



Nor-Tech Benefits of On Prem Over Cloud HPC

When it comes to High Performance Computing (HPC), on-premises data centers still offer several distinct benefits over cloud-based HPC. Here's a detailed look at the advantages of on-premises HPC over cloud-based HPC:

1. Predictable Performance and Low Latency

- **Consistent Performance:** On-premises HPC systems can provide more consistent and predictable performance since resources are dedicated to specific workloads. In contrast, cloud environments, which are shared, can experience variability in performance due to network congestion or competition for resources.
- **Low-Latency Networking:** HPC often requires low-latency communication between compute nodes, especially for applications like simulation or large-scale parallel processing. On-premises environments allow organizations to use high-performance interconnects like InfiniBand or high-bandwidth Ethernet that offer lower latency than cloud networks.
- **Custom Hardware Configurations:** On-premises infrastructure can be optimized specifically for HPC workloads, allowing for fine-tuned network, storage, and compute architectures that can be challenging to replicate in a generic cloud setup.

2. Cost Control for Long-Term, Intensive Workloads

- **Economies of Scale Over Time:** While cloud services operate on a pay-as-you-go model, which is ideal for burst workloads¹, on-premises HPC systems become more cost-effective over time, especially for organizations that run large-scale, consistent workloads. After the initial CAPEX² on infrastructure, the OPEX³ can be lower in the long run.
- **No Data Transfer Costs:** HPC workloads often generate massive datasets that need to be stored, transferred, and processed. Cloud providers typically charge for data ingress, egress, and storage, which can add significant costs, especially for applications involving large volumes of data (e.g., genomics, physics simulations). On-premises HPC eliminates these data transfer costs.
- **Avoiding Pay-Per-Use for Constant Workloads:** For organizations running HPC applications continuously, cloud pricing models get expensive. On-premises infrastructure, while requiring upfront investment, avoids the pay-per-use fees that accumulate over time in the cloud.

3. Data Security and Control

- **Full Control Over Data Security:** On-premises HPC environments give organizations full control over the security of their data, including access control, encryption, and physical security. This is particularly important for industries like defense, healthcare, finance, and scientific research, where data sensitivity is a significant concern.
- **Compliance and Data Sovereignty:** For companies bound by strict regulatory requirements (e.g., HIPAA⁴, GDPR⁵, ITAR⁶), it may be easier to manage compliance in on-premises environments.

¹ The extra workload transfers to external third-party cloud services.

² Capital expenditure

³ Operational expenditure

⁴ Health Insurance Portability and Accountability Act of 1996

⁵ General Data Protection Regulation: a 2016 European Union law to protect the privacy and security of personal data

⁶ The International Traffic in Arms Regulations is a U.S. regulation that controls the manufacture, sale, and distribution of defense and space-related products and services.

They can ensure that sensitive data never leaves their own controlled facilities, avoiding potential risks associated with multi-tenant cloud environments.

4. Customization and Optimization for Specific Workloads

- **Tailored Hardware for HPC:** On-premises setups allow organizations to customize hardware configurations specifically optimized for their HPC needs. For example, they can use GPUs, FPGAs, specialized CPUs, or even custom accelerators designed for specific types of computations, which are not as easily accessible or cost-effective in the cloud.
- **Storage Flexibility:** On-premises environments allow for high-performance storage systems like NVMe⁷, SSDs⁸, or specialized HPC file systems (e.g., Lustre, GPFS) that are finely tuned for the specific I/O patterns of HPC workloads. While some cloud providers offer high-performance storage, it's often at a premium and may not match the tailored configurations possible in-house.
- **Optimized Software Stack:** HPC environments often rely on a specific software stack for performance optimization, such as custom-built compilers, libraries, or middleware. In on-premises systems, you have full control over the software environment and the ability to deeply customize it to achieve maximum performance.

5. Predictable and Controlled Environment

- **Dedicated Resources:** In an on-premises setting, HPC workloads run on dedicated hardware with no competition for resources. In cloud environments, despite provisioning dedicated instances, underlying infrastructure can still be shared, leading to potential resource contention.
- **No Downtime for Scheduled Maintenance:** Cloud services may have scheduled maintenance windows, during which services can be interrupted or performance degraded. In an on-premises environment, you have full control over maintenance schedules, minimizing disruption to critical HPC workloads.
- **Energy and Cooling Efficiency Management:** Today's large HPC systems require advanced cooling to maintain efficiency. On-premises setups give organizations more control over energy and cooling, enabling optimization for specific environments, particularly in energy-sensitive or environmentally conscious industries.

6. Customization of Networking Topology

- **Custom Interconnects:** HPC workloads often demand high-speed, low-latency networking to maintain performance for parallel processing. On-premises HPC systems can be built with specialized interconnects (e.g., InfiniBand, 100G Ethernet) designed for optimal performance, which can be prohibitively expensive or unavailable in the cloud.
- **Network Topology Control:** In an on-premises environment, businesses can customize network topology to meet the needs of specific HPC workloads. This allows them to deploy architectures like fat-tree, butterfly, or torus networks, which are common in HPC but less flexible in cloud environments.

7. Data Privacy and IP Protection

- **Intellectual Property Concerns:** Companies handling sensitive intellectual property may prefer on-premises HPC because it avoids the perceived risk of entrusting valuable data to third-party cloud providers. This is especially important in industries like pharmaceutical, aerospace, or autonomous vehicle research, where proprietary algorithms and data are mission-critical.
- **Sensitive Research Applications:** For research institutions and government bodies working on confidential projects (e.g., national security or advanced scientific research), keeping workloads on-premises ensures data privacy and minimizes the risks of third-party access.

8. Superior Resource Management for Specialized Jobs

- **Batch Processing Control:** HPC workloads often use batch processing systems like SLURM, PBS, or HTCondor, which may be more robust or easier to manage in an on-premises environment. While cloud providers offer similar services, these environments may require additional configuration and costs.

⁷ Non-Volatile Memory Express

⁸ Solid-State Drive

- **Job Scheduling and Priority Management:** On-premises systems give organizations more granular control over how resources are allocated, prioritized, and scheduled. This can be crucial for organizations that need to run a mix of jobs with different priorities and execution times.

While cloud computing offers flexibility, scalability, and convenience, on-premises HPC has distinct advantages when it comes to performance, cost control for long-term workloads, data security, and customization. For companies with specific regulatory requirements, predictable workloads, and a need for high-performance, low-latency computing environments, on-premises HPC remains a competitive choice over cloud.

About Nor-Tech

Nor-Tech built its reputation on the industry's easiest-to-deploy cluster solutions and guaranteed no wait time support. The company designed and built the HPC cluster that enabled the first detection of a gravitational wave—a discovery destined to change history. In addition to HPC clusters, Nor-Tech's custom technology includes workstations, desktops, and servers for a range of applications including CAE, CFD, and FEA. Clients include some of the most respected organizations in the world. Nor-Tech engineers average 20+ years of experience and are responsible for significant high performance computing innovations. The company has been in business since 1998 and is headquartered in Burnsville, Minn. just outside of Minneapolis.