REDVET - Revista electrónica de Veterinaria - ISSN 1695-7504
Vol 25, No.2 (2024)
http://www.veterinaria.org
Article Received:10/07/2024 Revised:22/07/2024 Published:01/08/2024

Study and Analysis of Grid-Based Cloud Computing using different Computer Applications

Preeta Rajiv Sivaraman[1*], Dr. Rashi Agarwal[2], Dr. Renu Jain[3]

[1*]Ph.D. Research Scholar, Department of Computer Science, C. S. J. M. University, Kanpur, Uttar Pradesh, India, Email: Preetasiva.@gmail.com
[2]Faculty of Computer Science, Department of Computer Science, C. S. J. M. University, Kanpur, Uttar Pradesh, India.
[3]Head of the Department, Department of Computer Science, C. S. J. M. University, Kanpur, Uttar Pradesh, India.

***Corresponding Author:** Preeta Rajiv Sivaraman
*Ph.D. Research Scholar**,** Department of Computer Science**,** C. S. J. M. University, Kanpur**,** Uttar Pradesh, India, Email: Preetasiva.@gmail.com

Abstract
The research focuses on the perception of cloud web portal service data storage quality between companies for internet-based cloud data storage grids scheduling services. The study found significant variance between the two sample populations (Infosys and TCS) in attributes such as Tangibility, Reliability, and Responsiveness, Credibility, and Data security. The client response to reliability was more varied, with Infosys ranking higher than TCS. In responsiveness, TCS ranked higher than Infosys, suggesting smaller enterprises are more responsive. In credibility and Data security, both companies had a mean score above 10, indicating high levels of Cloud Storage credibility. ANOVA testing confirmed the findings, with no significant difference in the perception of credibility and security aspects between the two main Cloud Services Providing companies. However, there was no clear influence of one attribute over another in forming a client's perception of a company's Cloud web portal qualities observation. This thesis examines Data Grid Scheduling systems in Cloud web portals, focusing on their unique features such as heavy computing requirements, geographically distributed resources, and collaboration among users. It also examines the architecture of Data Grids used in cloud portals, data transport mechanisms, data replication systems, and resource allocation. The study demonstrates that considering data presence and computational resource availability improves application scheduling performance by improving job turnaround time. The Grid bus broker is effective for executing scientific applications on Grid resources, but there is still a question of whether newer application models can accommodate the current architecture. The Grid community has recently standardized on the Web Cloud Services Resource Framework (WSRF), which requires the broker to compose services based on their attributes and create service aggregations to achieve users' utility functions. The research findings indicate that client's Cloud web portal qualities perceive a difference in service offered by internet sites, with significant differences in tangibility, reliability, and responsiveness. Factors such as security and credibility are more relevant for Cloud web portal Data Grids or online cloud storage service companies. Future work aims to address new questions and improve understanding of Data Grid Scheduling in Cloud web portals.

Keywords: Deep learning, Cloud security, Artificial Intelligence, Block chain, Cryptography, Machine Learning.

Introduction

Cloud computing is a technology that allows users to access applications as utilities over the internet, allowing them to create, configure, and customize applications online [1]. The market for cloud computing services was $16 billion in 2008 and is expected to rise to $52 billion per year by 2014. Cloud computing offers cost advantages three to five times for business applications and more than five times for consumer applications. Enterprises are increasingly seeking to reduce computing costs, consolidating their IT operations and using virtualization technologies [2]. Cloud computing is a platform that supplies, configures, and reconfigures servers, while applications are extended to be accessible through the internet. Cloud computing differs from traditional computing paradigms in that it is scalable, can be encapsulated as an abstract entity, and can be dynamically configurable [3].

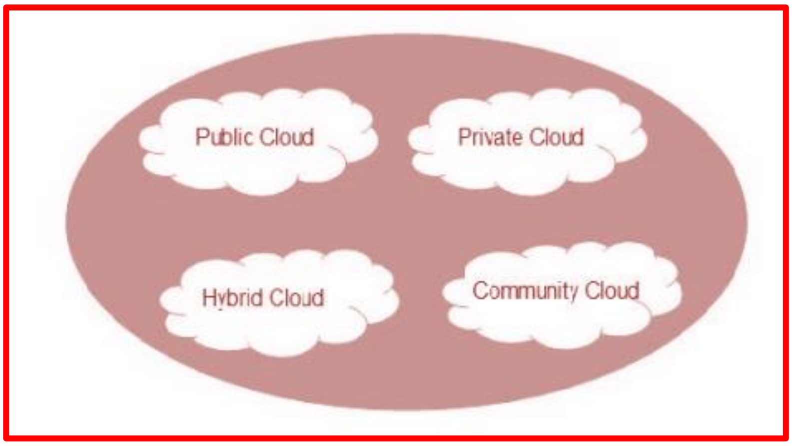

Fig.-1 Development Model of Cloud Computing

There are several deployment models and service models that make cloud computing feasible and accessible to end users [4]. These models include Public, Private, Hybrid, and Community. Public clouds allow systems and services to be easily accessible to the general public, while Private clouds provide increased security. Community clouds allow systems and services to be accessible by groups of organizations, and Hybrid clouds combine public and private clouds for critical activities [5].

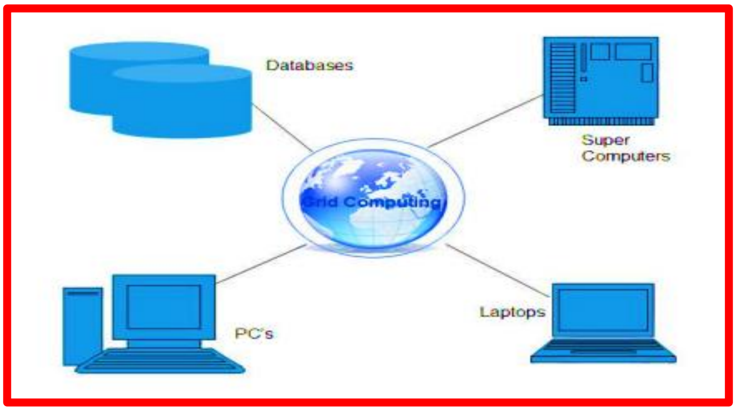

Fig.-2 worldwide Grid information services

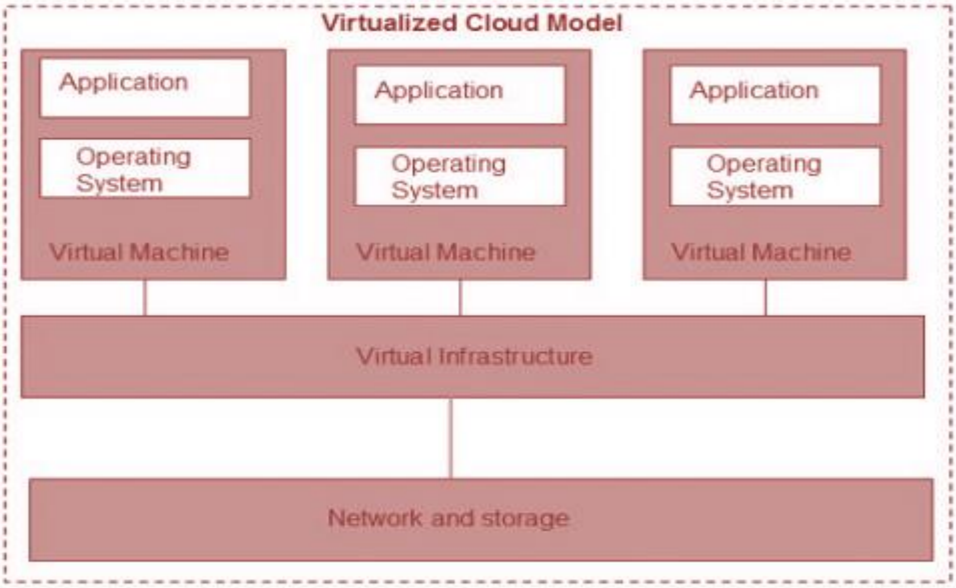

Fig.-3 Virtualized Cloud Computing Model

REDVET - Revista electrónica de Veterinaria - ISSN 1695-7504
Vol 25, No.2 (2024)
http://www.veterinaria.org
Article Received:10/07/2024 Revised:22/07/2024 Published:01/08/2024

Cloud computing is based on three service models: Infrastructure as a Service (IaaS), Platform as a Service (PaaS), and Software as a Service (SaaS). IaaS provides access to fundamental resources like physical machines, virtual machines, and virtual storage. PaaS provides the runtime environment for applications, development and deployment tools, and software applications as a service to end users [6]. The concept of cloud computing began in 1950 with the implementation of mainframe computers. It has evolved from static clients to dynamic ones, from software to services [7]. The cloud computing deployment phase focuses on selecting a cloud provider based on Service Level Agreements (SLA). Maintenance and technical services are provided by the cloud provider to ensure service quality [8]. Cloud computing platforms use technologies like virtualization, Service-Oriented Architecture (SOA), Grid Computing, and Utility Computing to make them flexible, reliable, and usable [9].

Fig.-4 cloud computing of the future, delivering cloud's computing

Research Methodology and Implementation

The emergence of Data Grids in scientific and commercial settings has led to various systems offering solutions for dealing with distributed data-intensive applications [10]. However, this has led to confusion in pinpointing their exact target areas. A taxonomy of Data Grids is provided to break down the overall research into specialized areas and categorize each one in turn [11]. The properties of a Data Grid are determined by its underlying organization, which includes the placement, replication, and propagation of data, as well as the interaction of users with the infrastructure [12]. Core mechanisms such as data transport, data replication, and resource management define the capabilities of a Data Grid. This taxonomy is split into four sub-taxonomies: Data Grid organization, transport technologies, metadata, and resource allocation and scheduling research [13]. Data Grid organization is based on various organizational characteristics of Data Grid projects. Models are the manner in which data sources are organized in a system, and four common models are: monadic, hierarchical, federation, and hybrid [14]. Monadic models have a central repository for accessing data, while hierarchical models have a single source for data distribution across collaborations worldwide. Federation models have a federation model for data sharing, while hybrid models have a hybrid model for data sharing [15].

Data Grids are systems that share data within existing databases, often referred to as federations. These systems allow researchers to request data from any database within the federation, provided they have proper authentication [16]. The degree of integration varies, with different types of federations using the Storage Resource Broker. Hybrid models, which combine these models, are emerging as Data Grids mature and become more widely used [17]. The scope of a Data Grid can vary, either restricted to a single domain or a common infrastructure for various scientific areas. Virtual Organizations (VOs) are the design of VOs, reflecting the social organization of the Data Grid [18]. This chapter discusses the methodology of data grids scheduling in cloud services quality. It aims to discuss the research purpose, approach, data collection and sample selection methods, research techniques, and data analysis methods [19]. The research focuses on factors such as technology adoption, perceived ease of use, and perceived usefulness of services. The data process involves introducing two data sets, analyzing their properties, and preparing them for analysis. The data is then divided into two equal parts, with the training set being the odd number data points and the test validation set being the even number data points. The model is built using various techniques, such as the Multiple Linear Regression, and validated using the test validation data set [20]. The research uses four strong model adequacy criteria to compare the methods of all techniques, ensuring the best prediction technique algorithm for each data set.

REDVET - Revista electrónica de Veterinaria - ISSN 1695-7504
Vol 25, No.2 (2024)
http://www.veterinaria.org
Article Received:10/07/2024 Revised:22/07/2024 Published:01/08/2024

Cloud computing environments offer virtualized resources that can be provisioned dynamically, but users are charged on a pay-per-use basis [21]. This can result in large data retrieval and execution costs when scheduled. To optimize execution time, data transfers between resources and execution costs must also be considered. Cloud computing is a new paradigm for distributed computing that delivers infrastructure, platform, and software as services, available as subscription-based services in a pay-as-you-go model [22]. Particle Swarm Optimization (PSO) is a self-adaptive global search-based optimization technique that relies on the social behavior of particles. PSO has become popular due to its simplicity and effectiveness in a wide range of applications with low computational cost. It has been applied to solving NP-Hard problems like scheduling and task allocation [23].
In India, new cloud computing providers are opening offices to provide cloud hosting services for small and medium businesses [24]. Top companies in India, such as TCS, Zenith Info Tech, Net Magic Solutions, Tata Communications, Synage, and Reliance Data Center, are offering PaaS, IaaS, and SaaS services at reasonable and low prices. This work uses two unique data sets, TCS and Infosys, and MATLAB software for all analyses [25].

Data Analysis and Report
The analysis of data and results using the Analysis of Variance (ANOVA) method. ANOVA is an advanced statistical method used to determine differences among several population means [26]. It requires independent random sampling from each population and assumes that the populations are normally distributed. The test statistic for ANOVA follows an F distribution with two degrees of freedom: numerator (r-1) and denominator (n-r). If the null hypothesis is rejected, further statistical tests may be needed [27].

Category	Var	Measure	Description
Input	X1	Products	Number of product page views
	X2	Lists	Number of product lists views
	X3	Personal	Number of personal list views
	X4	Space uses history	Number of space history page views
	X5	Search	Number of search conducted
	X6	Promotion	Number of promotional page views
	X7	Recipe	Number of recipe page views
	X8	Checkout	Number of checkout pages
	X9	Help	Number of help page views
Output	Y1	Basket size	Number of items at checkout

Fig.-5 Input and Output Variables for Cloud Web Portal Efficiency Measurement

The level of significance for the analysis is 0.05. The F distribution is used to determine differences between two population variances, with appropriate degrees of freedom for the numerator and denominator. The numerator is mean square treatment (MSTR) with r-1 degrees of freedom, and the denominator is mean square error (MSE) with n-r degrees of freedom. The test statistic for ANOVA is F(r-1, n-r) = MSTR / MSE. When the null hypothesis is true, all r population means are equal, with expected values of mean squares being E (MSE) = ó2 and E (MSTR) = ó2 + Óni (µi – µ)2 / (r-1). If differences exist among the r population means, MSTR will tend to be larger than MSE. The ANOVA test was conducted on five categories grouped by the two populations to find variance and determine the significance of the variance.

REDVET - Revista electrónica de Veterinaria - ISSN 1695-7504
Vol 25, No.2 (2024)
http://www.veterinaria.org
Article Received:10/07/2024 Revised:22/07/2024 Published:01/08/2024

	Efficiency Overall	*EfficiencyDesign1*	*EfficiencyDesign2*
Minimum	0.000	0.000	0.211
Maximum	1.000	1.000	1.000
Mean	0.648	0.663	0.766
Std Deviation	0.187	0.184	0.186
1st_Quartile	0.522	0.540	0.640
Median	0.644	0.796	0.928
3rd_ Quartile	0.785	0.660	0.782
Total	5383	4477	604

Fig.-6 Efficiency Scores Summary Statistics

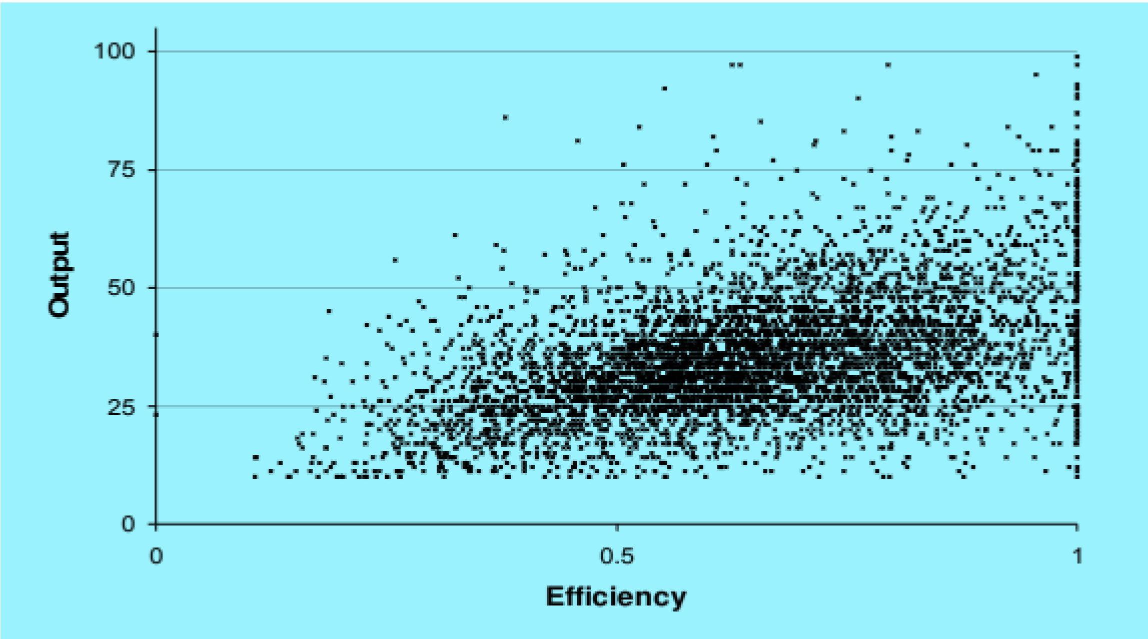

Fig.-7 Cloud Web Portal Qualities Efficiency Scores by Output

The results showed that there was a significant difference in the perception of the tangibility, reliability, responsiveness, credibility, and security aspects of the services provided by the two companies. The F ratio was higher than the critical value, indicating a rejection of the null hypothesis. The reliability data set showed a higher F ratio than the critical value, indicating a rejection of the null hypothesis. The credibility data set showed a lower F ratio than the critical value, indicating an acceptance of the null hypothesis. The security data set showed no significant difference in the perception

REDVET - Revista electrónica de Veterinaria - ISSN 1695-7504
Vol 25, No.2 (2024)
http://www.veterinaria.org
Article Received:10/07/2024 Revised:22/07/2024 Published:01/08/2024

of the services provided by the two companies. The results suggest that the tangibility, reliability, responsiveness, credibility, and security aspects of the services are not significantly different.

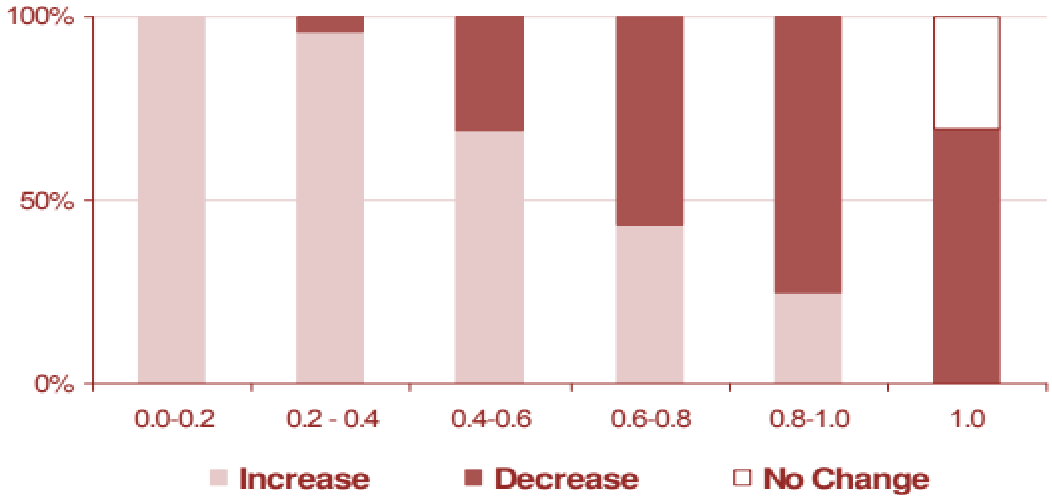

Fig.-8 Cloud Web Portal Qualities Efficiency Score Changes

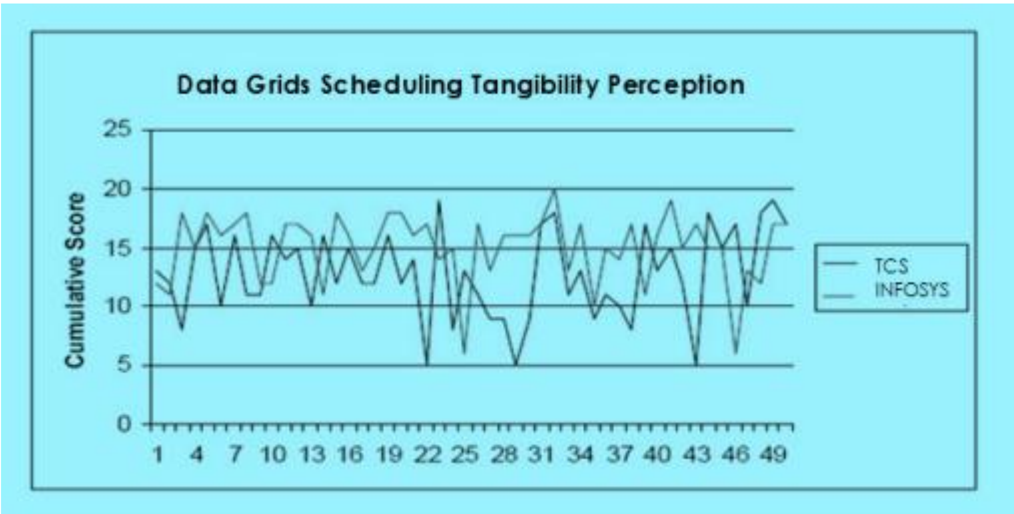

Fig.-9 Observation of Tangibility

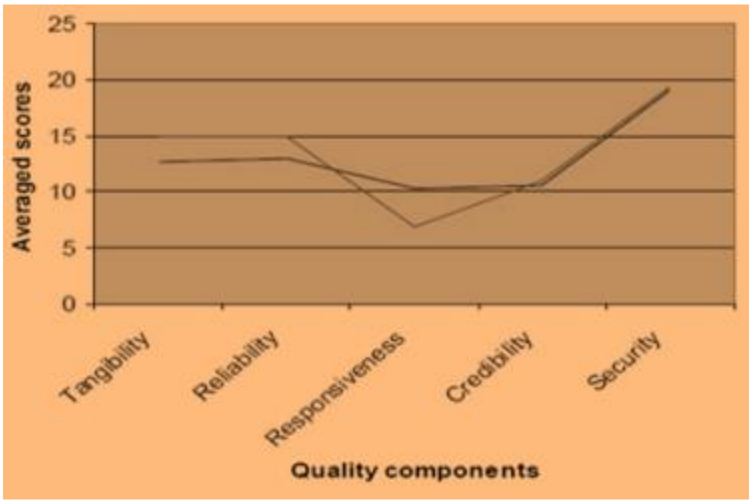

Fig.-10 Averaged Attributes scores

Data and findings were analysed using the Analysis of Variance (ANOVA) approach to identify variations in population means. The test statistic had two degrees of freedom and was distributed according to a F distribution. 0.05 was the significance level. Five categories were classified by the two populations and subjected to the ANOVA test. The findings revealed a notable disparity in how the two businesses' services were perceived in terms of tangibility, dependability, responsiveness, credibility, and security. The null hypothesis was rejected when the F ratio exceeded the crucial value. The null hypothesis was rejected based on the reliability data set's higher F ratio. There was no discernible difference in the security data set.

Conclusion

The study focuses on how businesses perceive the quality of data storage provided by cloud web portal services for internet-based cloud data storage grid scheduling services. The study discovered notable differences in characteristics including tangibility, reliability, responsiveness, credibility, and data security between the two sample populations (Infosys and TCS). Reliability had a more varied response from clients, with Infosys scoring higher than TCS. TCS outperformed Infosys in terms of responsiveness, indicating that smaller businesses are more receptive. Both businesses earned mean scores above 10 in data security and credibility, showing excellent levels of cloud storage credibility. The results were validated using ANOVA testing, which showed no discernible difference between the two primary Cloud Service Providers' perceptions of security and credibility. However, a client's opinion of a company's Cloud web portal attributes was not clearly influenced by any one attribute over another. With an emphasis on their distinctive characteristics—such as high processing demands, geographically dispersed resources, and user collaboration—this thesis investigates Data Grid Scheduling systems in Cloud web portals. Additionally, it looks at resource allocation, data replication systems, data transmission technologies, and the architecture of data grids used in cloud portals. The study shows that by enhancing work turnaround time, application scheduling performance is enhanced when data existence and computational resource availability are taken into account. It is currently unclear if more recent application models can support the current architecture, although the Grid bus broker works well for carrying out scientific applications on Grid resources. The Web Cloud Services Resource Framework (WSRF), which was recently standardized by the Grid community, calls for the broker to build service aggregations and compose services according to their properties in order to fulfil the utility functions of users. According to the study's findings, clients' Cloud web portal attributes show notable variations in tangibility, responsiveness, and dependability among the services provided by websites. For cloud web portal Data Grids or online cloud storage service providers, factors like security and credibility are particularly important. Future research attempts to answer new queries and advance knowledge of cloud web portal data grid scheduling.

Reference:

1. Youseff, L., Butrico, M. and Da Silva, D. (2008). Toward a Unified Ontology of Cloud Computing.In Grid Computing Environments Workshop (GCE '08),Austin, Texas, USA, November 2008, 1-10.
2. Qamar, S., Lal, N., Singh, M., (2010). Internet Ware Cloud Computing: hallenges.(IJCSIS) International Journal of Computer Science and Information Security, Vol. 7, No. 3, March 2010.
3. Wozniak, T., and Ristol, S., Grid and Cloud Computing A Business Perspective on Technology and Applications. Springer Berlin Heidelberg, 2009.
4. J. Sithiyopasakul, T. Archevapanich, S. Sithiyopasakul, A. Lasakul, B. Purahong and C. Benjangkaprasert, "Implementation of Cloud Computing and Internet of Things (IoT) by Performance Evaluation," *2024 12th International Electrical Engineering Congress (iEECON)*, Pattaya, Thailand, 2024, pp. 1-6, doi: 10.1109/iEECON60677.2024.10537945.
5. Kourpas E (2006) Grid Computing: Past, Present and Future – An Innovation Perspective.IBM white paper.
6. Joseph J, Ernest M, Fellenstein C (2004) Evolution of Grid Computing Architecture and Grid Adoption Models.IBM Syst. J. 43(4):624-644
7. S. Kushwaha and A. Rai, "Mobile Cloud Computing: The Future of Cloud," *2024 OPJU International Technology Conference (OTCON) on Smart Computing for Innovation and Advancement in Industry 4.0*, Raigarh, India, 2024, pp. 1-6, doi: 10.1109/OTCON60325.2024.10687896.
8. Foster I, Zhao Y, Raicu I, Lu S (2008) Cloud Computing and Grid Computing 360-Degree Compared. In: Grid Computing Environments Workshop (GCE'08).doi:10.1109/GCE.2008.4738445
9. Cloud Security Alliance. (2009). Security Guidance for Critical Areas of Focus in Cloud Computing.
10. M. Ansari, S. Arshad Ali and M. Alam, "Internet of things (IoT) fusion with cloud computing: current research and future direction", *International Journal of Advanced Technology and Engineering Exploration*, pp. 1812-1845, 2022.
11. Ozsu, M. T. and Valduriez, P. (1999). Principles of distributed database systems. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 2nd edition.
12. Moore, R., Rajasekar, A., and Wan, M. (2005). Data Grids, Digital Libraries and Persistent Archives: An Integrated Approach to Publishing, Sharing and Archiving Datas. Proceedings of the IEEE (Special Issue on Grid Computing), 93(3)

13. N. Kashyap, A. Rana, V. Kansal and Himdweep Walia, "Improve Cloud Based IoT Architecture Layer Security - A Literature Review", *2021 International Conference on Computing Communication and Intelligent Systems (ICCCIS) Greater Noida India*, pp. 112-115, 2021.
14. Mahajan, R., Bellovin, S. M., Floyd, S., Ioannidis, J., Paxson, V., and Shenker, S.(2002). Controlling High Bandwidth Aggregates in the Network. Computer Communications Review, 32(3):62–73.
15. Krauter, K., Buyya, R., and Maheswaran, M. (2002). A taxonomy and survey of grid resource management systems for distributed computing. Software: Practice and Experience (SPE), 32(2):135–164
16. M. Humayun, "Role of Emerging IoT Big Data and Cloud Computing for Real Time Application", *International Journal of Advanced Computer Science and Applications(IJACSA)*, vol. 11, no. 4, pp. 494-506, 2020.
17. Brady, M., Gavaghan, D., Simpson, A., Parada, M. M., and Highnam, R. (2003).Grid Computing: Making the Global Infrastructure a Reality, chapter eDiamond: AGrid-Enabled Federated Database of Annotated Mammograms, pages 923–943. WileyPress, London, UK.
18. Z. Ma, Y. Liu, X. Liu, J. Ma and F. Li, "Privacy-Preserving Outsourced Speech Recognition for Smart IoT Devices", *IEEE Internet of Things Journal*, vol. 6, no. 5, 2019.
19. P. S. Almeida, C. Baquero, N. Preguiça and D. Hutchison, Scalable Bloom Filters, vol. 101, no. 6, pp. 255261, 2007.
20. G. Cormode and S. Muthukrishnan, "An Improved Data Stream Summary: The Count-Min Sketch and its Applications", *Journal of Algorithms*, vol. 55, no. 1, pp. 58-75, 2005.
21. Fox, R. Griffith, A. D. Joseph, R. H. Katz, A. Konwinski, G. Lee, et al., *Above the Clouds*, 2009.
22. G. Malewicz, M. H. Austern, A. J. Bik, J. C. Dehnert, I. Horn, N. Leiser, et al., "Pregel: A System for Large-scale Graph Processing", *Proceedings of the 2010 ACM SIGMOD International Conference on Management of Data*, pp. 135-146, 2010.
23. Z. Zhang and Y. Zhang, "A Survey of Distributed File Systems", *Concurrency and Computation: Practice and Experience*, vol. 26, no. 12, pp. 18341852, 2014.
24. D. Ongaro and J. Ousterhout, *Search of an Understandable Consensus Algorithm. 2014 USENIX Annual Technical Conference (USENIX ATC 14)*, pp. 305-319, 2014.
25. T. Chen, M. Li, Y. Li, M. Lin, N. Wang, M. Wang, et al., MXNet: A Flexible and Efficient Machine Learning Library for Heterogeneous Distributed Systems, 2015.
26. J. Dean and L. A. Barroso, "The Tail at Scale", *Communications of the ACM*, vol. 56, no. 2, pp. 74-80, 2013.
27. J. Dean and S. Ghemawat, MapReduce: A Flexible Data Processing Tool, vol. 53, no. 1, pp. 72-77, 2010.