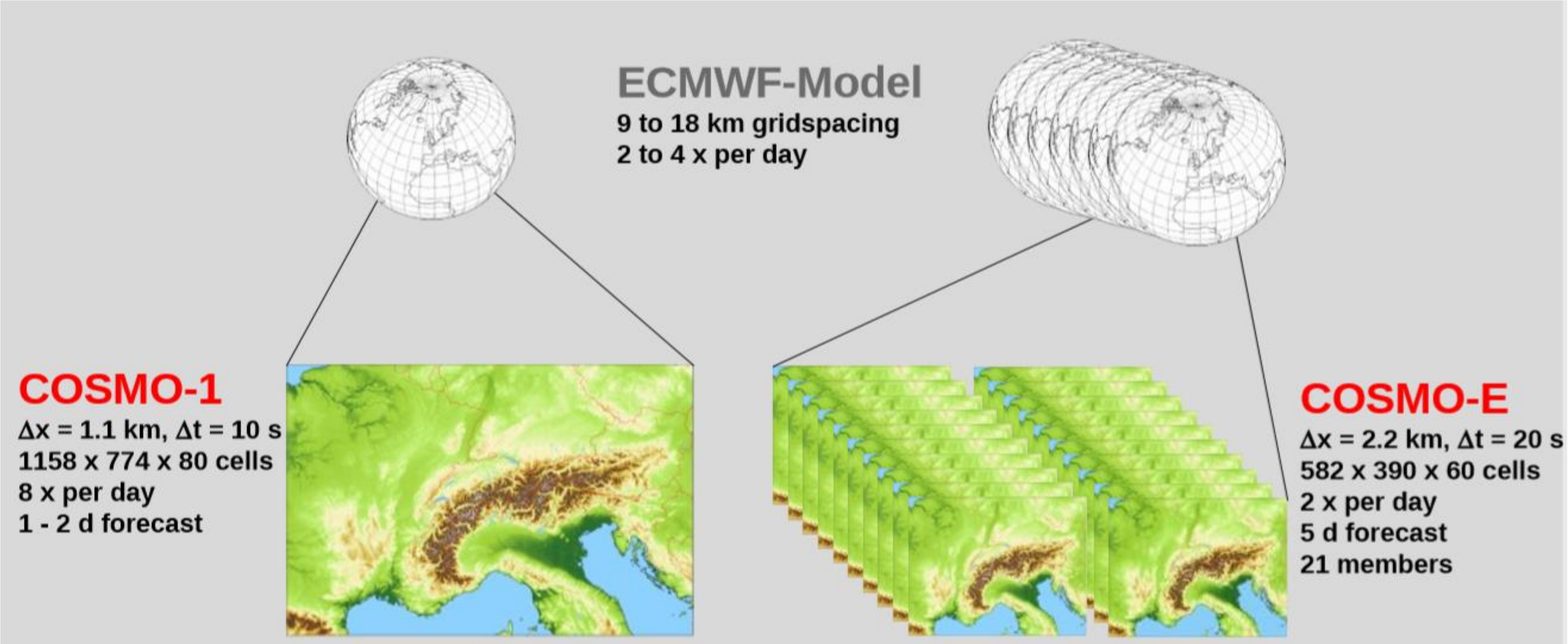
Abstract

We trained a large, deep convolutional neural network to classify the 1.2 million high-resolution images in the ImageNet LSVRC-2010 contest into the 1000 different classes. On the test data, we achieved top-1 and top-5 error rates of 37.5% and 17.0% which is considerably better than the previous state-of-the-art. The neural network, which has 60 million parameters and 650,000 neurons, consists of five convolutional layers, some of which are followed by max-pooling layers, and three fully-connected layers with a final 1000-way softmax. To make training faster, we used non-saturating neurons and a very efficient GPU implementation of the convolution operation. To reduce overfitting in the fully-connected layers we employed a recently-developed regularization method called "dropout" that proved to be very effective. We also entered a variant of this model in the ILSVRC-2012 competition and achieved a winning top-5 test error rate of 15.3%, compared to 26.2% achieved by the second-best entry.



"**Suddenly** people started to pay attention to Deep Neural Networks, not just within the AI community but across the technology industry as a whole." (The Economist, 25 June 2016)

Challenge #1: Growing HPC demands



Source: C. Osuna et al., Designs for Efficient Weather and Climate Models , ENES HPC Workshop

1km resolution requires:

- 97M node hours
  840 simulation days
- 68m CHF

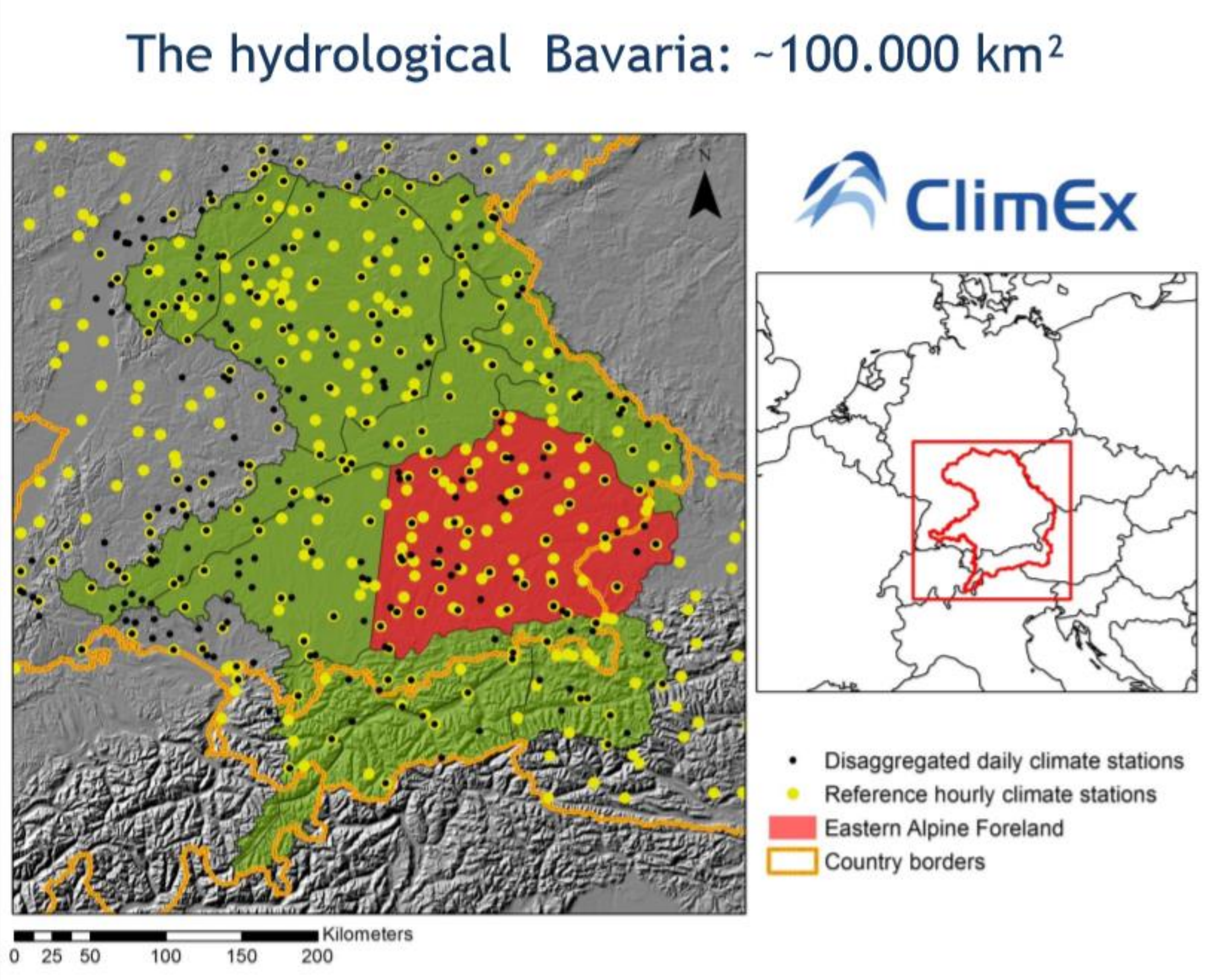
- 22GWh
- 3800 tons CO_2

Source: PASC 2018

Need for km-scale global climate simulations

2x resolution increase -> 8x computational effort

Challenge #1: Growing HPC demands



ClimEx: modeling the effects of climate change on meteorological and hydrological extreme events in Bavaria and Québec

Model: CanESM2 (ca. 200 km resolution) integrated with CRCM5 (ca. 11km resolution)

50 transient runs x 150 years -> 7500 years of modeled climate

Data: +500 TB

HPC: Leibniz Supercomputer Center

Challenge #2: Heterogeneity

- Sharing climate models and data
  - Different clusters based on different hardware
  - Hard maintenance: multiple branches, complex dev test procedures
  - Data movement/sharing challenge, portability challenge

What is your data strategy?
What are the challenges in such heterogenic and diverse ecosystem?
How do you expose/publish your models across Europe and the globe?
Where the models/data are located physically?

Challenge #3: Reproducibility

"14 percent of published research is false because data analysis was inaccurate or misinterpreted." [1]

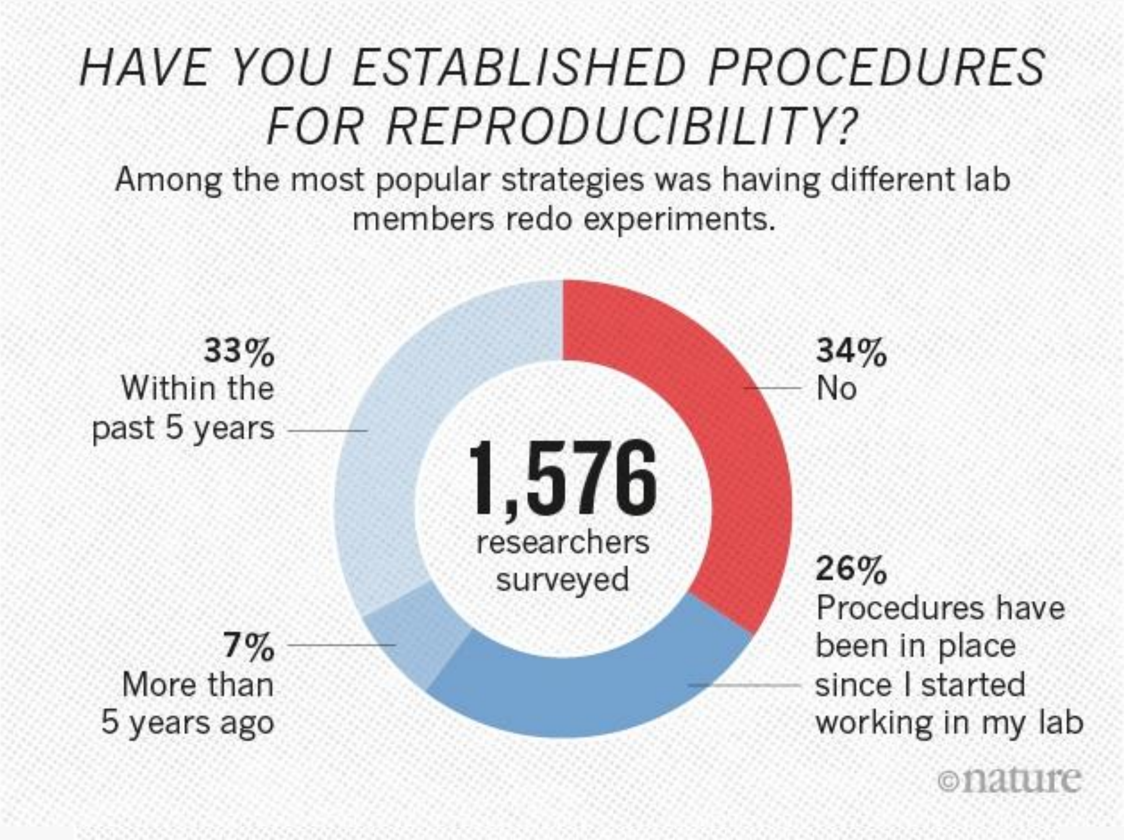
In 2011 Bayer conducted its own **reproducibility tests on 67** published experiments and failed to get similar results on **53**.

More than **70% of researchers** have tried and **failed** to reproduce another scientist's experiments, and **more than half have** failed to reproduce their own experiments [2,3].

"We need to archive, store and secure these data so that the steps taken in each experiment are traceable."
—Alex Kolodkin, chair of the Task Force on Research Reproducibility at Johns Hopkins Hospital.

[1] Could you Repeat That? Hopkins Medicine Magazine, Fall 2017.
[2] Scientific Rigour and Reproducibility, Special Edition, Nature, May 2017.
[3] 1,500 scientists lift the lid on reproducibility, Nature Survey, May 2016.

Challenge #3: Reproducibility



The Recomputation Manifesto

1. Computational experiments should be recomputable for all time [4]
2. Recomputation of recomputable experiments should be very easy
3. It should be easier to make experiments recomputable than not to
4. Tools and repositories can help recomputation become standard
5. **The only way** to ensure recomputability is to provide **virtual machines**
6. Runtime performance is a secondary issue

[3] 1,500 scientists lift the lid on reproducibility, Nature Survey, May 2016.
[4] Recomputation Intitiative

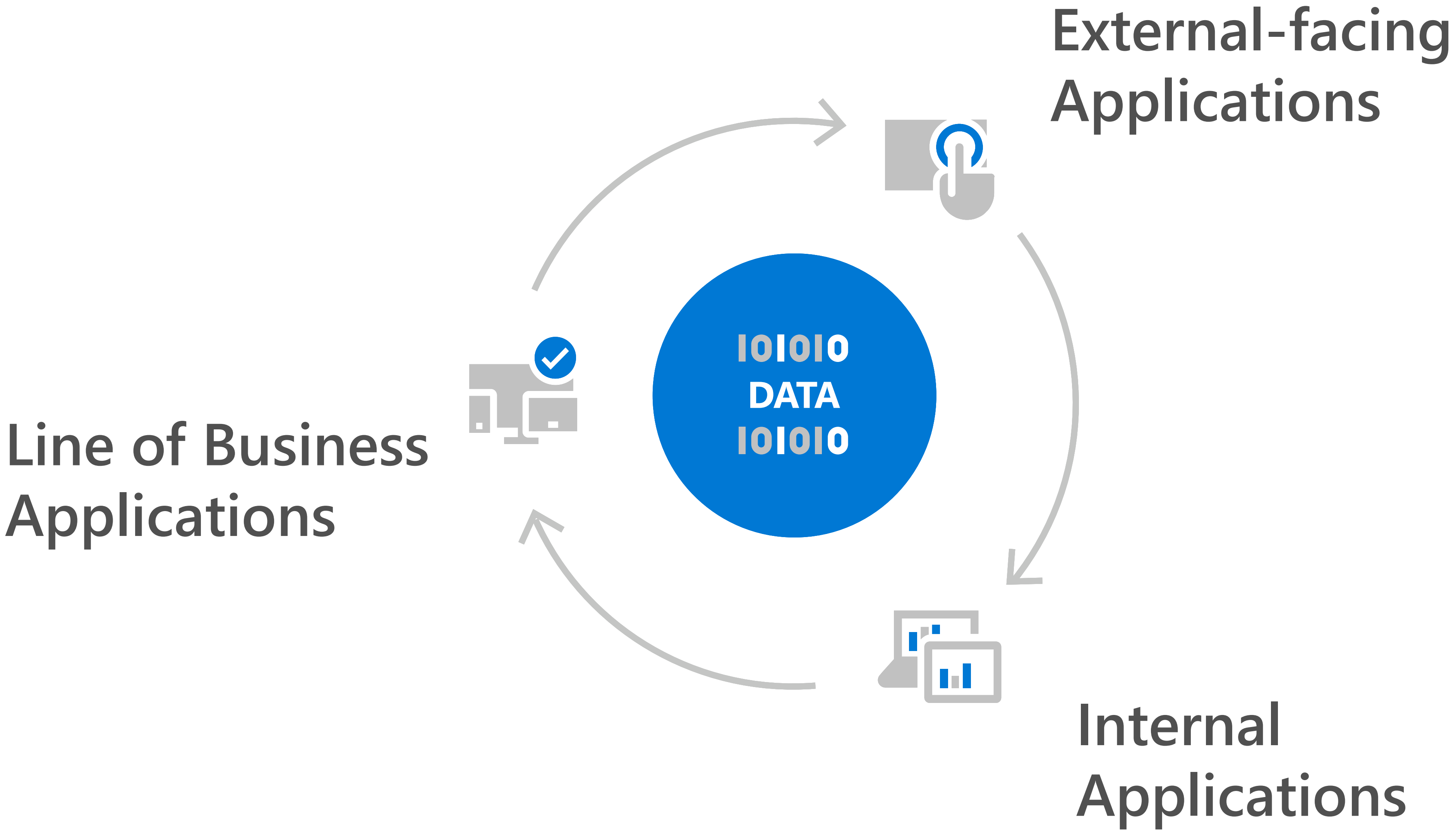
Microsoft

What is "The Cloud"?

“The Cloud” is a platform which allows you to

Write, run and **use** software

Business is powered by applications

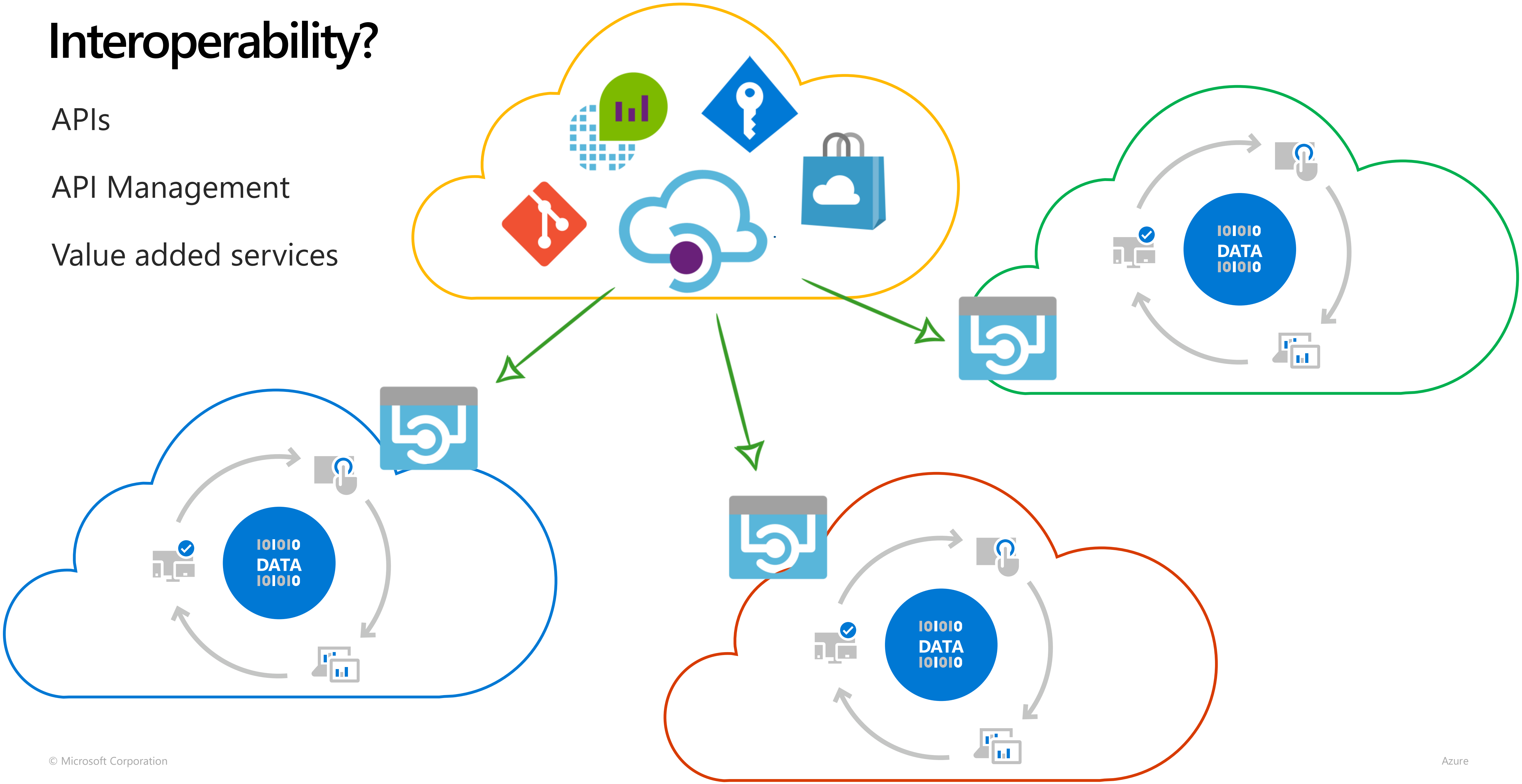


Interoperability?

APIs

API Management

Value added services



Application layers

Code

Data + Intelligence

Infrastructure

Challenges

Code

Application resiliency

Extensibility

Future proof

Data + Intelligence

Data growth

Performance

Rising costs

Confidentiality

Infrastructure

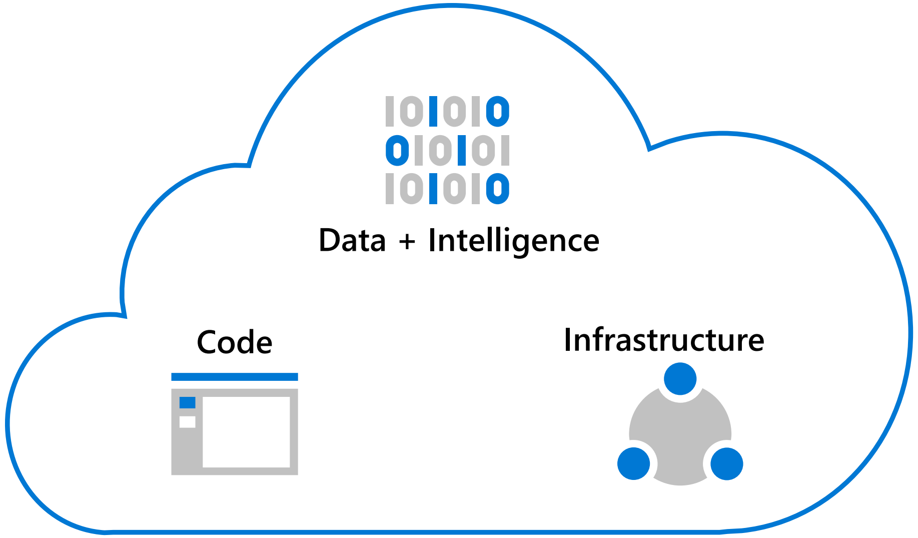
Hybrid/Global deployments

Security and management

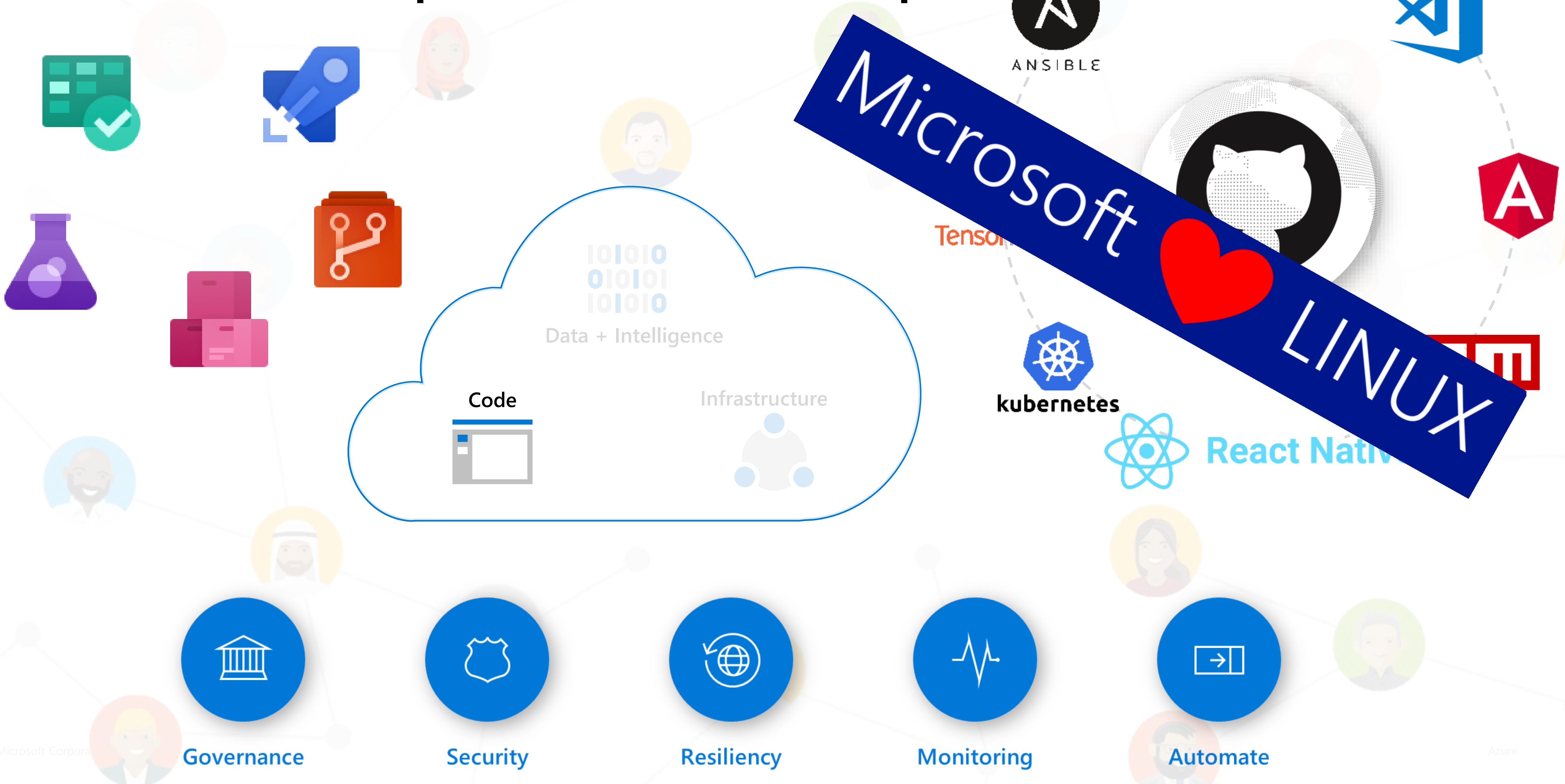
Continuous monitoring

Cost management

Microsoft Value Proposition – some examples



Microsoft Value Proposition – some examples



Microsoft Value Proposition – some examples

IoT

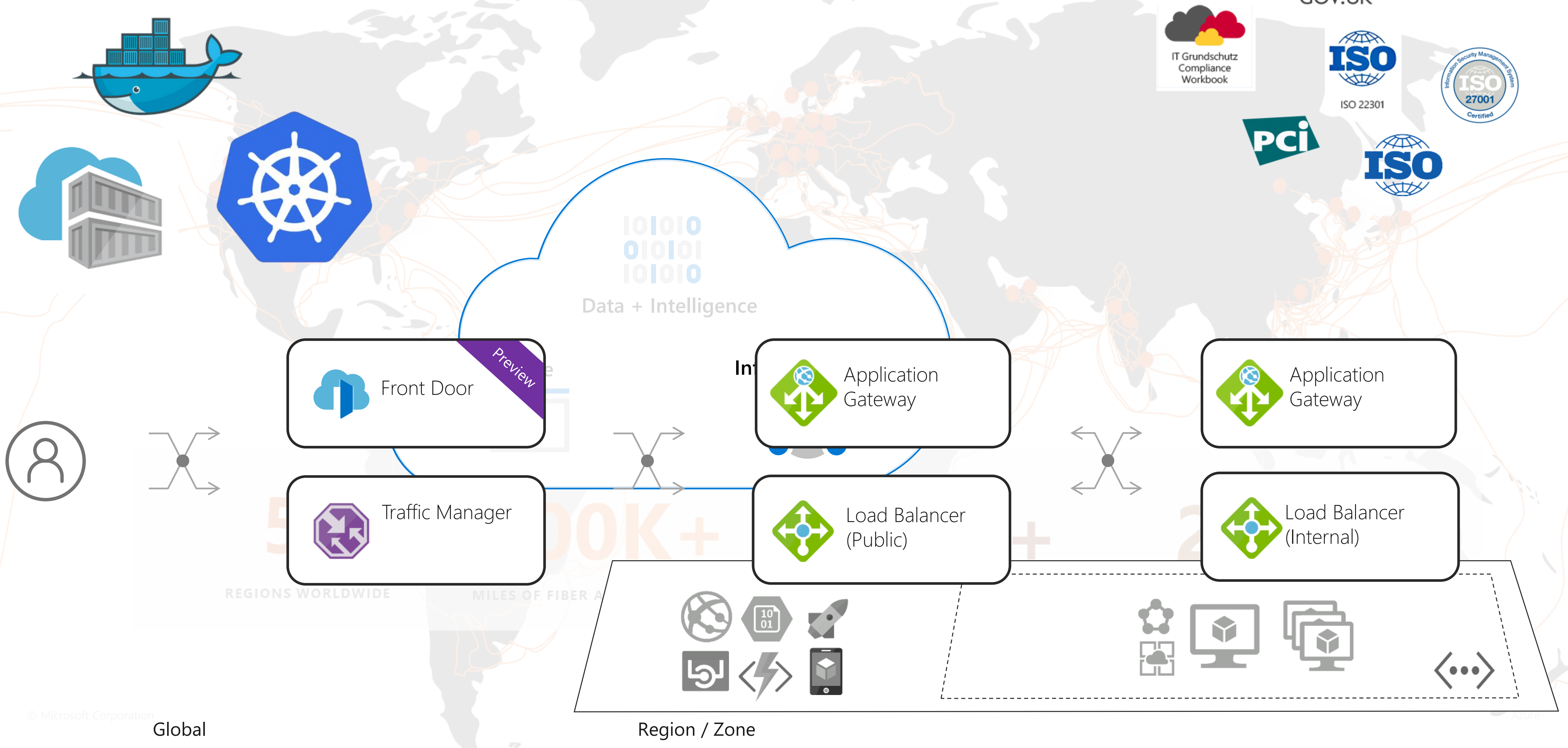
Machine Learning

Data + Intelligence

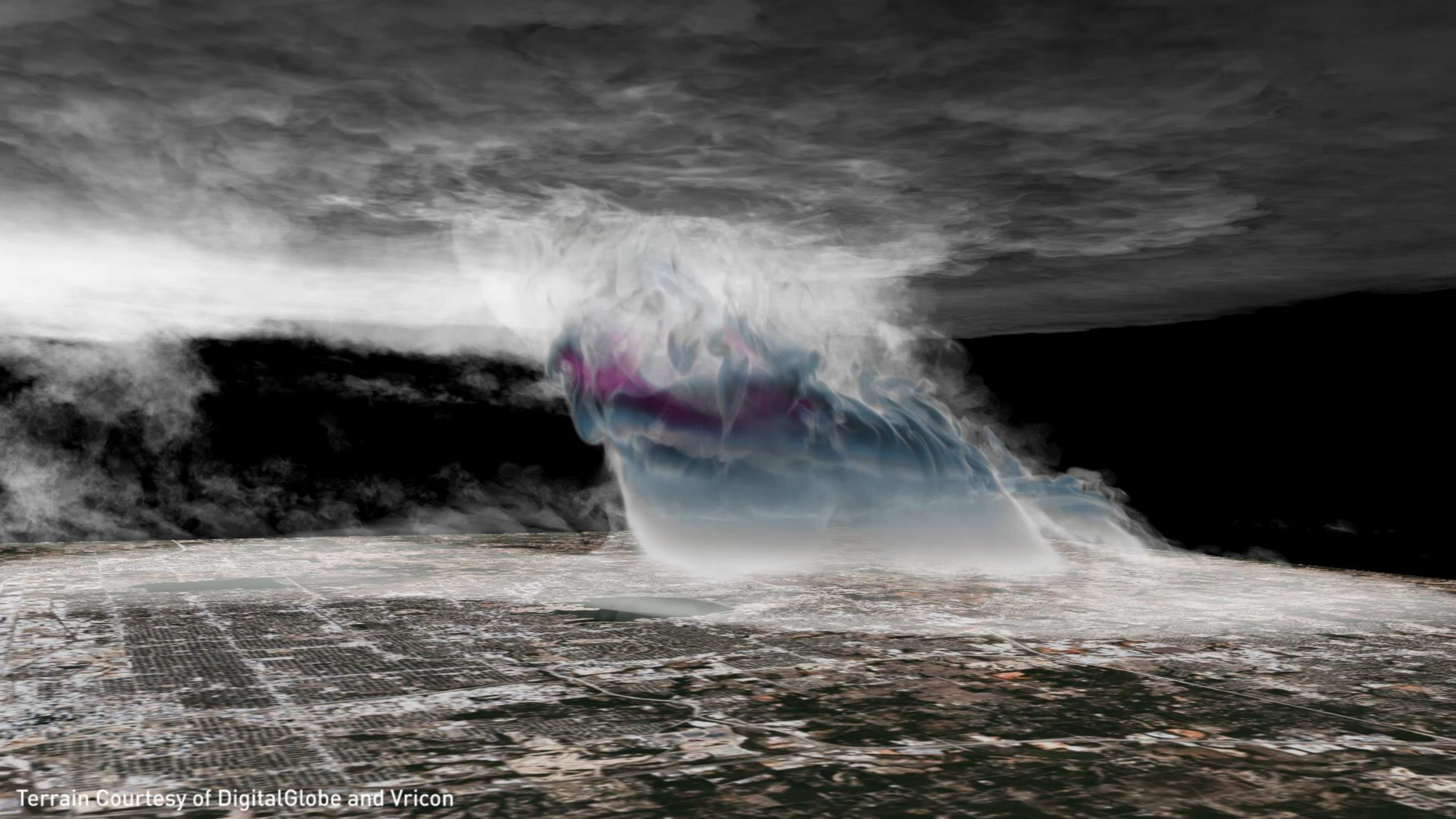
1. The data volumes are exploding, more data has been created in the past two years than in the entire previous history of the human race.
2. Data is growing faster than ever before and by the year 2020, about 1.7 megabytes of new information will be created every second for every human being on the planet.
3. By then, our accumulated digital universe of data will grow from 4.4 zettabyets today to around 44 zettabytes, or 44 *trillion* gigabytes.
4. Every second we create new data. For example, we perform 40,000 search queries every second (on Google alone), which makes it 3.5 searches per day and 1.2 trillion searches per year.

Microsoft Research
Project Silica

Microsoft Value Proposition – some examples



Use Cases


Terrain Courtesy of DigitalGlobe and Vricon

MeteoSwiss is interested in a failover and devtest infrastructure in the cloud

Challenge: MeteoSwiss needs a failover location with a specialized HPC architecture containing high-performance graphics processing units to enable weather simulations in a high availability mode.

Solution: We used ND40s with eight V100 to run a docker container with COSMO provided by NVIDIA. This setup simplified the GPUs were interconnected with NVLINK Kubernetes in Azure Container Service to handle back-end functions and increase the density of services running across compute nodes.

Outcome: With Azure, we managed to achieve very good performance. The simulation time is extremely close to the DGX1 system on prem with **2-3% variation**. MeteoSwiss engaged in follow up discussion on more workloads.

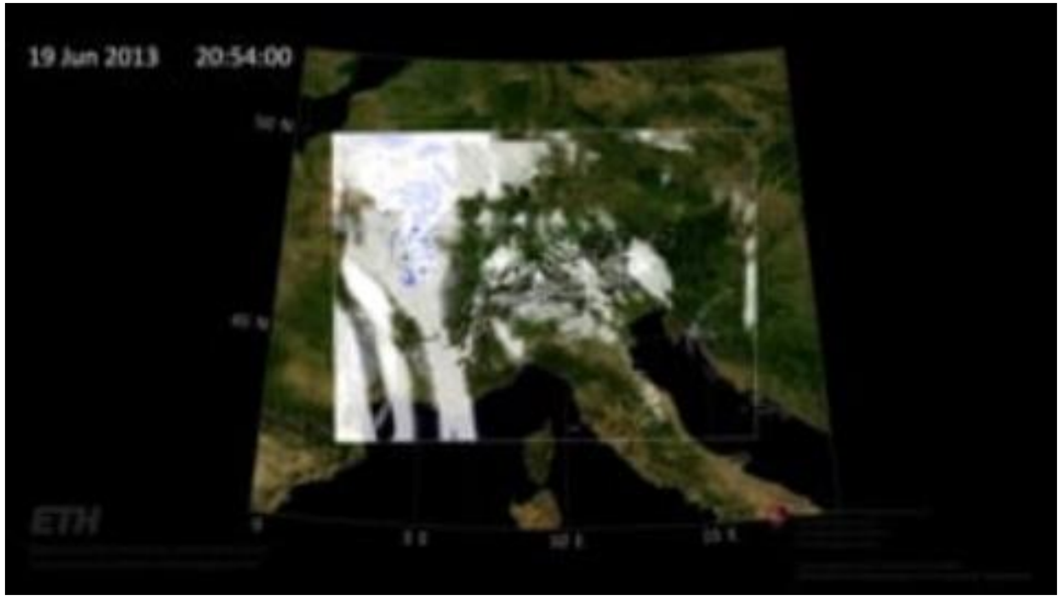


Fig. Visualization of COSMO-1 simulation.

Effort

5 commands, 1h work

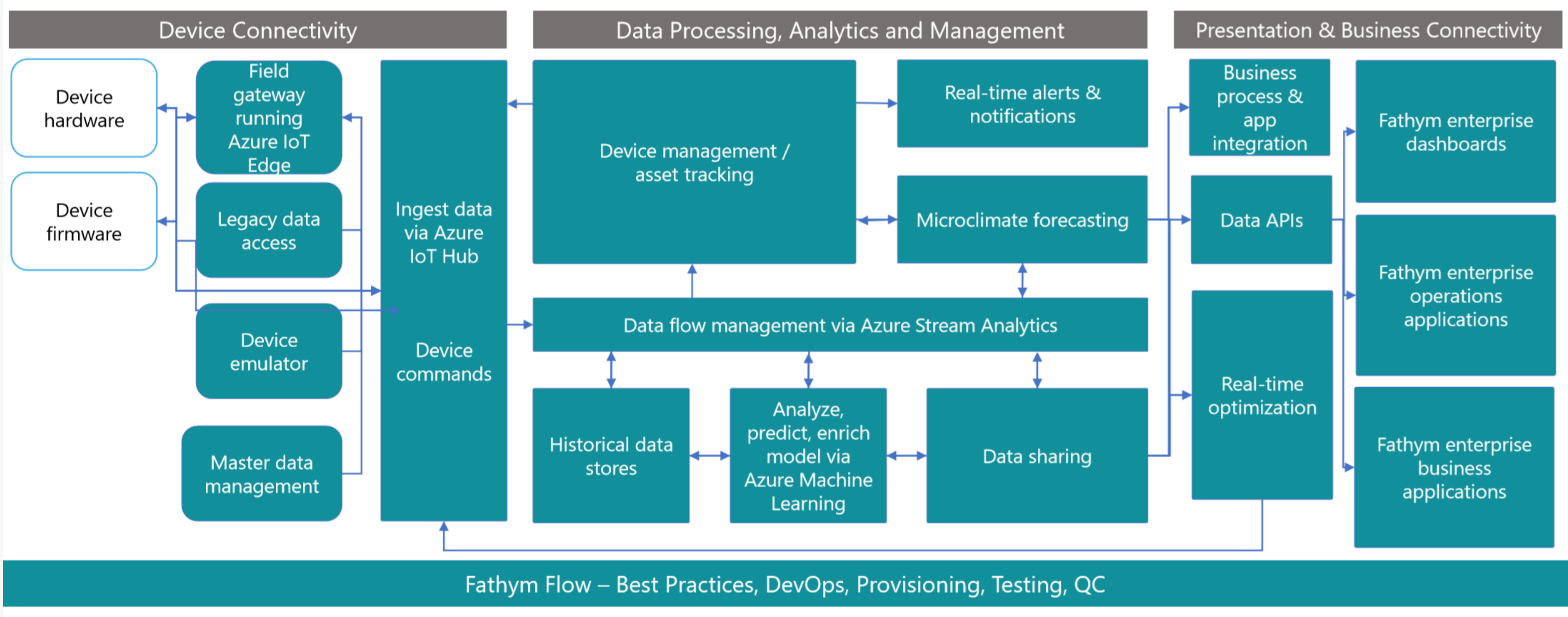
How Alaska outsmarts Mother Nature in the cloud

Alaskans are notoriously resilient people, and they don’t want to be sidelined by the weather. To keep Alaska’s highways open and safe during severe winter weather, the Alaska Department of Transportation & Public Facilities uses the Fathym WeatherCloud solution and Microsoft Azure IoT technologies to make better, hyper-local decisions about deploying road crews. By using the cloud to make better ground-level decisions, Alaska saves lives and significantly reduces road maintenance costs.

Road Weather Information System (IoT) data were transmitted to Decision Support System while mobile sensors placed on maintenance vehicles where transmitted to the cloud via mobile phones.

Products and Services	Organization Size	Industry	Country	Partner
Microsoft Azure Azure Data Lake Store Azure Event Hubs Azure IoT Hub Azure Stream Analytics	3,100 Employees	Government	United States	Fathym

Architecture



AI for Earth

AI for Earth puts Microsoft cloud and AI tools in the hands of those working to solve global environmental challenges.

Areas of focus

AI for Earth awards grants to projects that use artificial intelligence to address four critical areas that are vital for building a sustainable future.

Learn about AI for Earth grants >

Climate

The changing climate threatens human health, infrastructure, and natural systems. AI can give people more accurate climate predictions to help reduce the potential impacts.

Agriculture

By 2050, farmers must produce more food, on less arable land, and with less environmental impact to feed the world’s increasing population. AI can help people monitor the health of farms in real time.

Biodiversity

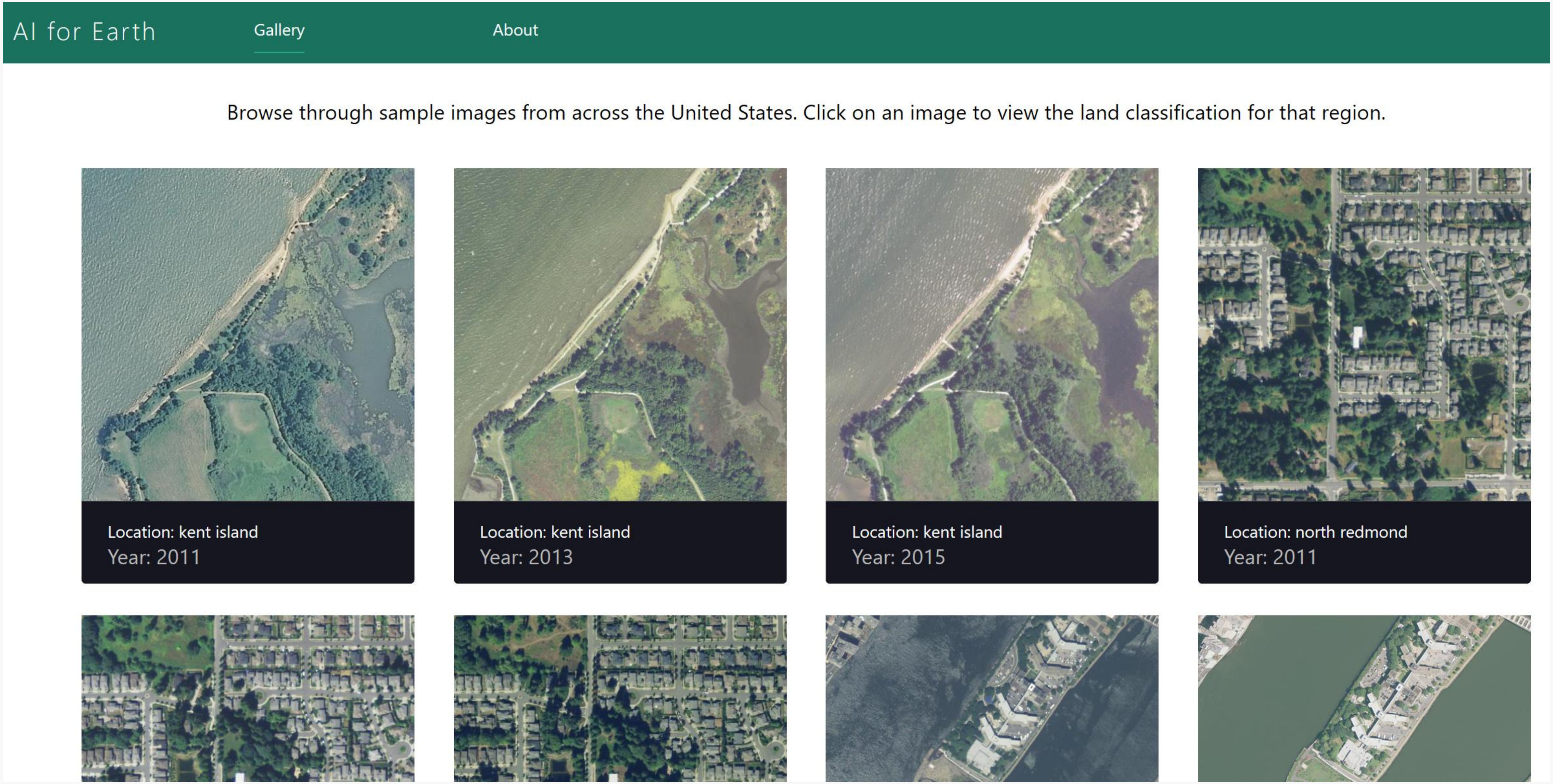
Species are going extinct at an alarming rate. AI can help people accelerate the discovery, monitoring, and protection of biodiversity across our planet.

Water

In the next two decades, demand for fresh water is predicted to dramatically outpace supply. AI can help people model Earth’s water supply to help us conserve and protect fresh water.

goto: aka.ms/aiforearth

Land Cover Mapping

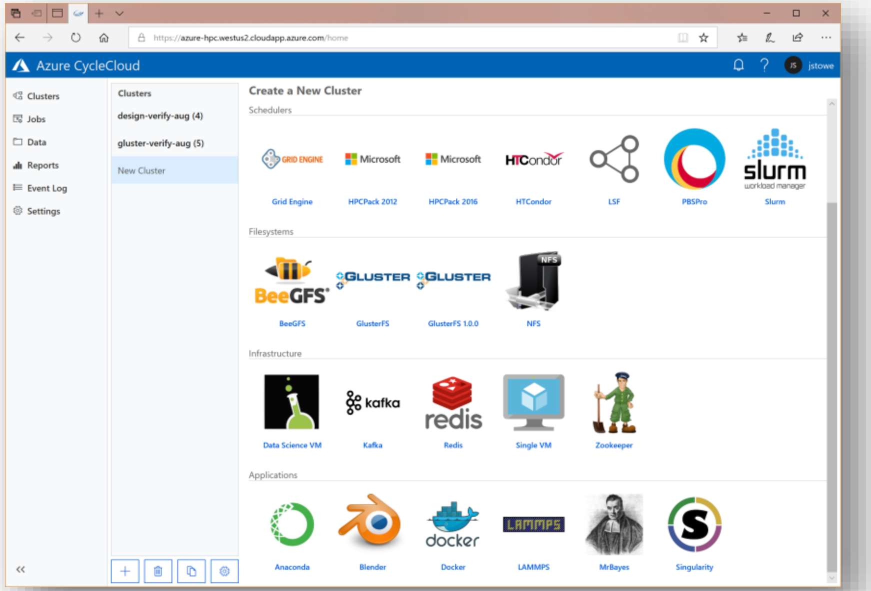


Microsoft

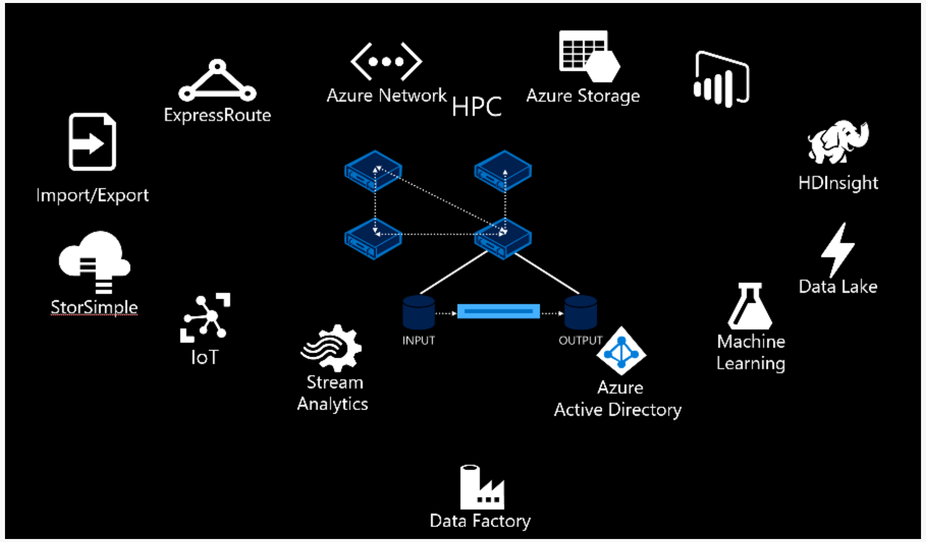
Enabling Capabilities

Next Generation HPC

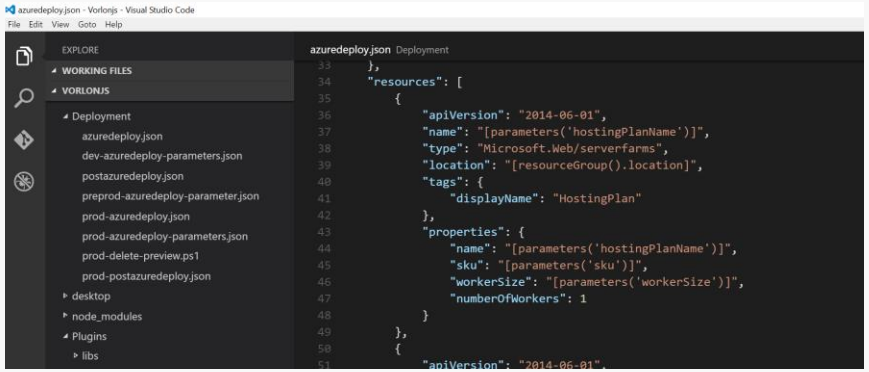
"A cluster (cloud!) for every HPC workload"



"HPC is one building block"



"HPC Infrastructure as code"



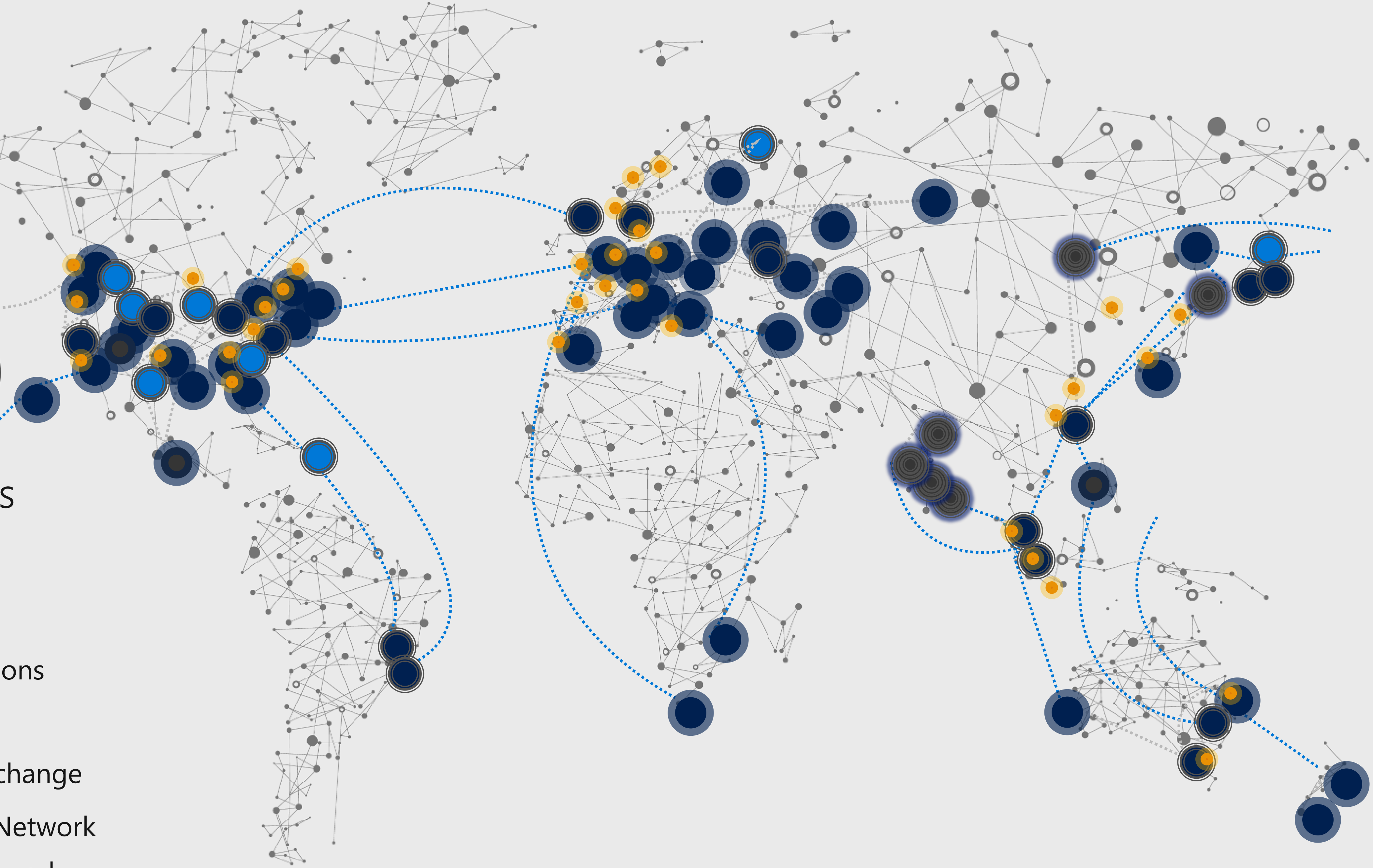
"AI, ML, Data Analytics"

Azure ML

Databricks for distributed deep learning (Horovod)
Tensorflow, Keras on Jupiter Notebooks

2.0

million miles
of fiber



- Datacenter
- CDN Locations
- Edge Node
- Internet Exchange
- Terrestrial Network
- Subsea Network

HPC in Azure- Azure adds value to HPC workloads

Most performant infrastructure

- Fastest HPC and GPU instances in the Cloud
- Specialized instances for AI Training, Remote visualization, Accelerated Analytics
- Only cloud providing ultra-fast, low latency networking with RDMA and Infiniband
- Azure capabilities for easily deploying and managing scale for large, parallel jobs

Open & Integrated

Industry specific

Life Sciences

CycleComputing, BizData, teradici, Workspot, rescale, Bright Computing

Finance

Willis Towers Watson, excelian, Milliman, rescale, ActiveViam, IHS Markit, kinetica

Manufacturing

Altair, d3View, JSOL Corporation, Dassault Systèmes, CD-adapco, esi, PBS Works, MSC Software, UberCloud, OpenFOAM, LSTC, Schlumberger, ANSYS, Halliburton, Siemens, Simulia

Software & Hardware

Languages

R, node, python, Ruby, Java, php

Operating Systems

CentOS, ubuntu, redhat, SUSE, FreeBSD, Windows Server

Infrastructure

excelian, CycleComputing, rescale, UberCloud, Fujitsu, intel, Mellanox, Bright Computing, teradici, Workspot

Granular cost control & Governance

- Flexible consumption and cost savings with low-priority VMs on Azure
- Per-minute billing for VMs
- Granular insights into HPC usage & costs helping with workload optimization
- Built-in policy based governance for richer collaboration
- Largest global footprint & compliance portfolio of any cloud

HPC in Azure: VMs with RDMA, GPU, FPGA + Cray

A

Entry Level VMs

Dev/Test
Workolads

D

General Purpose VMs

Common Applications,
Intel Haswell/Broadwell

F

Compute Optimized VMs

Gaming, Analytics

M

Large Memory VMs

Up to 128 vCPUs and
3.8 TB RAM

>80,000 IOPs
Premium Storage

Low latency, high
throughput apps

Infiniband

L

Storage optimized VMs

No SQL Databases
(Cassandra, MongoDB),
High Disk Throughput
and IO

H

High Performance VMs

Batch processing, fluid
dynamics, monte
carlo simulation

N

GPU-enabled VMs

NV -Graphic based applications

NC2 – Advanced Sim (P100-X)

ND1 – AI Inferencing (P40)

ND2* – AI Training (V100/V100
SXM)

FPGA*

Virtual Machines – HPC

FPGA Microservices –
AI/Edge

NEW!

Cray in Azure

Aries Connected
CPU/GPU/Storage
available in cloud

Cray in Azure

❶ A **dedicated** Cray XC or CS-Storm system with ClusterStor **built to your specifications**, running in an Azure data center, and delivered as managed service to **your virtual network**.

❷ A **scalable** and powerful infrastructure to support the needs of the most demanding codes in engineering, climate, energy, and others in a way that **no other public cloud can approach**.

❸ Use **Azure and Cray together**. This allows for smoother multi-stage **workflows** including dev/test, visualization, AI, analytics, and data commons.

❹ Because the Azure Cray system is supported and maintained by **Cray specialists**, you receive expert support on both Azure and Cray resources.

Challenges with on-premises HPC

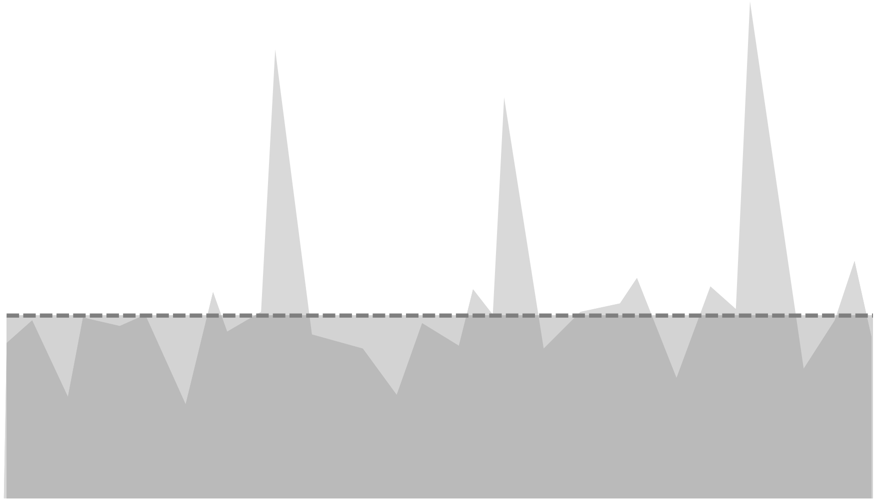
Finite resources do not scale with the business

Usage spikes creates inconsistent ROI that is challenging to forecast

Updating HW is time consuming and expensive

Regulations | AI | IoT | Data analytics

Random, unpredictable spikes in demand for HPC can come from any new or existing application workflow

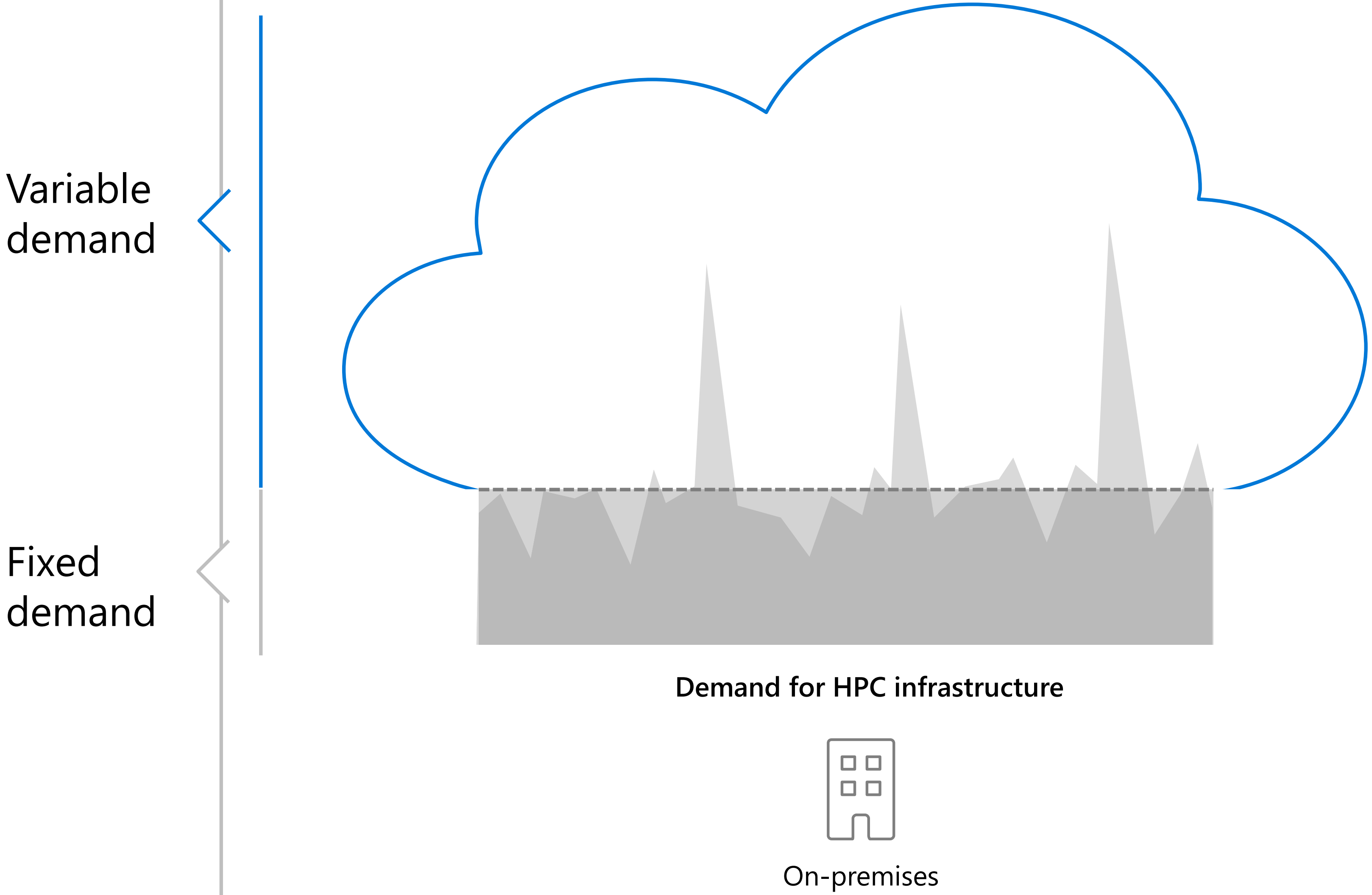


Demand for HPC infrastructure

On-premises

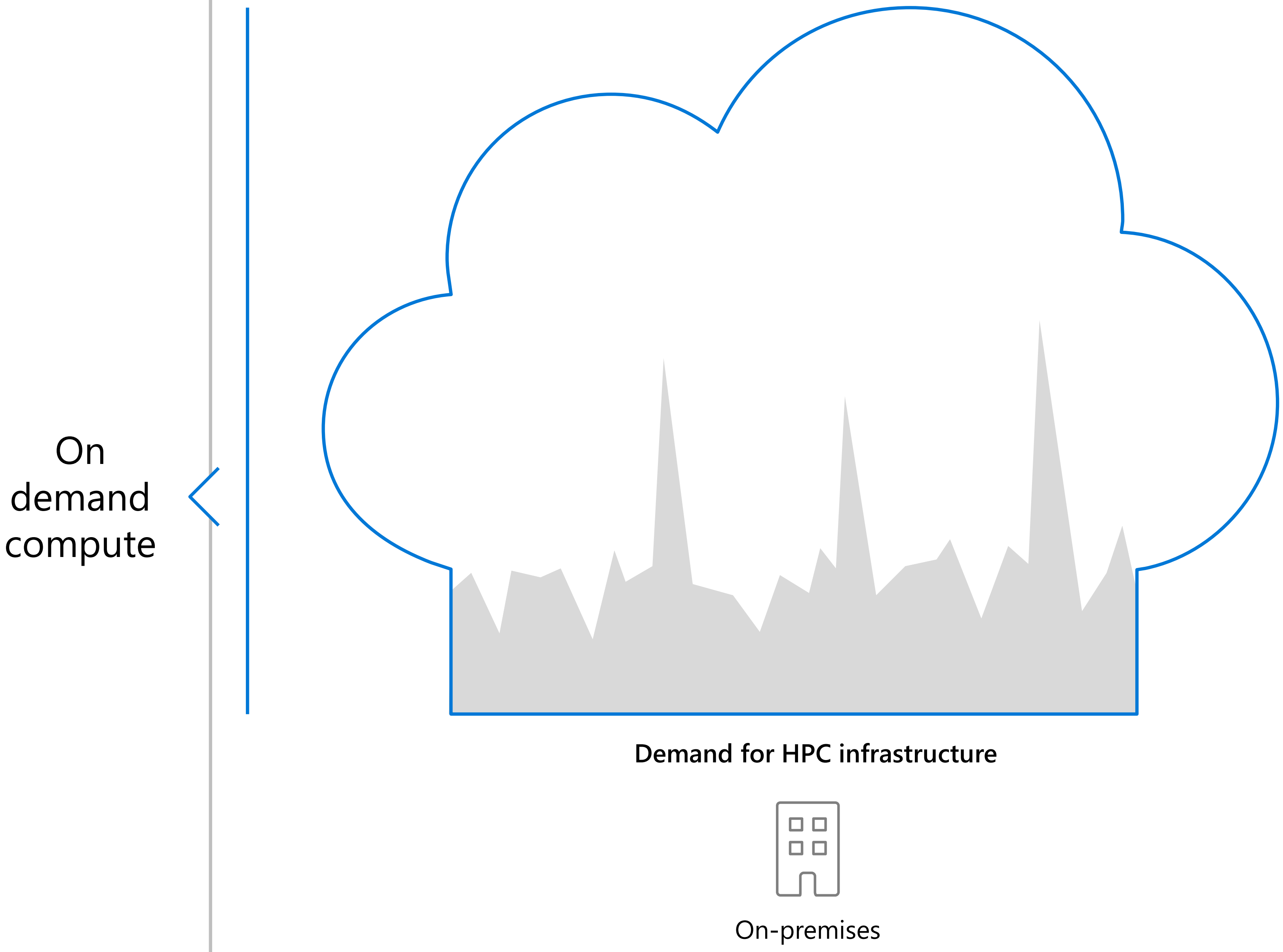
The HPC Cloud Opportunity

The **elasticity** of the cloud provides the perfect accommodation for variable demands of compute resources.



The HPC Cloud Opportunity

A fully native, cloud-based HPC environment ensures **optimized business agility** by accommodating an on-demand usage model for compute resources



Azure Big Compute Platform

HPC End-users, IT Staff, Line of Business

Developers

App Users

Cluster templates to run existing, on-prem HPC applications, schedulers

SaaS / Client Solution

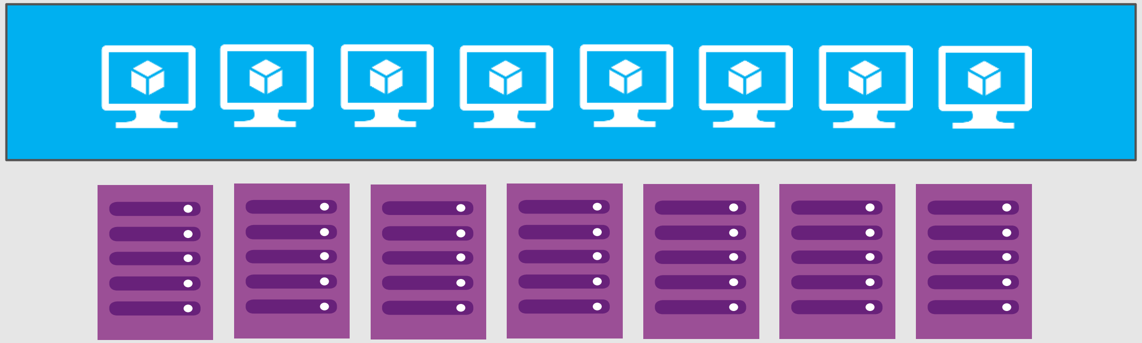
Azure Batch AI

Parallel R

VFX Plug-Ins

Azure CycleCloud
Hybrid & Cluster Manager for HPC/AI

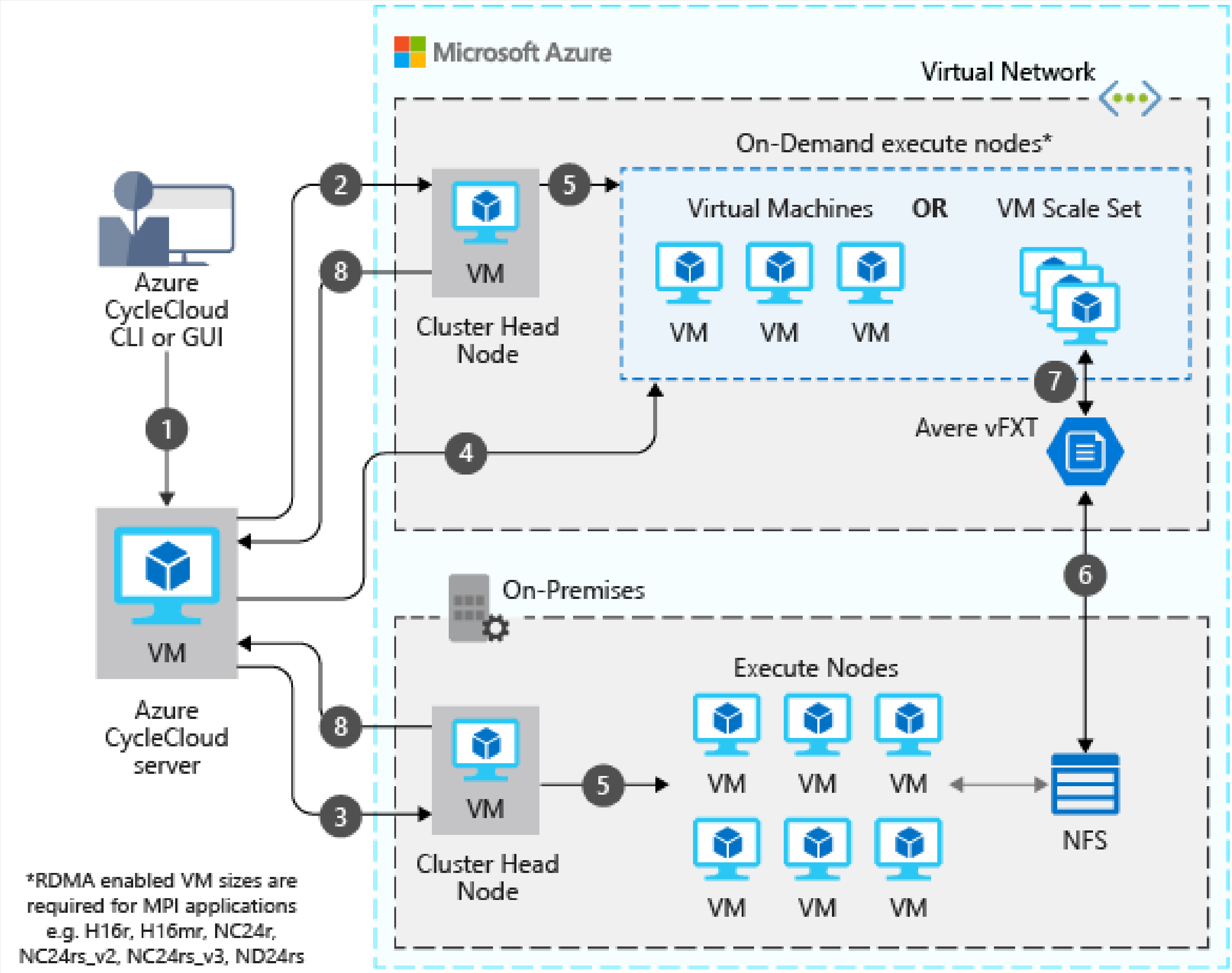
Azure Batch
VM Management & Job Scheduling



Cloud Services, VMs, VMSS

Hardware

Architecture



1. Connect to the Azure CycleCloud server to configure the cluster.

2. Configure and create the cluster head node, using RDMA enabled machines for MPI.

3. Add and configure the on-premises head node.

4. Autoscaling

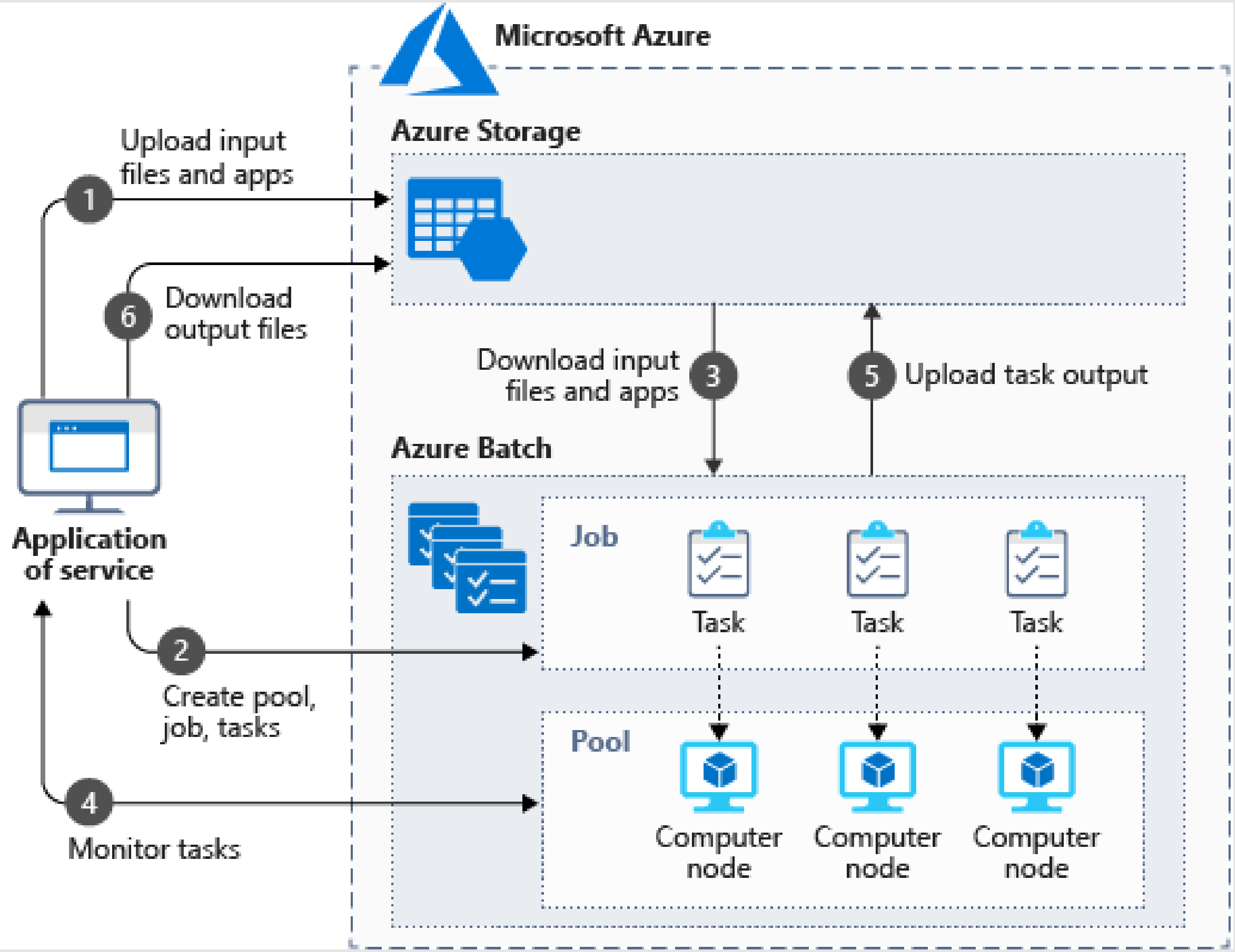
5. Tasks allocated to the execute nodes.

6. Data cached in Azure from on-premises NFS server.

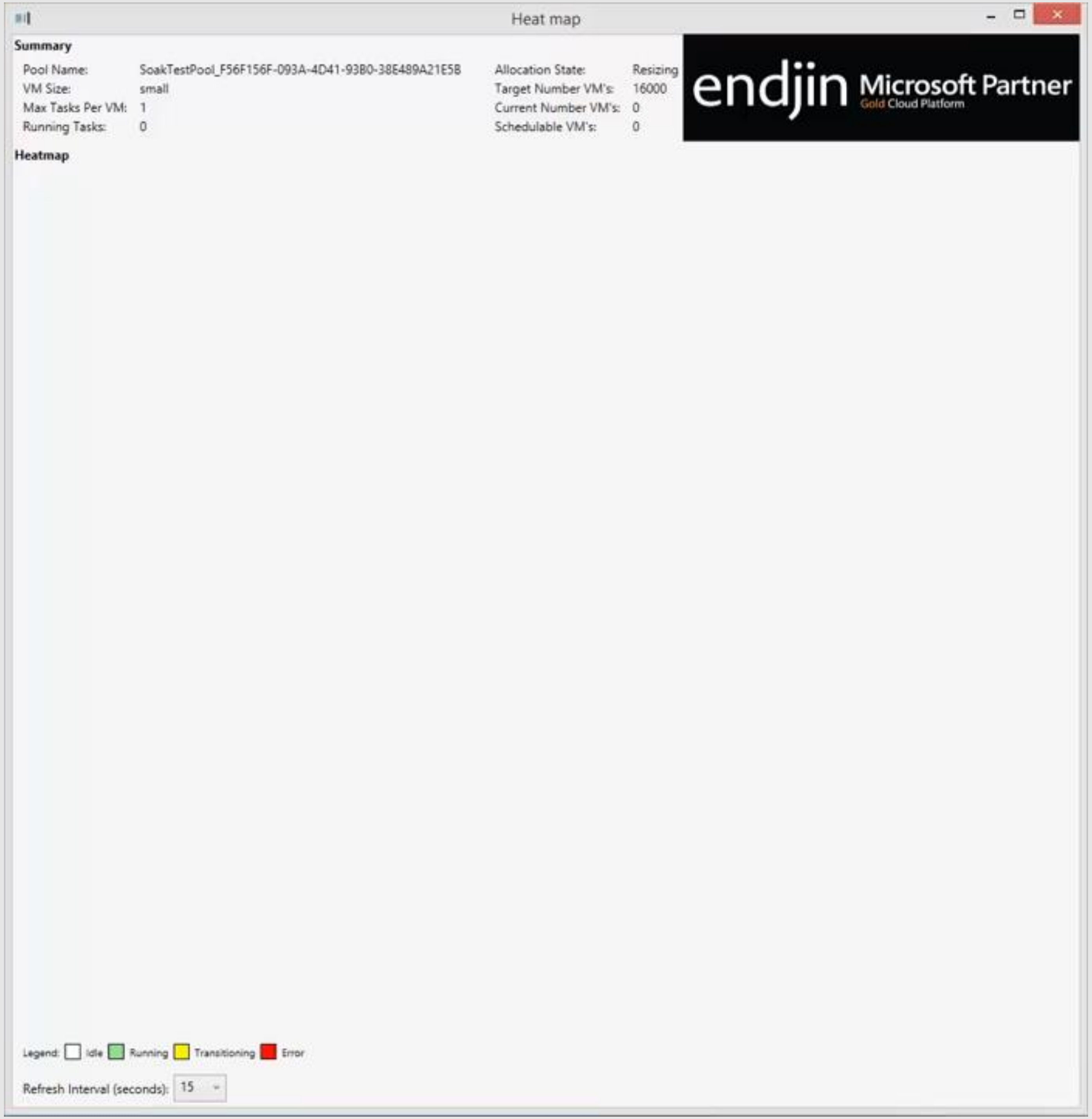
7. Data read in from the Avere vFXT for Azure cache.

8. Job and task information relayed to the Azure CycleCloud server.

Architecture



1. Upload input files and the applications to process those files to your Azure Storage account.
2. Create a Batch pool of compute nodes in your Batch account, a job to run the workload on the pool, and tasks in the job.
3. Download input files and the applications to Batch.
4. Monitor task execution.
5. Upload task output.
6. Download output files.


Heat map
Summary
Pool Name: SoakTestPool_F56F156F-093A-4D41-93B0-38E489A21E5B
VM Size: small
Max Tasks Per VM: 1
Running Tasks: 0
Allocation State: Resizing
Target Number VM's: 16000
Current Number VM's: 0
Schedulable VM's: 0
endjin Microsoft Partner
Gold Cloud Platform
Heatmap
Legend: Idle Running Transitioning Error
Refresh Interval (seconds): 15

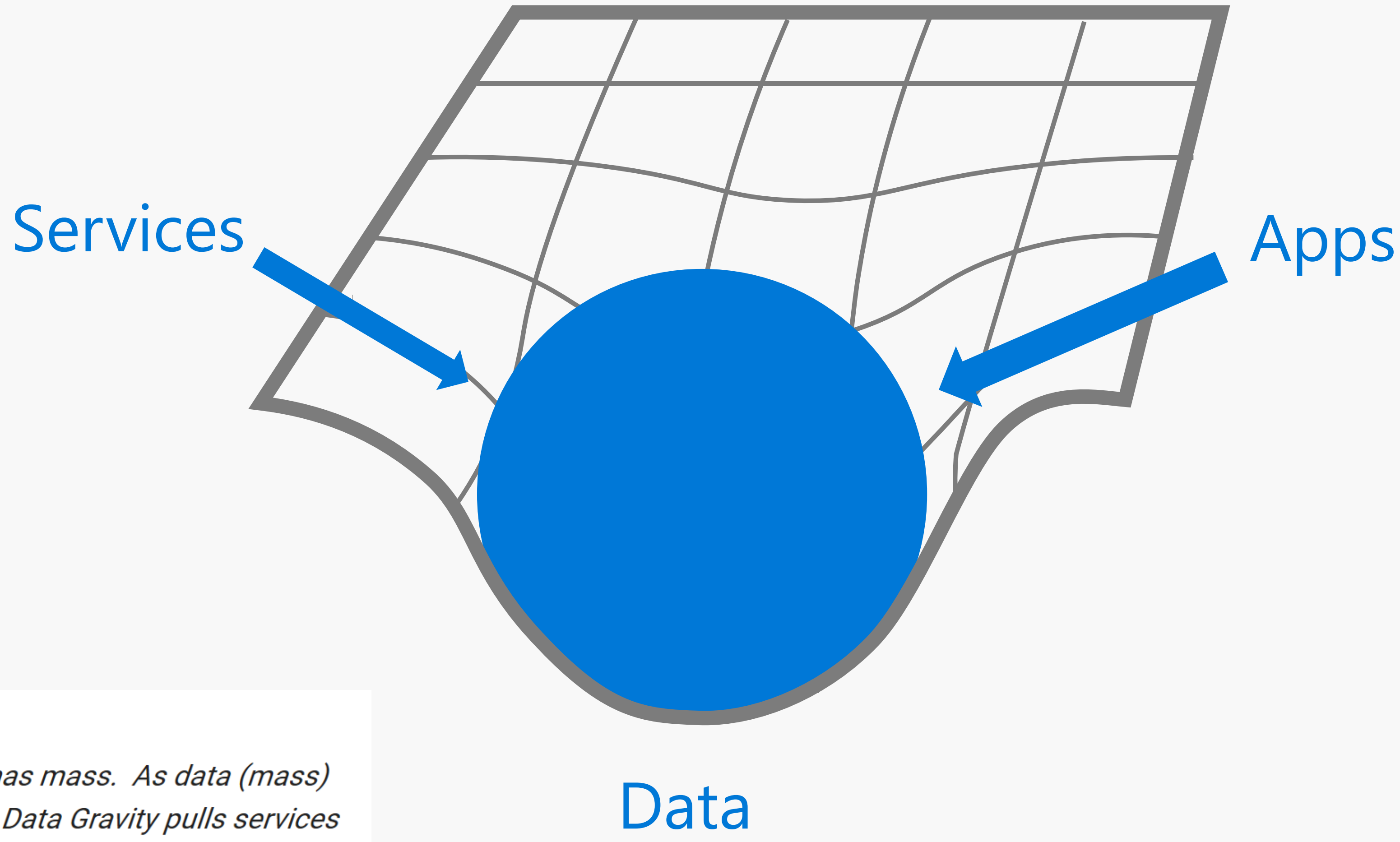
Conclusion

The concept of data gravity

Data Gravity

$$\frac{\left(\text{Data Mass} \times \text{Application Mass}\right) \times \text{Number of Requests per second}}{\left(\text{Latency in seconds} + \left(\text{Average Request Size in MBs} \,/\, \text{Bandwidth in MBs per second}\right)\right)^2}$$

McCrory's Original Equation



"

Data Gravity is a theory around which data has mass. As data (mass) accumulates, it begins to have gravity. This Data Gravity pulls services and applications closer to the data. This attraction (gravitational force) is caused by the need for services and applications to have higher bandwidth and/or lower latency access to the data.

Cloud Value for Weather Modeling

Scale

Elasticity

Pay for use

Reach & locality

Hardware selection

No infrastructure

Known costs

Microsoft