International Journal of Core Engineering & Management
Volume-5, Issue-11, February-2019, ISSN No: 2348-9510

SCALABILITY AND PERFORMANCE CONSIDERATIONS FOR BIG DATA ANALYTICS USING CLOUD COMPUTING

Sandeep Kumar Guduru
gudurusandeep@gmail.com

Abstract

An effective and scalable data analytics solution is necessary due to the ever-increasing amount of digital data generated by IoT devices, social media, and corporate systems. Bigdata Analytics, when integrated with Cloud Computing, provides a robust framework for managing and processing massive datasets in a cost-effective and flexible manner. This paper explores the scalability and performance considerations essential for optimizing Big Data Analytics within cloud environments. By leveraging virtualization, cloud resources provide on-demand access to computing power, storage, and analytics tools, addressing challenges such as data variety, velocity, and volume. Key advancements in partitioning, resource allocation, and load-balancing strategies enhance the processing capabilities of big data platforms. Furthermore, this study highlights the role of horizontal and vertical scaling in achieving efficiency and fault tolerance. The integration of frameworks like Apache Hadoop and Apache Spark exemplifies the adaptability of cloud-based Big Data solutions. This paper concludes with a discussion of optimization techniques, such as caching, data compression, and query optimization, emphasizing their impact on reducing latency and maximizing throughput.

Keywords: Big Data Analytics, Cloud Computing, Scalability, Performance Optimization, Virtualization, Horizontal Scaling, Vertical Scaling.

I. INTRODUCTION

The phrase "big data" describes techniques for exploring and systematically retrieving insights from datasets that are too large or complex for traditional data-processing technologies to manage. Big data, as the term implies, is just a massive quantity of data. The usual features of data may be described by 5V's. The term "cloud computing" describes the idea of making data storage and processing capacity available via a network of remote servers whenever needed [1]. Cloud computing usually enables users to collaborate with others while accessing, using, working on, and editing their work. With cloud computing, users may work whenever it's most convenient for them, and big data can help them get valuable insights.

Modern analytical solutions that can process enormous amounts of data are required due to the exponential increase in data produced by digital systems, IoT devices, and social media platforms. The administration of very large data sets is a fundamental problem in the context of cloud computing. This is true regardless of the shared resource type in the cloud; databases

International Journal of Core Engineering & Management
Volume-5, Issue-11, February-2019, ISSN No: 2348-9510

may be exposed to clients as part of the infrastructure or concealed behind service interfaces; it makes no difference [2].

The administration of very large data sets is a fundamental problem in the context of cloud computing. This is true regardless of the shared resource type in the cloud; databases may be exposed to clients as part of the infrastructure or concealed behind service interfaces; it makes no difference. Data must be partitioned and replicated across several internet data centers. Amazon and Google, two of the most popular search engines, have started to set up new data centers to provide cloud computing services. One of the most effective big data strategies is cloud computing, which takes advantage of these virtual machines.

There is a strong emphasis on creating scalable and on-demand data and resource availability while developing big data and cloud computing technologies. Through the use of virtualization, cloud computing unifies large datasets by providing on-demand access to programmable computer resources [3]. The capacity to provide resources on demand and pay for only the resources employed in product development is one of the several benefits of cloud computing. When it comes to big data and cloud computing, there are a lot of unanswered concerns and areas for further study. These include problems with data management, data velocity and variety, storage, processing, and resource management. Therefore, with the right infrastructure and tools, cloud computing may aid in the development of a business model for any kind of application [4].

Data analysis and development should be supported by big data apps that use cloud computing after analyzing and extracting useful insights from knowledge acquisition data, data scientists and business analysts need to be able to use cloud-based technologies that facilitate interactive and collaborative exploration of this data [5]. Options for cloud computing may be discussed within the paradigm of big data. The ability to store large amounts of data on the cloud is just one more perk of cloud computing [6].

A. Structure of the paper

Here is the outline of the paper: BigDataAnalytics on the Cloud is introduced in Section II. Descriptive, diagnostic, and predictive analytics are among the analytics techniques covered in Section III. Section IV addresses performance optimization and scalability challenges. Section V reviews relevant literature, and Section VI concludes with key findings and future research directions.

II. FUNDAMENTALS OF BIG DATA ANALYTICS IN CLOUD COMPUTING

Big data refers to datasets that are excessively big to be handled, processed, or analyzed by conventional means. Big Data necessitates methodologies and tools for analyzing and extracting patterns from massive data sets [7]. The term "Big Data Analytics" describes the steps used to find new insights in massive data sets by cleaning, organizing, and analyzing them. Big data analytics refers to a collection of tools and methods that, when combined, need novel integrations in order to unearth massive, previously unseen hidden values inside complicated, massively big databases. Better and more efficient solutions to existing issues, or problems with novel characteristics, are its primary emphasis [8].

International Journal of Core Engineering & Management
Volume-5, Issue-11, February-2019, ISSN No: 2348-9510

A. Types of Big Data Analytics Techniques

There are they discuss the big data analytics techniques, and these techniques are provided in four types, shown in Figure 1.

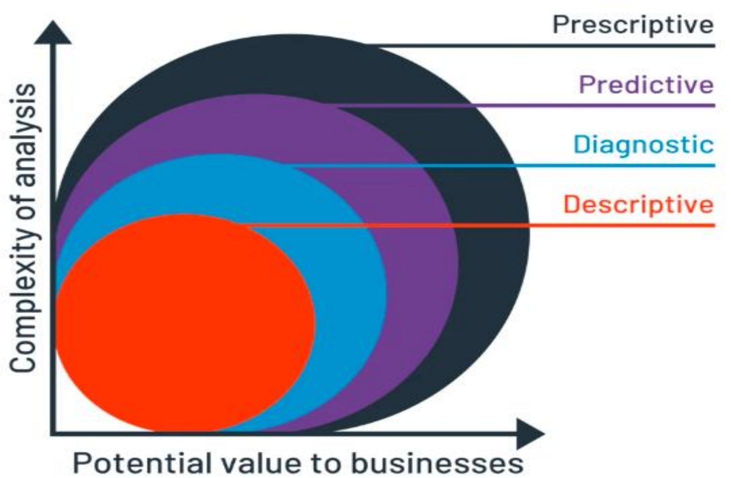

Fig.1. Types of Big Data Analytics

- Descriptive Analytics: Creating a collection of historical data is the first step in data processing. Data mining techniques sift through mountains of information in search of meaningful patterns. The purpose of descriptive analytics is to foretell potential outcomes by presenting patterns and probabilities [9].
- Diagnostic Analytics: Finding and fixing the root causes of problems is what diagnostic analytics is all about.. Its purpose is to ascertain the rationale for an occurrence. This kind looks for the reasons behind things and people's actions [10].
- Predictive Analytics: It is able to foretell the future by analyzing historical data. Predictions are the key. Prospective analytics, which analyses current data to produce possible outcomes, makes use of a wide variety of techniques, including data mining and AI [11].
- Prescriptive Analytics: Its sole function is to ascertain the optimal action to take. Descriptive analytics provides details about the past, whereas predictive analytics aids in predicting future events. Prescriptive analytics uses these parameters to choose the best course of action [12].

1) Key Features of Big Data Analytics

Volume, variety, value, velocity, and veracity are the five defining features of big data, as shown in Figure 2.

International Journal of Core Engineering & Management
Volume-5, Issue-11, February-2019, ISSN No: 2348-9510

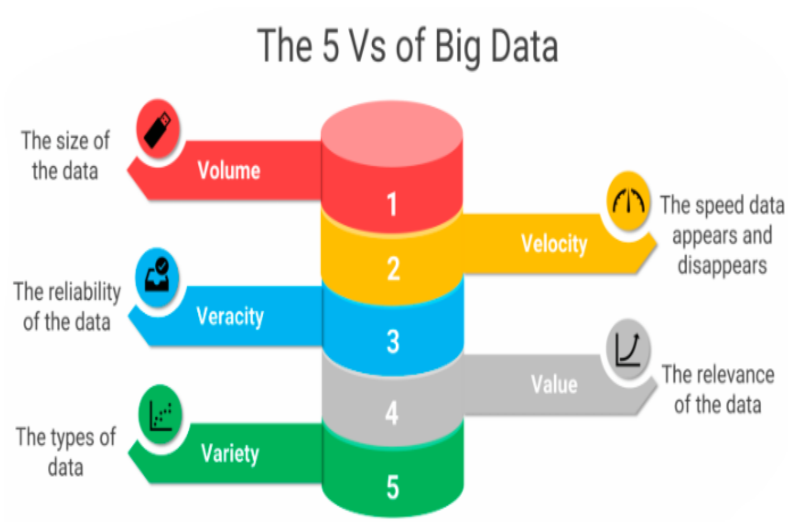

Fig.2. Big Data Characteristics

- Value: The usefulness, effectiveness, and practicality of big data systems are characterized by big data's value attribute. The other four Vs of big data systems are directly proportional to this quality, which is more prevalent in the results of data processing and analytics[13].
- Variety: Data may be organized or unstructured, and there are many different types of sources that contribute to its variety. Various forms of big data provide obvious challenges when it comes to data storage, mining, and analysis.
- Velocity: The word "velocity" stands for how quickly data can be processed. Real potential in data is defined by how quickly it is created and processed according to customer expectations. The rapidity with which data is created and the urgency with which it has to be processed [2].
- Volume: The term "volume" describes the ever-increasing quantity of data of all kinds that are created from various sources. The ability to uncover previously unseen patterns and information via data or information analysis is a major perk of collecting massive amounts of data. Longitudinal data collection requires substantial effort and initial financial commitment [14].
- Veracity: There is a deluge of data formats available nowadays. Statistical information is stored in conventional databases. Data generated by business-related programs. Data quality, dependability, and correctness must be guaranteed.

B. Big Data Analytics in The Cloud

The integration of cloud computing with big data analytics will result in its provision as a service rather than a standalone product. Using a big data solution on the cloud to store and analyze data gets rid of the requirement to build computers in-house [8]. Their strategy is based on a novel, highly scalable big data analytics algorithm that analyses massive amounts of data from all kinds of sources to find previously unidentified patterns and associations as well as other important information for business planning decision-making [15].

International Journal of Core Engineering & Management
Volume-5, Issue-11, February-2019, ISSN No: 2348-9510

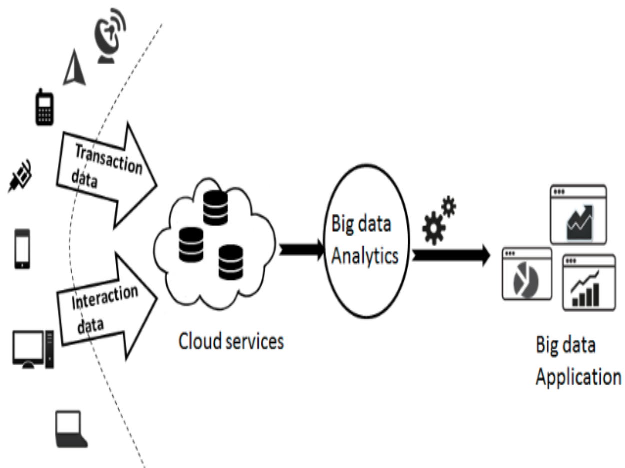

Fig.3. Big data analytics using cloud computing

The process of big data analytics in the cloud. Various sources, such as point-of-sale systems and user interactions, generate transaction and interaction data, respectively seen in Figure 3. This data is stored in cloud services. This data is further processed and analyzed using big data analytics techniques in order to uncover patterns and insights. Finally, these insights are visualized through dashboards and reports, providing businesses with actionable information for decision-making. Businesses may obtain a competitive edge by using the potential of big data using this technique [16]. The idea of cloud-based big data analytics programs. The components are as follows:

- Transaction Data: This data is created from sources such as point-of-sale systems, financial transactions, and e-commerce platforms.
- Interaction Data: This type of data comes from user interaction with devices and applications, for example, social media activity, browsing of websites, and mobile applications.
- Data Storage: The cloud acts as a central storage location for holding huge amounts of transaction and interaction data.
- Data Processing: It processes and analyzes the data stored in the cloud using different techniques to look for meaningful patterns and insights.
- Business Intelligence: Results of data analysis are portrayed to bring out meaningful insights to inform decision-making through dashboards and reports [17].

III. PERFORMANCE OPTIMIZATION FOR BIG DATA ANALYTICS IN THE CLOUD

Optimization of the performance of Bigdata Analytics systems is necessary for the effective processing and analysis of massive data sets. Data intake, storage, retrieval, calculation, and communication among distributed components are among the many aspects that might affect performance [18]. If analytics activities are to be finished within a reasonable amount of time, optimization methods are necessary to guarantee minimal latency, maximum throughput, and reduced resource utilization. In cloud environments, achieving optimal performance is critical due to the dynamic nature of workloads and infrastructure.

A. Low operating cost

C It is no longer necessary for a service provider to supply capacity according to the maximum load; instead, resources may be quickly assigned and deallocated on demand in a noisy environment. Because resources may be freed to minimize operational expenses when service demand is low, this results in considerable savings.

B. Load Balancing Strategies

Load balancing strategies are essential for distributing workloads effectively across cloud resources, ensuring system efficiency, fault tolerance, and high availability in big data analytics. Scalability relies on equally distributing the burden between nodes. When dealing with diverse workloads and different resource capabilities, load balancing becomes more complicated. Intelligent load-balancing methods and constant monitoring are required to balance the burden while taking into account aspects such as data locality, resource utilization, and node availability.

C. Resource Allocation and Management

To get the most out of cloud-based big data analytics solutions, proper administration and allocation of resources is essential. These processes involve efficiently distributing computational, storage, and network resources to meet varying workload demands. Dynamic resource allocation ensures scalability by provisioning or de-provisioning resources on demand, enabling systems to handle peak loads and conserve resources during idle periods. Techniques such as elastic scaling leverage the cloud's flexibility to automatically adjust resources based on workload patterns. Effective resource scheduling, including priority-based and fair scheduling algorithms, ensures that tasks are executed efficiently while maintaining a balance between performance and cost. Cost-aware allocation integrates budget constraints into resource management decisions, helping organizations achieve performance objectives without overspending.

D. Performance Optimization Techniques

For big data analytics solutions to scale, performance optimization is essential. The following methods may assist in increasing scalability and performance:

- Data Partitioning: Data may be partitioned among numerous nodes using a specified technique, allowing for fast data retrieval and parallel processing. Examples of such strategies are range-based and hash-based partitioning. To minimize data migration among nodes and provide a balanced distribution of data, partitioning should be carefully designed [19].
- Query Optimization: Performance optimization of queries is critical. To achieve this

goal, it is necessary to use suitable join algorithms, minimize unnecessary calculations, and rewrite queries to decrease data transfer [20].

- Data compression: Minimising data storage needs and enhancing data transit rates are both achieved via compression. Minimizing the data footprint without losing query efficiency is possible with various compression methods like Snappy or LZ4 [2].
- Caching: By storing intermediate results or frequently requested data, caching systems may lessen the need for recurrent calculations. Distributed caching frameworks like Redis or Apache Ignite are good examples of such mechanisms.
- Monitoring and Tuning: Performance bottlenecks may be identified with the use of continuous monitoring of system indicators, including CPU utilization, memory consumption, network traffic, and I/O activities. By examining these measurements, businesses can optimize system settings, distribute resources effectively, and tackle performance problems before they ever arise.
- Query Optimization: Performance optimization of queries is critical. This requires using suitable join algorithms, modifying queries to minimize data transfer, and minimizing duplicate calculations. Query optimization, predicate pushdown, and query rewriting are some of the techniques that may enhance the efficiency of query execution [18].

IV. SCALABILITY IN BIG DATA ANALYTICS USING CLOUD COMPUTING

To handle the growing amounts of data required for big data analytics, systems need to be scalable. The rise of cloud computing has made scalability a more practical and affordable goal to pursue. The capacity of a system to manage ever-increasing workloads and information is known as its scalability. Ideally, a scalable system would be able to handle increases in data volume or concurrent users without noticeably impairing performance.

A. Type of scalability in big data analytics using cloud

The most common forms of scalability are the horizontal and vertical varieties.

1) Horizontal Scaling

Horizontal scaling entails spreading the workload over several servers, some of which can be commodity computers. A related concept is "scale out," which refers to the practice of combining the processing power of several separate computers. In a typical setup, many OS instances are executed on distinct systems.

Key benefits of horizontal scaling include:

- Increased fault tolerance due to distributed architecture.
- Better alignment with modern cloud environments, where adding virtual instances is seamless.
- Enhanced capacity to handle large-scale and high-velocity data streams.

2) Vertical Scaling

Vertical scaling is the practice of adding more powerful hardware, such as additional processors, memory, and faster hardware, to an existing server. It sometimes includes running a

International Journal of Core Engineering & Management
Volume-5, Issue-11, February-2019, ISSN No: 2348-9510

single OS instance and goes by the name "scale up" as well. It is suitable for applications with monolithic architectures or those that require high-speed processing within a single node. Advantages of vertical scaling include:

- Reduced complexity, as there is no need to manage distributed resources.
- Improved performance for workloads that are CPU- or memory-intensive[21].

B. Examples of Scalable Big Data Frameworks

Scalable big data frameworks form the backbone of modern analytics platforms, enabling the processing of massive datasets with high efficiency.

1) Apache Hadoop

Distributed large data storage and processing is made easy using Hadoop, a popular platform. A sizable community backs Apache Hadoop, a popular Bigdata solution. It was crafted to sidestep the difficulties and poor performance that come with utilizing conventional technologies to handle and analyze massive data. The distributed file system and parallel clusters that makeup Hadoop allow it to handle massive data sets with ease. Hadoop's ability to add nodes dynamically makes it a robust solution for horizontal scaling[22].

2) Apache Spark

Apache Spark is a free and open-source framework for processing large amounts of data with an emphasis on simplicity, speed, and advanced analytics. Among big data processing technologies, Hadoop has been around for a decade and is still the go-to for handling massive datasets. Its Resilient Distributed Dataset (RDD) abstraction allows for fault-tolerant and scalable processing across large clusters [23].

- Data from streaming sources may be linked with data from historical sources, batch processing can be done using spark code, and ad hoc queries on streaming data can be executed.
- Hadoop applications may run faster and use less storage using Spark. With Spark, you can get interim results with fewer disc read/write operations. It stores information in RAM and performs disc IO operations only when absolutely necessary.

V. LITERATURE REVIEW

This section presents a literature review on Bigdata Analytics in Cloud Computing, highlighting best practices, challenges, and emerging solutions. A summary of the reviewed studies is provided in Table I for a concise overview.

In, Pawar, Kharat and Pardeshi (2016) explains the idea behind cloud computing's utilization of Big Data Analytics. The advent of cloud computing in the last decade has been a game-changer when it comes to meeting the demands of many businesses for flexible, on-demand data storage and processing, as well as an extensible computational platform. Nowadays, cloud computing is changing the way data is stored in the digital world; it is an example of a crucial technological trend. At the same time, other technologies, such as big data analytics, sift through mountains of data in search of previously unseen correlations, patterns, and insights[24].

In, Jain and Kumar (2015) explores a range of issues around big data processing and potential cloud-based solutions. There is an excessive amount of data, both structured and unstructured, being collected by businesses due to the proliferation of social media, the IoT, and multimedia. The data itself does not need to be relocated; rather, it is the analytical software that must be transferred, as data that is too large to handle cannot be transferred either. Cloud computing makes this feasible because it is more cost-effective for enterprises to analyze publicly available data sets in the cloud, including those from social media, financial markets, weather, genomes, and aggregated industry-specific datasets[25].

In, Yetis et al. (2016) methods for approaching data analytics on the cloud, one of the primary domains where Big Data finds application, are reviewed. Based on their findings, they suggest several avenues for further study into data-driven decision-making for Big Data analytics and cloud-based computing solutions to the scientific community. Data analytics use statistical methods in conjunction with cloud computing, ML, or AI to sift through massive amounts of data, extract useful information, create a knowledge base, and, finally, create a nonparametric model of the Big Data[26].

In, Manekar and Pradeepini (2016) centered on the question of how to integrate cloud computing and big data into a single development framework. Big data and cloud computing are two examples. There is a data tsunami coming from more than 5 billion people using mobile devices and about the same number using Facebook and other social media sites. Cloud computing, the paradigm of the next generation of information technology service, is also rapidly expanding as a means of providing this big data service. Both kinds of technology are always developing. In the long run, as cloud computing evolves, every leading company will consider creating a more nimble and efficient cloud environment[27].

In, Hari Baaskar, Sujitha and Praveen (2015) application-specific parameters are retrieved by simulation on the appropriate cloud environment. Incorporating every conceivable value for those characteristics into a big data analytics model will provide precise metrics and valuable business insights. The advent of cloud computing has been a paradigm shift in the IT industry, moving us closer to seeing computers for their practical applications. There is a dearth of reliable simulation methods for assessing the practicality, efficiency, and financial aspects of such applications. Conversely, a new class of data analytics methods has been made available by big data analytics[28].

International Journal of Core Engineering & Management
Volume-5, Issue-11, February-2019, ISSN No: 2348-9510

Table I. Presents the Comparative Table Based on Big Data Analytics Using Cloud Computing

Reference	Study On	Approach	Key Findings	Challenges	Limitations
[24]	Use of Big Data Analytics in cloud computing	Discusses how cloud computing provides a scalable and flexible infrastructure for big data analytics.	Cloud computing enables the handling of large datasets with scalability, flexibility, and efficiency.	Integration of diverse data formats; Real-time processing for decision-making.	Limited focus on specific cloud-based Big Data platforms or tools.
[25]	Problems and solutions in Big Data computation using cloud computing	Highlights the movement of analytical programs to data instead of transferring data due to its massive size.	Analyzing Big Data in the cloud reduces costs and enables better access to public datasets.	Scalability for growing datasets; Security of public and private datasets in the cloud.	Lack of discussion on hybrid cloud implementations or cost implications for SMEs.
[26]	Cloud environment approaches for Data Analytics	Surveys key applications of Big Data in cloud environments and provides future research recommendations.	Combines machine learning and computational intelligence to create models for managing Big Data.	Ensuring accuracy and reliability in nonparametric models for Big Data.	Limited real-world validation or application scenarios are provided.
[27]	Integration of Big Data and Cloud Computing	Explores how Big Data and cloud computing complement each other as evolving technologies for IT service delivery.	Cloud computing facilitates agile environments for processing Big Data generated by billions of users.	Rapid technological evolution; Integration challenges with legacy systems.	No detailed framework for implementing combined Big Data and cloud systems.
[28]	Simulation-based analysis of cloud environments for Big Data Analytics	Simulates cloud environments for application-specific analytics and parameter extraction.	Big Data Analytics improves feasibility, performance, and cost measures for cloud-based applications.	Accurate simulation of cloud-based systems; Handling diverse application-specific requirements.	Simulation results may not fully represent real-world cloud environments.

VI. CONCLUSION AND FUTURE WORK

A storage, processing, and analysis of data has been transformed by the combination of Bigdata Analytics and Cloud Computing. Cloud platforms offer scalable and cost-effective solutions for handling the massive amounts of data generated by various sources, including the IoT and social media. By leveraging virtualization, distributed storage, and advanced analytics techniques, organizations can gain actionable insights from massive datasets in real time. This paper explored various Big Data techniques, optimization strategies, and scalable frameworks like Hadoop and Apache Spark that facilitate efficient data management within cloud environments. The findings demonstrate that cloud-based Bigdata solutions offers a robust infrastructure for industries to enhance decision-making processes, improve operational efficiencies, and innovate new services.

Future work should focus on improving data security and privacy measures in cloud-based big data systems, especially as organizations adopt hybrid and multi-cloud environments. Developing efficient, real-time analytics models to handle diverse and dynamic data streams is another critical area for exploration. Additionally, the use of AI and ML techniques in Bigdata Analytics can be expanded to automate decision-making processes and uncover hidden insights from data.

REFERENCES

1. Amalarethinam and P. Hameed, “A Predictive Model to Forecast and Pre-Treat Diabetes Mellitus using Clinical Big Data in Cloud,” Int. J. Appl. Eng. Res., 2015.
2. T. S. Harsha, “Big Data Analytics in Cloud Computing Environment,” vol. 8, no. 8, pp. 159–179, 2017.
3. M. D. Assunção, R. N. Calheiros, S. Bianchi, M. A. S. Netto, and R. Buyya, “Big Data computing and clouds: Trends and future directions,” J. Parallel Distrib. Comput., 2015, doi: 10.1016/j.jpdc.2014.08.003.
4. I. A. T. Hashem, I. Yaqoob, N. B. Anuar, S. Mokhtar, A. Gani, and S. Ullah Khan, “The rise of ‘big data’ on cloud computing: Review and open research issues,” Information Systems. 2015. doi: 10.1016/j.is.2014.07.006.
5. D. P. and K. Ahmed, “A Survey on Big Data Analytics: Challenges, Open Research Issues and Tools,” Int. J. Adv. Comput. Sci. Appl., 2016, doi: 10.14569/ijacsa.2016.070267.
6. M. Sathyapriya and V. Thiagarasu, “Big Data Analytics Techniques for Credit Card Fraud Detection: A Review,” Int. J. Sci. Res., vol. 6, no. 5, pp. 2319–7064, 2015.
7. M. M. Najafabadi, F. Villanustre, T. M. Khoshgoftaar, N. Seliya, R. Wald, and E. Muharemagic, “Deep learning applications and challenges in big data analytics,” J. Big Data, 2015, doi: 10.1186/s40537-014-0007-7.
8. J. Verma and S. Agrawal, “Big Data Analytics: Challenges And Applications For Text, Audio, Video, And Social Media Data,” Int. J. Soft Comput. Artif. Intell. Appl., vol. 5, pp. 41–51, 2016, doi: 10.5121/ijscai.2016.5105.
9. H. Cao, M. Wachowicz, and S. Cha, “Developing an edge computing platform for real-time descriptive analytics,” in Proceedings - 2017 IEEE International Conference on Big Data, Big Data 2017, 2017. doi: 10.1109/BigData.2017.8258497.

International Journal of Core Engineering & Management
Volume-5, Issue-11, February-2019, ISSN No: 2348-9510

10. M. P. Vittoria and P. Napolitano, "Large cultural networks and smart specialization: What is new in regional policy diagnostic analytics?," Soc. Netw. Anal. Min., 2016, doi: 10.1007/s13278-016-0359-4.
11. E. T. Bradlow, M. Gangwar, P. Kopalle, and S. Voleti, "The Role of Big Data and Predictive Analytics in Retailing," J. Retail., 2017, doi: 10.1016/j.jretai.2016.12.004.
12. S. A. El-Seoud, H. F. El-Sofany, M. Abdelfattah, and R. Mohamed, "Big data and cloud computing: Trends and challenges," Int. J. Interact. Mob. Technol., 2017, doi: 10.3991/ijim.v11i2.6561.
13. M. H. ur Rehman, C. S. Liew, A. Abbas, P. P. Jayaraman, T. Y. Wah, and S. U. Khan, "Big Data Reduction Methods: A Survey," Data Sci. Eng., 2016, doi: 10.1007/s41019-016-0022-0.
14. A. Gandomi and M. Haider, "Beyond the hype: Big data concepts, methods, and analytics," Int. J. Inf. Manage., 2015, doi: 10.1016/j.ijinfomgt.2014.10.007.
15. C. W. Tsai, C. F. Lai, H. C. Chao, and A. V. Vasilakos, "Big data analytics: a survey," J. Big Data, vol. 2, no. 1, pp. 1–32, 2015, doi: 10.1186/s40537-015-0030-3.
16. B. M. Balachandran and S. Prasad, "Challenges and Benefits of Deploying Big Data Analytics in the Cloud for Business Intelligence," in Procedia Computer Science, 2017. doi: 10.1016/j.procs.2017.08.138.
17. H. Chen, R. H.L.Chiang, and V. C. Storey, "Business Intelligence and Analytics: From Big Data To Big Impact," MIS Q., vol. 36, no. 4, pp. 1165–1188, 2018.
18. L. E. Bautista Villalpando, A. April, and A. Abran, "Performance analysis model for big data applications in cloud computing," J. Cloud Comput., 2014, doi: 10.1186/s13677-014-0019-z.
19. M. P. Narooka and S. Chaodhary, "Paradigm Shift of Big-Data Application in Cloud Computing," vol. 5, no. 5, pp. 515–521, 2016, doi: 10.17148/IJARCCE.2016.55130.
20. S. Khan, K. A. Shakil, and M. Alam, "Big data computing using cloud-based technologies: Challenges and future perspectives," Networks Futur. Archit. Technol. Implementations, pp. 393–414, 2017, doi: 10.1201/b21234.
21. D. Singh and C. K. Reddy, "A survey on platforms for big data analytics," J. Big Data, vol. 2, no. 1, pp. 1–20, 2015, doi: 10.1186/s40537-014-0008-6.
22. A. Oussous, F. Z. Benjelloun, A. Ait Lahcen, and S. Belfkih, "Big Data technologies: A survey," Journal of King Saud University - Computer and Information Sciences. 2018. doi: 10.1016/j.jksuci.2017.06.001.
23. V. Rajaraman, "Big data analytics," Resonance, vol. 21, no. 8, pp. 695–716, 2016, doi: 10.1007/s12045-016-0376-7.
24. V. J. Pawar, K. D. Kharat, and S. Pardeshi, "Structuring cloud computing using big data analytics solution: A survey," in Proceedings of the International Conference on Communication and Electronics Systems, ICCES 2016, 2016. doi: 10.1109/CESYS.2016.7889827.

25. V. K. Jain and S. Kumar, "Big Data Analytic Using Cloud Computing," in Proceedings - 2015 2nd IEEE International Conference on Advances in Computing and Communication Engineering, ICACCE 2015, 2015. doi: 10.1109/ICACCE.2015.112.
26. Y. Yetis, R. G. Sara, B. A. Erol, H. Kaplan, A. Akuzum, and M. Jamshidi, "Application of Big Data Analytics via Cloud Computing," in World Automation Congress Proceedings, 2016. doi: 10.1109/WAC.2016.7582986.
27. A. K. Manekar and G. Pradeepini, "Cloud Based Big Data Analytics a Review," in Proceedings - 2015 International Conference on Computational Intelligence and Communication Networks, CICN 2015, 2016. doi: 10.1109/CICN.2015.160.
28. R. Hari Baaskar, K. Sujitha, and K. Praveen, "Analysing cloud simulation results using big data analytics model," in 2015 International Conference on Computer Communication and Informatics, ICCCI 2015, 2015. doi: 10.1109/ICCCI.2015.7218128.