A Comparative Analysis on Handling Big Data Using Cloud Services

Attique Ur Rehman[1], Sajid Mahmood[2], Ghulam Mustafa[2], Sumet Mehta[3]
[1]Department of Computer Science, School of Systems and Technology, UMT, Lahore, Pakistan
[2]Department of Informatics and Systems, School of Systems and Technology, UMT, Lahore, Pakistan
[3]Jiangsu University Zhenjiang, China
Email: F2019288002@umt.edu.pk, sajid.mahmood@umt.edu.pk, ghulammustafa@umt.edu.pk, msumet@outlook.com

ABSTRACT

In this era of technology, a lot of advancements have been done in almost every field such as medical, science, aerospace and other fields. With the increasing advancements in technology, a lot of data is being produced at the same time. For instances in the field of medicine there is a huge amount of data that is being generated as there are hundreds and thousands of patients who came for their checkup. So now the question arises where this huge amount of data is being stored. This huge amount of data is called as Big Data. And the major problem faced is how to manage and organize this huge amount of data along with its security and not being lost. Big data is used for extracting a lot of useful information but it is not easy to organize it. If the data is being lost than there are a lot of problems that can occur on a huge level as a lot of data being stored in big data is very confidential. This data can be stored on cloud which is the new advancement in the field of technology as it is highly reliable for huge amount of information. So, in this survey paper we will discuss about the solutions of organizing and handling big data proposed by different authors.

KEYWORDS: Big Data, Cloud Computing, Application, Data, Technology, Storage, Confidential, Security and Information.

1. INTRODUCTION

Data from different data sources have been generated at an enormous rate of growth in recent years because of the development and prevalence of information and communication technology (ICT). This vast volume of data is commonly named Big Data, and includes a huge range of unknown and rich data [1]. Big Data services are intended to create enormous potential for trade, business and research by revealing valuable information. Indeed, Big Data services involve a number of sophisticated processes. However, the development of large data services is hindered by a structural lacuna [2]. As data volumes grow exponentially, research and industry need better data management. This database, known as Big Data, is used by numerous organizations to gather precious details and can be computationally processed in order to demonstrate the human activity and actions of various organizational activities, events, developments and connections [3]. It is processed in the cloud owing to the vast volume of data and all processing will be carried out over massive data in the cloud because conventional solutions cannot hold such a large number of data [4].

However, to make accurate decisions, data must be structured, incorporated, protected, and visualized. It is not always beneficial to analyze large data volumes except when properly analyzed. Existing techniques are not enough to analyze the large data and identify the frequent cloud user services. Different services can be integrated for a better working environment. Thus, people are widely vulnerable to exposure with these services [5]. This means that more data is than required that can be collected which could lead to data leakage and therefore security concerns.

Cloud-based data management is a feasible choice for small and medium-sized businesses with respect to the usage of big data processing techniques [6]. Cloud storage provides access to computational services on demand that are often supplied by an external organization and involve little maintenance activities from the company. For cloud computing there are several architectures and deployment models, which can be used for other technologies and design methods. Owners of SMEs who cannot afford to implement distributed NAS technologies should use a variety of cloud infrastructure frameworks to satisfy their big data requirements [7].

2. BACKGROUND AND RELATED WORKS

One of the major sources of handling big data is cloud, which uses different methods of encryption for handling it. But in this advanced era the techniques used by cloud are not enough to secure the big data. And this problem will be solved by dividing this big data into small pieces which will be stored on several storage service providers of cloud. This solution results in saving time, money and cost as data is very big so the owners will save the cost of encryption by this method. Comparisons, theoretical evidence and simulations have demonstrated that the solution proposed is secure, effective and reliable to protect cloud customers from Big Data [1].

The era in which we are living, there are a lot of advancements being done in almost every field which means that a huge amount of data is also being generated known as big data which is a bit difficult to handle as it requires large space for being stored along with time. This paper provides various case studies on big data [7]. Differences between large databases and small databases, multiple systems, major database

frameworks, big data classes. This paper is dissected for the handling of big data by data cloud and layered architecture model. Big data technologies are often used as they are helpful in collecting knowledge by Tools that help to handle infrastructure [8].

Big data is a technology which is used frequently but due to its high cost and space taking problem some companies can't afford it due to lack of money and this problem is solved using cloud computing. Small companies can access to big data technology using cloud computing as it will reduce the cost of hardware and processing and will help in checking the efficiency of big data. This will help small companies to become better and launch new products [9].

Big data is not a novel concept for us, but it is challenging to handle a vast amount of data on a cloud storage network. The approach provided in order to solve the question is Map Reduce strategies and applications [10]. Map Reduce strategies and applications are the method recommended for resolving the matter.

Using big data through the path of cloud computing is good enough but these two technologies have their own rules and principles which are not the same. Moreover, companies also face security issues when they move to cloud. Hadoop and map reduce as big data systems, these two concepts will help in handling bulk amount of data [11]. We were able to get the benefits of data storage and processing power with minimal effort by using Hadoop, map reduction and integration. We can also reach optimum efficiency due to low complexity [12].

Map programming reduction does not bring big data much advantage. Spark may be used to overcome this problem. Much nice is found spark than map reduce programming [13].

With the increase in the advancement of technologies, the amount of data is also increasing and for its storage it's been loaded on cloud but the main issue is that this huge amount of data should not be jumbled up and should be maintained properly on cloud which is not easy along with the analysis of data which is another hard task and which can result in the leakage of data. The suggested solution was that we could use many services, such as chart, table, picture, etc. and the Map Reduce algorithm, to solve this problem [14]. To achieve a better result, various services and tools are combined. Also, for this purpose, Map Reduce Algorithm is used.

Big data known for storing huge amount of data and cloud computing which gives favors over internet are two major technologies being used frequently. BDAAS is another technology which give benefits over internet but to understand this technology we need a proper knowledge of it along with Hadoop system [15]. Basically, this technology BDAAS is used through Hadoop system by representing information through graphs. The approach proposed helps collect data through built-in graphs and social media to support our data management.

A lot of data is available today. This knowledge needs to be accurate and secure position to be rescued, VM Hadoop for this reason is being used but often people don't learn about it. We will train and help people learn about the advantages of VM Hadoop as a large-scale application creation tool. This study covers VM, MRPIV etc., which may lead to accelerating solutions by users and users can get maximum capacity for storage big data is helping in many fields but due to some gaps further development is getting difficuilt.to remove this wall of hindrance, a four-layer cloud-based network architecture is used which gives well organized perspective for accruing, storing and recovering of Big Data [16]. This has given researchers and engineers many opportunities in a lot of different fields too.in future they are going to merge this network with routing protocols and experiments will be carried out in Network Simulator 3.

In Big Data, data is stored over the internet and as a result several problems will be faced at the time of transfer. The solution proposed was to build problem model and then identify the problems and so we can control maximum issues [17].

Nowadays large amount of data is present and to store this data we need a trustworthy platform. Fault tree analyzer gives us reliable environment future FTA will be analyzed more deeply to ensure the reliability and data processing of BDS [18].

To manage and control data in real time, a network butterfly environment is made. In this environment different technologies are gathered and processed. We can also use this to improve employment, quality of life, reducing social anxiety, and get maximum productivity [19].

Face recognition and tagging is very popular now a days in social media, but we don't know how they work. So, for this purpose we will do thorough study [20]. Face tagging is done with the advancements of framework in face recognition which can be used for face authentication in future.

Present architecture of handling big data is not much advanced so we will educate people so they will make further tools for advancements in this field and hence best results will come out of it [21].

The systems efficiency is reduced when computationally intensive and IO intensive work together. This problem is solved by remodeling the big data systems based on cloud and hence result in efficiency of our data.

When the cloud's infrastructure is used by many customers at the same time then there are many chances that the performance of the system will be not as better. This problem is controlled by a virtual machine which uses balancing load and having large storage for big data and fulfill the requirements of multi tenants. Also, security issues will be handled in future [22].

3. BIG DATA CHALLENGES

With the increasing advancements in the technology in almost every filed around the world refers to the increase in data and information. And handling this data is not an easy task. Also, this data is not a normal or a usual data rather it is a very confidential and useful data which in case if leaked can cause a lot of damage. For instance, if we talk about the field of army there is a lot of confidential information which needs to be kept secret and private. In case of leakage of this data can cause harm to the whole country. Similarly, if we talk about the medical field, it contains data regarding the patient's history and disease and if it is being lost then the diagnosis of the patient cannot be done. So, this huge amount of data is known as Big Data. Big data exits in almost every field and it's not easy to handle. It needs a lot of security, protection and encryption so that if the data might get leak it is not decoded easily by any lay person.

4. COMPARATIVE ANALYSIS OF BIG DATA

Big data is a vast number of data which can't be stored on regular software's. Its size can be from few dozen terabytes to many petabytes of data in data centers which requires specialized security system to secure it.

Big data over cloud is not that much secured by using encryption. Big data can be secured by dividing them into parts and storing them on multiple cloud storage service providers [6]. This scheme will use trapdoor function. Analysis of this scheme proves that big data can be stored over cloud through this method and can be effective and feasible for cloud tenants to protect big data.

Advances in information technology and its broad growth are leading to explosion of information / data in several areas of business, engineering, medical and technical studies. Data management and processing is an emerging trend known as Big Data Computing; a new paradigm combining large-scale computing, new data-intensive techniques and mathematical models to build data analyses. The discovery and decision-making of this fast-growing voluminously data are a demanding task [9]. Big data mining requires a large amount of data gathering and retrieval capacity and networking that can be supported by on-site or cloud infrastructures.

In the fields of healthcare research, economic intelligence, social networking, and science discovery we have listed many case studies on Big Data. We also explain the disparity between Big Data databases and conventional databases and address BASE resources.

Big Data is a technique for data mining that is driven by modern technical and architectural innovations. Big data, however, involves a huge amount of hardware and processing resources, making it prohibitive for small and medium-sized enterprises to take on large-scale data technologies [11]. Cloud Calculation offers small and medium-sized businesses the promise of Big Data implementation.

Cloud storage enables SMEs, with a decreased investment to client capital, to incorporate Big Data Technologies. Cost savings on hardware, processing and ability to experience with the big data technology before taking a major share of company resources can be used to implement large-scale technology via cloud computing [14]. The businesses are willing to accept different types of cloud storage systems, each model providing a compromise of cost reduction and data protection and control failure.

This paper described a systematic survey flow in cloud computing for big data processing. We discussed key issues, including cloud storage and computer architecture, popular parallel workflows, major applications and Map Reduce optimization. Big Data is a really complex, not a fresh idea. It calls for a scalable storage index and for a distributed approach to get the necessary findings in real time. The fact that the data is too big for traditional processing is important. Nevertheless, Big Data is dynamic and constantly resides in all the great problems for humanity [16].

The 5V Big Data volume, velocity, variety, veracity, value Figure 1 make it a challenge for traditional data storage companies to manage and analyze the data. Cloud computing appears to be a perfect way to host large data workloads. But working on big data in the cloud is a challenge to reconcile two contradictory concepts. The philosophy of aggregation and resource pooling is focused on cloud storage, but massive data networks (such as Hadoop) are founded on the premise of nothing sharing, where every node is separate and autonomous.

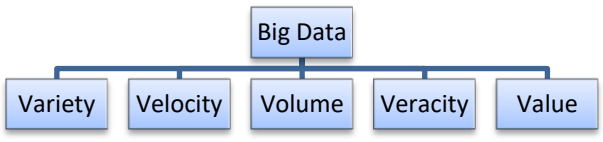

Figure 1: Vs of Big Data

This article introduced the basic principles and meanings of Big Data and explained how data storage transfers to the cloud. The paper presented the 'Map reducing' and the 'Hadoop' as big-data systems to enable large data sets to be processed in a distributed computer setting [18]. In this article, the attributes, patterns and problems of big data were also clarified.

Nowadays big data over cloud is being used. Big data is used to handle large amount of data through IT architecture. Many reports on big data have declared that it will grow between 40 to 60% every year.

Google was creating million searches on few million queries in 1998 and hence will create 1.2 trillion searches in one year. Netflix has also used big data firstly using crowd sourced algorithm and after that machine learning algorithms and hence got a boost by a factor of 4. Other companies are also adopting big data i.e., Harrah's or Tesco getting success by using Hadoop infrastructure on big data. Spark can also be discussed for the advancements in big data analytics.

Through all these studies we got a conclusion that big data has improved performance of companies and increased their profit rate. Also, that spark is better than map reduce programming.

Because of the large data volume, it is stored in the cloud. All analysis is done via big data via the cloud, because it is not possible to store large quantities of data on traditional systems. However, to make an effective choice, the data must be optimized, integrated, secured and visualized [19]. Analysis of the large data volume is not always helpful unless properly analyzed. Existing techniques are not enough to analyze big data and detect frequent cloud user's services.

Big Data Analytics utilizes predictive analytics to process broad and complex data comprising different content styles. The confidentiality and safety of these data is a priority. Powerful encryption protocols must then be enforced to ensure data is protected and none of the clusters and nodes can be hacked. Map Reduce Algorithm helps to keep the log and tells the user about the cloud service [20]. Time is successful.

A high-level cluster infrastructure that can handle high volume, high speed and different formats is required in the Big Data environment. Furthermore, the information collected is huge and not usually ready for analyzes. There are therefore a number of common steps, including data extraction, cleaning and modelling, carried out by all firms. In addition, Big Data tools are dynamically expanding and users might need to manage the environment flexibly.

This paper introduced a solution on the Hadoop ecosystem, which collects data and constructs graphs using the Graph Builder library from social media and databases. The application case we addressed involves the creation of a graph dependent on tweets for a deputy credibility. In the next step, following the review of existing techniques, we will try to improve semantic analysis and finish using Graph Lab (e.g., determine the area of the Deputy popularity) by adding a graphic analysis step [21].

Data volumes consumed by many applications have reached a high level of data through the rapid growth of information technology which significantly increases storage and computation capacity needs.

The current study is developing the personal, small sized, big data platform, which can replicate a VM Hadoop system effectively. The virtual environment technology of Virtual Box is designed and implemented by developers to easily design Hadoop Mapp / Reduce programming. The VM Hadoop, small cluster Hadoop and large Hadoop cluster of the NCHC, Braavos, make the benchmark for this study. Based on the tests, the VM Hadoop is an optimal framework to build and check the Map / Reduce application and the Braavos Hadoop cluster is most suited for manufacturing phase. The VM Hadoop is eligible for use for consumers in terms of a Large Data creation tool.

A huge source of knowledge and innovation, of high economic value, is the amount of data these things have. Global companies such as Google and Amazon have already shown that new popular services based on intelligent use of massive data sets can transform data to economic value. Due to its indistinctive nature and its sophisticated processes, it is difficult to collect, store, search, share, process, analyze, and visualize big data seamlessly and effectively.

Big data may be divided into two approaches in academia and companies. Firstly, how to collect and store it efficiently and secondly, how to understand it. The problem of data collection is certainly overlooked although both problems are of importance. Big data includes the 4Vs shown in Figure 2.

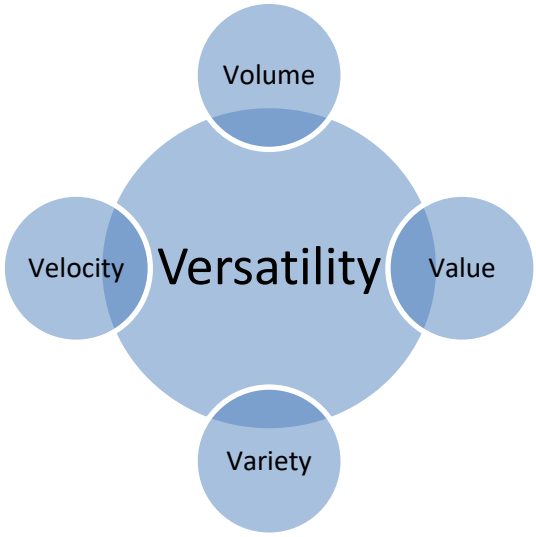

Figure 2: Vs of big data with the support of communication versatility.

The solution proposed consists of the four-layer Big Data services network architecture that outlines the process structure involved in Big Data services. The protocols and algorithms that can be built for each layer are added. It provided the general big data analysis and

services architecture and illustrated the multidisciplinary support requirements. A detailed guide to the four-layer network design is introduced, including data storage, data generation, data management and database generation, a large-scale server infrastructure and a large-scale data collection network architecture [22].

Big data applications store data sets by sharing data centers in the cloud environment, but dynamic change is required over time in big data applications. Faced with multiple data centers, such applications are confronting new challenges in data migration, mainly how to decrease the number of network accesses, reduce the total time consumption and improve efficiency by balancing the global burden of migration.

Given the Large Data Architecture issue in the cloud computing world, we have applied the dynamic data delivery approach and dynamic data transformation algorithm through environmental analysis and application analysis and then conducted simulation experiments in CloudSim. It significantly improves the efficiency of the task in comparison with the original distribution.

5. FAULT TREE ANALYSIS

Big data is becoming center of attraction after cloud computing. Big data is spreading very fast, it can tolerate many failures but still some failures are not acceptable. Reliability is one of them.

Jun Wang made the reliability model which can be used to analyze the rate of loss of data and repair of systems under different schemes of data layouts. Yoshinobu gives a model of jump diffusion which has 2-dimensional Wiener processes. Chen gives an assessment model to access IOT for the improvement of work of reliability.

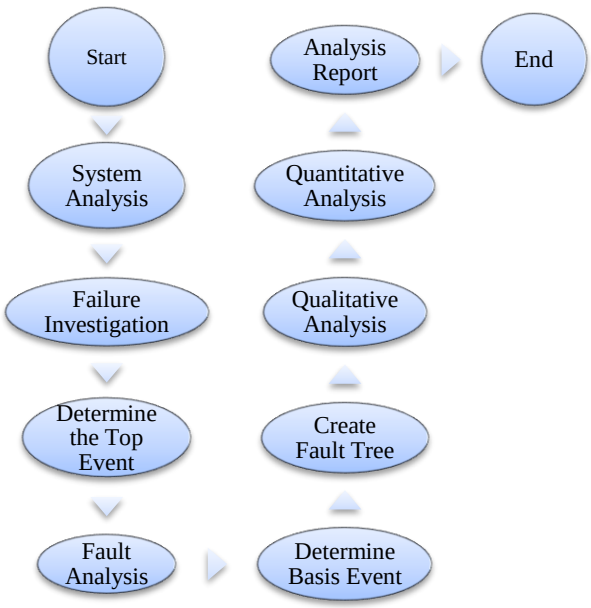

Figure 3: Flow chart of fault tree analysis.

Fault tree can be used to build reliability on cloud. FTA will be used for fault processing and quality assurance. Figure 3.

By utilizing analysis of fault tree based on reliability research, fault handling and reliability can be guaranteed.

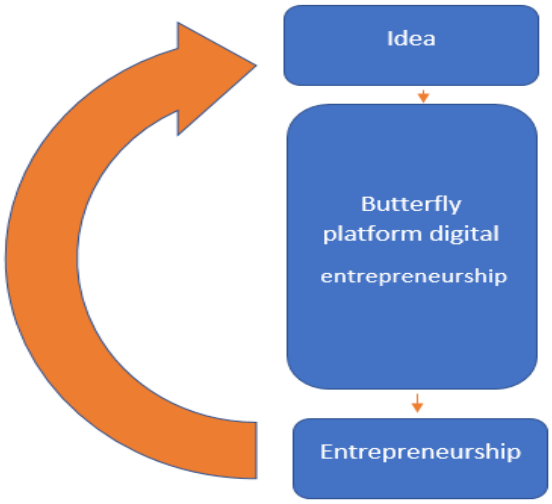

Figure 4: A flow chart diagram on digital entrepreneurship in cloud environment.

The world was confronted by the phenomena of big data in the development of communication technologies (internet technologies) and information technology (different types of electronic services). In addition, this data pool is controlled, managed and processed for a longer and longer term than the capability of software tools. The rise of cloud computing and the storing of cloud data is the key trigger of big data creation.

A network of modern company application infrastructure is provided in the cloud, focused on big data. Figure 4 With this cloud-based platform, the world's leading brands have a high value and have several thousand companies built in for a short time. The primary aim of the platform is to utilize data acquisition, retrieval, search and visualization technology, in order to capture, exploration and obtain exposure from highly varied data in order to create digital entrepreneurship to the secret trends and similarities. Facial recognition is getting famous from two decades. The human brain can recognize patterns which motivated scientist to used different techniques for FR. Figure 5 In this paper a behavioral perspective of FR is discussed to achieve novel FR based on deep machine learning to perform face tagging FR is done based on identification of nose, eyes, jawline etc. Samsung smart TVs and DARPA are also using this technology. Other FR can be done using this technique i.e., face authentication or access control.

IOT techniques when used with big data enhances decision making capability and gives profits in businesses. IOT allow users to make powerful industrial system by connecting large number of smart devices.

These devices can control processes, decisions, communication and cooperation. But a big challenge is waiting ahead in this enhancement of how to collect, process and analyze big data.

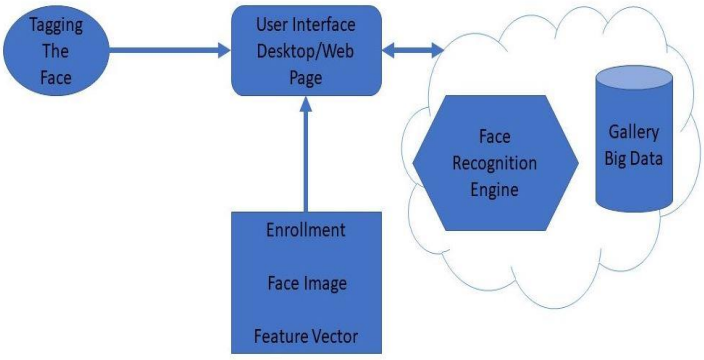

Figure 5: Diagram of the purposed methodology

For example, if big data is used effectively, in US a health care industry can alone make US$300 billion each year.

To handle all this data the industries should upgrade their systems i.e., those features that need to be developed for domain specific challenges in CPS deployment. The first and the most important step in this is handling and processing of big data.

Other important challenges are big data optimization and development tool designs which should be solved in future.

6. BIG DATA WITH CLOUD ENVIRONMENT

New advancements in big data over cloud computing of exploring datasets at cheap costs is causing large industries like Amazon, Google, Yahoo, Facebook and IBM to deploy big data and gain more customers. Small industries also want to enjoy the technologies of big data over cloud but are unable due to its cost. Deep Analysis of big data over cloud is suggested which can manage cost at a large scale. But still real-world applicability is unclear.

MapReduce is a pipeline-based framework which merges computations operation like mapping and reducing with input output operations like shuffling. These operations together will greatly affect the performance of virtual machines. Moreover 90% of our daily tasks involves these operations.

Initially cloud services use VMs but to handle the increasing load more servers are added. Moreover, virtualized services that are stable should be provided by the cloud providers but still it is unclear that big data application on cloud will be fine.

To overcome this problem, we will remodel the resource provision problem that will smartly differentiate the jobs of MapReduce and VMs. high level VMs can get better results than a large amount of low level VMs with same capacity.

In a process of 35.7 Gigabits of data on Amazon EC2 a conclusion is drawn that 2 extra-large instances can do MapReduce job 10 times faster than 16 small ones.

The cloud mostly doesn't face such ambiguities as a task that is both computational and input output intensive are very rare.

Evaluation of this work shows a positive result i.e., it will accurately estimate the time of job completion and improves user experience.

Cloud computing provides space to big data to store and process their large amount of data. Big data needs High-performance computing (HPC) which is also supported by Cloud computing. But due to multitenant environment sometimes it effects on systems performance and it can't be controlled by the user as it is not under his control. And as the time passes the available infrastructure will not be able to handle too much load. This is a very big challenge. Dynamic Scheduler (DySc) is used to deal with multitenant environment of cloud computing which will schedule and reschedule the policies to avoid exhausted resources.

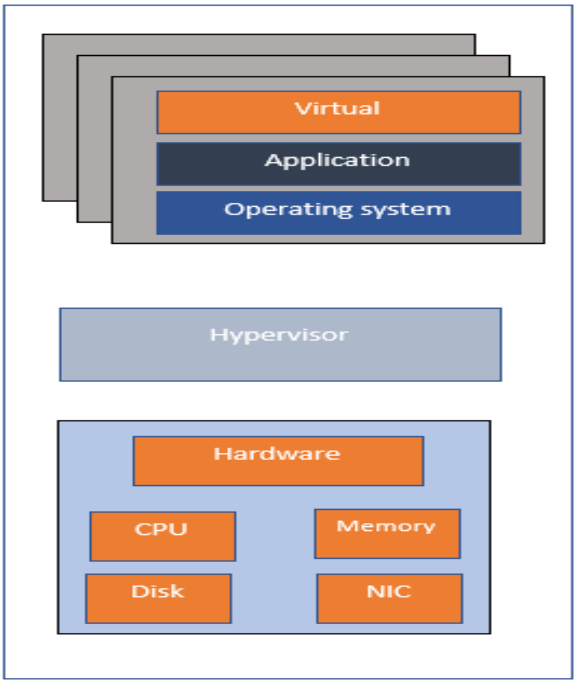

Figure 6: A commonly shared environment of cloud computing

SLA i.e., Service Level Agreement is an agreement made between customer and cloud services which means that the customer will only monitor virtually so if his system will face performance variability, he will not

be notified which is a very big issue.

To overcome the infrastructure complexity, we can provide horizontal resources and it will duplicate the cluster and hence reduce the complexity. But this approach will face other challenges i.e., client should know that different clusters will process data and will give accurate result and which cluster will perform load balancing is another challenge. Load balancing can be done through SDNs.

big data, Helion Eucalyptus i.e., cloud computing and Floodlight i.e., SDN. SDN will send the packet to controller and hence will define the network in real time as shown in Figure 6. By using this methodology, the system will overall give 22.04% effectiveness. Security is also a challenge which will be discussed in future.

7. CONCLUSION

So, this review paper was about the big data. That in today's era due to increase in technology in every field a lot of useful information and data is being generated which is known as big data. So, to handle this large amount of data, means big data some methodologies and solutions have been proposed. The major technology for handling big data is cloud computing. So, using cloud computing techniques and architectures, the handling of big data becomes easy.

In this review paper, we discussed different works of the authors who used cloud computing techniques for handling big data. We also discussed the challenges that were faced by them. And in the end the authors concluded that weather the proposed architecture helped in handling the big data or not.

REFERENCES

[1] H. Cheng, C. Rong, K. Hwang, … W. W.-C., and undefined 2015, "Secure big data storage and sharing scheme for cloud tenants," *ieeexplore.ieee.org*, Accessed: Mar. 11, 2022. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/7122469/.

[2] R. Kune, P. K. Konugurthi, A. Agarwal, R. R. Chillarige, and R. Buyya, "The anatomy of big data computing," *Softw. - Pract. Exp.*, vol. 46, no. 1, pp. 79–105, Jan. 2016, doi: 10.1002/SPE.2374.

[3] B. P.-J. of T. Research and undefined 2014, "Big data using cloud computing," *Citeseer*, Accessed: Mar. 11, 2022. [Online]. Available: https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.476.8238&rep=rep1&type=pdf.

[4] C. Ji, Y. Li, W. Qiu, U. Awada, K. L.-2012 12th international, and undefined 2012, "Big data processing in cloud computing environments," *ieeexplore.ieee.org*, 2012, doi: 10.1109/I-SPAN.2012.9.

[5] S. El-Seoud, H. El-Sofany, … M. A.-… J. of I., and undefined 2017, "Big Data and Cloud Computing: Trends and Challenges.," *researchgate.net*, 2017, doi: 10.3991/ijim.v11i2.6561.

[6] T. Perumal, K. B. Prakash, A. Rajaraman, T. Perumal, and P. Kolla, "Foundations to frontiers of big data analytics," *ieeexplore.ieee.org*, doi: 10.1109/IC3I.2016.7917968.

[7] Vistro, D. M., Rehman, A. U., Abid, A., Farooq, M. S., & Idrees, M. (2020). Cloud Based Architecture For Smart Educational System Using Modern Technology. vol, 7, 1493-1503.

[8] "Data privacy issue in Federated Learning Resolution using Block Chain," 2022, Accessed: Mar. 11, 2022. [Online]. Available: https://vfast.org/journals/index.php/VTSE/article/view/726.

[9] R. D.-2017 I. international conference on consumer and undefined 2017, "A review of big data analytics over cloud," *ieeexplore.ieee.org*, Accessed: Mar. 11, 2022. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/8307833/.

[10] S. Yousfi, D. C.-2015 I. C. on Cloud, and undefined 2015, "Big Data-as-a-service solution for building graph social networks," *ieeexplore.ieee.org*, 2015, doi: 10.1109/CloudTech.2015.7337009.

[11] A. Munawar, D. Mago Vistro, A. U. Rehman, S. Mehmood, and M. Idrees, "A Literature Review On Security Issues In Cloud Computing: Opportunities And Challenges," *researchgate.net*, 2020, doi: 10.31838/jcr.07.10.282.

[12] M. Zhao, A. Kumar, G. G. Md Nawaz Ali, P. Chong, and P. Han Joo Chong, "A cloud-based network architecture for big data services," *ieeexplore.ieee.org*, doi: 10.1109/DASC-PICom-DataCom-CyberSciTec.2016.118.

[13] Vistro, D. M., Rehman, A. U., Abid, A., Farooq, M. S., & Idrees, M. (2020). Analysis of cloud computing based blockchain issues and challenges. Journal of Critical Reviews, 7(10), 1482-1492.

[14] D. Jiaman, W. Sichen, … D. Y.-2016 17th I., and undefined 2016, "A dynamic migration method for big data in cloud," *ieeexplore.ieee.org*, Accessed: Mar. 11, 2022. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/7943339/.

[15] S. Ji, Q. Li, W. Cao, P. Zhang, H. M.-A.

Sciences, and undefined 2020, “Quality Assurance Technologies of Big Data Applications: A Systematic Literature Review,” *mdpi.com*, doi: 10.3390/app10228052.

[16] H. Beyadar, M. Askari, A. A.-2016 I. 10th International, and undefined 2016, “A Network platform for creating digital entrepreneurship in cloud environment based on big data,” *ieeexplore.ieee.org*, Accessed: Mar. 11, 2022. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/7991766/.

[17] A. Vinay, V. Shekhar, J. Rituparna, … T. A.-P. C., and undefined 2015, “Cloud based big data analytics framework for face recognition in social networks using machine learning,” *Elsevier*, Accessed: Mar. 11, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1877050915005967.

[18] B. Cheng, J. Zhang, G. P. Hancke, S. Karnouskos, and A. W. Colombo, “Industrial cyberphysical systems: Realizing cloud-based big data infrastructures,” *ieeexplore.ieee.org*, doi: 10.1109/MIE.2017.2788850.

[19] Y. Yuan, H. Wang, D. Wang, J. L.-2013 I. 21st, and undefined 2013, “On interference-aware provisioning for cloud-based big data processing,” *ieeexplore.ieee.org*, Accessed: Mar. 11, 2022. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/6550282/.

[20] Vistro, D. M., Rehman, A. U., Farooq, M. S., Abid, A., & Idrees, M. (2020). A Survey On Cloud Computing Security With Cross Platform. Journal of Critical Reviews, 7(10), 1439-1445.

[21] C. Vicentini, A. Santin, E. Viegas, V. A.-J. of N. and Computer, and undefined 2019, “SDN-based and multitenant-aware resource provisioning mechanism for cloud-based big data streaming,” *Elsevier*, Accessed: Mar. 11, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1084804518303643.

[22] D. Mago Vistro, M. Shoaib Farooq, A. U. Rehman, and S. Malik, “Smart Application Based Blockchain Consensus Protocols: A Systematic Mapping Study,” *atlantis-press.com*, 2021, Accessed: Mar. 11, 2022. [Online]. Available: https://www.atlantis-press.com/article/125960831.pdf.