

Azure HPC



Unlock your innovation with the leading HPC cloud solution

Azure high-performance computing (HPC) is a complete set of computing, networking, and storage resources integrated with workload orchestration services for HPC applications. With purpose-built infrastructure and optimized application services, Azure offers HPC customers a competitive price/performance against on premises with additional benefits. Additionally, Azure includes next-generation machine learning tools to drive smarter simulations and empower intelligent decision making.





Purpose-built

HPC







Workload Orchestration

End-to-end

workflow agility

using known,

familiar tools &

processes

A full range of CPU and GPU capabilities that

> scale to 80K+ cores

help applications

Intelligence Services

Al, machine learning, and deep learning at supercomputer scale Fast, Secure Networking

Fast InfiniBand in ter-connects as well as edge-tocloud connectivity High Performing Storage

A range of storage capabilities to support simpleto-complex storage needs

80,000

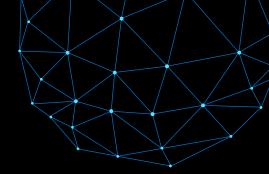
cores utilized for one tightly coupled job

1 billion

cell model of a sports car named after the famed 24 Hours of Le Mans race. 45%

more memory bandwidth means faster application performance

Azure HPC



High Performing H-Series VMs

	HBv2	НВ	НС	<u> </u>
Optimum Workload	Memory Bandwidth	Memory Bandwidth	Dense Compute	Large-Memory HPC
СРИ	AMD EPYC 2 nd Gen "Rome"	AMD EPYC 1st Gen "Naples"	Intel Xeon Platinum 1st Gen Intel Xeon E5 v3 "Skylake" "Haswell"	
Cores/VM	120	60	44	16
TeraFLOPS/VM (FP64)	4 TF	0.9 TF	2.6 TF	0.7 TF
Memory Bandwidth	353 GB/s	263 GB/sec 191 GB/sec		82 GB/s
Memory	4 GB/core, 480 total	4 GB/core, 240 total 8 GB/core, 352 GB		14 GB/core, 224 GB
Local Disk	900 GB NVMe	700 GB <u>NVMe</u>		2 TB SATA
InfiniBand	200 Gb HDR	100 Gb EDR		56 Gb FDR
Network	32 GbE	32 GbE		16 GbE

GPU-powered N-Series VMs

compute visualize learn

	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 <u>NVMe</u> Flash
Network	Azure Network + InfiniBand	Azure Network + InfiniBand EDR+ NVLink GPU interconnect	Azure Network + InfiniBand

	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)			

	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)			

Storage Options



Azure Blob Storage

For stage-able workloads, pleasantly parallelized for single node, single user configuration



Azure Data Lake

For HDFS/analytics workloads, pleasantly parallelized for single node and shared data access



Premium/Ultra SSD

For low scale NFS workloads with <1.5Gbps, <500 cores, 100TB SSD



Azure HPC Cache

For medium scale, read-heavy NFS workloads with <2Gbps Write and <14Gbps Read and 6-24 nodes



Azure NetApp Files

For balanced, write-heavy NFS workloads with <6.5 Gbps Write and <2Gbps Read and <4000 cores

Lustre



Orchestrated FS

For large-scale, parallelized workloads with <50 Gbps Write, <500 TB SSD and <50,000 cores



Cray ClusterStor

For bare metal HPC storage with >30 Gbps Write/Read, >500 TB, and >50,000 cores