International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 10 Issue: 12
Article Received: *25 July 2022* ***Revised:*** *12 September 2022* ***Accepted:*** *30 November 2022*

A Comprehensive Study on the Big Data Adaptation and Analytics of Cloud Computing Ecosystems

Priyam Vaghasia[1]
[1]Stevens Institute of Technology
Priyamvaghasia57@gmail.com[1]

Dhruvitkumar Patel[2]
[2]Staten Island Performing Provider System
pateldhruvit2407@gmail.com[2]

ABSTRACT:
Big Data, adaptation and analytics within Cloud Computing ecosystem are integrated in this study as a detail solution of their transformative role in contemporary business environment. Big Data and Cloud Computing convergence allows for the production of dynamic and scalable platforms that can process huge volumes of data for businesses which are able to make enhanced decision making, predictive capabilities and business operations. In this research we look at data storage, data processing frameworks and real time analytics independent of their integration into Cloud and analyze about the key components that make the difference when it comes to Big Data adaptartion. Finally, the study also points out that organizations will face challenges as well as opportunities that exist in the Cloud due to growing data management demands, with regards to issues of data security and privacy issues, and optimizing the infrastructure to support these data demands. It also describes various Cloud based analytical tools and technologies on which meaningful insights can be derived from large datasets. The book discusses key methodologies such as machine learning, artificial intelligence and data mining, which help to show how they are used in Cloud ecosystems to extract the business intelligence, that is, actionable information. Additionally, the paper explores the application of Big Data analytics and Cloud Computing in specific industry such as healthcare, finance, and retail. The study concludes by explaining how the integration of this nature will emerge in future years and how emerging technologies such as Edge Computing and Serverless Architecture are going to shape the next wave of the Big Data and Cloud Computing evolvement. The study presents an exhaustive framework of the intricate relationship between these two vital aspects, which collectively bring tremendous impetus toward the industrial revolution of the ages.

Keywords: Big Data, Cloud Computing, Data Analytics, Machine Learning, Predictive Analytics.

1. INTRODUCTION

Introduction

With the rise of technology, the spread of huge quantities of data has also occurred in many areas. As the definition of Big Data, its volume, variety and velocity, is big in the scope of business and technological landscapes, it can be considered one of the biggest challenges and opportunities. It has seen a tremendous transformation in the way organizations store, process and analyze data by way of it's integration with Cloud Computing ecosystems. By fusion, Big Data and Cloud Computing gives leverage to companies for developing resources for scalable, flexible and cost effective solutions for decision making, efficiency and competitiveness of their operation.

The Cloud Computing is the ideal platform, where we can manage and analyse inconsistent huge datasets with provision of on demand access of resource such as storage, processing power, and applications. The adaptation of Big Data in Cloud presents a number of complexities such as data storage in a secure environment, provision of real time data processing and provision of efficient analytics frameworks. This study focuses on exploring the interaction between Big Data and Cloud Computing to figure out how their coexistence modifies a new ecosystem for data and the problems that organizations deal with in adapting to this interaction and, in particular, the advantages that these organizations can acquire from it. To understand three V's of Cloud Computing refer to Figure 1.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 10 Issue: 12
Article Received: *25 July 2022* ***Revised:*** *12 September 2022* ***Accepted:*** *30 November 2022*

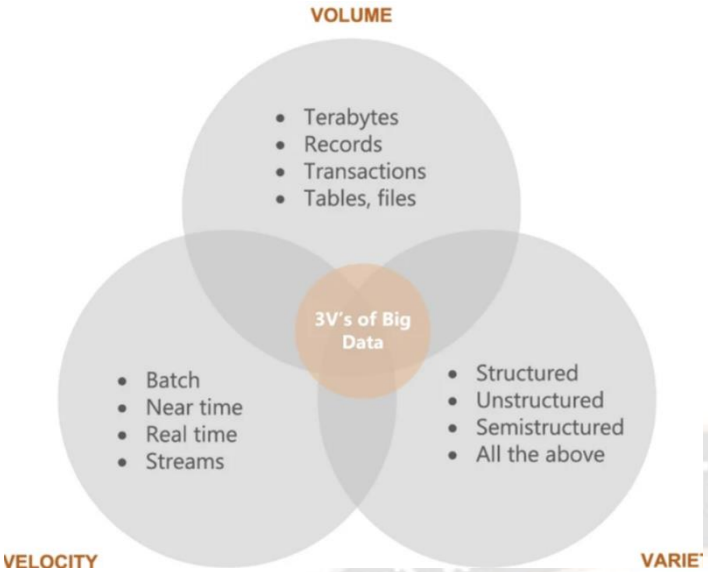

Figure 1: Three V's of Cloud Computing

1.1 Big Data in Cloud Computing

Big Data in the context of Cloud Computing refers to the large-scale management, storage, and processing of data through Cloud infrastructure. Cloud allows the businesses to achieve flexibility to scale up their data processing capacity as per their requirement. One significant advantage of having Big Data integrated into the Cloud ecosystem is that it allows storing and processing data at a much bigger scale than standard on premise systems can manage. Cloud platforms have the infrastructure for companies to manage the flood of information created from collecting more data, like unstructured data from different sources such as IoT devices, social media, and transaction records.

Additionally, Cloud computing gives you real time data processing necessary for the industries that need instant information, like budget, healthcare, and the retail industry, among other areas. The specialized tools, Apache Hadoop, Apache Spark, and NoSQL databases are available to the cloud providers to help handle and process Big Data more efficiently. Using the distributed computing concepts, these tools break the large datasets into smaller piece to be processed in parallel and dramatically speed up and enhance data analytics. Combining Big Data with the Cloud's computational power has brought us sophisticated analytics that inform organizations key, data driven decisions.

However, Big Data Contribution into Cloud is not 'rocket science' and despite so many opportunities it offers, there are certain challenges, the fact that the data security and privacy may be an issue along with regulations compliance that need to be considered. Data is usually kept in their multiple on premise and Cloud locations and organizations must have solid data governance that will secure data accessibility and privacy as well as integrity. To achieve this advanced encryption methods and integration of identity management are necessary to protect sensitive data from being accessed by unauthorized access. On the other hand, high cost for dealing with big scale data storage and processing in Cloud as well as alignment of investments in Cloud infrastructure with the business needs and objectives are another things that have to be taken care of by organizations in the Cloud.

1.2 Cloud-Based Analytical Tools and Technologies

The analytical tools accessed through the cloud have greatly contributed to the evolution of organizations working on data processing and analysis. Cloud-based tool is typically scalable, is automated and can be easily accessed compared to the traditional analytics systems which need significant manual intervention especially on-premise dedicated hardware. Data mining frameworks, machine learning platforms and data visualization software are some of the most common tools that make use of these Clouds and can easily combine with some of the Cloud's infrastructure to enable the organizations to find valuable pieces of information from their data. These are tools that help users to find the patterns, trends, relationships within the dataset and making good data based decision.

Leveraging machine learning algorithms is one of the important benefit of Cloud based analytics tools where Cloud will store all the data required for the analysis. The machine learning services available on the cloud such as Amazon Web Services (AWS), Microsoft Azure, and Google Cloud provide the organizations with the ability to build the predictive models, analyze trends, and forecast what the future outcomes will be based on such historical data. For industries that depend on forecasting such as retail and finance, this capability is very powerful as it enables you to predict what customer will do and what market trends will develop so that you can remain competitive. Furthermore, Cloud platform comes with inbuilt AI features that businesses can leverage to automate the decision making process, reducing the time as well as errors in the decisions taken.

Also, the Cloud analytic tools have a powerful data visualization capability that allows organizations to understand the complex data sets in an easy to digest format. These tools present insights through dashboards, charts, graphs and heatmaps, among other things. Data visualization does more than just have a picture of the data to see trends and anomalies; it also helps other teams to work together with data presented in a common understood fashion. But cloud ecosystems also set up an environment for collaboration and multiple users to work with data in real time regardless of their location or placement, cutting down silos and allowing for cross-functional teamwork. To understand efficient data processing for the power system refer to Figure 2.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 10 Issue: 12
Article Received: *25 July 2022* ***Revised:*** *12 September 2022* ***Accepted:*** *30 November 2022*

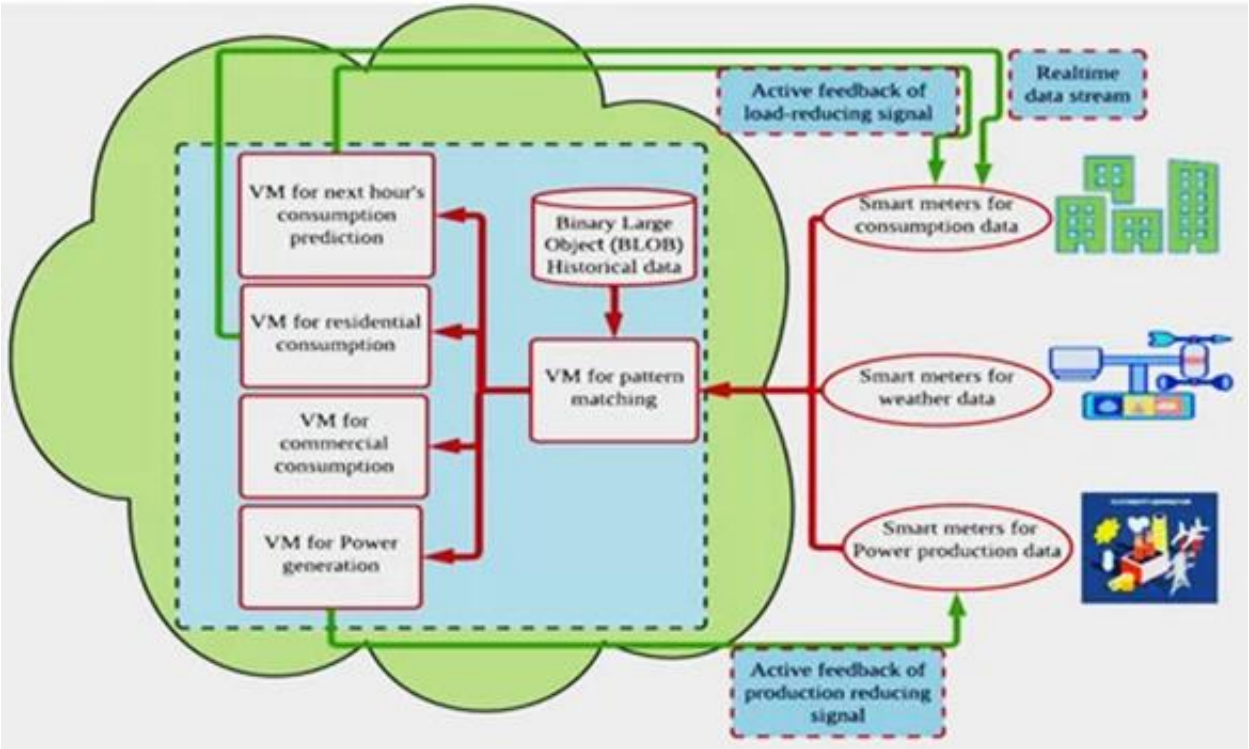

Figure 2: Efficient data processing for the power system.

1.3 Challenges in Big Data and Cloud Integration

The integration of Big Data and Cloud Computing has a lot of benefits but there are great challenges before an organization can leverage the full potential of Big data and cloud computing synergy. The main dilemma is data security. The more companies start storing their sensitive data in the Cloud, the more important it is to ensure that this data is confidential, intact and available. Cyberattacks, data breach, unauthorized access are significant concerns of handling personal data and financial information. However, these new security measures are being developed by cloud providers, and organizations need to put their own in place, also, encryption, multi-factor authentication and regular audits just to name a few, to protect their data.

The second issue is one of data governance, compliance. The number of regulations governing how an organization's data can be processed and stored increases, such as the General Data Protection Regulation (GDPR) in Europe and the California Consumer Privacy Act (CCPA), an organization's data processing and storage in the Cloud must comply with the law. To achieve this, requires developing strong data governance policies including data classification, access control, audit trails and so on. On top of that, managing data sovereignty, which is about moving to ensure that data always stays within a certain region or complies with local laws—another piece of difficulty that comes with adopting Cloud.

There are also several persistent scaling and cost optimization problems. It was Scalability of Cloud platforms but businesses mustn't use Clouds scales to unlimited without any thought to avoid significant expenses. Cloud services' pay as a you go model can create tremendous operational expenses very soon for a large amount of data and heavy analytical workloads. However, organizations have to adopt strategies that will enable them to keep the cost in check while using the Cloud in the optimum way and such strategies include using serverless computing or a hybrid Cloud to get the job done.

1.4 The Role of Machine Learning and Artificial Intelligence

The Big Data analytics on Cloud must follow the path of Machine learning (ML) and artificial intelligence (AI). Using these technologies systems are able to learn from the data, find patterns, and predict value without humans. When combined with Cloud ecosystems, ML algorithms enable businesses to analyze large dataset in quick and efficient manner and obtain insights otherwise nearly impossible via traditional methods. Supervised learning, unsupervised learning as well as reinforcement learning are largely used for such tasks as customer segmentation, fraud detection and predictive maintenance.

AI is being employed in Big Data analytics besides the mere automation. Now, AI models can work with unstructured data (for example, images, text, videos) that have been the problem to which data analysis. These complex models rely on computational power; this is what the cloud platforms give you. You can derive meaning out of new data sources with greater volume. There are many examples of using it, for instance, AI based sentiment analysis tools that help to

understand what customers say on Social media and give the direction to business to modify their strategy according to that. In addition, recommendation systems based in AI are commonly used in e-commerce platforms which analyzes customer behavior to recommend particular product based on customer behavior.

Organisations leveraging the coupling of Cloud based Big Data solutions with ML and AI can take advantage of new opportunity for innovation and efficiency. These technologies enable analyzing the data in real time to make decisions, adaptive systems able to adapt themselves over time, and continuous improvement of business processes itself. Models for predictive maintenance, such as equipment failures, can also be created so that manufacturers know precisely when these failures will occur and when equipment may be down and as a result, cause less downtime and enhance operational efficiency. With ML & AI stepping their game up, the more that it will be integrated with big data and cloud ecosystems will only get more sophisticated, increasing the automation and increasing the value that businesses are going to get from their data.

1.5 Future Prospects and Emerging Trends

With the future expected, Big Data and the integration of Cloud Computing are expected to grow, while so many trends emerge in the industry. Another major happening is the surge in edge computing, where data is processed nearer to it source, for example in IoT gadgets or neighborhood servers, as opposed to depending for all time on centralized Cloud foundation. an implementation of edge computing enables real time processing of data and ultimately helps businesses respond to the real life needs more effectively and with less latency. Edge computing will also become more important as the number of IoT devices continues to increase to allow IoT devices to manage and analyzed Big Data in a distributed manner.

Serverless computing is another trend — as it is totally on demand, businesses pay for the compute resources that they use, without the need to manage the servers. Serverless architecture makes the Cloud easy to adopt, removes server management and the burden of scaling it. It is in line with the rise in the need for more flexible and cost efficient Cloud solutions by organisations that want to minimise operational overheads while achieving the best results. Since Big Data tasks need to handle event-driven, intermittent processing, serverless computing well fits to the tasks that need to be handled.

Furthermore, the big data processing innovation may be transformed in the future with the advent of quantum computing. Current studies around Quantum computing, based on the axioms of quantum mechanics, would give it the capability to exponentially increase the computational power of organizations to process large datasets as never before. Being in its infancy, quantum computing has the potential to revolutionize Big Data analytics, be it in the fields such as drug discovery, financial modeling, and many optimization problems. While these technologies are still in the emerging phase, they will eventually lead to convergence between Big Data and Cloud Computing and ultimate detection of new opportunities in business space across the globe.

2. REVIEW OF WORKS

2.1 Evolution of Big Data and Its Impact on Cloud Computing

This Big Data has emerged due to the fast increase in data generation and now Big Data is playing a vital role in the cloud computing ecosystem. Hillbert and Lopez (2011) report that world technology storage and computation capacity has been increasing significantly over the past years, thus, the ability to process massive data has also increased significantly. These days, with the proliferation of connected devices IoT and digital technologies, the creation of data has made exponential growth and therefore for the management of such data is very important for the organizations. By 2025, they forecast Reinsel, Gantz, and Rydning (2017) that the data volume is going to grow exponentially, and most of it will be stored in cloud infrastructures. It brings the challenge and the opportunity of harnessing Big Data using the cloud based analytics platforms.

Combining big data with cloud computing gives a scalable answer for the management and evaluation of huge data. The accessibility to share resources at any time — also known as on demand — as offered by the cloud computing platforms make it inevitable that companies are moving to the Cloud platforms for their Big Data Analytics (Yadav & Sohal, 2017). Marr (2018) provides the analysis for the great amount of data (daily scale) generated and is needed for systems to efficiently store, process and analyze this data. This allows organizations to use Big Data analytics in order to make better decisions, gain business intelligence, and the analysis of queries to develop predictability (Hellerstein, 2019).

2.2 Key Characteristics of Big Data

The defining characteristics that often describe Big Data are the so called ‘three V’s’: volume, velocity and variety (Gewirtz, 2018). In this case, volume is the simply the sheer magnitude of data generated over multiple domains and data storage scales accordingly. WhishWorks (2019) suggest that vol is the most obvious one, as data continues to accelerate around the world. Velocity is the rate of the data production and processing. Since data streaming is an intrinsically real-world phenomenon, cloud platforms are optimized for the fast processing that are required (Weathington, 2012). Variety entails diverse types of data in the form of

structured data in databases and unstructured data from social media and sensors. Such diversity necessitates sophisticated data integration methods and systems for processing diverse types of data in an efficient way (Wikipedia, 2018).

Having these characteristics has necessitated advances in the area of data analytics technology. Also, the incompleteness of Big Data necessitates robust storage and highly specialized data processing tools as stipulated by Akhtar (2018). To address these issues, cloud computing system has evolved to store and manipulate large datasets with varied formats at high speed providing flexible and scalable infrastructure (Marr, 2018). Similarly, these systems make it possible for businesses to exploit the maximum value that Big Data brings with real-time analytics and insights that were previously out of reach (Kaisler, Armour, & Espinosa, 2013).

2.3 Big Data Analytics Techniques in Cloud Computing

Over the rising complexity of Big Data comes an urgency to be provided with more sophisticated analytics techniques in order to gain the acutest possible insights drawn from large data. In Yadav and Sohal (2017), various tools and methodologies employed in cloud computing for Big Data Analytics like, machine learning, statistical analysis and data mining techniques are explored. And these methodologies are crucial for finding patterns that you might not know the data will lead you to, or predicting, based on the known trends. Thanks to cloud support for distributed computing, data which is large enough to be processed in parallel on a single physical node can be processed in parallel in the cloud (Hellerstein, 2019). Google BigQuery is a cloud based solution which provides power analytics capability needed for big data processing in businesses where makes it possible to access fast and scalable solutions for processing Big Data (Google Cloud, 2020).

Additionally, ETL extract, transform, load process become important, especially when working with various advanced Big Data techniques this requires that their ETL processes be integrated with cloud platforms. LaprinthX (2018) compares ETL (Extract, Load, Transform) with the ELT which manages Big Data in the cloud environments in a way with more freedom. In applying ELT approaches, data processing is far more efficient as it will reduce the amount of time to transform data before storing them, which will speed up, and often with greater accuracy, analytics. Therefore, such practices will enable efficient processing of data and leverage the elasticity of these cloud resources to scale the analytics efforts (Kimball & Ross, 2013).

2.4 The Role of Cloud Computing Providers in Big Data Analytics

Specialized Big Data analytics platforms and services are provided by cloud service providers in order to help them facilitate Big Data analytics. AWS, Microsoft Azure, and Google Cloud all host the suite of high performance analytics tools, leading cloud platforms. WhishWorks (2019) also clearly pointed out that using cloud providers for Big Data handling enables high availability, scalability and economy of cost that is ideal for even the smallest business. The provided infrastructure as a service (IaaS) and platform as a service (PaaS) solutions help organizations to use Big Data capabilities without having much on prem hardware or too much IT expertise.

According to Reinsel et al. (2017), cloud computing not only allows to store and manage large datasets but also enhance analytics and AI driven insights that are able to be incorporated immediately into business processes. Businesses can get flexible and scalable environments for processing data, these are the optimized cloud solutions for Big Data workloads. With the adoption of cloud computing growing, companies will be able to exploit Big Data's full value in predictive analytics, customer insights, and operational optimization (Marr, 2018).

2.5 Challenges and Future Directions of Big Data in Cloud Ecosystems

Big Data has many advantages but bringing it into play with cloud computing is no walk in the park. In their discussion of several obstacles, which include concerns over data privacy, the potential security risks inherent in managing such large and diverse datasets, Kaisler et al. (2013) argue. Cloud computing is the distributed nature which makes issues with data governance and compliance and it concerns organization dealing with sensitive information. Furthermore, data management in multiple cloud zones comes with integration challenges that involve managing data in a synchronized fashion, and more advanced data synchronization and consistency techniques (Weathington, 2012).

The future of Big Data on the cloud ecosystems is only beginning to reach its height. Cloud platforms will advance AI and the ML to the most sophisticated level to handle complex Big Data tasks. Statista (2020) informs us that such numbers will increase and that more advanced Big Data analytics tools will be in high demand. In order to meet the increasing demand of the data in businesses, cloud computing providers are expected to develop more specialized services that will effectively handle the emerging needs such as real-time data processing and edge computing (Reinsel et al., 2017).

3. METHODOLOGY

The research methodology is qualitative with which the aims of this study is to examine the integration and the effect on cloud computing ecosystems through Big Data adaptation and analytics. Instead of spending time in

experimental testing, the research takes the approach of literary analysis of extant literature, industry reports and case studies to evaluate how organizations are using Cloud Computing for Big Data analytics. The study aims to make sense of how adoption and success of Big Data analytics depend on strategic and technological factors within cloud environments, as well as synthesizing secondary data.

The study relies on the secondary data obtained from different journals, books, industry reports, white papers and peer reviewed articles to gather the relevant information. Finally, the literature review has drawn the works of some famous authors from the area of cloud computing and Big Data, for example, Hillbert & Lopez (2011) and Reinsel et al. (2017). Case studies from well known companies and cloud service providers (e.g., Google Cloud (2020), Amazon Web Services) are also included, which illustrate in reality applications and problems. Data from Statista (2020), Forbes (2018) and the others reports are helpful to have on the market trend and the decision making via data to some extent in Big Data analytics.

The study makes use of thematic analysis and seeks to identify and classify some major themes taken from gathered materials to be used for data analysis. Thus, this is a systematic review to extract benefits and challenges and the technological aspects of integrating Big Data and cloud computing process. Some attention is paid to issues of scalability, data privacy, and processing techniques. It compares different cloud platforms and related Big Data solutions for their suitability and are they effective in thousands of data processing.

This study's primary limitation is based on the use of secondary data which limits the ability to acquire firsthand knowledge from the organization that are currently practicing Big Data analytics in cloud ecosystems. Secondary data only gives a generalized look into the trends and practices, but does not exhaustively yield the challenges and revolutions that are actually happening in that industry as they are happening. The study also concerns cloud computing platforms and is not focusing on other types of data management and non-cloud solutions. Although these limitations exist, the methodology guarantees that it provides valuable insight on present implementation of the adaptation of Big Data in cloud computing and creates the basis for future research on this emerging topic.

4. RESULTS AND DISCUSSION

4.1 Adoption of Big Data Analytics in Cloud Ecosystems

The results suggest a steady progression of its adoption in the cloud computing ecosystems. Cloud provides companies with scalability and flexibility of cloud environments which helps in storing, managing and analysing large datasets, therefore companies are increasingly relying on cloud platforms. Big Data solutions are suited through the use of distributed computing frameworks that cloud service providers, Amazon Web Services (AWS), Google Cloud, Microsoft Azure provide that businesses can use to effectively process large volumes of data. In industries like healthcare and finance, which are subjected to big datasets, also adopting these services has been exceptionally high.

4.2 Benefits of Big Data Analytics in Cloud Computing

An advantage of integrating Big Data analytics with cloud computing is the fact that they offer a variety of niches. The enhanced scalability one of the primary identified advantage is that an organization can scale its data storage and processing capabilities based on its needs. Apart from allowing for low initial cost investment in data management with pay per use models, cloud platforms allow for passing on cost rather than incurring them. It also makes it easy to perform big data analytics on the cloud which also allows for real time data processing; useful in businesses that require up to the minute insights on decisions. These advantages help improve the operational efficiency, improve customer experience and better focused marketing strategies.

4.3 Challenges in Integrating Big Data with Cloud Computing

However, there are several challenges in implementing Big Data analytics in the cloud environments. The biggest hurdle is protecting data security and privacy as ever sensitive materials might be stored and processed on external cloud servers. Also, businesses without the capacity to manage the complexity of the big data architectures at a large scale, are vulnerable. Data integration is a worry of the study when many organisations find it difficult to merge structured and unstructured data from many sources. Additionally, there are latency issues while processing real time data and businesses that require high performance will experience issues with it.

4.4 Key Technologies in Cloud-Based Big Data Analytics

It presents what is driving the success of Big Data analytics on the cloud. The distributed computing frameworks like MapReduce and Apache Hadoop are used to process the big amount of data through multiple servers parallely and provides faster results. Also, machine learning algorithms are within cloud ecosystem to enable predictive analytics and use of data to make decisions. Additionally, cloud storage technologies such as data lakes, NoSQL databases became easier to handle both structured and unstructured data. The speed and efficiency of Big Data analytics in the cloud would not be possible doing away with these technologies.

4.5 Industry Case Studies and Real-World Applications

In this study, several case studies are shown that validate the real world usages of Big Data analytics in cloud computing. To name a few, companies in the healthcare which uses cloud based Big Data analytics have been able to work on improving patient outcome by working with big data sets of medical records. Like this, platforms of e-commerce have been implemented the cloud solutions, to improve the customer affective personalization and optimize the inventory management by real time data processing. This series of case studies demonstrate how the cloud platforms allow different types of organizations to harvest actionable insights from the Big Data, and how it enhances their competitiveness.

4.6 Discussion

Overall, this study has gained findings that show an increasing usage of Big Data analytics in the cloud ecosystem quickly as cloud platforms offer the scalability, flexibility and cost efficiency. While there have been great benefits such as real time data processing and better decision making, data security, integration, and technical expertise issues remain which needs to be resolved. The identification of key technologies like distributed computing frameworks and machine learning algorithms, also support the making of cloud based Big Data analytics as a highly transformative technology. Finally, these results highlight the need to resolve the challenges to completely exploit the benefits of Big Data in cloud computing, and therefore the future research should be directed to create better security protocols and tools that can efficiently manage complex data architectures.

5. CONCLUSION

This study concludes with a strong argument of this big data combined with cloud computing leading into the effective adaptation and analytics. Cloud platforms are the most popular way of Big Data integration as the scalability, cost efficiency and flexibility of the cloud makes it suitable for organizations who want to use Big Data power to improve decision making processes and organizational efficiency. Data security, privacy, as well as technical complexity is a challenge to be faced when integrating Big Data analytics with the cloud ecosystem which provides a variety of advantages such as processing real time and efficient data management. This shows the necessity to find continuous innovations of cloud technologies and to create stronger solutions to create away with these troubles. Further looking into the future of Big Data analytics in cloud computing, the driving technologies and the complexity-facing barriers to overcome are still evolving, and it is no wonder why Big Data analytics in cloud computing looks bright for the future. According to businesses, using Big Data analytics on the cloud can be used to optimize internal processes as well as open up a market advantage with data-driven decisions. Once the ecosystem matures, there is a need for further research and development to perfect what is already working and to develop more secure, more efficient, and easier for businesses to integrate Big Data analytics into their current line of business without any conflicts.

REFERENCES

[1] Hillbert, M., & Lopez, P. (2011). The world's technological capacity to store, communicate, and compute information. *Science*, *III*, 62–65.

[2] Hellerstein, J. (2019, November 9). MapReduce leads the way for parallel programming. *Gigaom Blog*. https://gigaom.com/2008/11/09/mapreduce-leads-the-way-for-parallelprogramming/

[3] Statista. (2020). Worldwide data created. *Statista*. https://www.statista.com/statistics/871513/worldwide-data-created/

[4] Reinsel, D., Gantz, J., & Rydning, J. (2017). *Data age 2025: The evolution of data to-life critical*. International Data Corporation, Framingham.

[5] Marr, B. (2018, May 21). How much data do we create every day? The mind-blowing stats everyone should read. *Forbes*. https://www.forbes.com/sites/bernardmarr/2018/05/21/how-muchdata-do-we-create-every-day-the-mind-blowing-stats-everyone-shouldread/?sh=5936b00460ba

[6] Kaisler, S., Armour, F., & Espinosa, J. (2013). Big data: Issues and challenges moving forward. *Wailea, Maui, HI*, pp. 995–1004.

[7] Wikipedia. (2018). *Big data*. https://www.en.wikipedia.org/wiki/Bigdata/

[8] Gewirtz, D. (2018). Volume, velocity, and variety: Understanding the three V's of big data. *ZDNet*. https://www.zdnet.com/article/volume-velocity-and-varietyunderstanding-the-three-vs-of-big-data/

[9] Weathington, J. (2012). Big data defined. *Tech Republic*. https://www.techrepublic.com/article/big-data-defined/

[10] PC Magazine. (2018). *Big data*. http://www.pcmag.com/encyclopedia/term/62849/big-data

[11] Akhtar, S. M. F. (2018). *Big data architect's handbook*. Packt.

[12] WhishWorks. (2019). Understanding the 3 V's of big data: Volume, velocity, and variety. *WhishWorks*. https://www.whishworks.com/blog/data-analytics/understanding-the3-vs-of-big-data-volume-velocity-and-variety/

[13] Yadav, S., & Sohal, A. (2017). Review paper on big data analytics in cloud computing. *International Journal of Computer Trends and Technology*, *9*(3), 156–160.

[14] Kimball, R., & Ross, M. (2013). *The data warehouse toolkit: The definitive guide to dimensional modeling* (3rd ed.). John Wiley & Sons.

[15] Racharla, Mr Sathya Prakash, Mr Kontham Sridhar Babu, and Anil Kumar Jakkani. "An Iterative approach for the Restoration of Motion Blurred Images."

[16] Srivastava, Pankaj Kumar, and Anil Kumar Jakkani. "FPGA Implementation of Pipelined 8× 8 2-D DCT and IDCT Structure for H. 264 Protocol." *2018 3rd International Conference for Convergence in Technology (I2CT)*. IEEE, 2018.

[17] LaprinthX. (2018). Better, faster, smarter: ELT vs. ETL. *LaprinthX*. https://laptrinhx.com/better-faster-smarter-elt-vs-etl-2084402419/

[18] Xplenty. (2019). ETL vs. ELT. *XPlenty*. https://www.xplenty.com/blog/etl-vs-elt/#

[19] Forbes. (2018, November 6). Five benefits of big data analytics and how companies can get started. *Forbes*. https://www.forbes.com/sites/forbestechcouncil/2019/11/06/fivebenefits-of-big-data-analytics-and-how-companies-can-getstarted/?sh=7e1b901417e4

[20] EDHEC. (2019). Three ways educators are using big data analytics to improve the learning process. *EDHEC*. https://master.edhec.edu/news/three-ways-educators-are-using-bigdata-analytics-improve-learning-process#.

[21] Srivastava, P. Kumar, and A. Kumar Jakkani. "Android Controlled Smart Notice Board using IoT." *International Journal of Pure and Applied Mathematics* 120.6 (2018): 7049-7059.

[22] Mahajan, Lavish, et al. "DESIGN OF WIRELESS DATA ACQUISITION AND CONTROL SYSTEM USING LEGO TECHNIQUE." *International Journal of Advance Research in Engineering, Science & Technology* 2.5 (2015): 352-356.

[23] Google Cloud. (2020). *BigQuery*. https://cloud.google.com/bigquery