Azure supports a full spectrum of high performance computing scenarios whether you are interested in running powerful visual simulations in engineering; dynamic rendering for GPU intensive work or scientific and data scenarios using analytics languages like R.

If you have a compute-intensive problem, with HPC you can scale up or down your compute resources quickly according to the CPU, memory and GPU required.

Monte Carlo simulation in R Studio, running on Azure
**Azure HPC Infrastructure**

Azure offers virtual machines (VMs) with fast CPUs for simulations or analytics, large memory for parsing databases with in-memory support, or GPU support. Of course, all options support both Windows and Linux-based workloads as well as a variety of open source development, automation and management platforms.

---

**Graphics Intensive**

With graphics-intensive scenarios, Azure’s N-series virtual machines specifically are optimized for native-like access to GPU resources, offering near bare metal performance levels.

---

**NV-series VMs** are great for running computer-aided design and resource-intensive visual simulations, allowing engineers to access powerful, Azure-hosted compute from just about any workstation and without needing to move terabytes of modeling data between end points and servers.

---

**Cinematic, graphic intensive workload running on N series VMs in Azure**

---

**Azure-hosted compute: work from any device without moving terabytes of data**

---

**GPUs can of course be used for improved visualization, but the speed of the GPU also enables faster computation.**

---

**HPC infrastructure options in Azure**
Beyond graphical scenarios the ND series supports artificial intelligence workloads like model training and deep learning workloads.

<table>
<thead>
<tr>
<th>Size</th>
<th>vCPU</th>
<th>Memory: GiB</th>
<th>Temp Storage (SSID) GiB</th>
<th>GPU</th>
<th>Max data disks</th>
<th>Max NICs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard_ND6s</td>
<td>6</td>
<td>112</td>
<td>736</td>
<td>1</td>
<td>12</td>
<td>4</td>
</tr>
<tr>
<td>Standard_ND12s</td>
<td>12</td>
<td>224</td>
<td>1474</td>
<td>2</td>
<td>24</td>
<td>8</td>
</tr>
<tr>
<td>Standard_ND24s</td>
<td>24</td>
<td>448</td>
<td>2948</td>
<td>4</td>
<td>32</td>
<td>8</td>
</tr>
<tr>
<td>Standard_ND24rs*</td>
<td>24</td>
<td>448</td>
<td>2948</td>
<td>4</td>
<td>32</td>
<td>8</td>
</tr>
</tbody>
</table>

Additionally, NCv3 series VMs are powered by NVIDIA Tesla V100 GPUs. These GPUs can provide 1.5x the computational performance of the NCv2 series. Common scenarios that apply include: reservoir modelling, DNA sequencing, protein analysis and others.

<table>
<thead>
<tr>
<th>Size</th>
<th>vCPU</th>
<th>Memory: GiB</th>
<th>Temp Storage (SSID) GiB</th>
<th>GPU</th>
<th>Max data disks</th>
<th>Max NICs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard_NC6s_v3</td>
<td>6</td>
<td>112</td>
<td>736</td>
<td>1</td>
<td>12</td>
<td>4</td>
</tr>
<tr>
<td>Standard_NC12s_v3</td>
<td>12</td>
<td>224</td>
<td>1474</td>
<td>2</td>
<td>24</td>
<td>8</td>
</tr>
<tr>
<td>Standard_NC24s_v3</td>
<td>24</td>
<td>448</td>
<td>2948</td>
<td>4</td>
<td>32</td>
<td>8</td>
</tr>
<tr>
<td>Standard_NC24rs_v3*</td>
<td>24</td>
<td>448</td>
<td>2948</td>
<td>4</td>
<td>32</td>
<td>8</td>
</tr>
</tbody>
</table>

Because Azure is the only public cloud to provide InfiniBand and RDMA to connect thousands of instances across multiple GPUs and compute nodes, these services provide extremely fast compute times at low latency, especially compared to CPU-only architectures.

And if your jobs require the massive scale of a Cray supercomputer, we have an exclusive partnership with Cray that allows you to get a dedicated supercomputer – running within your Azure virtual network.

If you have an on-premises cluster for your HPC workloads that is too infrequent to justify the expense and management overhead, Azure HPC frees you of that burden, because you only pay for what you use.
Running HPC workloads in Azure

To make running your Azure workloads easier, there are a few core services that can help.

Azure Batch can run complex batch processes reasoning over large amounts of data and compute resources as a service, without the need to manage individual VMs. This means that you can place more focus on running your workload versus provisioning the compute to run it. It’s also worth noting that within Azure Batch, there are additional accelerators available for AI and rendering.

Azure provides granular monitoring of usage and associated costs. With per sub-second billing and Low-Priority virtual machines — which provides preemptible compute at up to 80% off — allowing you to only pay for what you need.

Of course, you can use virtual machine scale sets or Azure Resource Manager templates to help you deploy, manage and scale your solutions as well.

Once you’re up and running, Azure offers flexible consumption — you can change your infrastructure as your workload or business needs change.

To provision, monitor, auto-scale and manage the lifecycle of HPC compute clusters, Azure CycleCloud, allows IT groups to provide managed clusters in Azure.

If you are interested in the comprehensive set of compute options in Azure, there are dedicated topics for IaaS and modern compute in Azure Essentials. Of course, we are constantly adding new topics on Essentials, so please keep checking back for more at www.azure.com/essentials

Continued Learning

Microsoft Mechanics
Azure High Performance Computing
The latest updates to Microsoft’s Remote Desktop Services
Azure Essentials: High Performance Computing (HPC) options in Azure

Additional Azure Learning Resources
Virtual Machine Scale Sets Quickstarts
Batch Quickstart
Microsoft Azure Architecture - Getting Started

Azure Batch
Batch Rendering
Batch AI
Azure CycleCloud