Newest Trends in High Performance Data Analytics Security Concerns in Cloud and High-Performance Computing

Pratham Shrivastava
p.shrivastava@stud.uni-goettingen.de

Trevor Khwam Tabougua
trevor-khwam.tabougua@gwdg.de

March 31, 2024

Abstract

This paper provides a comprehensive examination of security concerns in Cloud Computing and High-Performance Computing (HPC) environments, emphasizing the key characteristics, real-world incidents, and effective security measures associated with each. The introduction talks about the fundamental features of cloud computing, categorizing services into SaaS, IaaS, and PaaS, while exploring deployment models such as public, private, and hybrid clouds. Subsequently, the focus shifts to HPC, detailing its distinctive attributes and applications across various fields.

Real-world examples of security breaches are analysed, drawing attention to incidents like the Capital One data breach in 2019 and the NERSC security incident of 2015, offering insights into the diverse security challenges faced by both cloud and HPC infrastructures. Major security concerns specific to cloud computing, such as virtualization, risks and issues related to data storage in public clouds and multitenancy, are then explored.

The paper delves into security methods employed in cloud computing, with a detailed examination of data encryption and cryptography techniques. In the context of HPC, security technologies such as local-level firewalls, antivirus software, and honeypots are discussed.

Furthermore, the paper provides practical security recommendations for both HPC and cloud computing, offering strategies to enhance resilience against evolving cyber threats. Lastly, the importance of benchmarking frameworks in evaluating and comparing the effectiveness of security measures is emphasized. Common frameworks like the Common Vulnerability Scoring System (CVSS) and the NIST Cybersecurity Framework are highlighted. This integrated perspective aims to contribute to a more holistic approach to security in the rapidly evolving landscapes of both cloud computing and HPC.

1 Introduction

1.1 Cloud Computing

Cloud Computing is a technology model that enables access to a shared pool of computing resources over the internet. Instead of owning and maintaining physical servers or computing infrastructure, users can rent or lease computing power, storage, and other resources on a pay-as-you-go basis from a cloud service provider. These resources are delivered over the internet, allowing users to access and manage them remotely[16].

1.1.1 Key characteristics of cloud computing include:

1. On-Demand Self-Service: Users can provision and manage computing resources as needed without requiring human intervention from the service provider.
2. Broad Network Access: Cloud services are accessible over the internet from a variety of devices, such as laptops, tablets, and smartphones.
3. Resource Pooling: Computing resources are shared among multiple users, and the provider dynamically allocates and reallocates resources based on demand.
4. Rapid Elasticity: Resources can be quickly scaled up or down to meet changing workloads, providing flexibility and cost efficiency.
5. Measured Service: Usage of cloud resources is monitored, controlled, and billed based on the actual consumption, allowing for transparent and cost-effective pricing.

1.1.2 Cloud computing services are typically categorized into three main models:

1. Infrastructure as a Service (IaaS): Provides virtualized computing resources over the internet. Users can rent virtual machines, storage, and networking components[16].
2. Platform as a Service (PaaS): Offers a platform that includes both the underlying infrastructure and tools to develop, deploy, and manage applications without worrying about the complexities of the underlying infrastructure[16].
3. Software as a Service (SaaS): Delivers software applications over the internet on a subscription basis. Users can access and use the software without managing or maintaining the underlying infrastructure.

1.1.3 Cloud computing deployment models include:

1. Public Cloud: Resources are owned and operated by a third-party cloud service provider and are made available to the public[14].
2. Private Cloud: Resources are used exclusively by a single organization. It may be hosted on-premises or by a third-party provider[14].
3. Hybrid Cloud: Combines elements of both public and private clouds, allowing data and applications to be shared between them[14].

Cloud computing provides numerous benefits, including cost savings, scalability, flexibility, and the ability to focus on core business activities without the need to manage and maintain complex IT infrastructure. Popular cloud service providers include Amazon Web Services (AWS), Microsoft Azure, Google Cloud Platform (GCP), and many others[6]. There are two stages when data has the most security concern, first is called as data at rest and secondly data in transit as shown in Fig 1. Data at rest refers to the situation when data is stored in the cloud and Data in transit refers to the phase when data is moving in or out of the cloud. While moving in or out of the cloud data can be encrypted or not. While data in transit is more prone to attacks as it is moving from 1 location to another.

1.2 High-Performance Computing (HPC):

HPC refers to the use of advanced computing technologies to solve complex problems and perform large-scale computations at significantly higher speeds and capacities than traditional computing resources. The primary goal of HPC is to deliver superior performance for demanding and resource-intensive applications[15].

Data at Rest
Data in Transit

Figure 1: Data at rest and in transit

1.2.1 Key characteristics of High-Performance Computing include:

1. Parallel Processing: HPC systems leverage parallel processing to break down complex problems into smaller tasks that can be solved concurrently. This is achieved by distributing the workload across multiple processors or nodes[4].
2. High-Speed Interconnects: HPC systems often use high-speed interconnects to enable efficient communication between processors. This is crucial for maintaining the overall performance of parallelized applications.
3. Large-Scale Storage: HPC environments require substantial storage capacity to handle vast amounts of data generated by simulations, analyses, and other scientific or engineering applications.
4. Specialized Hardware: HPC systems may use specialized hardware components, such as accelerators (e.g., GPUs or TPUs), to enhance computational performance for specific types of calculations[4].
5. Optimized Software: Applications in HPC are typically optimized for parallel execution and designed to take advantage of the specific architecture of the HPC system. This often involves the use of parallel programming models and libraries.
6. Scientific and Research Applications: HPC is widely used in scientific and research domains for simulations, modelling, data analysis, and other computationally intensive tasks. Common applications include weather forecasting, molecular modelling, financial modelling, and more.
7. High Throughput: HPC systems are designed to deliver high throughput and performance, allowing users to complete complex computations in a reasonable amount of time[4].

1.2.2 HPC is employed in various fields, including:

1. Scientific Research: Simulations and analyses in physics, chemistry, biology, and other scientific disciplines.
2. Engineering: Computational fluid dynamics, structural analysis, and optimization in engineering design.
3. Finance: Risk analysis, modelling, and algorithmic trading.
4. Weather Forecasting: Numerical weather prediction models.
5. Oil and Gas Exploration: Seismic data analysis and reservoir simulations.
6. Genomics and Bioinformatics: DNA sequencing analysis and protein folding simulations.
7. Aerospace: Aerodynamic simulations, structural analysis, and optimization.

HPC systems can be designed and deployed on-premises, in the cloud, or through a combination of both (hybrid environments). Supercomputers, which are among the most powerful

HPC systems, are often used for the most challenging computational tasks. Overall, High-Performance Computing plays a crucial role in advancing scientific research, engineering, and various other fields by providing the computational power needed to tackle complex problems and simulations.

2 Case Studies and Practical Examples:

Cloud Computing stores and manages the data of high confidentiality and hence providing security to data becomes one of the major responsibilities of the cloud service provider. If proper security majors are not kept in place, then it can lead to heavy loss for the companies who trusted the cloud service provider and kept their data using their cloud. Similarly using resource for High Performance Computing involves data which is highly confidential sometimes related to the security of a country or sometimes involves some research data that might be made public later, making it a prominent responsibility for HPC resource providers to maintain the security of that data. Security is a critical concern in both Cloud Computing and High-Performance Computing (HPC) environments. Here are some real-world examples of security in cloud computing:

2.1 Capital One Data Breach (2019):

Incident: A former employee of Amazon Web Services (AWS) exploited a misconfigured firewall to gain unauthorized access to Capital One's customer data stored in the AWS cloud.
Lesson Learned: Emphasizes the importance of proper configuration and monitoring of cloud security settings, especially when using Infrastructure as a Service (IaaS) platforms[13].

2.2 AWS S3 Bucket Misconfigurations (Various Incidents):

Incident: Numerous data breaches have occurred due to misconfigured Amazon S3 buckets, leading to the exposure of sensitive data. Organizations like Verizon, Dow Jones, and Accenture have experienced such incidents.
Lesson Learned: Stresses the significance of regularly auditing and configuring cloud storage settings correctly to prevent unauthorized access[8].

2.3 Code Spaces Shutdown (2014):

Incident: Code Spaces, a source code repository and project management service, suffered a massive DDoS attack followed by unauthorized access to their AWS control panel. The attackers deleted data, machine configurations, and backups, forcing Code Spaces to shut down.
Lesson Learned: Demonstrates the need for multi-layered security measures, including robust access controls and regular data backups.

2.4 Equifac Data Breach (2017):

Incident: Equifax failed to patch a known vulnerability in their web application software. Hackers exploited this vulnerability to access and steal sensitive data, including Social Security numbers, of over 147 million Americans.
Lesson Learned: Organizations need to prioritize timely patching of vulnerabilities in their

cloud-based software and applications. Implementing automated patching procedures and vulnerability scans can help ensure systems are kept up-to-date and secure[5].

2.5 Dropbox Breach (2014):

Incident: Attackers compromised a third-party app with access to Dropbox user accounts. This compromise exposed user data, including names and potentially some file contents.
Lesson Learned: Organizations using cloud services need to be aware of the security posture of third-party applications and integrations. Security assessments and proper vetting of third-party apps can help mitigate risks. The cloud security model is shared. While the cloud provider secures the underlying infrastructure, users are responsible for securing their applications and data. Understanding this model and implementing appropriate security measures is crucial[12].

High-Performance Computing (HPC) environments face unique security challenges due to their scale, complexity, and the sensitive nature of the data they process. Here are a couple of real-world case studies highlighting security incidents and lessons learned in HPC:

1. NERSC/ESnet Security Incident (2015):

Incident: The National Energy Research Scientific Computing Center (NERSC) and the Energy Sciences Network (ESnet), two major U.S. Department of Energy facilities, experienced a security breach. The incident involved unauthorized access to NERSC systems and resulted in a temporary shutdown of the center's computational resources.
Lesson Learned: This incident highlighted the importance of continuous monitoring, early detection of security threats, and the need for coordinated incident response in large-scale HPC environments.

2. Oak Ridge National Laboratory Titan Supercomputer Compromise (2018):

Incident: The Titan supercomputer at Oak Ridge National Laboratory, one of the most powerful supercomputers at the time, was temporarily taken offline due to a security incident. The compromise was related to an unauthorized attempt to access the system.
Lesson Learned: The incident underscored the necessity for robust access controls, regular security audits, and prompt response to potential security breaches in HPC environments[7].

3. ARCHER Supercomputer Compromise (2015):

Incident: The ARCHER supercomputer, located in the UK and used for scientific research, experienced a security incident that led to a temporary shutdown. The compromise was related to a security vulnerability in the system's login nodes.
Lesson Learned: This incident highlighted the importance of promptly patching and updating software in HPC environments to address vulnerabilities and minimize the risk of unauthorized access[11].

4. Barcelona Supercomputing Center (BSC) Malware Attack (2012):

Incident: The MareNostrum supercomputer at the Barcelona Supercomputing Center (BSC) was compromised by malware designed to steal scientific research data. The attackers gained access through a vulnerability in the job scheduler software and used it to deploy the malware across the HPC cluster.
Lesson Learned: HPC systems often prioritize performance over security. This incident highlights the importance of hardening HPC systems by following secure coding practices, keeping software up-to-date, and implementing network segmentation to isolate workloads. Job schedulers are critical components in HPC environments. Focusing on the security of these systems, including vulnerability management and access controls, is crucial[19].

5. Jülich Supercomputing Centre Breach (2016):

Incident: Hackers infiltrated the Jülich Supercomputing Centre in Germany, potentially gaining access to sensitive research data related to climate modeling and nuclear fusion. The exact attack vector remains unclear, but suspicions point towards exploiting a vulnerability in the user authentication system.
Lesson Learned: Implementing strong authentication measures like MFA for user access to HPC resources can significantly reduce the risk of unauthorized access, even if attackers obtain usernames and passwords. Monitoring user activity within the HPC environment can help detect suspicious behavior and potential breaches early on[1].

Revieing and witnessing above-mentioned real-life security threats that happened in past years related to cloud computing and High Performance Computing it is clear that security is a major concern and utmost responsibility to ensure proper security of data in cloud and HPC centres. In next sections I discuss some major security concerns and methods to ensure security in both cloud and HPC centres.

3 Security Concerns in Cloud Computing:

1. Virtualization: Virtualization in security in cloud computing refers to the use of virtualization technologies to enhance and strengthen the security posture of cloud-based systems. Cloud computing relies heavily on virtualization, which allows multiple virtual instances or machines to run on a single physical server. This technology provides flexibility, scalability, and resource efficiency, but it also introduces unique security challenges. The hypervisor, also known as the virtual machine monitor (VMM), is a critical component in virtualized environments. It manages and allocates resources to VMs. Ensuring the security of the hypervisor is crucial to prevent attacks that could compromise the entire virtualized infrastructure. Addressing the previously highlighted concerns involves implementing more effective planning for virtualization utilization. It is imperative to meticulously manage resources and ensure thorough authentication of data before reallocating these resources.

2. Storing Data in Public Cloud: In a public cloud environment, multiple tenants share the same infrastructure. Ensuring robust access controls and data segregation mechanisms is crucial to prevent unauthorized access to sensitive information. Misconfigurations or inadequate access controls can lead to data breaches. Encrypting data both in transit and at rest is essential to protect it from unauthorized access. While most reputable public cloud providers offer encryption features, organizations must configure and manage encryption keys appropriately to maintain control over their data. For some confidential data such as related to security

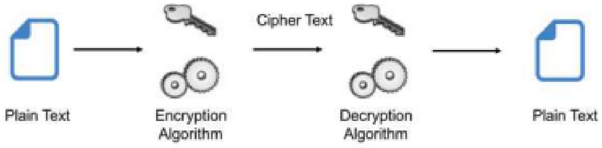

Figure 2: Basic Cryptographic Technique

concerns or some future technology it is highly advised to store in companies' private organizational cloud rather than storing at a public cloud.

3. Multitenancy: Shared access or multitenancy is recognized as a significant risk to data in cloud computing. The utilization of shared computing resources such as CPU, storage, and memory by multiple users poses a threat not only to individual users but to multiple users collectively. In such scenarios, there is a constant risk of inadvertent leakage of private data to other users. Exploits related to multitenancy can be highly precarious, as a single flaw in the system might grant unauthorized access to another user or hacker, potentially compromising all other data.

4. Insider Threat: In the world of cloud security, insider threats pose a significant risk. Unlike external attackers who have to breach firewalls and complex security systems, insiders already have authorized access to cloud resources. This makes them especially dangerous because they can bypass many security measures. Typr of Insider Threat include Malicious Insider, Negligent Insider and Careless Insider

5. Lack of Visibility: Unlike on-premises environments, organizations may have limited visibility into how their data is stored and secured in the cloud. This can make it difficult to detect and respond to security incidents.

6. Denial-of-Service (DoS) Attacks: Attackers can target cloud-based applications or services with DoS attacks, overwhelming them with traffic and making them unavailable to legitimate users.

Addressing such concerns involves implementing effective user authentication mechanisms. Proper authentication procedures are crucial to prevent unauthorized access to data. Various authentication techniques are employed to mitigate multitenancy issues in cloud computing, enhancing overall security measures.

4 Methods to Ensure Security in Cloud Computing:

1. Data Encryption: Different Encryption techniques can be used for two stages of data discussed above that is data at rest and data in transit. Key for when data is in transit can be short lived while the key for data when at rest might retain for longer time[1]. Various cryptographic techniques are employed to encrypt data in contemporary settings. Cryptography plays a crucial role in enhancing data protection by ensuring content integrity, authentication, and availability. In its fundamental application, plaintext undergoes encryption into ciphertext using an encryption key. Subsequently, the generated ciphertext is decrypted using a decryption key, as depicted in Fig 2.

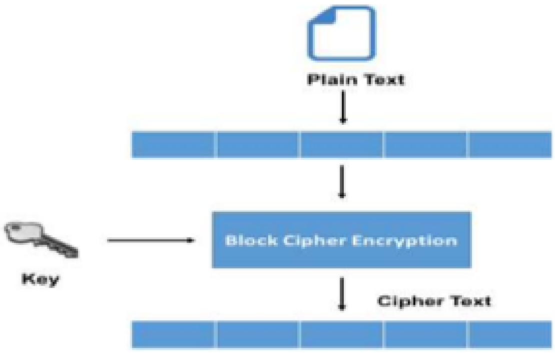

Figure 3: Block Cipher Mechanism

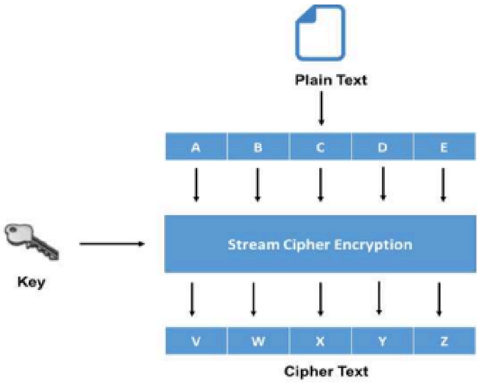

Figure 4: Stream Cipher Mechanism

Different cryptography techniques are used, such as:

A. Block cipher is a cryptographic algorithm that operates on fixed-size blocks of data, typically blocks of a specific number of bits. In contrast to stream ciphers that encrypt data one bit or byte at a time, block ciphers process data in fixed-size blocks. The block cipher algorithm takes a fixed-length block of plaintext bits as input and transforms it into a block of ciphertext bits of the same length[1]. As shown in Figure 3, in block cipher mechanism, plain text is divided into equal number of blocks and then encrypted using an encryption key. And decryption occurs in the exact same way.

B. Stream Cipher: A stream cipher is a type of cryptographic algorithm that encrypts plaintext one bit or byte at a time. Unlike block ciphers, which process data in fixed-size blocks, stream ciphers encrypt data continuously as a stream of bits. Stream ciphers are often employed in scenarios where a continuous flow of data needs to be encrypted and decrypted efficiently. As depicted in Figure 4, a stream cipher employs an encryption key to encrypt individual bits rather than blocks of text. The output, or ciphertext, is a continuous stream of encrypted bits. To retrieve the original plaintext, the ciphertext is decrypted using the decryption key[1].

C. Hash Function: In this method, a mathematical operation known as a hash function is employed to transform an input text into an alphanumeric string. Typically, the resulting alphanumeric string has a fixed size. This approach ensures the uniqueness of alphanumeric strings, preventing any two different input strings from generating the same alphanumeric out-

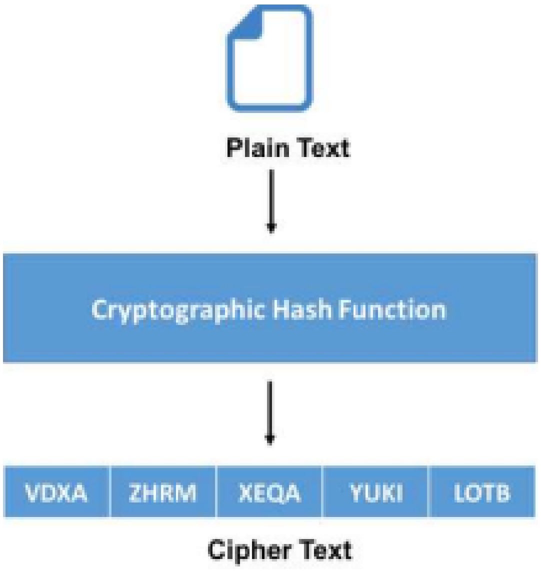

Figure 5: Cryptographic Hash Function Mechanism

put. Even with slight differences in the input strings, the resulting alphanumeric strings can exhibit significant distinctions. Figure 5, depicted below, illustrates the workings of hash function cryptography[1].

All of above-mentioned encryption techniques are widely used in Cryptography, use depends on scenario to scenario. Whichever technique is used, security and integrity of data at both private and public cloud is prioritised.
Some other measures that should be taken to keep security in place includes:

2. Use of Firewall: Use firewalls and network security groups to control incoming and outgoing traffic. Segment your network to minimize the impact of a security breach and isolate sensitive data.
3. Data Backups: Regularly back up your data and ensure that backup processes are functioning correctly. In the event of data loss or a security incident, having reliable backups is crucial for recovery.
4. Security Patching and Updates: Keep all software, including operating systems and applications, up to date with the latest security patches. Regularly update and patch systems to address known vulnerabilities.
5. Access Controls: Implement robust access controls to restrict who can access specific resources. Use principles like the principle of least privilege to ensure users and applications have the minimum access necessary to perform their tasks.

IT security is a multifaceted and intricate challenge. Potential attack vectors range from internal threats posed by malicious employees attempting to sabotage the network to large-scale distributed denial-of-service (DDoS) attacks originating from thousands, or even millions, of hosts. Moreover, Bruce Schneier, a renowned IT security expert, emphasized that security should be viewed as a continuous process rather than a static product. This implies that achieving absolute security for an HPC infrastructure is an ongoing endeavour, acknowledging the ever-evolving nature of security threats.

There will never be a flawless solution to security issues. Dependence on the security offered by a singular product is not advisable. The crucial approach to achieving an adequate level of security is known as defence-in-depth. This strategy involves incorporating multiple layers of security controls throughout the IT system. Even if one or two layers fail—for instance, due to a poorly coded web application enabling an SQL injection attack and unauthorized access to sensitive data—other layers remain in place to potentially thwart the attack. For example, an intrusion detection system might identify anomalies in the typical HTTP request pattern to a vulnerable URL, leading to the firewall blocking the source IP address.

5 Security technology used in High Performance Computing:

1. Local Level Firewall: Firewalls play a crucial role in enhancing security in High-Performance Computing (HPC) environments. Here are several ways in which firewalls contribute to securing HPC systems. The most basic form of a firewall is a packet filter, which functions at the network and transport layers of the ISO/OSI model. It possesses information regarding the source and destination IP addresses as well as TCP/UDP ports. Advanced firewalls, beyond the capabilities of packet filters, go a step further by maintaining the state of network connections, a process known as stateful packet inspection[2].

2. Antivirus software: Antivirus software plays a critical role in enhancing security in High-Performance Computing (HPC) environments by protecting against malicious software, commonly referred to as malware. Here are several ways antivirus software contributes to HPC security: Malware Detection, Real Time Scanning, Heuristic Analysis, Signature-Based Detection, Email and Web Protection, etc[2].

3. Honeypot: is a deceptive and controlled system or network designed to attract and detect attackers or unauthorized access attempts. The primary purpose of a honeypot is to study and understand the tactics, techniques, and procedures (TTPs) employed by attackers, as well as to gather information about potential security threats.

4. Data Loss Prevention (DLP): is a crucial aspect of security in High-Performance Computing (HPC) environments, where the protection of sensitive data is of paramount importance. DLP measures aim to prevent the unauthorized access, transmission, or leakage of sensitive information[2].

5. Authentication: based on a combination of a username and a static password is a widely used method for validating user identity. Despite its popularity, this approach is not considered the most secure. Several alternative and more secure authentication methods exist, with two-factor authentication (2FA) being a prevalent choice. In the 2FA model, merely possessing a password (referred to as "something you know") is insufficient for user authentication. An additional factor, commonly a possession factor like a private and public key (referred to as "something you have"), is required.

A notable technology gaining increased attention is the one-time password (OTP) method. An OTP is a password valid for a single login session, addressing some of the weaknesses associated with static passwords, such as susceptibility to replay attacks. If an attacker obtains the password, it becomes invalid and cannot be reused. The implementation of OTP typically involves the use of a special device to generate passwords. While there is an associated cost with acquiring such devices, it is generally reasonable and justifiable given the enhanced security afforded by OTP.

6 BENCHMARKING FRAMEWORKS:

Benchmarking frameworks play a crucial role in assessing and comparing the effectiveness of different security measures. These frameworks help organizations evaluate the performance of security solutions, identify weaknesses, and make informed decisions about their cybersecurity strategy. Here are some established benchmarking frameworks used for assessing security measures:

1. Cloud Standards Customer Council (CSCC) Cloud Service Measurement Framework (CSMF):
Purpose: Provides a standardized approach to measure various cloud service attributes.
Use Case: Organizations can compare offerings from different cloud providers based on performance, security, and portability metrics defined by CSMF. This helps in selecting the most suitable cloud service for their specific needs.

2. Common Vulnerability Scoring System (CVSS):
Purpose: CVSS provides a standardized method for assessing and prioritizing vulnerabilities. It assigns a severity score to vulnerabilities based on various metrics such as exploitability, impact, and access complexity[**?**].
Use Case: Useful for comparing and prioritizing vulnerabilities to address the most critical security issues first.

3. Open Web Application Security Project (OWASP) Application Security Verification Standard (ASVS):
Purpose: ASVS provides a framework for designing, building, and testing modern web applications and web services. It includes security controls that focus on authentication, session management, access control, and data protection[18].
Use Case: Ideal for benchmarking the security of web applications against a set of standardized security controls.

4. NIST Cybersecurity Framework:
Purpose: Developed by the National Institute of Standards and Technology (NIST), this framework provides a risk-based approach to managing and improving cybersecurity. It consists of functions such as Identify, Protect, Detect, Respond, and Recover[9].
Use Case: Suitable for assessing an organization's overall cybersecurity posture and identifying areas for improvement.

5. SPECcloud Benchmarks:
Purpose: Measure the performance of cloud infrastructure for various workloads.
Use Case: Organizations can utilize SPECcloud benchmarks to evaluate the suitability of a cloud provider for their specific workload types like web applications, high-performance computing, or big data analytics. This helps assess if the cloud infrastructure can handle their processing demands efficiently.

6. BSIMM (Building Security In Maturity Model):
Purpose: BSIMM is a software security framework that helps organizations assess and improve their software security initiatives. It provides a set of best practices observed across various industries.

Use Case: Useful for benchmarking and improving an organization's software security practices.

7. High-Performance Linpack (HPL):
Purpose: Measures the floating-point computing performance of HPC systems.
Use Case: HPC centers can use HPL to compare the raw processing power of different HPC systems. This benchmark is particularly useful for workloads heavily reliant on floating-point operations, such as scientific simulations involving complex mathematical calculations[3].

8. High-Performance Computing Benchmark (HPCG):
Purpose: Focuses on sparse matrix computations commonly used in scientific simulations and engineering applications.
Use Case: Research institutions and engineering firms can leverage HPCG to evaluate how well an HPC system performs on workloads involving sparse matrices, which are frequently encountered in these domains[10].

9. Security Content Automation Protocol (SCAP):
Purpose: SCAP is a suite of specifications for standardizing the format and nomenclature by which security software communicates information about software vulnerabilities, security policies, and security configurations.
Use Case: Useful for automating security measurement and compliance checking across different systems.

10. High Performance Storage System (HPSS) Benchmark:
Purpose: Measures the performance of parallel file systems used in HPC environments.
Use Case: HPC administrators can utilize HPSS benchmarks to assess the efficiency of their parallel file systems in storing and retrieving large datasets. This helps ensure smooth data management and minimizes bottlenecks related to data access within the HPC environment[17].

When selecting a benchmarking framework, organizations should consider their specific needs, the nature of their systems, and the relevant threat landscape. Combining multiple frameworks may provide a more comprehensive assessment of security measures.

7 Security Recommendations for HPC and Cloud Computing:

Cloud Security:

1. MFA Everywhere: Enforce Multi-Factor Authentication for all cloud resource access. (Strengthens login security)
2. Least Privilege Access: Grant users access only to the resources they need (Minimizes damage from compromised credentials).
3. Encrypt Sensitive Data: Encrypt data at rest and in transit (Protects data confidentiality even if accessed by attackers).
4. Regular Security Audits: Conduct periodic audits to identify and address vulnerabilities (Proactive approach to security).
5. Cloud Security Certifications: Choose providers with security certifications like SOC 2 (Provides assurance of robust security practices).

HPC Security:

1. Harden HPC Systems: Focus on secure coding practices and software updates (Reduces exploitable vulnerabilities).
2. Secure Job Schedulers: Implement strong access controls for job schedulers (Protects job execution from unauthorized manipulation).
3. Network Segmentation: Isolate workloads within the HPC environment (Minimizes impact of a security breach).
4. Monitor User Activity: Track user activity within HPC for suspicious behavior (Helps detect potential insider threats).
5. Background Checks and Access Control: Conduct thorough background checks and enforce least privilege for HPC user access (Reduces risk of malicious insiders).

8 Conclusion:

In conclusion, this paper has navigated through security in both Cloud Computing and High-Performance Computing (HPC), providing a comprehensive understanding of their distinctive characteristics, real-world incidents, and mitigation strategies. Beginning with an introduction of cloud computing, including its key services and deployment models, the exploration extended to the realm of HPC, highlighting its key attributes and diverse applications. Real-world security breaches, such as the Capital One data breach and the NERSC incident, underscored the critical importance of robust security measures.

The paper delved into major security concerns in cloud computing, scrutinizing challenges associated with virtualization, data storage in public clouds, and multitenancy. Thorough discussions on security methods, encompassing data encryption and cryptography, offered actionable insights to fortify cloud security. Similarly, security technologies in HPC, such as data loss prevention, authentication, were examined, providing a holistic view.

Providing practical security recommendations for both HPC and cloud computing, the paper emphasized the significance of periodic audits, stringent device testing, and the implementation of stateful firewall networks. The discussion on benchmarking frameworks, including SCAP and the NIST Cybersecurity Framework, highlighted their pivotal role in evaluating and enhancing security postures.

In a rapidly evolving digital landscape, this paper aims to equip organizations with the knowledge and tools necessary to navigate and fortify their cloud and HPC environments against an ever-expanding array of security challenges.

References

[1] Johannes Buchmann. *Introduction to cryptography*, volume 335. Springer, 2004.

[2] Ramesh Bulusu, Pallav Jain, Pravin Pawar, Mohammed Afzal, and Sanjay Wandhekar. Addressing security aspects for hpc infrastructure. In *2018 International Conference on Information and Computer Technologies (ICICT)*, pages 27–30. IEEE, 2018.

[3] Teresa Davies, Christer Karlsson, Hui Liu, Chong Ding, and Zizhong Chen. High performance linpack benchmark: a fault tolerant implementation without checkpointing. In *Proceedings of the international conference on Supercomputing*, pages 162–171, 2011.

[4] Roberto R Expósito, Guillermo L Taboada, Sabela Ramos, Juan Touriño, and Ramón Doallo. Performance analysis of hpc applications in the cloud. *Future Generation Computer Systems*, 29(1):218–229, 2013.

[5] Gregory S Gaglione Jr. The equifax data breach: an opportunity to improve consumer protection and cybersecurity efforts in america. *Buff. L. Rev.*, 67:1133, 2019.

[6] Chunye Gong, Jie Liu, Qiang Zhang, Haitao Chen, and Zhenghu Gong. The characteristics of cloud computing. In *2010 39th International Conference on Parallel Processing Workshops*, pages 275–279. IEEE, 2010.

[7] Dan Huang, Zhenlu Qin, Qing Liu, Norbert Podhorszki, and Scott Klasky. A comprehensive study of in-memory computing on large hpc systems. In *2020 IEEE 40th International Conference on Distributed Computing Systems (ICDCS)*, pages 987–997. IEEE, 2020.

[8] Anton Jäger. Finding and evaluating the effects of improper access control in the cloud, 2021.

[9] Barbara Krumay, Edward WN Bernroider, and Roman Walser. Evaluation of cybersecurity management controls and metrics of critical infrastructures: A literature review considering the nist cybersecurity framework. In *Secure IT Systems: 23rd Nordic Conference, NordSec 2018, Oslo, Norway, November 28-30, 2018, Proceedings 23*, pages 369–384. Springer, 2018.

[10] Vladimir Marjanović, José Gracia, and Colin W Glass. Performance modeling of the hpcg benchmark. In *High Performance Computing Systems. Performance Modeling, Benchmarking, and Simulation: 5th International Workshop, PMBS 2014, New Orleans, LA, USA, November 16, 2014. Revised Selected Papers 5*, pages 172–192. Springer, 2015.

[11] Gihan R Mudalige, IZ Reguly, Michael B Giles, AC Mallinson, WP Gaudin, and JA Herdman. Performance analysis of a high-level abstractions-based hydrocode on future computing systems. In *High Performance Computing Systems. Performance Modeling, Benchmarking, and Simulation: 5th International Workshop, PMBS 2014, New Orleans, LA, USA, November 16, 2014. Revised Selected Papers 5*, pages 85–104. Springer, 2015.

[12] SAMUEL KINUTHIA NDURA. *ANALYSIS OF SECURITY VULNERABILITY IN DROPBOX CLOUD DATA EXCHANGE AND STORAGE.* PhD thesis, UNIVERSITY OF SCIENCE AND TECHNOLOGY, 2016.

[13] Nelson Novaes Neto, Stuart Madnick, Moraes G de Paula, Natasha Malara Borges, et al. A case study of the capital one data breach. *Stuart E. and Moraes G. de Paula, Anchises and Malara Borges, Natasha, A Case Study of the Capital One Data Breach (January 1, 2020)*, 2020.

[14] Hiral B Patel and Nirali Kansara. Cloud computing deployment models: A comparative study. *International Journal of Innovative Research in Computer Science & Technology (IJIRCST)*, 2021.

[15] Thomas Sterling, Maciej Brodowicz, and Matthew Anderson. *High performance computing: modern systems and practices*. Morgan Kaufmann, 2017.

[16] William Voorsluys, James Broberg, and Rajkumar Buyya. Introduction to cloud computing. *Cloud computing: Principles and paradigms*, pages 1–41, 2011.

[17] Richard W Watson. High performance storage system scalability: Architecture, implementation and experience. In *22nd IEEE/13th NASA Goddard Conference on Mass Storage Systems and Technologies (MSST'05)*, pages 145–159. IEEE, 2005.

[18] Shao-Fang Wen and Basel Katt. A quantitative security evaluation and analysis model for web applications based on owasp application security verification standard. *Computers & Security*, 135:103532, 2023.

[19] Katinka Wolter, Alberto Avritzer, Marco Vieira, and Aad Van Moorsel. Resilience assessment and evaluation of computing systems. 2012.