Quest Journals
Journal of Software Engineering and Simulation
Volume 4 ~ Issue 2 (2018) pp: 39-45
ISSN(Online) :2321-3795 ISSN (Print):2321-3809
www.questjournals.org

Research Paper

The Positive and Negative Characteristics of Cloud Computing-Based Big Data Analytics for Business Intelligence.

Madhu N Y [1], Latha P H [2], Savitha N [3]
[1,2] *Department of Computer Science and Engineering, Government CPC polytechnic, Mysore, Karnataka, India.*
[3]*Department of Computer Science and Engineering, Government polytechnic, Channapatna, Karnataka, India.*

ABSTRACT
In recent years, cloud computing and big data analytics have become key technologies in the IT business. Surprisingly, combining the two technologies yields significant benefits for organizations. Cloud computing is transforming IT service delivery and user interaction with IT resources. Big Data is a data analysis approach made possible by recent advancements in information and communication technologies. Big data analysis demands significant computational resources, making it unaffordable for small and medium organizations. In this article, we discuss the benefits and pitfalls of delivering big data analytics using cloud computing. We claim that cloud computing can meet the storage and computational needs for big data analytics. Consolidating these two main technologies can boost big data mining and help organizations make better decisions. We emphasize the dangers and concerns to consider while implementing a cloud-based service model (CLaaS).
Keywords: *Cloud computing; Big Data; Business; Intelligence; CLaaS.*

I. INTRODUCTION

The business Intelligence (BI) encompasses tools and processes for collecting, integrating, analyzing, and presenting corporate information. company Intelligence aims to improve and accelerate company decision-making. Organizations must acquire, analyze, and use data to enhance decision-making and operations. In an ever-changing business environment, many firms are now under increasing pressure to build and scale up their business intelligence operations rapidly and affordably in order to remain competitive. Cloud computing, which has just evolved, is transforming the way organizations deliver IT services as well as how they interact with IT resources. It marks a paradigm change, introducing flexible service models to which businesses may subscribe on a pay-as-you-use basis.

The world's data is rising at an exponential rate. Big data is a dynamic word that refers to any massive volume of organized, semi-structured, and unstructured data that has the potential to be mined for relevant information. Big data refers to data that surpasses the processing capabilities of standard databases. The data is too large to process on a single system. Big data analytics is a rapidly developing area that studies massive volumes of data in search of hidden patterns, correlations, and other insights. Big data technology is now possible because to recent advances in computer technology, as well as algorithms and methodologies designed to manage large amounts of data.

The purpose of this study is to explore the effects of cloud computing and big data on organizations, as well as to assess the benefits and problems they present. First, we will discuss the ideas, challenges, and technologies of cloud computing and big data separately. We then provide a framework that integrates these two technologies to create an optimal platform for e-commerce. We examine how big data might help improve e-commerce's primary functional areas such as customer management, marketing, payments, supply chain, and management.

II. LITERATURE REVIEW

Cloud computing's popularity has sparked a number of academic and industrial projects to investigate its capabilities and advancements. One of the important study topics is comparing the value proposition of cloud computing to on-premise expenditures. There are various programs aimed especially at addressing security concerns and obstacles in cloud computing.

Several academic efforts have been launched to investigate the e-business model features of cloud computing. Aydin [2] explains his study on E-Commerce Based on Cloud Computing. Dan and Roger [4] investigated several cloud providers, including Google App Engine, Amazon EC2, and Microsoft Azure, to provide recommendations on cost, application performance (and restrictions), and deployment options. Agarwal et al. [1] propose a variety of strategies for dealing with large data analysis difficulties using the Map Reduce framework and the Hadoop Distributed File System. In this work, Map Reduce methods are used for Big Data analysis with HDFS. Yadav et al. [20] provide an overview of the architecture and methods utilized in big datasets. These algorithms explain numerous structures and methods used to manage Big Data, and this article covers the tools built to analyze them. It also discusses the many security challenges, applications, and trends observed in a huge data set. Fan and Bifet [6] provide an overview of big data mining, highlighting its current state, controversies, and future prospects. This article also covers a variety of intriguing and cutting-edge issues in Big Data mining. Sharma and Navdeti [15] explore large data security at the environmental level, as well as the testing of built-in defenses. It also discusses some of the security concerns that we are now facing and suggests security solutions and commercially available strategies to overcome them. The article also discusses all security strategies for securing the Hadoop environment. They also present an outline of big data, its relevance in our lives, and various tools for dealing with it. Jassena and David [8] address the concerns, challenges, and solutions to big data mining. Padgavankar and Gupta [14] present a detailed study of the issues associated with massive data storage and provide ways to address them. Jayasree [9] contrasts big data technologies, including MapReduce and Hadoop, to classic data mining approaches. Zulkernine et al [22] provide a conceptual architecture for cloud-based analytics as a service (CLaaS).

III. THE FRAMEWORK OF CLOUD COMPUTING

3.1 What is Mean by Cloud Computing

Many scholars describe cloud computing differently. NIST provides a widely accepted definition. According to the NIST definition [11], "Cloud computing provides on-demand access to configurable computing resources, such as networks, servers, storage, applications, and services, that can be quickly provisioned and released with minimal management effort or service provider interaction. The cloud model consists of 5 basic features, 5 service types, and 4 deployment models".

3.2 Features of Cloud Computing

Cloud computing possesses five fundamental traits. The features include on-demand capabilities, extensive network access, resource pooling, quick flexibility, and measurable service. These are the traits that set it apart from previous computer paradigms.

On-Demand Self-Service: Users can provision computing resources, such as processing power, storage, and networking, as needed without requiring human intervention from the service provider. This capability enables users to scale resources up or down dynamically based on demand, leading to greater flexibility and efficiency.

Broad Network Access: Cloud services are accessible over the internet from a variety of devices, including desktop computers, laptops, smartphones, and tablets. Users can access cloud resources and applications from anywhere with an internet connection, enabling remote access and collaboration.

Resource Pooling: Cloud providers aggregate computing resources, such as servers, storage, and networking infrastructure, into a shared pool that multiple users can access as needed. This pooling of resources allows for greater utilization efficiency and economies of scale, reducing costs for both providers and consumers.

Rapid Elasticity or Quick Flexibility: Cloud resources can be rapidly provisioned and released to quickly scale up or down in response to changing demand. This elasticity enables organizations to adapt to fluctuations in workload without overprovisioning or underutilizing resources, thereby optimizing resource usage and cost-effectiveness.

Measured Service or Pay-Per-Use: Cloud computing resources are typically metered and billed based on actual usage, allowing users to pay only for the resources they consume. This pay-per-use model offers cost transparency and flexibility, as users can scale resources according to their needs and avoid upfront capital expenditures on hardware and infrastructure.

These five characteristics collectively define the essence of cloud computing and distinguish it from traditional computing paradigms. By offering on-demand capabilities, extensive network access, resource pooling, quick flexibility, and measurable service, cloud computing delivers scalability, accessibility, efficiency, and cost-effectiveness, empowering organizations to innovate and compete in today's digital economy.

3.3 Models to obtain Cloud Deployment

Cloud deployment models are classified into four types: private, public, community, and hybrid cloud. Private clouds are the most secure approach to use cloud computing. The cloud infrastructure is dedicated to a single company with various users (e.g., business units). It can be owned, managed, and operated by the organization, a third party, or a mix of both. It can be located on or off premises. The community cloud is only used by a group of users from companies with similar interests. It may be owned, managed, and administered by community organizations, a third party, or a mix of both, and can be located on or off premises. The public cloud is designed for open usage. It might be owned, managed, and operated by a corporation, an academic, a government body, or a mix of these. It exists on the cloud provider's premises. A hybrid cloud is a combination of two or more independent cloud infrastructures (private, communal, or public) that remain separate entities but are linked together by standardized or proprietary technology that allows data and application mobility [2, 3].

3.4 Models for Delivering Cloud Computing Services

Cloud-based services are widely classified into four models: data as a service (DaaS), software as a service (SaaS), platform as a service (PaaS), and infrastructure as a service (IaaS). Software as a Service (SaaS) is a paradigm that gives users access to pre-built developer apps that are operating in the cloud. Access is provided via cloud clients, and cloud users do not maintain the infrastructure in which the program lives, removing the requirement for the cloud user to install and execute the application on their own machines. Platform as a Service (PaaS) is a paradigm that provides users with development environments in which they may create and execute in-house applications. The services might comprise an operating system, a programming language execution environment, databases, and web servers. Infrastructure as a Service (IaaS) is a concept that offers the customer with virtual infrastructure, such as servers and data storage space. Virtualization is critical in this mode because it enables IaaS-cloud providers to deliver resources on-demand from vast pools placed in data centers. Data as a Service (DaaS) is a paradigm that makes data easily available via a cloud-based platform. Simply said, DaaS is a new method of accessing business-critical data from an existing data center. Figure 1 depicts the standard cloud computing architecture.

Cloud Computing

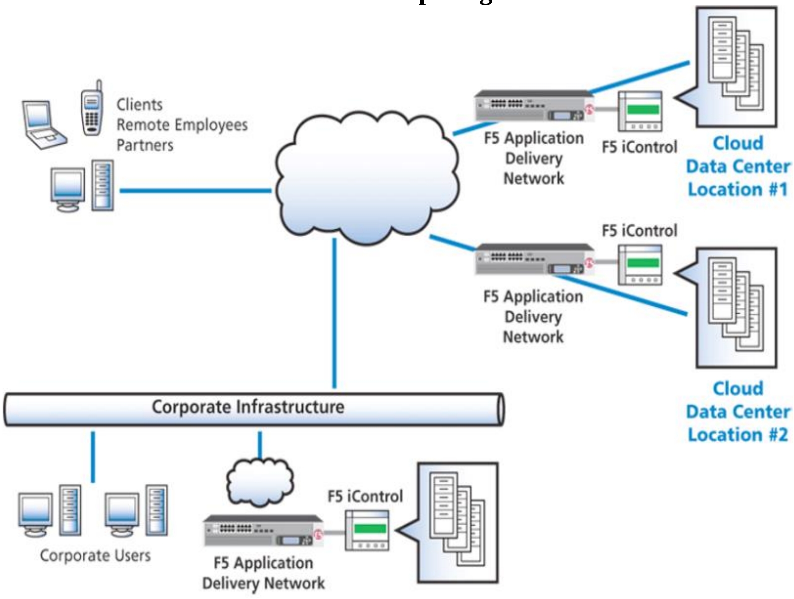

Fig. 1 An Illustration of a Cloud Computing Architecture

3.5 Benefits of Cloud Computing

Cloud computing offers a wide range of benefits that have transformed the way organizations operate and deliver IT services. Some of the key benefits include:

Cost Efficiency: Cloud computing eliminates the need for organizations to invest in and maintain on-premises hardware and infrastructure. Instead, they can access computing resources on a pay-per-use basis, reducing capital expenditure and shifting IT spending to operational expenses. This model allows organizations to scale resources up or down as needed, optimizing costs based on actual usage.

Scalability: Cloud computing provides virtually limitless scalability, enabling organizations to quickly scale computing resources up or down in response to changing demand. This elasticity allows businesses to accommodate fluctuations in workload without the need for upfront investments in additional hardware or infrastructure.

Flexibility and Agility: Cloud computing offers unparalleled flexibility, allowing organizations to deploy, manage, and customize applications and services with ease. Cloud platforms provide a wide range of tools and services, enabling developers to rapidly build, test, and deploy applications, accelerating time-to-market and enabling faster innovation.

Accessibility and Remote Collaboration: Cloud computing enables users to access applications and data from anywhere with an internet connection, facilitating remote work and collaboration. This accessibility allows employees to work from any location and on any device, increasing productivity and flexibility while reducing dependency on physical office space.
Reliability and Disaster Recovery: Cloud providers typically offer high levels of reliability and uptime through redundant infrastructure and data centers. This reliability ensures that applications and services remain available even in the event of hardware failures or outages. Additionally, cloud platforms often include built-in disaster recovery capabilities, enabling organizations to quickly recover data and applications in the event of a disaster or outage.
Security: Cloud providers invest heavily in security measures to protect data and infrastructure from cyber threats. These measures include encryption, access controls, identity management, and continuous monitoring. By leveraging the expertise and resources of cloud providers, organizations can enhance their security posture and better protect sensitive data.
Automatic Updates and Maintenance: Cloud providers handle the maintenance and updates of underlying infrastructure and software, reducing the burden on IT teams. This ensures that organizations have access to the latest features, security patches, and performance improvements without the need for manual intervention.
Environmental Sustainability: Cloud computing can contribute to environmental sustainability by enabling resource optimization and energy efficiency. By consolidating workloads onto shared infrastructure and leveraging virtualization techniques, cloud providers can achieve higher levels of resource utilization, reducing energy consumption and carbon emissions compared to traditional on-premises deployments.
Overall, cloud computing offers a compelling array of benefits, including cost efficiency, scalability, flexibility, accessibility, reliability, security, and environmental sustainability. These benefits have made cloud computing a cornerstone of digital transformation initiatives across industries, empowering organizations to innovate, compete, and thrive in today's rapidly evolving business landscape.

IV. ANALYTICS OF BIG DATA

4.1 Big Data Technologies
To enable big data analytics, a computer platform must match the following three requirements, known as the 3 Vs, as shown in Figure 2.
Variety: The platform supports a wide range of data and allows businesses to handle it both in its native format and with sophisticated transformation capabilities to convert it to other forms. ***Velocity:*** The platform can handle data at any velocity, including low-latency streams like sensor or stock data, as well as enormous amounts of batch data.
Volume: The platform can manage large amounts of at-rest or flowing data.

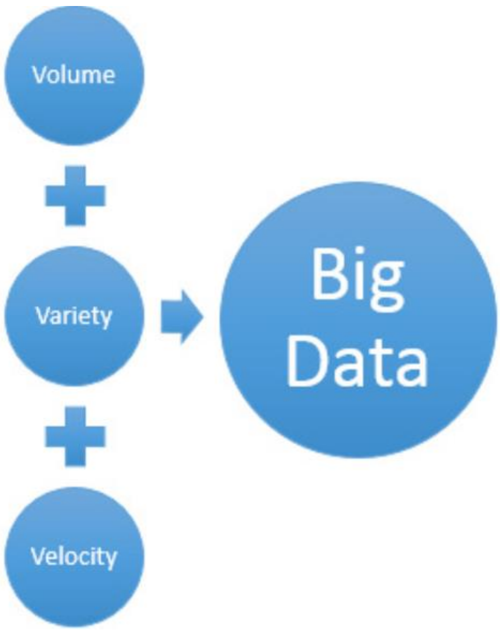

Fig. 2 An Illustration of a Cloud Computing Architecture

Traditional data mining focuses on identifying patterns in datasets, whereas big data analytics includes storing and analyzing vast amounts of data. Hadoop and MapReduce are widely used for large data analytics.

More tools and technologies are being developed for large data processing. Examples of hosted BI data warehouses include Amazon's Redshift, Google's BigQuery, IBM's Bluemix cloud platform, and Amazon's Kinesis. The future of big data will be a combination of on-premises and cloud solutions. NoSQL databases, an alternative to traditional SQL-based relational databases, are becoming increasingly popular for large data analytic applications.

V. EMPLOYING CLOUD-BASED BIG DATA ANALYTICS

Cloud-based big data analytics is a service paradigm in which components of the big data analytics process are delivered over a public or private cloud [18, 20]. It employs a variety of analytical tools and approaches to assist organizations in extracting information from large amounts of data and presenting it in an easily categorizable and accessible manner via a web browser. Cloud-based data analytics products and services are often priced on a subscription or utility basis (pay-per-use). This service paradigm is known as Cloud Analytics as a Service (CLAaaS) [22]. In this paradigm, analytics are easily available via a cloud computing platform. A cloud-based data analytics solution will allow firms to automate activities at any time and from any location. Cloud-based analytics products and services include hosted data warehouses, SaaS BI, and cloud-based social media analytics. Data saved in a cloud database can aid firms in their decision-making processes. Cloud-based big data provides analysts with not just additional data to work with, but also the processing ability to handle massive quantities of entries with many properties. This has the potential to improve predictability. The combination of big data and cloud computing also enables analysts to investigate fresh behavioral data, such as websites visited or location on a regular basis.

5.1 Major Benefits for Business Organisations

Business organizations stand to gain numerous benefits from adopting cloud computing. Some of the major advantages include:

Cost Savings: Cloud computing eliminates the need for upfront investments in hardware, infrastructure, and IT personnel. Instead, businesses can access computing resources on a pay-per-use basis, reducing capital expenditure and shifting IT spending to operational expenses. This model allows organizations to scale resources up or down as needed, optimizing costs based on actual usage.

Scalability and Flexibility: Cloud computing provides virtually limitless scalability, allowing businesses to quickly scale computing resources up or down in response to changing demand. This elasticity enables organizations to accommodate fluctuations in workload without the need for upfront investments in additional hardware or infrastructure. Additionally, cloud platforms offer flexibility in deploying and managing applications and services, enabling faster innovation and time-to-market.

Improved Collaboration and Productivity: Cloud computing enables remote access to applications and data from anywhere with an internet connection, facilitating remote work and collaboration. This accessibility allows employees to work from any location and on any device, increasing productivity and flexibility while reducing dependency on physical office space. Cloud-based collaboration tools also streamline communication and collaboration among teams, enhancing efficiency and teamwork.

Reliability and Business Continuity: Cloud providers typically offer high levels of reliability and uptime through redundant infrastructure and data centers. This reliability ensures that applications and services remain available even in the event of hardware failures or outages. Additionally, cloud platforms often include built-in disaster recovery capabilities, enabling organizations to quickly recover data and applications in the event of a disaster or outage, thus ensuring business continuity.

Enhanced Security: Cloud providers invest heavily in security measures to protect data and infrastructure from cyber threats. These measures include encryption, access controls, identity management, and continuous monitoring. By leveraging the expertise and resources of cloud providers, organizations can enhance their security posture and better protect sensitive data.

Automatic Updates and Maintenance: Cloud providers handle the maintenance and updates of underlying infrastructure and software, reducing the burden on IT teams. This ensures that organizations have access to the latest features, security patches, and performance improvements without the need for manual intervention. Automatic updates also help organizations stay compliant with regulatory requirements and industry standards.

Business Agility and Innovation: Cloud computing enables rapid experimentation, innovation, and time-to-market for new products and services. By providing access to a wide range of tools and services, cloud platforms empower developers to build, test, and deploy applications more quickly and efficiently. This agility allows businesses to respond faster to market changes, customer demands, and competitive pressures, driving innovation and growth.

Environmental Sustainability: Cloud computing can contribute to environmental sustainability by enabling resource optimization and energy efficiency. By consolidating workloads onto shared infrastructure and leveraging virtualization techniques, cloud providers can achieve higher levels of resource utilization, reducing energy consumption and carbon emissions compared to traditional on-premises deployments.
Overall, cloud computing offers a multitude of benefits for business organizations, including cost savings, scalability, flexibility, improved collaboration and productivity, reliability, security, agility, innovation, and environmental sustainability. These benefits have made cloud computing a strategic imperative for businesses looking to drive digital transformation, compete more effectively, and achieve long-term success in today's dynamic business landscape.

5.2 Major Benefits for Business Organisations

Certainly, the decision to store enterprise data outside the corporate firewall, particularly in cloud environments, raises several serious concerns and challenges. Some of the most common challenges are discussed below:

5.2.1 **Data Security and Privacy**

Data Breaches: Storing data outside the corporate firewall increases the risk of data breaches and unauthorized access. Cloud providers typically implement robust security measures, but breaches can still occur due to various factors such as misconfigurations, insider threats, or sophisticated cyber attacks.

Compliance: Organizations may face challenges in ensuring compliance with data protection regulations and industry standards when data is stored in the cloud. Different jurisdictions may have varying data privacy laws, and organizations must ensure that their data handling practices comply with relevant regulations such as GDPR, HIPAA, or CCPA.

5.2.2 Data Loss and Availability

Service Outages: Cloud service providers may experience downtime or service disruptions, impacting the availability of enterprise data. While reputable cloud providers strive to maintain high levels of uptime, occasional outages can still occur due to factors such as infrastructure failures, maintenance activities, or natural disasters.

Data Loss: Despite the redundancy and backup mechanisms implemented by cloud providers, data loss can still occur due to hardware failures, software bugs, or other unforeseen events. Organizations must have robust data backup and recovery strategies in place to mitigate the risk of data loss in the cloud.

5.2.3 Vendor Lock-in

Dependency on Cloud Providers: Organizations that heavily rely on cloud services may face challenges related to vendor lock-in, where switching to alternative providers or migrating data back on-premises becomes difficult and costly. This dependency can limit flexibility and negotiation power, as well as increase the long-term costs associated with cloud usage.

Interoperability: Integrating cloud services from different providers or transitioning between providers can be complex due to differences in APIs, data formats, and service offerings. Lack of interoperability standards and vendor-specific technologies can hinder portability and interoperability between cloud environments.

5.2.4 Data Governance and Control

Loss of Control: Storing data in the cloud may lead to a perceived loss of control over data governance and management processes. Organizations may have concerns about visibility, transparency, and control over how their data is stored, processed, and accessed by cloud providers and third-party services.

Data Sovereignty: Organizations operating in multiple jurisdictions may face challenges related to data sovereignty, where regulations require data to be stored or processed within specific geographical boundaries. Cloud providers may have data centers located in different regions, raising concerns about compliance with local data protection laws and regulations.

5.2.5 Network Connectivity and Performance

Network Latency: Accessing data stored in the cloud over the internet may result in network latency and performance issues, particularly for latency-sensitive applications or large volumes of data. Organizations operating in remote or underserved areas may face additional challenges related to network connectivity and bandwidth limitations.

Dependency on Internet Connectivity: Cloud-based applications and services rely on stable internet connectivity for access and operation. Any disruptions or fluctuations in internet connectivity can impact the availability and performance of cloud services, affecting business operations and productivity.

VI. CONCLUSIONS

Businesses employ data analytics to optimize earnings and make informed decisions. Cloud computing and big data technologies are expected to significantly impact the corporate sector. Cloud computing has become a reality for the whole technology sector, rather than simply a buzzword. Cloud computing's big data technology enables organizations to make informed decisions by forecasting future trends and behaviors. Businesses may store data remotely and access it from anywhere, anytime. Cloud-based data analytics reduces the need for enterprises to construct their own infrastructure.

Cloud computing offers cost-effectiveness, resource pooling, on-demand services, quick flexibility, and simplicity of management, enabling organizations to remain competitive in addition to data analytics. Although there are benefits, there are certain obstacles and cons, especially with privacy and security concerns. To invest in cloud-based big data analytics, organizations must first understand the full scope of the project. Investing in cloud analytics may be rewarding, but requires good strategy to cover all aspects of analytics.

REFERENCES

[1]. Agarwal, D., Das, S. and Abbadi, A. (2011). Big Data and Cloud Computing: Current State and Future Opportunities. ACM 978-1-4503-0528-0/11/0003. Retrieved from: http://www.edbt.org/Proceedings/2011-Uppsala/papers/edbt/a50-agrawal.

[2]. Aydin, N. (2015). Cloud Computing for E-Commerce, Journal of Mobile Computing and Application. Vloume 2, Issue, 1, pp 27-31.

[3]. Barthelus, L. (2010). adopting cloud computing within the healthcare industry: opportunity or risk? Online Journal of Applied Knowledge Management, Volume 4, Issue 1.

[4]. Dan, S and Roger, C. (2010). Privacy and consumer risks in cloud computing, Computer Law and Security Review, Vol 26, pp: 391-397.

[5]. Fan, J., Han, F. & Liu, H., 2013. Challenges of Big Data Analysis. ResearchGate, 1(1), pp.1-38.

[6]. Ilieva G, Yankova T. and Klisarova, S. (2015). Big Data Based System Model of Electronic Commerce, Trakia Journal of Sciences, Vol. 13, Suppl. 1, pp 407-413.

[7]. Fan, W. and Bifet, A. “Mining Big Data, Current Status and Forecast to the Future”. SIGKDD, Explorations, 14(2), pp:1-5.

[8]. Jaseena K.U and David J.M. (2014). ISSUES, CHALLENGES, AND SOLUTIONS: BIG DATA MINING, Natarajan Meghanathan et al. (Eds) : NeTCoM, CSIT, GRAPH-HOC, SPTM – 2014, pp. 131–140.

[9]. Jayasree, M. (2013). Data Mining: Exploring Big Data Using Hadoop and MapReduce, International Journal of Engineering Science Research, IJESR, Volume 04, Issue 01.

[10]. Khan, I., Naqvi, S.K. Alam, M. Rizvi, S.N.A. (2015). Data model for Big Data in cloud environment. Computing for Sustainable Global Development (INDIACom), 2015 2nd International Conference. pp. 582 -585.

[11]. Mell, P, and Grance, T. (2011). The NIST Definition of Cloud Computing, National Institute of Standards and Technology Special Publication.

[12]. Neaga, I. and Hao, Y. (2014). A Holistic Analysis of Cloud Based Big Data Mining. International Journal of Knowledge, Innovation and Entrepreneurship. Volume 2 No. 2, 2014, pp. 56—64. Retrieved from: http://ijkie.org/IJKIE_December2014_IRINA&HAO.pdf

[13]. NESSI. (2012). Big Data: A New World of Opportunities. Retrieved from: http://www.nessi-europe.com/Files/Private/NESSI_WhitePaper_BigData.pdf

[14]. Padgavankar M.H and Gupta, S.R. (2014). Big Data Storage and Challenges, International Journal of Computer Science and Information Technologies, Vol. 5 (2) , 2218-2223.

[15]. Sharma, P, Navdeti, C. (2014). “Securing Big Data Hadoop: A Review of Security Issues, Threats and Solution”, IJCSIT, Vol 5(2), 2126-2131.

[16]. Puneet Singh Duggal, Sanchita Paul, “Big Data Analysis : Challenges and Solutions”, International Conference on Cloud, Big Data and Trust 2013, Nov 13-15, RGPV

[17]. Rachana, R.C and Guruprasad, H.S (2014). Emergency Security Issues and Challenges in Cloud Computing, International Journal of Engineering Science and Innovative Technology (IJESIT) , Vol 2, Issue 2, pp 485-490.

[18]. Pareek,A and Gupta, M. (2012). Review of Data Mining Techniques in Cloud Computing Database, International Journal of Advanced Computer Research Volume-02, Number-02, Issue-01.

[19]. Shukla, D.P., Patel, S.B, and Sen, A.K. (20114). A Literature Review in Health Informatics Using Data Mining Techniques, International Journal of Software and Hardware Research in Engineering, Volume 2, Issue 2, pp.123-129.

[20]. Talia, D. (2013). Clouds for Scalable Big Data Analytics. Published by IEEE Computer Society.

[21]. Yadav, C. Wang, S. and Kumar M. (2013). “Algorithm and Approaches to handle large Data- A Survey”, IJCSN, Vol 2, Issue 3, ISSN: 2277-5420.

[22]. Zulkernine, F. Bauer, M. and Aboulnaga,A.(2013). Towards Cloud-based Analytics-as-a-Service (CLAaaS) for Big Data Analytics in the Cloud, 2013 IEEE International Congress on Big Data.