Microsoft

Big Compute: HPC on Azure

The performance and scalability of a world-class supercomputing centre is now available to everyone, on demand, in the cloud! Azure provides high memory, High Performance Computing (HPC) resources that enable you to run large parallel and batch compute jobs. Extend your on-premises HPC cluster to the cloud when you need more capacity, or run work entirely in Azure. Scale up and down based upon what you need and pay only for what you use to reduce costs.

Learning Objectives:

This workshop will familiarize you with the "Best Practices" for running Big Compute solutions on Azure and provide you with an opportunity to provision a cluster of machines capable of running parallel workloads.

Target Audience:

People who use HPC (scientists, engineers, developers). People who support HPC (IT infrastructure, architects or sysadmins). Anyone who needs to calculate, simulate, or compute! Linux or Windows. All welcome, but particularly from industries: Finance, Insurance, Manufacturing & Engineering, Life Science, Oil & Gas, Media/Rendering. L300 Hands-On.

Pre-Requisites:

Basic HPC knowledge (people who have worked or are working with HPC). Attendees must bring their own laptop, and have access to an Azure subscription.

Agenda:

Introduction: **(1 hour)**
- HPC on Azure (presentation/demo)

Fundamentals: **(3 hours)**
- Azure IaaS: My first VM, My first Scale Set
- HPC IaaS: Creating HPC clusters
- HPC PaaS: Azure Batch (Batch AI + Batch Render)

Specifics: **(2 hours)**
- Remote Visualisation
- HPC Containers
- Industry examples (Rendering, Genomics, AI)

Labs available online at:
https://github.com/azurebigcompute/Labs

Who are we? (HPC Global Black Belt team, EMEA)

Karl Podesta (IE)
Western Europe

Hugo Meiland (NL)
Western Europe

Mike Kiernan (UK)
United Kingdom

Tomasz Jozefiak (PL)
Central Eastern Europe

Karlheinz Pischke (DE)
Germany

Gabriel Sallah (DB)
Middle East & Africa

- 15+ years working as a Linux & HPC techie (engineer, consultant, trainer, architect, programme manager)
- Worked with Bloomberg, Tullow Oil, Securelinx, National Institute for Cellular Biotechnology (DCU)
- Certified by Red Hat (RHCDSS, RHCSA, RCHE, OpenStack), SUSE (CLP/CLE), LPI, Panasas, TOGAF
- 1 year with Microsoft
- Specialist for HPC on Azure
- Helping customers, partners, colleagues
- Doing workshops, presentations, POCs, demos, training – and customer work
- Covering Western Europe (7 countries – Spain, Italy, Portugal, Switzerland, Belgium, Luxembourg, Ireland)
- Helping YOU get what you can from Azure HPC

Who are we? (HPC Global Black Belt team, EMEA)

Karl Podesta (IE)
Western Europe

Hugo Meiland (NL)
Western Europe

Mike Kiernan (UK)
United Kingdom

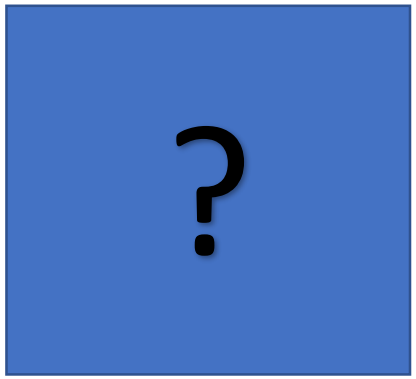
Tomasz Jozefiak (PL)
Central Eastern Europe

Karlheinz Pischke (DE)
Germany

Gabriel Sallah (DB)
Middle East & Africa

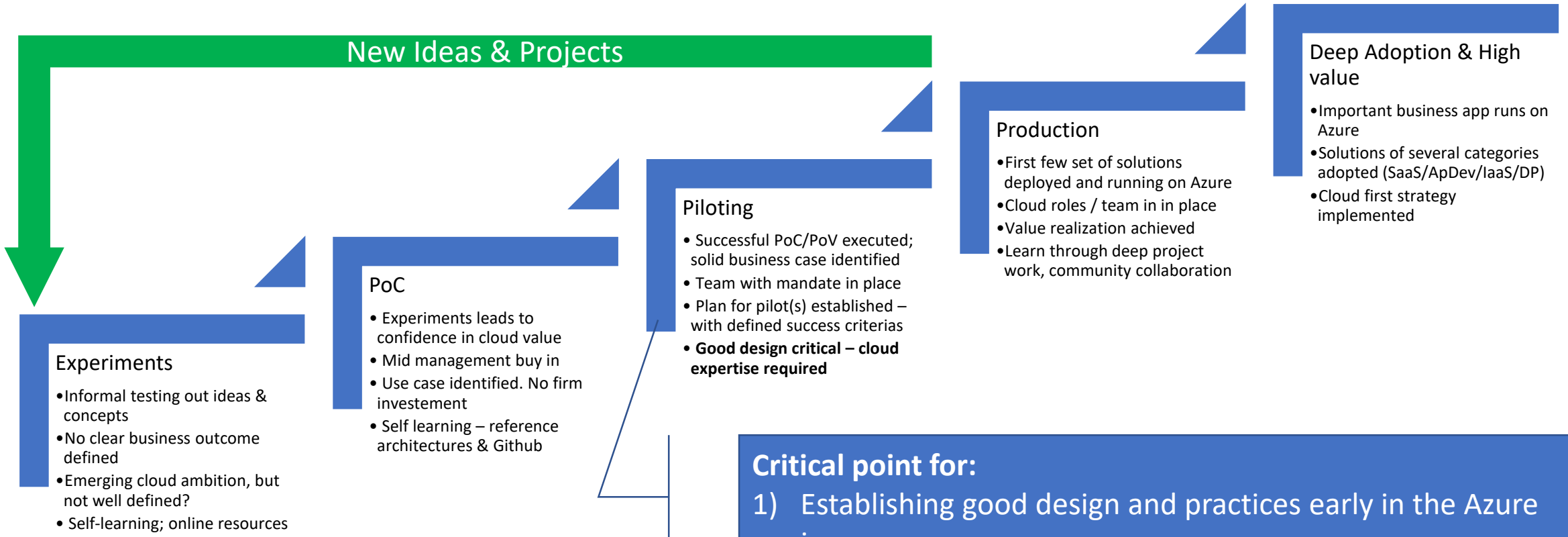
- 15+ years working as a Linux & HPC techie (consultant, admin, manager, support)
- Worked with Leiden University (NL), Bull (FR), Atos (EU)
- Certified by Red Hat (RHCE in the RH7 days)
- 6 months with Microsoft
- Specialist for HPC on Azure
- Helping customers, partners, colleagues
- Doing workshops, presentations, POCs, demos, training – and customer work
- Covering Western Europe (6 countries – Norway, Sweden, Finland, Denmark, Netherlands, Austria)
- Helping YOU get what you can from Azure HPC

Who are you?



- Name
- Role
- Company
- What you would like to learn / leave with today?

Azure Adoption Journey



Critical point for:

1) Establishing good design and practices early in the Azure journey
2) Accelerate projects; minimize time to value

Microsoft's comittment:

Help reduce time to value for prioritized clients through systematic and high quality topic based workshops

Microsoft Cloud Workshops Menu – Spring 2018

Cloud Infrastructure

- SAP@Azure
- Big Compute
- Azure Stack

Data & AI

- Database Modernization & Migration
- Azure Advanced Analytics & AI
- Cognitive Services & Deep Learning with Azure Data Services
- Spark on Microsoft Azure Databricks

IoT

- Azure IoT Solution Walk Through

Mobile App Innovation

- Mobile Applications Devops
- Data Driven Intelligent Apps

Application modernization

- Modernize your IT with Red Hat on Azure
- Modernize and Cloud-enable LOB Applications with Kubernetes & Microservices

Cloud Adoption Fundamentals: On-boarding, Storage, Networking, Security & Governance

Agenda for Today

Introduction: (90 mins) – presentation/videos/demos
- Why Cloud HPC?
- Why Azure?
- Solution Overview
- Our Partners + their solutions
- Our Customers + industries

Fundamentals: (3 hrs) – everyone does these labs
- Azure IaaS – Portal/CLI/PowerShell/SDK (CLI preferred) – create a VM, then create a scale set
- HPC IaaS – Azure ARM or CLI / CycleCloud / Microsoft HPC Pack
- HPC PaaS – Azure Batch
- HPC SaaS – Azure Batch Rendering/Genomics/AI, and Partner Solutions (UberCloud, Rescale, Altair)

Specifics: (2 hrs) – attendees can choose their lab – some examples below:
- Remote Visualisation (Linux/VNC on NV, running glxgears or other 3D interactive)
- HPC containers (try an example from Batch shipyard, discuss/show Singularity)
- Industry examples (Rendering, Genomics, AI)
- End: Challenges, Guidance, Resources

Course Agenda

09:00 - 09:30	Coffee and Registration
09:30– 09:45	Welcome and Logistics
09:45 - 10:15	Why HPC and why on Azure?
10:15 – 11:15	Overview partners and customers
11:15 - 11:45	Morning Break
11:45 – 12:15	Azure IaaS: VM and ScaleSet
12:15 – 13:15	HPC IaaS: ARM / Cycle / HPC Pack
13:15 – 14:00	Lunch and Group Discussion
14:00 – 14:45	HPC PaaS: Azure Batch CLI and Python SDK
14:45 – 15:15	HPC SaaS: Azure Batch and Ubercloud / Rescale
15:15 – 15:45	Afternoon Break
15:45 – 16:45	Labs: pick your track!
15:45 – 16:45	Labs: pick your track!
16:45 – 17:00	Discussion
17:00	Event Closes

Introduction

Karl Podesta (@karlpodesta)
kapodest@microsoft.com
https://github.com/azurebigcompute

Why Cloud HPC?

Karl Podesta (@karlpodesta)
kapodest@microsoft.com
https://github.com/azurebigcompute

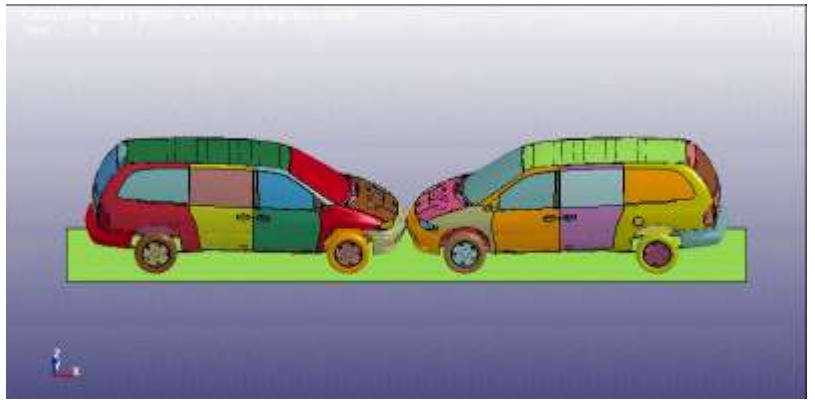
Heard about Big Data? … What about **Big Compute?**

How can our banks understand risk? Can we have safer cars?
How will global weather change affect us? Can we find & cure disease?

We use High Performance Computing (HPC) to **compute, simulate, learn, predict**

Karl Podesta (@karlpodesta)
kapodest@microsoft.com
https://github.com/azurebigcompute

Cloud HPC is about doing even more….

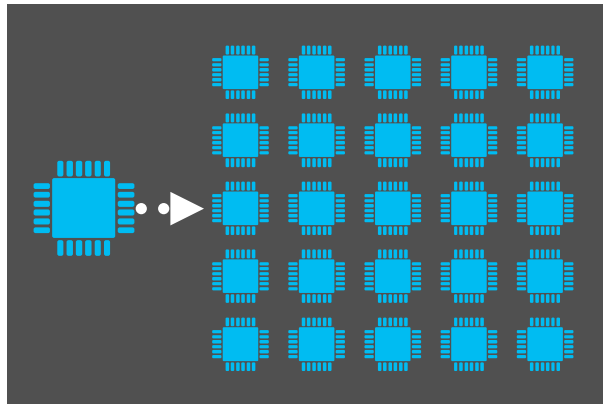


…. by many more. The cloud brings HPC to everyone!

What would you do with 100x the scale?

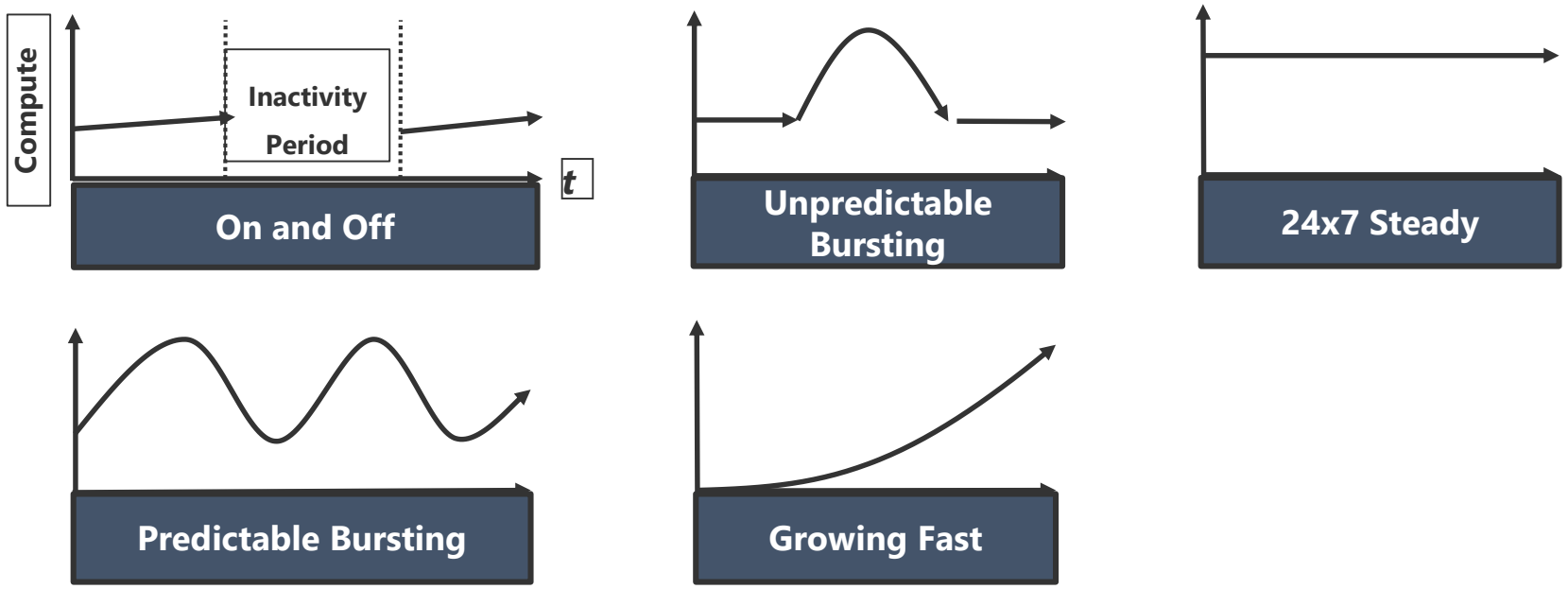
Would you do more?
- Service more users?
- Run more projects?
- Run larger simulations?
- Get results faster?
- Explore new insights?

Would you remove limitations?
- Modify more parameters?
- Analyse more complex models?
- Visualise larger results?
- Run more iterations?
- Generate higher fidelity results?

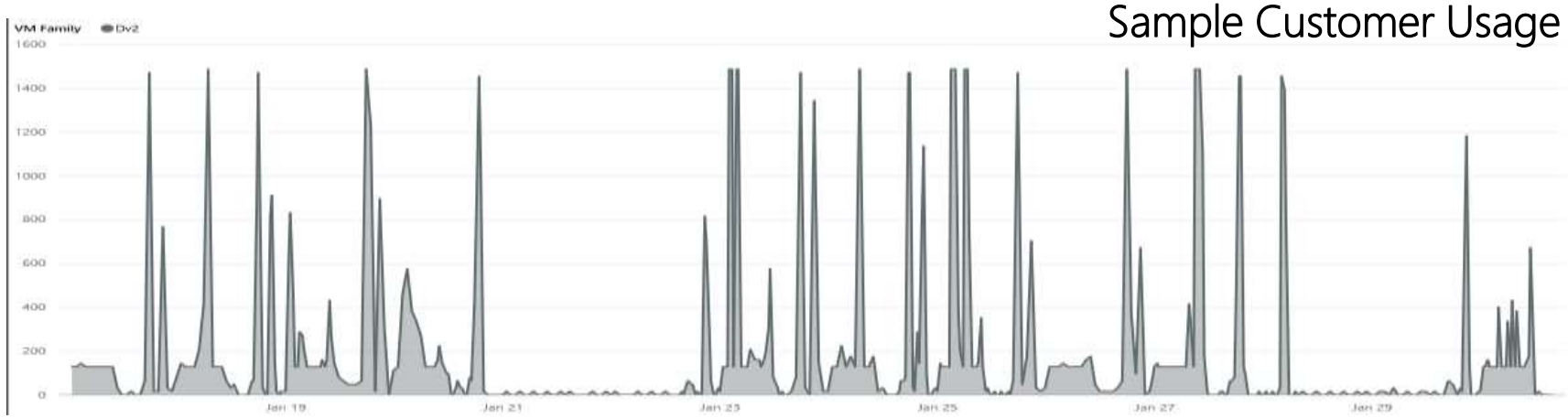


Karl Podesta (@karlpodesta)
kapodest@microsoft.com
https://github.com/azurebigcompute

Bursting & Workload Management



Sample Customer Usage



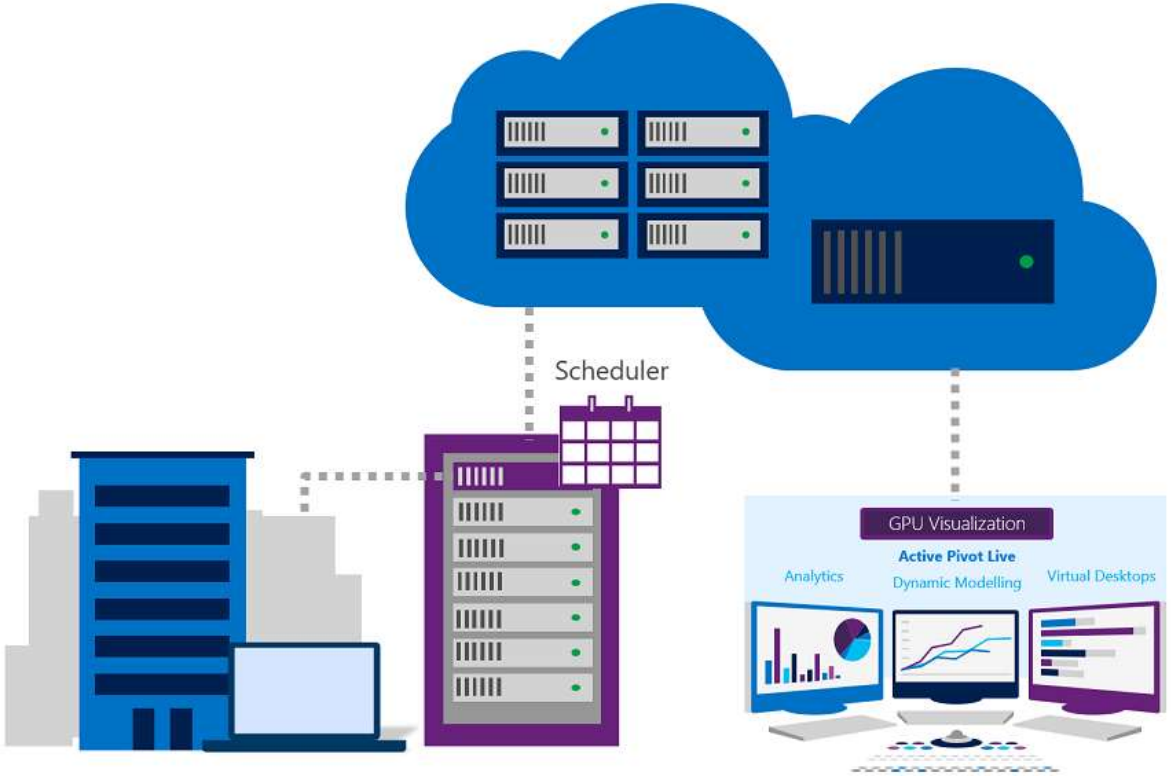
Karl Podesta (@karlpodesta)
kapodest@microsoft.com
https://github.com/azurebigcompute

Hybrid HPC – complement your capability

What you need, when you need it

Expand your existing capability

Collaborate better across locations; let users come to cloud to see & work with results



Microsoft

Cloud Workshops 2018
HPC – Big Compute

INTRODUCTION

Why Cloud HPC?

Why Azure?

Our Solutions

Our Partners

Our Customers

Our Value Add

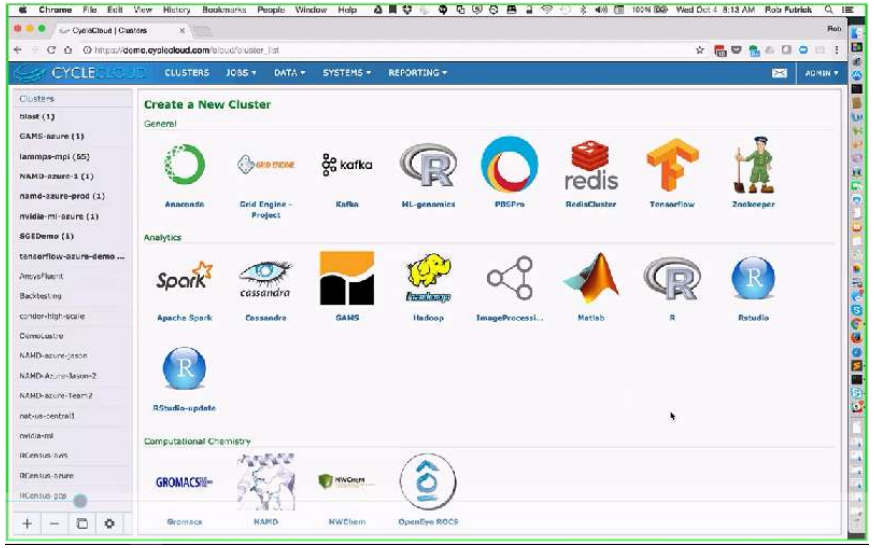
FUNDAMENTALS

SPECIFICS

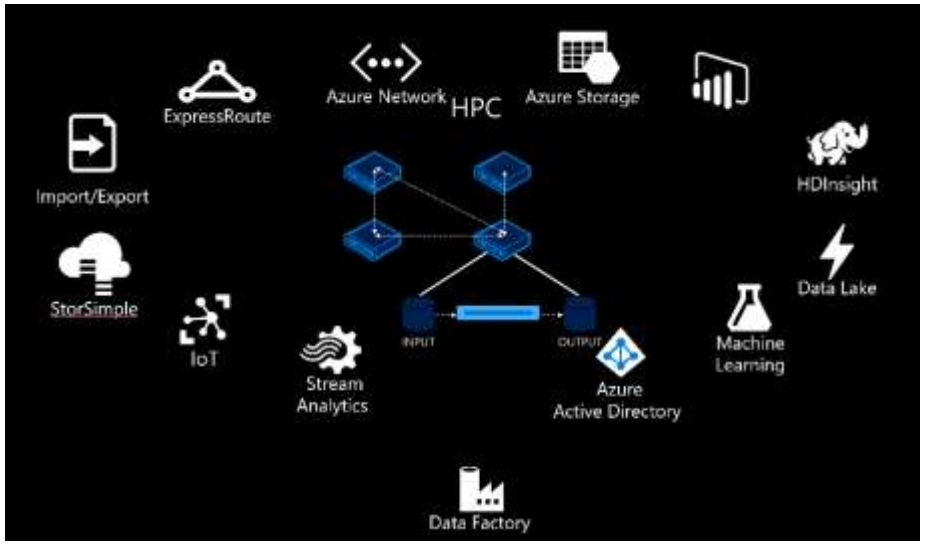
Karl Podesta (@karlpodesta)
kapodest@microsoft.com
https://github.com/azurebigcompute

Next Generation HPC – thinking differently with public cloud

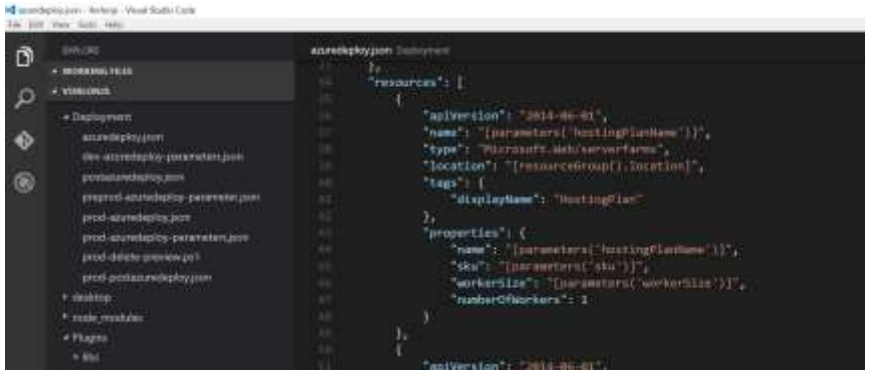
"A cluster (cloud!) for every HPC workload"



"HPC is one building block"



"HPC Infrastructure as code"



"AI, ML, Data Analytics"

Karl Podesta (@karlpodesta)
kapodest@microsoft.com
https://github.com/azurebigcompute

Why Azure?

Karl Podesta (@karlpodesta)
kapodest@microsoft.com
https://github.com/azurebigcompute

Proof points – the reasons customers choose us

Productive Hybrid Intelligent Trusted

Karl Podesta (@karlpodesta)
kapodest@microsoft.com
https://github.com/azurebigcompute

Productive – build & innovate with choice and flexibility

Integrated tooling

Visual Studios
3rd parties | DevOps

100+ services

Azure functions
Kubernetes | Logic apps

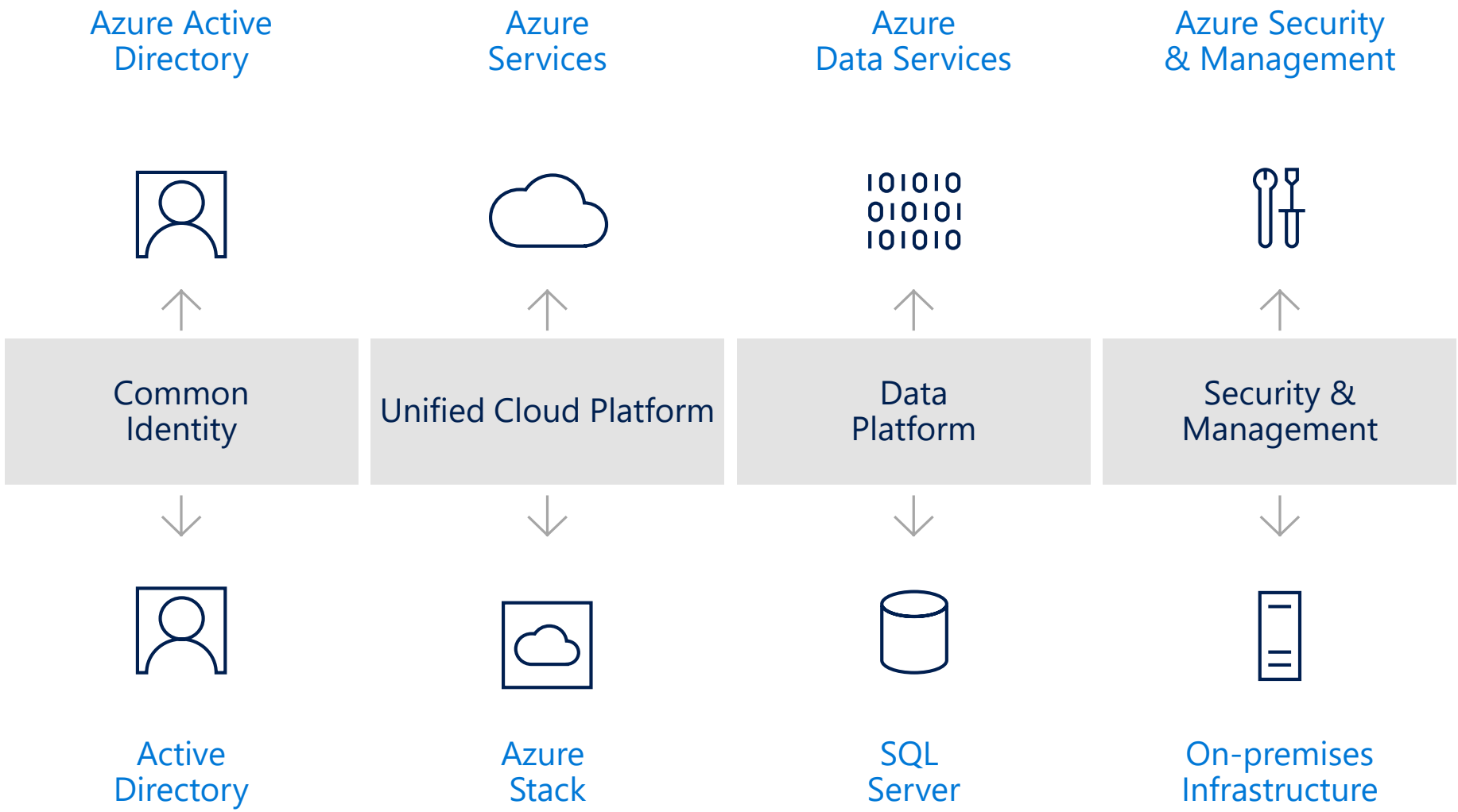
Unified management

Single cloud | Policy

4000+ solutions in Azure Market-Place

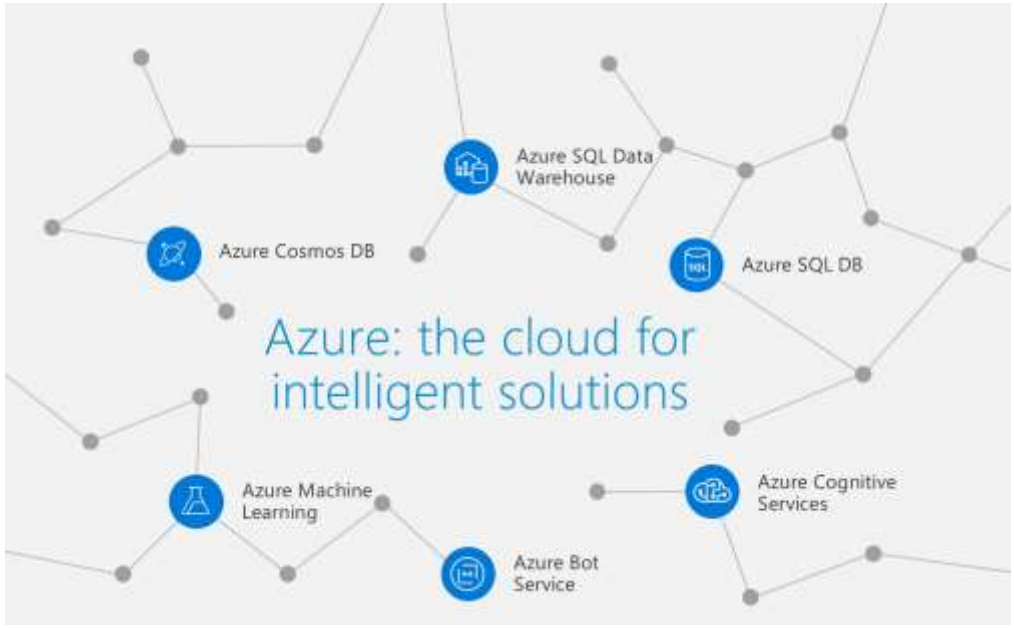
Karl Podesta (@karlpodesta)
kapodest@microsoft.com
https://github.com/azurebigcompute

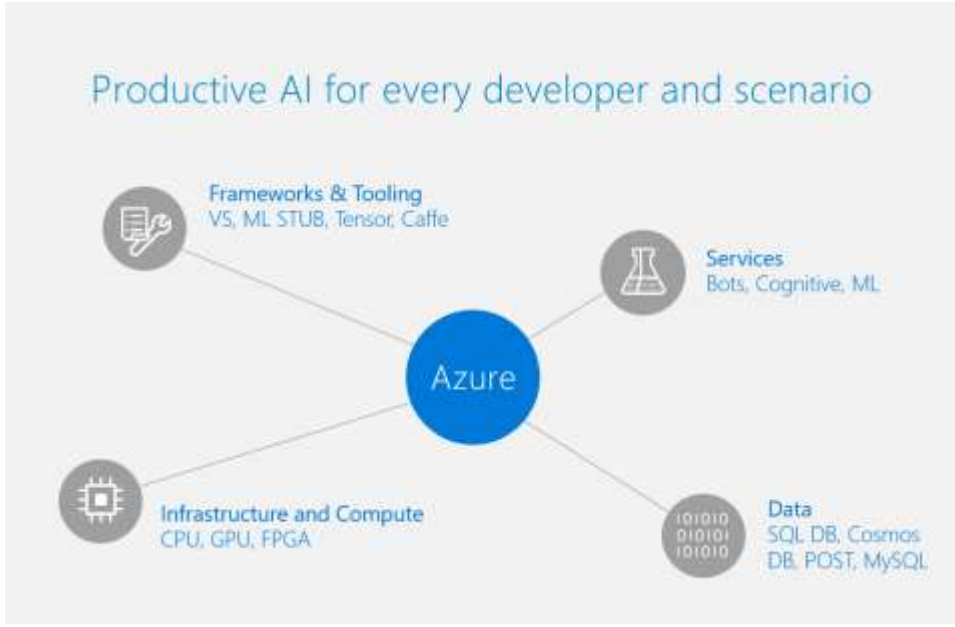
Hybrid – the only consistent, hybrid cloud

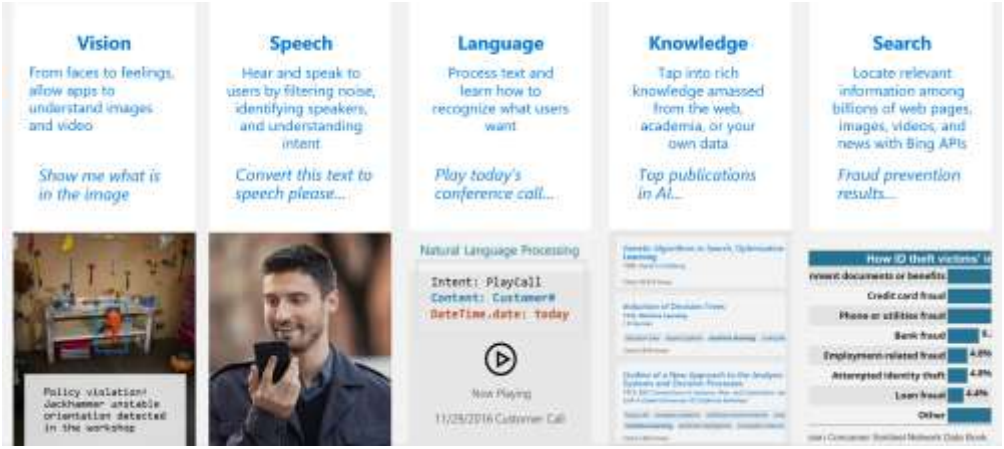


Karl Podesta (@karlpodesta)
kapodest@microsoft.com
https://github.com/azurebigcompute

Intelligent – Data, AI, Frameworks, Services, IoT







Karl Podesta (@karlpodesta)
kapodest@microsoft.com
https://github.com/azurebigcompute

Trusted – 71+ compliance offerings

Global

- ☑ ISO 27001:2013
- ☑ ISO 27017:2015
- ☑ ISO 27018:2014
- ☑ ISO 22301:2012
- ☑ ISO 9001:2015
- ☑ ISO 20000-1:2011
- ☑ SOC 1 Type 2
- ☑ SOC 2 Type 2
- ☑ SOC 3
- ☑ CSA STAR Certification
- ☑ CSA STAR Attestation
- ☑ CSA STAR Self-Assessment
- ☑ WCAG 2.0

US Gov

- ☑ FedRAMP High
- ☑ FedRAMP Moderate
- ☑ EAR
- ☑ DoD DISA SRG Level 5
- ☑ DoD DISA SRG Level 4
- ☑ DoD DISA SRG Level 2
- ☑ DFARS
- ☑ DoE 10 CFR Part 810
- ☑ NIST SP 800-171
- ☑ NIST CSF
- ☑ Section 508 VPATs
- ☑ FIPS 140-2
- ☑ ITAR
- ☑ CJIS
- ☑ IRS 1075

Industry

- ☑ PCI DSS Level 1
- ☑ GLBA
- ☑ FFIEC
- ☑ Shared Assessments
- ☑ FISC (Japan)
- ☑ FCA (UK)
- ☑ MAS + ABS (Singapore)
- ☑ 23 NYCRR 500
- ☑ HIPAA BAA
- ☑ HITRUST
- ☑ 21 CFR Part 11 (GxP)
- ☑ MARS-E
- ☑ NHS IG Toolkit (UK)
- ☑ NEN 7510:2011 (Netherlands)
- ☑ FERPA
- ☑ CDSA
- ☑ MPAA
- ☑ FACT (UK)
- ☑ DPP (UK)
- ☑ SOX

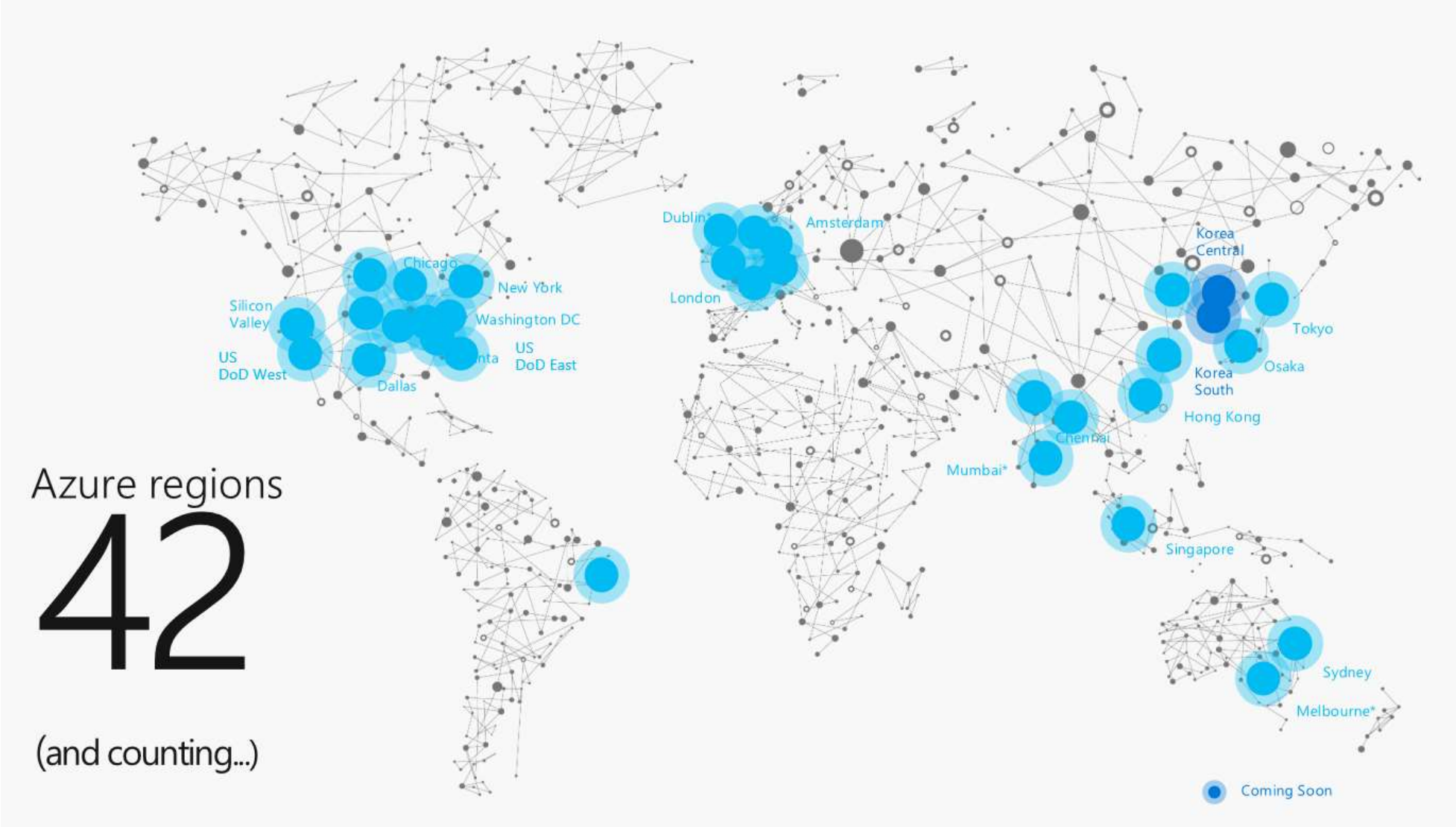
Regional

- ☑ Argentina PDPA
- ☑ Australia CCSL / IRAP
- ☑ Canada Privacy Laws
- ☑ China GB 18030:2005
- ☑ China DJCP (MLPS) Level 3
- ☑ China TRUCS / CCCPPF
- ☑ EN 301 549
- ☑ EU ENISA IAF
- ☑ EU Model Clauses
- ☑ EU – US Privacy Shield
- ☑ Germany IT-Grundschutz workbook
- ☑ Germany C5
- ☑ India MeitY
- ☑ Japan CS Mark Gold
- ☑ Japan My Number Act
- ☑ Netherlands BIR 2012
- ☑ New Zealand Gov CIO Fwk
- ☑ Singapore MTCS Level 3
- ☑ Spain ENS
- ☑ Spain DPA
- ☑ UK Cyber Essentials Plus
- ☑ UK G-Cloud
- ☑ UK PASF

https://aka.ms/AzureCompliance

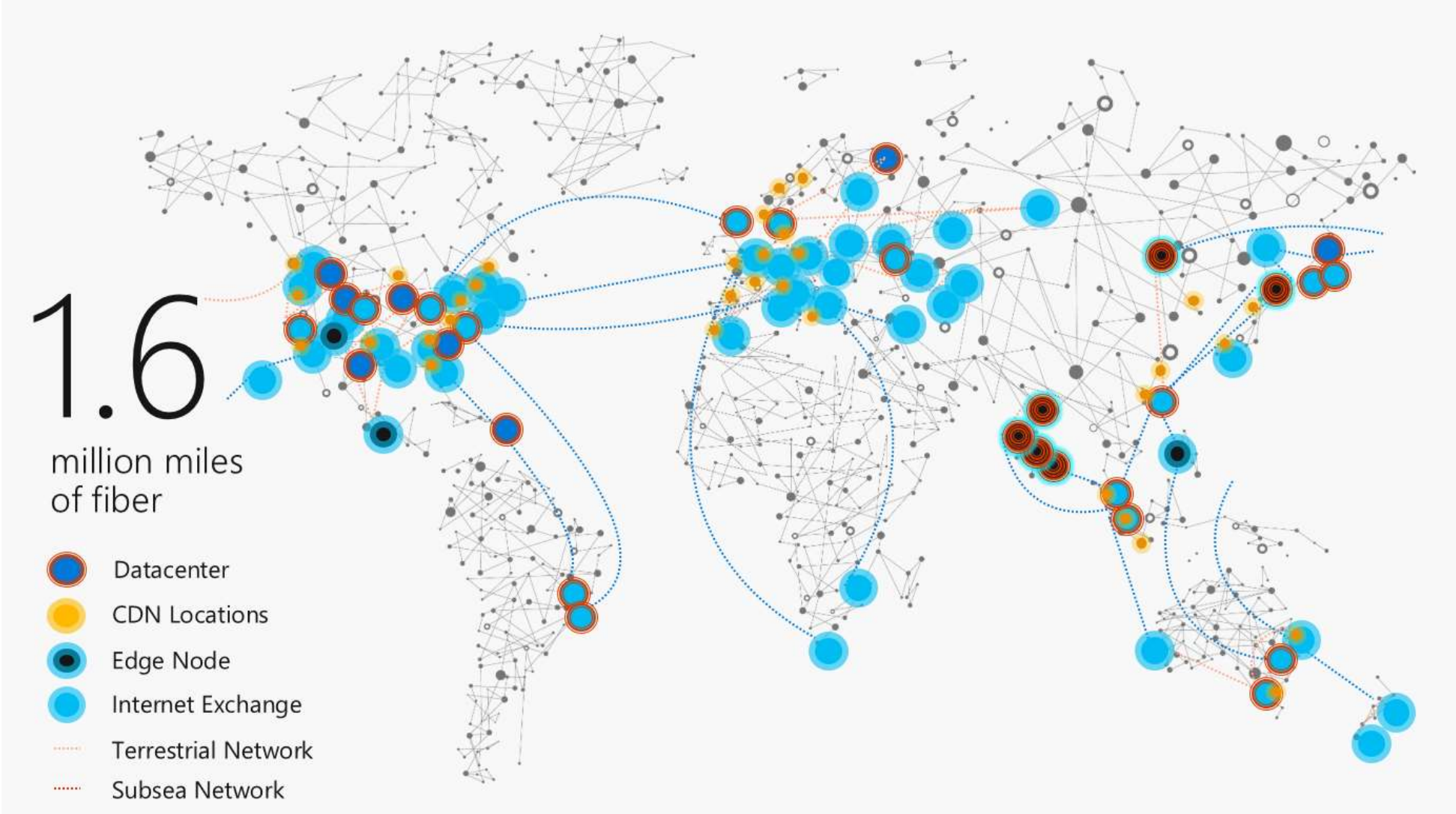
Karl Podesta (@karlpodesta)
kapodest@microsoft.com
https://github.com/azurebigcompute

We are **Global** – more locations than anyone else (& growing)



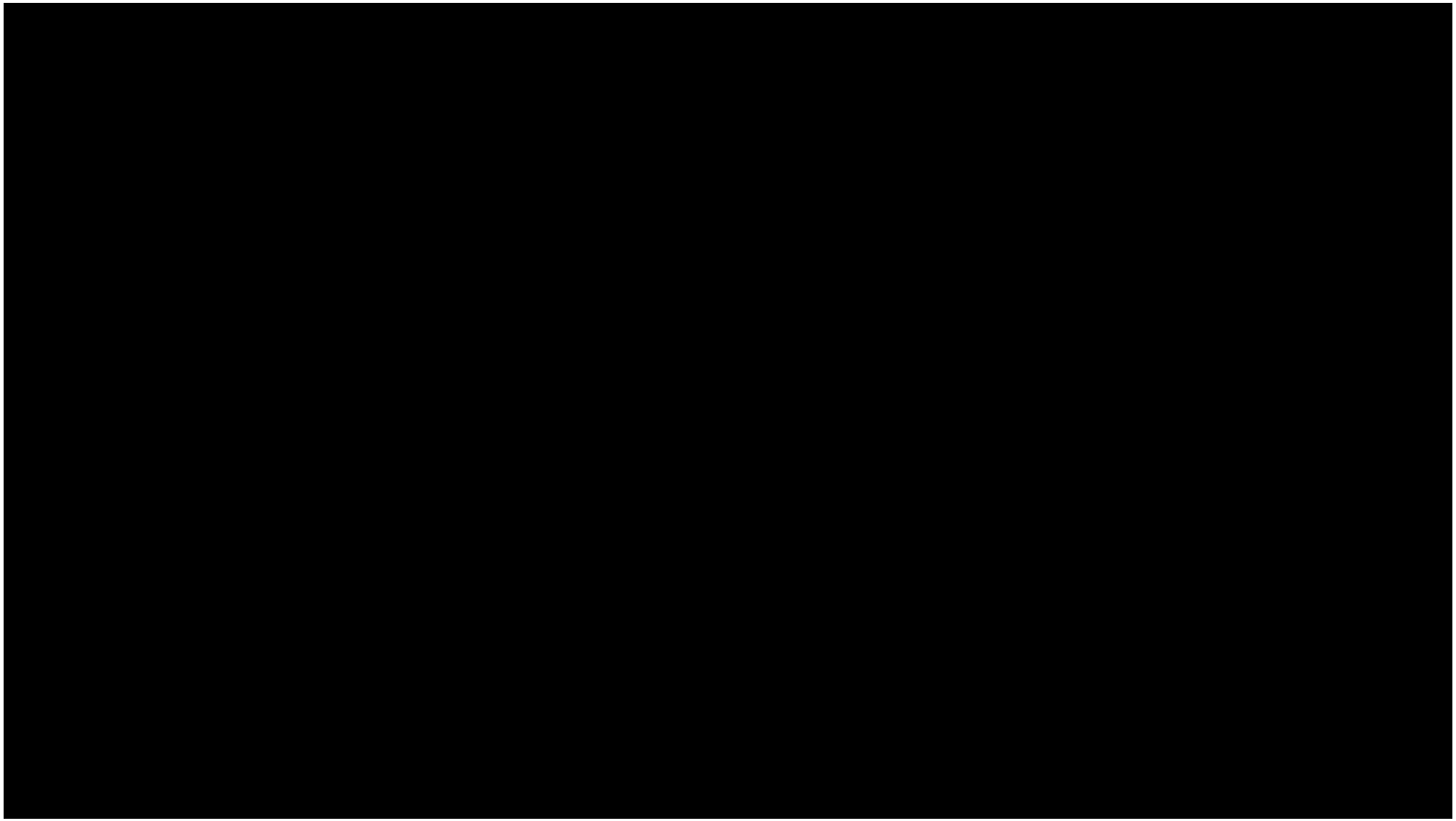
Karl Podesta (@karlpodesta)
kapodest@microsoft.com
https://github.com/azurebigcompute

We are **Global** – more locations than anyone else (& growing)



Karl Podesta (@karlpodesta)
kapodest@microsoft.com
https://github.com/azurebigcompute

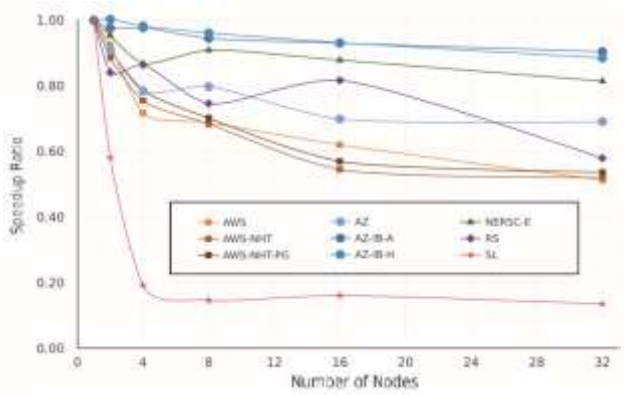
Azure Datacentres video – 1 min 30 sec



Karl Podesta (@karlpodesta)
kapodest@microsoft.com
https://github.com/azurebigcompute

We Scale

Proper supercomputer-class scaling



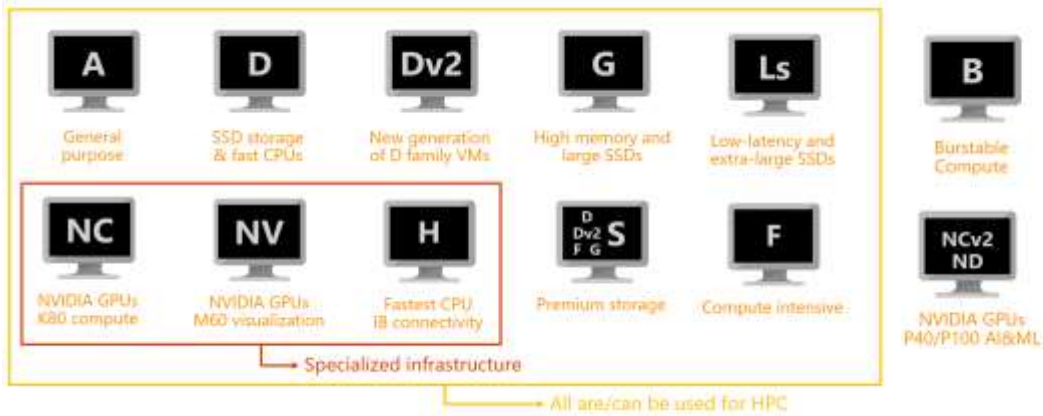
Source: Mohammadi, M., Bazhirov, T. (Feb 9, 2017) Exabyte Inc. "Comparative benchmarking of cloud computing vendors with High Performance Linpack" (https://arxiv.org/pdf/1702.02968.pdf)

Spinning up 16000 A1 VMs on Azure Batch

https://blogs.endjin.com/2015/07/spinning-up-16000-a1-virtual-machines-on-azure-batch/

We Perform

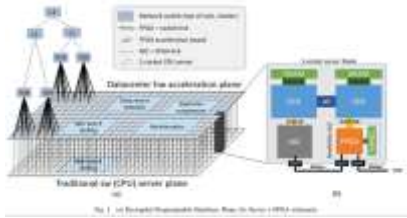
Choose to optimise for CPU, RAM, Disk, Network, Cost



GPUs (NC, NV, ND)

Infiniband Networking

FPGA Networking



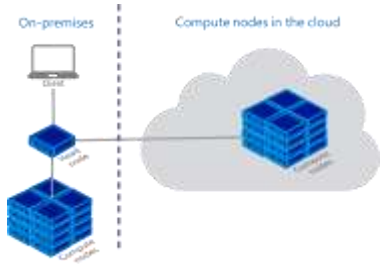
Karl Podesta (@karlpodesta)
kapodest@microsoft.com
https://github.com/azurebigcompute

Solution Overview

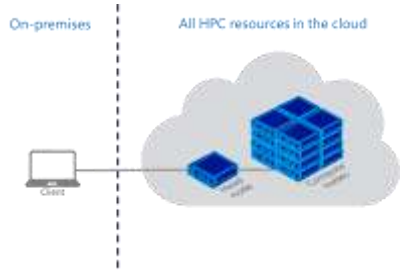
Karl Podesta (@karlpodesta)
kapodest@microsoft.com
https://github.com/azurebigcompute

Self-managed / Hybrid

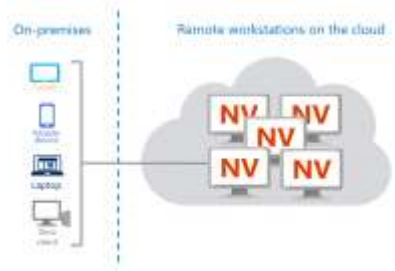
Cloud burst



Cluster on the cloud

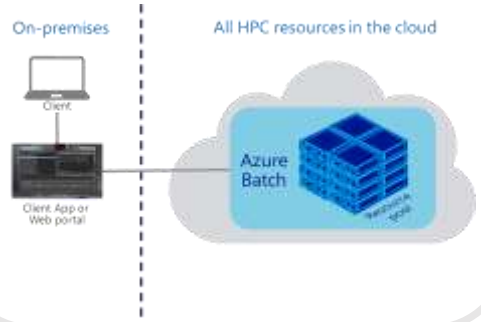


Remote Workstations



Fully-managed

HPC as a service



Add cloud resources to your cluster, on demand

- Burst to cloud to add resources when needed
- Continue using your existing on-premises infrastructure
- Run workloads on Windows and Linux, on Azure and on-premises
- Cover peaks in demand or special projects
- Pay only for what you use
- Microsoft HPC Pack, Univa Grid Engine, and Altair PBS Pro already have this capability (and more are coming!)

Provision one (or more) new clusters in minutes

- Deploy a complete HPC cluster, all in the cloud, in just minutes
- Use templates, scripts, and images to quickly deploy at any scale
- Use your current HPC scheduler
- Shift existing applications, scripts and tools to cloud
- Deploy as many clusters as you need!

Deploy powerful GPU-enabled workstations

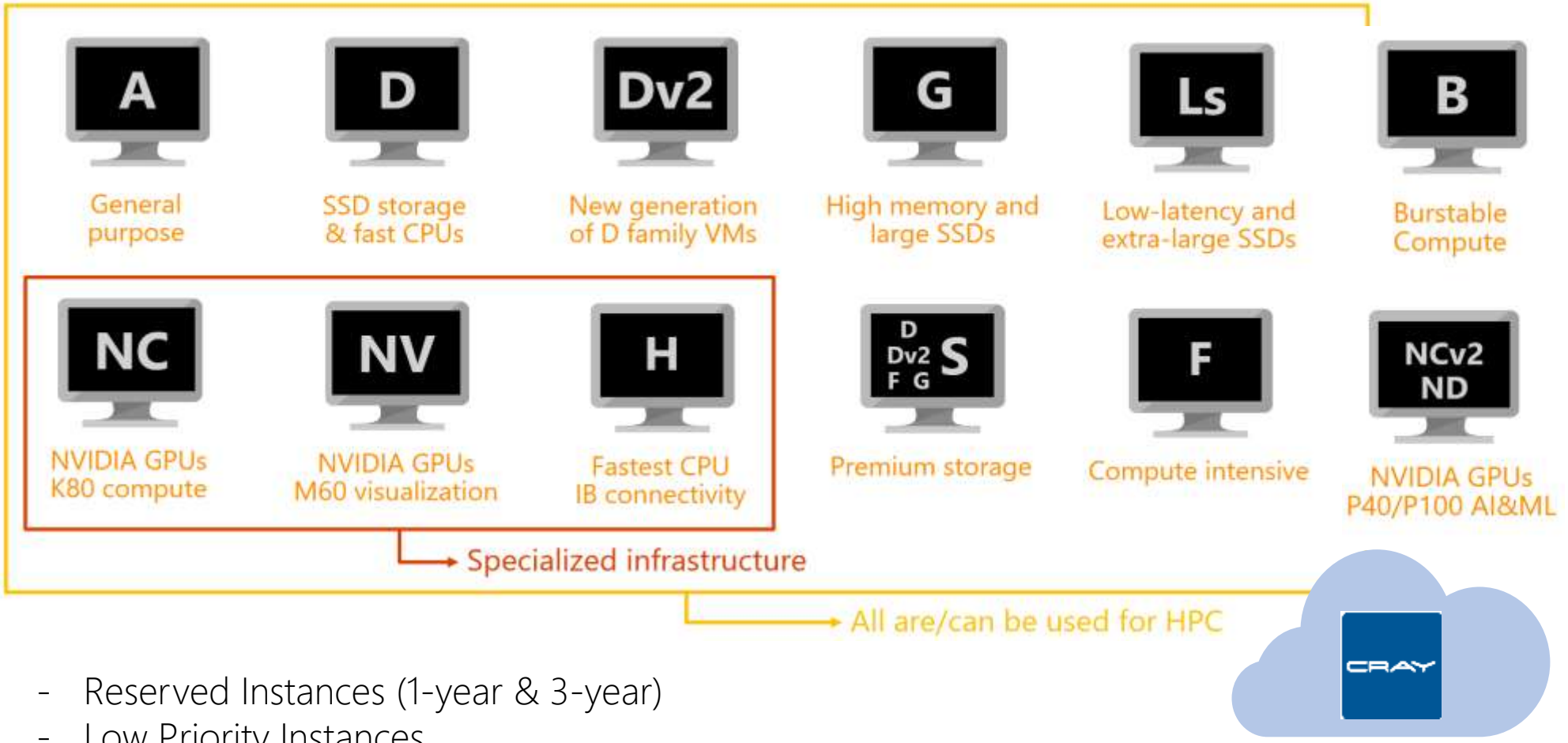
- Deploy one or more GPU-enabled workstations on the cloud
- Connect from any laptop or device (iOS, Android, Windows)
- Collaborate, share, explore.
- With or without optimized remote visualization platforms (e.g., Teradici PCoIP or Citrix HDX)

Run at scale directly from your application

- Integrate with Azure Batch, directly from a client application (GUI or CLI) or online portal
- Batch abstracts resource management and scheduling completely
- Supports small to extremely large deployments and can deploy any VM size
- Provides auto-scaling and stopping of resources
- Run HPC jobs at scale on Docker containers
- Using Batch is free, you only pay for the underlying compute!

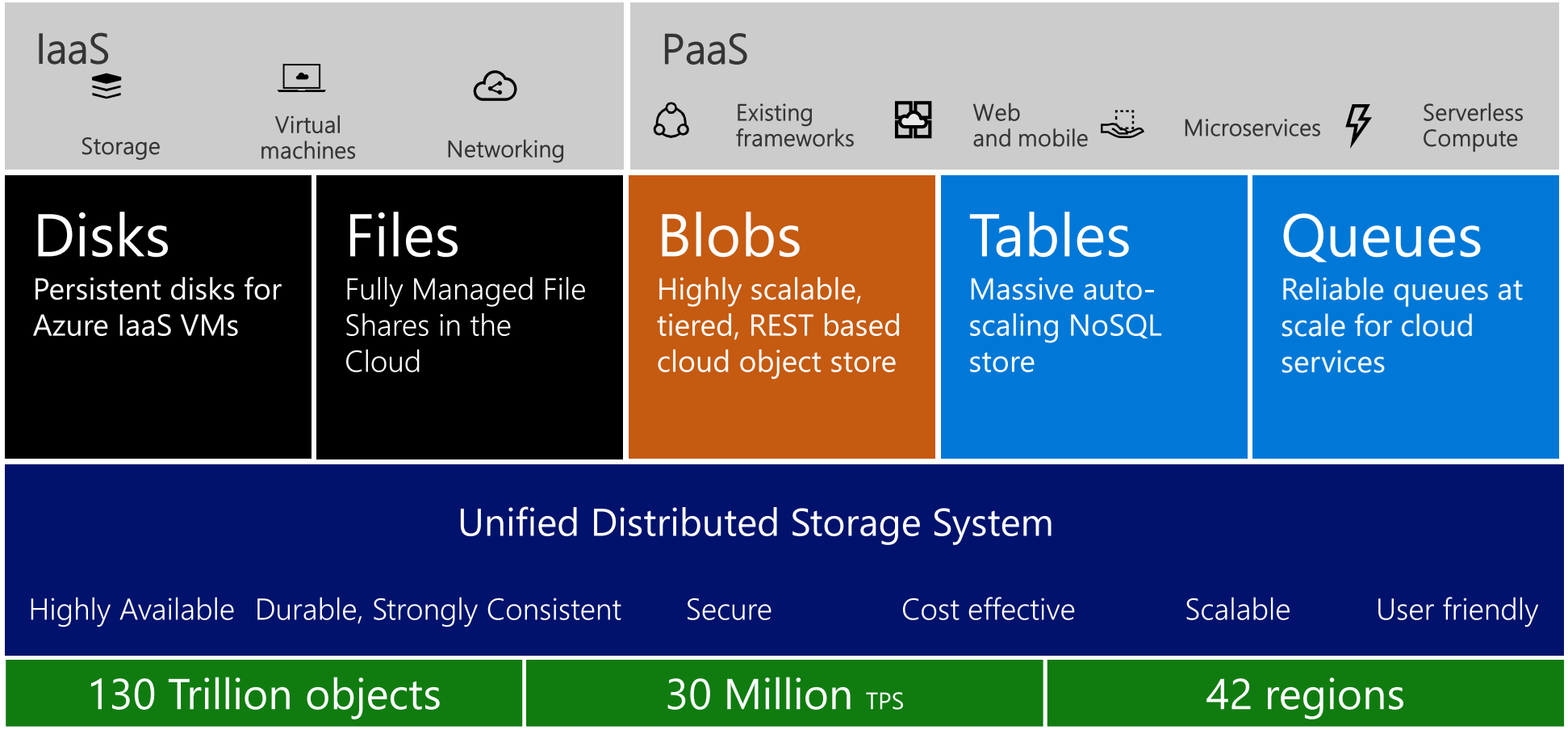
Karl Podesta (@karlpodesta)
kapodest@microsoft.com
https://github.com/azurebigcompute

Azure Compute



- Reserved Instances (1-year & 3-year)
- Low Priority Instances
- Pay as you go Instances

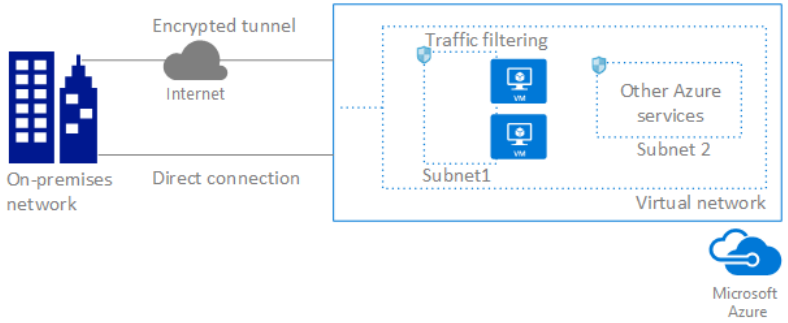
Azure Storage



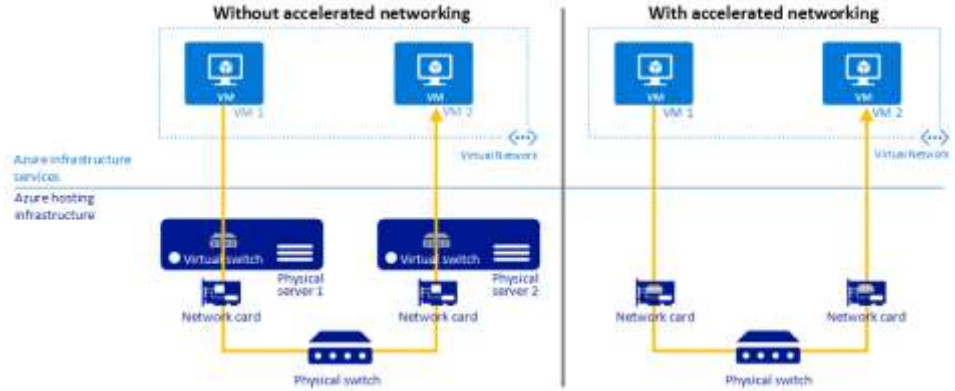
Karl Podesta (@karlpodesta)
kapodest@microsoft.com
https://github.com/azurebigcompute

Azure Networking

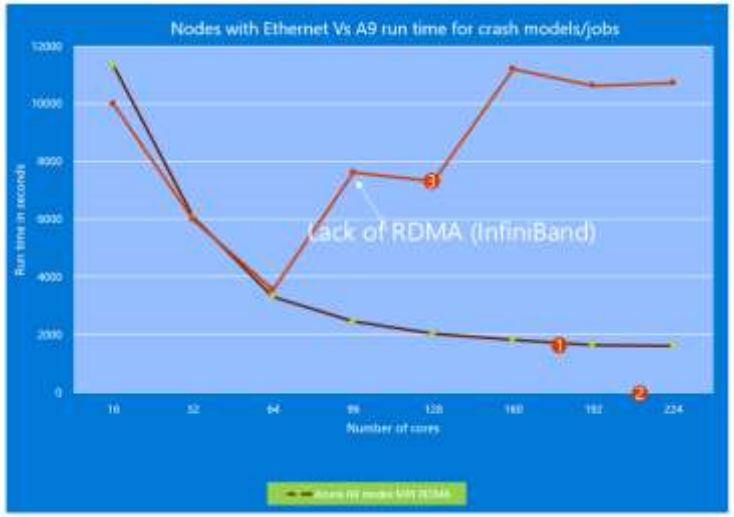
Networking Basics



Accelerated Networking between VMs



Infiniband Networking: true HPC networks



Karl Podesta (@karlpodesta)
kapodest@microsoft.com
https://github.com/azurebigcompute

Karl Podesta (@karlpodesta)
kapodest@microsoft.com
https://github.com/azurebigcompute

Azure all up.... 100+ services

Web & Mobile

- Web Apps
- Mobile Apps
- Logic Apps
- API Apps
- Content Delivery Network
- Media Services
- Search

Container Services

- Azure Container Service
- Azure Container Registry

Databases

- SQL Database
- SQL Data Warehouse
- SQL Server Stretch Database
- DocumentDB
- Redis Cache
- Data Factory

Data & Analytics

- HDInsight
- Machine Learning
- Cognitive Services
- Azure Bot Service*
- Data Lake Analytics
- Power BI Embedded
- Azure Analysis Services

Internet of Things (IoT)

- Azure IoT Hub
- Event Hubs
- Stream Analytics
- Notification Hubs
- BizTalk Services
- Service Bus
- Data Catalog

Security & Identity

- Security Center
- Key Vault
- Azure Active Directory
- B2C
- Domain Services
- Multi-Factor Authentication

Developer Tools

- Visual Studio Team Services
- Azure DevTest Labs
- VS Application Insights
- API Management
- HockeyApp
- Developer Tools
- Service Profiler*

Monitoring & Management

- Azure Portal
- Azure Resource Manager
- Azure Advisor
- Azure Monitor
- Log Analytics
- Automation
- Scheduler

Karl Podesta (@karlpodesta)
kapodest@microsoft.com
https://github.com/azurebigcompute

Build in intelligence with AI: our Cognitive Services API

Vision

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

Show me what is in the image

Speech

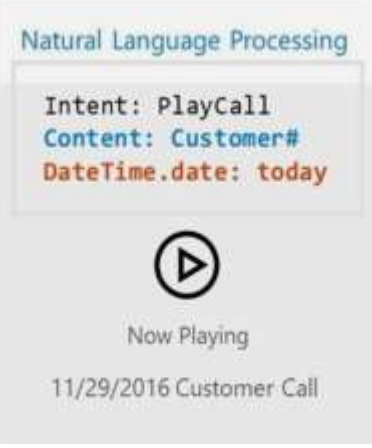
Hear and speak to users by filtering noise, identifying speakers, and understanding intent

Convert this text to speech please...

Language

Process text and learn how to recognize what users want

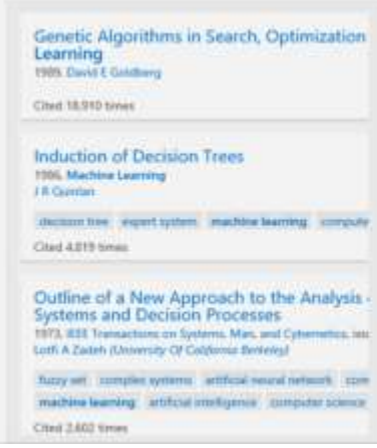
Play today's conference call...



Knowledge

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

Top publications in AI...



Search

Locate relevant information among billions of web pages, images, videos, and news with Bing APIs

Fraud prevention results...

Simplify IoT with Pre-configured solutions

Azure IoT Suite

- Device connectivity & management
- Data ingestion and command & control
- Stream processing & predictive analytics
- Workflow automation and integration
- Dashboards and visualization
- Preconfigured solutions
  - Remote monitoring
  - Predictive maintenance
  - Connected factory (new)

Karl Podesta (@karlpodesta)
kapodest@microsoft.com
https://github.com/azurebigcompute

Karl Podesta (@karlpodesta)
kapodest@microsoft.com
https://github.com/azurebigcompute

Doing it together: our partners

Karl Podesta (@karlpodesta)
kapodest@microsoft.com
https://github.com/azurebigcompute

We Partner and we Work Together

Choose OS

BYOScheduler

BYO Parallel FS

BYO Platform

CYCLE COMPUTING

BROAD INSTITUTE

- 2 weeks, from concept to completion
- Accessing 50,000 cores
- 30 years of compute in 6 hours for less than $5,000

"Azure basically gives us a solid, rock solid performance - no questions asked."

—Rolf Seuster: HEPnet Technical Manager
University of Victoria

"We believe that the Siemens teamplay framework and Microsoft Azure platform offer the potential to fundamentally change the way healthcare services are delivered."

—Robert Day: Chief Operating Officer
Zwanger – Pesiri Radiology

"We can spin up virtual machines when we need them, and we don't have to pay for them otherwise. Pay-as-you-go is really important to minimizing our costs."

—Jussi Mattila: Head of Research and Development
Combinostics

Karl Podesta (@karlpodesta)
kapodest@microsoft.com
https://github.com/azurebigcompute

What have others done?
Our Customers

Karl Podesta (@karlpodesta)
kapodest@microsoft.com
https://github.com/azurebigcompute

Bar Ilan Experiment – Cryptanalysis Research

Insurance – modelling complex natural disasters with HPC

Global insurance firm models complex natural disasters with cloud-based HPC

AXA Global P&C manages reinsurance programs for the AXA Group, a global insurance provider based in Paris, France. To create complex catastrophe models for floods and other natural disasters, a team of actuaries created a high-performance computing (HPC) solution based on the Microsoft Azure platform and Azure HPC Pack. Now, AXA Group can improve insurance services with more accurate, detailed information about events ranging from floods to hurricanes.

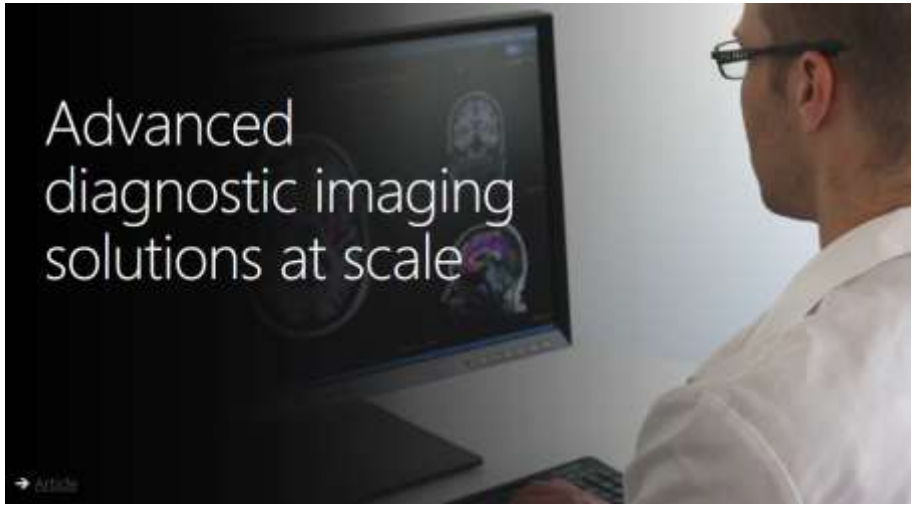
AXA GLOBAL P&C
redefining / standards

Products and Services	Organization Size	Industry	Country	Business Need
Azure Batch Azure Storage Azure Virtual Machines Microsoft HPC Pack	165,000 employees	Insurance	France	Business agility

Karl Podesta (@karlpodesta)
kapodest@microsoft.com
https://github.com/azurebigcompute

Karl Podesta (@karlpodesta)
kapodest@microsoft.com
https://github.com/azurebigcompute

Advancing Scientific Research



Karl Podesta (@karlpodesta)
kapodest@microsoft.com
https://github.com/azurebigcompute

Our Value Add: Microsoft Research

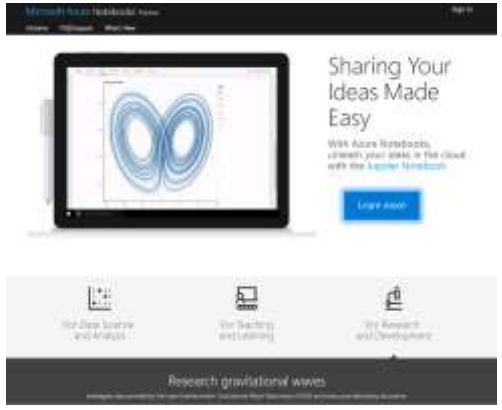
Karl Podesta (@karlpodesta)
kapodest@microsoft.com
https://github.com/azurebigcompute

We **Research** and we **Add Value**

Grants for non-profit research

Microsoft Genomics Service

Azure Notebooks



We will bring Quantum to everyone

Top contributors on GitHub; Solutions/Recipes shared

Fundamentals

- Azure IaaS
- HPC IaaS
- HPC PaaS
- HPC SaaS

Karl Podesta (@karlpodesta)
kapodest@microsoft.com
https://github.com/azurebigcompute

Pre-Requisites

Getting Access:

- Azure Pass – https://www.microsoftazurepass.com
- Free Account – https://azure.microsoft.com/free
- OR – your own company's Azure account

Ways to Access:

- Azure Portal – https://portal.azure.com
- Azure CLI – Windows/Linux/Mac – az cli
- PowerShell
- Code: SDK (Java, C#, Python, etc)

Changing core quota:

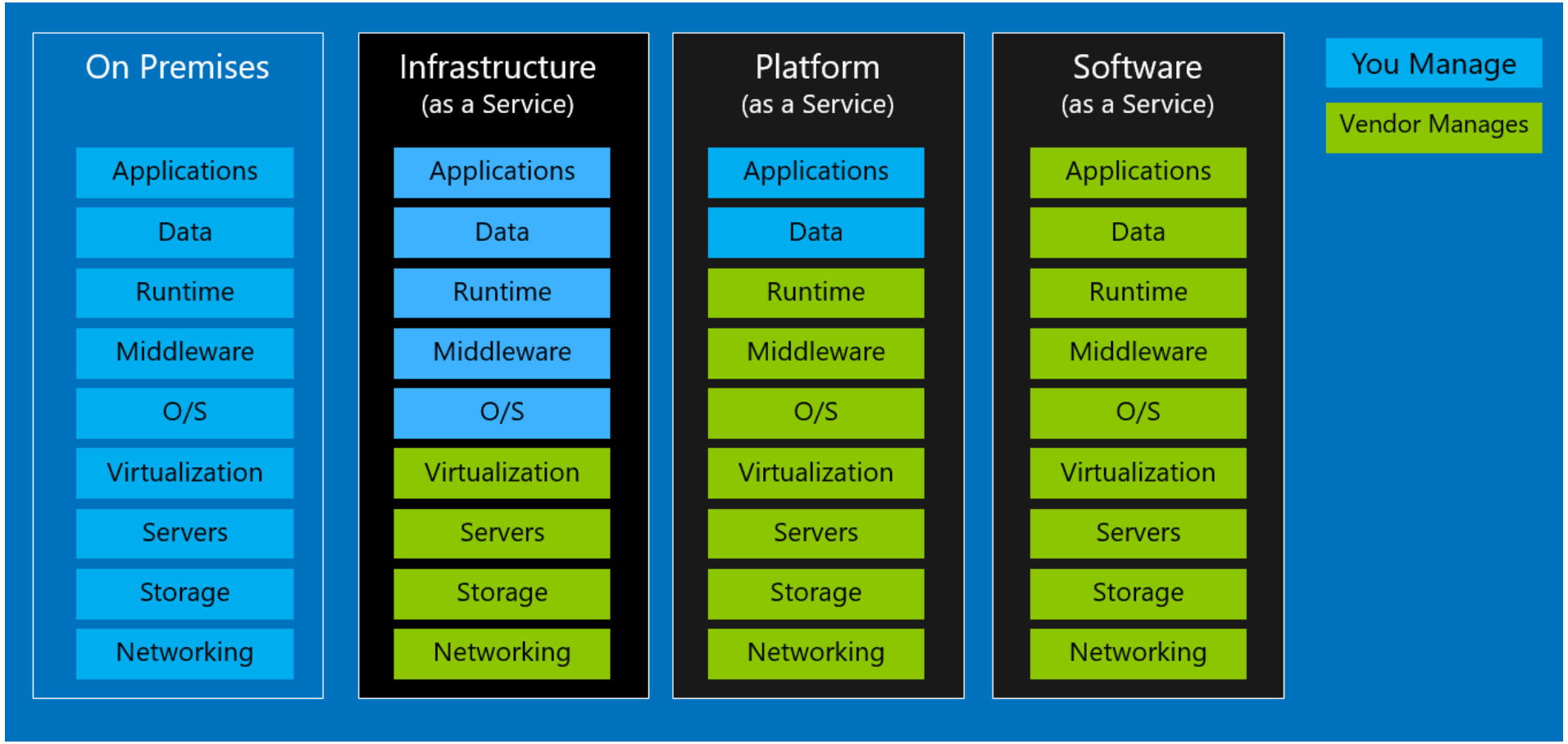
- Raise a support ticket in Azure portal

Karl Podesta (@karlpodesta)
kapodest@microsoft.com
https://github.com/azurebigcompute

Azure IaaS

Karl Podesta (@karlpodesta)
kapodest@microsoft.com
https://github.com/azurebigcompute

Level Set: Cloud Basics



Karl Podesta (@karlpodesta)
kapodest@microsoft.com
https://github.com/azurebigcompute

Level Set: Azure Basics

EA – Accounts – Subscriptions – Resource Groups

Compute – VMs also called Instances

Network – VNets, Network Security Groups (NSGs)

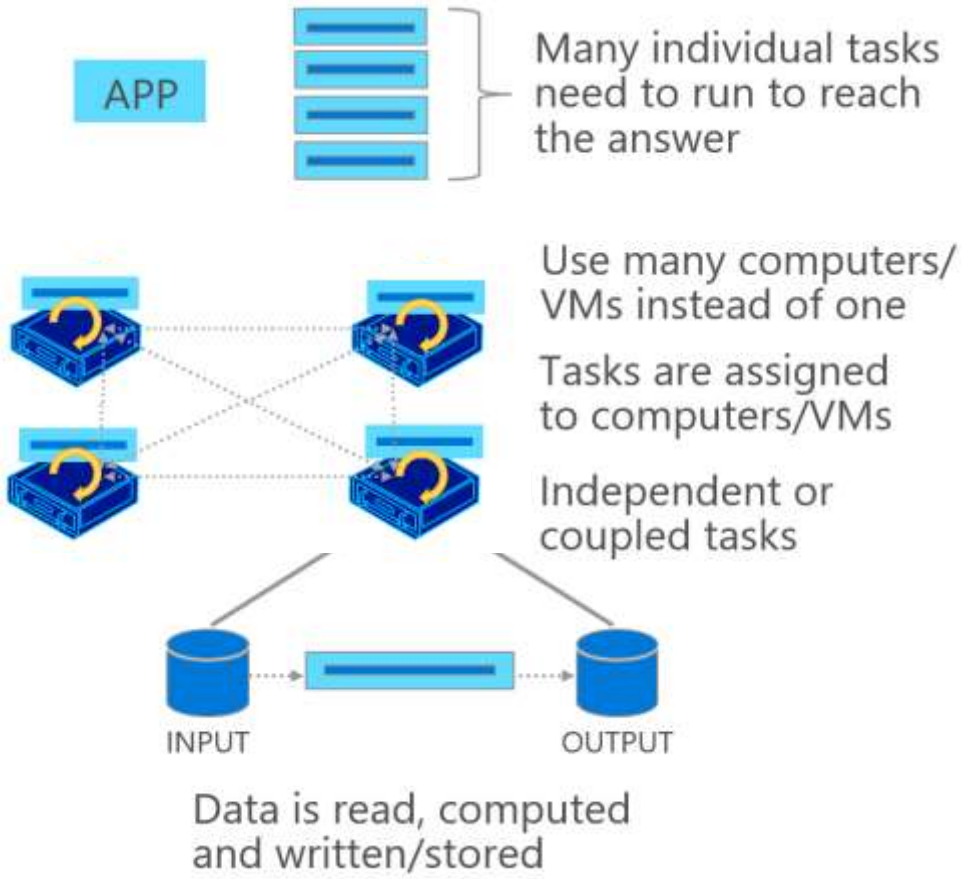
Storage – Blobs, Files, Disks

Permissions

Karl Podesta (@karlpodesta)
kapodest@microsoft.com
https://github.com/azurebigcompute

Level Set: HPC Basics

- HPC Clusters
  - Head node
  - Compute nodes
  - Login nodes
  - Visualisation nodes
  - Storage nodes
  - Networking
- Jobs
- Queues
- Licenses



Microsoft

Cloud Workshops 2018
HPC – Big Compute

INTRODUCTION

FUNDAMENTALS

Azure IaaS

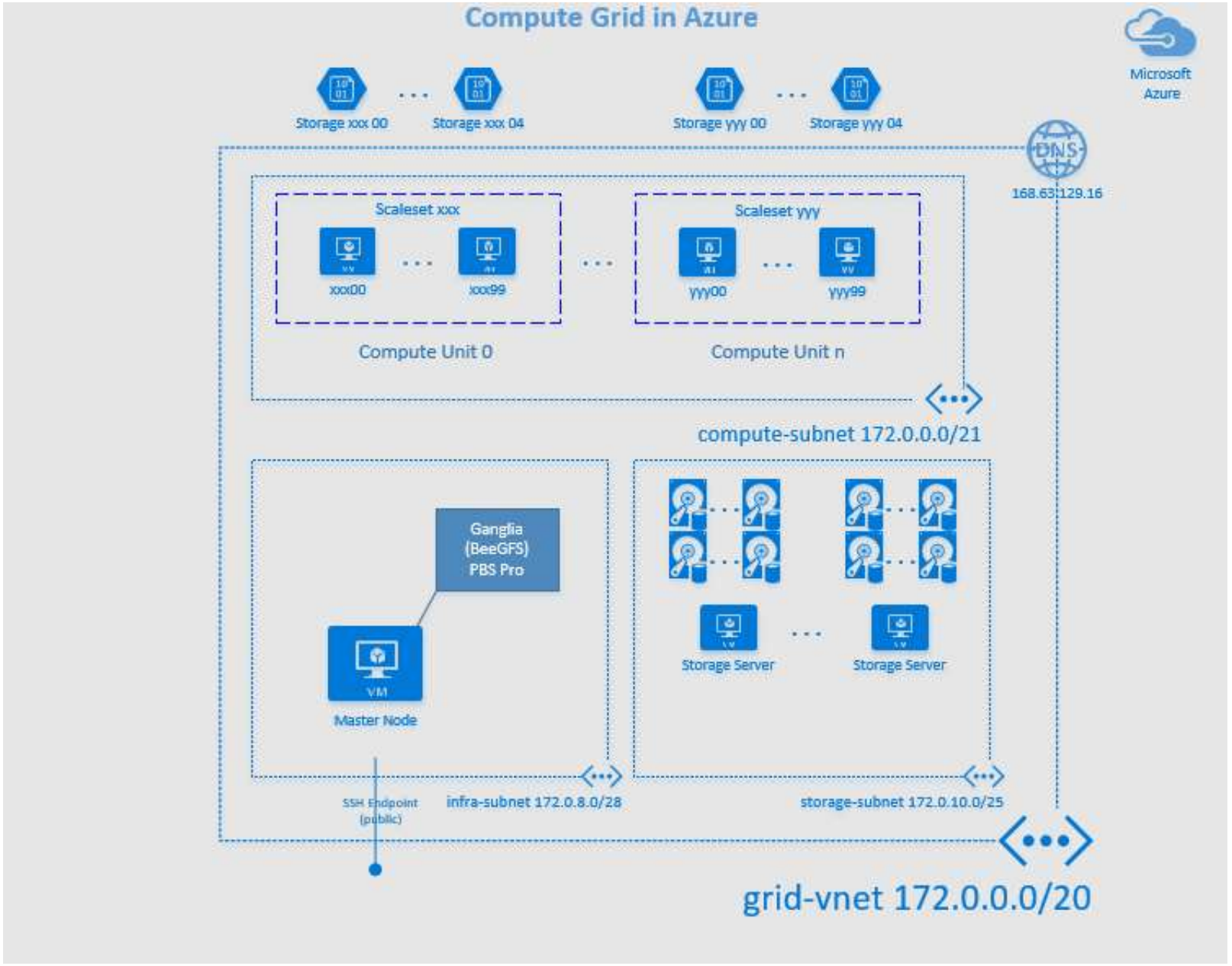
HPC IaaS

HPC PaaS

HPC SaaS

SPECIFICS

Karl Podesta (@karlpodesta)
kapodest@microsoft.com
https://github.com/azurebigcompute



Karl Podesta (@karlpodesta)
kapodest@microsoft.com
https://github.com/azurebigcompute

Logging in & looking around:

Connecting to Azure:

```
az login
```

Choosing to work inside a chosen subscription:

```
az account list
az account set -s <subscription id here>
```

Listing VM images:

```
az vm image list --publisher OpenLogic --all | grep HPC
az vm image list --publisher OpenLogic --all | grep CentOS
az vm image list --publisher Microsoft
```

https://github.com/azurebigcompute/BigComputeLabs/tree/master/Azure_Iaas

Karl Podesta (@karlpodesta)
kapodest@microsoft.com
https://github.com/azurebigcompute

My first VM

Making a Resource Group:

```
az group create -n hpclab -l westeurope -o table
```

Making a VM:

```
az vm create --name golden01 --resource-group hpclab --image
OpenLogic:CentOS:7.4:7.4.20180118 --size Standard_A4 --storage-
sku Standard_LRS --generate-ssh-keys -o table

(or)

az vm create --name golden02 --resource-group hpclab --image
OpenLogic:CentOS-HPC:7.1:7.1.20170608 --size Standard_H16r --
storage-sku Standard_LRS --generate-ssh-keys -o table
```

Showing (& connecting to) a VM:

```
az vm list-ip-addresses -o table
ssh <ip address>
```

https://github.com/azurebigcompute/BigComputeLabs/tree/master/Azure_Iaas

Karl Podesta (@karlpodesta)
kapodest@microsoft.com
https://github.com/azurebigcompute

My first VM Scale Set (VMSS)

Making a scale set:

```
az vmss create --name lab-vmss --resource-group hpclab --image
OpenLogic:CentOS-HPC:7.1:7.1.20170608 --vm-sku Standard_H16r --storage-
sku Standard_LRS --instance-count 2 --generate-ssh-keys --disable-
overprovision
```

Showing a scale set:

```
az vmss list-instance-connection-info --name lab-vmss --resource-group
hpclab
```

Connecting to the scale set:

```
ssh <ip-address> -p 50000
```

Growing/Shrinking:

```
az vmss scale --new-capacity 4 --name lab-vmss --resource-group hpclab
-o table
```

https://github.com/azurebigcompute/BigComputeLabs/tree/master/Azure_Iaas

Karl Podesta (@karlpodesta)
kapodest@microsoft.com
https://github.com/azurebigcompute

Testing Infiniband (RDMA/MPI)

Set your environment:

```
export INTELMPI_ROOT=/opt/intel/impi/2017.2.174
export I_MPI_FABRICS=shm:dapl
export I_MPI_DAPL_PROVIDER=ofa-v2-ib0
export I_MPI_ROOT=/opt/intel/compilers_and_libraries_2017.2.174/linux/mpi
source /opt/intel/impi/2017.2.174/bin64/mpivars.sh
```

Provision to arrange SSH between nodes:

```
provision.sh
```

Run a “ping-pong” test:

```
mpirun --hosts node1,node2 IMB-MPI1 pingpong
```

https://github.com/azurebigcompute/BigComputeLabs/tree/master/Azure_Iaas

Karl Podesta (@karlpodesta)
kapodest@microsoft.com
https://github.com/azurebigcompute

Cleaning up… (deleting)

Deleting a VM:

```
az vm delete --name golden01 --resource-group hpclab -o table
az vm list-ip-addresses -o table
```

Deleting a VM Scale Set:

```
az vmss delete --name lab-vmss --resource-group hpclab -o table
```

Deleting a Resource Group:

```
az group delete --name lab-vmss
```

https://github.com/azurebigcompute/BigComputeLabs/tree/master/Azure_Iaas

Karl Podesta (@karlpodesta)
kapodest@microsoft.com
https://github.com/azurebigcompute

Something a bit more packaged?

Try the HPC Lab from our friends at Microsoft Research:

- Create a SLURM cluster using an Azure Resource Manager template
- Copy local resources to a SLURM cluster
- Remote into a SLURM cluster
- Run jobs on a SLURM cluster
- Start and stop nodes in a SLURM cluster
- Use the Azure Resource Manager to delete a SLURM cluster

https://github.com/MSRConnections/Azure-training-course/tree/master/Content

https://github.com/MSRConnections/Azure-training-course/blob/master/Content/High-Performance%20Computing/SLURM%20Linux%20Cluster%20HOL.md

Karl Podesta (@karlpodesta)
kapodest@microsoft.com
https://github.com/azurebigcompute

HPC IaaS

Microsoft

Cloud Workshops 2018
HPC – Big Compute

INTRODUCTION

FUNDAMENTALS

Azure IaaS

HPC IaaS

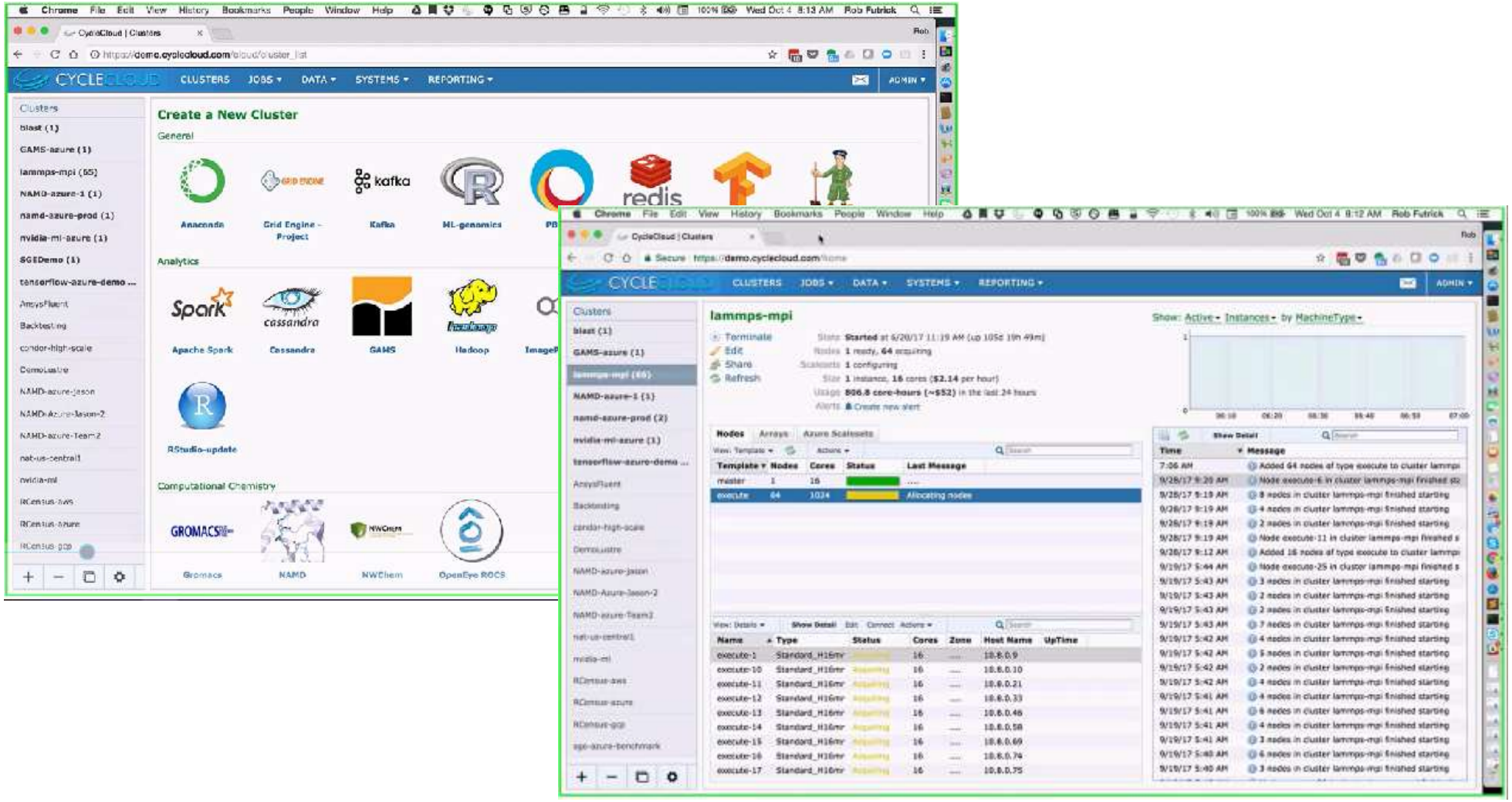
HPC PaaS

HPC SaaS

SPECIFICS

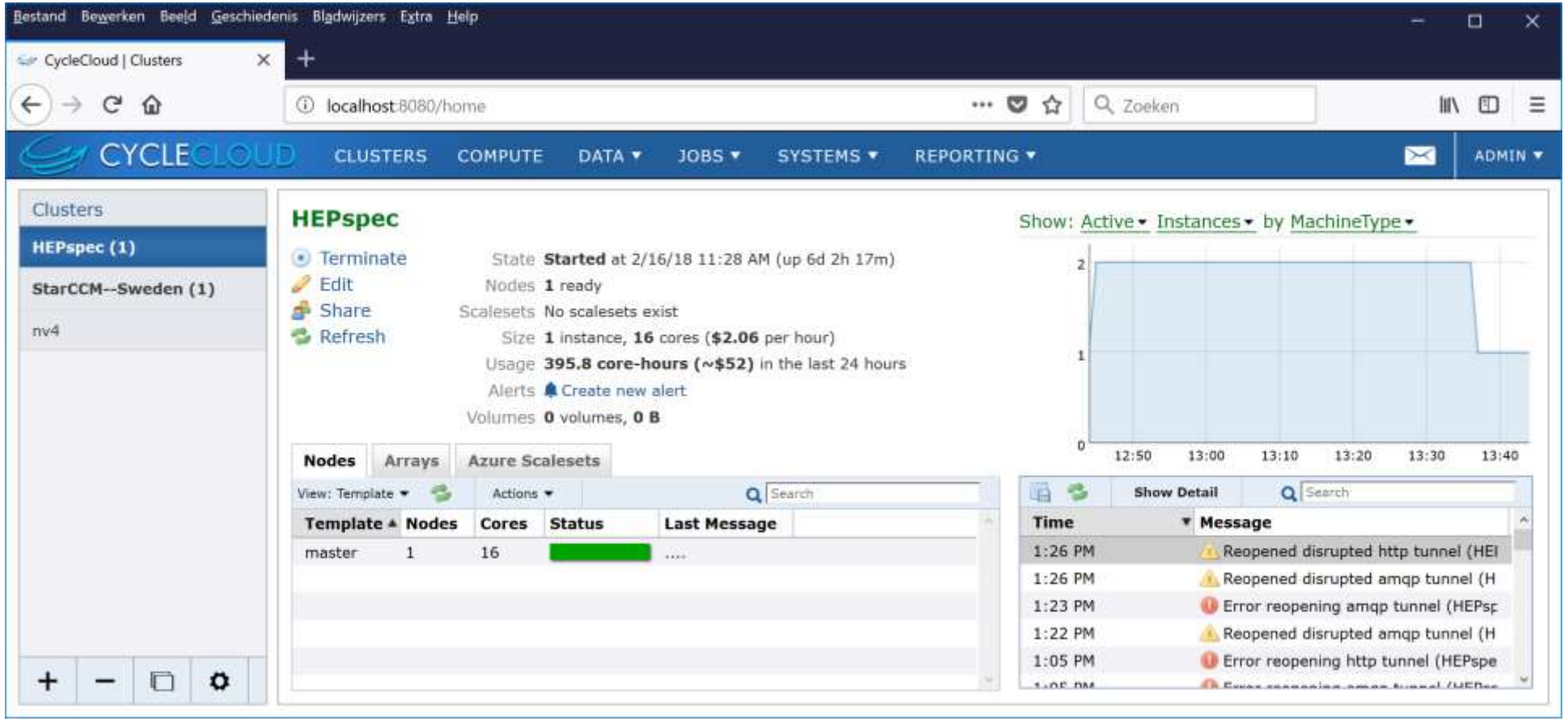
Karl Podesta (@karlpodesta)
kapodest@microsoft.com
https://github.com/azurebigcompute

CycleCloud – a HPC cluster for every workload



Karl Podesta (@karlpodesta)
kapodest@microsoft.com
https://github.com/azurebigcompute

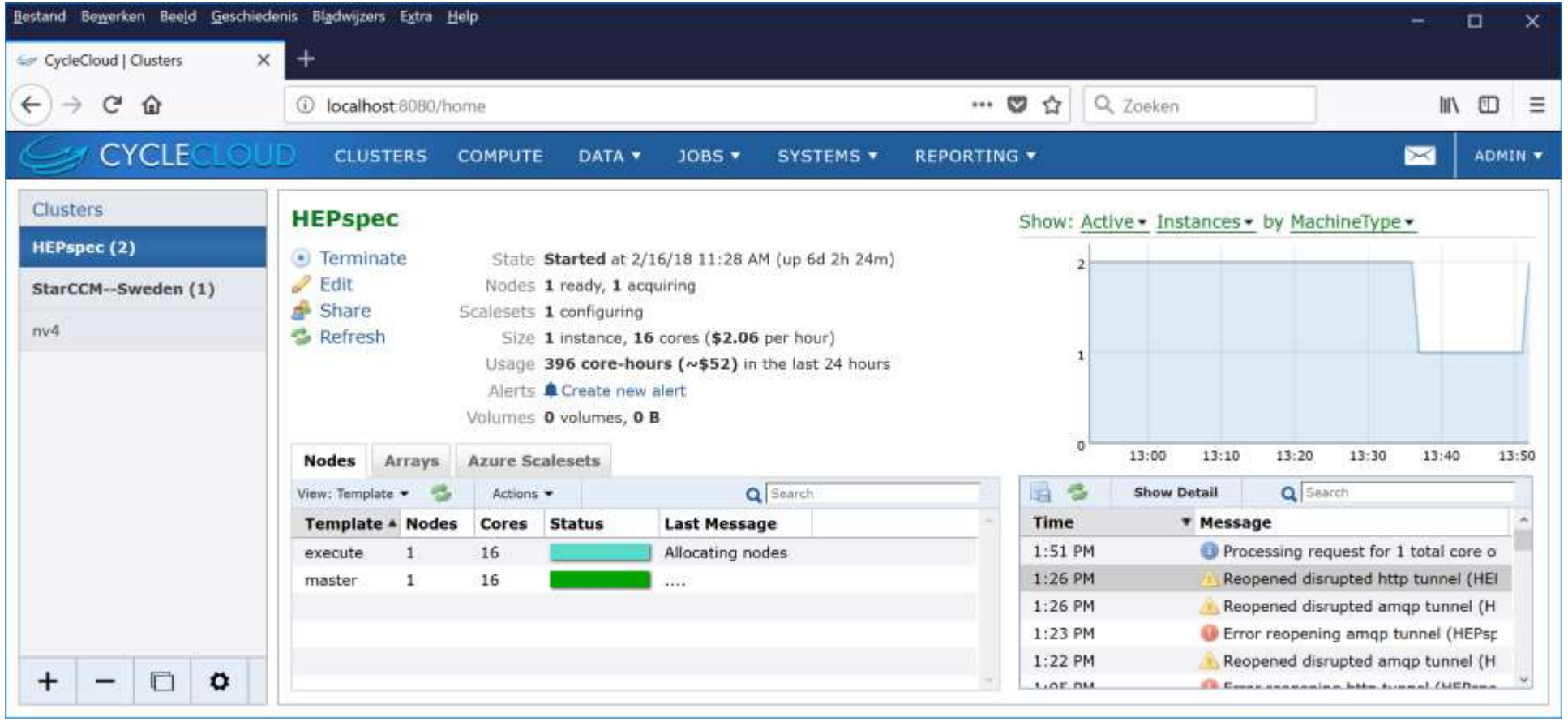
CycleCloud




```
cyclecloud connect master -c HEPspec
sudo su - cluster.user
qsub testjob.sh
```

Karl Podesta (@karlpodesta)
kapodest@microsoft.com
https://github.com/azurebigcompute

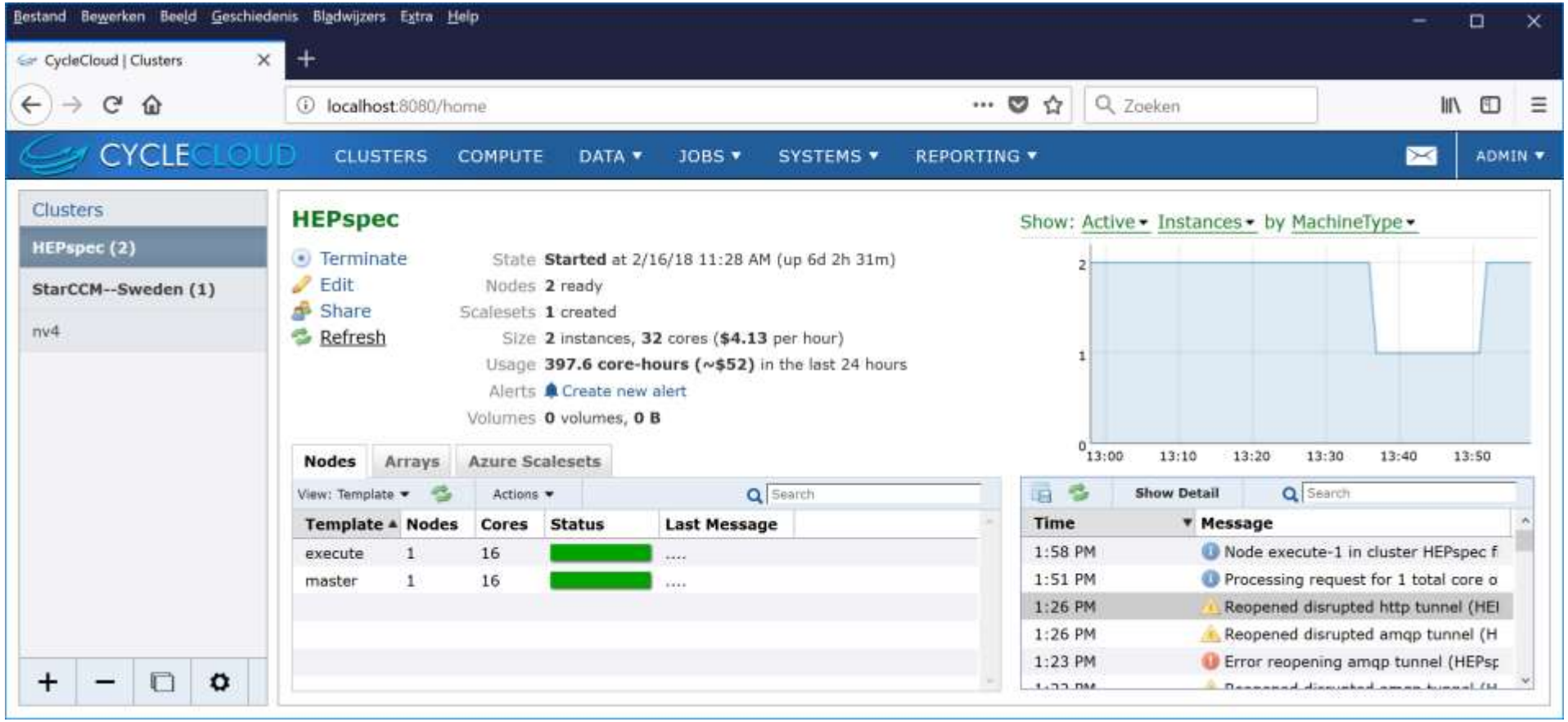
CycleCloud




```
[cluster.user@ip-0A000008 ~]$ qstat
job-ID  prior   name       user         state submit/start at     queue                          slots ja-task-ID
-----------------------------------------------------------------------------------------------------------------
2 0.00000 testjob.sh cluster.user qw    02/22/2018 12:36:48                                        1
```

Karl Podesta (@karlpodesta)
kapodest@microsoft.com
https://github.com/azurebigcompute

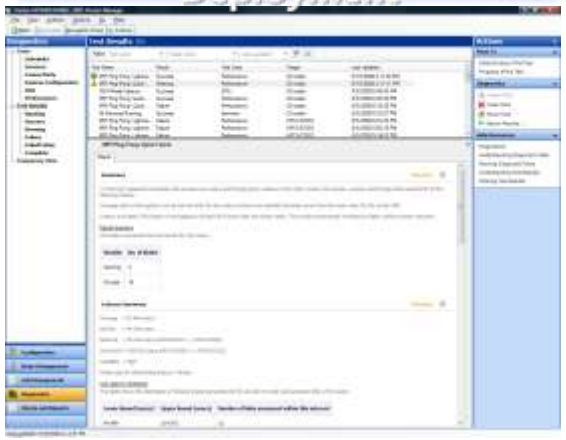
CycleCloud



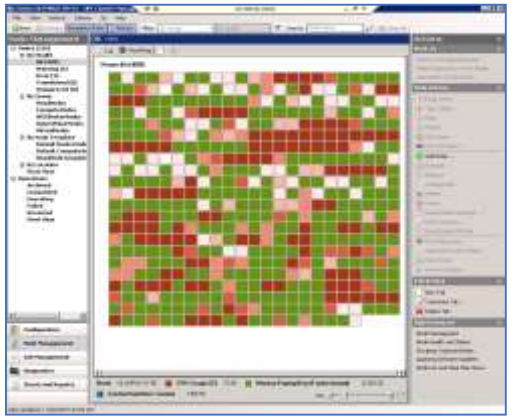

```
[cluster.user@ip-0A000008 ~]$ qstat
[cluster.user@ip-0A000008 ~]$ qhost
HOSTNAME                ARCH         NCPU  LOAD  MEMTOT  MEMUSE  SWAPTO  SWAPUS
----------------------------------------------------------------------------------
global                  -               -     -       -       -       -       -
ip-0A000005             linux-x64      16  0.49  110.2G 1011.7M  500.0M     0.0
ip-0A000008             linux-x64      16  0.30  110.2G    1.2G  500.0M     0.0
```

HPC Pack – traditional scheduler & resource management

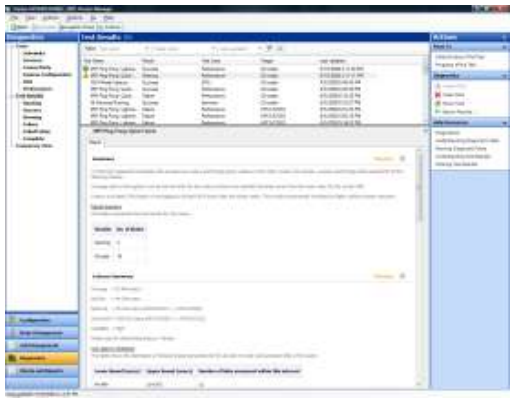
Configuration & Deployment



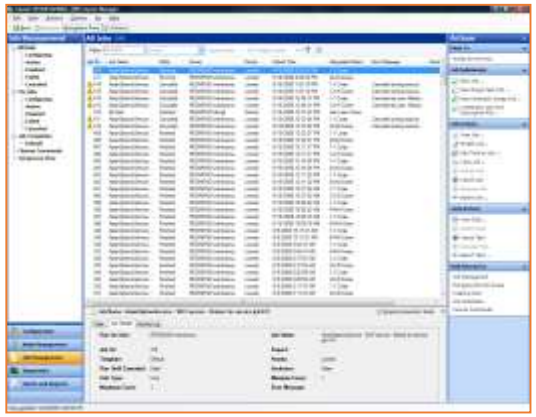
Monitoring



Node Management

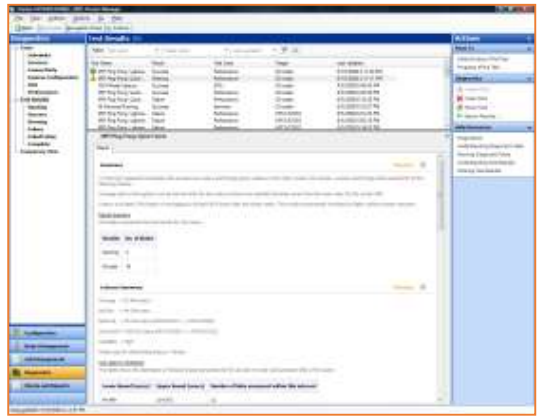


Job Management



Reporting

Diagnostics



Karl Podesta (@karlpodesta)
kapodest@microsoft.com
https://github.com/azurebigcompute

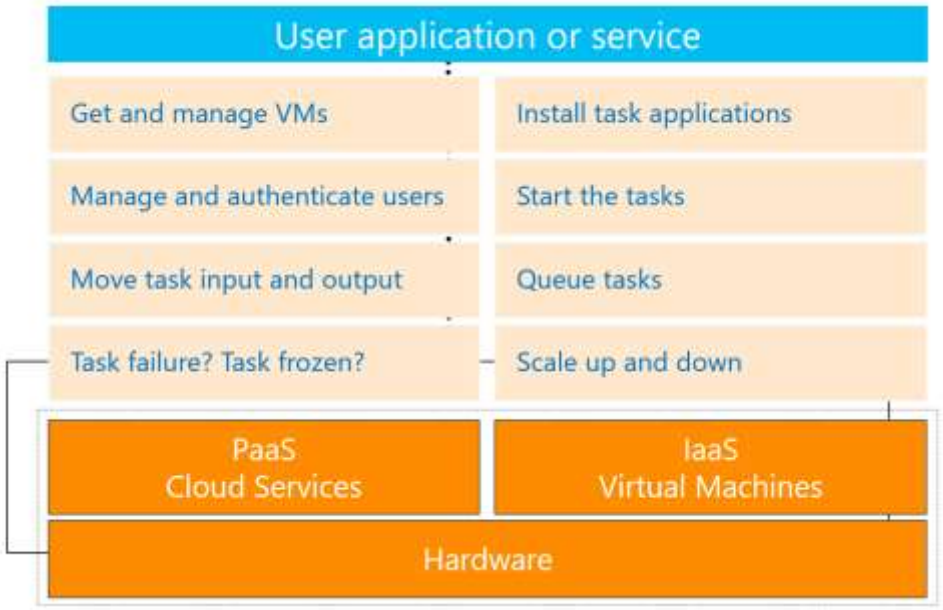
Karl Podesta (@karlpodesta)
kapodest@microsoft.com
https://github.com/azurebigcompute

HPC PaaS

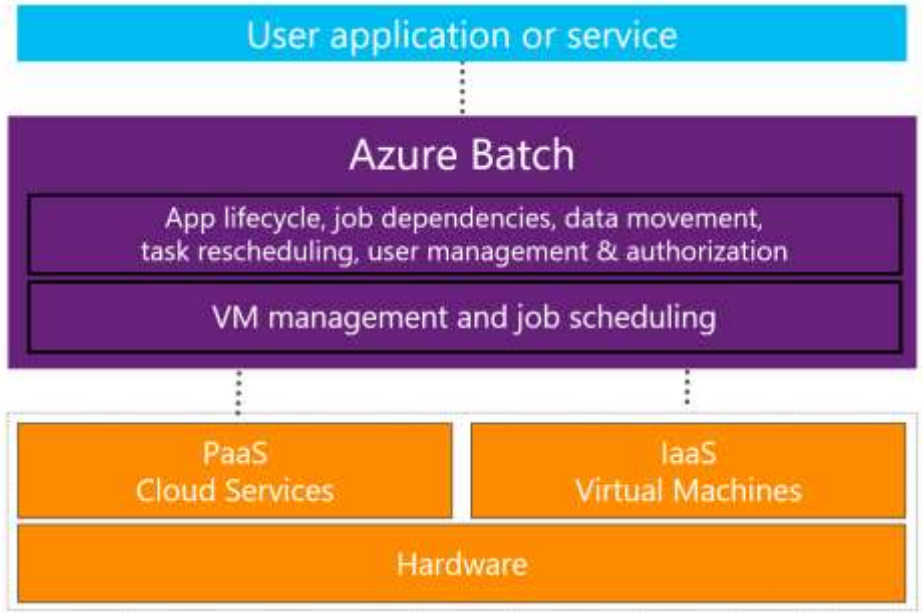
Karl Podesta (@karlpodesta)
kapodest@microsoft.com
https://github.com/azurebigcompute

Azure Batch – Concepts

Before….



After…

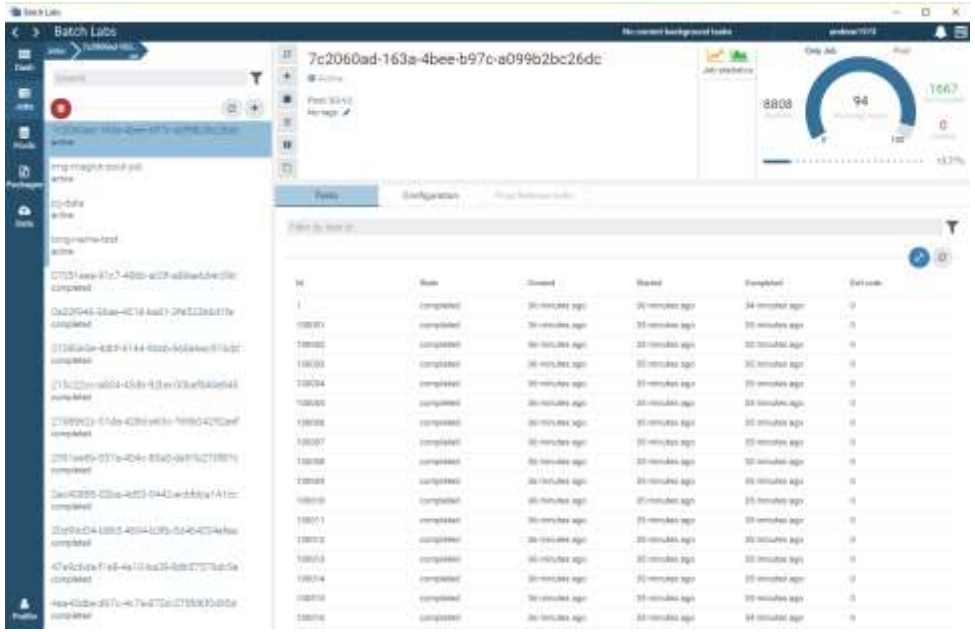


- Code directly against this service (it's an API, and it's free)
- Let Azure Batch manage the resources & scheduling
- Use any resources (any VM instance, GPUs, RDMA/MPI, etc)
- Popular with developers, software vendors (ISVs)
- Create "Pools", submit "Jobs" made up of "Tasks"
- Mix & match, dynamically resize, include Low Pri VMs to save!

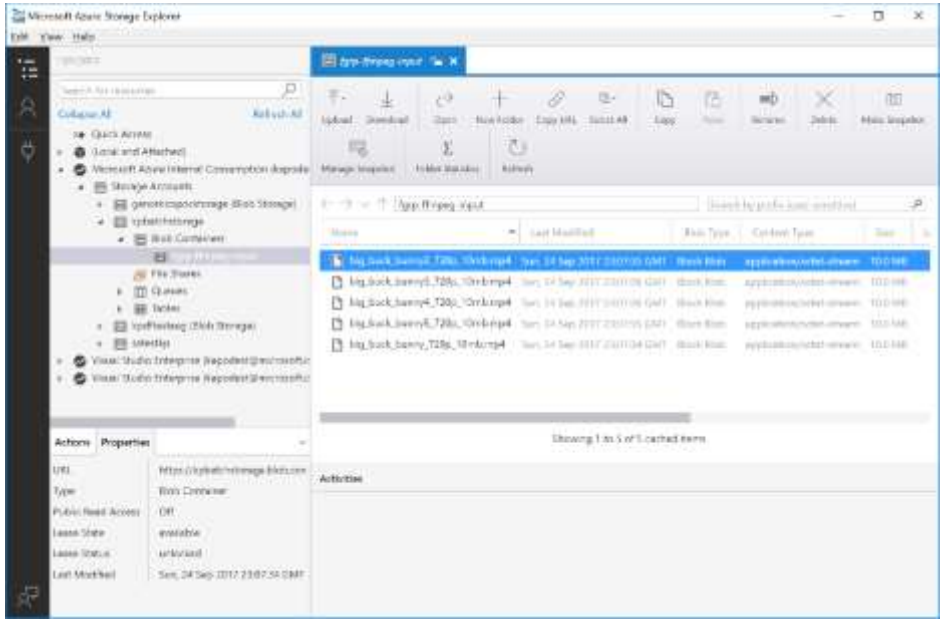
Karl Podesta (@karlpodesta)
kapodest@microsoft.com
https://github.com/azurebigcompute

Azure Batch – Lab setup

Azure Batch Labs



Azure Storage Explorer



Batch AI Lab:
https://github.com/azurebigcompute/CloudWorkshops/blob/master/BatchAIWorkshop.md

Batch with FFMPEG Lab:
https://github.com/azurebigcompute/Labs/tree/master/Azure%20Batch%20Masterclass%20Labs

Karl Podesta (@karlpodesta)
kapodest@microsoft.com
https://github.com/azurebigcompute

HPC SaaS

Karl Podesta (@karlpodesta)
kapodest@microsoft.com
https://github.com/azurebigcompute

HPC SaaS solutions from our partners (a few examples)

UberCloud – Click & Go from Marketplace

Rescale – submit input, scale, get output

Altair – pbscloud.io & solidThinking Unlimited

Oil & Gas: Schlumberger, Landmark, Paradigm

Schlumberger

Specifics

- Lab Choices
- Challenges
- Guidance
- Resources

Karl Podesta (@karlpodesta)
kapodest@microsoft.com
https://github.com/azurebigcompute

You have a choice of Labs!

- **Remote Visualisation** – run a remote Linux or Windows desktop on our GPUs
- **HPC Containers** – Singularity, Docker, Azure Batch Shipyard
- **HPC Storage** – deploy a BeeGFS shared storage system
- **HPC Data Transfer** – methods for transferring data to/from/inside Azure
- **HPC Pack** – expand your Excel workbook calculations to a cloud cluster!
- **Batch AI/Render** – dip into AI training or Video Rendering with Azure Batch
- **Industry Scenario** – Bio/Genomics
- **Infrastructure Tooling** – see how to use Chef, Puppet, Ansible, Salt, Chocolatey
- **Solution/Architecture/Design Lab** – consider alternative design scenarios

Karl Podesta (@karlpodesta)
kapodest@microsoft.com
https://github.com/azurebigcompute

Challenges:

What about **cost**?

What about **scale**?

What about **data movement**?

What about **3rd party software licensing**?

What about **openness & transparency**?

What about **security, privacy, and trust**?

Karl Podesta (@karlpodesta)
kapodest@microsoft.com
https://github.com/azurebigcompute

Challenges: What about cost?

Pricing of Azure:

- Reserved Instances (1yr & 3yr)
- Burstable Instances
- Low Priority Instances
- Pay as you go

Buying from Microsoft:

- Sign up with Credit Card
- Negotiate an EA
- Buy through a 3rd party or CSP

Cost insight

- Cloudyn, aka Azure Cost Management

Cost control

- Controls in Cycle Computing
- Controls in Azure

Karl Podesta (@karlpodesta)
kapodest@microsoft.com
https://github.com/azurebigcompute

Challenges: What about scale?

SMALL	MEDIUM	LARGE
1-2000 cores	2,000-20,000 cores	20,000-200,000 cores

Considerations:

- Test, Measure, Plan
- Start small - then work upwards or outwards
- VM spin-up time, and overprovisioning
- VM placement & availability sets
- Azure limits – storage accounts, cores, defaults
- Really big? Tell us what's going on

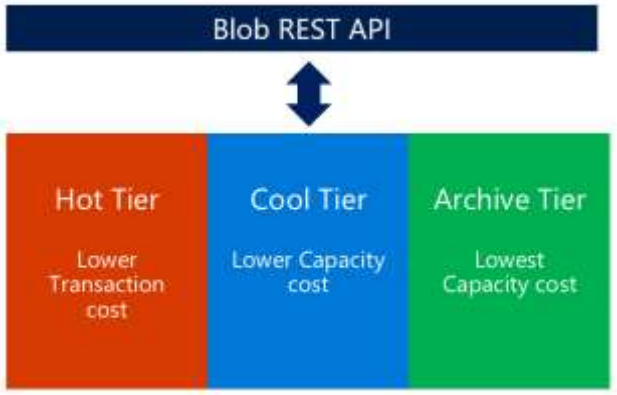
Challenges: What about data movement?

Getting data in or out:

- Standard Internet Connection
- VPN
- ExpressRoute
- Azure Transfer Software
- 3rd Party Transfer Software
- Azure Data Box

Moving data around:

- Between storage
  - disk <-> share <-> blob <-> archive
- Between regions



Karl Podesta (@karlpodesta)
kapodest@microsoft.com
https://github.com/azurebigcompute

Karl Podesta (@karlpodesta)
kapodest@microsoft.com
https://github.com/azurebigcompute

Challenges: What about software licenses?

- With software vendors: it varies! They are transforming, like all of us
- Don't underestimate your power as a customer
- Connecting to on-premise license servers via VPN
- Making license servers in the cloud
- Software/License Tokens
- Pay up front, or pay as you go

Challenges: What about skills?

- Documentation: good and getting better
- Online courses from Microsoft
- Online courses from others
- Solutions on GitHub
- Keeping up – twitter, stackoverflow

Karl Podesta (@karlpodesta)
kapodest@microsoft.com
https://github.com/azurebigcompute

Karl Podesta (@karlpodesta)
kapodest@microsoft.com
https://github.com/azurebigcompute

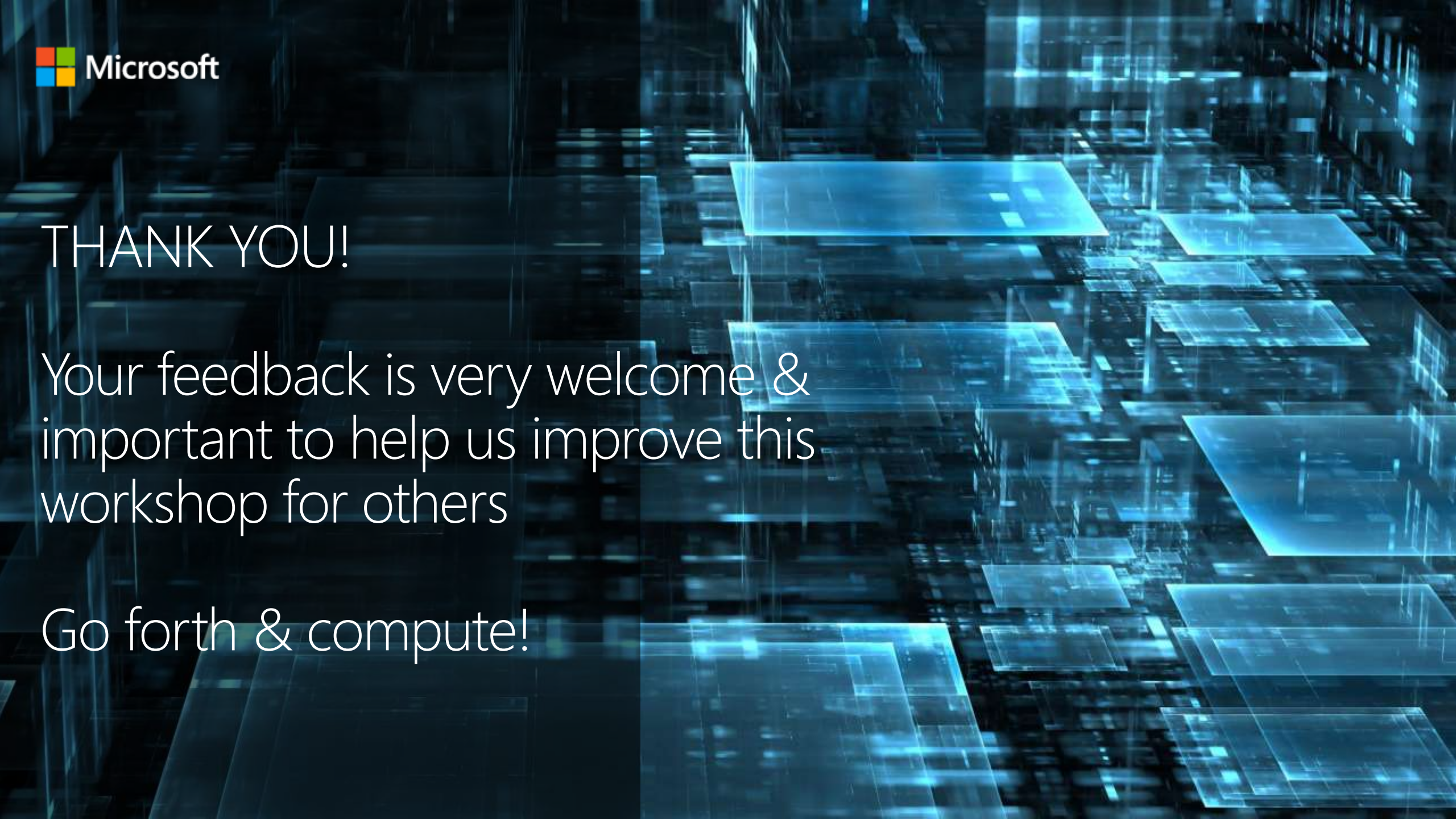
Guidance – let's share some experiences ☺

- Using IaaS, PaaS, SaaS, or a combination?
- Choose one app, one user, or one scenario – try it and see
- Identify existing pain points, problems, potentials
- "Cloud thinking" – designing for change
- Azure Architecture/Solution Centre – designing for Azure
- When to: HPC Pack, CycleCloud, Azure Batch, ARM, Terraform, etc.
- Skilling up
- Keeping up

Resources

- Azure Big Compute – GitHub Repository
  - https://github.com/azurebigcompute
- Azure Documentation
  - https://docs.microsoft.com/en-us/azure/
- Videos online
- Contact us!

Karl Podesta (@karlpodesta)
kapodest@microsoft.com
https://github.com/azurebigcompute


Microsoft
THANK YOU!
Your feedback is very welcome & important to help us improve this workshop for others
Go forth & compute!