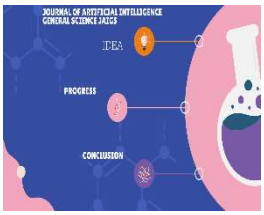
Vol.2, Issue 01, February 2024
Journal of Artificial Intelligence General Science JAIGS

https://ojs.boulibrary.com/index.php/JAIGS



Securing Multi-Cloud Architectures: A Machine Learning Perspective

Sundeep Reddy Mamidi

Dallas, TX, USA.

ARTICLEINFO
Article History:
Received:
01.02.2024
Accepted:
10.02.2024
Online: 29.02.2024
Keyword: Cloud Security, AI-Based Security, Dynamic Security Policies, Machine Learning, Threat Detection

ABSTRACT

Multi-cloud computing, the utilization of multiple cloud computing services in a single heterogeneous architecture, has gained significant traction in recent years due to its potential for enhancing flexibility, resilience, and performance. This paper provides an overview of multi-cloud computing, exploring its key concepts, advantages, challenges, and best practices. It examines the motivations behind adopting multi-cloud strategies, the various deployment models, management approaches, and emerging trends. Additionally, the paper discusses the implications of multi-cloud computing for security, interoperability, and vendor lock-in. Through a comprehensive analysis, this paper aims to offer insights into the complexities and opportunities associated with multi-cloud environments.

Introduction:

In recent years, cloud computing has emerged as a fundamental paradigm for delivering computing resources and services over the internet, enabling organizations to scale their operations, enhance agility, and reduce infrastructure costs. While the adoption of cloud computing continues to accelerate, organizations are increasingly exploring strategies that involve the use of multiple cloud computing services simultaneously, known as multi-cloud computing. This approach offers several potential advantages, including increased flexibility, resilience, and performance optimization. However, it also introduces complexities related to management, security, and interoperability.

This paper provides an overview of multi-cloud computing, aiming to explore its key concepts, benefits, challenges, and best practices. By examining the motivations behind adopting multi-cloud strategies, exploring different deployment models and management approaches, and discussing emerging trends, this paper seeks to offer insights into the intricacies and opportunities associated with multi-cloud environments. Additionally, the implications of multi-cloud computing for security, interoperability, and vendor lock-in are addressed to provide a comprehensive understanding of this evolving landscape.

Objectives:

Objective 1: To examine the key concepts and principles underlying multi-cloud computing, including its definition, characteristics, and components.

Objective 2: To analyze the benefits and challenges associated with adopting a multi-cloud approach, exploring factors such as flexibility, resilience, performance optimization, management complexities, security concerns, and interoperability issues.

Objective 3: To identify and discuss best practices and emerging trends in multi-cloud computing, aiming to provide practical insights and recommendations for organizations considering or already implementing multi-cloud architectures.

Research Method:

The research method employed for this study involves a comprehensive review and analysis of existing literature, academic papers, industry reports, and case studies related to multi-cloud computing.

1. Literature Review: A systematic review of academic articles, conference papers, and books on multi-cloud computing is conducted to gather foundational knowledge and insights into the topic. This includes exploring scholarly databases such as IEEE Xplore, ACM Digital Library, and Google Scholar to identify relevant literature.

2. Data Collection: Relevant information and data pertaining to multi-cloud computing, including definitions, concepts, benefits, challenges, best practices, and emerging trends, are collected from various sources. This includes academic publications, industry reports, whitepapers, and reputable online resources.

3. Analysis: The gathered information is systematically analyzed to identify key themes, trends, and patterns related to multi-cloud computing. This involves categorizing and synthesizing the collected data to extract meaningful insights and perspectives.

4. Synthesis: The findings from the literature review and analysis are synthesized to develop a coherent overview of multi-cloud computing. This includes identifying common trends, discussing recurring themes, and integrating diverse viewpoints to provide a comprehensive understanding of the topic.

5. Interpretation: The synthesized information is interpreted to draw conclusions and implications regarding multi-cloud computing. This involves critically assessing the strengths and limitations of multi-cloud approaches, as well as identifying opportunities for future research and practice.

Overall, the research method adopted for this study aims to provide a rigorous and evidence-based exploration of multi-cloud computing, drawing upon a diverse range of scholarly and industry sources to inform the discussion.

Literature Review:

Securing multi-cloud architectures through a machine learning perspective is crucial in today's complex network environments. Leveraging machine learning algorithms for intrusion detection in cloud systems enhances security by detecting patterns of malicious activities [1]. Additionally, incorporating machine learning techniques in network threat detection frameworks, such as the SmartX Multi-Sec framework, can effectively recognize hidden and complex patterns of network traffic, especially in the presence of distributed edge nodes [2] [3]. Furthermore, utilizing AI-based algorithms like KNN and neural network procedures can significantly improve cloud security by encrypting data stored in cloud servers, thus enhancing information literacy levels for cloud re-appropriation [4]. By amalgamating these approaches, organizations can fortify their multi-cloud architectures against evolving cyber threats effectively.

Background:

According to the National Institute of Standards and Technology (NIST), cloud computing is defined as a model that provides users with access to a shared pool of configurable computing resources, including networks, servers, storage, applications, and services. These resources can be rapidly provisioned with minimal management or interaction with the service provider [25]. While cloud computing isn't a novel technology itself, it represents a novel approach to leveraging existing technologies [26]. Essentially, cloud computing allows for the remote access and management of computing resources, transforming computation into a utility comparable to water, electricity, gas, and telephony [6]. This technology has been widely integrated into businesses, enabling them to focus on core activities without the burden of maintaining complex computing infrastructures.

One of the significant advantages of cloud computing is its ability to optimize the utilization of computational resources, leading to faster results through economies of scale. For instance, while the cost of operating one server for 1000 hours may be equivalent to operating 1000 servers for one hour, the latter scenario leverages the collective processing power to significantly reduce task completion time [6].

Cloud computing exhibits five essential characteristics that enhance its utility for consumers [7]. Firstly, it offers on-demand self-service, allowing users to remotely access resources without requiring human interaction. Additionally, cloud networks can be accessed through various interfaces, including mobile phones, laptops, and desktop computers, facilitating broad network access. Cloud Service Providers (CSPs) typically employ multi-tenant models, pooling resources to serve multiple customers effectively. Moreover, cloud systems offer rapid elasticity, allowing resources to be provisioned and released according to consumer needs. Lastly, cloud services feature metering capabilities, providing transparent feedback for both providers and users [10].

These characteristics make cloud computing particularly appealing for small businesses, as it eliminates the need for upfront investment, reduces operational costs, and offers easy accessibility. By outsourcing infrastructure management to providers, companies can concentrate on core business functions [22].

Presently, three primary cloud computing service models dominate the industry: Software as a Service (SaaS), Platform as a Service (PaaS), and Infrastructure as a Service (IaaS). In SaaS, consumers can access software applications hosted on the provider's cloud infrastructure over the internet. PaaS allows consumers to utilize cloud platforms to develop or run their own applications. Lastly, IaaS provides consumers with virtual storage and computing resources [20, 21].

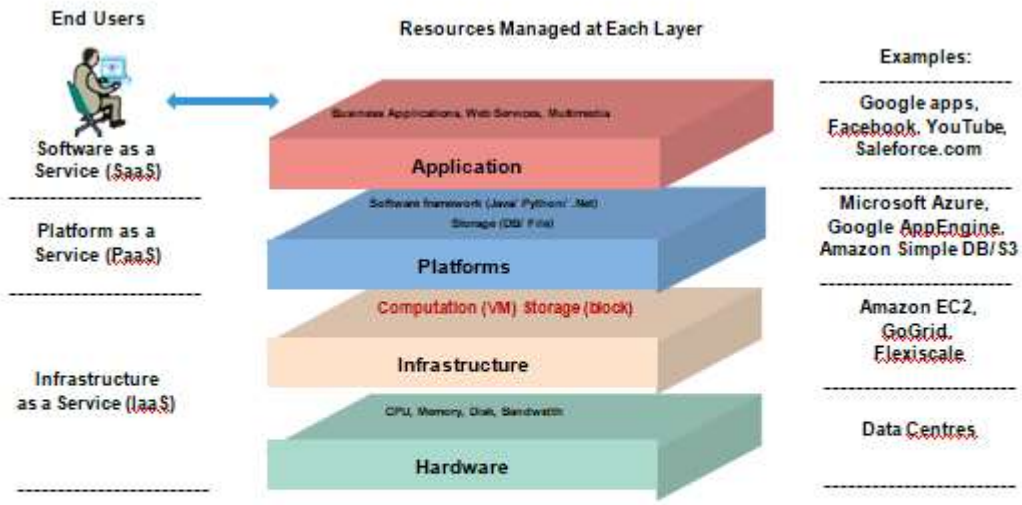


Fig. 1 Cloud Computing Architecture, based on [20]

The architecture of cloud computing, depicted in Figure 1, comprises four layers, commencing with the Hardware Layer. At this foundational level, the Hardware Layer encompasses the physical components, such as servers, switches, power systems, and cooling mechanisms. Typically, these components are housed within data centers housing thousands of servers. Moving up the hierarchy, the Infrastructure Layer utilizes virtualization technologies to partition compute and storage resources. The Platforms Layer consists of operating systems and application frameworks, primarily aimed at simplifying the deployment of applications within virtual machine containers. Finally, the Application Layer encompasses the cloud computing applications, including Software as a Service (SaaS) applications, which are built upon the lower layers (IaaS and PaaS) of the cloud network. For instance, applications available on platforms like Google Play or other app stores leverage cloud infrastructure to execute certain features, facilitating ease of access and reducing computational burden on user devices by offloading operations to the cloud.

Deployment models are crucial to consider in cloud computing. There are four primary deployment models. In the private model, the cloud is provisioned for exclusive use by a single organization with multiple users, such as business consumers. The organization may own and manage the system entirely, engage a third party for services, or employ a combination of both approaches. The community cloud model involves multiple organizations sharing cloud provisioning responsibilities. Similar to private cloud computing, one or more of the organizations utilizing the community cloud may own and operate it, rely on third-party providers, or adopt a hybrid approach. A public cloud, on the other hand, is accessible to the general public and is typically hosted on the premises of the cloud provider, such as a university campus, unlike private or community clouds, which may also be accessible off-premises. Lastly, hybrid cloud systems combine two or more of the aforementioned deployment models, remaining

distinct entities but unified by underlying technologies that enable data and application portability between the units [11].

Challenges in Cloud Computing

After establishing the fundamentals of cloud computing, it is pertinent to delve into the challenges encountered by this technology. While cloud computing offers significant benefits to businesses, such as eliminating the need for extensive provisioning and enabling scalability aligned with business growth, it also presents various issues and obstacles that warrant further investigation [26]. As cloud computing gains widespread adoption, researchers and businesses alike have begun to scrutinize potential challenges associated with the technology.

Security emerges as a primary concern cited by numerous authors, with data security and privacy being particularly emphasized [6, 15, 26, 27]. A study by [15] identified security concerns as a predominant issue in 66 research papers reviewed, surpassing other challenges such as infrastructure and data management. Practitioners interviewed in the study echoed these concerns, underlining the practical significance of security in cloud computing for many stakeholders.

Security issues in cloud computing can be segmented into various subcategories, including safety mechanisms, cloud server monitoring, data confidentiality, and mitigation of malicious operations [27]. Challenges related to data storage in cloud computing networks encompass aspects like data integrity, confidentiality, availability, and privacy. Ensuring data integrity involves safeguarding against unauthorized deletion, modification, or fabrication of data. Data confidentiality poses challenges in protecting against both insider threats and external breaches, particularly concerning the storage of sensitive data like medical records or government files [33]. Data availability refers to the ability to recover data, while data privacy pertains to the selective sharing of information among individuals or groups.

Despite these apprehensions, [6] argues that with proper preparation, cloud computing may potentially offer enhanced security compared to traditional methods. The authors assert that obstacles encountered in cloud computing are not unique and can be addressed using existing technologies such as data encryption, virtual local area networking, firewalls, and more.

Legal Considerations:

Legal concerns, although closely intertwined with security issues, represent a separate and discernible challenge [15, 27]. Regulations and laws governing data storage vary significantly depending on geographical location [6]. While service level agreements (SLAs) between providers and consumers have been established, the absence of standardized regulations poses a notable challenge. Adapting cloud computing storage practices to align with local laws may present future challenges, necessitating meticulous planning and coordination during system design.

Data Management:

Another significant concern revolves around the fate of data stored in cloud computing systems in the event of catastrophic data loss. For instance, if a cloud provider were to declare bankruptcy, the potential loss of data could prove insurmountable. This scenario poses a considerable dilemma for large organizations heavily reliant on data storage [15, 33].

Interoperability:

Interoperability, defined as the ability of diverse systems and organizations to collaborate seamlessly, is identified as a critical challenge by [23], ranking second only to data security and trust issues in cloud computing systems. Achieving interoperability is paramount as it enables users to evade vendor lock-in, wherein dependency on a single Cloud Service Provider (CSP) inhibits the migration of services to alternative platforms or clouds. The absence of standardized protocols complicates data transfer between disparate cloud systems.

Latency Challenges:

Cloud computing grapples with the issue of high latency, characterized by delays between initiating data transfer and its actual commencement. In cloud computing, this delay arises from the necessity of communication among various nodes within the cloud infrastructure. Two potential solutions have emerged to address this challenge: fog computing and edge computing.

Fog computing involves a scenario where a multitude of heterogeneous and decentralized devices, including wireless and autonomous ones, communicate and potentially collaborate among themselves and with the network to execute storage and processing tasks independently of third-party intervention [31]. Another definition proposed by [35] describes fog computing as a geographically distributed computing architecture comprising one or more ubiquitously connected heterogeneous devices situated at the network's edge, complemented by cloud services. This network collectively delivers elastic computation, storage, and communication within isolated environments to numerous nearby clients.

To mitigate latency issues, [17] proposed a fog-based opportunistic spatio-temporal event processing system, which anticipates the future query region for mobile consumers. It is noteworthy that several types of fog networks exist.

Edge computing, on the other hand, shares similarities with fog computing but focuses on distributed computing paradigms where computation predominantly occurs on edge devices such as smartphones and laptops, rather than solely relying on the central cloud network. Computation tasks are shifted towards the periphery of the system to enhance efficiency and reduce latency.

Vendor Lock-in:

The issue of vendor lock-in, extensively discussed by numerous authors, describes a scenario where clients become reliant on a particular provider, making the transition to alternative providers costly due to legal complications or technical incompatibilities. This challenge is exacerbated by interoperability issues arising from the existence of multiple incompatible cloud systems [21]. The lack of standardization across operating platforms, Application Programming Interfaces (APIs), Service Level Agreements (SLAs), and cloud semantics contributes significantly to interoperability concerns, hindering resource migration between cloud providers [21]. Consequently, companies may find themselves dependent on a single provider, raising concerns about data security and continuity in the event of provider bankruptcy [20].

While some businesses have embraced cloud technology, the vendor lock-in problem remains a significant barrier to adoption for others, driven by a lack of trust in single Cloud Service Providers (CSPs). For instance, if a business contracts a CSP that subsequently goes bankrupt, the risk of losing all data stored on that cloud is substantial.

Various solutions have been proposed to address this challenge. Notably, recent literature suggests standardization efforts in edge computing as a potential remedy [28]. Initiatives like the EdgeX Foundry project, launched in 2017, aim to establish a vendor-neutral framework for Internet of Things (IoT) computing. Similarly, the Open-Fog consortium, a collaboration between Princeton University and leading tech companies such as Dell and Microsoft, seeks to standardize fog computing. Additionally, projects like MELODIC (Multi-cloud Execution-ware for Large-scale Optimized Data-Intensive Computing) [11, 12] aim to provide vendor-independent middleware for cloud applications. These efforts strive to mitigate vendor lock-in by promoting interoperability and standardization across cloud platforms.

Multi-Cloud Solutions

In response to the challenges highlighted earlier, the evolution of cloud computing has led to the emergence of multi-cloud computing—a paradigm where organizations leverage multiple cloud networks and services simultaneously [13]. Simply put, multi-cloud systems utilize more than one Cloud Service Provider (CSP), although they encompass various subcategories, as outlined below. This concept bears resemblance to cross-cloud architectures, defined as systems that span across multiple provisioning boundaries [14]. While the precise differentiation between multi-clouds and cross-clouds remains somewhat nebulous (explained in greater detail below), it is generally understood that in multi-cloud environments, users or businesses leverage distinct cloud services for different applications within their operations. For instance, they might opt to store data on a private cloud, share documents via the Google Cloud platform, and conduct data analysis on yet another cloud. Conversely, cross-cloud architectures aim to facilitate seamless data transfer and application utilization across multiple clouds, ensuring greater coherence [14].

The adoption of multi-cloud or cross-cloud approaches stems from various considerations, many of which aim to address the challenges highlighted in the preceding section.

According to [24], multi-cloud computing directly addresses 10 key issues, some of which have been mentioned above. Furthermore, it is important to realize that these issues are both shared and specific to the parties involved, including individual customers, company-level customers, and the Cloud Service Providers (CSPs). The 10 key issues addressed by multi-cloud computing are as follows:

1. Managing peaks in service/resource requests by leveraging external resources on demand.
2. Optimizing costs or enhancing the quality of services.
3. Responding to changes in offers by the providers.
4. Adhering to constraints such as new locations or regulations.
5. Ensuring high availability of resources and services.
6. Avoiding dependence on a single external provider.
7. Ensuring backups to handle disasters or scheduled inactivity.
8. Serving as an intermediary.
9. Enhancing own Cloud service/resources offers based on agreements with others.
10. Utilizing different services for their specific features not available elsewhere.

For instance, if a company provides an online service and overestimates user engagement, scaling down by shifting from lifelong virtual machines to ephemeral ones, which charge by the minute, can reduce costs. Similarly, consolidating teams within a company may require significant overhauls to the underlying logic if different branches or teams rely on different cloud infrastructures.

Another advantage of multi-cloud computing is avoiding long-term commitment to a single CSP, mitigating interoperability and vendor lock-in issues. By facilitating communication between different cloud components, multi-cloud platforms may introduce new methods of operability, either through increased standardization or by devising new ways for clouds to share data universally. Developing technologies enabling service transfer to other CSPs would enhance flexibility within businesses.

It should be noted that ambiguity exists in the literature regarding cloud computing models. Some authors, like [13], categorize multi-clouds, hybrid clouds, and federation clouds under cross-cloud computing, while [24] distinguishes them under "multiple cloud" computation. For clarity, this document treats hybrid clouds, multi-clouds, and federation clouds as distinct methods. Future research may benefit from concrete definitions akin to those provided by NIST for cloud computing as a whole [25].

Multi-Cloud

Multi-cloud refers to cloud systems where applications are hosted across a heterogeneous network of different cloud providers. Unlike hybrid clouds, where various deployment methods are used, multi-cloud systems consist of unique cloud services utilized for distinct roles within a single organization. This approach reduces the trust requirements on any single Cloud Service Provider (CSP) by distributing responsibilities among multiple providers .

In a multi-cloud environment, a business or organization uses different CSPs for various needs, each with differing levels of application and Service Level Agreements (SLAs). For example, a company might use one cloud service for document storage and sharing, another private cloud for sensitive company data, and a third for data analysis. Each of these clouds may offer different SLAs, costs, and degrees of utilization by the company.

The motivations for adopting a multi-cloud strategy are diverse. Legal considerations may necessitate the use of different clouds to comply with data storage regulations across jurisdictions. Avoiding dependency on a single CSP is another significant reason. For instance, [9] proposes an architectural framework for Programmable Network Clouds hosting Software Defined Networking (SDN) and Network Function Virtualization (NFV) across geographically distributed multi-cloud environments. This framework focuses on cost and SLA-aware resource provisioning and scheduling to minimize operating costs while adhering to SLAs. Future research suggested by the authors includes optimization techniques that simultaneously optimize VM/container placement and traffic consolidation, enhancing SLA satisfaction and reducing costs.

One of the most valuable and well-researched applications of multi-cloud architecture is enhanced security. Studies such as those by [5, 7, 18, 30] describe multi-cloud architectures with a primary focus on security, particularly for

sensitive data like medical records. This approach ensures that security measures are distributed and reinforced across multiple platforms, reducing the risk of data breaches and improving overall data protection.

Enhanced Security Mechanisms in Multi-Cloud Environments

This paper discusses four mechanisms designed to enhance security in multi-cloud environments. The first mechanism is application imitation, where different clouds within the multi-cloud system "double-check" each other. By comparing results from different clouds, data integrity is ensured. The second mechanism, layer-wise application partition, involves separating the application's logic from its data. This separation protects against logic flaws, ensuring that the application provider cannot reconstruct the entire logic from user data unless the execution occurs entirely on the user's system.

Partitioning into segments is the third mechanism, where the application is divided into smaller segments, each run on different clouds. This process involves two phases: in the first phase, trusted private clouds handle small computational parts, while untrusted public clouds manage higher loads. In the second phase, the computation is distributed among the untrusted clouds. The final mechanism is distribution of chunks, where data is split among different clouds within the multi-cloud system. Various cryptographic data segmentation methods have been trialed for secure data storage and retrieval.

Several models have been proposed for these mechanisms. For instance, a cooperative provable data possession (CPDP) scheme was tested by [36], but it had a security flaw where the parameter π' could be forged, bypassing authentication . More recently, integrity checking methods using the Co-Check scheme, based on Boneh-Lynn-Shacham (BLS) signatures and homomorphic tags, have been explored . Homomorphism, as defined by [8], is a property allowing a problem in one algebraic system to be converted, solved in another system, and converted back, enabling third parties to view and use encrypted information if they have the right tools. Homomorphic encryption allows operations or calculations on encrypted data without needing to decrypt it first.

In the Co-Check framework, users do not need to retrieve the entire data file during challenge-response and integrity checking stages. Instead, users generate challenges for audits using parts of the metadata stored on the client side, enhancing audit efficiency and ensuring that malicious CSPs cannot bypass the check . Other authors have also explored homomorphic encryption tools for multi-cloud security. Future research directions for multi-cloud models will likely include experimenting with security mechanisms, as homomorphic encryption is still in its infancy regarding cloud computing applications.

Hybrid Cloud

Hybrid clouds are a prevalent type of multi-cloud framework, combining private cloud deployment methods with one or more public clouds. Unlike multi-clouds, hybrid clouds integrate private and public clouds into a unified system rather than using entirely distinct clouds. This approach is particularly useful for businesses that need to share some data with various user types while keeping other elements confidential. For instance, hybrid clouds are being explored in the medical field, where the availability of big data necessitates secure and efficient solutions. A recent study by [19] examined security solutions for healthcare data using a hybrid approach that combines linear network coding and re-encryption based on ElGamal cryptography. The authors implemented a linear network

coding mechanism and employed the ElGamal re-encryption scheme to securely encode the key matrix, enhancing security and fault tolerance for cloud storage.

A significant difference between hybrid and multi-clouds lies in the nature of the tasks performed. Multi-clouds, using distinct clouds, can handle a variety of tasks and services for a company. Conversely, hybrid clouds, with their unified framework, are typically employed to manage a single task, such as securely storing and distributing medical records.

Federated Cloud Computing

Federated cloud computing is characterized by CSPs agreeing to share resources under a single federation. This system allows CSPs to share client loads, providing clients with access to a comprehensive catalog of services and resources. This setup enhances interoperability and application portability. An example of federated cloud computing is the EU-based EGI Federated Cloud, which unites more than 20 cloud providers and 300 data centers. Federated clouds address the vendor lock-in problem by facilitating easier migration of applications and data from one cloud to another, thanks to the underpinning SLAs. This model is particularly beneficial for smaller providers, as they can compensate for each other's weaknesses by pooling resources and capabilities.

Cross-Cloud Computing

According to [14], a cross-cloud application utilizes more than one cloud API within a single version of the application. Cross-cloud architecture can involve either dynamically interchangeable APIs or non-dynamically interchangeable APIs, provided the APIs are distinct. The literature on cross-cloud systems is less extensive compared to other frameworks, potentially due to the lack of a clear definition. In the [14] paper, the author categorized hybrid clouds, multi-clouds, and federated clouds as subcategories of cross-cloud computing, rather than multi-cloud computing, which raises questions about the precise distinction between these terms and what exactly constitutes a "cross-cloud."

Despite this ambiguity, recent work proposes various cross-cloud architectures for different purposes, sometimes merging the concept of cross-cloud systems with other multi-cloud models. A recurring idea in cross-cloud computing is the notion of the "cloud broker"—an entity that manages the use, performance, and delivery of cloud services and negotiates relationships between CSPs and consumers [25]. For instance, [2] proposed a broker-based cross-cloud federation manager as a business model for federated cloud networks. More recently, cross-cloud computing has been highlighted as a potential solution for managing Big Data in the Internet of Things (IoT) [29].

Research Directions and Considerations for the Future

In the previous sections, we discussed how multi-cloud computing frameworks address several challenges in cloud computing, such as data security and vendor lock-in avoidance. Beyond these issues, multi-cloud computing networks have numerous other applications. Two major themes currently being explored in the literature are cloud

computing and Big Data, as well as Machine Learning. Below, we will describe some of the challenges and current projects in these fields.

Big Data

Data management has always been a critical issue in cloud computing, and the need to process vast quantities of data is a major focus of current research. One significant application of Big Data is in the medical field, where enormous amounts of patient information must be stored and processed across large cloud computing networks. This issue is further complicated by the introduction of medical wearables, which gather and transmit even more data to cloud networks, decentralizing data sources.

The interconnectivity between household electronics, smart devices, computers, vehicle monitoring equipment, and other data-collecting devices is referred to as the Internet of Things (IoT). The IoT presents a significant challenge for data analysis because the data comes from various sources. Traditional analytic methods cannot keep up with the sheer volume of existing and generated data . Efficient processing, horizontal and vertical analysis, and data transfer between different cloud systems are essential. For example, in the case of medical records, a centralized system might work well within a nationalized healthcare system, but it must also accommodate peripheral services, such as private practices, which might use different cloud systems. Private firms need the capability to integrate medical information into larger, mainstream sources.

One proposed solution to the Big Data storage problem is the "rain cloud," a multi-cloud model where each member cloud has an SLA with other member clouds to collaborate when data becomes too large for any single cloud to handle . Hybrid clouds are also considered useful for Big Data management because the combination of private and public deployments allows for confidentiality when different parties require varying levels of data access. Businesses might need parts of their data and infrastructure to remain behind the corporate firewall due to industry standards, legal regulations, or a desire for privacy .

Other approaches involve distributing data across multiple clouds for rapid storage. For example, Hadoop, an open-source implementation of Google's data storage method, breaks large data volumes into manageable chunks distributed across thousands of computers. A parallelized programming API is then used to distribute computations to where the data is located and aggregate the results. This method is highly applicable to bioinformatics and genomics, although technical expertise and the lack of parallel-running bioinformatics tools have been barriers until recently . Researchers have reviewed various Big Data cloud computing frameworks, including Hadoop, Spark, and Flink, and it will be interesting to see how multi-cloud computing architectures will integrate these approaches. Additionally, Amazon Elastic MapReduce, IBM BigInsights, and Microsoft Azure HDInsight are mentioned as methods for facilitating large-scale data processing frameworks, such as Apache Hadoop and Apache Spark .

Machine Learning

Machine learning, while related to Big Data, presents its own unique set of challenges and applications within cloud computing. The vast amounts of data stored in clouds, such as photographs and videos, provide a rich source of information that can be processed by machine learning algorithms. For instance, virtual machine optimization techniques for healthcare services are being trialed and show promise, but they need to be tested with other diseases

to ensure consistency . In the oil and gas industry, researchers have explored using hybrid cloud storage combined with machine learning to handle technical documents, which often contain valuable information from geoscience and engineering disciplines and are typically stored in an unstructured format . To enhance data extraction and utilization, a machine-learning-enabled platform has been proposed. This platform employs a sequence of algorithms developed as a hybrid cloud container, automatically reading and understanding technical documents with minimal human supervision. Users can upload raw data to the platform, which is stored on a private local server. The platform then generates structured data, which is pushed to a search engine accessible via the cloud. This system allows users to quickly identify important parts of technical documents, automate the extraction of relevant data, present it meaningfully for further analysis, and easily share and port it to other platforms.

Future directions for machine learning and multi-cloud computing research include a shift towards widespread adoption of auto-tuners, particularly for the SaaS layer of the cloud. Researchers also anticipate the development of new automated tools that allow cloud users to benefit from the experiences of others through partially automated application builders, automated database sharers, query optimizers, smart load balancers, and service replicators . Essentially, the interface between the cloud, applications, and users is expected to become more streamlined and intuitive based on user experiences. Security, a critical concern for any cloud network, could also benefit from novel machine learning methods, although the specific nature of these methods and their implementation remain an area for future research.

Conclusion

Cloud computing is a rapidly evolving technology with diverse applications across various industries, particularly in remote computing and storage. Major concerns such as vendor lock-in and cybersecurity can potentially be mitigated by employing hybrid clouds, multi-clouds, and federated clouds. These models offer users alternatives during scheduled maintenance, security breaches, or service shutdowns, each with distinct advantages and disadvantages. Hybrid clouds are highly customizable for specific applications but are less transferable and typically used for singular tasks. In contrast, multi-clouds and federated clouds are better suited for businesses that require multiple tasks or services.

Future research should focus on integrating multi-cloud paradigms with other emerging technologies like machine learning and big data. These integrations could address existing challenges in cloud computing, such as enhancing security, and pave the way for new analytical methods and applications.

References List:

[1]. Malaiyappan, J. N. A., Karamthulla, M. J., & Tadimarri, A. (2023). Towards Autonomous Infrastructure Management: A Survey of AI-driven Approaches in Platform Engineering. Journal of Knowledge Learning and Science Technology ISSN: 2959-6386 (online), 2(2), 303-314.

[2]. Talati, D. (2023). AI in healthcare domain. Journal of Knowledge Learning and Science Technology ISSN: 2959-6386 (online), 2(3), 256-262.

[3]. Talati, D. (2023). Telemedicine and AI in Remote Patient Monitoring. Journal of Knowledge Learning and Science Technology ISSN: 2959-6386 (online), 2(3), 254-255.

[4]. Talati, D. (2024). Virtual Health Assistance–AI-Based. Authorea Preprints.

[5]. Talati, D. (2023). Artificial Intelligence (Ai) In Mental Health Diagnosis and Treatment. Journal of Knowledge Learning and Science Technology ISSN: 2959-6386 (online), 2(3), 251-253.

[6]. Talati, D. (2024). Ethics of AI (Artificial Intelligence). Authorea Preprints.

[7]. Talati, D. V. AI Integration with Electronic Health Records (EHR): A Synergistic Approach to Healthcare Informatics December, 2023.

[8]. Singla, A., & Malhotra, T. (2024). Challenges And Opportunities in Scaling AI/ML Pipelines. Journal of Science & Technology, 5(1), 1-21.

[9]. Singla, A., & Chavalmane, S. (2023). Automating Model Deployment: From Training to Production. Journal of Knowledge Learning and Science Technology ISSN: 2959-6386 (online), 2(3), 340-347.

[10]. Gehrmann, S., & Rončević, I. (2015). Monolingualisation of research and science as a hegemonial project: European perspectives and Anglophone realities. Filologija, (65), 13-44.

[11]. Roncevic, I. (2021). Eye-tracking in second language reading. Eye, 15(5).

[12]. Šola, H. M., Gajdoš Kljusurić, J., & Rončević, I. (2022). The impact of bio-label on the decision-making behavior. Frontiers in sustainable food systems, 6, 1002521.

[13]. Sirigineedi, S. S., Soni, J., & Upadhyay, H. (2020, March). Learning-based models to detect runtime phishing activities using URLs. In Proceedings of the 2020 4th international conference on compute and data analysis (pp. 102-106).

[14]. Verma, V., Bian, L., Ozecik, D., Sirigineedi, S. S., & Leon, A. (2021). Internet-enabled remotely controlled architecture to release water from storage units. In World Environmental and Water Resources Congress 2021 (pp. 586-592).

[15]. Soni, J., Sirigineedi, S., Vutukuru, K. S., Sirigineedi, S. C., Prabakar, N., & Upadhyay, H. (2023). Learning-Based Model for Phishing Attack Detection. In Artificial Intelligence in Cyber Security: Theories and Applications (pp. 113-124). Cham: Springer International Publishing.

[16]. Verma, V., Vutukuru, K. S., Divvela, S. S., & Sirigineedi, S. S. (2022). Internet of things and machine learning application for a remotely operated wetland siphon system during hurricanes. In Water Resources Management and Sustainability (pp. 443-462). Singapore: Springer Nature Singapore.

[17]. Soni, J., Gangwani, P., Sirigineedi, S., Joshi, S., Prabakar, N., Upadhyay, H., & Kulkarni, S. A. (2023). Deep Learning Approach for Detection of Fraudulent Credit Card Transactions. In Artificial Intelligence in Cyber Security: Theories and Applications (pp. 125-138). Cham: Springer International Publishing.

[18]. Biswas, A., & Talukdar, W. (2024). Intelligent Clinical Documentation: Harnessing Generative AI for Patient-Centric Clinical Note Generation. arXiv preprint arXiv:2405.18346.

[19]. Talukdar, W., & Biswas, A. (2024). Synergizing Unsupervised and Supervised Learning: A Hybrid Approach for Accurate Natural Language Task Modeling. arXiv preprint arXiv:2406.01096.

[20]. Karamthulla, M. J., Malaiyappan, J. N. A., & Tillu, R. (2023). Optimizing Resource Allocation in Cloud Infrastructure through AI Automation: A Comparative Study. Journal of Knowledge Learning and Science Technology ISSN: 2959-6386 (online), 2(2), 315-326.

[21]. Tembhekar, P., Malaiyappan, J. N. A., & Shanmugam, L. (2023). Cross-Domain Applications of MLOps: From Healthcare to Finance. Journal of Knowledge Learning and Science Technology ISSN: 2959-6386 (online), 2(3), 581-598.

[22]. Talati, D. (2024). AI (Artificial Intelligence) in Daily Life. Authorea Preprints.