International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Tackling the Challenges of Distributed Data Management in Cloud Computing - A Review of Approaches and Solutions

Ali Hikmat Ibrahem[1*], Subhi R. M. Zeebaree[2]

Submitted: 12/12/2023 **Revised**: 23/01/2024 **Accepted**: 01/02/2024

Abstract: Cloud computing is a contemporary endeavour to provide computing resources, such as hardware or software, as a service across a network. Cloud computing is a current IT trend that involves shifting computation and data storage from desktop and portable PCs to enormous data centre's that can store massive amounts of information, measured in peta-bytes. Cloud Computing encompasses multiple facets, including availability, scalability, virtualization, interoperability, quality of service, and the delivery types of the cloud, which are private, public, and hybrid. Cloud databases are mostly utilized for data-intensive applications, such as data warehousing, data mining, and business intelligence. A cloud database is necessary to efficiently accelerate the process of reducing the burdens associated with routing configuration. This article conducts a thorough analysis of challenges in cloud computing data management, focusing on dimensions such as consistency, scalability, security, interoperability, migration, and latency. Scholarly investigations address distributed databases, consensus algorithms, encryption, access control, auditing, and the development of harmonious ecosystems for diverse cloud environments. Emphasis is placed on automated migration tools, best practices, and methodologies for smooth transitions, as well as innovative solutions for minimizing latency in real-time applications. The overarching goal is to advance data confidentiality, integrity, system security, and long-term advancements in cloud computing.

Keywords: *Data Management, Storage, cloud computing, challenges, consistency.*

1. Introduction

The exponential surge in data generation over the past decade has outpaced the capabilities of conventional institutional data management. In response to this challenge, cloud-based data management has emerged as a transformative solution, facilitating efficient scalability of resources. This approach involves institutions renting storage and computing power, alleviating the need for substantial in-house capital investments in infrastructure. Despite the initially touted economic benefits by public cloud providers, the adoption of cloud-based systems has been restrained, with apprehensions revolving around long-term costs, security, control, and legal safeguards for data and services. Cloud computing, as a transformative innovation, encompasses not only technological advancements but also signifies a paradigm shift in the business landscape of computing. Its evolutionary trajectory can be traced from the era of large mainframes through the age of personal computers to the present landscape of compact personal devices. Beyond a mere technological progression, cloud computing introduces a novel approach to Information Technology (IT) provisioning, holding the potential to revolutionize how computing resources and applications are deployed. This innovation disrupts traditional value chains, carving out space for innovative business models [1].

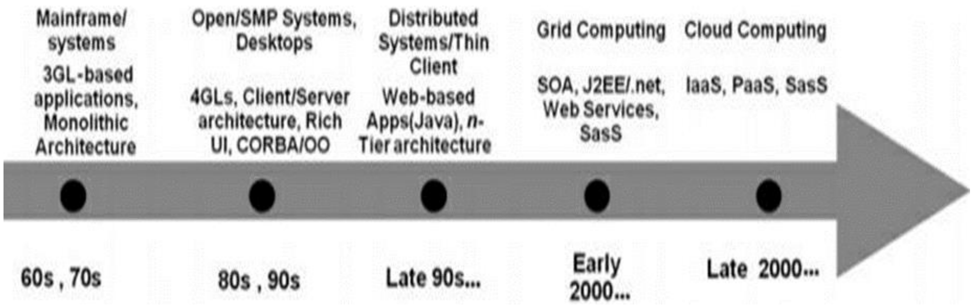

Fig (1): Evolution of computing Architecture.

[1]*Information Technology Dept., Technical College of Informatics-Akre, Akre University of Applied Science, Duhok, Iraq. ali.hikmat@auas.edu.krd*
[2]*Energy Eng. Dept., Technical College of Engineering, Duhok Polytechnic University, Duhok, Iraq, subhi.rafeeq@dpu.edu.krd*

Key industry players such as IBM, Amazon, Google, Microsoft, Salesforce, and Sun have solidified their positions as leaders in the cloud computing market, providing robust platforms and infrastructure services. Despite its evident business potential, cloud computing remains somewhat elusive and ambiguous to many. The transition to this new paradigm has already commenced, with enterprises adopting a phased "crawl, walk, run" approach, gradually progressing towards comprehensive cloud implementations [2].

The multifaceted landscape of cloud-based data management is marked by intricate challenges and opportunities. Institutions grapple with concerns encompassing costs, security, and control, as the narrative of cloud computing continually transforms IT provisioning and business models. Within the realm of cloud-based data management, systems navigate the norm of component failures on clusters of storage nodes, demanding robust fault tolerance and recovery mechanisms. Portability across diverse hardware and software platforms becomes imperative, while the management of terabytes of data necessitates optimization of I/O operations and block sizes for large-scale datasets. A Cloud Database Management System (DBMS) positioned to tackle these challenges must exhibit elasticity to accommodate varying workloads, scalability to address expanding data volumes, fault tolerance for unwavering operation, and self-manageability for streamlined administration. Google Cloud Platform (GCP) emerges as a frontrunner in data management, presenting tools that facilitate seamless cloud-based data access, processing, and sharing. This comprehensive overview underscores key facets of GCP's prowess in navigating the complexities of contemporary data management [3].

2. Background Theory

2.1. Public Datasets

GCP's Public Datasets open the doors to a vast reservoir of diverse datasets hosted in BigQuery and Cloud Storage. With a repertoire exceeding 100 datasets cutting across various industries, it emerges as a dynamic playground for researchers. The allure lies not just in the breadth of datasets but in the simplicity, it affords, coupled with complimentary storage and a generous monthly data access limit of up to 1TB, all provided at no cost. Accessibility is a key feature, facilitated through various methods including the user-friendly Google Cloud Console and the versatile gsutil. These tools empower researchers to navigate and utilize datasets with ease, fostering a collaborative environment where knowledge can be shared and research endeavours can be undertaken collectively. This initiative not only democratizes access to valuable datasets but also serves as a catalyst for impactful research. By removing barriers such as cost and storage constraints, GCP's Public Datasets contribute to the acceleration of knowledge discovery and innovation within the research community. It stands as a testament to GCP's commitment to advancing collaborative research through accessible and powerful data resources [4].

2.2. Data Storage

GCP's Cloud Storage emerges as a cornerstone, providing globally accessible and exceptionally durable object storage that seamlessly scales to exabytes. The introduction of storage classes—Standard, Nearline, Coldline, and Archive—reflects a nuanced approach, catering to diverse data needs. Delving into the intricacies of each class becomes pivotal in optimizing costs based on specific data access requirements [5]. The robust access control mechanism through Cloud Storage Buckets, complemented by the strategic implementation of Object Lifecycle Management, plays a crucial role in ensuring long-term cost efficiency. This dynamic combination empowers users to tailor storage solutions to the lifecycle of their data, aligning access levels and costs with the evolving nature of information. In essence, GCP's Cloud Storage not only provides a scalable and durable foundation for data but also offers a sophisticated toolbox for users to finely tune their storage strategy, balancing accessibility, durability, and cost-effectiveness in tandem[6].

2.3. Data Discovery

In the expansive landscape of GCP's Big Data solutions, the architecture facilitates not just the efficient handling but the robust capture and analysis of data at a scale that reaches into the petabytes. The role played by the Cloud Data Catalog becomes even more pivotal as a centralized metadata service, orchestrating a systematic approach to discovering and understanding data assets across the platform. BigQuery, standing as a fully managed analytics warehouse, offers a user-friendly avenue for navigating the vast sea of data through familiar SQL queries. Meanwhile, Cloud Life Sciences introduces advanced genomics processing tools, marking a significant stride in the fusion of cutting-edge technology with cloud-based data management practices. Within the contours of this research initiative, our lens focuses sharply on unravelling the intricate challenges woven into the fabric of data management within cloud computing ecosystems. Our ambition is not merely to understand but to extract meaningful insights, offering innovative solutions that stand to enhance the efficiency and fortify the security of data within this dynamic and ever-evolving landscape. This journey aims to contribute not just to the academic discourse but to the practical implementations and advancements within the realm of cloud-based data management [4].

2.4. Security

The primary security challenge within the realm of cloud computing research is intricately tied to the imperative of ensuring data confidentiality, integrity, and availability in an environment characterized by dynamism, distribution, and multi-tenancy Within the academic sphere of cloud computing research, a central and intricate security challenge revolves around ensuring the confidentiality, integrity, and availability of data. This challenge unfolds within an environment characterized by the dynamic nature of cloud systems, their widespread distribution, and the shared tenancy model. At its core, the fundamental concern arises from the inherent trade-off associated with relinquishing control over data when stored remotely in the cloud. Despite the compelling advantages of cloud computing, such as on-demand self-service and cost efficiency, it introduces a spectrum of security threats to the data outsourced by users. The centralization or outsourcing of data to the cloud translates into users ceding ultimate control, exposing them to risks ranging from potential data integrity issues due to internal and external threats to the spectre of outages and security breaches [6].

Traditional cryptographic primitives encounter significant challenges in ensuring data security within this context. The impracticality of downloading all data for integrity verification, due to elevated I/O and transmission costs, compounds the issue. Moreover, the sheer volume of outsourced data renders auditing tasks both formidable and expensive for researchers and academics. In response to these challenges, the necessity for public auditing services becomes apparent, allowing researchers to engage third-party auditors for periodic checks on data integrity in a more accessible and cost-effective manner. Privacy concerns introduce another layer of complexity into this intricate landscape. Existing public auditability schemes often overlook the imperative of safeguarding users' data privacy against external auditors. The potential leakage of user data information during the auditing process emerges as a formidable security risk. Encryption, while a critical component, proves insufficient in addressing this issue, underscoring the need for a privacy-preserving third-party auditing protocol. The proposed solution takes a pioneering stride by integrating homomorphic linear authenticators with random masking, offering a pathway to privacy-preserving public auditing. This innovative approach ensures that auditors cannot derive user data content even when armed with knowledge of linear combinations of the same data blocks. The incorporation of a public key-based homomorphic linear authenticator not only supports public auditability but also positions the protocol to be lightweight and efficient in its design. This comprehensive exploration, rooted firmly in the academic context of cloud computing research, delves into the depth of challenges and underscores promising avenues for innovative solutions within this critical domain. [7]

Amidst the dynamic evolution of cloud computing, the imperative of bolstering security has risen to the forefront, with cryptography standing as a pivotal solution. Robust encryption schemes are deployed to uphold the principles of confidentiality, integrity, and availability (CIA) in this ever-changing cloud environment. Cryptography assumes a central role in safeguarding shared storage by implementing stringent access controls and ensuring secure data backup practices. In the realm of academic research, this cryptographic approach significantly contributes to the discourse on cloud security, offering insights into its nuanced applications and highlighting its foundational role in establishing a trusted computing environment. The integration of cryptography not only addresses current security challenges but also proactively anticipates and adapts to the evolving threats within the digital ecosystem [8]. Within the scholarly investigation into cloud computing security, an additional facet deserving meticulous scrutiny pertains to the customary framework governing backup and recovery practices. Situated within the intricacies of a multi-tenant environment, a nexus of inquiries arises, encompassing the periodicity of backup procedures, custodial responsibility for storage, and the harmonization of such methodologies with the idiosyncratic requisites delineated by individual clientele. The imperative of ensuring judicious preservation of logical data segregation during the orchestration of backup processes assumes paramount importance in this milieu. This necessitates an unwavering commitment to the incorporation of encryption mechanisms within backup media, elucidating its criticality as an indispensable facet instrumental in augmenting the fortification of comprehensive security postures within the cloud computing paradigm. As the discourse unfolds, it becomes evident that the confluence of these factors not only accentuates the need for heightened diligence but also underscores the nuanced intricacies integral to safeguarding the integrity and confidentiality of data within the multifaceted landscape of cloud computing [9].

Within the expansive domain of the security landscape in cloud computing, another discernible facet emerges in the realm of data format standards. Acknowledged as pivotal elements for safeguarding data availability, these standards introduce an additional stratum of intricacy to the ongoing discourse on security. Despite commendable endeavours, such as the Cloud Computing Effort orchestrated by DMTF, striving to grapple with this exigency, the absence of a universally embraced data format standard poses a formidable impediment. This hinders the seamless exchange of data across heterogeneous cloud service providers. Consequently, the challenge at hand resides in the establishment of standardized formats that not only facilitate the fluid portability of data but also foster interoperability. This approach addresses the inherent complexities associated with diverse cloud environments [10]. In summation, the security challenge in cloud computing research is a multifaceted endeavour that necessitates adept navigation through the complexities inherent in guaranteeing the confidentiality, integrity, and availability of data. Whether grappling with cryptographic intricacies, devising effective backup and recovery strategies in a multi-tenant landscape, or striving for standardized data formats, the pursuit of robust security in cloud computing research demands a comprehensive understanding and strategic approach. Each facet contributes uniquely to the broader goal of fortifying the cloud environment against emerging threats and ensuring the resilience of data in the digital landscape [11].

2.5. Scalability and Consistency in Distributed Database Systems

2.5.1. Scalability in Distributed Database Systems:

Scalability within distributed database systems holds a pivotal role in efficiently managing the escalating demands of augmented data volumes, user loads, or transaction rates while concurrently preserving or enhancing overall performance. This critical aspect becomes particularly evident in contemporary systems like Google's Bigtable [12], PNUTS [13], and Amazon's Dynamo where the implementation of strategic design decisions significantly contributes to their scalability prowess. The adoption of a key-value pair data model is a foundational choice that streamlines data access. By restricting operations to single keys, these systems enhance simplicity in transactions, a key factor in facilitating scalability. Horizontal partitioning is another strategic decision that enables the effective distribution of data across multiple servers. This not only prevents bottlenecks but also ensures seamless accommodation of growing datasets [14]. A noteworthy design decision involves capitalizing on the loose coupling of data in applications featuring single-object semantics. This enhances the adaptability of these systems, allowing for the independent scaling of different components. Consequently, increased workloads or data volumes can be accommodated without compromising the overall performance, marking a significant achievement in the evolution of distributed database architecture [15,16].

2.5.2. Consistency in Distributed Database Systems:

Consistency within distributed database systems is a paramount goal, aiming to ensure the accuracy and reliability of data across diverse nodes within the system [17]. To address the intricate challenges associated with consistency, contemporary distributed systems deploy an array of sophisticated mechanisms. One prevalent strategy involves adopting a key-value pair data model and restricting operations to single keys. This approach enables manageable consistency at the granularity of individual items, effectively mitigating the complexities inherent in distributed environments. In the pursuit of high availability, these systems frequently turn to replication as a fundamental tactic. Replicating data across multiple nodes enhances fault tolerance and guarantees continuous access to information, even in the event of node failures. Furthermore, the arena of consistency management witnesses the implementation of various strategies, including eventual consistency and timeline consistency [18]. These approaches aim to synchronize replicas, albeit with varying degrees of stringency, thereby establishing distinct levels of data consistency [19]. The nuanced interplay of these mechanisms not only addresses the challenges associated with maintaining data consistency in distributed systems but also underscores the adaptability and sophistication of contemporary approaches. The incorporation of diverse strategies reflects the ongoing efforts within the field to strike a balance between ensuring high availability and upholding the accuracy and reliability of data across the distributed landscape. In the dynamic realm of distributed database systems, the historical landscape grappled with notable performance challenges arising from partial failures and the inherent overhead associated with synchronization [20, 21]. However, a transformative shift occurred with the advent of contemporary and scalable systems, exemplified by the likes of Google's Bigtable, PNUTS, and Amazon's Dynamo. These innovative solutions strategically addressed the aforementioned challenges by embracing key-value pair data models, where operations on single keys were rendered both consistent and atomic [22, 23].

This departure from traditional approaches marked a paradigmatic evolution, as these modern systems not only prioritized scalability but also introduced novel methodologies to enhance performance. The adoption of key-value pair data models facilitated streamlined and efficient operations on individual keys, effectively mitigating the complexities associated with managing distributed datasets. This architectural shift played a pivotal role in addressing the intricacies of partial failures, ushering in a new era characterized by improved fault tolerance and reliability in distributed databases [24]. The confluence of these advancements represents a significant milestone in the ongoing evolution of distributed database systems, offering a more robust and scalable foundation for contemporary data management challenges. Google's Bigtable, serving as a trailblazer in this domain, laid the foundational groundwork for these significant advancements by introducing a pioneering distributed storage system explicitly designed to manage vast amounts of data across clusters. This innovative approach marked a departure from traditional paradigms and set the stage for subsequent transformative developments [25]. Building upon Bigtable's breakthroughs, subsequent systems like PNUTS and Amazon's Dynamo expanded and refined the principles of key-value pair data models. The pivotal contribution of these systems lies in their ability to enable consistent and atomic operations at the granularity of single keys. This strategic move not only surmounted historical performance bottlenecks but also paved the way for the establishment of scalable, reliable, and high-performance distributed databases In essence, the evolution from historical distributed database systems to their contemporary counterparts represents a strategic and responsive approach to the challenges posed by partial failures and synchronization overhead. The transition to key-value pair data models, as exemplified by the notable contributions of Bigtable, PNUTS, and Dynamo, stands as a testament to the innovation required to meet the demands of a rapidly evolving digital landscape [26].

2.6. Interoperability

In the academic examination of cloud computing interoperability, the concept refers to the ability of diverse cloud systems to operate seamlessly. This includes the exchange of critical elements such as application formats, Service Level Agreement (SLA) templates, authentication and authorization token formats, and attribute data. The objective is to facilitate efficient collaboration among various cloud platforms and address challenges like vendor lock-ins. The focus is on enabling users to transition their cloud deployments between providers while maintaining compatibility with essential components and standards [27]. The challenges encountered in achieving interoperability within cloud computing span multiple levels. The complexities in transitioning from Infrastructure-as-a-Service (IaaS) to Platform-as-a-Service (PaaS) highlight the intricacies of different cloud-based environments. Concurrently, challenges at the provider-to-provider level shed light on the difficulties tenants face when migrating resources across distinct cloud service providers [28].

An organized categorization of interoperability challenges identifies four main levels: consensus on data formats and communication protocols, addressing service provider reluctance to adopt common standards, establishing horizontal or vertical interoperability across diverse deployment levels, and effectively managing interactions, considering real-time services or asynchronous patterns between services. These categories provide a systematic framework for understanding the multifaceted nature of interoperability challenges, emphasizing the need for cohesive standards and nuanced management strategies [29].

Interoperability challenges within cloud computing services encompass both technological and managerial dimensions, categorized into three types: differences in hypervisors, platforms, and policies in cloud management. The technological challenge arises from variations in hypervisors for virtualization, impacting cooperation between Infrastructure-as-a-Service (IaaS) providers [30]. The impracticality of standardizing the structure for all hypervisor virtual machines necessitates solutions such as restructuring virtual machines or employing interoperability standards like the Open Virtualization Format (OVF). Similarly, platform challenges manifest during application development in cloud environments, where disparate Application Programming Interfaces (APIs) hinder interoperability at both the application and platform layers. Initiatives like the Cloud Computing Interoperability Forum (CCIF) aim to standardize cloud interfaces and unify semantic cloud data models. The Unified Cloud Interface (UCI), facilitated by a cloud broker, mitigates these challenges by abstracting the usage of diverse cloud APIs, providing a unified interface through semantic web technologies. These solutions collectively address the intricate challenges associated with achieving interoperability in cloud computing services [31].

The discourse on standards in cloud computing is of paramount importance, particularly in the context of addressing interoperability challenges. Interoperability, within the cloud-computing community, is defined as the seamless movement of workloads and data between cloud providers or between private and public clouds [32]. The apprehension of vendor lock-in, where organizations face potential difficulties or significant costs when switching providers, serves as a notable barrier to cloud adoption. This concern is amplified by the associated risks of reduced negotiation power, encompassing potential price increases and service discontinuation. To tackle interoperability challenges, the utilization of open standards is a widely advocated strategy, endorsed by various sources, including the military, which promotes an open-standards approach to enhance cloud computing adoption [33]. The Open Cloud Manifesto outlines key principles, including the use of open standards, emphasizing collaboration within the industry. However, it is acknowledged that standards alone constitute just one facet of the cloud interoperability puzzle. Achieving interoperability may also necessitate sound architecture principles and dynamic negotiation between cloud providers and users. This report delves comprehensively into the role of standards in cloud-computing interoperability, aiming to provide insights into areas where standards can enhance interoperability. It also identifies areas where standards may require further development to offer increased value [34].

2.7. Migration

In the domain of cloud computing, migration encompasses the intricate process of transitioning diverse computing resources, spanning applications, data, and services, from an organizations on premises infrastructure to the cloud environment. This strategic shift leverages the capabilities of the cloud to provide flexible and scalable solutions accessible over the internet. The fundamental premise of cloud migration lies in optimizing the utilization of computing resources, encompassing both hardware and software, in a manner that is not only efficient but also cost-effective [35]. This optimization involves the configuration and, when necessary, reconfiguration of machines, which can manifest as either physical or virtual entities, tailored to meet the specific demands of users. The overarching goal is to harness the benefits of cloud infrastructure, granting organizations access to computing resources for tasks such as data processing and storage without the necessity for extensive on-site hardware investments. Within the realm of migration challenges, security emerges as a pivotal concern, underscored by the inherent lack of visibility into the security measures of the cloud platform. The dynamic nature of cloud environments necessitates cloud service providers to continually verify the effectiveness of their security protocols, thereby elevating security to a paramount consideration in the selection of a provider. The need for robust security measures is accentuated by the potential risks associated with unauthorized access, data breaches, and other cyber threats. Ensuring data integrity throughout the migration process, particularly for sensitive data, constitutes another formidable challenge. This demands the utilization of migration tools equipped to verify the integrity of data and the implementation of comprehensive risk assessments. The imperative is not only to facilitate a seamless transition but also to safeguard the confidentiality, availability, and integrity of data throughout the migration journey [36].

The complexity of choosing the right cloud vendor further amplifies the challenges in migration. Factors such as vendor lock-in and portability become critical considerations, necessitating thorough research and evaluation of data migration tools provided by key industry players like Google Cloud Platform, Amazon Web Services, and Microsoft Azure [35]. The decision-making process involves a nuanced examination of each vendor's capabilities, pricing structures, and compatibility with the organization's specific requirements. Vendor lock-in, in particular, poses a multifaceted challenge, as organizations must weigh the benefits of a chosen provider against the potential challenges of transitioning to an alternative provider in the future. This intricate decision-making process requires careful consideration of long-term business objectives, technology roadmaps, and the evolving landscape of cloud services. In conclusion, migration challenges in the context of cloud computing extend beyond the technical aspects, delving into critical considerations such as security, data integrity, and vendor selection. A comprehensive approach to addressing these challenges involves a strategic blend of technological solutions, risk assessments, and informed decision-making to navigate the complexities inherent in migrating computing resources to the cloud [36]. The process of migrating data to the cloud entails numerous challenges, with a pivotal aspect being the selection of an appropriate cloud provider. This decision involves a nuanced evaluation of support mechanisms, available tools, and migration approaches. Compounding this complexity are concerns related to vendor lock-in and portability, necessitating careful consideration of long-term implications and potential transition challenges [37, 38].

The human factor introduces a significant dimension to the migration process, as adapting to new systems and software platforms may incur additional costs and resistance from employees [39]. Overcoming this challenge requires not only technological solutions but also strategic change management practices to facilitate a smooth transition. To address these multifaceted challenges, effective migration strategies and models assume paramount importance. Diverse cloud migration strategies, including re-hosting (lift and shift), re-platforming, repurchasing, re-architecting, and retiring, provide organizations with tailored approaches aligned with their specific needs [40, 41]. These strategies ensure a structured and planned migration, contributing to a smoother transition to the cloud. The intricacies of data migration demand meticulous planning and execution. Strategies encompass thorough analysis of existing data, accurate mapping between systems, and the development of robust conversion programs with rigorous testing. Security challenges are addressed through encryption and access controls, while data quality is enhanced using cleansing techniques. Flexible migration approaches, such as re-hosting or re-platforming, are adopted based on organizational requirements. Evaluation of cloud computing models, including IaaS, PaaS, and SaaS, is conducted to determine suitability. A well-defined strategy covers aspects such as legacy data, data quality, tools, and environment. Continuous monitoring, robust backups, employee training, and post-migration activities contribute to the overall success of the migration initiative [42].

Vendor management assumes a critical role in ensuring clear Service Level Agreements (SLAs), while agile principles allow for an iterative and adaptable migration process. Collectively, these solutions navigate the complexities inherent in data migration, enhancing the likelihood of a successful and seamless transition to the cloud [43]. The hesitancy exhibited by financial institutions, particularly banks, in embracing cloud computing security is deeply rooted in apprehensions surrounding the safeguarding of sensitive information. Despite technological advancements, the perceived probability of a successful cyberattack remains a significant deterrent for these entities. While cloud migration promises benefits such as cost reduction and scalability, persistent security concerns continue to influence decision-making. In addressing these apprehensions, strategic initiatives and technologies are employed within the financial sector. Systems such as Google File System (GFS) and Hadoop Distributed File System (HDFS) are designed to optimize data access, prioritizing efficiency and reliability. Notably, Prediction-Based Encryption (PBE) emerges as a key strategy, providing fine-grained access control and serving as a solution for secure multicasting. The operational framework of this proposed system involves encryption, decryption keys, and a secure channel, as exemplified by the HDFS layer [44].

2.8. Latency

Cloud computing has revolutionized the utilization of computing infrastructure by businesses, relying on prominent providers like Google Cloud Platform and Amazon EC2 [45]. Despite the flexibility to customize computing resources, the assurance of guaranteed network latency for applications remains a persistent challenge. This paper delves into the impact of network latency on various cloud applications, employing an experimental methodology that involves artificially injecting latency. The study encompasses applications ranging from DNS to machine learning workloads, capturing diverse complexities. Leveraging a bespoke hardware appliance, NRG, the experimental setup enables precise control over latency. By measuring application performance under varying network latency scenarios, the research aims to furnish insights into the nuanced relationship between latency and the efficiency of cloud applications. The contributions of the paper include a detailed description of the experimental methodology and an extensive measurement study, offering valuable insights into the dynamic interplay of cloud computing and latency [46].

Performance monitoring, particularly with Infrastructure Response Time (IRT), emerges as a vital aspect for meeting service level agreements. Security measures, including firewalls, play a crucial role in mitigating user-end latency issues such as Distributed Denial of Service (DDoS) attacks. The proposed architecture comprehensively addresses security challenges like botnet attacks through the strategic deployment of firewalls and continuous monitoring. Overall, a holistic approach that encompasses technology enhancements, network optimization, performance monitoring, and robust security measures is deemed essential for ensuring an efficient and secure cloud computing experience [47]. This underscores the multifaceted considerations and strategic measures required to navigate the complexities of cloud computing infrastructure. In the burgeoning landscape of Internet of Things (IoT) applications, the proliferation of internet-connected devices has resulted in the generation of time-sensitive data, commonly referred to as "Big Data." Traditional Cloud computing, although encountering latency issues in processing real-time data, faces potential bottlenecks. Conversely, Fog computing, leveraging its proximity advantages, contends with challenges in optimizing latency, particularly concerning data expiration time. In response to these challenges, a proposed solution involves the implementation of a hybrid Fog-Cloud architecture. This architectural framework is specifically designed to address latency concerns by formulating the scheduling problem for IoT requests utilizing Integer Linear Programming (ILP). Additionally, an adaptive genetic algorithm (GA) is developed as a heuristic approach aimed at efficiently minimizing service time for IoT requests [48]. The genetic algorithm is intricately customized, incorporating a specialized representation of solutions (chromosomes) and a well-designed crossover operator. To ensure adherence to problem constraints, a penalty mechanism is integrated, discouraging the selection of infeasible solutions. This tailored GA solution is then rigorously evaluated through a network simulator to quantify its impact on delay performance [49].

Embracing Infrastructure Response Time (IRT) as a novel and crucial approach, this paradigm serves as a metric to assess infrastructure performance by measuring the time taken for an application to request work within the virtual environment and for the virtual environment to complete the request. In this holistic evaluation, Application Response Time becomes a key indicator, offering insights into the duration required for applications to respond to user requests. The scope extends to virtualization performance, encompassing critical parameters like the number of Virtual Machines (VMs), time for new VM creation, duration for application migration between VMs, and resource allocation time. Advocating for IRT as a pivotal performance metric, the establishment of benchmarks becomes paramount to delineate optimal performance standards. Continuous monitoring tools for infrastructure, applications, and virtualization performance are deemed essential components of this comprehensive approach. Moreover, adaptive strategies are recommended to dynamically respond to fluctuations in latency and network conditions, ensuring the resilience and responsiveness of the computing environment. [50, 51].

3. Literature Review

Cloud computing, a phenomenon capturing widespread attention across diverse domains, including individual users and government agencies, is characterized by a nuanced definition. Essentially, it functions as a subscription-based service, providing access to networked storage space and computer resources. In drawing a contrast, traditional documents created using word processing software typically reside on the device unless intentionally moved [52]. This distinction becomes evident when considering an email client as an exemplar of cloud computing. In this scenario, data is not confined to the device but is seamlessly accessible through networked services. However, the scope of cloud computing transcends the confines of emails; it extends to encompass a myriad of data and applications, offering users the flexibility to select and access information within the expansive realm of the cloud. Public clouds, managed externally, usher in a host of challenges that reverberate across the realms of data sovereignty and compliance. The intricate landscape of the hybrid cloud model further amplifies concerns, placing a premium on achieving seamless data integration to navigate the complexities of this multifaceted paradigm. Infrastructure as a Service (IaaS) encounters its share of obstacles, particularly in grappling with issues related to data portability and the potential entanglements of vendor lock-in. Successfully addressing these challenges demands a strategic planning approach that meticulously charts the course forward [53].

Moving into the Platform as a Service (PaaS) domain, unique challenges come to the forefront. Here, version control becomes a focal point, accompanied by the ongoing pursuit of maintaining data consistency. Navigating these intricacies requires a strategic approach that aligns with the dynamic nature of cloud computing evolution. On the Software as a Service (SaaS) front, models face their own set of challenges, particularly in the domains of data customization and control, highlighting the ongoing dialogue around striking a delicate balance between user flexibility and centralized management. Within the overarching concern of security, a spectrum of threats demands vigilant attention. From guest-hopping attacks to network-level threats, the imperative for robust virtualization security measures becomes apparent. Incorporating tools such as forensics and High Assurance Platforms becomes essential. The network-level security attacks underscore the fundamental importance of deploying robust firewalls, encryption protocols, and authentication mechanisms to fortify the overall security posture in the ever-evolving landscape of cloud computing. Comprehensively, this understanding extends to service models like Infrastructure as a Service (IaaS), Platform as a Service (PaaS), and Software as a Service (SaaS). Within these models, SaaS stands out for delivering business processes and applications through web browsers, exemplified by services like Gmail. PaaS steps in to provide a computing platform over the internet, streamlining the development and deployment of web applications. Meanwhile, IaaS virtually delivers computing resources, affording clients the flexibility to lease resources as needed [54].

This foundational understanding not only demystifies the intricacies of cloud service models but also lays the groundwork for a nuanced exploration of deployment models. The Private Cloud model caters to secure and exclusive environments, ensuring a tailored approach to data and applications. On the other hand, the Community Cloud model fosters shared benefits among a specific user community. Public Cloud, in contrast, is characterized by widespread service delivery to the general public. Finally, the Hybrid Cloud model emerges as a versatile solution, facilitating data and application portability across various environments. Amidst the pervasive integration of cloud computing into diverse domains, a tapestry of challenges persists, encompassing pivotal aspects such as confidentiality, security, and performance optimization. These challenges underscore the evolving landscape where the benefits of cloud adoption coexist with the imperative to address and mitigate associated complexities [55].

In tackling these challenges systematically, a noteworthy contribution comes from a comprehensive literature review. This review not only identifies but categorizes these challenges, presenting a structured taxonomy that serves as a valuable framework for international discourse. By delineating and organizing these challenges, the literature review not only offers a snapshot of the current landscape but also becomes a catalyst for fostering collaborative dialogue on innovative solutions and advancements within the dynamic realm of cloud computing [56].

In the realm of hosting and delivering services over the Internet, cloud computing stands as a transformative paradigm, garnering significant attention for its appeal to business owners. This technology eliminates the need for customers to plan extensive provisioning, offering organizations the flexibility to commence operations on a small scale and scale up only in response to increasing service demand. Despite its vast prospects for the IT sector, cloud computing technology is still in its nascent stages, accompanied by numerous challenges that necessitate careful consideration. This literature review provides a comprehensive survey of cloud computing, encompassing fundamental concepts, architectural principles, cutting-edge implementations, and prevailing research problems [57].

Furthermore, the pay-as-you-go nature of cloud operations introduces a crucial dimension. Each configuration, characterized by factors such as cluster size and VM type, incurs both operating costs and execution times. Consequently, optimizing a task becomes a strategic choice between achieving the lowest cost or the shortest execution time—two distinct configurations with critical implications for service quality and commercial competitiveness. The selection of an appropriate cloud configuration becomes paramount, considering that a faulty setup can lead to costs up to 12 times higher for the same performance target. The potential savings from effective cloud design are particularly pronounced for recurring operations with similar workloads. However, the challenge lies in choosing the optimal cloud configuration, as achieving the delicate balance between high accuracy, minimal overhead, and adaptability for diverse applications simultaneously proves to be a complex endeavour [58].

Kamal et al. studied the challenges in distributed data management in cloud computing, they highlighted the challenges in managing multi-tenant systems [59]. Chandaran et al. were mainly investigating on data security challenges in the data management and cloud computing, in addition they presented the solutions to tackle these challenges [60]. Horey et al. proposed the use of clout-get to simplify the management of distributed services in cloud computing environment [61]. Forell et al. discussed the importance of effective management of shared resources and services for cloud computing and outlined of challenges in cloud management [62]. el-khameesy et al. proposed an effective distributed scheme to handle security aspects of data storage [63]. Ooi presented the opportunities and challenges of developing a scalable cloud data management system and also discussed the aspects of the future of the cloud computing [64]. Sakr et al. explored the option of building parallel data management systems capable of serving petabytes of data [65]. Zhang et al. discussed the challenges faced by cloud computing due to the heterogeneous and dynamic nature of resources and workloads, as well as the proposed solutions to reduce data center energy consumption and improve application performance [66]. Tameberg et al. proposed an algorithm that works towards meeting applications performance objectives [67]. Lehner et al. discussed the importance of efficient database management, the consideration of SaaS as an alternative to on-premise data management, and the emphasis on virtualization

[68]. George et al. we're focusing on exploring replication procedures in the cloud environment [69]. Chhabra et al. provide a comprehensive overview of the prevalent problems with sharing and communication in organizations of all sizes in the context of cloud computing [70]. Sen et al. discuss Security concerns that exist in handling of data and computations by third-party infrastructure [71]. Sundareswaran et al. Propose a highly decentralized information accountability framework to overcome fears in losing control of own data.

4. Discussion and Comparison

In this section, we provide a discussion about the reviewed papers. We have chosen the most recent papers for this study, mainly focusing on how to tackle the challenges of distributed data management in cloud computing as well as presenting the approaches and the solutions to it. In addition, we summarized the results and findings of the review papers in the comparison table 1 below which extracts the objectives of each research paper as well as its results.

Table 1: literature review in summary to solve the Challenges of Distributed Data Management in Cloud Computing

Authors	Region	Author's objectives	Main findings (Results)
[58], 2014	Australia	Distributed data management systems in the cloud promise rapid elasticity and horizontal scalability so that cloud applications can sustain enormous growth in data volume, velocity, and value.	1. Cloud computing has changed the payment model, distributed data management systems offer scalability. 2. Challenges in managing multi-tenant systems are discussed.
[59], 2017	India	Data security and data management are the key issues that challenge trusted cloud computing.	1. Discussion of diverse types of data and their specific attributes2. 2. In-depth literature on the taxonomy of data management, issues, and approaches to solutions in cloud computing. 3. Attention to the entities and actors responsible for addressing the issues related to data security and management in cloud computing.
[60], 2012	United States	Distributed services should be viewed as a single application consisting of virtual machines.	1. Infrastructure-as-a-Service and Big Data platforms have revolutionized data storage and processing, but the lack of readily available solutions for customizing or creating new Big Data stacks is a significant challenge. 2. The introduction of Cloud-Get simplifies the management of distributed services in a cloud computing environment.
[61], 2011	USA	The research community can address challenges in cloud management.	1. Importance of effective management of shared resources and services for cloud computing, the outline of challenges in cloud management, and the description of new management

			architectures to address these challenges. 2. The paper also provides open issues and challenges in cloud management for the research community to address.
[62], 2012	Egypt	The proposed distributed scheme achieves the integration of storage correctness insurance and data error localization.	1. The focus on security aspects of data storage. 2. The proposal of an effective distributed scheme with unique features.
[63], 2009	Singapore	Building a scalable data management system on existing commercial cloud platforms poses a grand challenge.	1. Cloud Computing is changing the IT industry by providing computing utility as a service. 2. Building a scalable data management system on existing commercial Cloud platforms poses a grand challenge. 3. The paper presents the opportunities and challenges of developing a scalable Cloud data management system, including the anatomy of a Cloud data-intensive system and three main challenges posed by a scalable Cloud data processing system. It also discusses possible solutions, current practices, and speculates the future of Cloud Computing.
[64], 2011	World wide	The database research community has started to investigate cloud computing as a cost-effective option to build scalable parallel data management systems capable of serving petabytes of data for millions of users.	The exploration of cloud computing as a cost-effective option for building scalable parallel data management systems to handle vast amounts of data for data-intensive applications.
[65], 2014	Canada	Effective workload management in cloud computing environments remains to be a difficult challenge.	1. The challenges faced by cloud computing due to the heterogeneous and dynamic nature of resources and workloads. 2. Proposed solutions to reduce data center energy consumption and improve application performance.
[66], 2017	Sweden	A distributed algorithm works towards meeting individual application's performance objectives.	1. The presented algorithm quickly converges and performs near optimal in terms of system-wide operational cost and application performance, outperforming similar naïve and random methods.

			2. The algorithm was evaluated over two different types of topologies with varying degrees of heterogeneity and was found to quickly and consistently converge despite a high degree of heterogeneity in the system.
[67], 2013	Germany	The need and desire to provide database functionality in the context of the Database as a Service paradigm for database outsourcing are fundamental challenges posed by the need and desire to provide database functionality in the context of the Database as a Service paradigm for database outsourcing.	1. The importance of efficient database management, the consideration of 'Software-as-a-Service' as an alternative to on-premise data management, and the emphasis on virtualization as a key technique for establishing the 'Everything-as-a-Service' paradigm. 2. The paper also outlines challenges and open research questions in providing a data management solution following this paradigm.
[68], 2017	India	Replication is a powerful and effective strategy to surpass the performance, availability and reliability in the distributed framework by making different duplicates of the same information.	1. Replication is a powerful strategy to improve performance, availability, and reliability in the distributed framework. 2. The study aims to explore current replication procedures in the cloud environment.
[69], 2022	India	Cloud computing has become the backbone of the computing industry.	1. The paper provides a comprehensive overview of the prevalent problems with sharing and communication in organizations of all sizes in the context of cloud computing. - Cloud computing has experienced significant growth and has made numerous noteworthy contributions, making it a huge success. 2. Cloud computing is getting attention as a way to lower capital expenditure and boost system effectiveness.
[70], 2015	India	Cloud computing offers different service models and provides an on-demand, pay-per-use, multitenant, pervasive model which makes it suitable to adapt in the fast-growing industries.	1. Cloud computing is an emerging field with different service models and a pervasive model suitable for fast-growing industries. - Security concerns exist due to the handling of data and computations by third-party infrastructure. 2. The paper provides a distributed solution to implement trusted cloud computing platforms to address infrastructure level attacks.

[71], 2012	USA	Users' fears of losing control of their own data can become a significant barrier to the wide adoption of cloud services.	Include the proposal of a highly decentralized information accountability framework, the utilization of JAR programmable capabilities for authentication and logging, and the provision of distributed auditing mechanisms.

5. Extracted Statistics

According to the reviewed papers there are a worldwide interest in studying and investigating the process of tackling the challenges of distributed data management in cloud computing and finding practical approaches to solve the challenges in this area. We have noticed that the majority of research about this important topic has been carried on in India and United States as shown in figure 2.

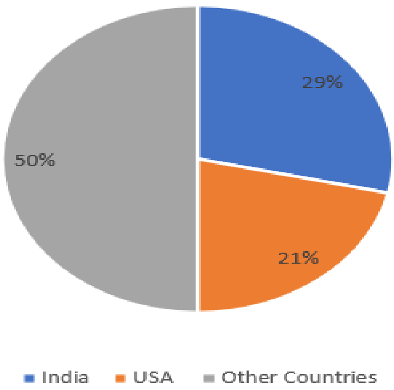

Fig (2): Distribution of the researches carried out in this topic by region.

6. Recommendations

Due to the importance of distributed data management in cloud computing and the continuous updates in this field, there are always challenges that come up and there is always a need for solutions to solve these challenges. Therefore, we recommend more researches to be done on this field to address the challenges in early stages as the cloud computing emerges which will help in discovering the solutions more easily. In addition, we recommend the focus of the researches be more on the security of data management in cloud computing as it is a critical aspect for the organizations and end users.

7. Conclusion

In conclusion, the challenges in cloud computing, spanning security, scalability, interoperability, migration, and latency, demand a nuanced approach. Security focuses on cryptography to ensure confidentiality, integrity, and availability. Privacy-preserving auditing integrates homomorphic linear authenticators and random masking. Scalability and consistency in distributed databases evolve with key-value pair models and replication strategies. Interoperability issues emphasize open standards and collaborative efforts like the Cloud Computing Interoperability Forum (CCIF). Migration complexities require strategic approaches, including re-hosting and risk assessments. Latency concerns are addressed through innovative solutions like a hybrid Fog-Cloud architecture. In essence, the dynamic landscape of cloud computing research requires a comprehensive understanding. From cryptographic innovations to interoperability challenges, migration strategies to latency optimization, the pursuit of robust security and efficiency shapes the future of cloud computing ecosystems through technological advancements and collaborative standards.

References

[1] Vankudre, A. "Data Management in Cloud Computing" Adarsh Institute Of Technology.

[2] Vela, F. "Cloud Computing Management" Masaryk University Faculty of Informatics, 2013, pp. 12-13.

[3] Mohammad, S., Breß, S., Schallehn, E. "Cloud Data Management: A Short Overview and Comparison of Current Approaches."

[4] "Data Management Solutions in the Cloud: Accessing, Sharing, and Processing Data on Google Cloud Platform (GCP)."

[5] S. R. M. Zeebaree, H. Shukur, L. Haji, and R. Zebari, "Characteristics and Analysis of Hadoop Distributed Systems," *Technol. Reports Kansai Univ.*, vol. 62, no. 4, pp. 1555–1564, 2020.

[6] Wang, C., Wang, Q., Ren, K., Lou, W. "Privacy-preserving public auditing for data storage security in cloud computing," IEEE INFOCOM, 2010, pp. 1-9.

[7] P. Y. Abdullah, S. R. M. Zeebaree, K. Jacksi, and R. R. Zeabri, "an Hrm System for Small and Medium Enterprises (Sme)S Based on Cloud Computing Technology," *Int. J. Res. - GRANTHAALAYAH*, vol. 8, no. 8, pp. 56–64, 2020, doi: 10.29121/granthaalayah.v8.i8.2020.926

[8] Kaufman, L. M. "Data Security in the World of Cloud Computing," IEEE Security & Privacy Magazine, 2009.

[9] Security Guidance for Critical Areas of Focus in Cloud Computing," Cloud Security Alliance, 2010.

[10] P. Abdullah, H. Shukur, K. Jacksi, P. Y. Abdullah, S. R. M. Zeebaree, and H. M. Shukur, "HRM System using Cloud Computing for Small and Medium Enterprises (SMEs) Head of IT Department View project Client-Server and Video Broadcasting View project HRM System using Cloud Computing for Small and Medium Enterprises (SMEs)," *Tkru*, vol. 62, no. May, 2020, [Online]. Available: https://www.researchgate.net/publication/341883552.

[11] [Gartner Report](http://www.gartner.com/DisplayDocument?id=685308).

[12] S. R. M. Zeebaree, R. R. Zebari, K. Jacksi, and D. A. Hasan, "Security Approaches For Integrated Enterprise Systems Performance: A Review," 2019, [Online]. Available: https://api.semanticscholar.org/CorpusID:219780910.

[13] Chang, F., et al. "Bigtable: A Distributed Storage System for Structured Data," OSDI, 2006, pp. 205–218.

[14] Cooper, B.F., et al. "PNUTS: Yahoo!'s Hosted Data Serving Platform," VLDB Endowment, 2008, pp. 1277–1288.

[15] S. R. M. Zeebaree, A. B. Sallow, B. K. Hussan, and S. M. Ali, "Design and Simulation of High-Speed Parallel/Sequential Simplified DES Code Breaking Based on FPGA," in *2019 International Conference on Advanced Science and Engineering (ICOASE)*, 2019, pp. 76–81, doi: 10.1109/ICOASE.2019.8723792.

[16] DeCandia, G., et al. "Dynamo: Amazon's Highly Available Key-Value Store," SOSP, 2007, pp. 205–220.

[17] Lindsay, B.G., et al. "Computation and Communication in R*: A Distributed Database Manager," ACM Trans. Comput. Syst., 1984, pp. 24–38.

[18] S. R. M. Zeebaree, "DES encryption and decryption algorithm implementation based on FPGA," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 18, no. 2, pp. 774–781, 2020, doi: 10.11591/ijeecs.v18.i2.pp774-781.

[19] Rothnie Jr., J.B., et al. "Introduction to a System for Distributed Databases (SDD1)," ACM Trans. Database Syst., 1980, pp. 1–17.

[20] Bernstein, P.A., Hadzilacos, V., Goodman, N. "Concurrency Control and Recovery in Database Systems," Addison Wesley, 1987.

[21] H. Malallah *et al.*, "A Comprehensive Study of Kernel (Issues and Concepts) in Different Operating Systems," *Asian J. Res. Comput. Sci.*, vol. 8, no. 3, pp. 16–31, 2021, doi: 10.9734/ajrcos/2021/v8i330201.

[22] Weikum, G., Vossen, G. "Transactional Information Systems: Theory, Algorithms, and the Practice of Concurrency Control and Recovery," Morgan Kaufmann, 2001.

[23] H. Shukur, S. Zeebaree, R. Zebari, O. Ahmed, L. Haji, and D. Abdulqader, "Cache Coherence Protocols in Distributed Systems," *J. Appl. Sci. Technol. Trends*, vol. 1, no. 3, pp. 92–97, 2020, doi: 10.38094/jastt1329.

[24] Harsh, P., Dudouet, F., Cascella, R. G., Jegou, Y., Morin, C. "Using Open Standards for Interoperability Issues, Solutions, and Challenges facing Cloud Computing," Inria Rennes - Bretagne Atlantique.

[25] Schubert, L., et al. "The Future of Cloud Computing: Opportunities for European Cloud Computing Beyond," European Commission, 2010.

[26] S. R. Zeebaree *et al.*, "Multicomputer Multicore System Influence on Maximum Multi-Processes Execution Time," *TEST Eng. Manag.*, vol. 83, no. 03, pp. 14921–14931, 2020.

[27] Petcu, D. "Portability and Interoperability between Clouds: Challenges and Case Study, Towards a Service-Based Internet," Lecture Notes in Computer Science, 2011, vol. 6994, pp. 62-74.
[28] Rashidi, B., Sharifi, M., Jafari, T. "A Survey on Interoperability in the Cloud Computing Environments," International Journal of Modern Education and Computer Science, 2013, pp. 17-23.
[29] Z. N. Rashid, S. R. M. Zeebaree, M. A. M. Sadeeq, R. R. Zebari, H. M. Shukur, and A. Alkhayyat, "Cloud-based Parallel Computing System Via Single-Client Multi-Hash Single-Server Multi-Thread," in *2021 International Conference on Advance of Sustainable Engineering and its Application (ICASEA)*, 2021, pp. 59–64, doi: 10.1109/ICASEA53739.2021.9733077.
[30] Lewis, G. A. "The Role of Standards in Cloud Computing Interoperability," October 2012.
[31] Rao, R., Prakash, P. "Improving Security for Data Migration in Cloud Computing using Randomized Encryption Technique," 2013.
[32] Z. N. Rashid, S. R. M. Zeebaree, R. R. Zebari, S. H. Ahmed, H. M. Shukur, and A. Alkhayyat, "Distributed and Parallel Computing System Using Single-Client Multi-Hash Multi-Server Multi-Thread," in *2021 1st Babylon International Conference on Information Technology and Science (BICITS)*, 2021, pp. 222–227, doi: 10.1109/BICITS51482.2021.9509872.
[33] Amin, R., Vadlamudi, S., Rahaman, M. M. "Opportunities and Challenges of Data Migration in Cloud."
[34] Cranford, N. "Five Challenges of Cloud Migration," RCR Wireless News, 2017.
[35] L. M. Haji, S. R. M. Zeebaree, O. M. Ahmed, M. A. M. Sadeeq, H. M. Shukur, and A. Alkhavvat, "Performance Monitoring for Processes and Threads Execution-Controlling," in *2021 International Conference on Communication & Information Technology (ICICT)*, 2021, pp. 161–166, doi: 10.1109/ICICT52195.2021.9568445.
[36] Cardoso, A., Moreira, F., Simões, P. "A Survey of Cloud Computing Migration Issues and Frameworks."
[37] M. A. M. Sadeeq and S. R. M. Zeebaree, "Design and implementation of an energy management system based on distributed IoT," *Comput. Electr. Eng.*, vol. 109, p. 108775, 2023, doi: https://doi.org/10.1016/j.compeleceng.2023.108775.
[38] "Pros and Cons of Cloud Data Migration," Flatworld Solutions, [Online] Available: [link](https://www.flatworldsolutions.com/datamanagement/articles/pros-cons-cloud-data-migration.php).
[39] Uggirala, A. "Five Cloud Migration Strategies for Applications," Imperva Blog, 2018.
[40] Furfaro, A., Garro, A., Tundis, A. "Towards Security as a Service (SecaaS): On the Modeling of Security Services for Cloud Computing," 2014 Int. Carnahan Conf. Secur. Technol. ICCST.
[41] M. A. M. Sadeeq and S. R. M. Zeebaree, "DPU-ALDOSKI dataset of Monitoring and Controlling distributed far distances energy consumed system based on Internet of Things," *Data Br.*, vol. 49, p. 109455, 2023, doi: https://doi.org/10.1016/j.dib.2023.109455.
[42] Abdou Hussein, A. "Data Migration Need, Strategy, Challenges, Methodology, Categories, Risks, Uses with Cloud Computing, and Improvements in Its Using with Cloud Using Suggested Proposed Model (DMig 1)," Journal of Information Security, 2021.
[43] Kushwah, V. S., Saxena, A. "A Security Approach for Data Migration in Cloud Computing," 2013.
[44] Moore, A. W. "Characterizing the Impact of Network Latency on Cloud-Based Applications' Performance," 2017.
[45] L. M. Haji, S. R. M. Zeebaree, Z. S. Ageed, O. M. Ahmed, M. A. M. Sadeeq, and H. M. Shukur, "Performance Monitoring and Controlling of Multicore Shared-Memory Parallel Processing Systems," in *2022 3rd Information Technology To Enhance e-learning and Other Application (IT-ELA)*, 2022, pp. 44–48, doi: 10.1109/IT-ELA57378.2022.10107953.
[46] Bali, M. S., Khurana, S. "Effect of Latency on Network and End User Domains in Cloud Computing," Dec 2013.
[47] Aburukbaa, R. O., AliKarrara, M., Landolsia, T., El-Fakih, K. "Scheduling Internet of Things Requests to Minimize Latency in Hybrid Fog-Cloud Computing."
[48] Y. S. Jghef, S. R. M. Zeebaree, Z. S. Ageed, and H. M. Shukur, "Performance Measurement of Distributed Systems via Single-Host Parallel Requesting using (Single, Multi and Pool) Threads," in *2022 3rd Information Technology To Enhance e-learning and Other Application (IT-ELA)*, 2022, pp. 38–43, doi: 10.1109/IT-ELA57378.2022.10107923.
[49] Bali, M. S., Khurana, S. "Effect of Latency on Network and End User Domains in Cloud Computing."

[50] Z. S. Ageed and S. R. M. Zeebaree, “Distributed Systems Meet Cloud Computing: A Review of Convergence and Integration,” *Int. J. Intell. Syst. Appl. Eng.*, vol. 12, no. 11s, pp. 469–490, 2024, [Online]. Available: https://ijisae.org/index.php/IJISAE/article/view/4468.

[51] Kamal, J.M.M. and Murshed, M., 2014. Distributed database management systems: architectural design choices for the cloud. *Cloud Computing: Challenges, Limitations and R&D Solutions*, pp.23-50.

[52] S. R. M. Zeebaree , *et al.*, “Single-threading Based Distributed-multiprocessor-machines Affecting by Distributed-parallel-computing Technology,” *J. Duhok Univ.*, vol. 26, no. 2, pp. 416–426, 2023, doi: 10.26682/csjuod.2023.26.2.39.

[53] Huth, A., Cebula, J. "The Basics of Cloud Computing."

[54] Turab, N. M., Talebn, A. A., Masadeh, S. R. "Cloud Computing Challenges and Solutions."

[55] Mell, P., Grance, T. "The NIST Definition of Cloud Computing, National Institute of Standards and Technology," 2009.

[56] Ghanam, Y., Ferreira, J., Maurer, F. "Emerging Issues & Challenges in Cloud Computing—A Hybrid Approach."

[57] Hassan, A. "Cloud Computing: Literature Review," 2022, [Online] Available: [link](https://mars.gmu.edu/bitstream/handle/1920/11608/hassan_cloud.pdf?sequence=1).

[58] "Architecture of Cloud Computing," GeeksforGeeks, 2022, [Online] Available: [link](https://www.geeksforgeeks.org/architecture-of-cloud-computing/amp/).

[59] Chandran, A. and Shyamala, C.K., 2017, January. Data management issues in cloud integrated computing: A big picture. In *2017 4th International Conference on Advanced Computing and Communication Systems (ICACCS)* (pp. 1-8). IEEE.

[60] Horey, J., Begoli, E., Gunasekaran, R., Lim, S.H. and Nutaro, J., 2012. Big data platforms as a service: challenges and approach. In *4th USENIX Workshop on Hot Topics in Cloud Computing (HotCloud 12)*.

[61] Forell, T., Milojicic, D. and Talwar, V., 2011, May. Cloud management: Challenges and opportunities. In *2011 IEEE international symposium on parallel and distributed processing workshops and Phd forum* (pp. 881-889). IEEE.

[62] el-Khameesy, N. and Rahman, H.A., 2012. A proposed model for enhancing data storage security in cloud computing systems. *Journal of Emerging Trends in Computing and Information Sciences*, *3*(6), pp.970-974.

[63] Ooi, B.C., 2009, October. Cloud data management systems: Opportunities and challenges. In *2009 Fifth International Conference on Semantics, Knowledge and Grid* (pp. 2-2). IEEE.

[64] Sakr, S., Liu, A., Batista, D.M. and Alomari, M., 2011. A survey of large scale data management approaches in cloud environments. *IEEE communications surveys & tutorials*, *13*(3), pp.311-336.

[65] Zhang, Q. and Boutaba, R., 2014, May. Dynamic workload management in heterogeneous cloud computing environments. In *2014 IEEE Network Operations and Management Symposium (NOMS)* (pp. 1-7). IEEE.

[66] Tärneberg, W., Papadopoulos, A.V., Mehta, A., Tordsson, J. and Kihl, M., 2017, May. Distributed approach to the holistic resource management of a mobile cloud network. In *2017 IEEE 1st International Conference on Fog and Edge Computing (ICFEC)* (pp. 51-60). IEEE.

[67] Lehner, W., Sattler, K.U. and Sattler, K.U., 2013. *Web-scale data management for the cloud* (Vol. 5). Berlin: Springer.

[68] George, S. and Edwin, E.B., 2017, December. A review on data replication strategy in cloud computing. In *2017 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC)* (pp. 1-4). IEEE.

[69] Chhabra, S. and Singh, A.K., 2022. A comprehensive vision on cloud computing environment: Emerging challenges and future research directions. *arXiv preprint arXiv:2207.07955*.

[70] Sen, P., Saha, P. and Khatua, S., 2015, February. A distributed approach towards trusted cloud computing platform. In *2015 Applications and Innovations in Mobile Computing (AIMoC)* (pp. 146-151). IEEE.

[71] Sundareswaran, S., Squicciarini, A. and Lin, D., 2012. Ensuring distributed accountability for data sharing in the cloud. *IEEE transactions on dependable and secure computing*, *9*(4), pp.556-568.