Acknowledgments

Lead Authors
Andrew Howard
Guan-Sin Ong

Contributors
Hing-Yan Lee
Michael Roza
Jernej Zidar

CSA Global Staff
Ekta Mishra
Hing-Yan Lee
Claire Lehnert

About the High Performance Computing (HPC) Cloud Security Working Group
The crossing of cloud and HPC environments often leads us to question how security in an HPC cloud environment can be implemented, enforced, and ensured without compromising performance. This working group strives to provide recommendations that can answer these questions.
Introduction

According to the 2020 Research and Markets report, the High Performance Computing (HPC) market is projected to grow from $37.8 billion USD in 2020 to USD $44.98 billion USD in 2025. The demand for HPC is driven by the increasing need for efficient computing, enhanced scalability, reliable storage, the emerging need for accurate high-speed data processing, and adoption of HPC in the cloud.¹

The European Commission defines High Performance Computing, also known as supercomputing, as computing systems with extremely high computational power that can solve hugely complex and demanding problems. The EU considers HPC to be core to major advances and innovation and is a strategic resource for Europe’s future.²

In 2015, President Obama issued Executive Order 13702 issued by “Creating a National Strategic Computing” Initiative, in which the term “high-performance computing” refers to systems that, through a combination of processing capability and storage capacity, can solve computational problems that are beyond the capability of small- to medium-scale systems. The USA considers HPC to be key to economic competitiveness and scientific discovery.³

The EU, USA, and other global powers, as shown in Figure 1, see HPC as crucial to solving problems such as climate change, renewable energy, evolution, the origin of the universe, prevention and treatment of epidemics and diseases, among many others.

The World’s Top Supercomputers

Computational performance of the most powerful supercomputers, in teraFLOPS*²

*FLOPS= floating point operations per second, i.e. the number of basic mathematical operations a computer can perform in a second. Tera= one trillion

Figure 1: The World’s Top Supercomputers

1 https://www.researchandmarkets.com/reports/5230462/global-high-performance-computing-hpc-market-by
Additionally, The International Data Corporation (IDC) estimated that 1.2 zettabytes (1.2 trillion gigabytes) of new data were created in 2010, up from 0.8 zettabytes the year before. The amount of the newly created data in 2020 was predicted to grow 44x to reach 35 zettabytes (35 trillion gigabytes). Two years ago, we were already at 33 zettabytes, leading IDC to predict that in 2025, 175 zettabytes (175 trillion gigabytes) of new data will be created around the world.4

Some of the primary uses and generators of data are:5

- **Financial services**: monitor real-time stock trends and automate trading.
- **Research labs**: develop novel renewable energy sources, accelerate material design, improve weather forecasts and climate change-related phenomena, and study universe evolution.
- **Healthcare**: design of new drugs, improve accuracy of the diagnostics
- **Entertainment industry**: edit feature films, stream live events, and render special effects or create a whole movie inside a computer.
- **Oil and gas**: identify where to drill for new wells and increase production from existing wells.
- **Machine Learning and Artificial Intelligence**: improve cancer and drug screening methods, detect credit card fraud, offer self-assisted technical support, and teach self-driving vehicles.
- **Industry**: improve products, reduce the time and cost to develop.
- **Big Data**: As our ability to gather information increases, high-performance computing systems are crucial in making sense of the data.

**Goal**

To present an overview of what to consider to ensure the proper selection, design, and implementation of an HPC solution to satisfy business, security, and compliance objectives.

**Audience**

This document is aimed at customers, IT executives, legal and compliance management, and IT operations and technical managers involved in selecting, designing, and implementing an HPC solution.

---


5 [https://www.acecloudhosting.com/blog/high-performance-computing/](https://www.acecloudhosting.com/blog/high-performance-computing/)
Cloud Architecture

Workload Manager/Scheduler

A workload manager/scheduler is a program that aims to provide the system administrator with control over how processes can have access to resources such as computation, storage, I/O, and so on. It can initiate and manage jobs submitted to it based on instructions (or job control language), monitor performance, and notify job completion.

Virtual Compute Cluster

A compute cluster typically refers to a group of physical machines, whereas a virtual machine cluster works by protecting the physical machine from hardware and software failures. When a physical node fails, the virtual machine can access another node, with little or no time lag. In a sense, it provides a dynamic backup.

Storage

HPC storage systems allow computers to operate continuously while data is written to, or read from, the cluster. The requirement for continuous operation and manipulation of a massive amount of data requires a different approach to HPC storage.

Types of Workloads

Loosely coupled vs. tightly coupled, data-intensive vs. less data-intensive

Tightly Coupled

Workloads can be tightly-coupled or loosely coupled. Tight coupling refers to an arrangement where a group of computations is highly dependent on one another. This has the idea of binding resources to specific purposes and functions. Tightly-coupled components may bind each resource to a particular use case, a specific interface, or a specific front end. A tightly coupled system is purpose-built, and every custom deviation from the standard comes with its resources and integrations. Tight coupling brings clear losses to software extensibility and scalability.

Tightly coupled workloads consist of parallel processes that are dependent on each other to carry out the calculation. All processes of a tightly coupled simulation iterate together and require communication with one another. An iteration is defined as one step of the overall simulation. The failure of one node usually leads to the failure of the entire calculation. To mitigate the risk of complete failure, application-level checkpointing can occur during a computation to allow for a simulation’s restart from a known state. Examples of tightly coupled HPC workloads include computational fluid dynamics, weather prediction, and reservoir simulation.
Loosely Coupled

Loose coupling is achieved by using a design that promotes single-responsibility and separation of concerns. A loosely-coupled computation can be consumed and tested independently of other (concrete) computations. In such a system, the components are detached from each other. Each resource could have multiple front ends or applications. All systems can work independently as part of a larger group of systems. Loose coupling benefits from extensibility and scalability but may suffer from increased overall complexity.

Visualization

Visualization of scientific data explains the phenomenon or data being studied. It is about the graphical representation of data to better communicate the voluminous amount of data by exploiting parallel algorithms to process the data quantitatively and qualitatively to enhance and interpret it. The representation is a mapping between the original data and graphic elements (such as lines, colors, or points in a chart).

Intelligent Computing

Intelligent computing research aims at bringing intelligence, reasoning, perception, information gathering, and analysis to computer systems. It has also been referred to as artificial intelligence, knowledge-based systems, expert systems, and so on, depending on the dominant underlying approaches. In its popular usage today, it often refers to computational systems that are machine learning (ML) centric; these systems derive their processing capability by crunching huge data sets using different approaches (including inductive, neural networks, deep learning, and signal processing algorithms). The huge datasets used for training and testing these intelligent systems would consume immense processing power to ‘batch learn’ or ‘learn incrementally.’

Software

A typical user of an HPC cloud service expects that the software running the application is available, either because (a) he has the license (e.g., open source software) or has written the software himself, or (b) a commercial license is available as SaaS to run on the HPC cloud.

Option (a) may require the user to test the software to run properly on the HPC cloud service and yield the expected results obtained previously.

Option (b) may require assurance that the commercial software available under SaaS is the same version previously used, as an earlier or later version may require familiarization or checking that the results produced can be compared with earlier obtained results. SaaS pricing may be an issue, as the price for using multiple copies may be heftier than expected.
Performance

Let’s extract the best performance of the underlying platform to effectively and efficiently use the facility.

Many elements combine to determine the performance efficiency of workloads. The highly parallel nature of HPC workloads requires efficient coding techniques to effectively utilize the underlying capabilities of the hardware. Specialized hardware like GPUs may also be used to accelerate workloads, and the tuning of the workload I/O patterns are critical elements of extracting the maximum performance while managing the energy requirements to execute the workload.

Due to the high-performance requirements of workloads, ‘close to metal’ operations are often demanded, stretching the processor’s core physical computing resource to its utmost capabilities. The apprehension lies in that running on a virtualized hypervisor may cause performance to suffer.

Running Benchmarks

This aspect connects software and hardware to performance. Check the availability of appropriate benchmark tests from the HPC communication service providers to understand potential bottlenecks, software incompatibilities and possible unseen issues that can delay and increase costs. If not, consider installing your software and running several basic tests. The results provide indicators of what is being tested and understanding of the application and associated workflow. These actions will help to eliminate faulty assumptions.

Productivity

The value of an HPC system to a user includes many factors, such as execution time on a particular problem, software development time, direct hardware costs, and indirect administrative and maintenance costs. The topic of productivity in HPC involves defining several characteristic HPC workflows that are useful for understanding how users exploit HPC systems and discusses the role of activity and purpose benchmarks in establishing an empirical basis for HPC productivity.6

Security and Risk Management

This section addresses the most common risks when selecting HPC Cloud as the solution to objectives.

The concerns for HPC are further complicated by the complex and ever-evolving threat landscape. As we increasingly see cases of pure HPC bare metal infrastructure interacting with the cloud such as I/O interfaces and processes, it brings more opportunities for malicious attacks. While this should be

---

considered and integrated into security policies and guidelines, performance faces the peril of being compromised as precious resources are carved out for security protocols and processes.

The crossing of cloud and HPC environments often leads us to questions of how security in an HPC cloud environment can be implemented, enforced and ensured without the need to compromise performance.

**Architecture**

Due to the nature of high-performance requirements, high-speed interconnect is an element encompassed by HPC for fast communication.

**Access**

How do we ensure a secure channel from existing HPC centers into the cloud?

**Runtime**

The dependencies of workload may change or evolve over time; so it is also important to isolate the runtime environment of the workload by new tech, such as container technology (Singularity, runc). In the case of container technologies, it is necessary to clearly define how such containers are built, stored, and used, to prevent resource abuse and limit the potential attack surface.

**Ecosystem**

When migrating the workload to the cloud, it’s hard to know infrastructure details, such as Common Vulnerabilities and Exposures (CVE) systems. Ecosystem or open-source (with vendor enhancement) is one option for the end-user.

**Storage**

Storage is also a factor that influences high-speed communications requirements. HPC applications used by researchers frequently handle a colossal volume of data that are pulled from consistent storage. For lower-latency access, data is often divided between multiple storage instances, putting further strain on I/O environments.

HPC involves large datasets going in and out of compute capacity, in this case, the cloud. There is a need to stage a huge amount of data into the cloud to be crunched and at the same time ensure to provide a mechanism to manage the cost for both storage and data leaving the cloud.

**Backup**

HPC backup and recovery requires high performance that can handle vast amounts of data with ease. For HPC environments that have petabytes of data to store, tape can still be the best option for long term archival and meeting Recovery Point Objective (RPOs). It is also very effective regarding retention requirements and media preservation.
Disaster Recovery (DR)

In a non-cloud setting, organizations invest in DR in the form of a cluster located in an offsite datacenter awaiting activation in the event of a disaster. This option is expensive, with the attendant problem of keeping the back-up insync with the production. With the availability of public clouds, such hefty upfront investments can be reduced. An organization can also leverage public cloud storage as a tier of their storage infrastructure. It is important to be able to move one’s data easily to the cloud in the event of disaster, or to ensure the right data is being periodically fed out to the public cloud. When an emergency occurs, such data is already in the public cloud. Mirroring an existing HPC infrastructure in the public cloud provides multiples of one’s HPC system on the public cloud for different requirements, including DR, busting capacity during planned downtime, testing, development or expansion.

Talent

The shortage of HPC skills is a critical cause for the slow adoption of HPC by industry. Today’s graduates lack basic HPC competency, partly because supercomputing is sometimes mistakenly seen as an old technology. The HPC slowdown as regular uncertainty in government funding, alongside the explosive growth of Internet companies, create the image of HPC as maturing and even dying. With a high proportion of today’s graying HPC professionals within retirement age and not enough HPC-trained graduates being produced to replace them, the shortage of HPC skills has become dire. The much-demanded HPC skills are mathematics, engineering, or the physical sciences; these folks become computer scientists, operating systems experts, parallel programmers, programmers for heterogeneous (CPU-accelerator) systems, systems administrators, and tools developers.

Governance

When a customer is looking to move an HPC workload to the cloud, in the case of sensitive data like genomics, the first step would be infrastructure compliance (i.e., SOC3) or any other standards that will be a baseline to comply with.

Secondly, an HPC Cloud has a well-architected framework for security, tools for analysis for customers requirements to fine tune to.

Thirdly, an HPC Cloud offers cloud services that are tuned to a specific industry or nation. (For example, countries like Singapore have established policies regarding the use or processing of genomics data.)

Lastly, the host country and the HPC Cloud are signatory to global standards, that is, Global Alliance for Genome and Health (GA4GH).7

---

7 [https://ga4gh-duri.github.io/](https://ga4gh-duri.github.io/)
Cost/Benefits

The cost/benefits from moving resource-intensive application workloads from on-premise HPC environments to public HPC clouds.

Barrier to entry to HPC Usage

Without a hefty upfront investment as a cost barrier to entry, there will likely be more HPC users encouraged by the relatively lower cost involved than before. This is because either a physical HPC capability or a cloud HPC capability is most effectively governed by the cost to run a specific workload.

The aspects of HPC that many newcomers are challenged by are the highly parallel nature of the systems and the most effective way to express their workload within those environments.

Alternative

HPC resources are scarce and very expensive; they are often very well used. HPC clouds that are suitable for running lower priority workloads can address the pent-up demands.

Elasticity

As demand for HPC resources varies with HPC users’ computation loads, the HPC clouds can run the workloads to ride the peaks and troughs.

CAPEX to OPEX

On-premise HPC resources are very high-expense investments (millions of dollars) that require a cycle of justifications as legacy systems are retired at the end of their life. The shift to paying on a per-use basis reflects reality and is in step with HPC resource usage. This means HPC users can use HPC cloud services from their research budget (without asking for the moon).

Agility

The access of HPC cloud services does not face challenges associated with procurement (associated with a resource and time-consuming tender evaluation), site preparation, machine installation, and so on.

Time to Market

Easy access to HPC clouds and the relatively lower cost of usage will encourage certain industry sectors to develop new products and/or services than previously possible.
Multi-Cloud

The on-premise HPC cluster may not work because of bugs/issues; with multi-cloud, the end-user can switch to different cloud vendors to avoid business outages.

An ecosystem is important in such scenarios; the workload may not be able to shift to other cloud vendors if the interfaces are different.

Conclusion

Current access to HPC resources has been the exclusive privilege of a small group of IT users. The reasons include the hefty investment in capital assets and operations, specialized knowledge and expertise requirements, and the ability to harness such enormous computational resources properly to solve appropriate problems. The emergence of HPC cloud services by commercial CSPs has democratized this access, allowing a much larger group of users to harness such vast computational resources, without having to make a hefty upfront investment and employing specialized operation staff. However, other requirements remain; the HPC user has to gain competence in understanding the problems and the type of workloads that can benefit from HPC. This report aims to provide a starting point for such users.

The next report will focus on the security considerations of using an HPC cloud service, since almost all current HPC resources are on-premises, with the usual security measures in place. With commercial HPC cloud services maturing, lessons and best practices will become available leading to an increase in such usage. This next report will be useful to both (a) new HPC users; and (b) experienced on-premises HPC users.
References

UberCloud Customer Onboarding: https://info.theubercloud.com/ubercloud-customer-onboarding


Azure Subscription Requirements for The UberCloud Platform, accessed April, 05, 2021: https://info.theubercloud.com/azure-subscription-requirements

Glossary

**Artificial Intelligence (AI)** - the application of “advanced analysis and logic-based techniques, including machine learning, to interpret events, support and automate decisions, and take actions.

**Cluster** - Clusters are groups of servers that are managed together and participate in workload management. A cluster can contain nodes or individual application servers. A node is usually a physical computer system with a distinct host IP address that is running one or more application servers.
[https://www.ibm.com/docs/en/was-nd/8.5.5?topic=servers-introduction-clusters](https://www.ibm.com/docs/en/was-nd/8.5.5?topic=servers-introduction-clusters)

**Cluster Interconnect (CI)** - The cluster interconnect is the physical configuration of devices that are used to transfer cluster-private communications and data service communications between cluster nodes.

**Daemon** - is a Unix/Linux program that executes in the background, ready to operate when required. The daemon functions like an extension to the operating system. A daemon is usually an unattended process that is initiated at startup. Typical daemons are print spoolers and email handlers or schedulers that start up another process at a designated time.
[https://www.pcmag.com/encyclopedia/term/daemon](https://www.pcmag.com/encyclopedia/term/daemon)

**HPC Workloads** - require the performance of significant computational work and, typically, demand a large amount of processor (CPU) and storage resources to accomplish demanding computational tasks within a limited timeframe, even in real-time.
[https://searchdatacenter.techtarget.com/definition/workload](https://searchdatacenter.techtarget.com/definition/workload)

**Machine Learning (ML)** - is an application of artificial intelligence (AI) that allows systems to learn and improve from experience without being explicitly programmed. ML focuses on developing computer programs that can access data and use it to learn for themselves.
[https://www.expert.ai/blog/machine-learning-definition/](https://www.expert.ai/blog/machine-learning-definition/)

**SOC 3** - A Service Organization Control 3 (SOC 3) report outlines information related to a service organization's internal controls for security, availability, processing integrity, confidentiality or privacy. These five areas are the focuses of the AICPA Trust Services Principles and Criteria.

A SOC 3 reports on the same information as a SOC 2 report. The main difference between the two is that a SOC 3 is intended for a general audience. These reports are shorter and do not include the same details as a SOC 2 report, which is distributed to an informed audience of stakeholders. Due to their more general nature, SOC 3 reports can be shared openly and posted on a company's website with a seal indicating their compliance.
[https://searchcloudsecurity.techtarget.com/definition/Soc-3-Service-Organization-Control-3](https://searchcloudsecurity.techtarget.com/definition/Soc-3-Service-Organization-Control-3)
Zettabyte - is 1021 or 1,000,000,000,000,000,000,000 bytes. One zettabyte (abbreviated “ZB”) is equal to 1,000 exabytes and precedes the yottabyte unit of measurement. Zettabytes are slightly smaller than zebibytes, which contain 1,180,591,620,717,411,303,424 (270) bytes. A single zettabyte contains one sextillion bytes, or one billion terabytes.  
https://techterms.com/definition/zettabyte

Acronyms

AI - Artificial Intelligence
CI - Cluster Interconnect
EU - European Union
GA4GH - Global Alliance for Genome and Health
HPC - High-Performance Computing
IDC - International Data Corporation
I/O - Input / Output
LAN - Local Area Network
MC - Memory Channel
ML - Machine Language
NIST - National Institute of Standards and Technology
SOC3 - Service Organization Control 3
USA - United States of America
Annex 1 - Top 8 Supercomputers at 11/2021

<table>
<thead>
<tr>
<th>#</th>
<th>Category</th>
<th>Cores</th>
<th>Rmax (TFlop/s)</th>
<th>Rpeak (TFlop/s)</th>
<th>Power (kW)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Supercomputer Fugaku - Supercomputer Fugaku, A64FX 48C 2.2GHz, Tofu interconnect D, <strong>Fujitsu</strong> RIKEN Center for Computational Science <strong>Japan</strong></td>
<td>7,630,848</td>
<td>442,010</td>
<td>537,212</td>
<td>29,899</td>
</tr>
<tr>
<td>2</td>
<td>Summit - IBM Power System AC922, IBM POWER9 22C 3.07GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband, <strong>IBM</strong> DOE/SC/Oak Ridge National Laboratory <strong>United States</strong></td>
<td>2,414,592</td>
<td>148,600</td>
<td>200,794</td>
<td>10,096</td>
</tr>
<tr>
<td>3</td>
<td>Sierra - IBM Power System AC922, IBM POWER9 22C 3.1GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband, <strong>IBM / NVIDIA / Mellanox</strong> DOE/NNSA/LLNL <strong>United States</strong></td>
<td>1,572,480</td>
<td>94,640</td>
<td>125,712</td>
<td>7,438</td>
</tr>
<tr>
<td>4</td>
<td>Sunway TaihuLight - Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway, <strong>NRCPC</strong> National Supercomputing Center in Wuxi <strong>China</strong></td>
<td>10,649,60</td>
<td>93,014</td>
<td>125,435</td>
<td>15,371</td>
</tr>
<tr>
<td>5</td>
<td>Perlmutter - HPE Cray EX235n, AMD EPYC 7763 64C 2.45GHz, NVIDIA A100 SXM4 40 GB, Slingshot-10, <strong>HPE</strong> DOE/SC/LBNL/NERSC <strong>United States</strong></td>
<td>761,856</td>
<td>70,870.0</td>
<td>93,750.0</td>
<td>2,589</td>
</tr>
</tbody>
</table>

---

8 [https://www.top500.org/lists/top500/2021/11/](https://www.top500.org/lists/top500/2021/11/)
<table>
<thead>
<tr>
<th>Rank</th>
<th>System Name</th>
<th>Description</th>
<th>MIPS</th>
<th>FLOPS</th>
<th>Memory</th>
<th>Bandwidth</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>Selene - NVIDIA DGX A100, AMD EPYC 7742 64C 2.25GHz, NVIDIA A100, Mellanox HDR Infiniband, <strong>Nvidia</strong></td>
<td>United States</td>
<td>555,520</td>
<td>63,460</td>
<td>79,215</td>
<td>2,646</td>
</tr>
<tr>
<td>7</td>
<td>Tianhe-2A - TH-IVB-FEP Cluster, Intel Xeon E5-2692v2 12C 2.2GHz, TH Express-2, Matrix-2000, <strong>NUDT</strong> National SuperComputer Center in Guangzhou <strong>China</strong></td>
<td>China</td>
<td>4,981,760</td>
<td>61,444.5</td>
<td>100,678.7</td>
<td>18,482</td>
</tr>
<tr>
<td>8</td>
<td>JUWELS Booster Module - Bull Sequana XH2000 ,AMD EPYC 7402 24C 2.8GHz, NVIDIA A100, Mellanox HDR InfiniBand/ParTec ParaStation ClusterSuite, <strong>Atos</strong> Forschungszentrum Juelich (FZJ) <strong>Germany</strong></td>
<td>Germany</td>
<td>449,280</td>
<td>44,120.0</td>
<td>70,980.0</td>
<td>1,764</td>
</tr>
</tbody>
</table>