# INVESTIGATION TO IMPROVE PERFORMANCE OF UNSTRUCTURED DATA FROM CLOUD ENVIRONMENT

**Sasmita Pradhan** Asst. Prof. Einstein Academy of Technology and Management, Bhubaneswar,
**Debasish Pradhan** Asst. Prof. Einstein Academy of Technology and Management, Bhubaneswar,
**Swapnarani Barik** Student, Einstein Academy of Technology and Management, Bhubaneswar,

**Huge volume of digital information build up are of different types and compound that is categorized as Big Data. The information increased from Big Facts adds value through Big Data Analytics (BDA).. As the established methods of data management not properly reinforced by the cloud.. Therefore there is a need to develop Routine Optimization algorithms for the NoSQL database under consideration. The main objective of planned study is to improve performance of unstructured data stored in the cloud environment using some methodologies tools.**

**KEYWORDS digital information, Big Data Analytics (BDA), Routine Optimization algorithms ,**

## INTRODUCTION:

In today era two mainstream technologies Big Data and Cloud Computing are the center of concern in IT. Basically, Big data is all about dealing with the huge level of data whereas Cloud computing is about infrastructure. However, the simplification offered by Big data and Cloud technology is the major purpose for their huge enterprise acceptance.

Usages of both big data and cloud computing gives valuable outcome for the organizations. Not to mention, both the technologies are in the stage of evolution but their combination influences scalable and cost-effective resolution in big data analytics. But managing and gaining insights from the overflow of information stored is becoming a challenge.

The two main requirements of big data analytics solutions are
- extensible storage that can have room for the growing data and huge processing capacity that can run complex analytical tasks in predetermined and acceptable time [1]. The process of data retrieving, acquisition, archiving, and analyzing might become very luxurious from the cloud environment except handle and analyze big data. The ability of storing huge amount of data in variation and process it in very large speeds will answer in data that can guide education institutions, business and other industries to grow fast [2]. Because of this, the enterprise dealing with big data from the cloud environment are nowadays searching for methods to reduce their cost by optimizing their infrastructure. The goal of research is aim at design algorithms which can handle big data in the cloud in an optimized manner.

## AVAILABILITY OF INFORMATION :

With the age of Big Data, cloud computing will also grow and so that with the age of Big Data, the survey report proposed some of the key challenges existing in the field of cloud computing. The survey report proposed some of the key challenges existing in the field of cloud computing. With the existing tool and techniques it is not sufficient to remain all the challenges relating to big volume of data. Even it is not sufficient to provide better data quality with the existing technology. Privacy is again big problem with cloud data. So to process

 

streaming data we need some novel algorithms and some efficient tools.[3]

Hence there is a need to develop Performance Optimization algorithms for the database under consideration. The main objective of proposed study is to improve performance of unstructured data stored in the cloud environment using some methodologies tools.

## SIGNIFICANCE OFTHE STUDY:

According to Khan, Nawsher & Yaqoob currently, over 2 billion people worldwide are connected to the Internet, and over 5 billion individuals own mobile phones. According to them by 2020, 50 billion devices are expected to be connected to the Internet. At this point, predicted data production will be 44 times greater than that in 2009 [6]. In the digital world, knowledge is generated and collected at a rate that quickly go beyond the limit. As information is moved on and shared at light speed on optic fiber and wireless networks, the volume of data and the speed of market growth increase [6]. So, the huge growth rate of such huge data generates numerous challenges, like the rapid growth of data, transfer speed, diverse data, and security. Unstructured data, such as text messages, location information, videos, and social media data, are data that do not follow a specified format. Considering that the size of this type of data continues to increase through the use of smart phones, the need to analyze and understand such data has become a challenge.

## REVIEW OFTHE LITERATURE IN AREA OF RESEARCH:

The authors Chen, Gang [4], in today's era enterprises are increasingly transferring their data processing to the cloud, due to of cost, scalability, and convenience, among others. Though,hosting multiple applications and storage systems on the same cloud introduces resource sharing and heterogeneous data processing challenges. In addition to, according to Chen [4], real clouds are never perfectly symmetric - there often are differences between individual processors in their capabilities and connectivity.

According to the author Leavitt, Neal [5], many organizations collect huge amounts of customer, scientific, sales, and other data for future analysis. By tradition, most of these organizations have stored structured data in relational databases for subsequent access and analysis. So, a mounting number of developers and users have begun revolving to various types of No relational, called NoSQL-databases. No relational databases, including hierarchical, graph, and object-oriented databases-have been since the late 1960s. However, new types of NoSQL databases are being developed. Varieties of NoSQL databases take different approaches. The only common thing between them is not relational. Main advantage of that, they can handle unstructured data such as word-processing files, e-mail, multimedia, and social media efficiently.

The author Samiya Khan [1] described view of the challenges and issues that there is a need for a cloud-based algorithms to facilitate advanced analytics. The analytical workflow is

 

collected of several steps, which contain data acquisition, storage,processing,and analytics.

The author Santosh Kumar Majhi [3] predicted cloud computing will grow and with the age of BigData. With the existing tool and techniques it is not sufficient to adhere all the challenges relating to big volume of data. It is not feasible to provide better data quality with the existing technology and again privacy is big problem with cloud data.

## RESEARCH METHODOLOGY:

Discover appropriate NoSQL database to persist data and analyze Big data a performance metrics-based study on the selected NoSQL is to be done and analyze the Results and Record the conclusions.

## CONCLUSION:

The main goal of this research is results the various research problems transaction with the various perception of handling unstructured data such as - technique to efficiently handle unstructured data, archive huge volume/Sized unstructured data, to share large Sized unstructured data, Optimize performance for retrieving unstructured data.

## FUTUREWORK:

To develop algorithms for performance optimization of unstructured data in Cloud environments, identification of the accurate data model suitable to manage a diversity of Data and determine the exact database to manage unstructured data in the cloud. Research can then recognize the way to share huge sized unstructured data. Discover resource optimization algorithm based on the performance results

## REFERENCES

1. Khan, Samiya & Shakil, Kashish & Alam, Mansaf. (2017). Big Data Computing Using Cloud-Based Technologies:Challenges and Future Perspectives.
2. El-Seoud, Samir & El-Sofany, Hosam & Abdelfattah, Mohamed & Mohamed, Reham.(2017). Big Data and Cloud Computing:Trends and Challenges.
3. Majhi, Santosh &Shial, Gyanaranjan. (2015). Challenges in Big Data Cloud Computing And Future Research Prospects: A Review.
4. Chen, Gang & Jagadish, H.V. & Jiang, Dawei & Maier, David & Chin Ooi, Beng & Tan, Kian-Lee & Tan, Wang-chiew. (2014). Federation in Cloud Data Management: Challenges and Opportunities.
5. Leavitt, Neal.(2010).Will NoSQL databases live up to their promise?.
6. Khan, Nawsher & Yaqoob, Ibrar & Hashem, Ibrahim & Inayat, Zakira & Kamaleldin, Waleed & Alam, Muhammad & Shiraz, Muhammad & Gani, Abdullah. (2014). Big Data: Survey, Technologies, Opportunities, and Challenges.The ScientificWorld Journal.2014. 18. 10.1155/2014/712826.

 