International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 07 Issue: 04 | Apr 2020 www.irjet.net p-ISSN: 2395-0072

The Upsurge of Cloud Computing and Big Data

Mr. Vinay Sripathi[1], Mr. Miheer Samant[2]

[1]*Department of Electronics and Telecommunication Engineering, Rajiv Gandhi Institute of Technology, Mumbai, Maharashtra, India*
[2]*Department of Electronics and Telecommunication Engineering, Ramrao Adik Institute of Technology, Mumbai, Maharashtra, India*

Abstract - *Big Data has emerged on the scene in the initial years of 21st century. Companies like Google, LinkedIn, Facebook, eBay were created around big data from their beginning. It has become an uprising Model which provides large data and opportunities to progress and enables research, decision making in almost all the branches of study. While allowing for all these opportunities it is difficult for firms to store, process, transport, mine and serve the data. To minimize these difficulties the Cloud computing is introduced, which basically provides essential support to address the challenges with shared resources such as computing, storages, networking and software based on analytics. This paper investigates the two edges – Big Data and cloud computing – and analyses the advantages and significances of utilizing cloud computing to tackling Big Data in the digital world and applicable science provinces. While Big Data is accountable for data storage and processing, the cloud provides a dependable, accessible, and scalable environment for Big Data systems to operate. Big Data is defined as the amount of digital data produced from diverse sources of technology, for example, sensors, digitizers, scanners, numerical modeling, mobile phones, Internet, videos, social networks. Cloud Computing and Big Data are corresponding to each other. Swift development in Big Data is observed as a problem. Clouds are developing and providing solutions for the suitable environment of Big Data while outdated storage cannot meet the necessities for dealing with Big Data, in addition to the essential for data exchange between several scattered storage sites. Cloud Computing provides solutions and addresses difficulties with Big Data. Big data and Cloud computing both the technologies are respected on its own. Furthermore, many trades are targeting to combine the two techniques to gain more business benefits. Both the technologies intention to enhance the profits of the company while reducing the investment cost. While Cloud achieves the local software, Big data helps in business decisions. In paper presents the connection between Big Data and Cloud Computing, Cloud Computing role of Big Data, advantages of Big Data and Cloud computing, cloud architecture, importance of Cloud Computing.*

Key Words: Big Data, Cloud Computing, Hadoop, Analytics, NoSQL, IT

1. INTRODUCTION

Big data fundamentally mentions, sets of data that are huge in size and cannot be treated through traditional application software. The term big data is not new, as it has been around for a long time and there are many new thoughts related to the term. Albeit, the concept is not new in the industry, there is a lot of misunderstanding around the exact meaning of what big data actually is. When you work on a specific principle and start gathering knowledge on it, you start producing data that will be useful for you in the future to examine and get further understandings. Cloud computing propose the finest technology with a wide range of applications for numerous purposes in the most cost-effective way. Big data and cloud computing go hand in hand because there is a great deal of data — and only cloud computing can deliver that kind of compute power to process the data. Whatever we do nearly leaves a digital track, as we generate data whenever we are on the internet. As cloud computing is transforming IT, huge quantities of compute power are desired with the help of the internet to store and examine this data. Cloud computing has redesigned the way computers are being used to process data. Cloud has made it very simple for data storage compared to outdated data storage [2]. Cloud computing offers scalable resources on request and has changed the way data is stored and handled. Cloud computing is an influential approach to examine data provided and has become vital in the growth of big data in multiple industries.

Today, two mainstream technologies are the epicenter of concern in IT those are Big Data and Cloud Computing. Primarily different, Big data is all about dealing with the enormous scale of data whereas Cloud computing is about infrastructure. Yet, the simplification offered by Big data and Cloud technology is the main reason for their huge enterprise adoption. For example, Amazon "Elastic Map Reduce" reveals how the power of Cloud Elastic Computes is leveraged for Big Data processing. The amalgamation of both yields beneficial outcome for the organizations [3]. Not to mention, both the technologies are in the stage of progress but their combination leverages scalable and cost-effective solution in big data analytics.

With an ever-escalating reach to the masses, a substantial reduction in cost and a copious choice of applications

available, you can be truly empowered to leverage the best of existing technology, often without spending a penny on initial investment. The democratization of information technology has not only affected the cloud space, but big data as well. Implementation of open source Hadoop is growing at a fast pace and the ability to accomplish analytics on non-proprietary and inexpensive hardware is becoming more ubiquitous. Along with this phenomenon, we are now observing a burst of information generated through social media, messaging, emails and more. Organizations and individuals are directing a maze of ever-increasing data that can be tough to roam through, let alone control and dissect [5]. This surge in the volume of data is now presenting a challenge to the cloud. Establishments have built their data architecture, storage policies and best practices mainly working with structured data, whereas the unstructured data does not fit the traditional relational database management system (RDBMS) framework.

The matter is how to manipulate and extract the core of the data rather than simply storing and retrieving it. As pointed out during IBM InterConnect 2013, businesses can get amplified value from data insights gained through big data analytics supported by a cloud infrastructure. The burst in unstructured data means that ways to harness the benefits of hybrid cloud and big data are more vital than ever. The hybrid cloud model can assist organizations in addressing security concerns in their private cloud, while leveraging the public cloud infrastructure for analytics facilities. It comes as no surprise that there has been a keen curiosity from government agencies, government bodies and other organizations to try and extract expressive insight from this maze of data, whether it be security associated or simply about patterns of consumers. On average, 2.5 billion gigabytes of data are generated daily, containing of 200 million tweets and 30 billion pieces of content shared on Facebook each month. Looking at the available forecasts, the amount of data created by the year 2021 will reach a staggering 60 trillion gigabytes, with six billion people in possession of cell phones. Cloud computing and big data, while still in continuous development, are proving to be the superlative combination. Together, they provide a cost effective and scalable infrastructure to bolster big data and business analytics.

2. Evaluation of Big Data

"You can have data lacking information, but you cannot have information lacking data." – Daniel Keys Moran

The above quote describes the importance of data. Disregarding the importance of big data can lead to be a very costly error for any kind of business in today's world. If data is that vital then using effective analytics or big data tools to unlock the hidden power of data becomes imperious. Here we will discuss the assistances of using cloud computing for big data. If you have followed our previous blogs, we have discussed at length the value of big data and here we will explore it even further [8].

Today, every association, government, IT firm and political party considers data as a new and very useful currency. They readily invest resources to unlock insights from gathered data in their individual fields which can be profitable if it is satisfactorily mined, stored and analyzed [9]. The initial stages of using big data were mostly based around packing the data and applying some basic analytics modules. Now, as the exercise has evolved, we have implemented more advanced approaches of modeling, transforming, and extracting on a much larger scale. The field of big data now has the volume for a globalized infrastructure. Internet and social media goliaths such as Google and Facebook were the forerunners of big data when they began uncovering, assembling and analyzing information collected by their users. Back then companies and researchers entered worked with externally sourced data, which was fundamentally drawn from the "internet" or "public data sources". The term "big data" wasn't created until 2010 roughly when they comprehended the power, need and importance of this information. Given the room of information, the term "big data" come into the picture. And with that, the advent of newly developed technologies and processes to help companies to turn the data into insight and profit [10, 11].

The term Big Data is being swiftly used almost everywhere across the planet – online and offline. Before that, information stored on your servers or computers was only sorted and filed. But today, all data converts big data no matter where you have stored it or in which format [12].

Fig.1- Big Data Infographic

3. The Relationship Between Big Data and Cloud Computing

With the generation of a vast amount of data, cloud computing is playing an important role in the storage and management of that data. It's not only about the growth of big data but also the extension of data analytics platforms like Hadoop. As a result, it is creating new openings in Cloud computing. Hence, the service providers like AWS, Google and Microsoft are offering their own big data systems in a cost-efficient manner which is accessible for businesses of all sizes.

This, in turn, has led to a new service model which is known as *Analytics as a Service* (*AaaS*). This will provide a quicker and scalable way to integrate different types of structured, semi-structured and unstructured data, analyze them, transform and envision them in real time [13].

Additionally, Big data cloud computing relationship can be evaluated from below viewpoints and benefits –

1. A cloud computing environment generally has several user terminals and service providers. From the collection terminals, the user gathers the data using the big data tools. On the contrary, from the service provider ends saves, stores and processes the big data. Hence, cloud computing provides a big data infrastructure. The infrastructure must deliver on-demand resources and services to ensure continuous service.
2. Since the cloud environment is scalable, hence it can provide adequate data management solution irrespective of the volume of the data. If the necessary cloud computing service provider can also offer security policies as per the user demands.
3. Identity management and access control are two main concerns while dealing with confidential company data. Cloud computing can meet this security necessity using a simple software interface by abstracting internal details of the information. Additionally, this guarantees complete confidentiality of user data and only offers access to the authorized users.
4. Big data for data processing can be situated across the global locations and preserving such huge servers in different locations is a costly portion for an organization. As cloud computing can store and procedure data through geographically detached and as well as virtual servers it reduces the cost of big data processing significantly.
5. Cloud computing uses high-level software and applications which do not depend on the efficiency of the user devices. Additionally, it depends on the network servers and their strength. On the contrary, if we use personal resources for big data that will be dependent on the user device. Hence, big data cloud computing service is helpful.
6. Cloud computing permits high-speed data flow over the network. As a result, it causes faster big data processing [14-15].

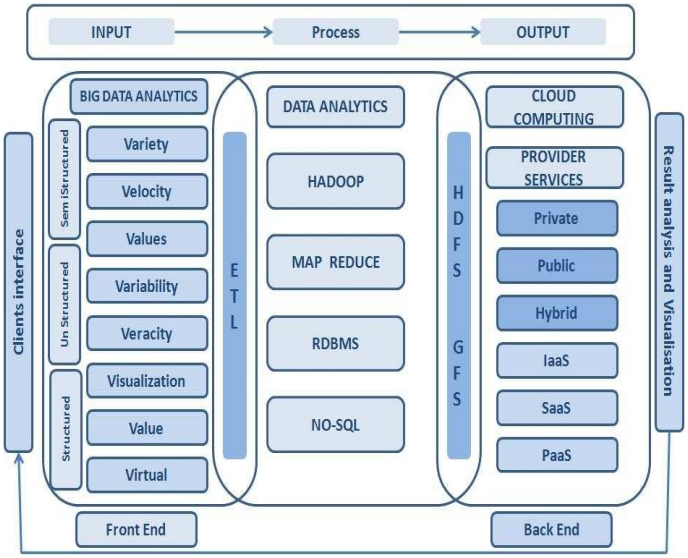

Fig.2- Connection between Big Data and Cloud Computing.

4. Cloud Computing Role for Big Data

Big data and Cloud computing relationship can be categorized based on service types:

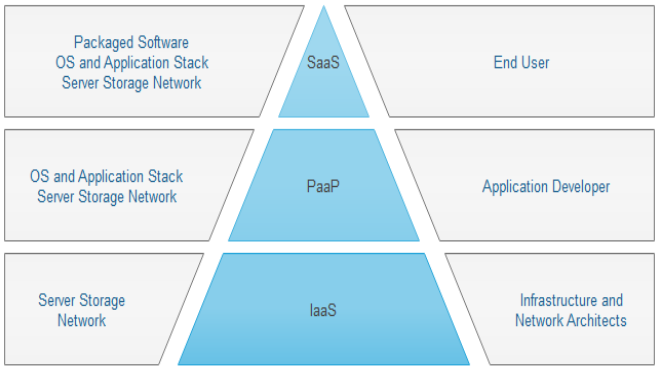

Fig.3- Cloud computing role on Big Data

4.1. IAAS in Public Cloud

IaaS is a cost-effective answer and employing this Cloud service, Big Data facilities enable people to access limitless storage and compute power. It is a very cost-effective solution for enterprises where the Cloud provider bears all the expenditures of handling underlying hardware.

4.2. PAAS in Private Cloud

PaaS sellers incorporate Big Data technologies into their accessible service. Hence, they eradicate the need for dealing with the difficulties of managing single software and hardware elements which is a real worry while dealing with terabytes of data.

4.3. SAAS in Hybrid Cloud

Examining social media data is nowadays an essential parameter for companies for business analysis. In this framework, SaaS vendors provide an excellent platform for conducting the analysis [16].

5. Benefits of Big Data Analysis in Cloud

Hence, from the above explanation, we can see that Cloud enables "As-a-Service" pattern by conceptualizing the challenges and difficulty through a scalable and elastic self-service application. Big data requirement is same where distributed processing of huge data is abstracted from the end users.

There are multiple benefits of Big data analysis in Cloud.

5.1. Enhanced Analysis

With the development of Cloud technology, big data analysis has become further improved causing improved results. Hence, companies choose to perform big data analysis in the Cloud. Moreover, Cloud helps to integrate data from numerous sources.

5.2. Easy Infrastructure

Big Data analysis is an incredible strenuous job on infrastructure as the data comes in huge volumes with varying speeds, and types which traditional infrastructures frequently cannot keep up with. As the Cloud computing provides flexible infrastructure, which we can scale according to the needs at the time, it is easy to manage workloads [17].

5.3. Security and Privacy

Data security and privacy are two main concerns when dealing with enterprise data. Moreover, when your application is presented on a Cloud platform due to its open environment and limited user control security becomes a key concern. On the other hand, being an open source application, Big data answer like Hadoop uses a lot of third-party services and infrastructure. Hence, these days system integrators bring in Private Cloud Solution that is Elastic and Scalable. Likewise, it also leverages Scalable Distributed Processing.

Besides that, Cloud data is stored and treated in a central location commonly known as Cloud storage server. Along with it the service provider and the customer sign a service level agreement (SLA) to gain the trust among them. If require the provider also leverages essential advanced level of security control. This enables the security of big data in Cloud computing covering the following issues:

1. Protecting big data from progressive threats.
2. How Cloud service providers sustain storage and data.

There are rules associated with service level agreements for shielding. Those are Data, Capacity, Security, Scalability, Privacy, Availability of data storage and data growth.

On the other hand, in many organizations, big data analytics is operated to detect and prevent advanced threats and malicious hackers.

5.4. Virtualization

Infrastructure plays a vital role to bolster any application. Virtualization technology is the perfect platform for big data. Virtualized big data applications like Hadoop deliver multiple benefits which are not available on physical infrastructure, but it simplifies big data Management [18]. Big data and Cloud computing point to the junction of various technologies and trends that makes IT infrastructure and related applications more dynamic, more expendable and more modular and. Hence, Big data and Cloud computing projects rely heavily on virtualization.

6. Advantages of Cloud Computing and Big Data

Cloud computing and big data is a perfect amalgamation as it delivers a solution which is both scalable and accommodating for big data and business analytics. Visualize a world where all the information resources are easily reachable, and each aspect of life can benefit from this information. Let's take a look at these advantages in detail:

6.1. Agility

The traditional technique of storing and managing data is becoming out-of-date quickly. Setting up an infrastructure is not only expensive but also time-consuming, as installing and running a server can take weeks. With cloud computing, it's likely to offer any infrastructure with all the obligatory resources almost instantly. A good cloud provider will company to ensure that their work is constantly on the go without any complications.

6.2. Elasticity

A cloud podium can dynamically enlarge to afford storage for ever cumulative data. Once the company gets the essential insight from the data, storage space can be increased or reduced to accommodate the data as per the condition.

6.3. Data Processing

A huge capacity of data leads to the matter of how to process it efficiently. Social media alone produce an enormous amount of unstructured data in various forms. With Big Data platforms, cloud computing makes the entire process easier and accessible to small, medium and large enterprises.

6.4. Cost cutting with Big Data in the Cloud

Cloud computing is a great solution for enterprises that wish to have state of the art technology running their operations under a restricted budget. Preserving a big data center to perform Big Data analytics can swiftly drain an IT budget. Currently, companies have the option to evade investing heavily in setting up the IT department and preserving hardware infrastructure. With the cloud computing, the concern shifts to the cloud providers and the company only have to pay for the storage space and power consumption.

6.5. Reduced Complexity

Any execution of big data solution needs quite a few components and integrations. Cloud computing delivers the option to automate these components, thus dropping complexity and enhancing the productivity of the Big Data analysis team [19-20-21].

7. Cloud Big Data Challenges

Vertical scaling attains elasticity by totaling additional instances with each of them serving a part of the demand.

Software like Hadoop are precisely designed as distributed systems to take benefit of vertical scaling. They process minor independent tasks in huge parallel scale. Distributed systems can also assist as data stores like NoSQL databases, e.g. Cassandra or HBase, or filesystems like Hadoop's HDFS. Substitutes like Storm provide coordinated stream data processes in near real-time through a cluster of machines with complex workflows.

The inter-changeability of the resources composed with distributed software design counters failure and equivalently scaling of virtual computing instances unperturbed. It is conceivable to crawl massive web sites with millions of pages in days or hours for a few hundred dollars or less. Inexpensive tiny virtual instances with minimal CPU resources are ideal for this purpose since the majority of crawling the web is spent waiting for IO resources. Instantiating thousands of these machines to reach millions of requests per day is easy and often costs less than a fraction of a cent per instance hour [23].

Of course, such mining operations should be mindful of the resources of the web sites or application interfaces they mine, respect their terms, and not hinder their service. A poorly deliberated data mining operation is equivalent to a denial of service attack. Lastly, cloud computing is naturally a good fit for storing and processing the big data accumulated form such operations [24-25].

8. Cloud Architecture

The 3 main cloud architecture models have developed over time; private, public and hybrid cloud. They all share the idea of reserve commodification and to that end usually virtualize computing and abstract storage layers.

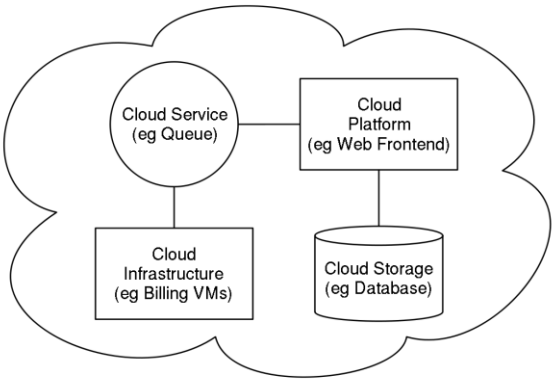

Fig.4- Cloud Computing Architecture Diagrams.

8.1. Private Cloud

Private clouds are devoted to one organization and do not share physical resources. The resource can be provided in-house or outside. A typical primary requirement of private cloud arrangements are security requirements and regulations that need a strict separation of an organization's data storage and processing from accidental or malevolent access through shared resources [26]. Private cloud setups are stimulating since the economic advantages of scale are usually not achievable within most projects and organizations regardless of the utilization of industry standards. The return of investment associated to public cloud offerings is rarely obtained and the operational overhead and risk of failure is significant.

Additionally, cloud providers have captured the inclination for increased security and provide special environments, i.e. dedicated hardware to rent and encrypt virtual private networks as well as encrypted storage to report most security concerns. Cloud providers may also offer data storage, transfer, and processing limited to specific geographic regions to ensure agreement with local privacy laws and regulations [27].

Alternative motive for private cloud deployments are legacy systems with special hardware needs or exceptional resource demand, e.g. extreme memory or computing cases which are not available in public clouds. These are effective concerns though if these demands are extraordinary the question if a cloud architecture is the accurate solution has to be raised. One reason can be to launch a private cloud for a transitionary period to run legacy and challenging systems in parallel while their services are ported to a cloud environment culminating in a switch to a cheaper public or hybrid cloud.

8.2. Public Cloud

Public clouds share physical resources for data transfers, storage, and processing. Though, customers have private visualized computing environments and isolated storage. Safekeeping concerns, which lure a few to adopt private clouds or custom deployments, are for the vast majority of customers and projects inappropriate. Visualization makes access to other customers' data very difficult.

Real-world problems around public cloud computing are more ordinary like data lock-in and fluctuating performance of individual instances. The data lock-in is a soft measure and works by making data inflow to the cloud provider free or very inexpensive. The replication of data out to local systems or other providers is often more expensive. This is not an insoluble problem and in practice encourages to employ more services from a cloud provider instead of moving data in and out for different services or processes. Typically, this is not sensible anyway due to network speed and difficulties around dealing with multiple platforms.

The changing performance of instances stems typically from the dependency on what kind of load other customers produce on the shared physical infrastructure. Secondly, over time the bodily infrastructure providing the virtual resources changes and is updated. The accessible resources for each customer on a physical machine are usually throttled to ensure that each customer receives a guaranteed level of performance. Larger resources generally deliver very expectable performance since they are much closer aligned with the physical instance's performance. Horizontally scaling projects with small instance should not rely on an exact performance of each case but be adaptive and focus on the average performance vital and scale according to need.

8.3. Hybrid Cloud

The hybrid cloud architecture combines private and public cloud deployments. This is frequently an attempt to achieve security and elasticity, or deliver cheaper base load and burst capabilities. Some organizations experience short periods of extremely high loads, e.g. as an outcome of seasonality like black Friday for retail, or marketing events like sponsoring a popular TV event. These events can have enormous economic influence to organizations if they are serviced poorly.

The hybrid cloud delivers the chance to serve the base load with in-house services and rent for a short period a multiple of the resources to service the extreme demand. This requires a great deal of operational capability in the organization to flawlessly scale between the private and public cloud. Tools for hybrid or private cloud deployments exist like Eucalyptus for Amazon Web Services. On the long-term the supplementary expense of the hybrid approach often is not justifiable since cloud providers offer big discounts for multi-year commitments. This makes moving base load services to the public cloud striking since it is escorted by a simpler deployment strategy [28].

9. Importance of Cloud Computing for Big Data

There are numerous reasons for having a big data on cloud. Some of them are deliberated below:

9.1. Immediate Infrastructure

One of the key aids of a cloud-based method to big data analytics is the skill to establish big data infrastructure as quickly as likely with a scalable environment. A big data cloud service delivers the infrastructure that companies would otherwise have to build up themselves from scratch.

Big data offers all analytics needs in a single roof. It is significant to note that cloud-based big data analytics success is dependent on many crucial factors. Most significant of these is the quality and dependability of the solution provider. The seller must combine robust, extensive expertise in both the big data and cloud computing sectors.

9.2. Cutting Costs with Big Data in the Cloud

Putting the big data analytics on the cloud, permits firms to cut costs in terms of acquiring equipment, cooling machines and warranting security, while also letting them to keep the most sensitive data on-premise and have the full control on it.

9.3. Fast Time to Value

A modern data-management platform carries together master data management and big data analytics abilities in the cloud so that business can create data-driven applications using the dependable data with relevant insights. The principal advantage of this combined cloud platform is faster time-to-value, keeping up with the pace of business [29]. Whenever there is a need for a new, data-driven decision management application, you can make one in the cloud quickly. There is no need to set up infrastructure [hardware, operating systems, databases, application servers, analytics], create new integrations, or define data models or data uploads. In the cloud, necessities are already set up and available [30].

10. Conclusion

Big Data has arisen in the past few years as a new standard providing abundant data and opportunities to progress and/or enable research and decision-support applications with extraordinary value for digital earth applications

including business, sciences and engineering. At the same time, Big Data presents tasks for digital earth to store, transport, process, mine and serve the data. Cloud computing provides essential support to address the challenges with shared computing resources counting computing, storage, networking and analytical software; the application of these resources has fostered impressive Big Data improvements. Cloud computing is an influential technology to achieve massive-scale and complex computing. It removes the need to maintain expensive computing hardware, devoted space, and software. Massive evolution in the scale of data or big data produced through cloud computing has been detected. Addressing big data is a challenging and time-demanding task that needs a large computational infrastructure to confirm successful data processing and analysis. Cloud computing delivers enterprises a cost-effective & flexible way to access a huge volume of information we call the Big Data. Because of Big Data and cloud computing, it is now much easier to start an IT company than ever earlier.

REFERENCES

[1] Villars, R. L., Olofson, C. W., & Eastwood, M (June, 2011). Big data: What it is and why you should care. IDC White Paper. Framingham, MA: IDC.

[2] Shakil, K. A.; Sethi, S.; Alam, M., (2015). An effective framework for managing university data using a cloud-based environment, Computing for Sustainable Global Development (INDIACom), 2nd International Conference on, vol., no., pp. 1262, 1266, 11-13.

[3] K. Das and P. M. Kumar, Big data analytics: A framework forum structured data analysis, International Journal of Engineering and Technology, 5 (1) (2013), pp. 153-156.

[4] T. K. Das, D. P. Acharjya and M. R. Patra, Opinion mining about aproduct by analyzing public tweets in twitter, International Conference on Computer Communication and Informatics, 2014.

[5] M. K. Kakhani, S. Kakhani and S. R. Biradar, Research issues in bigdata analytics, International Journal of Application or Innovation in Engineering & Management, 2 (8) (2015), pp. 228-232.

[6] A. Gandomi and M. Haider, Beyond the hype: Big data concepts, meth-ods, and analytics, International Journal of Information Management, 35 (2) (2015), pp. 137-144.

[7] C. Lynch, Big data: How do your data grow? Nature, 455 (2008), pp. 28-29.

[8] X. Jin, B. W. Wah, X. Cheng and Y. Wang, Significance and challenges of big data research, Big Data Research, 2 (2) (2015), pp. 59-64.

[9] R. Kitchin, Big Data, new epistemologies and paradigm shifts, Big Data Society, 1 (1) (2014), pp. 1-12.

[10] C. L. Philip, Q. Chen and C. Y. Zhang, Data-intensive applications, challenges, techniques and technologies: A survey on big data, Infor-mation Sciences, 275 (2014), pp. 314-347.

[11] K. Kambatla, G. Kollias, V. Kumar and A. Gram, Trends in big data analytics, Journal of Parallel and Distributed Computing, 74 (7) (2014), pp. 2561-2573.

[12] S. Del. Rio, V. Lopez, J. M. Bentez and F. Herrera, On the use of mapreduce for imbalanced big data using random forest, Information Sciences, 285 (2014), pp. 112-137.

[13] MH. Kuo, T. Sahama, A. W. Kushniruk, E. M. Borycki and D.
K. Grunwell, Health big data analytics: current perspectives, challenges and potential solutions, International Journal of Big Data Intelligence, 1 (2014), pp. 114-126.

[14] R. Nambiar, A. Sethi, R. Bhardwaj and R. Vargheese, A look at challenges and opportunities of big data analytics in healthcare, IEEE International Conference on Big Data, 2013, pp. 17-22.

[15] Z. Huang, A fast clustering algorithm to cluster very large categorical data sets in data mining, SIGMOD Workshop on Research Issues on Data Mining and Knowledge Discovery, 1997.

[16] Chang, J. Dean, S. Ghemawat, W. C. Hsieh, D. A. Wallach, M. Burrows, T. Chandra, A. Fikes, and R. E. Gruber. Bigtable: A Distributed Storage System for Structured Data. In OSDI, pages 205–218, 2006.

[17] Rajeev Gupta, Himanshu Gupta, and Mukesh Mohania, "Cloud Computing and Big Data Analytics: What Is Newfrom Databases Perspective?". S. Srinivasa and V. Bhatnagar (Eds.): BDA 2012, LNCS 7678, pp. Springer-Verlag Berlin Heidelberg 42–61, 2012.

[18] Curino, C., Jones, E. P. C., Popa, R. A., Malviya, N., Wu, E., Madden, S., Balakrishnan, H., Zeldovich, N.: Realtional Cloud: A Database-as-a-Service for the Cloud. In: Proceedings of Conference on Innovative Data Systems Research, CIDR- 2011.

[19] Alberto Ferandez, Sara del R, Victoria opez, Abdullah Bawakid, Maria J. del Jesus, Jose M. Benitez, and Francisco Herrera. "Big Data with Cloud Computing: an insight on the computing environment, Map Reduce, and programming frameworks". doi: 10.1002/widm.1134. WIREs Data Mining Knowl Discov, 4:380–409, 2014.

[20] Lu, Huang, Ting-tin Hu, and Hai-shan Chen. "Research on Hadoop Cloud Computing Model and its Applications". Hangzhou, China: 2012, pp. 59 – 63, 21-

24 Oct. 2012.

[21] Wie, Jiang, Ravi V. T, and Agrawal G. "A Map-Reduce System with an Alternate API for Multi-core Environments.". Melbourne, VIC: 2010, pp. 84-93, 17-20 May. 2010.

[22] K, Chitharanjan, and Kala Karun A. "A review on hadoop - HDFS infrastructure exten-sions.". JeJu Island: 2013, pp. 132-137, 11-12 Apr. 2013.

[23] F. C. P, Muhtaroglu, Demir S, Obali M, and Girgin C. "Business model canvas perspective on big data applications." Big Data, 2013 IEEE International Conference, Silicon Valley, CA, Oct 6-9, p. 32 – 37, 2013.

[24] Castelino, C., Gandhi, D., Narula, H. G., & Chokshi, N. H. (2014). Integration of Big Data and Cloud Computing. International Journal of Engineering Trends and Technology (IJETT), 100-102.

[25] Chandrashekar, R., Kala, M., & Mane, D. (2015). Integration of Big Data in Cloud compu-ting environments for enhanced data processing capabilities. International Journal of Engineering Research and General Science, 240- 245.

[26] James Kobielus, I., & Bob Marcus, E. S. (2014). Deploying Big Data Analytics Applica-tions to the Cloud: Roadmap for Success. Cloud Standards Customer Council.

[27] M. Herland, T. M. Khoshgoftaar and R. Wald, A review of data mining using big data in health informatics, Journal of Big Data, 1 (2) (2014), pp. 1-35.

[28] X. Y. Chen and Z. G. Jin, Research on key technology and applicationsfor internet of things, Physics Procedia, 33, (2012), pp. 561-566.

[29] N. Mishra, C. Lin and H. Chang, A cognitive adopted frameworkfor iot big data management and knowledge discovery prospective, International Journal of Distributed Sensor Networks, 2015, (2015), pp. 1-13.

[30] Z. Hongjun, H. Wenning, H. Dengchao and M. Yuxing, Survey of research on information security in big data, Congresso da sociedada Brasileira de Computacao, 2014, pp. 1-6.

BIOGRAPHIES

VINAY SRIPATHI
Currently pursuing under graduation final year in the branch of EXTC at Rajiv Gandhi Institute of Technology, Versova.

MIHEER SAMANT
Currently pursuing under graduation final year in the branch of EXTC at Ramrao Adik Institute of Technology, Nerul.