Challenges of Security Concerns Associated with Big Data in Cloud Computing

[1]Anusha Prem I, [2]Hajeera Begum R
[1]Assistant professor, [2]Student,
[1,2]Department of Computer Science, St. Joseph's College of Arts and Science for Women, India

Abstract- Currently, various organizations, industries, business relishes the value of big data applications. This paper emphasis the security issues in cloud computing that are associated with data. IT societies were progressed due to the involvement of cloud computing in a prominent manner. Cloud computing with big data application is more beneficial and plays a vital role in all support services to end users. Moreover, Cloud computing is an open environment, offers all services over the internet, but still ambiguity persists regarding the insecurity and unauthorization to end users. In this regard, the cloud computing can be enhanced with proper techniques and methodologies to secure data with its effective stores and to provide confidentiality. This paper strives to offer leading research opportunities and challenges in security services in adoption of cloud computing with big data applications and also with solutions, which can help the stakeholders to better use of cloud computing service.

Keywords: Security issues, Networking, Big Data, Cloud Computing

I. INTRODUCTION

Big Data and cloud are currently important in the territory of computer science. These two topics have drawn the attention of researchers as well as designers and businessmen.

The aim of the researchers is to identify the best methodology of extracting information through enormous information of various types of data. Organizations need to improve their frameworks for various applications. This enthusiasm from various groups will decide the development in research in enormous information applications. Big data is specific and it differs depending on the activity domain:

1. Healthcare
2. Science
3. Education
4. Governance
5. Sporting events
6. Banking processes.

Almost all of these fields need some data to be public, while the rest must be privacy-protected to support customer confidence and companies during the use of big data. In another words, big data is changing the very way of business, from social insurance to retail and farming. The rate at which information is collected on every conceivable activity means that there are increasing opportunities to fine-tune procedures and operations to squeeze out every last exclusion of efficiency. Of course, in business once an item has been developed or manufactured it should be sold and distributed [1]. The amount of client information, including you and me, effectively assembled by big retailers lets them know who will need to purchase what, where and when. Also, the cloud changes the ways data is stored and gives new views on information management, while applications are becoming easier to be used by anyone.

II. CLOUD COMPUTING

Earlier, IT sector faced difficulty in allocation of resources in multi user environments. Therefore, many stakeholders were not in position to finish their software projects in time due to resources unavailability. As a result, cloud computing came into existence and provided a solution on persisting resource sharing problems. A cloud offers a resources as pay-per- use technique, which can be reconfigured instantly according to use requirements [5]. Due to this, cloud computing was deployed in infrastructural environments, share and bind more nodes together at a time. Cloud is collection of computing devices and nodes that shares services publically on the World Wide Web. It is implemented and managed via cloud applications by gathering all computing resources all together. It also maintains user's history information in their internet usages, processes and offer accurate information, so that it offers intellectual support service to IT users [3]. Cloud users can be easily connected via web services in different networking environments. Progression of cloud computing increased due to its benefits such as global access, standard platforms, massive scalability, dynamic infrastructure, Administrative functions and majorly its minimal usage costs. IT sector benefited from cloud computing with its high level services with reduction in IT resource deployments. By deploying cloud computing, third party users become the owner of the resources needs to deliver services and manage their infrastructures according to their own requirements. Furthermore, cloud computing offers many features such as easy data maintenance, minimum effort for deploying software applications, offers sophisticated IT services and Minimal IT infrastructure [4].
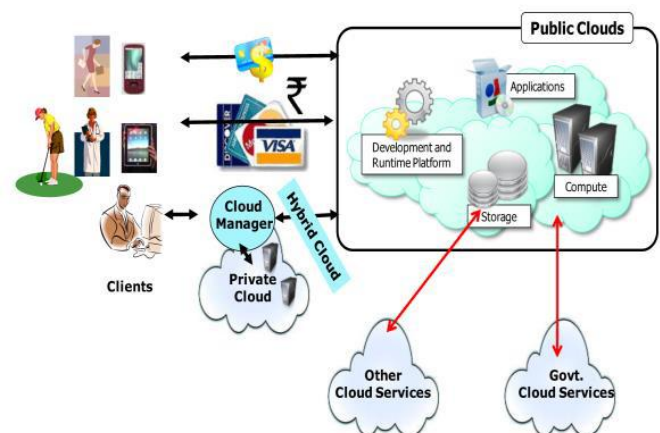

Figure 1: Cloud Computing services in IT Structure

IT society authorized the cloud computing with ownership of establishing IT infrastructures, due to its risk free implementation [3] [4] . Henceforth, various stakeholders benefited since they paid only for what they utilized. Seamless IT resources are offered to global users as a support service via internet, helped many major and minor enterprises grow. Cloud computing offers IT solutions with proper access rights in minimal cost with increased flexibility.

III. COMPELLING COMBINATION OF BIGDATA AND CLOUD COMPUTING

Nowadays, more IT societies and IT technologies supports in storing and processing huge amount of data effectively. Cloud computing enabled IT society to handle large amount of information instantly, with the integration of Big Data. Handling Social application information's such as Facebook, twitter, whatsApp, Visa Customer transactions, Credit Card processing in bank applications are some of the real time examples illustrating the usage of Big Data in Cloud Computing [2].

Figure 2: Big Data Support Services in Cloud

The term "Big Data" originated from Internet with Mining Transactions to manage very large distributed information. Big data offers cloud computing with increased storage capability to handle massive volumes of information, which can be integrations of audio, video, pictures and plain text. In real time, Big Data has three properties, namingly V3 indicating Volume, Variety and Velocity and two more Dimensions, namingly Variability and Complexity. Volume indicates the size of Data like storing Socio - Media Data, Live Streaming of Data, etc. variety indicates the type of data or format of data like audio, video, emails, transactions, etc. Velocity It indicates the processing technique of data to meet the user's on demand requirements. Variability indicates the inconsistency of data flows in processing. Complexity indicates the type of processing the data like mapping, linking, cleansing and transformations.

IV. BIG DATA IN CLOUD COMPUTING

Benefits-

In cloud computing, many software packages are offers big data as an service, where end user can integrate structured and unstructured data as per his/her requirements from the entirety of its organization's storage capacity, thus the end user with analysis capability to handle the threats [8]. Thisis a greater advantage of big data for an individual to detect the sensitive data and transform them according to their necessity, also provides the individual to have their control on their data. Many

IT Industries and organizations benefited from the assistance of big data with its Capacity, Scalability and velocity of cloud storage. Furthermore, new business opportunities were created in IT Society as the end users has the possibility of visualizing their data in various dimensions. Big data already provides data analytics, allows end user to personalize their information. Additionally, it also provides predictive analytics, allows end user to predict their data for information processing for their business applications.

A. Security Challenges-

Even though stakeholders have the entire control of data, cloud computing technologies doesn't have control on stored data. Consider an example, end user doesn't have the knowledge on where the data is stored and processed in the cloud. Since data will be migrating over the internet, stakeholders will not have control over the data. Furthermore, internet is open environment it is susceptible to more security threats. If security threat occurs in big data, it leads to information leakage and loss of reputation to that IT industry. Due to loss of control in cloud computing and big data, there is need to concentrate on security of data or it will lead to legal complications to the organization on data. In cloud computing, the biggest challenge is to detect and prevent malicious hackers and to overcome security threats. Most of the big data techniques helps in detecting threats of fraudulent activities occurs over internet. Another biggest challenge is information privacy issues in IT societies. Due to increase of Big Data Utilization in cloud computing, more IT industries and IT Business suffers from privacy issues. Instead of assuring security policies and privacy techniques on IT industries, it is better to revise the policies and regulations of cloud computing, so that there will be a balance between data security and privacy polices provide to global user [6] [7].

B. Security Issues-

IT Vendors offers storage, computation, application hosting services with backed performance and charge on pay-per-use techniques via cloud service providers (CSP). But also, cloud computing undergoes several security issues and defined below:

1. Data occurrences of end user data has to be integrated, linked, mapped and transformed from other occurrences of another end user data.
2. Data migration over the internet exposes sensitive information to malicious intrudes leads to loss of privacy, legal liability and reputational loss to the IT organizations.
3. Internet is open to global users, an intruder can exploit a software bug to sniff trustworthy information, to misuse IT resources, or generates a pathway leading to advanced threats and attacks.
4. Outsourcing information must be limited to a authorized users with access rights, so to avoid the abuse of malicious access.
5. Lack of Data replication and recovery technique in cloud computing leads to loss of data.
6. Cloud services must be made accountable mandatory as an application requirement.

C. Research Opportunities-

Across the globe, two IT innovations are ruling the IT Industries, namingly Cloud computing and Big Data Analytics. Both innovations lead IT services with enhanced business productivity, providing valuable awareness generating new innovations wit International conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) reduced costs [9]. The evolution of cloud computing lead IT enterprises with more agile and effective cloud environments offering increased IT support services. But still, IT organizations started questioning on how to store big data and what techniques to be employed to address data analytics to meet business requirements. As a consequence, IT organizations started constructing cloud computing integrated with Big Data [10]. The problem still persists with big data environments, which need networking servers to support their large volumes, high velocity and numerous Big data formats. Nowadays, clouds are designed with large networking servers, increased storage capacity and reduced infrastructure enabling modifications as needed by customers. As a results, it was proved that cloud computing is a big solution to integrated big data technologies with advanced analytic tools, which enhances business significance.

Table1: Research Challenges And Oppurtunities

Research Area	Opportunities	Level
Data Access	The ownership of data and its storage area is still unknown.	Content Generation and Authentication
Data Management	Selection of cloud providers with evaluation criteria's Risk pertains in storage of data, maintenance and its analysis by multiple data providers with an authorization.	Content Storage, Data Accomplishment
Data Integration	Various formats of data persists in cloud, how to integrate those formats. What costs to be Imposed	Data Analysis and Combinations
Data Streaming	Real time data streaming requires large infrastructures to manage. Non –stop data streaming may be handled in-house.	Data Communication and Data loss management
Distributed data	Migrating data from and to the cloud leads to more cost. The amount of data and cost for migrating has to be computed for cost analysis.	Distributed Cloud Computing Environments Distributed Data Analysis and Management
Security	Clouds can be designed with built-in security infrastructure. Cloud platforms can also be designed to support security infrastructure to provide greater assurance.	Security Issues and Challenges
Distributed Nodes	Identified difficulty in finding the precise location of Node computations. Difficulty in building security to the computational node.	Distributed Networking
Internode Communication	TCP/IP employs Remote procedure calls for data transfer between nodes, so intruder can modify or tap the node to hack the data.	Communication and Networking
Data Protection	Many cloud environments doesn't have data protection mechanisms, therefore there is an chance for hacker to hack the cloud node to hack critical data.	Security Issues and Analytics
Node Authentication	Independent nodes can join the network for providing simultaneous operations. Since there is not authentication, Third party node can join the network to steal the data or disrupt the operation of users.	Node Management

Information Resources	Data arrives from diverse sources	Assurance of Availability, Integrity and Confidentiality

V. SECURITY AND RESEARCH CHALLENGES

Systematic Analysis

Cloud Computing Security is one of the most important concern to be achieved in data privacy protection. But, it's more challenging in safeguarding cloud access in distributed environment from various threats due to data distribution. Therefore, Data leak occurs due to unauthorized access in shared environment and leads to violation in use of cloud computing resources [11] [12]. Data segregation is the proposed solution to resolve this challenge, where every client's data can be segregated from other client's data. Another challenges to be resolved is data migration, where data transfer occurs between various countries and again leads to violation of data regulations. To prevent this, data anonymity is the resolution by authorizing client's data privacy and security. Nowadays, big data integrated with cloud computing is growing but along with that privacy is also a biggest research area growing with that. Still many researches going on to assure peak data privacy and data access to achieve client's objectives. Another major privacy challenge in cloud computing is the nature of information source. In Big data, data are generated from various sources in various formats leading to data to be secured and managed in caution [10]. At the same time, data should be protected from unauthorized users from changing the originality of it. And also, client's privacy should be sustained, while collecting, managing and communicating data's. These challenges are very crucial to the privacy of cloud and are mentioned as assurance of availability, confidentiality, and integrity. Big data always deals with data collection from various sources and is one of the major challenged faced in cloud computing, since it leaks client's sensitive information. This type of data migration may lead to violation of user's privacy policy. Another challenge lies in cloud computing is user's identification and authentications and need to be resolved in networking [6]. The challenges and their possible countermeasures are mentioned in the below table.

Table-2: Privacy And Security Concerns In Cloud Computing

Privacy Challenges	Security Concern	Countermeasures
Protecting cloud computing resources	Information will be exposed to the risk of authorized access in Shared Environment	Data Isolation
Data Migration	Data may transfer between different borders and face violation of policies	Data Segregation
Business Adoption of Computing Assets in IT infrastructures	Risk of losing control or governance by organizations leading to privacy assurance.	Specify resource requirements, compliance reporting

The cloud entities are majorly targeted in their functionality of data migration, data storage and their maintenance. The cloud privacy is imposed on the cloud entities by considering the mentioned challenges to address the user's objectives and authorizations. The relationship between cloud entities and cloud privacy is depicted in the below.
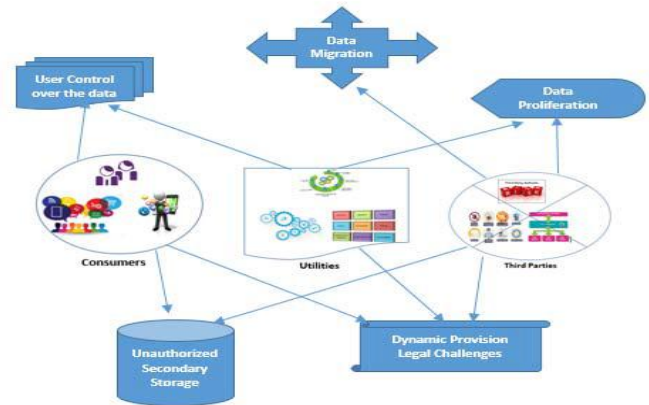

Figure 3: Relationship between Cloud Entities and Cloud Privacy Issues

CONCLUSION

Security and privacy concerns are considered as real challenges in this cloud computing and big data fields. This papers discuses on these two issues and also research challenges related to cloud computing. The cloud users are excited in utilizing the opportunities provided by big data, but still they are more curious about their privacy and security issues. More researches are still in progress in identifying security challenges and suggest resolutions for them. This paper focuses on privacy concerns identifies as an associated concern with security flaws. Security issues were discussed and few solutions were also proposed. Many IT organizations were still working on ensuring the security of outsourced data in an untrusted cloud data center and how to achieve the privacy of user's data. The cloud computing area defined in this paper is defined with the integration of Big Data as an application to cloud users provides more benefits stands as a winning condition. Moreover, Many IT infrastructures faces economic crises due to minimal business budgets, therefore there is a need to identify opportunities in minimizing functional costs, so that simultaneously risk of technology usage also reduces.

This work can be focused in two major tracks in future and are cloud computing in terms of minimal economy and real-time end user objectives. The second scope can be achieved in proper business plans and the decisions of IT Management to follow the right path. To design this Innovation Diffusion Theory (IDT) will be suitable tools that contributes to understanding the adoption process of organizations. To the overall, the privacy and security concerns defined here must be discovered more functionally through the field research that investigates the perceptions of organizations after adopting cloud computing and how they evaluate such experience.

References

[1]   D.Wang, Zh.Han, (2015), "Sublinear Algorithms for Big Data Applications", ISBN 978-3-319-20448-2 (eBook).

[2]   M. Armburst et al., "Above the Clouds: A Berkeley View of Cloud Computing", Tech. report, Univ. of California, Berkeley, 2009.

[3]   A Platform Computing Whitepaper, "Enterprise Cloud Computing: Transforming IT", Platform Computing, pp6, viewed 13 March 2010.

[4]   R. Buyya, Y. Chee Shin, S. Venugopal, "Market-Oriented Cloud Computing: Vision, Hype, and Reality for Delivering IT Services as Computing Utilities", In 10th IEEE Conference on High Performance Computing and Communications, 2008.

[5]   MafruzZamanAshrafi, David Taniar, and Kate A. Smith, "A Data Mining Architecture for Distributed Environments" pages 27-34, Springer-VerlagLondon,UK 2007.

[6]   Xiao, Z., & Xiao, Y. "Security and Privacy in Cloud Computing" IEEE Communications Surveys & Tutorials, Vol. 15(2), pp. 843-859, 2013.

[7]   Neela, T. J., & Saravanan, N. "Privacy Preserving Approaches in Cloud: a Survey" Indian Journal of Science and Technology, Vol. 6(5), pp. 4531-4535, 2013.

[8]   Ashktorab, V., & Taghizadeh, S. R "Security Threats and Countermeasures in Cloud Computing" International Journal of Application or Innovation in Engineering & Management (IJAIEM), Vol.1 (2), pp. 234-245, 2012.

[9]   Hao, Chen, and Ying Qiao. "Research of Cloud Computing based on the Hadoop platform." Chengdu, pp. 181 – 184, 21- 23 China: Oct 2011.

[10]   A, Katal, Wazid M, and Goudar R.H. "Big data: Issues, challenges, tools and Good practices." pp. 404 – 409, 8-10, Noida: Aug. 2013.

[11]   Lu, Huang, Ting-tin Hu, and Hai-shan Chen. "Research on Hadoop Cloud Computing Model and its Applications.". Hangzhou pp. 59 – 63, 21-24, china: Oct. 2012.

[12]   P.R , Anisha, Kishor Kumar Reddy C, Srinivasulu Reddy K, and Surender Reddy S. "Third Party Data Protection Applied To Cloud and Xacml Implementation in the Hadoop Environment With Sprawl." pp 39-46, Jul – Aug. 2012.