



Azure for Academic Research

Charles Rice
[linkedin.com/in/cricex](https://www.linkedin.com/in/cricex)
crice@microsoft.com

Chris Akin
[linkedin.com/in/akinchris](https://www.linkedin.com/in/akinchris)
christopher.akin@microsoft.com



what do higher education researchers care about most?



Higher education researchers may have varying areas of interest depending on their specific field of study, but some topics that are commonly of interest to them include:



1. **Student Success:** Researchers in higher education are often interested in understanding what factors contribute to student success, such as academic and non-academic factors, and how to best support students in achieving their academic goals.
2. **Access and Equity:** Many higher education researchers are also interested in understanding and addressing issues related to access and equity in higher education, including issues of affordability, diversity, and inclusion.
3. **Teaching and Learning:** Higher education researchers may also be interested in studying various aspects of teaching and learning, such as effective instructional strategies, student engagement, and the use of technology in the classroom.
4. **Institutional Effectiveness:** Researchers in higher education may also focus on issues related to institutional effectiveness, such as strategic planning, resource allocation, and performance measurement.
5. **Policy and Governance:** Finally, higher education researchers may be interested in studying various aspects of policy and governance in higher education, including issues related to accreditation, accountability, and the role of government in higher education.



tell me a joke about academic research & cloud computing



Sure, here's a joke for you:

Why did the academic researcher decide to move their data to the cloud?

Because they wanted to reach new heights in their research...literally!

Hook 'em!



Chris Akin

University of Texas Lead



Megan Davis

Cloud Specialist



Charles Rice

Principal Data & AI Architect



Azure Hour for Education



<https://aka.ms/edu/AzureHour>

- App Dev/DevOps Office Hours
- Data & AI Office Hours
- Research Office Hours
- VDI Office Hours
- Security Office Hours
- Azure Updates, Open Q&A
- Topic of The Week
- Teams community
- Fridays @ 1:00pm CST



Azure training and certifications

<https://esi.microsoft.com>

Role-based		Apps & Infra		Data & AI	Specialty
Technical skills required to perform a job					Deep technical skills managing industry solutions
Expert	Azure Solutions Architect (AZ- 303 & AZ 304)				Azure for SAP Workloads (AZ- 120)
	DevOps Engineer (AZ- 400)				
Associate	Azure Administrator (AZ-104)		Azure Database Admin (DP-300)		Azure IoT Developer (AZ- 220)
	Azure Developer (AZ-204)		Azure Data Scientist (DP-100)		Migrate SQL Workloads to Azure (DP-050)
	Azure Security Engineer (AZ-500)		Azure AI Engineer (AI-100)		Migrate NoSQL Workloads to Azure Cosmos DB (DP-060)
			Azure Data Engineer (DP-200 & DP-201)		Migrate Open Source Data Workloads to Azure (DP-070)
			Data Analyst (DA-100)		
Fundamentals	Azure Fundamentals (AZ-900)		Azure Data Fundamentals (DP-900)		Migrate Application Workloads to Azure (WS-050)
	Foundational understanding of technology		Azure AI Fundamentals (AI-900)		

 NIH**STRIDES**

**20% Discount on any Azure
Service for NIH-sponsored
research**

<http://aka.ms/UT/GetAzure>



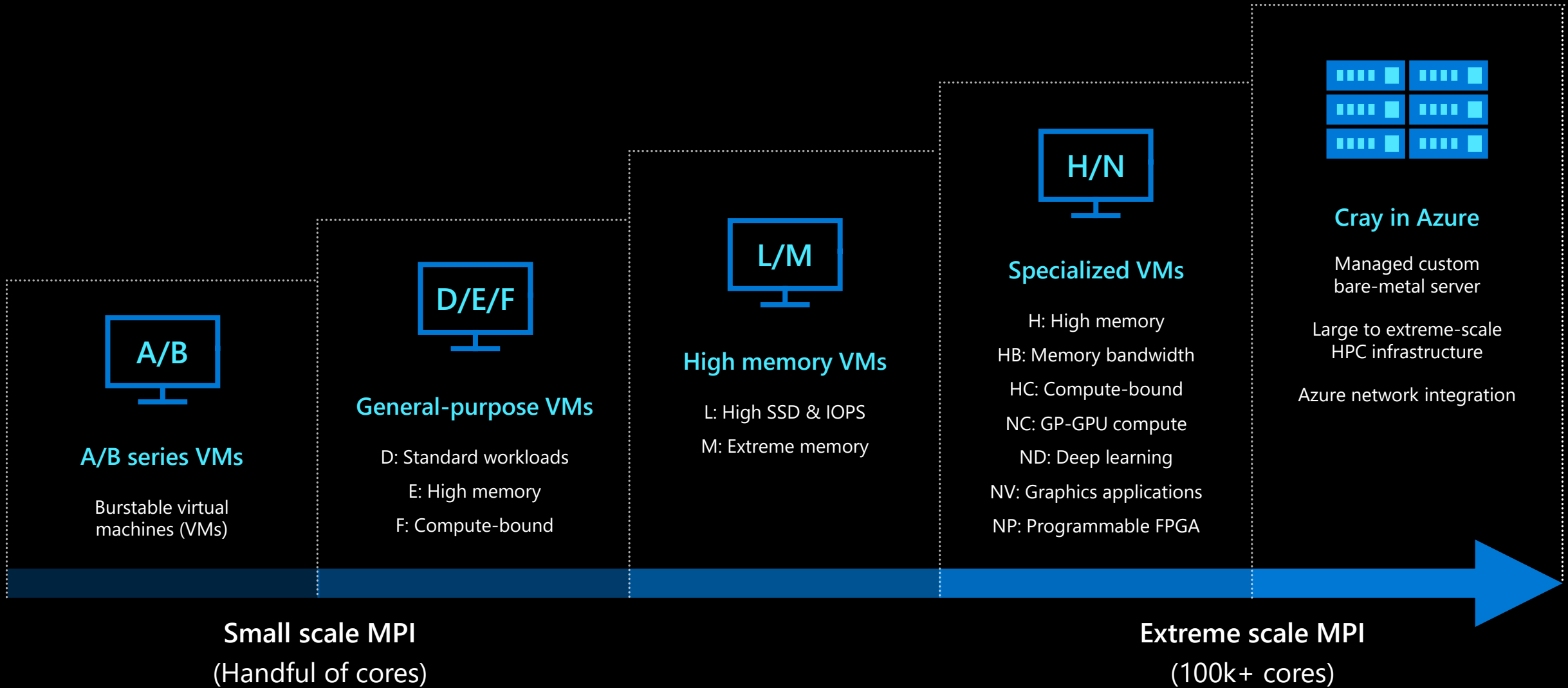
Research Table Stakes:

**Optimized
Compute**

**High
Performing
Storage**

**Workload
Orchestration**

Solve any HPC, AI, Research workload—at any scale



GPUs for Compute (NC) & Visualization (NV)







Azure Instance →	NC	NCv2	NCv3	NCasT4_v3	NC A100 v4
Cores	6, 12, 24	6, 12, 24	6, 12, 24	4, 8, 16, 64	24, 48, 96
GPU	Tesla K80	Tesla P100	Tesla V100	Tesla T4	A100 Tensor Core
Memory	56/112/224 GB	112/224/448 GB	112/224/448 GB	28/56/110/440 GB	220/440/880 GB
Local Disk	340/680/1440 GB SSD	736/1474/2948 GB SSD	736/1474/2948 GB SSD	180/360/2880 GB SSD	1123/2246/4492 GB
Network	Azure Network + InfiniBand (largest size only)				



Azure Instance →	NV	NVv3	NVv4	NVads A10 v5
Cores	6, 12, 24	12, 24, 48	4, 8, 16, 32	6, 12, 18, 36, 72
GPU	Tesla M60	Tesla M60	Radeon Instinct MI25	A10 Tensor Core
Memory	56,112,224 GB	112/224/448 GB	14/28/56/112 GB	55/110/220/440/880 GB
Local Disk	340/680/1440 TB SSD	320/640/1280 GB SSD	88/176/352/704 GB	180/320/720/1400 GB
Network	Azure Network			



















GPUs for Deep Learning (ND Series)

	 ND	 NDv2	 ND A100 v4	 NDm A100 v4
CPU Cores	6,12,24	40	96	96
GPU	1x, 2x, or 4x P40 GPUs	8x V100 32 GB (NVLink) GPUs	8x A100 40 GB GPUs	8x A100 80 GB GPUs
Memory	12/224/448 GB	672 GB	900 GB	1900 GB
Local Disk	736/1474/2948 GiB SSD	2948 GiB SSD	6 TB SSD	6.4 TB SSD
Network	Azure Network + InfiniBand EDR	Azure Network + InfiniBand EDR + NVLink GPU Interconnect	Azure Network + InfiniBand EDR + NVLink GPU Interconnect	Azure Network + InfiniBand EDR + NVLink GPU Interconnect



deep learning

Confidential computing @ Azure: Protect your research data while in use

Developer tools	 Azure Machine Learning OSS	 CCF SDK OSS	 Open enclave SDK OSS	 Mystikos LibOS OSS	
Azure Confidential Services	 MSFT SQL Azure GA	 Azure Key Vault M-HSM GA	 Microsoft Azure Attestation GA	 Azure Kubernetes Service (AKS) GA	 Azure Confidential Ledger Public preview
Virtual machines and edge	 DCsv2 SGX VMs GA	 ECa/DCa SEV-SNP VMs Limited preview	 Trusted Launch VMs Public preview	 IoT Edge Device GA	
Innovative new hardware	 GA	 Limited preview	 GA		
Industry leadership and standardization	 Co-founded	 Stewardship			



Azure confidential computing offerings cover not just VMs, but also Azure PaaS/SaaS services. Choose a 'most-secure' route with control over every line of code, or an 'easy button' route to lift n' shift existing apps to be confidential.

High Performing Storage



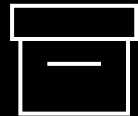
Azure Storage



Disk storage

Ultra
Premium
Standard

Reliable, persistent, high performing storage for Virtual Machines



Object storage

Azure Blobs
Azure Datalake

Massively-scalable, secure, centralized storage target for unstructured data



File storage

Azure Files
Azure NetApp Files

Simple, distributed, cross-platform file system



Data transport

Azure Import/Export
Azure DataBox

Move or migrate petabytes of data into Azure

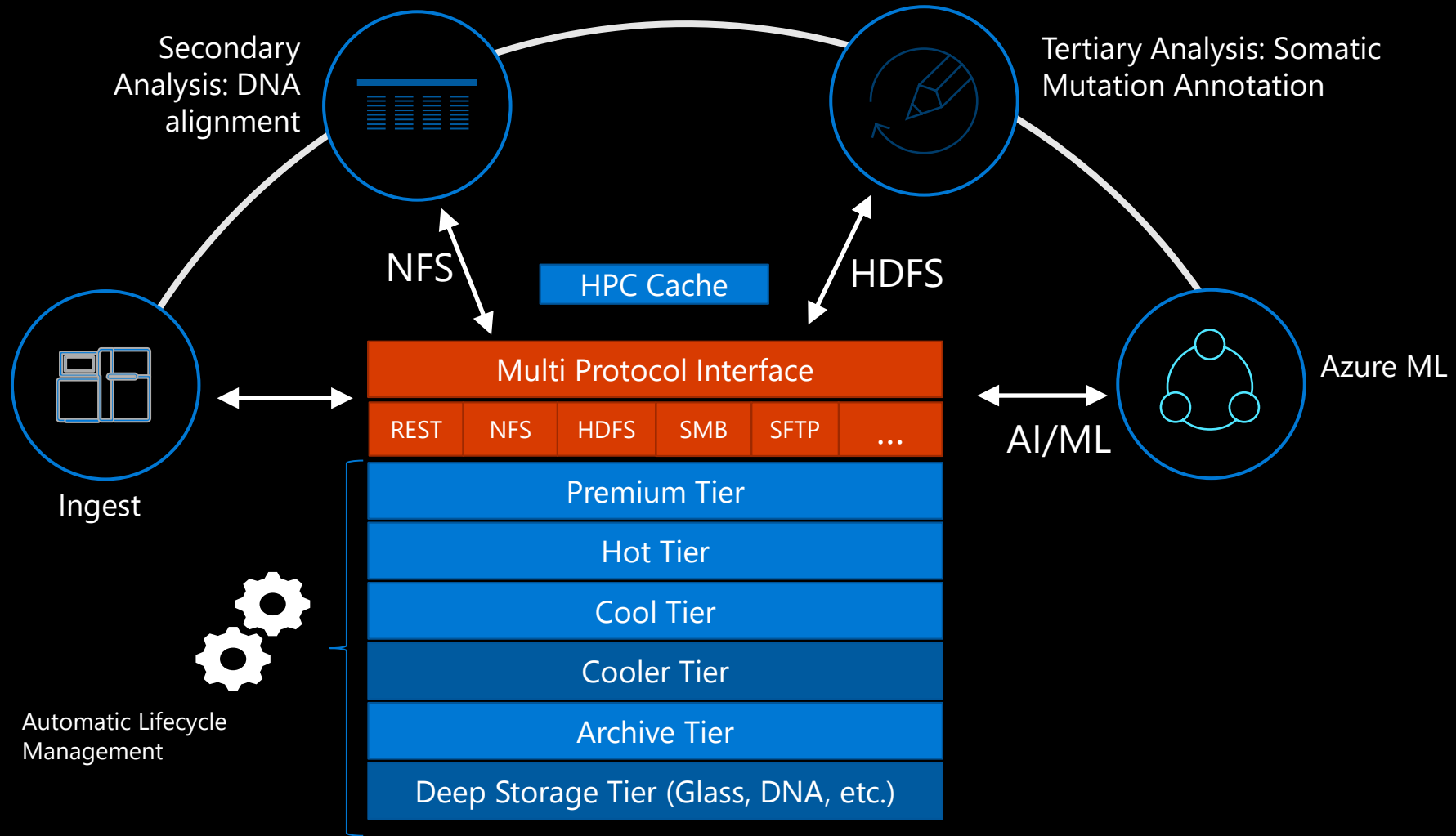


Hybrid storage

Azure Files Sync
Azure HPC Cache
Azure Managed Lustre
Azure Cray Clusterstore
Azure Databox Gateway

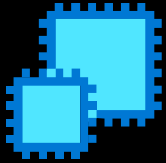
Secure, intelligent data tiering between on-premises and cloud storage

Azure Blob/ADLS Gen2: Multi-Protocol Enables End to End Workflow



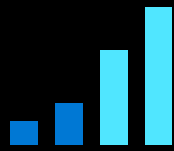
Introducing Azure HPC Cache

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

Azure Cray 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 of 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

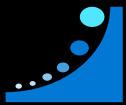
- Parallel file system available either to support compute on Cray systems through Infiniband or support compute execute on Azure VMs through ethernet connection

Azure Managed Lustre



(Public Preview)

Scalable, powerful, cost-effective, high-performance storage for the most demanding workloads in engineering, climate, energy, scientific research, and more



Accelerate HPC workloads: Provides a high-performance distributed parallel file system solution delivering hundreds of GBps storage bandwidth and solid-state disk latency, with several performance options to choose from.



Protect application investments: Enables HPC applications in the cloud without breaking application compatibility by providing familiar Lustre parallel file system functionality, behaviors and performance, securing long-term application investments.



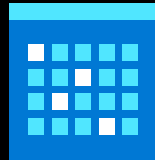
Azure Blob integration: Connect file systems with Azure storage containers for high -performance processing of blob data and archiving processed data into tiered blob storage for optimal data placement and cost management.



Managed solution: Enables the benefits of a Lustre parallel file system for HPC workloads with the expected agility of cloud services, available in most Azure regions and without management headaches. Leverage the latest features and performance improvements with Lustre 2.15 (LTS)



Azure Kubernetes Services: Containerize workloads using the available Azure Kubernetes Services compatible CSI Driver




Public Preview


February 22 '23

General Availability

Late May '23

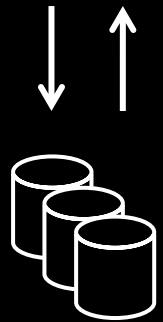
	SSD Performance (MB/s TiB)	Capacity (TiB)*	Throughput (GB/s)*	Pricing ¹ (GB/month)
	250	768	Up to 192	\$0.210
	125	768	Up to 96	\$0.145

*Default Capacity limit = 128TiB
¹EastUS LIST Price

	US	US	Other
	East US	West US 2	Australia East
	East US 2	West US 3	Canada Central
	Central US*	Europe	Central India
	South Central US	North Europe*	Southeast Asia*
		West Europe	Sweden Central*
			UK South

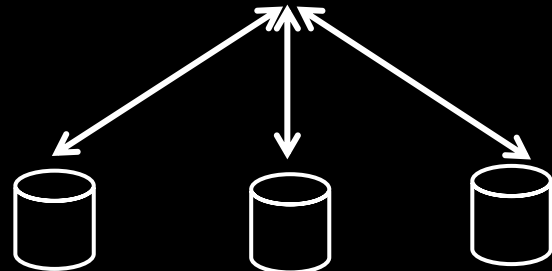
*Coming Soon

Azure Storage: Built in Replication



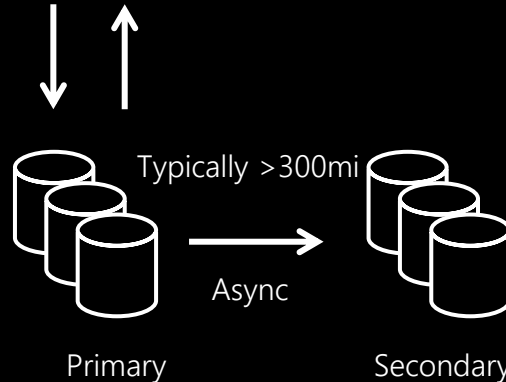
LRS (11 9s)

3 replicas, 1 region
Protect against disk, node, rack failures
Write is ack'd when all replicas are committed
Superior to dual-parity RAID



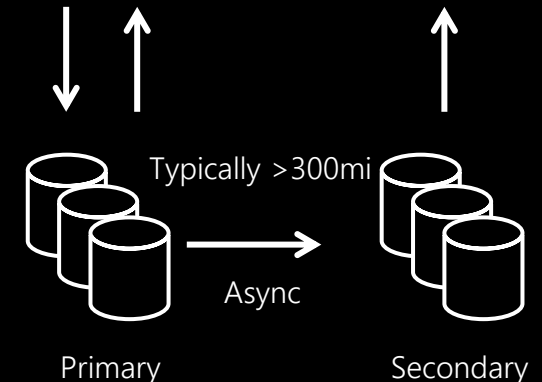
ZRS (12 9s)

3 replicas across 3 Zones
Protect against disk, node, rack and zone failures
Synchronous writes to all 3 zones
Available in 8 regions



GRS (16 9s)

6 replicas, 2 regions (3/region)
Protects against major regional disasters
Asynchronous to secondary



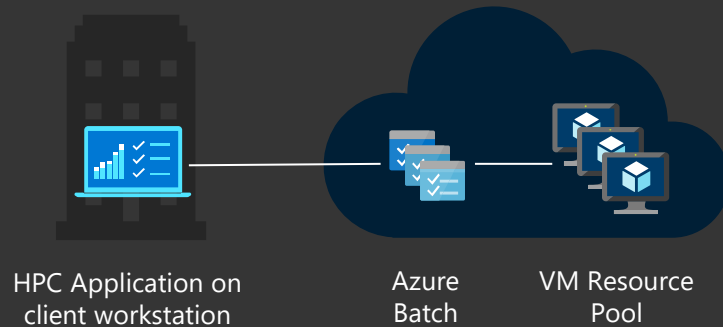
RA-GRS (16 9s)

GRS + Read access to secondary
Separate secondary endpoint
RPO delay to secondary can be queried

Services for HPC Workload Management

Azure Batch

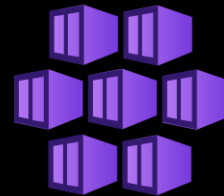
Cloud-native job scheduling



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

Azure Kubernetes Service (AKS)

Fully managed Kubernetes

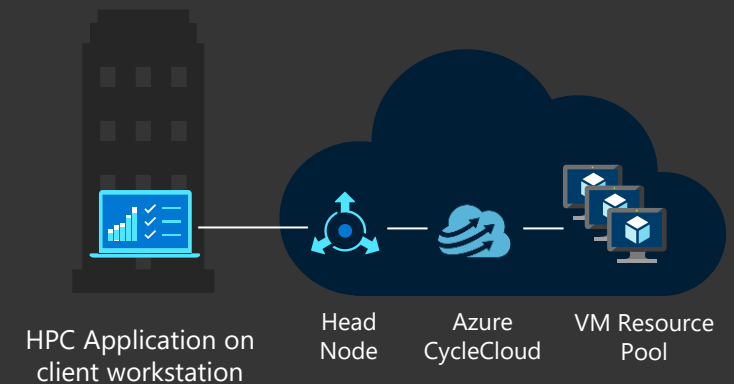


- At no charge* and cost efficient for your applications
- Fully managed K8s in 4 min
- 100% vanilla K8s
- Enterprise-grade security

Azure CycleCloud

Traditional cluster scheduler orchestration

Hybrid Bursting or Cloud native model



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

Support for traditional job schedulers



Unmatched security

- > \$1B annual investments
- > 8500 security experts
- > 43 Trillion daily security signals



"Make no mistake - Microsoft is a security company now"

FORRESTER®

Azure covers over 100 compliance offerings

Azure has the deepest and most comprehensive compliance coverage in the industry

Global	US Gov	Industry		Regional	
<ul style="list-style-type: none">• CIS Benchmark• CSA STAR attestation• CSA STAR certification• CSA STAR self-assessment• ISO 20000-1• ISO 22301• ISO 27001• ISO 27017• ISO 27018• ISO 27701• ISO 9001• SOC 1 Type 2• SOC 2 Type 2• SOC 3• WCAG 2.0 (ISO 40500)	<ul style="list-style-type: none">• CJIS• CNSSI 1253• DFARS + CMMC• DoD IL2• DoD IL4• DoD IL5• DoD IL6• DoE 10 CFR Part 810• EAR• FedRAMP• FIPS 140-2• IRS 1075• ITAR• NIST 800-171• NIST 800-53• NIST CSF• Section 508 VPATs	<ul style="list-style-type: none">• Automotive<ul style="list-style-type: none">• TISAX (Germany)• Education<ul style="list-style-type: none">• FERPA (US)• Energy<ul style="list-style-type: none">• NERC (US)• Financial Services<ul style="list-style-type: none">• 23 NYCRR 500 (US)• AFM + DNB (Netherlands)• AMF + ACPR (France)• APRA (Australia)• CFTC 1.31 (US)• EBA (EU)• FCA + PRA (UK)• FFIEC (US)• FINMA (Switzerland)• FINRA 4511 (US)• FISC (Japan)• FSA (Denmark)• GLBA (US)• KNF (Poland)• MAS + ABS (Singapore)• NBB + FSMA (Belgium)• OSFI (Canada)	<ul style="list-style-type: none">• Financial Services (Cont.)<ul style="list-style-type: none">• OSPAR (Singapore)• PCI DSS Level 1• PCI 3DS• RBI + IRDAI (India)• SEC 17a-4 (US)• SEC Regulation SCI (US)• Shared assessments (US)• SOX (US)• TruSight• Healthcare and Life Sciences<ul style="list-style-type: none">• ASIP HDS (France)• GxP (FDA 21 CFR Part 11)• HIPAA (US)• HITRUST• MARS-E (US)• NEN 7510 (Netherlands)• Media and Entertainment<ul style="list-style-type: none">• CDSA• DPP (UK)• FACT (UK)• MPA• Telecommunications<ul style="list-style-type: none">• GSMA	<ul style="list-style-type: none">• Americas<ul style="list-style-type: none">• Argentina PDPA• Canada Privacy Laws• US CCPA• Asia Pacific<ul style="list-style-type: none">• Australia IRAP Unclassified• China GB 18030:2005• China DJCP (MLPS)• China TRUCS/CCCPPF• China TCS• India MeitY• Japan CS mark gold• Japan My Number act• Korea K-ISMS• New Zealand ISPS• Singapore MTCS Level 3	<ul style="list-style-type: none">• Europe and Middle East<ul style="list-style-type: none">• EU EN 301 549• EU ENISA IAF• EU GDPR• EU model clauses• Germany C5• Germany IT—Grundschutz workbook• Netherlands BIR 2012• Russia personal data law• Spain ENS High• Spain LOPD• UAE DESC• UK cyber essentials plus• UK G-Cloud• UK PASF

<https://aka.ms/AzureCompliance>



Azure global compliance



The following compliance standards apply globally

- [CIS Benchmark](#)
- [CSA-STAR attestation](#)
- [CSA-STAR certification](#)
- [CSA-STAR self-assessment](#)
- [ISO 20000-1:2011](#)
- [ISO 22301](#)
- [ISO 27001](#)
- [ISO 27017](#)
- [ISO 27018](#)
- [ISO 27701](#)
- [ISO 9001](#)
- [PCI DSS](#)
- [SOC](#)
- [WCAG](#)

Azure regions

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

Over
60 regions
worldwide

140 available in
140 countries

- Available region
- Announced region
- ◆ Availability zones

Azure Policy

Active control and governance at scale for your Azure resources



Enforcement and compliance

Enforce policies to help ensure cloud compliance and avoid misconfigurations.



Apply policies at scale

Reduce the time needed to audit your environments by having all your compliance data in a single place.



Remediate and automate

Minimize drift by configuring automated remediation tasks and alerts.





Microsoft Defender for Cloud | Regulatory compliance

Showing 8 subscriptions

Search

- Download report
- Manage compliance policies
- Open query
- Compliance over time workbook
- Audit reports
- Compliance offerings

General

- Overview
- Getting started
- Recommendations
- Security alerts
- Inventory
- Cloud Security Explorer (Preview)
- Workbooks
- Community
- Diagnose and solve problems

Cloud Security

- Security posture
- Regulatory compliance**
- Workload protections
- Firewall Manager
- DevOps Security (Preview)

Management

- Environment settings
- Security solutions
- Workflow automation

Microsoft cloud security benchmark (preview)

33 of 62 passed controls

Lowest compliance regulatory standards

Show all 3

NIST SP 800 53 R5	293/333
HIPAA HITRUST	491/531

Audit reports

Stay up to date on the latest privacy, security, and compliance-related information for Microsoft's cloud services.

[Open](#)

- Microsoft cloud security benchmark
- HIPAA HITRUST
- NIST SP 800 53 R5

Recommendations from Microsoft Defender for Cloud - Regulatory Compliance should not be interpreted as a guarantee of compliance. It is up to you to evaluate and validate the effectiveness of customer controls per your regulatory environment. These services are subject to the terms and conditions in the [licensing terms](#).

NIST SP 800 53 R5 is applied to 3 subscriptions

Expand all compliance controls

AC. Access Control

- AC-1. Access Control Policy and Procedures [Control details](#)
- AC-2. Account Management
- AC-2*. Additional assessments for AC-2 - Account Management

Automated assessments	Resource type	Failed resources	Resource compliance status
A maximum of 3 owners should be designated for subscriptions	Subscriptions	3 of 3	<div style="width: 100%; height: 10px; background-color: red;"></div>
Managed identity should be used in function apps	Web applications	3 of 4	<div style="width: 75%; height: 10px; background-color: red; background-color: limegreen;"></div>
SQL servers should have an Azure Active Directory administrator provisioned	SQL servers	1 of 4	<div style="width: 25%; height: 10px; background-color: red; background-color: limegreen;"></div>
Deprecated accounts with owner permissions should be removed from subscriptions	Subscriptions	0 of 3	<div style="width: 100%; height: 10px; background-color: limegreen;"></div>
External accounts with owner permissions should be removed from subscriptions	Subscriptions	0 of 3	<div style="width: 100%; height: 10px; background-color: limegreen;"></div>

Full-stack coverage with dedicated detections

Compute

- Any server
- Unmanaged K8s
- App Services
- Azure VMSS
- Azure K8s

Service Layer

- Resource Manager
- Key Vault
- Network Layer V1
- Azure DNS

AWS workloads

- Amazon EKS
- Amazon EC2
- Unmanaged Kubernetes
- Unmanaged SQL

GCP workloads

- GKE clusters
- Google Compute
- Unmanaged Kubernetes
- Unmanaged SQL

On-premise workloads

- Kubernetes
- SQL Servers
- Servers

Databases and Storage

- Blob storage
- File storage
- Maria DB
- Azure Cosmos DB
- Azure SQL
- MySQL
- Postgres SQL
- Unmanaged SQL



AI Portfolio Overview

Citizen
Developers



AI Builder

Azure AI

Scenario specific AI services



Bot Service



Cognitive Search



Form Recognizer



Video Indexer



Metrics Advisor



Immersive Reader

Customizable AI models
Cognitive Services



Vision



Speech



Language



Decision

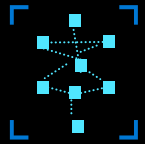
ML platform



Azure Machine Learning

Developers &
Data Scientists

Most comprehensive set of AI capabilities



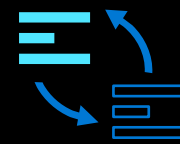
Vision

Image understanding
Text extraction
Image captioning
Form recognition
Video indexing
Facial recognition
Spatial analysis



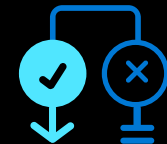
Speech

Speaker recognition
Speech to text
Text to speech
Custom neural voices
Speech translation



Language

Entity extraction
Sentiment analysis
Intent understanding
Text and document translation
Q & A creation



Decision

Anomaly detection
Root-cause analysis
Metric monitoring
Personalization

Machine Learning on Azure



Domain-specific pretrained models

Vision

Speech

Language

Search



Familiar data science tools

Visual Studio Code

Azure Notebooks

Jupyter

Command Line



Popular frameworks

PyTorch

TensorFlow

Scikit-Learn

ONNX



Productive services

Azure DataBricks

Azure Machine Learning

ML VMs



Powerful infrastructure

CPU

GPU

FPGA



*Ensure that artificial
general intelligence (AGI)
benefits humanity*



*Empower every person and
organization on the planet
to achieve more*

GPT-3.5 and GPT-4

Text

ChatGPT

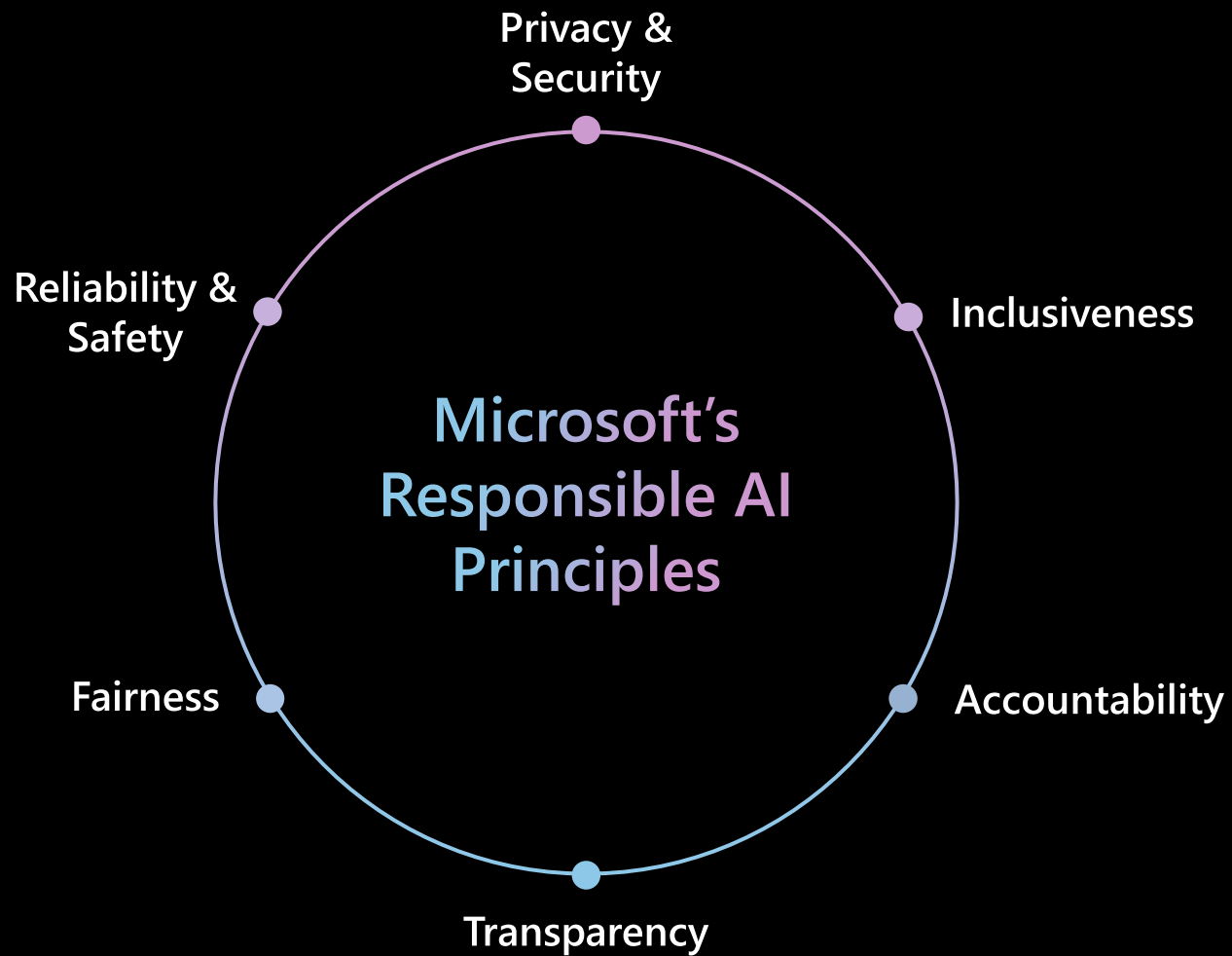
Conversation

Codex

Code

DALL·E 2

Images



Building blocks to enact principles



Tools and processes



Training and practices



Rules



Governance

| Microsoft Azure Cloud

Runs on trust

Your data is your data

Data is stored encrypted in *your Azure* subscription

Your data from any fine-tuning is not used to train the foundation AI models

Azure OpenAI Service provisioned in *your Azure* subscription

Model fine tuning stays in *your Azure* subscription and never moves into the foundation AI models

Your data is protected by the most comprehensive enterprise compliance and security controls

Encrypted with Customer Managed Keys

Private Virtual Networks, Role Based Access Control

Soc2, ISO, HIPPA, CSA STAR Compliant

Thank you & hook 'em!