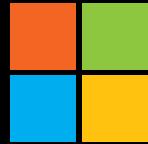




SDSC HPC/CI Training Series

Intro to Azure HPC

Paul Yu
Cloud Solution Architect
pauyu@microsoft.com



Microsoft Azure

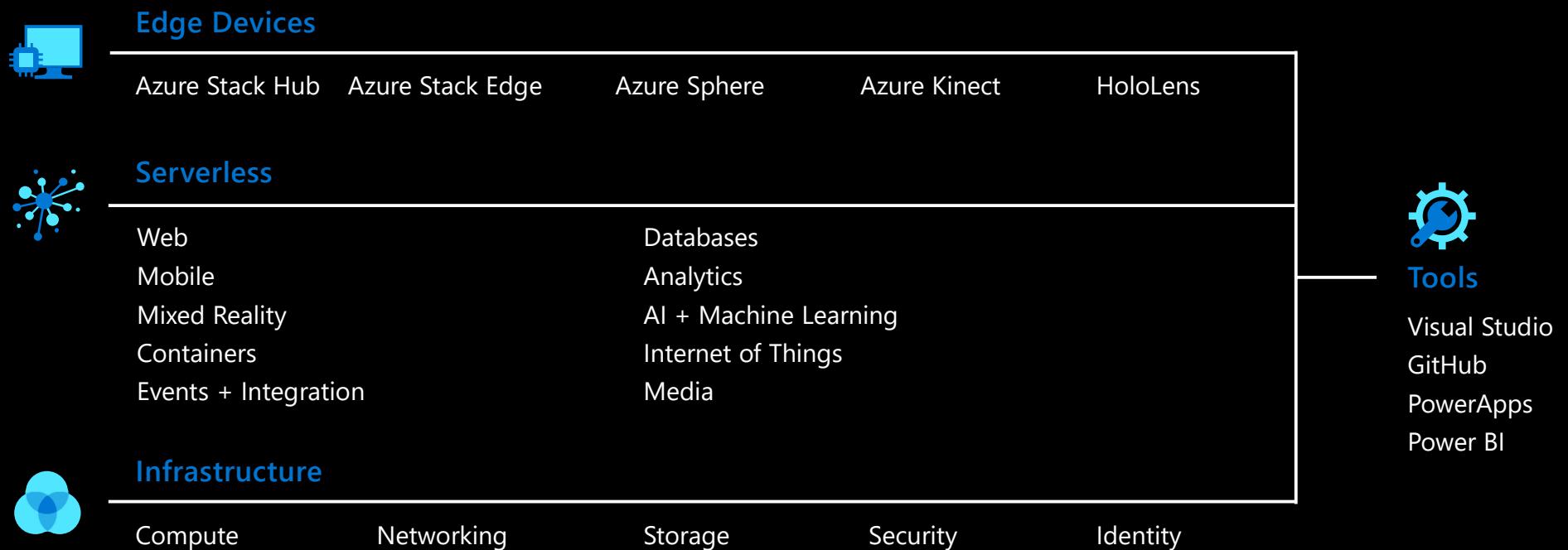
Be future
ready

Build on
your terms

Operate hybrid
seamlessly

Trust
your cloud

What is Azure?

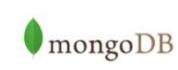


Works with popular tools and frameworks

Management



Databases & middleware



Applications



Infrastructure



App frameworks



>95% of Fortune 500
use Microsoft Azure

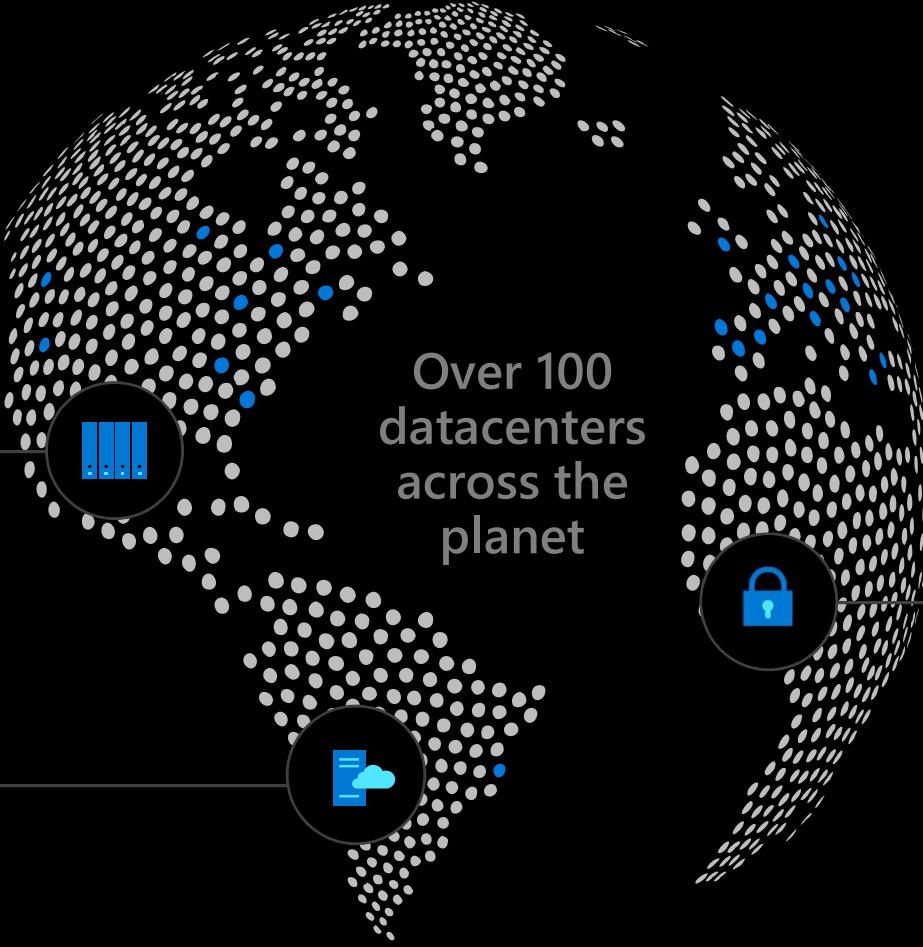


BANK OF AMERICA



A secure foundation at global scale

Each **physical datacenter**
protected with world-class,
multi-layered protection



Over 100
datacenters
across the
planet

Global cloud infrastructure
with custom hardware and
network protection

Secured with cutting-edge
operational security

- Restricted access
- 24x7 monitoring
- Global security experts



Global



60+ Azure regions

Largest geographical footprint of any cloud provider with more than **60** Azure regions



Secure



Microsoft Cyber Defense Operations Center

>3,500 full-time security professionals

6.5 trillion global signals daily

\$1 billion annual cybersecurity investment



Compliant

91

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
- CIS Benchmark
- CSA STAR Certification
- CSA STAR Attestation
- CSA STAR Self-Assessment
- WCAG 2.0 (ISO 40500:2012)

U.S. GOVT

- FedRAMP High
 - FedRAMP Moderate
 - EAR
 - ITAR
 - 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
 - CJIS
 - IRS 1075
 - CNSSI 1253
- PCI DSS Level 1
 - GLBA (US)
 - FFIEC (US)
 - Shared Assessments (US)
 - SEC 17a-4 (US)
 - CFTC 1.31 (US)
 - FINRA 4511 (US)
 - SOX (US)
 - 23 NYCCR 500 (US)
 - OSFI (Canada)
 - FCA + PRA (UK)
 - APRA (Australia)
 - FINMA (Switzerland)
 - FSA (Denmark)
 - RBI + IRDAI (India)
 - MAS + ABS (Singapore)
 - NBB + FSMA (Belgium)
 - AFM + DNB (Netherlands)
 - AMF + ACPR (France)
 - KNF (Poland)
 - European Banking Authority (EBA)
 - FISC (Japan)
 - HIPAA BAA (US)
 - HITRUST Certification
 - GxP (FDA 21 CFR Part 11)
 - MARS-E (US)
 - NHS IG Toolkit (UK)
 - NEN 7510:2011 (Netherlands)
 - FERPA (US)
 - CDSA
 - MPAA (US)
 - FACT (UK)
 - DPP (UK)

INDUSTRY

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

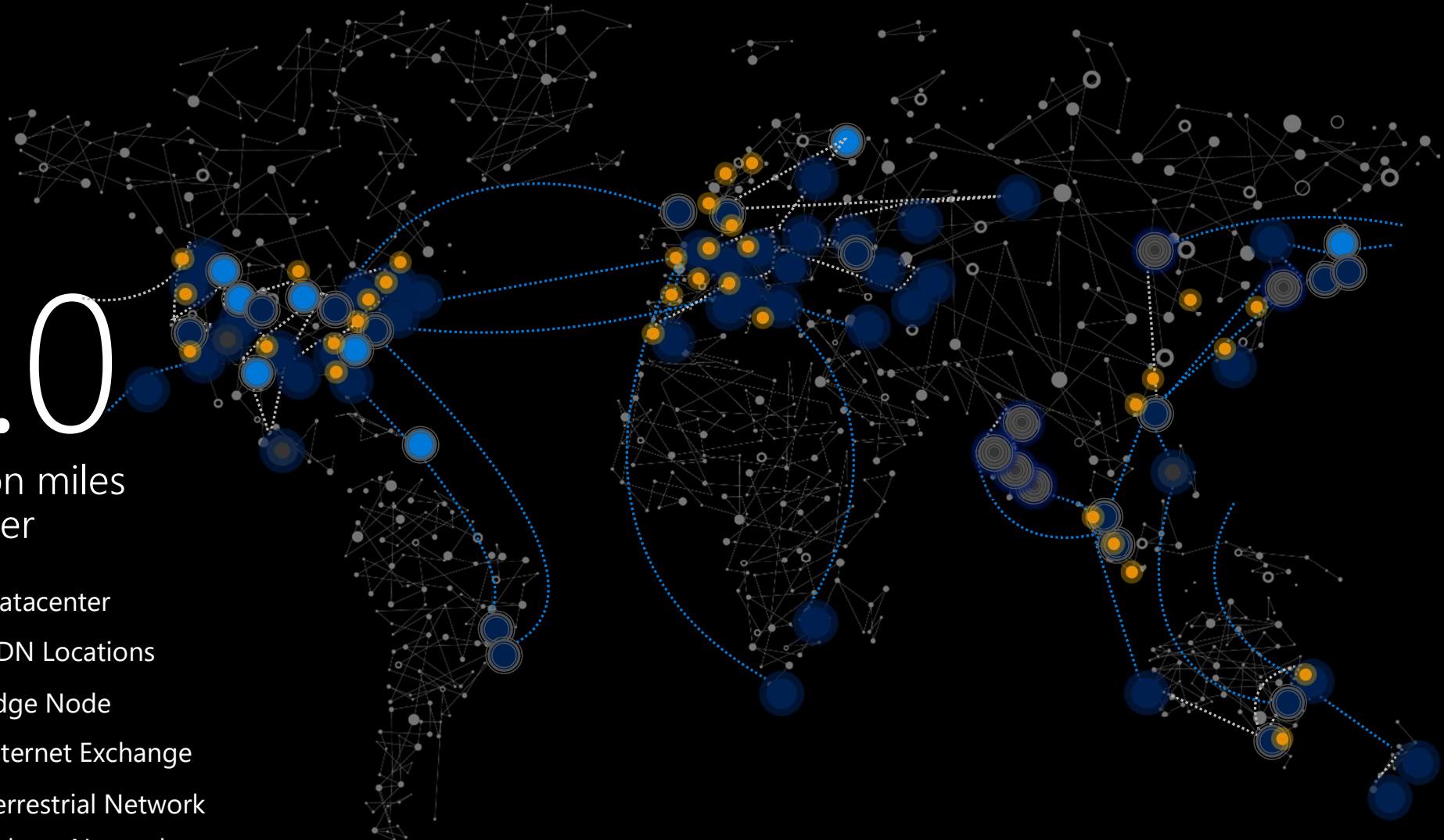
REGIONAL

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

2.0

million miles
of fiber

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



Azure Regions



Azure Fundamentals



Identity

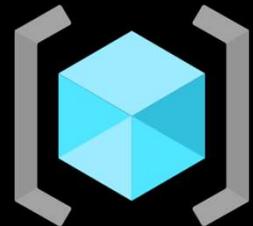
- Azure Active Directory is the identity store for both Office 365 and Azure



Subscription

- Billing boundary for Azure and where you deploy resources into

Azure Fundamentals



Resource Groups

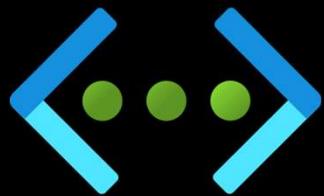
- Logical container for resource organization



Identity and Access Management

- Role Based Access Control
- Assignments can be scoped at the following levels:
 - Management Groups
 - Subscriptions
 - Resource Groups
 - Resource

Azure Fundamentals



Virtual Network

- RFC1918 address space
 - 10.0.0.0 - 10.255.255.255 (10.0.0.0/8)
 - 172.16.0.0 - 172.31.255.255 (172.16.0.0/12)
 - 192.168.0.0 - 192.168.255.255 (192.168.0.0/16)
- Create subnets to further divide vnet address space
- Inter-vnet communication available by default



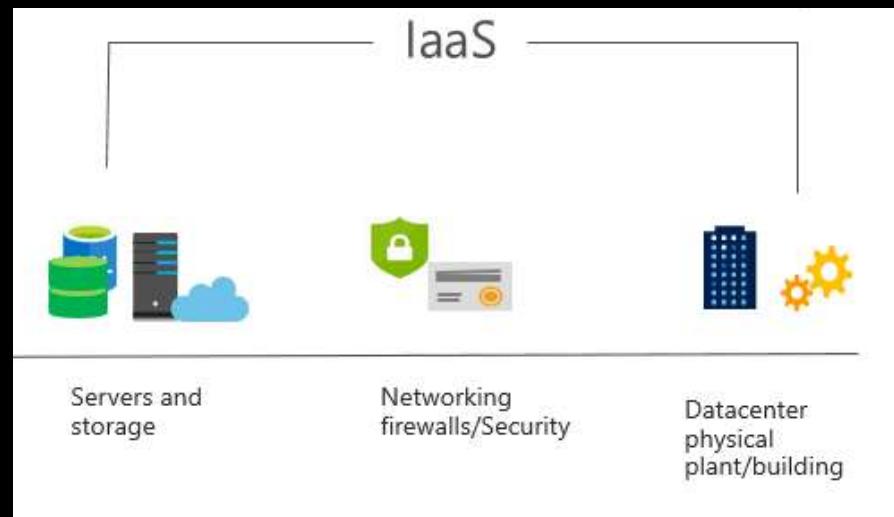
Network Security Group

- Stateless firewall – allow or deny traffic using IP/Port

Ways to deploy resources to Azure

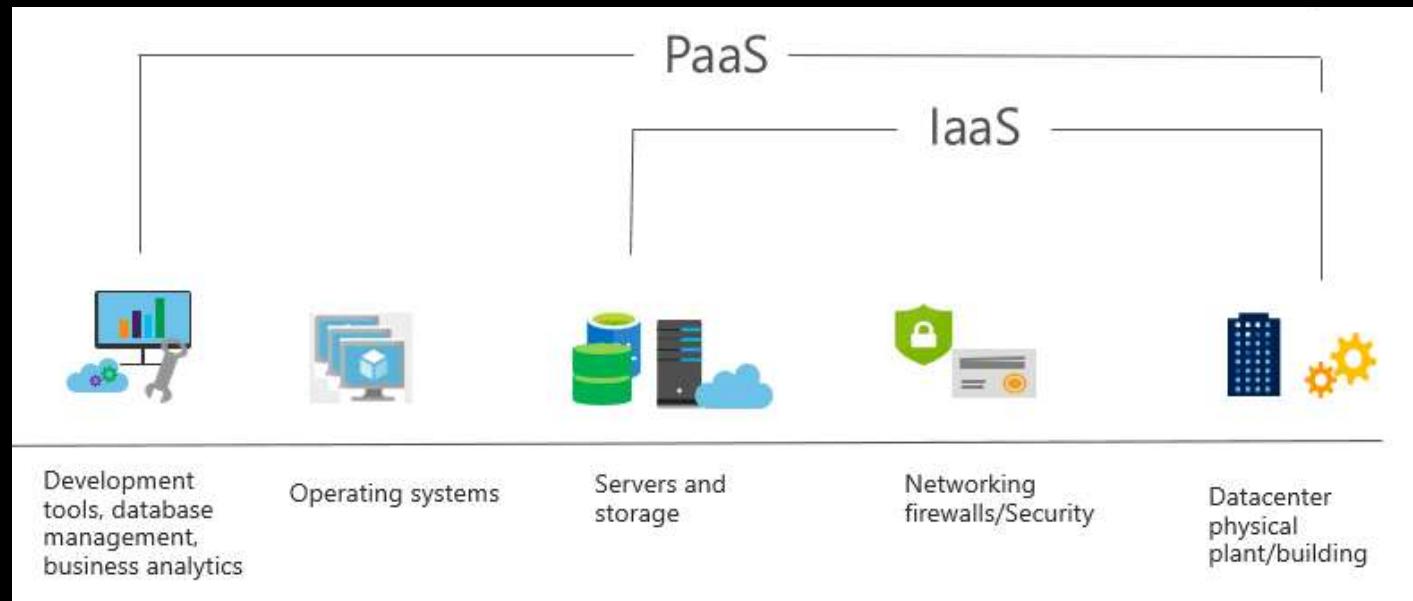
- Azure Portal
 - Web UI – <https://portal.azure.com>
- Azure Resource Manager Templates
 - Declarative JSON templates
- Command Line
 - Azure PowerShell
 - Azure CLI
 - Azure CloudShell
- Azure REST API
- 3rd party DevOps tools
 - Terraform
 - Ansible

Infrastructure as a Service (IaaS)



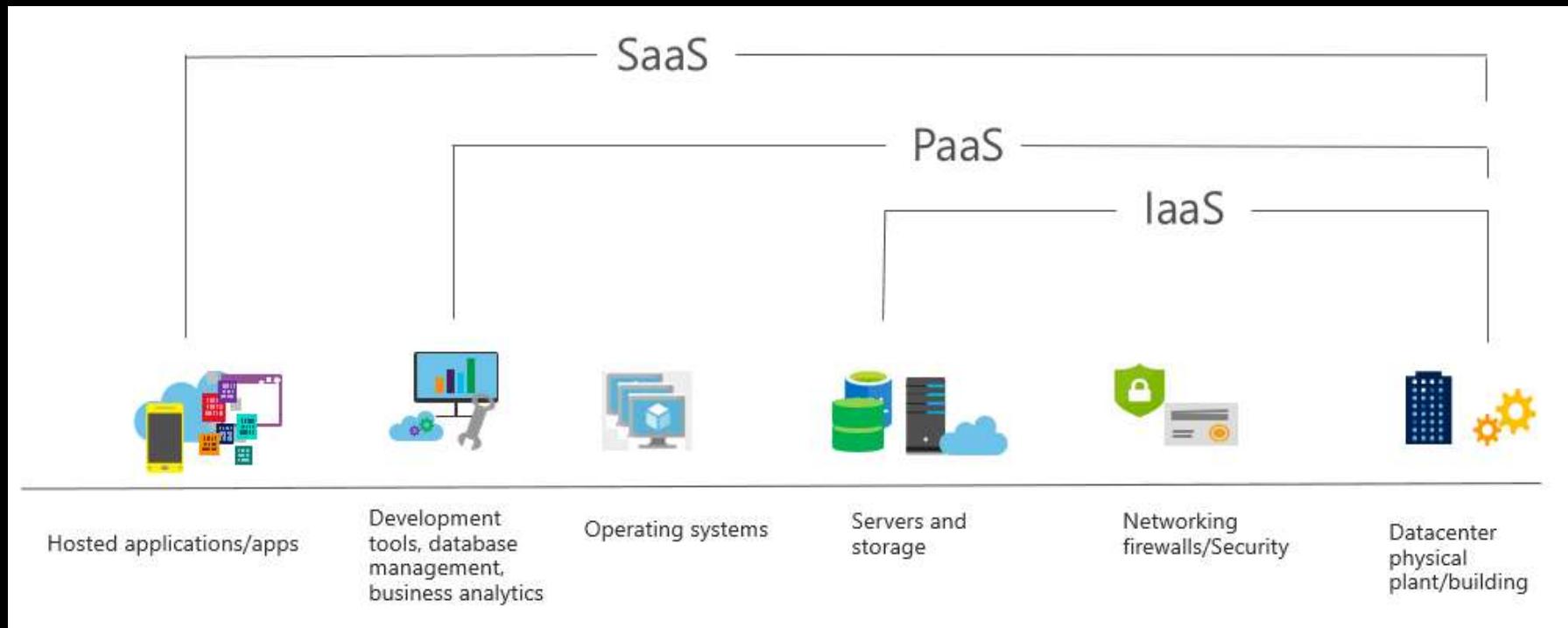
Build pay-as-you-go IT infrastructure by renting servers, virtual machines, storage, networks, and operating systems from a cloud provider.

Platform as a Service (PaaS)



Provides environment for building, testing, and deploying software applications; without focusing on managing underlying infrastructure.

Software as a Service (SaaS)



Users connect to and use cloud-based apps over the internet: for example, Microsoft Office 365, email, and calendars.

Compare cloud services

IaaS

- The most flexible cloud service.
- You configure and manage the hardware for your application.

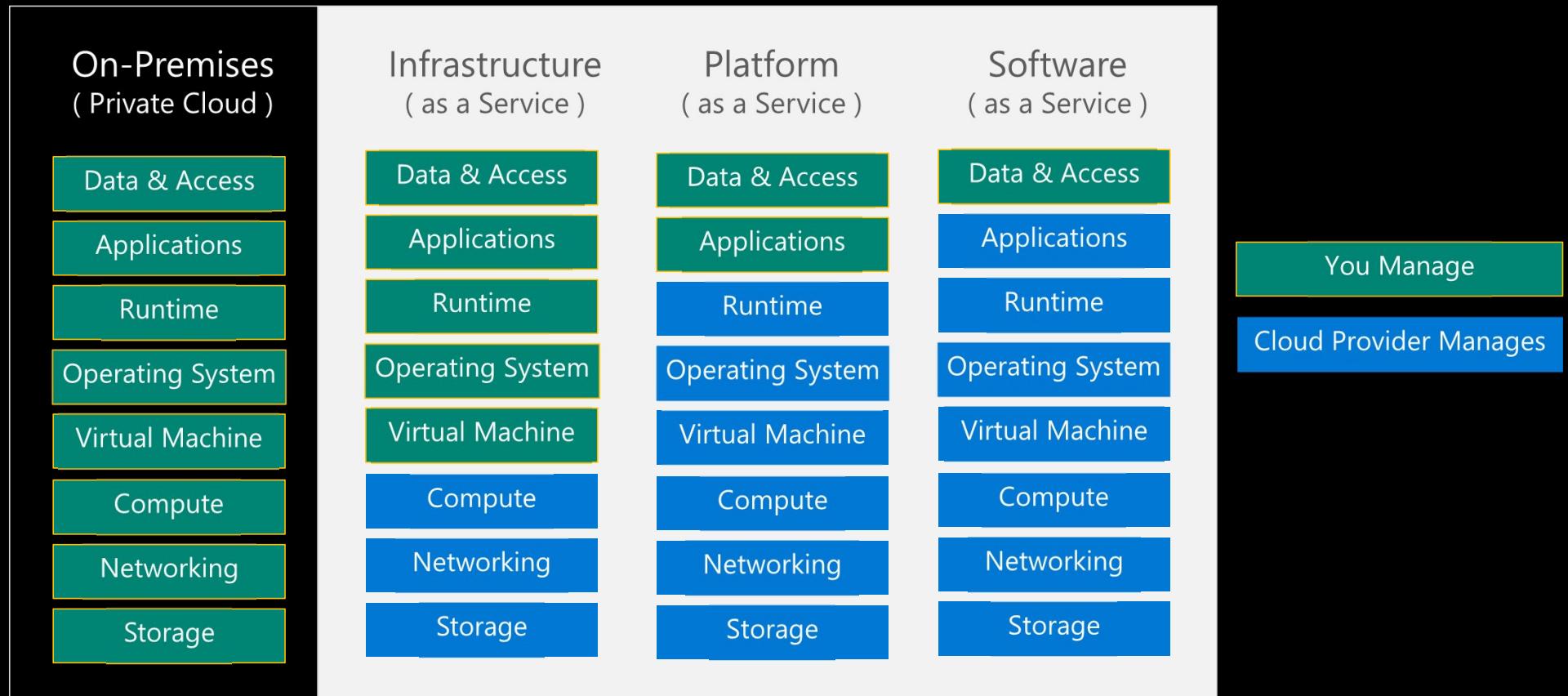
PaaS

- Focus on application development.
- Platform management is handled by the cloud provider.

SaaS

- Pay-as-you-go pricing model.
- Users pay for the software they use on a subscription model.

Shared responsibility model



Azure HPC Platform



HPC Resource Stack on Azure



Transformative Services

Azure Machine Learning

Azure Data Lake

Azure ML Compute



Workload Orchestration

VM Scale Sets

Azure Batch

Azure CycleCloud



Fast, Secure Networking

ExpressRoute

InfiniBand



High Performing Storage

Azure HPC Cache

Azure NetApp Files

Clusterstore

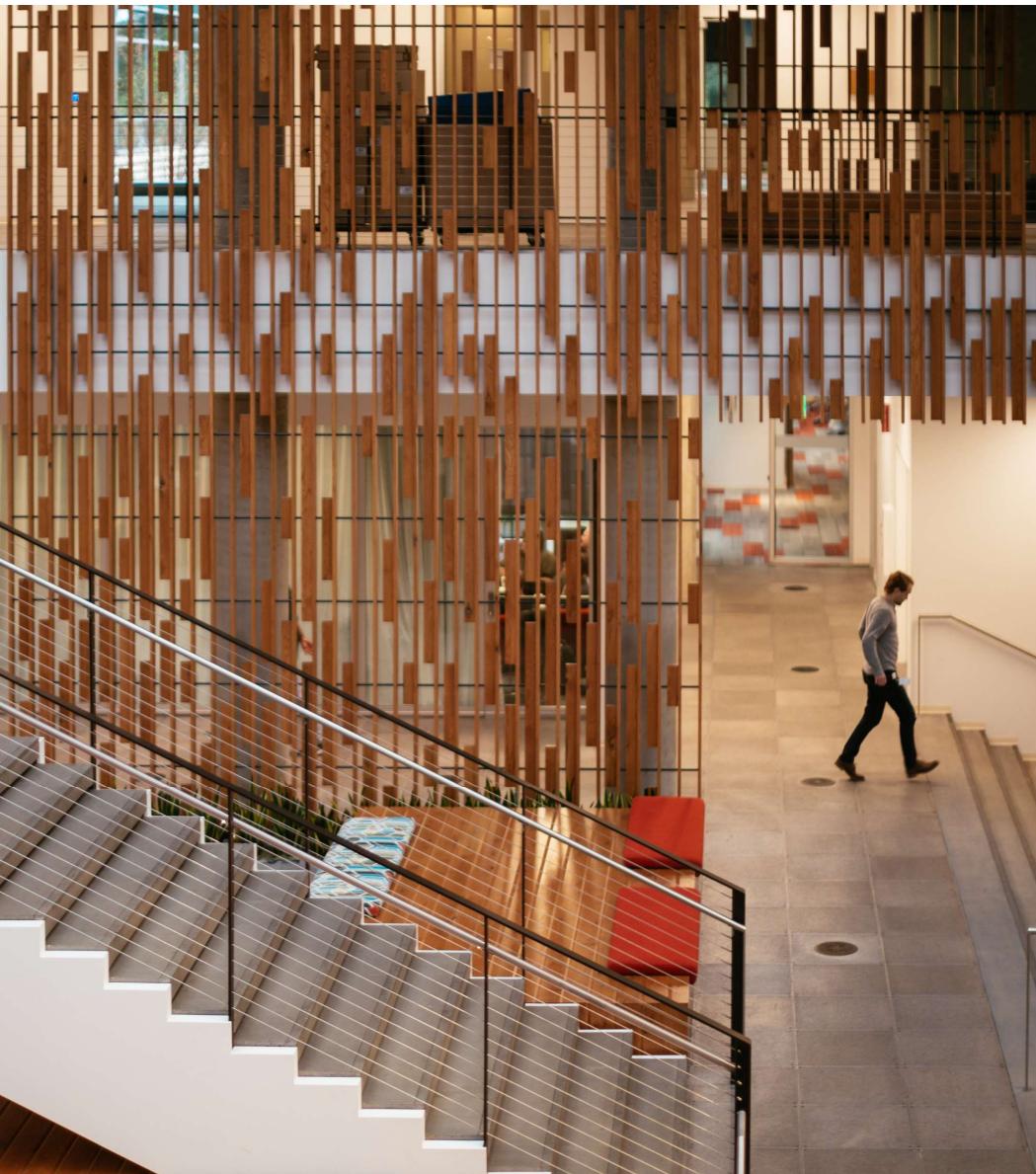


Optimized Compute

H-Series

N-Series

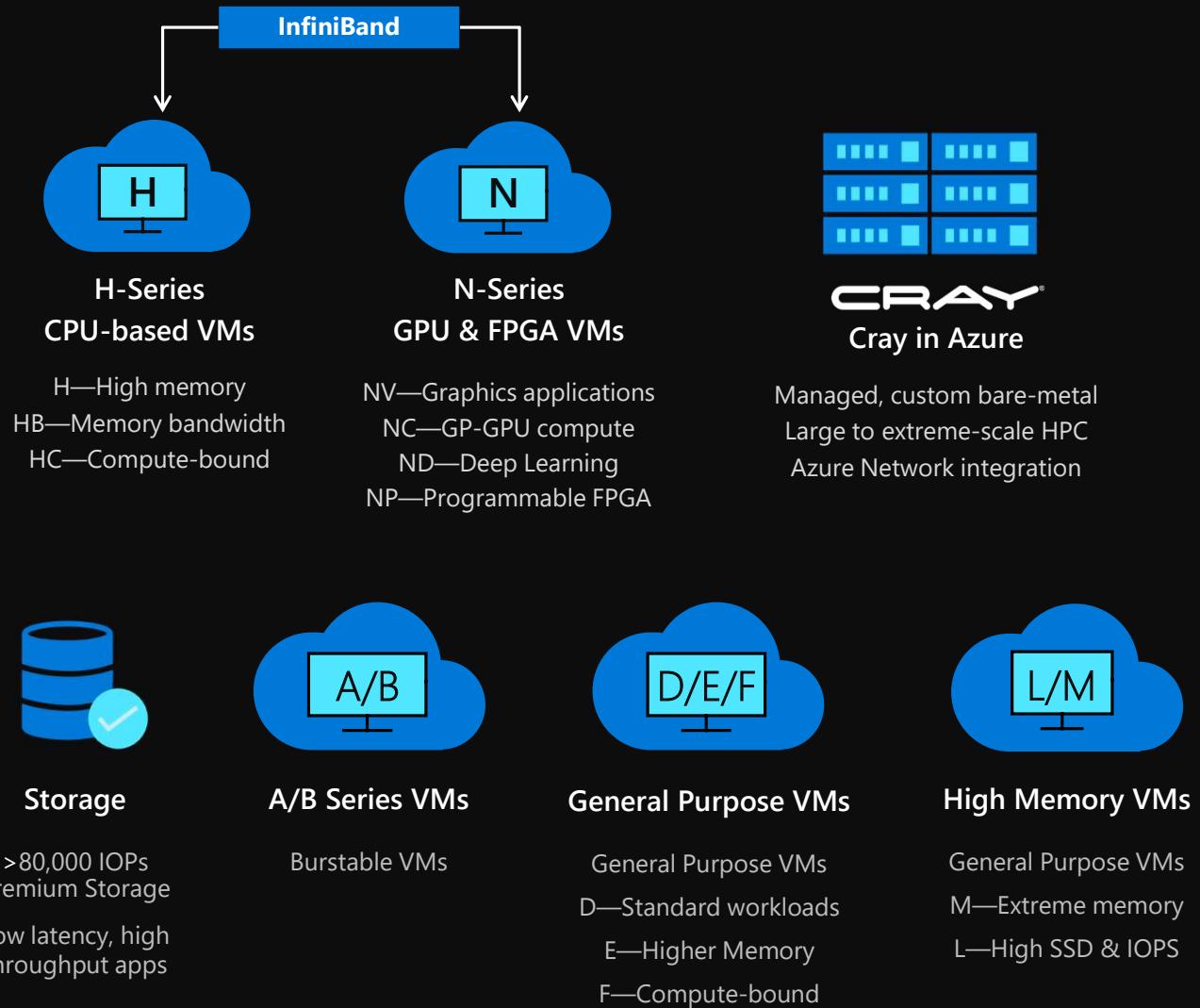
Cray



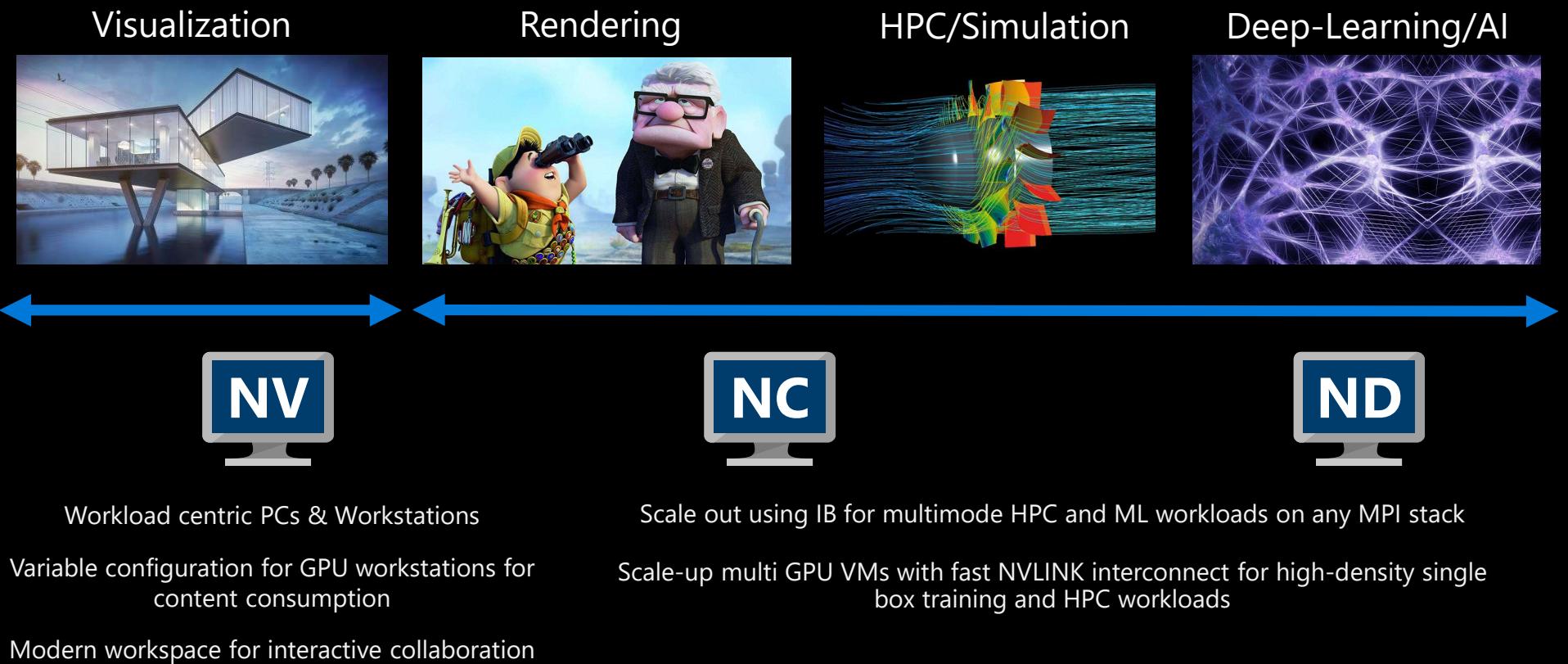
Optimized
Compute



Azure Specialized Compute for HPC, AI, and Viz



GPU Computing



GPUs for Compute (NC) & Visualization (NV)



	NC	NCv2	NCv3
Cores	6, 12, 24	6, 12, 24	6, 12, 24
GPU	K80	P100	V100
Memory	56/112/224 GB	112/224/448 GB	112/224/448 GB
Local Disk	~380/~680/~1.5 TB SSD	~700/~1.4/~3 TB SSD	~700/~1.4/~3 TB SSD
Network	Azure Network + InfiniBand (largest size only)		

	NV	NVv2	NVv3	NVv4
Cores	6, 12, 24	6, 12, 24	12, 24, 48 (24, 32 HT)	4, 8, 16 Partial, 32 Full
GPU	K80	P100	M60	Radeon Mi25
Memory	56/112/224 GB	112/224/448 GB	112/224/448 GB	14/28/56/112 GB
Local Disk	~380/~680/~1.5 TB SSD	~700/~1.4/~3 TB SSD	~700/~1.4/~3 TB SSD	~88/~176/~352/~700 GB
Network	Azure Network + InfiniBand (largest size only)			



GPUs for Deep Learning (ND Series)

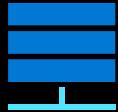


	NDv1	NDv2	NDv3
Cores	6, 12, 24	40	8, 16 (IPUs)
GPU	1, 2, or 4 P40 GPU	8 V100 SXM 32 GB GPU	8 x Graphcore C2
Memory	112/224/448 GB	768 GB	672 DDR4
Local Disk	~700/~1.4/~3 TB SSD	~1.3 TB SSD	6 TB NVMe Flash
Network	Azure Network + InfiniBand	Azure Network + InfiniBand EDR+ NVLink GPU interconnect	Azure Network + InfiniBand

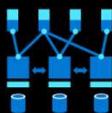
deep learning

What is Cray in Azure?

Scalable, powerful infrastructure for the most demanding workloads in engineering, climate, energy, and scientific research



A dedicated, single-tenant supercomputing resource with complete control and security over your applications, data, and hosting environment



Integrated ClusterStor™ high performance storage that centralizes your data resource and eliminates spinning up additional data repositories



Built to your custom specifications so you can get the scale of the public cloud without compromising specific application, compliance, or regulatory requirements



Fully managed service and support by Cray specialists, integrated alongside Azure support resources for the best of both worlds



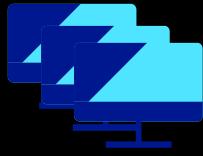
CRAY



Workload Orchestration

VM Scale Sets

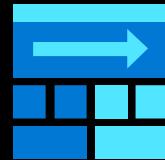
Simple ways to scale up your environment



Deploy & manage
groups of identical VMs
at very large scale



Enjoy steep
discounting with low
priority VMs



Scale multiple VM
sets utilizing
different VM types

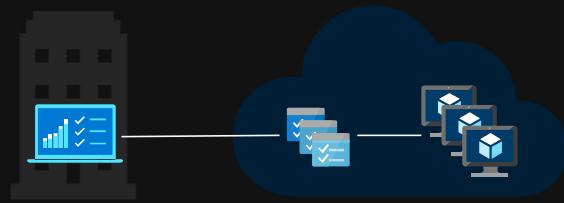


Automatically scale
by policy without
pre-provisioning

Services for HPC Workload Management

Azure Batch

Cloud-native job scheduling



HPC Application on client workstation

Azure Batch

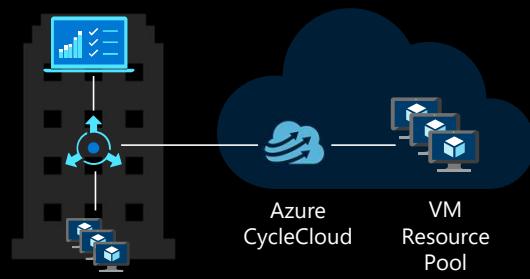
VM Resource Pool

- HPC-as-a-Service Model
- All HPC resources are cloud-based

Azure CycleCloud

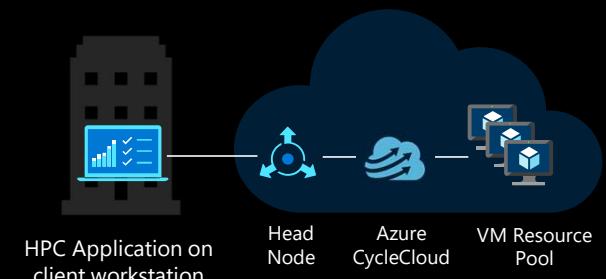
Traditional cluster scheduler orchestration

Hybrid / cloud bursting model



HPC App, Head node
and on prem compute

Cloud native model



HPC Application on client workstation

Head Node

Azure CycleCloud

VM Resource Pool

- Support for third party schedulers
- Traditional HPC scaling methodology, but using Azure

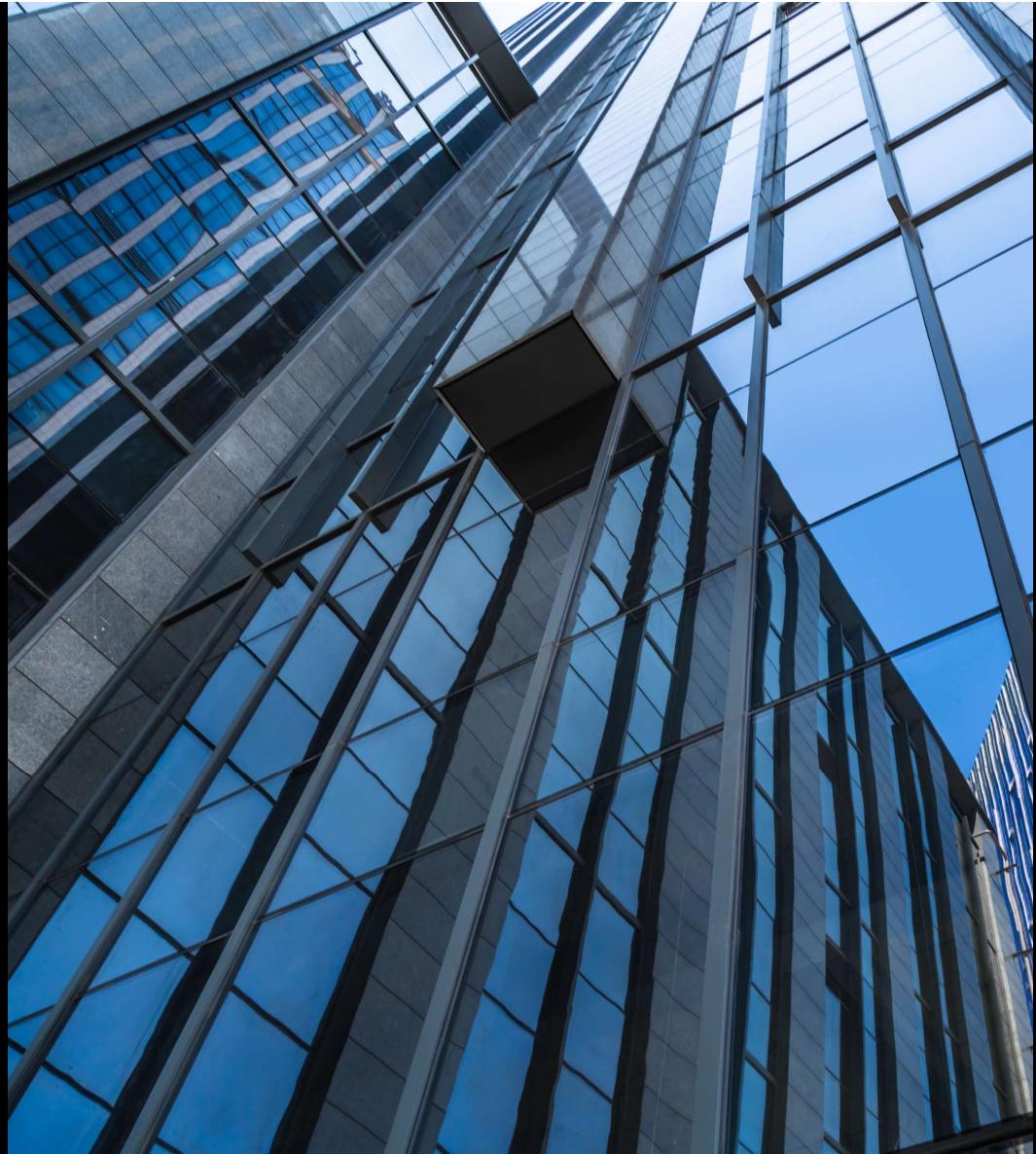
Azure Batch

Batch pools

- Configure and create VMs to cater for any scale: tens to thousands
- Automatically scale the number of VMs to maximize utilization
- Easy low-priority and VM sizing, suited to your application

Batch jobs and tasks

- Task is a unit of execution; task = application command line (EXE, BAT, CMD, PS1, etc.)
- Jobs are created and tasks are submitted to a pool. Next, tasks are queued and assigned to VMs
- Any application, any execution time; run applications unchanged
- Automatic detection and retry of frozen or failing tasks



Azure Batch Capabilities



Job scheduling

Supports both EP & MPI jobs

Run > 1 task in parallel per node

Detect and retry failed tasks

Task dependencies

Job prep and cleanup tasks



Rich app management

Get apps from blobs, Batch app packages, package managers, custom VM images

Docker container images



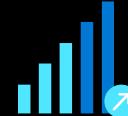
Choice of VMs

Windows or Linux

Standard or custom images

Windows pool can use AHUB

Use low-priority VMs



Monitoring

VM monitoring and auto-recover

Metrics and logs available via Portal and API



Developer access

.NET, Java, Node.js, Python, REST

PowerShell, x-plat Azure CLI

Azure Portal, Batch Labs x-plat client UI

Azure CycleCloud

User empowerment

- Able to cloud-enable existing workflows and schedulers
- Enable instant access to resources
- Provide auto-scaling, error handling

IT management

- Link workflows for internal and external clouds
- Use Active Directory for authentication and authorization
- Provide secure and consistent access

Business management

- Able to link usage to spend
- Provide tools to manage and control costs

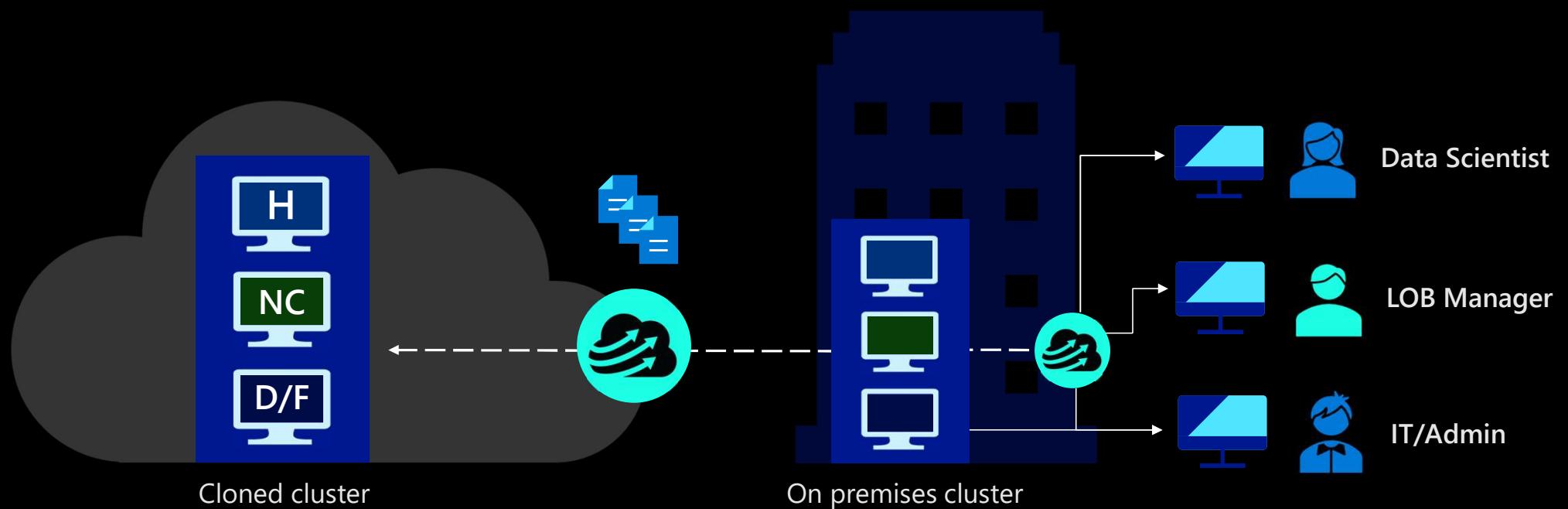


Azure CycleCloud Scenario

Filling the gaps in moving a strategic HPC application environment to Azure

Using templates to mirror the on-premises environment,
CycleCloud provisions an identical environment into
which the application can be loaded without rewriting

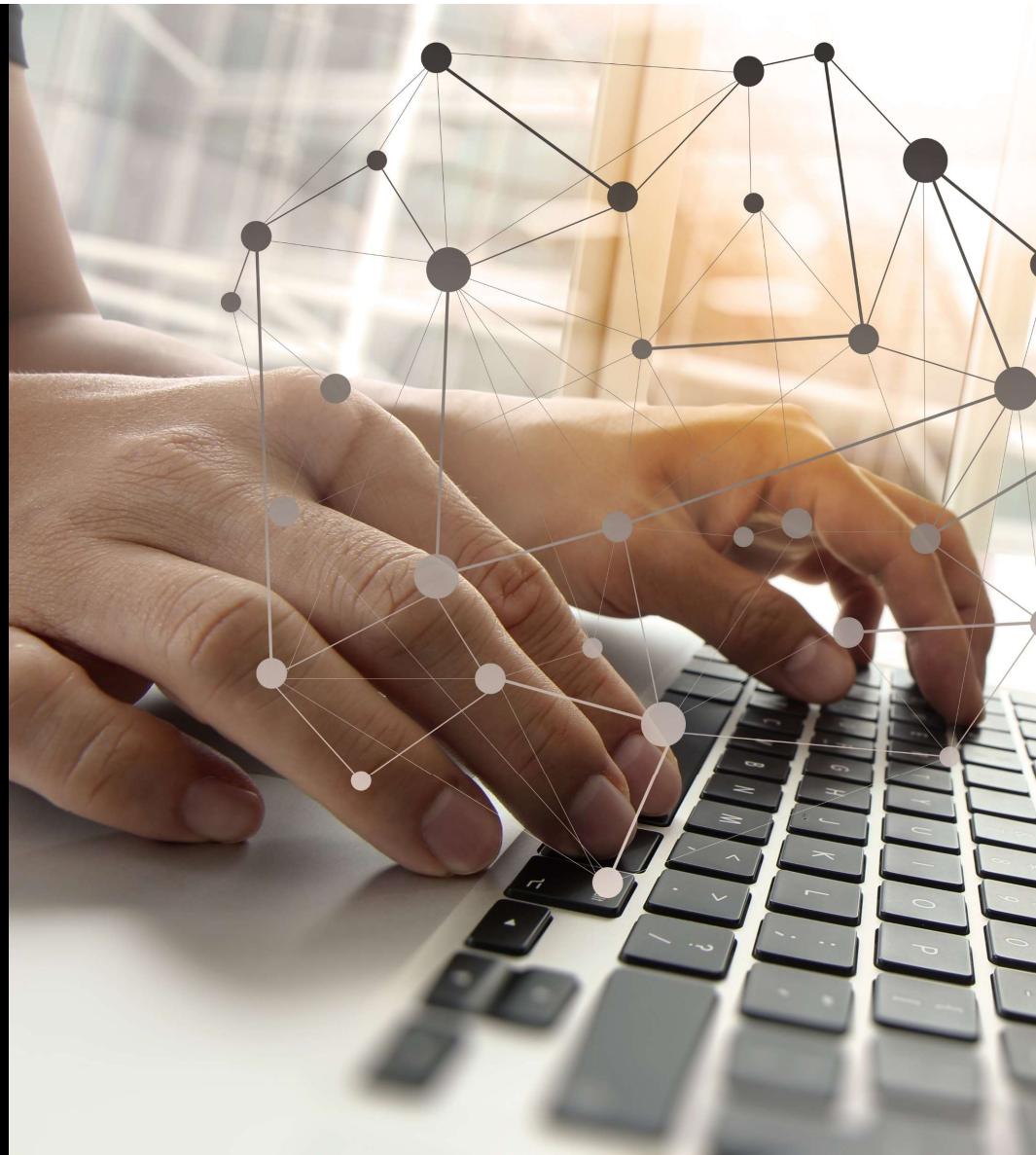
CycleCloud solves governance issues by providing a
policy-based system for access, authorizations, cost
controls, and compliance audit reporting.



Support for traditional job schedulers



High Performing Storage



Drivers of HPC Storage Choices



Performance & Scale

File sizes and quantities
(Larger # of small files == greater Ops requirement)

Concurrent access
(# of clients, rate of requests, avoid hotspots)

Latency
effects on initial load, writes, metadata (e.g., Large # small files + WAN)

Job Run Time
(the whole job, including storage components)



Access

Local (Disk)

Shared NAS

BLOB

multi-source
(combined sources)

Fault Tolerance



Protocol

Block

NFS/SMB

Rest (BLOB)

Parallel FS
(GPFS, Lustre, BeeGFS)

HDFS



Costs

Actual cost of storage
Complexity

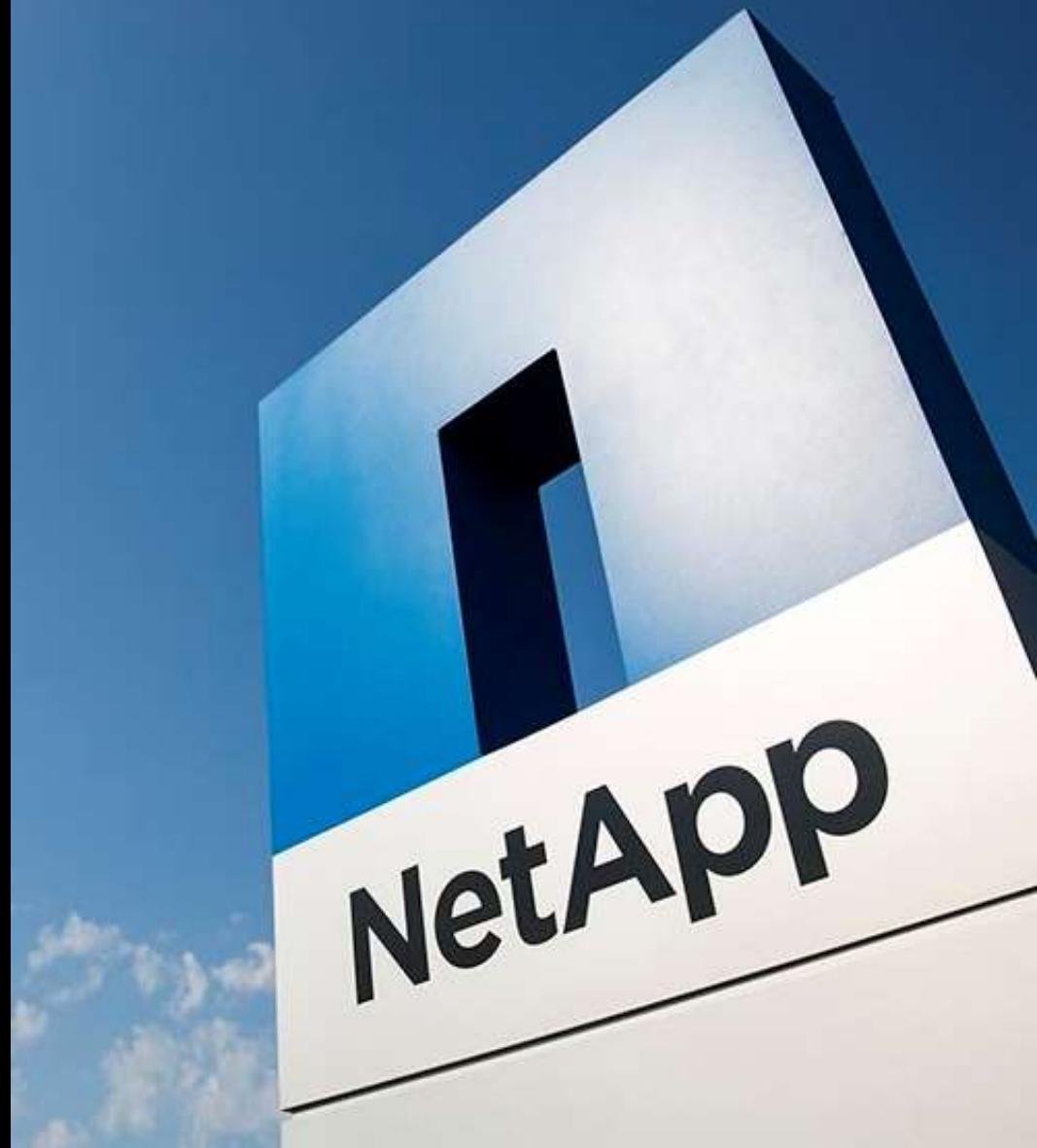
Azure NetApp Files

Key Considerations

- Capacity and performance are dynamically adjustable, readily burstable
- Excels at IOPS, including small/mixed IO
- Inode scaling
- Vnet IP address limit
- Volume limit (100TiB)
- Max bandwidth limit

Futures

- SMB multi-channel enabled by default
- Simultaneous multi-protocol access ([Mn] deliverable)
- In-place tier change ([Mn] deliverable)





Azure NetApp Files

- Widest choice of file protocols (NFSv3, NFSv4.1, SMB3) in the public cloud
- On-premises class performance with multiple tiers (Ultra, Premium, Standard)
- Azure consistent experience, support, and billing

Workloads

Enterprise File Workloads HPC
Database VDI
Containerized Applications

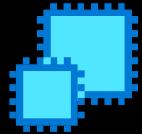
Verticals and Industries

Semiconductor EDA
Oil and Gas
Retail and E-Commerce

NetApp ONTAP technology as a fully managed Azure service

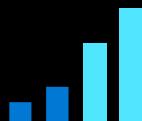
Introducing Azure HPC Cache

Flexible File System Caching for your computational workloads



High-Performance

High throughput, low latency with scale-out performance



Big Scale

Provide continuity to cloud workloads



Flexible

Burst file data to your applications



Simple

Easily enable computational workloads of any size

Choose from three Performance SKUs

Create a hot cache of shared file data

Highly-Available Distributed Scale

Support 10s or 10s of 1000s of clients/cores

Flexible Access

Use your on-prem NAS data, cloud-based data, or both in a single namespace

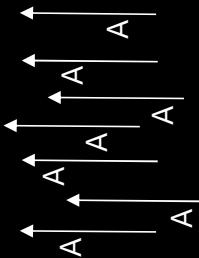
Easy to Integrate

Start in minutes
Use Azure APIs or Portal

Data and metadata at scale

HPC Cache replicates data and metadata

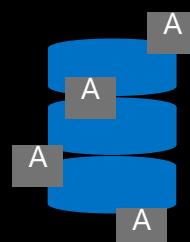
HPC Farm



File A

Read $x \times 1000s$ of times

Azure HPC Cache



HOT file data AND metadata replicated across the service instance

Only blocks of data that clients are reading are cached and replicated

Avoids “hot-spotting” by sharing the load across the service instance

Clustered NAS



File A

Read minimal number of times from the filer

Periodic attribute checking ensures that the cached version is the latest version

Clusterstore Offer Value



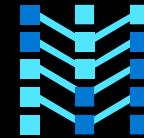
Dedicated Lustre Solution

- Single-tenant, fully managed Lustre HPC storage in Microsoft Azure
- Tailored to the customer HPC storage needs with three configs:
Small / Medium / Large
- 3 years RI for high utilization not just uncommitted demand or burst needs.
- Parallel file system available either to support compute on Cray systems through Infiniband or support compute execute on Azure VMs through ethernet connection



HPC Value Proposition target

- Azure HPC customers looking for improved performance on storage and expand their footprint
- S400 Azure customer with HPC on-prem but not in Azure targeting Manufacturing, Automotive and public sector (e.g. Federal/Government, Weather, etc.)

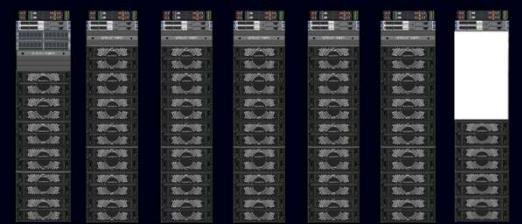


Azure Networking

- Parallel file system available either to support compute on Cray systems through Infiniband or support compute execute on Azure VMs through ethernet connection

Clusterstore Configurations

Every instance of Clusterstore supports a dedicated parallel Lustre file system, fully managed as a service from Azure

	Small	Medium	Large
			
Aggregate throughput speed	60 GBps	160 GBps	460 GBps
Raw storage capacity	5 PB	15 PB	45 PB

Learn Resources

[Azure Fundamentals Learning Path](#)

[Introduction to HPC on Azure](#)

[Run high-performance computing \(HPC\) applications on Azure](#)

[Create, customize and manage an HPC cluster in Azure with Azure CycleCloud](#)

[Customize high-performance computing clusters deployed with Azure CycleCloud](#)

[Optimizing tightly coupled HPC applications on HBv2, HC, and HB series virtual machines](#)