ISSN: 2350-0328

International Journal of Advanced Research in Science, Engineering and Technology

Vol. 6, Issue 8 , August 2019

Architecture for Data Integration in Big Data Analytics

Rahul Kumar * , Dr. Subodh Kumar, Dr. Saurabh Gupta

Research Scholar Department of Computer Applications, IET Mangalayatan University Aligarh-U.P. (India)
Associate Professor Department of Computer Applications, IET Mangalayatan University Aligarh-U.P. (India)
Deputy Director General Naitional Informatics Center, Delhi. (India)

ABSTRACT: Big data explore new opportunities to modern era for discovering new information and knowledge for better understanding and rapid decision making. Big data refers to the bang of existing information. The big data movements are obsessed by the very large amounts of high-dimensional or unstructured data which is continuously generated and stored with a much cheaper cost for further use. There is massive increase in the data size in every sector. The term big data became very popular now days due to large amount of information's are added every second. With the increase of data, the efficient and critical analysis of data is required for its better understanding and decision making.
In this paper, authors have proposed architecture for big data integration processes. Authors have also been discussed the challenges of big data analysis and different types of data integration. The key challenges of big data integration are schema mapping, record linkage and data fusion.

GENERAL TERMS: Big data Integration, Distributed Data, Unstructured Data, Data Warehouse.

KEYWORDS: Big Data, Hadoop, Hbase, HDFS, Streaming,Kafka.

1. INTRODUCTION

Big data is a term applied to data sets whose size or type is beyond the ability of traditional relational databases to capture, manage, and process the data with low-latency. Big data analytics is an advanced analytic techniques to analyze the very large, diverse data sets that include different types of data such as structured/unstructured and streaming/batch and different sizes from terabytes to zettabytes. The Hadoop Distributed File System (HDFS) is designed to store very large data sets reliably, and to stream those data sets at high bandwidth to user applications. In a large cluster, thousands of servers both host directly attached storage and execute user application tasks. By distributing storage and computation across many servers, the resource can grow with demand while remaining economical at every size. We describe the architecture of HDFS and report on experience using HDFS to manage 40 Petabytes of enterprise data at Yahoo. Now days, data is boosting very fast and hence some efficient tools and techniques are required to process and interpret big data as conventional tools are not efficient enough to handing big data smoothly. In this paper authors have proposed architecture for data integration in big data environment for efficient execution and handling of big data. The organization of present paper is in six different sections namely section-1: Introduction, Section-2: related work, section-3: Data Integration Challenges of Big Data, section-4: Flow of Data Integration in Big Data, section-5: Conclusion and lastly paper ended with references and authors brief profile.

II. RELATED WORK

In present time, the digital universe comprise all structured and unstructured data spanning from videos, movies, surveillance video, photographs, data recorded through sensors, connected devices etc. and it is expected to increase up to 44 zettabytes by the 2020 [1]. Till now numbers of studies have been conducted in the field of big data at national and international platform. This section covers the related work conducted so far in this field as follows:
Lu,W. et al. [2] implemented the K nearest neighbor on Hadoop to speed up the processing time by using multiple computers working together over a network. They concluded that the map reduce K nearest neighbor outperform the sequential K nearest neighbor with big data.

ISSN: 2350-0328

International Journal of Advanced Research in Science, Engineering and Technology

Vol. 6, Issue 8 , August 2019

A map reduce support technique namely MRSMO (Map Reduce base distributed SVM algorithm for automatic image annotation) used by Liu,Y., et al. [3] in their study. The results of this study indicated that their proposed technique gives significant reduction in the training time and a high accuracy in both binary and multiclass classification.

Costa C. and Santos M.Y [4] presented an architecture to store and process the data of city energy consumption with the help of big data technologies. They used Hadoop framework to process this data. In this framework they used HDFS as file structure and apache Hbase used to store relational data Hive used for querying purpose and PIG used for data streaming.

Girtelschmid S., Steinbauer M., Kumar V., Fensel A., and Kotsis G et al. [5] proposed and presented an architecture for efficient processing of sensor data from large scale smart building installation. They have addressed this model with ontology reasoning and big data processing. They used big data streaming clusters for basic processing.

Simmhan Y., Aman S., Kumbhare A., Liu R., Stevens S., Zhou Q., and Prasanna V et al. [6] discussed about the solution of big data analytics with the help of smart Gird. They used smart grid in cloud technologies, by this way they can provide the infrastructure as per they needed. This platform provides an adaptive information integration pipeline for ingesting dynamic data

Ke Xu, [7] used map reduce based parallel support vector machine for classification of emails. They have evaluated many techniques including vector machine based on map reduce and their results indicated that the vector machine outperforms the all other techniques.

Liang Jiajun et al. [8] have discussed about the big data applicability in dropout ratio from education. This research work had helped for management to identify the individual students' dropout from courses.

Zin thi, and Hamehiromitueal [9] defined their research work for analysis the consumer behavior using the big data techniques. In this work they used technological data structured co –occurrences methodologies and Marko chain theory.

Riedel M. et al. [10] proposed a model to analytics the big data for high productivity data process analytics, the author used the Hadoop for storing the big data and they used HPC (high performance computing) and HTC (High Throughput computing) for powerful processing architecture. They also used SVM (Support Vector Machine) to parallel processing and Speedup the processing. And for data replication purpose they used B2SHARE.

Dem C.Y. et al. [11] proposed a big data framework to analysis the big data with the use of HPC (High Processing Cluster). They used Hadoop open source tools and it's supported tools for data streaming and visualizations. They used Microsoft Azure for cloud computing platform. They had done descriptive study of automation process of big data analytics. They took a case study on bioinformatics application deployed on cloud. They used VPC (virtual private cloud) to provide the dynamic allocation of infrastructure.

Xu Brain [12] had proposed a big data framework with machine learning algorithm to improving the prediction for health monitoring, in this framework; they used HDFS, Hbase, MapReduce, Mahout and zookeeper. They also used ML tools Rapid Miner and Weka to designing the Machine Learning algorithms. They used linear Regression, Logistic Regression, Neural networks, Support Vector Machine, Naïve bayes, Nearest Neighbor Decision tree, Random forest and gradient boosted tree.

Gu Lin [13] suggested algorithms to minimize the cost of communication to process the big data analytics. They used network function virtualization technology (NFV) to increase the networking efficiency, flexibility and scalability. By using this proposed a low complexity heuristic algorithm.

Wang Ziqi and Zhao Haihui et al. [14] have done the empirical study of big data tools and techniques. They used these tools for a private Chinese firm data to improvements in business process. They also used cloud computing to perform big data tools and techniques. The main focus of authors was to make business strategies by using big data at operational level data. The size of big data is increasing day to day and presently it ranging from a few dozen terabytes (TB) to many Peta bytes (PB) of data in a single data set. Apart of this there are many challenges and difficulties related to storage, capture, retrieve, share, analytics, and visualization of big data. Today, many organizations are using large size of data to discover facts. Hence, big data analytics is required to analyze the big data sets.

Govindarajan K. et al. [15] proposed a big data framework for analysis. In this framework they used HDFS as file system and other it's supported tools. The main focus of this project was SDN (software defined network). By this the network bandwidth can allocated to big data component as per the needed.

Dubey V. and Pradhan C. [16] illustrated the big data applicability for monitoring big data of project work. They proposed an open source model by using HDFS as file system and Hbase for storing relational database and sqoop use for data streaming between oracle and HDFS. And they also use Map Reduce for analysis the data and extract the report. For the purpose of authentication they us LDAP (Light weight Directory Access Protocol).

ISSN: 2350-0328

International Journal of Advanced Research in Science, Engineering and Technology

Vol. 6, Issue 8 , August 2019

Bakshi K. [17] explored the architecture of big data technology the author discussed about the theoretical architecture of big data solutions using open source tools. He used HDFS as file system and Hbase for storing relational data and legacy data and MapReduce used for analysis the stored data and also try to make common architecture for big data industries.

Deshmukh S. and Sumeet S. [18] discussed about how to use the big data technology in public cloud infrastructure, many public infrastructure are available in market (for example Amazon EMR, Amazon EC2). They used these services to implement big data tools and analysis big data.

Phaneendra S. Vikram,Reddy E. Madhusdhana et al. [19] discussed that in older days the data was less can easily handle by traditional RDBMS tools from the decades the data had generated by very speed form of structured and unstructured i.e. big data. In this study the authors define the properties of big data i.e. volume, velocity, verity and value. They also illustrate the big data architecture using Hadoop, that consist the name node, data node and also define HDFS file format to store the large data. The authors also focused on the applicability of big data techniques in industries i.e. financial, retail, health care etc.

Reddi Kumar Kiran, Indira Dnvsls et al. [20] Illustrate that the transfer of big data on internet is not an easy task the authors proposed an algorithms to transfer the structured and unstructured data, in this algorithms they used stored and forward technique. Under this model you can use only when the demand of data is very low. The band width can be repurposed for transmission without impacting others users in system.

Fan Wei,Bifet Albert et al. [21] explained about the big data mining and the capability of big data set extracting by using existing tools and techniques is not an easy task due to its big volume, variety and velocity. The author had discussed about the existing big data handling tools like Hadoop, strome apache s4. Specific tools for graph mining were Pegasus and Graph.

Bifet Albert et al. [22] did the study about the streaming of big data at real time they had discussed about the problem of streaming of large data set at real time by which an organization can take the necessary decision at real time. The tools used for data mining are apache Hadoop, apache pig, Strome, Habase apache Mahout and R etc. They had discussed about the ability to handle many Exabyte of data that have rich variety data set, techniques.

Purcell Bernice et al. [23] illustrated that the big data is an poses opportunities and challenges for business. They did the study about the traditional tools that are not able to handle the structured and unstructured data sets. The Hadoop architecture used to process the big data sets and used Map Reduce to querying the real time data.

Agarwal Sameer, Mozafari Barzan , Panda Aurojit, Milner Henry, Sarmuel Madden, Stoica Ion et al. [24] discussed BlinkDB, that is an approximate query engine that is running S&L queries on large volume of data. It is based on parallel processing Blink used two key ideas(i) multi dimensional sampling strategy that builds and maintain variety of samples. (ii) a run time dynamic based sample selection to choose the best sample for satisfy query constraints. They also elaborate a query to process in 2 sec on 17 TB data with 90 to 98 % accuracy.

Sharam ashaish, Viyas Snehalata et al. [25] discussed about the Hadoop 2, In Hadoop 2 overcome the problems of hadoop 1 and the authors focused on the YARN (yet another resource negotiator) that provides the resource to hadoop so also called resource manager. In Hadoop 2 YARN becomes aspirate part or can say individual tool and increase the performance and one is most important things is that it is reliable to existing applications that is running in Hadoop1 YARN act alike a operating system for hadoop because it provide the resources, security and compatibility.

Eldawy Ahamed, Mokbel Mohamed et.al. [26] proposed a fully fledged map reduced framework with native support of spatial data. Spatial Hadoop is a comprehensive extension of hadoop that pushed the spatial data inside the core functionality of Hadoop. They described the spatial hadoop model in different-2 Layers i.e Language, storage, MapReduce and operation. In language layer a non-technical can use any high level language to analysis the data. In storage layer, they define two level indexes, global and local at node level. At Map Reduce layer they used basic spatial components and operational level used three basic operations: range query, K-NN queries and spatial join.

Dean Jeffery, Ghemawat Sanjay et al. [27] demonstrated the implementation of Map Reduce on large clusters on commodity hardware to analysis the big data. The Map Reduce can analysis may terabytes of data on thousands of machines, hundreds of Map reduce program have been implemented and thousands of Map Reduce Job are executed on Google's clustered every day. The authors have learned many things from this work first restricting the programming model makes it easy to distributed computations and second network bandwidth is a very valuable resources in distributed computations, by the map reduce job sharing the network and increasing the availability of network. Third, redundant execution can be reducing the impact of slow machine and handle machine failure and data loss.

Condie Tyson,Conway Neil, Alvaro Peter, Hellerstein M. Joseph et al. [28] conducted a study of MapReduce framework that is a popular model for large data set for parallel programming by using this property, he proposed a modified model of MapReduce, in which intermediate data is pipelined between operations. To check this model the

ISSN: 2350-0328
International Journal of Advanced Research in Science, Engineering and Technology
Vol. 6, Issue 8 , August 2019

author design a model i.e hadoop online prototype (HOP), a pipeline version of Hadoop, this extends the MapReduce programming model beyond batch processing can reduce the completion time and utilization of system. HOP supports for event monitoring and streaming processing.
Lemeniz Infortech [29] proposed a model of mapReduce cloud services. i.e Cura to provide cost effective mapReduceservice in a cloud over the existing cloud service Cura have multiple benefits i.e first: minimize the infrastructure cost, minimize the number of servers requirements and resource utilization because proposed system can configure automatically configurations of required system as per the requirements.
Mridual Mrigank, Khajuria Akashdeep, Kumar Dutta Snehasish, M.R Prasad et al. [30] did the exploratory study about the solution of big data analysis with the use of Hadoop and Map Reduce. The authors had seen the existing problems in business and many others fields the growing of data is very fast and the native tools of the data handling are not sufficient to handle the unstructured and structured data. In this study they proposed a model to analysis the big data with the use of Hadoop and MapReduce. The most sources of big data are web logs, RFID generated data, sensors network data social network data, Internet search indexing etc.
Aarnio Helsinki [31], present the study of computational problem with two functions, Map and Reduce. The actual parallel processing define in the Map Reduce library that assign the Map and Reduce job to the dedicated nodes to process the job. The map reduce programming is used to represents the problems of parallelization, load balancing, fault tolerance and by using this the excellent performance could be achieve with single computer.
He Chen, Lu Ying, David Swanson et al. [32] proposed a new MapReduce program to schedule the map reduce task. The authors introduced this program with default Hadoop FIFO scheduler and Hadoop fair scheduler. They have compared this experiment with MapReduce scheduling technique and without MapReduce scheduling technique. Experimental result demonstrated that their matchmaking algorithm obtain the highest data locality rates and the lowest average response time for map task.

III. DATA INTEGRATION CHALLENGES OF BIG DATA

The data Integration is a big challenge for any organization to provide the data as high availability due to dynamic generation of big data from distinguish sources in various formats. In general, organizations are required to integrate data of various departments to a central place for their better decision making. To integrate the data at a central point a specific architecture or framework is required which can maintain the data as well as can do the integration of data if required simultaneously.
The following key challenges of data integration

1) Volume of data is very huge.
2) Format of data (structured and unstructured)
3) Heterogeneous data sources.
4) Maintenance of data (Live Streaming).
5) Incremental backup.

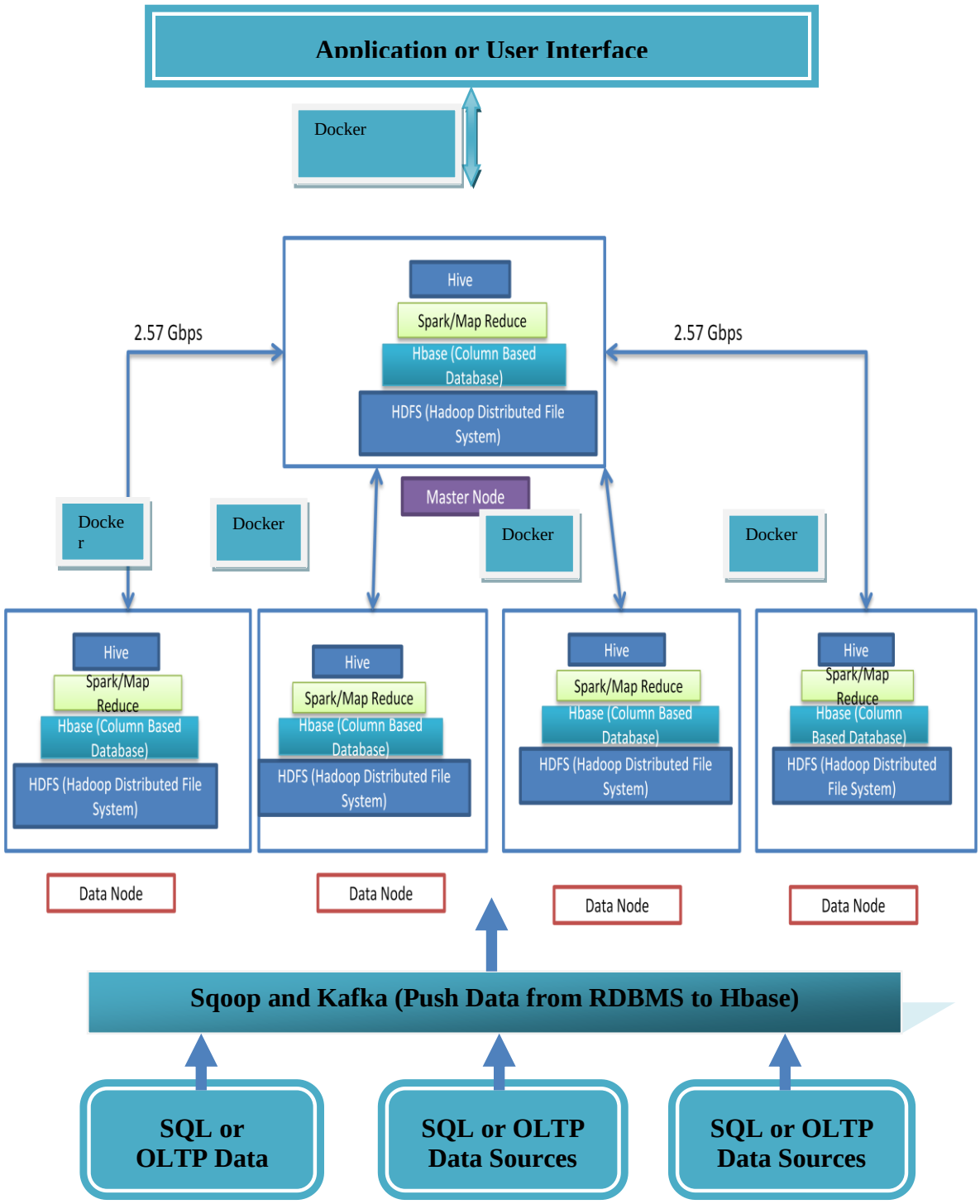
To overcome these challenges there is a need to develop framework or any architecture. There are many open source tools available in market to integrate big data like apache Sqoop, apache Kafka etc. With the help of these tools I have define a architecture to integrate and streaming big data.

IV. FLOW OF DATA INTEGRATION IN BIG DATA

To integrate structured and unstructured data in big data firstly it is important to know about the sources of data and formats of data. After identification of sources of data, the following steps are to follow:
Step 1: Extract the data from different-2 sources.
Step 2: Convert format of source data to big data formats.
Step 3: Store formatted data in big data warehouse.

After data integration in one instance the next challenge is to stream the data.

ISSN: 2350-0328
International Journal of Advanced Research in Science, Engineering and Technology
Vol. 6, Issue 8 , August 2019



I.

Figure 1: Data Integration Architecture of Big Data

In above figure-1, the data integration architecture of big data with help of open source tools has been laid out. In this architecture the data will continue replicate from online data transactions including legacy data replication.

ISSN: 2350-0328

International Journal of Advanced Research in Science, Engineering and Technology

Vol. 6, Issue 8 , August 2019

In architecture Sqoop and Kafka tools has been used for continuous replication of legacy data and after replication this data will be stored in Hbase. However, the **docker** has been used as container for clustering purpose in the above architecture. On other hand, the light weight resource utilization technique has been used to minimize the dependency. In proposed architecture Sparks and Map Reduce have been used for processing of data

V. CONCLUSION

Now day's data is generating very fast from various sources in different formats such as text data, audio and video data, image data etc. Hence, data integration is the biggest challenge in any big data environment. From the previous studies, it is clearly indicated that the existing integration techniques are not capable enough to integrate the different types of data in big data environment in efficient way or there is no standard framework or architecture available to support the different types of data integration in big data environment. Therefore, some framework or architecture is required to support the data integration. In the present study an attempt has been made to propose architecture for different type of data integration in big data environment. In the present study authors have discussed the process flow of big data integration through the proposed architecture for data integration of big data. The proposed architecture contains docker as container for clustering and Sparks, Map Reduce have been used for processing of data.

REFERENCES

[1] IDC-2014, "The Digital Universe of Opportunities: Rich Data and the Increasing Value of the Internet of Things", 2014.
[2] Lu, W., et al., "Efficient processing of k nearest neighbor joins using Map Reduce, Proceedings of the VLDB Endowment", 5(10), pp. 1016-1027, 2012.
[3] Liu, Y., et al., "HSim: A Map Reduce simulator in enabling cloud computing". Future Generation Computer Systems, vol. 29, no.1: pp. 300-308, 2013.
[4] Costa C. and Santos M.Y., "Improving Cities Sustainability through the Use of Data Mining in a Context of Big City Data," in Lecture Notes in Engineering and Computer Science: Proceedings of The World Congress on Engineering 2015, London, U.K., pp. 320–325, 2015.
[5] Girtelschmid S., Steinbauer M., Kumar V., Fensel A., and Kotsis G., "Big Data in Large Scale Intelligent Smart City Installations," in Proceedings of International Conference on Information Integration and Web-based Applications; Services, New York, NY, USA, pp. 428-432, 2013.
[6] Simmhan Y., Aman S., Kumbhare A., Liu R., Stevens S., Zhou Q., and Prasanna V., "Cloud-Based Software Platform for Big Data Analytics in Smart Grids," Computing in Science Engineering, vol. 15, no. 4, pp. 38–47, 2013.
[7] Ke Xu, et al., "A MapReduce based Parallel SVM for Email Classification. Journal of Networks",vol. 9, 6: pp. 1640-1647, 2014.
[8] Liang Jiajun, Yang Jian, "Big Data Application in Education Dropout Prediction in Edx MOOCs", IEEE Second International Conference on Multimedia Big Data, 2016.
[9] Zin thi thi,Tin Pyke, Toriu Takashi and Hamehiromitueal "A Big Data Application Framework for Consumer Behavior Analysis" IEEE 2nd Global Conference on Consumer Electronics, 2013.
[10] Riedel M., Memon AS. and M.S., "High Productivity Data Processing Analysis Methods with Application" MIPRO 2014,pp. 26-30, May 2014.
[11] Dem Chinko Yuri,Turkmen Fataih,Laat dechees and Blanchet Christophe, "Cloud Based Big Data Infrastructure: Architectural Components and Automated Provisioning" ,978-15090-2088-1/16 IEEE, 2016.
[12] Xu Brain, Kumar Alam Palayam, Satish, "Big Data Analytics Framework for System Health Monitoring", 2015 IEEE International Congress on Big Data, 2015.
[13] Gu Lin, Tao Sheng, Zang Deze and Hai Jin, "Communication Cost Efficient Virtualized Network Function Placement for Big Data Processing", IEEE INFOCOM First International Workshop on Big Data Science, Technologies and Applications (BDSTA), 2016.
[14] Wang Ziqi and Zhao Haihui, "Empirical Study of Using Big Data for Business Process Improvement at Private Manufacturing Firm in Cloud Computing", 2016 IEEE 3rd International Conference on Cyber Security and Cloud Computing, 2016.
[15] Govindarajan K., Somasundaram S.T., Boulanger D., and Kumar S.V., "A Framework for Scheduling and Managing Big Data Applications in a Distributed Infrastructure", 2015 Seventh International Conference on Advance Computing (ICoAC), 2015.
[16] Dubey V., and Pradhan C., "Project Monitoring System for Big Data", 2016 Symposium on Colossal Data Analysis and Networking (CDAN), 2016.
[17] Bakshi K., "Considerations for Big Data: Architecture and Approach", 978-1-4577-05557-1/12 IEEE, 2012.
[18] Deshmukh S. and Sumeet S., "Big Data Analytics using Public Cloud Infrastructure: Use cases and Cost Economics", 2015 International Conference on Computational Intelligence and Communication Networks, 2015.
[19] Phaneendra S.Vikram & Reddy E.Madhusudhan "Big Data- solutions for RDBMS problems- A survey" In 12th I EEE/IFIP Network Operations & Management Symposium (NOMS 2010) (Osaka, Japan, Apr 19, 2013.
[20] Kumara Reddi Kiran & Indira Dnvsl "Different Technique to Transfer Big Data : Survey" IEEE Transactions on vol. 52, no. 8, , pp. 2348-2355, Aug.2013.
[21] Fan Wei, Bifet Albert "Mining Big Data:- Current status and forecast to the future", Huawei Noah's Ark Lab,Yahoo! Research Barcelona Av. Diagonal 177 Barcelona, Catalonia, Spain.
[22] Bifet Albert "Mining Big Data In Real Time" Yahoo! Research Barcelona Avinguda Diagonal 177, Catalonia.
[23] Purcell Bernice "The emergence of "big data" technology and analytics" Journal of Technology Research 2013.

ISSN: 2350-0328

International Journal of Advanced Research in Science, Engineering and Technology

Vol. 6, Issue 8 , August 2019

[24] Agarwal Sameer, MozafariX Barzan, Panda Aurojit, Milner† Henry, MaddenXSamuel, Stoica Ion "BlinkDB: Queries with Bounded Errors and Bounded Response Times on Very Large Data" Copyright © 2013i ACM 978-1-4503-1994.
[25] Sharma Ashish, Vyas Snehlata "Hadoop2 Yarn" Computer Science Jodhpur National University , Jodhpur, Rajasthan, India.
[26] Eldawy Ahmed, Mokbel Mohamed F. "A Demonstration of SpatialHadoop:An Efficient MapReduce Framework for Spatial Data" Proceedings of the VLDB Endowment, vol. 6, no. 12 Copyright 2013 VLDB Endowment 21508097/13/10.
[27] Dean Jeffrey and Ghemawat Sanjay "MapReduce: Simplified Data Processing on Large Clusters" OSDI 2010.
[28] Condie Tyson, Conway Neil, Alvaro Peter Hellerstein M. Joseph "Online Aggregation for Large MapReduce Jobs" June 10, 2010, Inianappolis, Indiana USA. Copyright 2010 ACM 978-1-4503-0032- 2/10/06.
[29] Lemeniz Infotech Cost-effective Resource Provisioning for MapReduce in a Cloud" gartner report, 25, 2010.
[30] Mridul Mrigank, Khajuria Akashdeep, Dutta Snehasish, N Kumar " Analysis of Bidgata using Apache Hadoop and Map Reduce" vol. 4, no.5, May 2014.
[31] Aarnio Helsinki, "Parallel Data Processing with MapReduce: A Survey" Helsinki University of Techology, TKK T-110.5190 Seminar on Internetworking, pp. 4-27, 2009.
[32] He Chen, Lu Ying, Swanson David "Matchmaking: A New MapReduce Scheduling" in 10th IEEE International Conference on Computer and Information Technology (CIT'10), pp. 2736–2743, 2010.

AUTHOR'S PROFILE

Rahul Kumar is Sr. Software Engineer, National Informatics Center New Delhi. He is pursuing his Ph.D. from Mangalayatan University, Aligarh and the Master degree (MCA) had done from there with 89.10%. He is also working as new technology advisor and also working with Big data Analytics team at National Transport Project, National Informatics Center HQ Delhi. He had 3+ Year of extensive experience of Java developer.

Dr. Subodh Kumar is Head in Department of Computer Applications, Institute of Engineering and Technology, Mangalayatan University, Aligarh. He has obtained his Ph. D. from Mangalayatan University, Aligarh and the Master degree in computer application from Dr. A.P.J. Abdul Kalam Technical University, Uttar Pradesh, Lucknow (formerly Uttrapradesh Technical University). He has been meticulously carried out research activities since last 9 years in the field of Mobilie Ad-hoc Network, Big Data, Software Engineering and Database. He has taught various courses at graduate as well as at postgraduate level and supervised number of projects and dissertations of BCA & MCA students. Presently he is supervising five Ph. D scholars in the field of Big Data, Mobile Ad-hoc Network, and Cyber Security. He has published 15 research papers in International Journals of repute, and has been part of many National and International conferences.

Dr. Saurabh Gupta is Senior Technicial Director and SIO -UP, National Informatics Center, and Ph.D. in Computer Engineering, Degree in Management and Law with Certification in Management Development Programmed from IIM, Bangalore and Strategic management from IIM Lucknow. Besides this he has many certification such as Synthetic Aperture Ruder Technology (SAR) from IIT- Bombay, Certification in "Certified Software Quality Professional" from STQC-Delhi, Certification in "Internal Quality Auditor for ISO" from STQC , Certification in "Accounts and Administration" from ISTM Delhi and Certification on Decentralized Planning under 73rdAmendment of Constitution in SIRD, Lucknow.