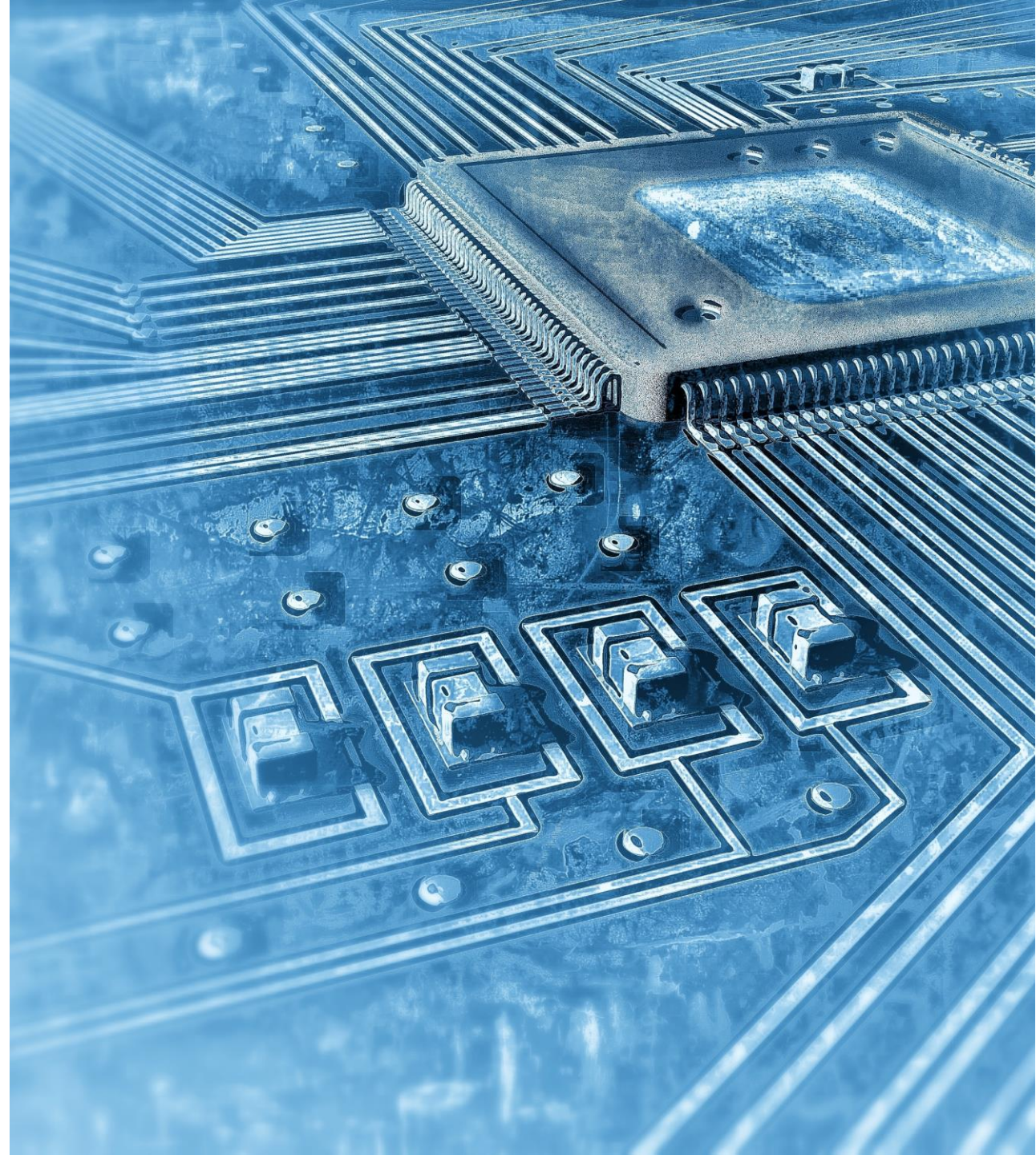




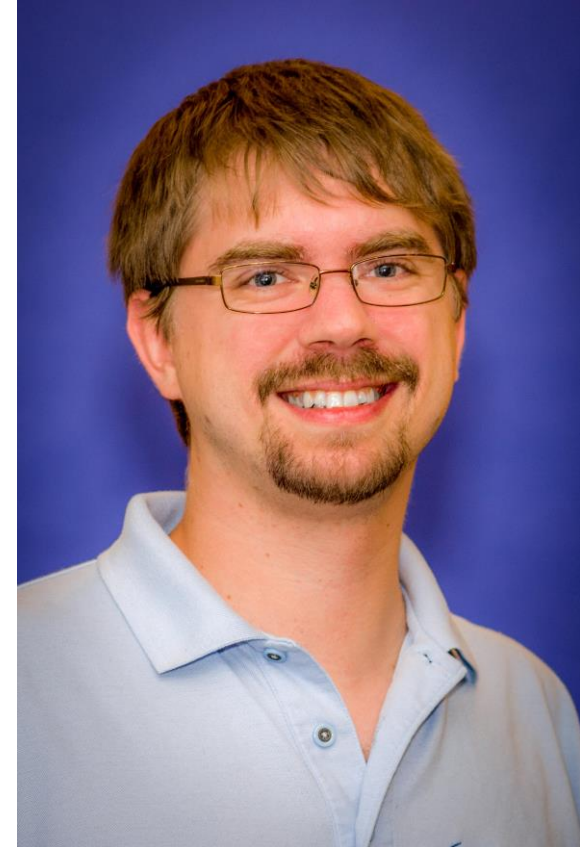
SC20 vSCC Cloud Webinar

September 2, 2020



Welcome and Introduction

- This will be a short(ish) presentation, followed by a longer Q&A session – please put questions in chat
- We're recording this session, the recording and these slides will be posted on the webinars page
- There will be follow-up conversations/webinars/tutorials with more details about Azure and the cloud component of the competition as they become available



Andy Howard
Azure HPC

Agenda

- Overview of HPC on Azure
- Testing vs Competition budgets
- Access and Quotas
- Q&A



HPC on Azure

Accelerate | Connect | Excite



Azure HPC Strategy

Accelerate “time to solutions” by enabling easy on-ramp to differentiated HPC cloud solution



- Process more data
- Accelerate a backlog of simulations
- Analyze more models in parallel
- Run larger analyses or simulations

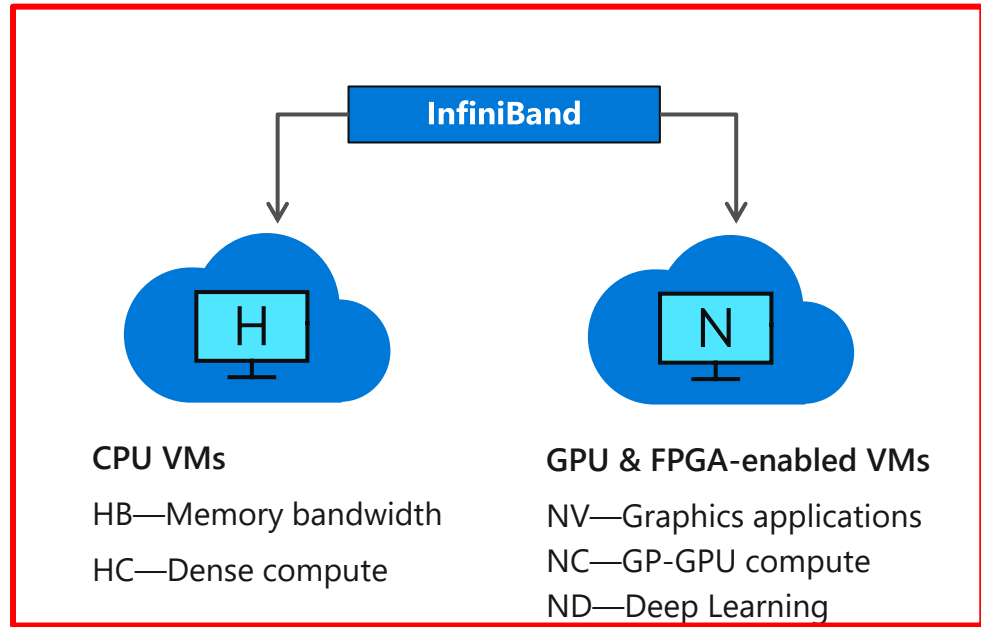


- Designed based on deep understanding of HPC workloads
- Focus on delivering HPC optimized cloud not cloud-optimized HPC
 - Goal: increased Perf/\$ at scale, TTM with new tech
 - Architected around end-to-end workflows
- Tested on real world applications vs. synthetic benchmarks



- HPC expertise in product, engineering, CAT, Sales
- Bring AI expertise to accelerate convergence of HPC and AI
- Solutions build with trusted industry partners
- Flexible cloud options (e.g., RI)

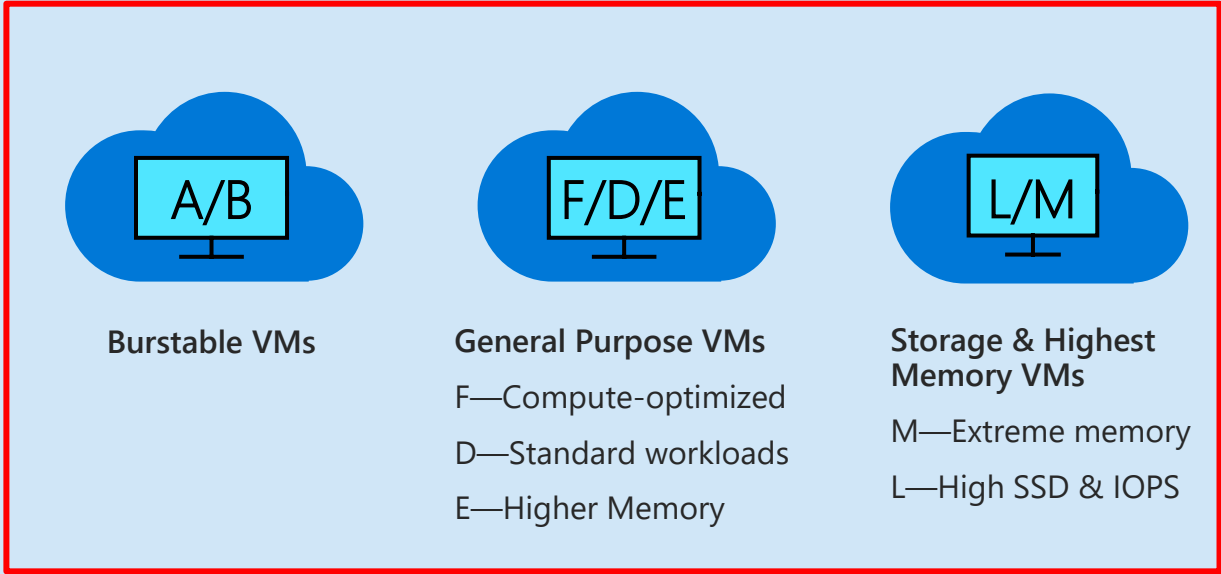
Azure Specialized Compute for HPC, AI, and Viz



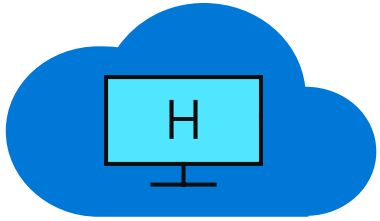
Cray in Azure
Managed, Custom bare-metal
Large to extreme-scale HPC
Azure Network integration



>80,000 IOPs
Premium Storage
Low latency, high
throughput apps



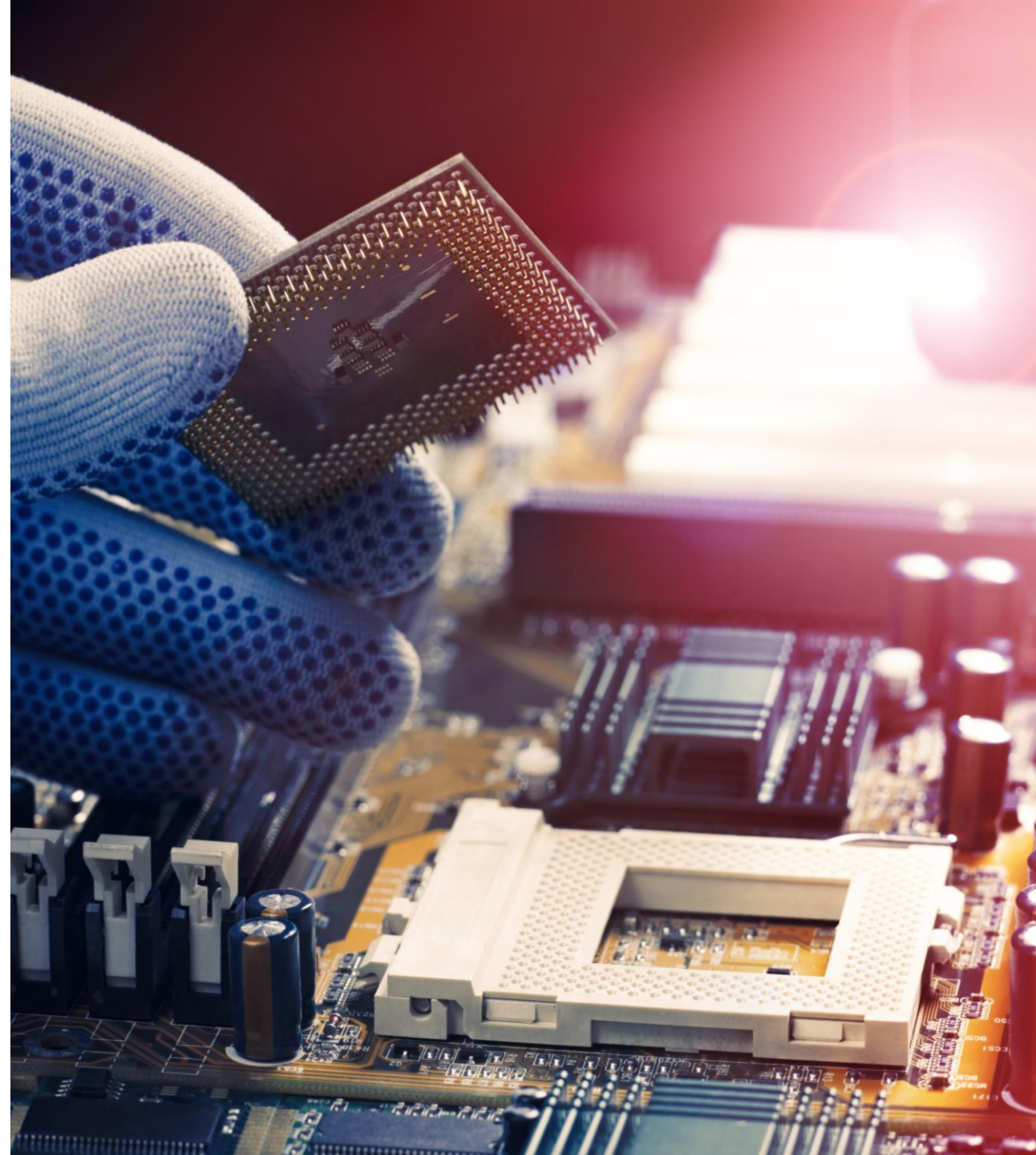
H Series VMs



CPU VMs

HB—Memory bandwidth

HC—Dense compute



High-Performance Computing VMs (H)

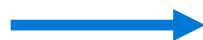
	Available Now	Available Now	Available Now	Available Now
	HBv2	HB	HC	H
Workload Optimized	Memory Bandwidth	Memory Bandwidth	Dense Compute	Large-Memory HPC
CPU	AMD EPYC 2 nd Gen "Rome"	AMD EPYC 1 st Gen "Naples"	Intel Xeon Platinum 1 st Gen "Skylake"	Intel Xeon E5 v3 "Haswell"
Cores/VM	120	60	44	16
TeraFLOPS/VM (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 NVMe		2 TB SATA
InfiniBand	200 Gb HDR	100 Gb EDR		56 Gb FDR
Network	32 GbE	32 GbE		16 GbE



On-Prem Compete

90-95% of market

Dell PowerEdge R740 HPE ProLiant DL380



V.

Azure HBv2



~\$50k 5-year cost
1x HPC performance

~\$50k 5-year cost
1.4-2.2x HPC performance



Cloud Compete

5-10% of market

AWS max of ~1,200
cores w/ 100 GbE

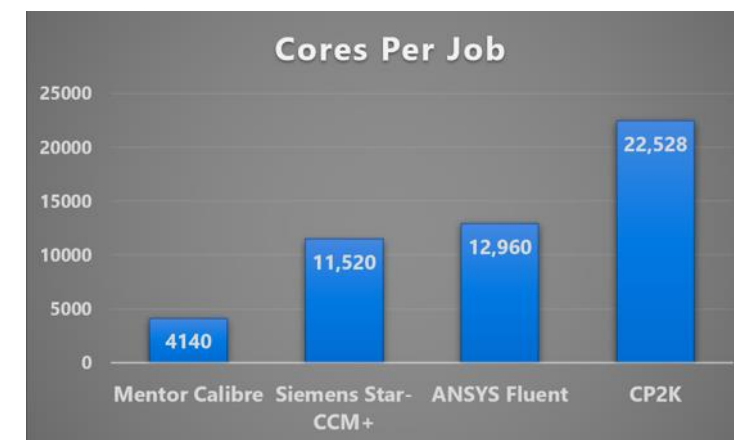


CFD++ (Metacomp Technologies)

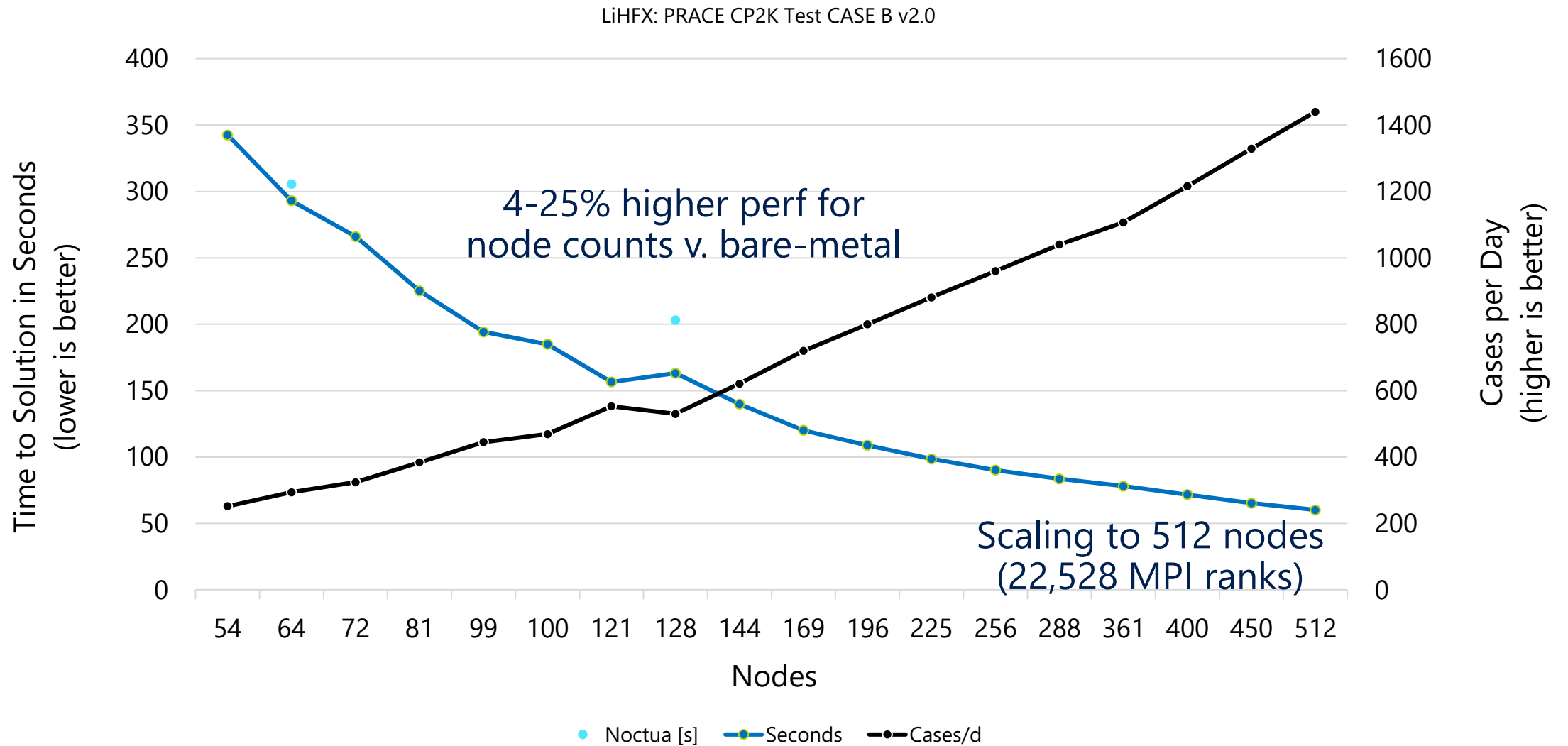


V.

Azure scales **17x higher**, with
target of 60x higher (HBv2) in 2020

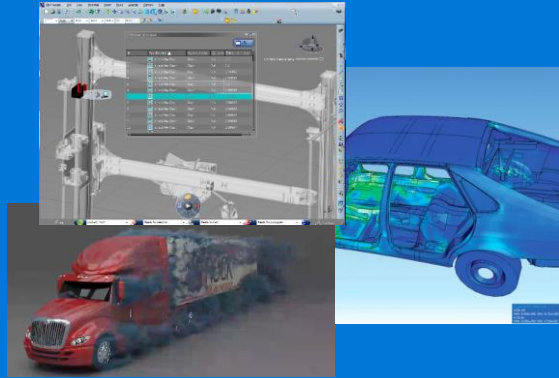


CP2K on Azure HC



GPU Computing

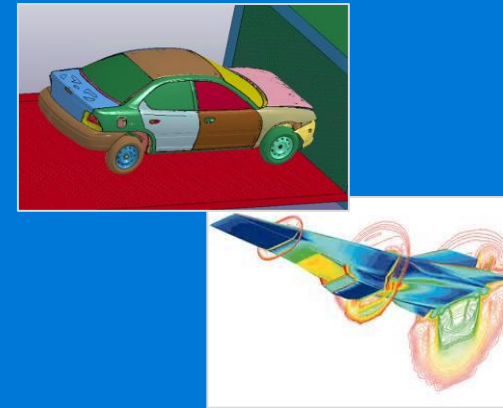
Visualization & Gaming



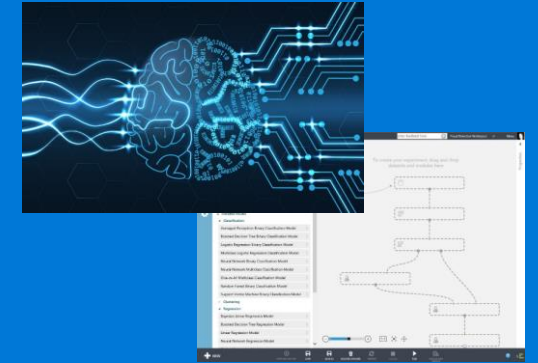
Rendering



HPC/Simulation



Deep-Learning/AI



NV

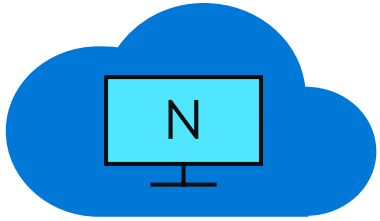
NC

ND

- Workload centric PCs & Workstations
- Variable configuration for GPU workstations for content consumption
- Modern workspace for interactive collaboration

- Scale out using IB for multimode HPC and ML workloads on any MPI stack
- Scale-up multi GPU VMs with fast NVLINK interconnect for high-density training and HPC workloads

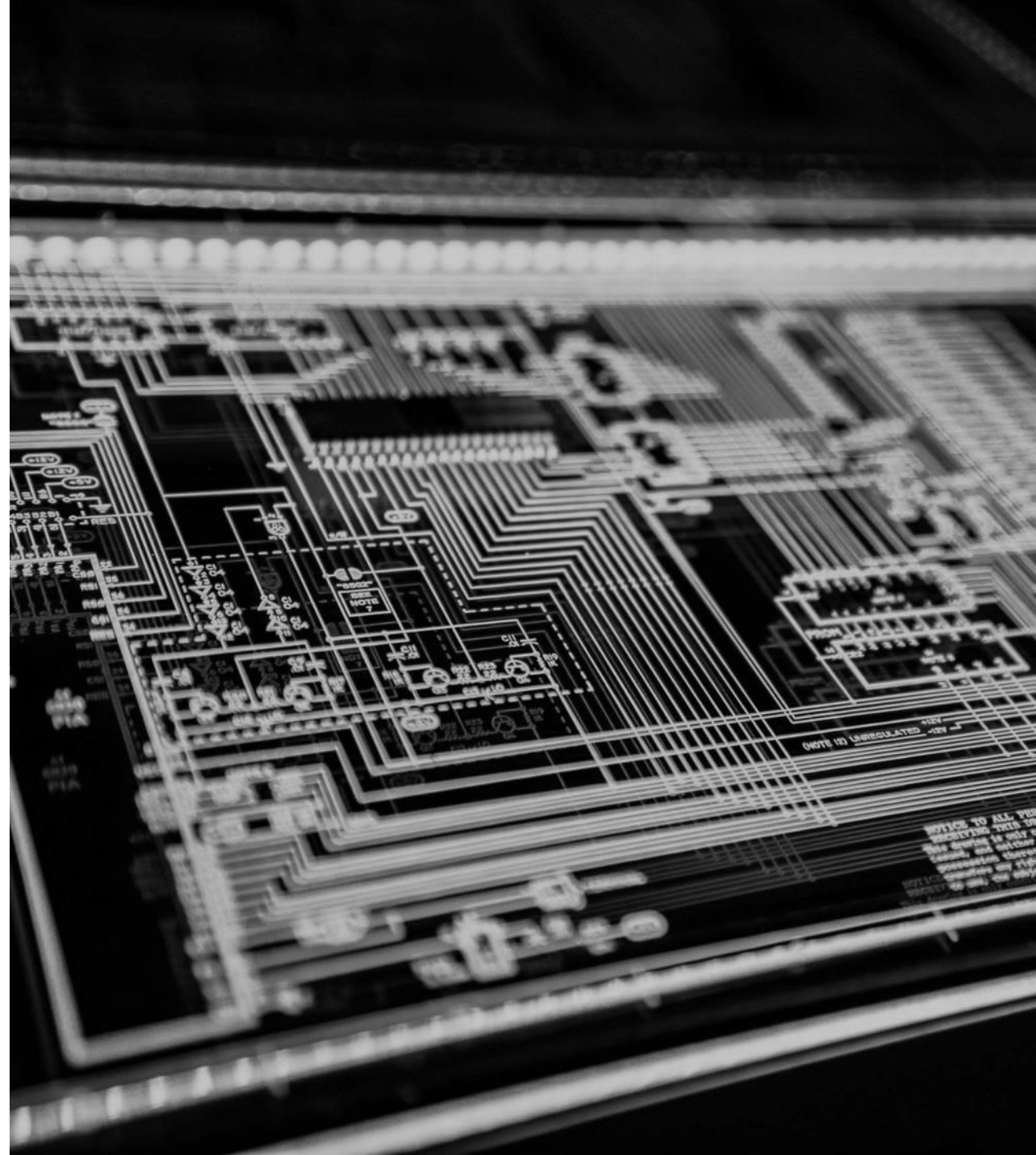
Nc/Nd Series VMs



GPU enabled VMs

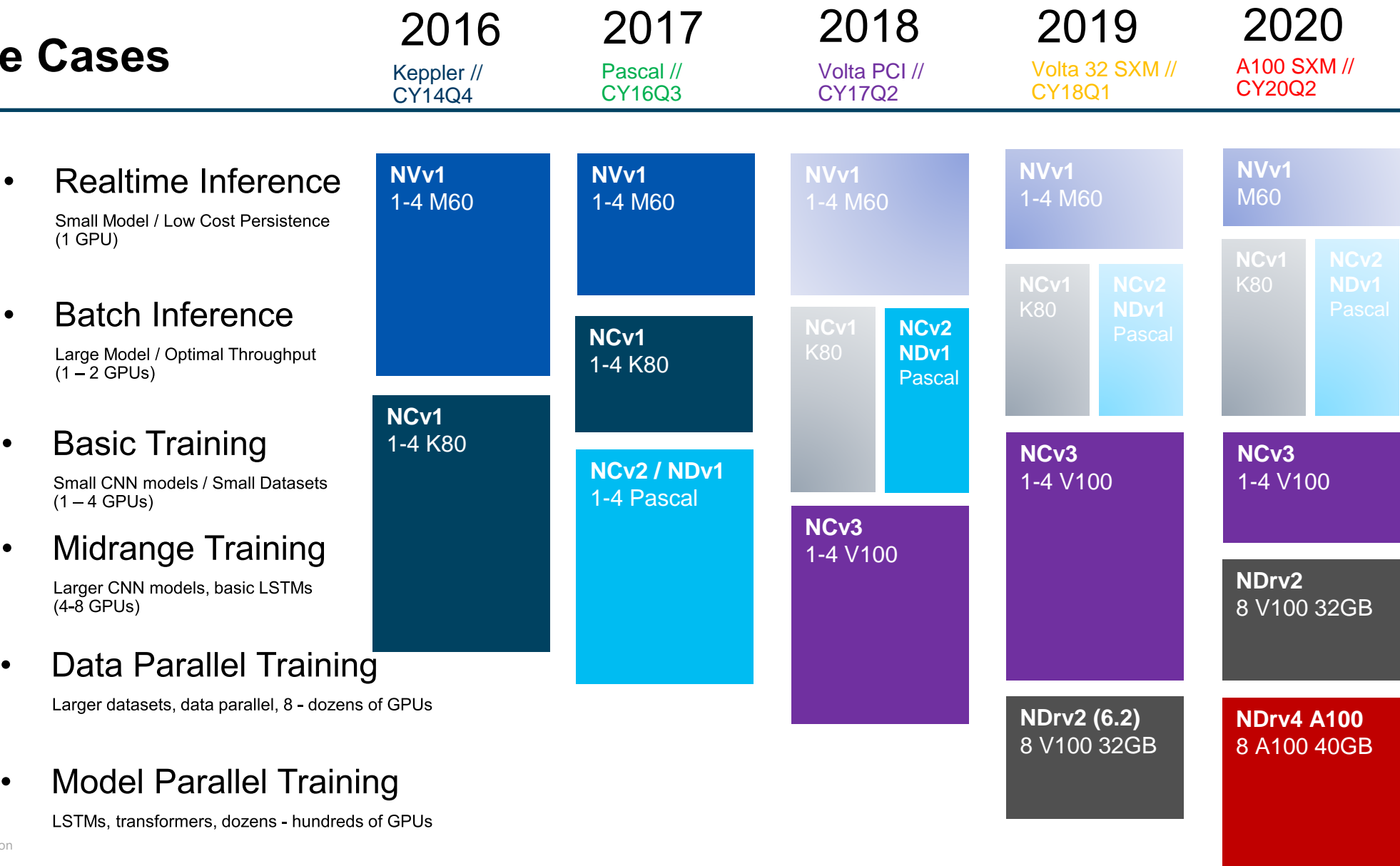
NC—GP-GPU compute

ND—Deep Learning



GPU AI Compute Workloads Are Continuously Evolving

Use Cases



NDR v2 – Volta Generation GPU Compute

Flagship offering accelerating machine learning and HPC workloads

Volta SXM GPU instances – 8X NVIDIA V100 GPUs interconnected with NVLink, modified hypercube topology

Tensor Core technology to deliver over 100 TFLOPS of deep learning performance

Skylake based processor with premium storage support (SSD backed) and InfiniBand Cluster Interconnect

Specs:

5120 CUDA Cores / 640 NVIDIA Tensor Cores per GPU

FP64 - 7.8 TFLOPS of double precision floating point performance

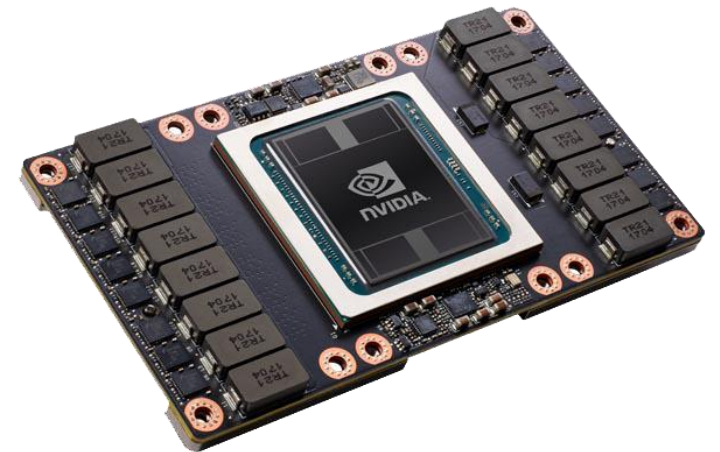
FP32 – 15.7 TFLOPS of single precision performance

GPU Memory 32 GB

300 GB/s GPU interconnect from each Nvlink lane, up to 6 lanes per GPU

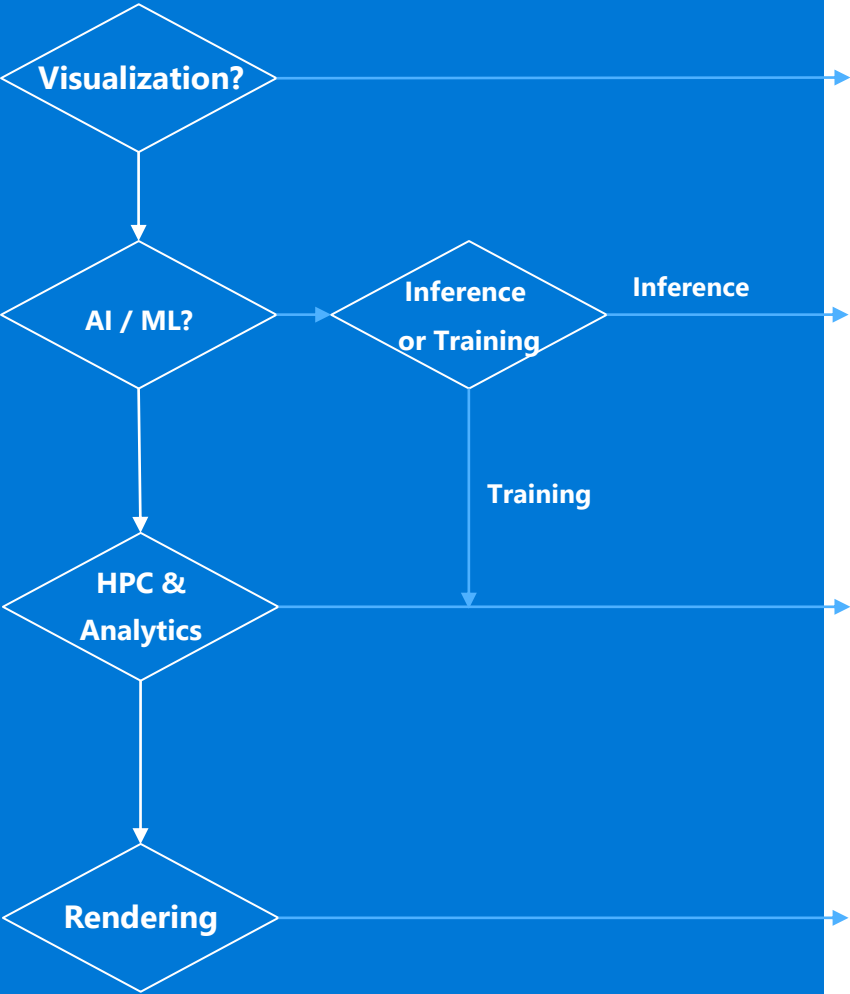
100 Gigabit InfiniBand EDR with Mellanox OFED on each VM

	ND40s_v3
Cores	40 cores (Skylake 2 x Intel Xeon 8168)
GPU	8 x V100 SXM
Memory	768 GB
Local Disk	~1.3 TB SSD
Network	Azure Network + NVLink + Infiniband



GPU VM Triage

Start at top and work down to find a GPU VM Solution



Workload Fit:

VM / GPU Solution:



Large Dataset (CFD / FEA):
Conventional CAD / Modeling:

NV_v3 NV M60
M60



Large Model:
Large Batch Size:
General Purpose:
Simple Models:

NC_v3 V100 PCIe
ND P40
NC_v2 P100
NC K80



Cost-effective development VM:
Cost-effective deployment:
Multi-GPU optimized (6-8 GPUs):
Large jobs (8-500 GPUs):
Exploration & Education:

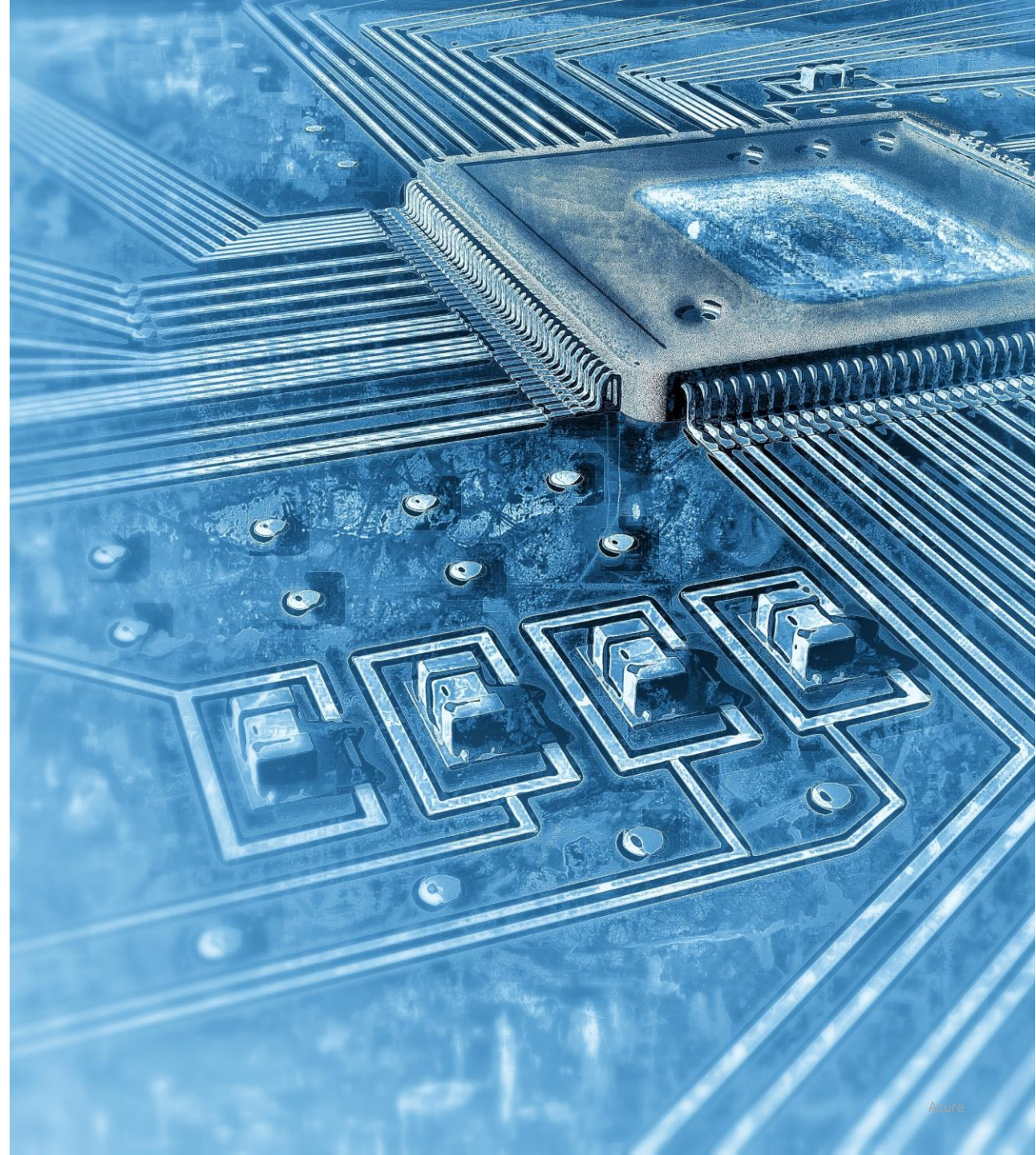
NC_v2 P100
NC_v3 V100 PCIe
NDR_v2 V100 SXM + EDR
NCr_v3 V100 PCIe + FDR
NC K80



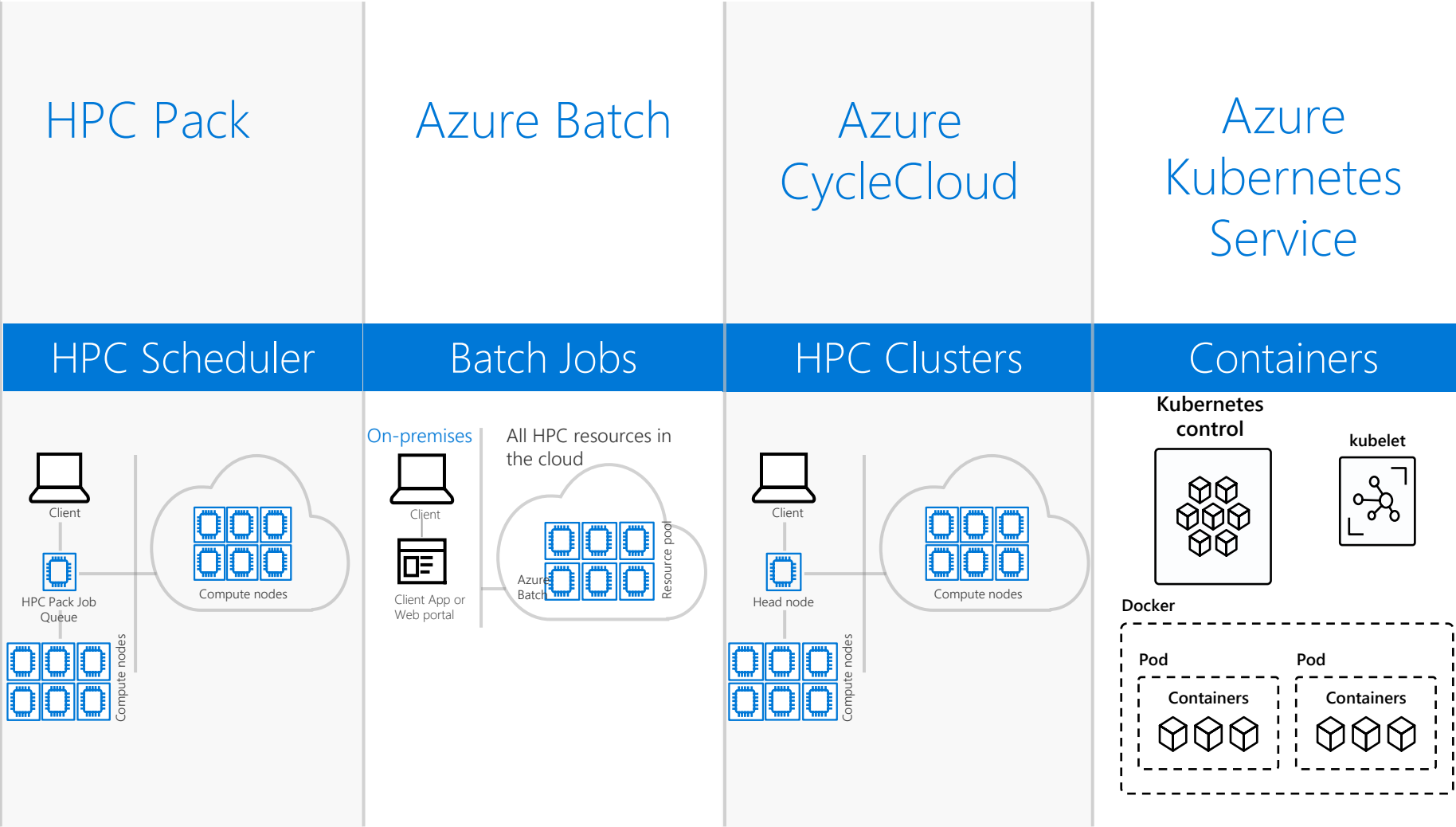
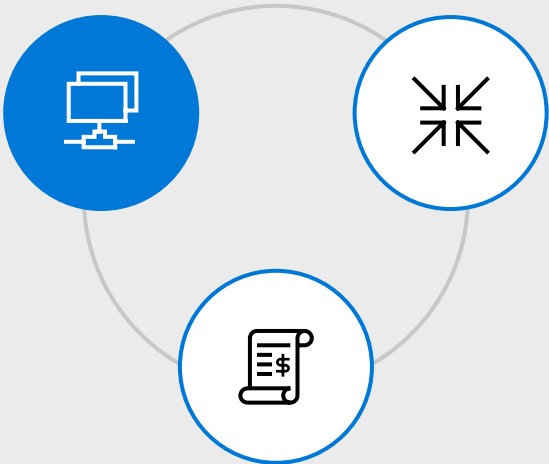
General Purpose:
Large Textures & High Resolution:

NC_v2 P100
ND P40

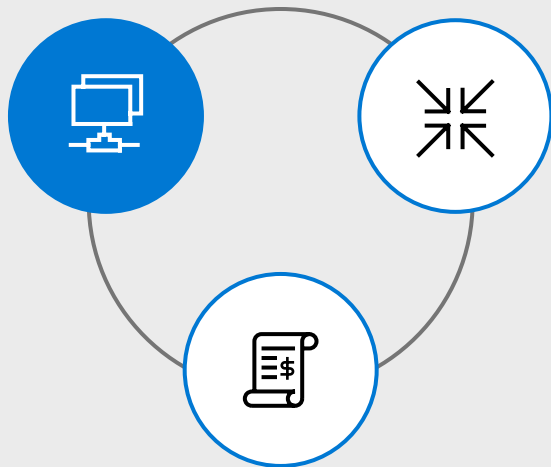
HPC Software Platform



Services for Workload Management



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

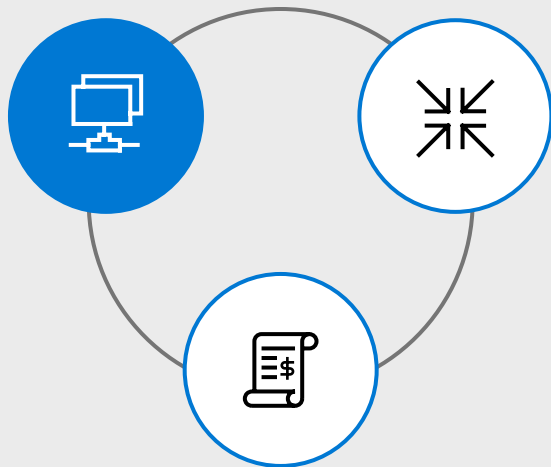


Business management

Able to link usage to spend

Provide tools to manage and control costs

Azure CycleCloud



Traditional Scheduler Orchestration

Scheduler Support

Provides autoscaling and
orchestration for:

Slurm

PBS Professional

IBM Spectrum LSF

IBM Spectrum Symphony

(Univa) Grid Engine

+ others



Competition budgets & access

Accelerate | Connect | Excite



Competition Budgets

- Different testing budgets for each month leading up to the competition
- Will not be as large as the competition budget
- May go up and down depending on the month
- Monthly testing budgets will be announced by Sept. 11 once the committee has finalized them
- Two great ways to approximate price of a cluster:
 - Azure CycleCloud pricing information
 - Azure Pricing Calculator

<https://azure.microsoft.com/en-us/pricing/calculator>

The image shows two overlapping screenshots from the Azure portal. The top screenshot is the 'Select one or more machine types' page, which includes filters for Compute Type, Availability, Networking Support, Disk Support, Ephemeral Support, and Low Priority Support. It displays a table of machine types with columns for SKU, Tier, Cores, Memory, \$/Hour, \$/Core, Quota, Available, Low Prio, SR-IOV, RDMA, Ephemeral, Data Disks, and Storage. Two machine types are listed: HB60rs and HB120rs_v2. The bottom screenshot shows the 'pbstest' cluster overview, including buttons for Terminate, Edit, Access, and Refresh, and a summary of the cluster's state, nodes, users, size, usage, and alerts.

SKU	Tier	Cores	Memory	\$/Hour	\$/Core	Quota	Available	Low Prio	SR-IOV	RDMA	Ephemeral	Data Disks	Storage
HB60rs	Standa...	60	223.52 GB	\$2.51	\$0.042	12000	12000	yes	no	yes	yes	4	700.00 GB
HB120rs_v2	Standa...	120	468.75 GB	\$3.96	\$0.033	92000	91040	yes	no	yes	yes	8	937.50 GB

pbstest

[Terminate](#) [Edit](#) [Access](#) [Refresh](#)

State **Started** at 8/24/20 12:17 PM (up 8d 22h 49m) - [View in Portal](#)

Nodes **1** ready

Users **1** admin [Show](#)

Size **1** instance, **4** cores (**\$0.33** per hour)

Usage **96.0** core-hours (~\$9) in the last 24 hours

Alerts [Create new alert](#)

Issues No issues found

Access to Azure

- Each team will receive login information for a dedicated CycleCloud install/bastion VM
 - ***Note: It is highly recommended that you don't enable public IPs or password logins on your clusters!!!***
- Access will be restricted to a single Resource Group in Azure and dedicated VNets/Subnets
- VM Family quotas will be set ahead of the competition to ensure fair access to resources
 - Quotas for HPC VM types will only be increased if teams ask for them!
 - During testing, some reasonable quotas will be set, but will likely be lower than during the actual competition
- Team advisors will get login information by Sept. 21

Why CycleCloud?

- Easier for committee to setup and manage environments
- Easier for teams to get started without having to learn intricacies of Azure
- Out-of-the-box autoscaling capabilities to keep costs down
- Realtime cost reporting across clusters managed by CycleCloud down to the minute
 - A special plugin will be installed to allow teams to query their Azure spend, both the total for each month and the current hourly and minute burn rates
 - Getting started resources for CycleCloud are available on the Microsoft Docs site:
<https://docs.microsoft.com/en-us/azure/cyclecloud/?view=cyclecloud-8>



Q&A

Accelerate | Connect | Excite



