Seventh International Conference in Information and Communication Technologies, Lusaka, Zambia
15th to 16th October 2025

A Systematic Literature Review on inhibitors and challenges in Cloud Computing as a Platform for Big Data Storage and Analysis

Calvin Swatulani Silwizya
School of ICT. Zambia University College of Technology, Ndola, 10101, Zambia
E-mail: csilwizya@zut.edu.zm

Naomi Mutampuka Mwaba
School of MNS. Mukuba University Kitwe, 10101, Zambia
E-mail: naomimutampukamwaba@gmail.com

Gilbert Samukonga
School of MNS. Mukuba University Kitwe, 10101, Zambia
E-mail: gsamukonga@mukuba.edu.zm

Abstract - The ability to handle large volumes of data at the same time, allowing a flexible, reliable, and a more secure access to this data, has proved that cloud computing is appropriate even more in Big Data analysis. Effective and efficient organization of large-scale data analysis illustrates a critical challenge. Lately, cloud computing has played a significant role in the storage and processing of large data sets. This systematic literature review (SLR) aims to review the existing research studies on cloud computing as a platform for Big Data analysis. The search was done through IEEExplore, Springer, Science Direct, Arxiv, ACM, Tandfonline, DAOJ, Emerald insight and hindawi databases in order to sprout out any article related either directly or indirectly to cloud computing. This SLR examined the research studies published between 2020 and 2025 within the popular digital libraries. About one hundred and fifty (150) articles were identified but twenty (20) papers were selected after a meticulous screening of published works to answer the proposed research questions. This systematic literature review presents the key issues of cloud computing in big data processing, including elements of platforms, cloud architecture, cloud database and data storage scheme. Finally, it discusses the open issues and challenges, and deeply explores the research directions in the future on big data analysis in cloud computing platforms.

Keywords: Analysis, Cloud Computing, Big Data, Cloud Models, challenges,

[1] INTRODUCTION

The rapid increase in organisations data creation poses significant challenges of managing vast complex datasets that hinges on data analysis [1]. Cloud computing service models are systematically elucidated with reference to exhaustive examination of deployment models, this involves public, private, hybrid, and community clouds [2]. Big data" refers to a collection of datasets that are so large and intricate that they are difficult to handle with conventional data processing software or database management tools. The issues associated with big data processing include sharing, transferring, storing, searching, analysing, and visualizing the data. For competitive advantage, innovation, and well-informed decision-making in a variety of industries, nominal processing and analysis of large datasets are essential. Government, business, and academic community are paying more attention to big data in order to improve decision-making and processes [3]. Large-scale data computing platforms have become indispensable instruments to tackle these problems.

Large amounts of data are produced by sectors like financial services, retail, healthcare, pharmaceuticals, and telecommunications. In today's operations, it is more important than ever to extract valuable business insights through the examination and analysis of these enormous datasets. But when confronted with the problem of handling ever-increasing data volumes, current data warehouses and associated solutions usually fail to deliver effective response times. Businesses are frequently forced to perform analytics that take days to process large amounts of data or to perform transactions on smaller datasets in a matter of seconds. [4] Addressing these constraints by guaranteeing real-time or nearly real-time responsiveness when managing large data sets has become essential due to changing demands.

Cloud Computing is based on pay-per-use services for enabling convenient, on-demand network access to a shared pool of configurable computing resources such as servers, networks, and services that can be rapidly provisioned and released with minimal management effort or service provider interaction”. Studies have acknowledged some of the limitations when dealing with data within a cloud, including: availability, accessibility, data migration, data resilience, scalability, performance, efficient multi-tenancy, privacy and security, yet few have attempted to quantify and investigate these issues and explore potentially inter-related implications. When the volume of data at rest or in motion makes data management a key factor in system architecture design, the new big data paradigm emerges [5]. Better understanding of the inter-relations between various cloud components will potentially enable the innovation of new systems based on existing components and systems for enhancing flexibility, extendibility, availability, optimization and cost efficiency.

II. RESEARCH METHOD

The methods and recommendations that were put forward by [6] were followed in conducting this systematic literature review, and they were combined with [7]. This literature review's primary goal was to conduct a thorough search for all published works that are pertinent to the research issue. The researcher was able to identify the problem statement, collect data, identify relevant research papers from trustworthy databases, analyse and evaluate the essential data acquired, and present the findings in accordance with the Kitchenham requirements.

A. Research Questions

The purpose of this study is to address the following queries:

1. What techniques and technologies are employed in applying cloud computing as a platform for big data storage and analysis?

2. What elements influence the choice and application of cloud computing as a platform for big data storage and analysis?
3. What are the inhibitors and challenges that would prevent cloud computing from being employed as big data storage and analysis?
4. How can cloud computing be accepted as a Platform for Big Data Storage and Analysis?

B. Search Strategy

The following databases were searched in order to find any relevant articles that could be directly or indirectly related to cloud computing: IEEExplore, Springer, Science Direct, Arxiv, ACM, Tandfonline, DAOJ, Emerald insight, and Hindawi. This was done in order to provide appropriate answers to the research questions. The subject of this SLR looked at research papers that were published in the aforementioned digital libraries from 2013 and 2023. After a thorough review of published publications to address the suggested study issues, almost 150 articles were found, however only 20 papers were chosen. To expand the search radius, the subsequent search query was employed: Cloud computing for the study of large data.

C. Inclusion Criteria

As proposed by Kitchenham in the systematic literature review guidelines (Barbara Kitchenham, 2009), the primary research studies that directly address the formulated research questions were included in this study based on the following criteria provided:

i. Studies conducted since 2020
ii. Studies discussing cloud computing as a platform for Big data storage and analysis only
iii. Articles published in the nine (9) popular digital databases that is IEEExplore, Springer, Science Direct, Arxiv, ACM, Tandfonline, DAOJ, Emerald insight and hindawi databases
iv. Publications directly or indirectly related to cloud computing and Big data
v. Studies written in English language

D. Exclusion Criteria

i. Studies published before 2020
ii. Studies whose discussion is out the topic area
iii. Studies that does not answer to the research questions
iv. Studies conducted from outside the topic of this systematic literature review

E. Data Extraction and Synthesis

The aim of the data extraction stage is to gather primary studies that utilize different research methods such as quantitative, qualitative, mixed methods including the literature reviews from studies. The data extracted from the primary sources was arranged as below:

i. Authors of the primary studies
ii. Year published in the Journal
iii. Electronic source
iv. Topic Area
v. Research methodology
vi. Data collection method
vii. Summary of the findings.

The findings of the study are now presented with respect to the research questions that guided the execution of the systematic literature review. This segment was achieved using the search strategy discussed and a total of 150 articles were retrieved from the electronic databases identified in the search strategy. After applying the inclusion and exclusion criteria method, only 20 articles were included in this study to be reviewed as indicated in table 2.5.1.

TABLE 2.5.1 ELECTRONIC ARTICLES RETRIEVED

Digital Library	IEEExplore	Springer	Science Direct	Arxiv	ACM	T & F	DAOJ	Emerald insight	Hindawi	Total
Total No. of Studies	29	35	12	21	32	5	3	3	9	150
Included No. of Studies	4	5	4	2	2	1	0	0	2	20
Excluded No. of Studies	25	30	8	19	30	4	3	3	7	130

Table 2.5.2 outlines the range of the primary studies which were considered for adoption in this study and the papers formed the bank for the data collection process and provided the answers to the research questions.

TABLE 2.5.2 PAPERS ADOPTED FOR REVIEW

ISBN: 978-9982-95-500-3

Article ID	Author	Topic Area	Year	Digital Library
A1	Robb Shawe	Cloud Computing: Purpose and Future	2024	Hindawi
A2	A. K. Sandhu	Big data with cloud computing: Discussions and challenges	2021	IEEE
A3	Gautami et	Cloud Algebra: An Innovative Approach for Managing Resources, Services and Big Data on Clouds	2020	Hindawi
A4	Fernández et al	Real-time big data processing for instantaneous marketing decisions: A problematisation approach	2014	Science Direct
A5	Gibson et al.	Benefits and challenges of three cloud computing service models	2021	Arxiv
A6	Gudadhe & Agrawal	Analysis of Data Replication Strategies in Cloud Environment.	2022	Science Direct
A7	Abel Adane	Cloud Services Quality Enhancement through Control & Monitoring of Service Level Agreements at Server and Consumer Side	2020	Hindawi
A8	Silvano Gai	Building a Future-Proof Cloud Infrastructure: A Unified Architecture for Network, Security, and Storage Services Building a Future-Proof Cloud Infrastructure: A Unified Architecture for Network, Security, and Storage Services	2020	ACM
A9	Manjunath et al	Big Data MapReduce Hadoop Distribution Architecture for Processing	2021	Science Direct
A10	Mell & Grance	The NIST definition of Cloud Computing	2022	T & F
A11	Mushtaq et al	Cloud Computing Environment and Security Challenges	2021	Science Direct
A12	Palos-Sanchez et al	Adoption of SaaS as a Strategic Technology	2022	IEEE
A13	Sallehudin et al	Factors Influencing Cloud Computing Adoption in the Public Sector	2023	ACM
A14	Wook et al	Big data traits and data quality dimensions for big data analytics	2023	Springer
A15	Yang et al	Big data enriches critical review of Automatic Identification System (AIS) data applications.	2020	IEEE
A16	Wang et al	Efficient allocation of cloud resources	2022	Springer
A17	Jabalia et al	Ensures availability of cloud resources	2022	Springer
A18	Cui et al	Cost of scheduling in Cloud Computing	2021	IEEE
A19	Patel et al	Cloud services perfomance	2020	Springer
A20	Singh et al	Migration of Cloud services	2020	Springer

For a complete and thorough analyse, the findings of the primary studies as guided by [6], the results were examined by identifying whether the findings are consistent with one another or are inconsistent. The guidelines also suggest that results can be recorded in the tables for examination in terms of consistency or inconsistency relations.

III. RESULTS

All of the evaluated articles were categorized into four groups: techniques and technologies used for cloud computing as a platform for big data storage and analysis, elements that influence the selection of a cloud computing platform, inhibitors and challenges preventing use of a cloud computing platform as well as the strengths of cloud computing as a platform for big data storage and analysis.

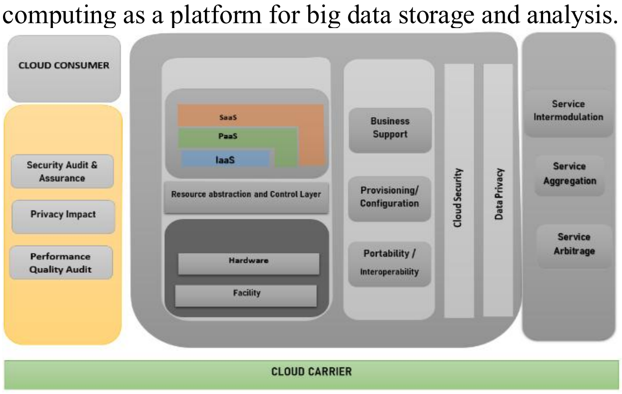

FIGURE 3.0 CLOUD COMPUTING REFERENCE ARCHITECTURE

A. Techniques and technologies used for implementing cloud computing as a platform for big data storage and analysis

Cloud Computing is the delivery of computing services such as servers, storage, databases, networking, software, analytics etc., over the Internet ("the cloud") with the aim of providing flexible resources, faster innovation and economies of scale [8]. About seven (7) papers stated that Cloud Computing provides combination and integration function as well, but it provides the function in the background of Web [9]. Further the papers have outlined different technologies used in different cloud computing service types, which is Infrastructure as a Service (IaaS), Platform as a Service (PaaS), and Software as a Service (SaaS) [10].

Infrastructure as a Service

IaaS is the cloud computing service delivered in the form of a platform in a virtual environment [11] . Clients are not required to purchase servers, data centers, network equipment or space.

Software as a Service (SaaS)

The SaaS model facilitates users to access the software and other programs in a cloud. Using the SaaS solution eliminates the need for in-house applications, data storage, and support for the application administration. Companies pay to use the SaaS resources on a user basis [12].

Platform as a Service (PaaS)

PaaS is a cloud computing service that supports a full software life cycle and allows users to develop cloud applications and services [13]. Programmers and developers do not need to purchase their equipment; instead, they use intermediary equipment and deliver the developed applications to clients over the internet. Platforms in these fields include Google App Engine, Microsoft Azure, Map Reduce and Hadoop for big data analysis. MapReduce/Hadoop is a data processing and analytics technology that has been revolutionary in the realm of computer science and is one of the hottest technologies in the big data space [14]. Table 3.1 shows cloud computing technologies.

In the selection and implementation of cloud computing technologies there are cardinal factors that are considered. These decisive factors must be systematically evaluated prior to making the decision to adopt cloud-based solutions. This review has established through a number of papers that the following factors affect the selection and implementation of cloud technologies. The benefits of Cloud Computing must be greater than the costs of this adoption and Security, privacy and ethics concern. [15]. The need for robust security measures, including encryption, access controls, and monitoring, can influence the choice of cloud provider and service type. With Cloud Computing there is a reduction of energy costs and environmental costs [16]. Maintenance costs of Cloud Computing are very low. Selecting and

 ISBN: 978-9982-95-500-3

implementing cloud computing technologies requires careful consideration of various factors, including business needs, cost, security, performance, scalability, and technical expertise. A well-rounded approach that balances these considerations can help organizations choose the right cloud solutions, ensuring they meet both current and future requirements while mitigating risks.

TABLE 3.1 CLOUD COMPUTING TECHNOLOGIES

Service type	IaaS	PaaS	SaaS
Service category	VM Rental, Online Storage	Online Operating Environment, Online Database, Online Message	Application and Software Rental
Service resource optimization	Network Virtualization, Server Virtualization, Storage Virtualization	Large-scale Distributed File System, Database, Middleware etc	Multi-tenancy
Service monitoring	Physical Resource Monitoring	Logic Resource Monitoring	Application Monitoring
Service security	Storage Encryption and Isolation, VM Isolation	Data Isolation, Operating Environment Isolation,	Data Isolation, Operating Environment Isolation, Web Authentication and Authorization

B. Elements that influences the choice and application of cloud computing as a platform for big data storage and analysis

The rapid expansion and complexity of monitoring data pose significant challenges to the online monitoring systems for power equipment [17]. These challenges include issues such as real-time processing, swift variations, high precision, diverse application scenarios, abnormal data induced by power fluctuations, and handling large-scale, high-dimensional datasets with intricate structures [18]. Broadly, research efforts in addressing these challenges can be categorized into two approaches: traditional detection methods and data mining-based detection methods.

Traditional methods for abnormal detection in power load data rely heavily on heuristic approaches grounded in human experience, state estimations, load curve analysis, similarity assessments, and rate of load change. While these methods are effective in identifying anomalous data with pronounced deviations, they suffer from limitations such as low efficiency, high subjectivity, and susceptibility to human error [18]. Consequently, they are inadequate for detecting subtle anomalies in power load data.

Advancements in artificial intelligence and cluster analysis theory have significantly enhanced the capabilities of data mining techniques for detecting anomalies in power load datasets [19]. These techniques can be broadly classified into two categories: supervised learning-based methods and unsupervised learning-based methods. Supervised learning approaches involve training models using segments of historical power system load data and employing algorithms, such as support vector machines, artificial neural networks, and decision trees, to ensure the trained models achieve the desired outputs [20].

In contrast, unsupervised learning methods do not require preselected training samples from historical datasets. These methods typically include techniques such as density analysis and cluster analysis [21]. For instance, distance-based outlier detection algorithms—an unsupervised approach—are not only intuitive and interpretable but also highly effective in processing medium to high-dimensional data. However, density-based anomaly detection methods often require the specification of certain parameters based on subjective human judgment, which may compromise their precision.

Lastly, clustering algorithms offer significant promise for abnormal load data detection by effectively distinguishing between normal and anomalous data points within an electric load dataset [22]. These algorithms leverage the distinctive characteristics inherent to power load data to enhance accuracy compared to traditional and density-based techniques. However, addressing challenges related to parameter selection remains vital for improving the robustness of anomaly detection systems in power monitoring applications.

C. The inhibitors and challenges that prevents cloud computing from being employed as big data storage and analysis

One of the limitations of cloud computing and its technologies is resource allocation, load balancing, scheduling, bandwidth and quality of networks [23]. As the consumers can access cloud from anywhere anytime and cannot access the cloud resources directly, instead there are different APIs which help with requesting for storage and higher processing power . The cloud support on-demand flexible resource allocation, but this can lead to non-optimal resource allocation. The restrictions such as the response time that is measuring the effectiveness of resource allocation irrespective of the category of resources allocated to the customers.

Cloud computing tools and technologies offer significant strengths in scalability, cost-efficiency, accessibility, and advanced analytics capabilities. However, these benefits come with limitations, including security concerns, performance issues, cost management challenges, vendor lock-in, and data governance complexities [24]. To fully leverage the potential of cloud computing, organizations must carefully assess these strengths and limitations, adopting strategies that address the specific needs of their operations and the inherent risks of cloud environments.

Network latency, bandwidth limitations, and the performance variability of virtualized resources can impact the efficiency of cloud-based applications, particularly for real-time data processing or applications with high-performance requirements.

D. Acceptance of Cloud Computing as a Platform for Big Data Storage and Analysis

The user or organization acceptance of new technology mainly depends on how technology can help them to perform tasks more quickly and efficiently. The ability and greatness of technology are seen as a tool to facilitate the task, but its effectiveness and benefits for users are subjective in many issues and circumstances [25]. Many studies have been conducted in measuring users' acceptance of the implementation and use of cloud computing platforms and technologies in both individual and organisation levels of data analysis and organizational levels. Furthermore, these models and theories introduce factors that can affect user acceptance. Technology-Organization-Environment (TOE) framework by Tornatzky. The framework sees a direct connection between the adoption of IT technology with three contexts (technological, organizational, and environmental) by organizations. Diffusion of Innovation (DOI) theory by Rogers - Theory of DOI explains how innovation is acceptable to users, how it is disseminated, and to what extent is it acceptable to it. The acceptance of cloud computing as a platform for big data storage and analysis is a testament to its scalability, flexibility, and cost-effectiveness. The combination of advanced analytics capabilities, global accessibility, and improved security has made cloud computing an attractive option for organizations of all sizes [26]. However, successful adoption requires careful consideration of data governance, cost management, and the potential for vendor lock-in. As cloud computing continues to evolve, it is likely to play an increasingly central role in big data strategies across industries.

Table 3.4 Identifying issues and solutions presented in reviewed articles

Article ID	Issues Presented	Solutions Presented
A1	Cost challenges in using Cloud Computing tools with high cost of software licenses.	Provides a dynamic infrastructure that provides reduced cost and improved services with less development and maintenance cost.
A2	Requires high speed network and connectivity constantly. Privacy and security is not good	unlimited processing, storage, networking etc. in an elastic way
A3	Disastrous situation is unavoidable and recovery is not possible always.	Can recover rapidly and has improved restore capabilities for improved business resiliency
A4	Data analysis processes and users have external dependency for mission critical applications.	environment friendly computing as it only uses the server space required by the application.
A5	the cloud must provide, improved performance when a user moves to cloud computing infrastructure.	Implementations of open source and Virtualisation
A6	lack of proper resources disk space, limited bandwidth, lower CPU speed, memory, network connections	With Cloud Computing there is a reduction of energy costs and environmental costs
A7	Lack of Reliability and Availability	Firewall configuration, server patching, and placement of the intrusion-detection system are enforced.
A8	Lack of Scalability	adding more nodes to the system such as adding a new computer to an existing service provider system
A9	Elasticity	adding of resources to a single node in the system such as memory
A10	Interoperability	open APIs and standard interfaces for VM formats and service deployment interfaces.
A11	Abuse, insecure interfaces, and nefarious usage were the vulnerable threats.	Extracting the device information, followed by the ID decryption in the second step. The final step involves the decryption
A12	Resource Management	Algorithms for managing turnaround time, response time, waiting time, throughput and resource utilization
A13	System Scheduling	Processing jobs execution order to finish job execution and optimize some parameters
A14	Bandwidth Cost	Hosting cloud computing services in data centers
A15	Portability	developing open APIs and standard interfaces for VM formats and service deployment interfaces.
A16	Uploading bulk data of unevenly sized cause uneven distribution for jobs	Bulk data uploads are normally followed by the creation of new tags or the modification of existing ones.
A17	Uneven tasks, process and data cause outliers in job completion times	Disco has proved to be a robust and flexible platform on which to build a data analytics stack.
A18	Cloud Computing Service attacks	Secure from denial of service attacks, performance slowdowns, equipment outages and natural disasters.
A19	Lack of load balancing	An efficient load balancing to maximize the cloud provider's network throughput, avoid the overload and reduce the waiting time
A20	The level of difficulty experienced by users to learn and use cloud computing services.	Provision of video tutorials for users to learn how to access cloud services

IV. DISCUSSION AND OPEN ISSUES

From the overview in this paper, that basic functions of Cloud computing, such as development models and basic services are well documented and from the reviewed papers the comparatives were done in terms of challenges and solutions presented

Comparing Challenges

With regard to the results of the analysis described above in table 3.4 the twenty (20) challenges presented from the articles pertaining to the use of cloud computing as a platform in big data storage and analysis. Article A1, A2 and A14 shows the cost challenges in using Cloud Computing tools and technologies with high cost of software licenses as well as bandwidth, servers and network equipment especially in developing countries. Also the need for high speed network and connectivity constantly in order to access data at a fast rate to trigger response. The two articles also emphasize privacy and security.

Study A7, A8, A9, A10, A11 and A18 brings out that security is a major factor in cloud computing without firewall configuration, server patching, and placement of the intrusion-detection system are enforced the services can't be accessible or can be breached. Lack of reliability, availability, scalability, elasticity and interoperability attributes to cloud services not being effective and these

 ISBN: 978-9982-95-500-3

factors affect the selection and implementation of cloud computing tools and technologies for data storage and analysis.

B. Comparing Solutions

The reviewed articles A20, A19, A18, A12 and A11 present solutions to marked challenges in security and provide an efficient load balancing to maximize the cloud provider's network throughput, avoid the overload and reduce the waiting time. Provision of video tutorials for users to learn how to access cloud services enables users accepts new technologies and cloud computing services. Extracting the device information, followed by the ID decryption is one of the secure step in implement API's (Stilgoe, 2018). The papers also show that developing algorithms for managing turnaround time, response time, waiting time, throughput and resource utilization

V. CONCLUSIONS

In this Systematic Literature Review the literature on cloud computing and Big data analysis topics have been reviewed, including cloud security threats and their mitigation strategies. Cloud computing tools and technologies were identified with several security risks to cloud computing. Data tampering and leakage. This paper also discusses the limitations and strengths of these tools and technologies and factors that affect the selection and utilization of technologies and tools used in cloud computing. The paper has also established the challenges that affects the implementation of cloud computing in the storage and analysis of Big data

A. The Complexities of Cross-Cloud Data Analysis

1. *Data Silos:* Every cloud service has different methods for managing and storing data. Because of this, data frequently gets segregated, which makes it challenging to access and analyse across platforms. This can be resolved by transferring data between cloud providers smoothly utilizing a cloud data integration solution like Talend or Apache NiFi. Furthermore, data latency can easily happen during cloud data transfers, which might affect real-time data analysis and decision-making [27]. Nevertheless, use edge computing to analyse data in real time and lessen the requirement for cross-cloud data transfers.

2. *Diverse Data Formats:* A data transformation layer that can dynamically convert data formats is necessary because different cloud providers may employ different data formats and storage technologies, which makes it difficult to standardize data for analysis [28]. Consequently, more tools—like Apache—are required, which may have an expense.

3. *Security and Compliance:* A centralized identity and access management (IAM) solution policy and encryption key management may be required to standardize security protocols. Ensuring data security and compliance with various providers' policies is a tangle, and various providers have unique encryption techniques, identity management systems, and compliance standards.

4. The Cross-Cloud Data Analysis challenges could be minimised by:

Data Virtualisation: Platforms for data virtualization, such as Cisco Data Virtualization, can abstract the underlying complexity and offer a consistent view of data across many clouds. Google Cloud into a unified virtualized dataset for further examination.

Containerisation and Orchestration: Using containerisation tools like Kubernetes and many others could deploy data analysis applications consistently across various cloud environments. For example, machine learning models developed on Google Cloud can be containerised and deployed in AWS or Azure without significant modification.

Cross-Cloud ETL Tools: ETL (extract, transform, load) technologies, such as Talend or Stitch Data, facilitate smooth data transfer between clouds by automating the operations of data extraction and transformation. Formats for Interoperable Data: To ensure cross-platform portability, choose widely-used data formats for processing and storing, such as Parquet, Avro, or Apache Arrow. For example, by storing data in Apache Parquet format, a data analytics team can easily analyze the data in Google BigQuery, AWS Redshift, or Snowflake.

Edge Computing and Hybrid Cloud: Utilise edge computing for real-time data analysis at the source, reducing the need for cross-cloud data transfers and minimising latency.

To sum up, cross-cloud data analysis has many benefits, but there are drawbacks as well. Data portability and interoperability must be smooth as more organizations and corporations implement multi-cloud strategies. Organizations may fully utilize their data to make informed decisions and achieve a competitive edge in the data-driven world by comprehending and tackling these difficulties. Not to mention, having up-to-date workforce is essential in today's market to handle all the complexities in cloud computing.

The systematic literature review on inhibitors and challenges in cloud computing as a platform for big data storage and analysis reveals a multifaceted landscape of concerns that need to be addressed to fully leverage cloud-based solutions. Security and privacy emerge as the most significant inhibitors, with data breaches, unauthorized access, and compliance with regulations being persistent issues. These challenges are compounded by the inherent complexity of cloud environments, where data is distributed

 ISBN: 978-9982-95-500-3

across multiple locations and managed by third-party providers, raising concerns over control and data governance.

Performance-related challenges, such as latency, bandwidth limitations, and resource allocation, also pose significant barriers. The scalability and elasticity of cloud platforms are crucial for handling the vast volumes of data characteristic of big data applications. However, the unpredictability of resource demand and the need for efficient data processing frameworks require sophisticated management to avoid bottlenecks and ensure optimal performance.

Cost is another critical factor, where the pay-as-you-go model of cloud services can lead to unexpectedly high expenses, especially when dealing with large-scale data storage and processing. This is further exacerbated by the need for continuous monitoring and optimization to balance cost with performance and resource utilization.

Interoperability and vendor lock-in present additional challenges, as organizations seek to integrate cloud services with existing on-premise systems and avoid dependency on a single provider. The lack of standardized protocols and the diverse nature of cloud offerings make it difficult to achieve seamless integration and migration between different platforms.

In conclusion, while cloud computing offers significant advantages for big data storage and analysis, such as scalability, flexibility, and cost-effectiveness, these benefits are often offset by challenges that need careful consideration and management. Addressing the inhibitors identified in this review—security, performance, cost, and interoperability—requires a strategic approach that involves selecting the right cloud architecture, implementing robust security measures, and adopting best practices in cloud management. As the cloud computing landscape evolves, ongoing research and innovation will be essential in overcoming these challenges and unlocking the full potential of cloud-based big data solutions.

REFERENCES

[1] "Machine Learning and Deep Learning for Big Data Analytics: A Review of Methods and Applications | Partners Universal International Innovation Journal." Accessed: Aug. 22, 2025. [Online]. Available: https://www.puiij.com/index.php/research/article/view/144

[2] Ü. Demirbaga, G. S. Aujla, A. Jindal, and O. Kalyon, "Cloud Computing for Big Data Analytics," in *Big Data Analytics: Theory, Techniques, Platforms, and Applications*, Ü. Demirbaga, G. S. Aujla, A. Jindal, and O. Kalyon, Eds., Cham: Springer Nature Switzerland, 2024, pp. 43–77. doi: 10.1007/978-3-031-55639-5_4.

[3] J. Wang, Y. Yang, T. Wang, R. S. Sherratt, and J. Zhang, "Big Data Service Architecture: A Survey," *J. Internet Technol.*, vol. 21, no. 2, pp. 393–405, Mar. 2020.

[4] R. Gupta, H. Gupta, and M. Mohania, "Cloud Computing and Big Data Analytics: What Is New from Databases Perspective?," in *Big Data Analytics*, S. Srinivasa and V. Bhatnagar, Eds., Berlin, Heidelberg: Springer, 2012, pp. 42–61. doi: 10.1007/978-3-642-35542-4_5.

[5] G. Agapito and M. Cannataro, "An Overview on the Challenges and Limitations Using Cloud Computing in Healthcare Corporations," *Big Data Cogn. Comput.*, vol. 7, no. 2, p. 68, Jun. 2023, doi: 10.3390/bdcc7020068.

[6] A. Carrera-Rivera, W. Ochoa, F. Larrinaga, and G. Lasa, "How-to conduct a systematic literature review: A quick guide for computer science research," *MethodsX*, vol. 9, p. 101895, Jan. 2022, doi: 10.1016/j.mex.2022.101895.

[7] "Automating Systematic Literature Reviews with Natural Language Processing and Text Mining: A Systematic Literature Review | SpringerLink." Accessed: Aug. 22, 2025. [Online]. Available: https://link.springer.com/chapter/10.1007/978-981-99-3243-6_7

[8] A. K. Sandhu, "Big data with cloud computing: Discussions and challenges," *Big Data Min. Anal.*, vol. 5, no. 1, pp. 32–40, Mar. 2022, doi: 10.26599/BDMA.2021.9020016.

[9] R. Shawe, "Cloud Computing: Purpose and Future," *J. Softw. Eng. Appl.*, vol. 17, no. 10, pp. 763–769, 2024, doi: 10.4236/jsea.2024.1710041.

[10] M. Ahad, G. Tripathi, and S. Paiva, "Cloud Algebra: An Innovative Approach for Managing Resources, Services and Big Data on Clouds," *EAI Endorsed Trans. Energy Web*, vol. 0, no. 0, p. 162687, Jul. 2020, doi: 10.4108/eai.13-7-2018.162687.

[11] "A Study of Data Storage Security Issues in Cloud Computing - ProQuest." Accessed: Oct. 27, 2023. [Online]. Available: https://www.proquest.com/openview/49d9946f86316032cd86d4cdf5100eb2/1?pq-origsite=gscholar&cbl=5314840

[12] "Cloud Computing: Concepts, Technology, and Architecture," abe.pl. Accessed: Aug. 22, 2025. [Online]. Available: https://www.abe.pl/en/book/9780138052256/cloud-computing-concepts-technology-and-architecture

[13] "Bloom Filter: A Data Structure for Computer Networking, Big Data, Cloud Computing, Internet of Things, Bioinformatics and Beyond," abe.pl. Accessed: Aug. 22, 2025. [Online]. Available: https://www.abe.pl/en/book/9780128235201/bloom-filter-a-data-structure-for-computer-networking-big-data-cloud-computing-internet-of-things-bioinformatics-and-beyond

[14] K. Hsu, "Big data analysis and optimization and platform components," *J. King Saud Univ. - Sci.*, vol. 34, no. 4, p. 101945, Jun. 2022, doi: 10.1016/j.jksus.2022.101945.

[15] "Evaluating the benefits of Cloud Computing in Small, Medium and Micro-sized Enterprises (SMMEs) -

ScienceDirect." Accessed: Aug. 22, 2025. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S187705092100274X

[16] "Energy efficiency in cloud computing data centers: a survey on software technologies | Cluster Computing." Accessed: Aug. 22, 2025. [Online]. Available: https://link.springer.com/article/10.1007/s10586-022-03713-0

[17] W. Lin, F. Shi, W. Wu, K. Li, G. Wu, and A.-A. Mohammed, "A Taxonomy and Survey of Power Models and Power Modeling for Cloud Servers," *ACM Comput Surv*, vol. 53, no. 5, p. 100:1-100:41, Sep. 2020, doi: 10.1145/3406208.

[18] H. Ben Hassen, N. Ayari, and B. Hamdi, "A home hospitalization system based on the Internet of things, Fog computing and cloud computing," *Inform. Med. Unlocked*, vol. 20, p. 100368, Jan. 2020, doi: 10.1016/j.imu.2020.100368.

[19] "Anomaly Detection of Energy Consumption in Cloud Computing and Buildings Using Artificial Intelligence as a Tool of Sustainability: A Systematic Review of Current Trends, Applications, and Challenges | SpringerLink." Accessed: Aug. 23, 2025. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-031-45214-7_9

[20] I. I. Imran, "Exploiting Anomalies with Data Mining Techniques to Enhance Cloud Security. | EBSCOhost." Accessed: Aug. 23, 2025. [Online]. Available: https://openurl.ebsco.com/contentitem/doi:10.18280%2Fmmep.120227?sid=ebsco:plink:crawler&id=ebsco:doi:10.18280%2Fmmep.120227

[21] J. Schmitt, J. Bönig, T. Borggräfe, G. Beitinger, and J. Deuse, "Predictive model-based quality inspection using Machine Learning and Edge Cloud Computing," *Adv. Eng. Inform.*, vol. 45, p. 101101, Aug. 2020, doi: 10.1016/j.aei.2020.101101.

[22] "Research on Parallel Adaptive Canopy-K-Means Clustering Algorithm for Big Data Mining Based on Cloud Platform | Journal of Grid Computing." Accessed: Aug. 22, 2025. [Online]. Available: https://link.springer.com/article/10.1007/s10723-019-09504-z

[23] A. Jabbar, P. Akhtar, and S. Dani, "Real-time big data processing for instantaneous marketing decisions: A problematization approach," *Ind. Mark. Manag.*, vol. 90, pp. 558–569, Oct. 2020, doi: 10.1016/j.indmarman.2019.09.001.

[24] S. Gathu, "High-Performance Computing and Big Data: Emerging Trends in Advanced Computing Systems for Data-Intensive Applications," *J. Adv. Comput. Syst.*, vol. 4, no. 8, pp. 22–35, Aug. 2024.

[25] B. Berisha, E. Mëziu, and I. Shabani, "Big data analytics in Cloud computing: an overview," *J. Cloud Comput.*, vol. 11, no. 1, p. 24, Aug. 2022, doi: 10.1186/s13677-022-00301-w.

[26] F. M. Awaysheh, M. N. Aladwan, M. Alazab, S. Alawadi, J. C. Cabaleiro, and T. F. Pena, "Security by Design for Big Data Frameworks Over Cloud Computing," *IEEE Trans. Eng. Manag.*, vol. 69, no. 6, pp. 3676–3693, Dec. 2022, doi: 10.1109/TEM.2020.3045661.

[27] H. Allam, "Cross-Cloud Chaos: Strategies for Reliability Testing in Hybrid Environments," *Int. J. Emerg. Trends Comput. Sci. Inf. Technol.*, vol. 4, no. 3, pp. 61–70, Oct. 2023, doi: 10.63282/3050-9246.IJETCSIT-V4I3P107.

[28] "Using Cloud Computing to Analyze Model Output Archived in Zarr Format in: Journal of Atmospheric and Oceanic Technology Volume 39 Issue 4 (2022)." Accessed: Aug. 23, 2025. [Online]. Available: https://journals.ametsoc.org/view/journals/atot/39/4/JTECH-D-21-0106.1.xml