IMPACT OF BIG DATA AND CLOUD ON INDIAN INDUSTRY

Pratima P. Vishwakarma

Master of Computer Application, IMCOST, Mumbai University, Mumbai, India.

Abstract-
Cloud computing plays a vital role in protecting data, applications and the related infrastructure with the help of policies, technologies, controls, and big data tools. In this paper, we discuss impact of cloud Computing Big data Environment on Indian industry. The main focus is on security, Accessibility, Technical, Business, Attacks issues in cloud computing that are associated with big data. Big data applications are a great benefit to organizations, business, company and many large scale and small scale industries. We discuss various possible solutions for the issues in cloud computing security and Hadoop .Cloud computing security is developing at a rapid pace which includes computer, network, Information security and data privacy .Moreover, cloud computing and its applications, advantages are like to represent the most promising new frontiers in science. In this paper, we come up with some approaches in providing security. A system that can scale to handle a large number of sites and also be able to process large and massive amounts of data .In this paper, we characterize the problems in addition, we describe our experience and lessons learnt in construction of a cloud computing Platform.

Keywords-
Cloud Computing, Big Data.

I. INTRODUCTION

Cloud "Cloud" computing – a relatively recent term, builds on decades of research in virtualization, distributed computing, utility computing, and more recently networking, web and software services. Cloud Computing is a more recent and latest version of Grid Computing. It implies a service oriented architecture, reduced information technology overhead for the end-user, great flexibility, reduced total cost of ownership, on demand services and many other things. loud Computing, the long-held dream of computing as a utility, has the potential to transform a large part of the Indian it industry, making software even more attractive as a service and shaping the way for designing IT hardware. Cloud Computing can be considered now as a Pay and Use service.

Cloud computing contains software applications and data storage services that are delivered in real time over a network, usually the Internet. These services include basically anything that you can do on a PC: e-mail, data storage, and communications and productivity applications. The benefits of cloud computing may contain lower costs, greater mobility and enhanced collaboration. Additionally, by leveraging cloud services small business can reduce the need for in house exchange servers and IT staff, thereby potentially reducing operation expenditures.

In order to analyze complex data and to identify patterns it is very important to secure store, Manage and share large amounts of complex data. Cloud comes with an explicit security challenge i.e. Data owner might not have any control of where the data is placed. The reason behind this issue is that if one wants to get the benefits of cloud computing, he/she must also utilize the allocation of resources and also the scheduling given by the controls. Hence it's required to protect the data in the midst of untrustworthy processes .Since cloud involve extensive complexity, we believe that rather than providing a holistic solution to securing the cloud, it would be ideal to make noteworthy enhancements in securing the cloud that will ultimately provide us with a secure cloud.

II. CLOUD COMPUTING

Cloud computing is a new purpose Internet based technology through which information is stored in

servers and provided as a service and on demand to clients.

Cloud computing it's a technology which depends on sharing of computing resources than having Local servers or personal devices to handle the applications. In Cloud Computing, the word "Cloud" means "The Internet", so Cloud Computing means a type of computing in which services are delivered through the Internet. The goal of Cloud Computing is to increasing Computing power to execute millions of instructions per second.

Cloud Computing uses large group of networks servers with specialized connections to distribute data processing among the Servers. Instead of installing a software suite for each computer, this technology requires to install single software in each computer that allows users to log into a Web-based service and which also hosts all the programs required by the user. There's a significant workload shift in a cloud computing system. Local computers no longer have to take the entire burden when it comes to running applications. Cloud computing technology is used to minimize the usage cost of computing resources. The cloud network, consisting of networks of computers, handles the load instead the cost of software and hardware on the user end decreases. The only thing that must be done at the users end is to run the cloud interface software to connect to the cloud.

A. Features of Cloud Computing:

- Cloud provides resources which are on demand as there is isolation so no need to actual sharing.
- It is heterogeneous in nature.
- It adds the virtualization to the data and hardware resources too.
- It deals with end user security.

B. Characteristics of Cloud Computing

1. High Scalability:

Cloud environments enable servicing of business requirements for large audiences, through high scalability.

2. Cost Savings:

A company reduces their capital expenditures and use operational expenditures for increasing their computing capability.

3. Multisharing:

The cloud working in distributed and shared mode. Multiple users and applications can work more efficiently with cost reductions by sharing common infrastructure.

4. On-Demand Self-Service:

The on-demand and self-service of cloud computing mean that a consumer can use cloud services as needed without any human interaction with the cloud provider.

C. TYPES OF CLOUD COMPUTING ENVIRONMENTS:

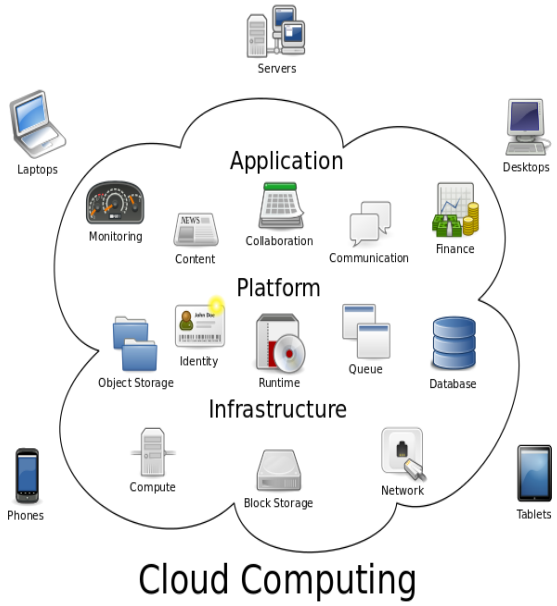
- **PUBLIC CLOUDS**: - This cloud infrastructure is available to the public on a commercial basis by a cloud service provider.
- **PRIVATE CLOUDS:-**This type of cloud infrastructure has been deployed, and is maintained and operated for a specific organization.
- **HYBRID CLOUDS: -**It's a combination of both private (internal) and public (external) cloud computing environments.
- **COMMUNITY CLOUD: -** In type of cloud infrastructure is shared among a number of organizations with similar interests and requirements.

D. ISSUES AND CHALLENGES

Cloud computing comes with numerous issues because it encompasses many technologies including networks, databases, operating systems, virtualization , resource scheduling, transaction management, load balancing, concurrency control and memory management. Hence, security issues of these systems, technology are relevant to cloud computing. For example, it is very important for the network which interconnects to the systems in a cloud to be secure.

Also, virtualization paradigm in cloud computing results in several security Concerns.
For example, mapping of the virtual machines to the physical machines has to be performed very securely.

Data security not only involves the encryption of the data , but also ensures that applicable policies are enforced for data sharing . In addition, resource allocation and memory management algorithms also have to be secure .The big data issues are most acutely felt in certain industries ,such as telecoms , web marketing and advertising , retail and financial services , and certain government activities . The data explosion is going to make life difficult in many industries , and the company will gain significant advantage which is capable to adapt well and gain the ability to analyze such data explosions over those other companies . Finally, in cloud data mining method can be used in the malware detection.

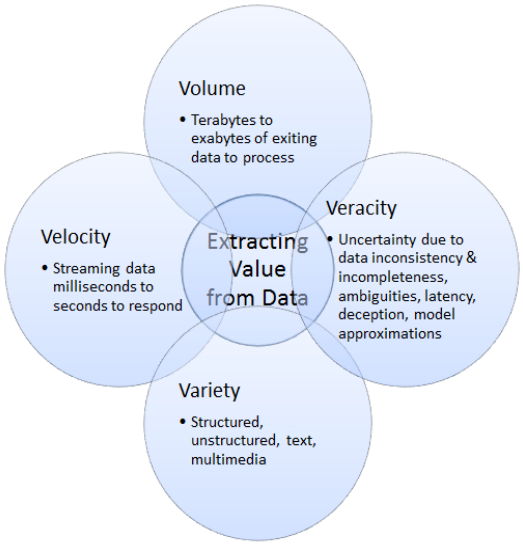


III. BIG DATA

Big Data is word used to describe massive volumes of structured and unstructured data that are so large that it is very difficult to process this data using traditional databases and software technologies. The term "Big Data" is companies who had to query loosely structured very large distributed data. Large amounts of data, a no of sources, high speed production, but also high speed processing the amount of data that is generated and collected in each second grows exponentially. Big Data contains huge volume, high velocity and extensible variety of data. The data in it will be of three types.

- **Structured data**: Relational data.
- **Semi Structured data**: XML data.
- **Unstructured data:** Word, PDF, Text, Media Logs.

The management of Big Data, the intelligent use of large, heterogeneous data sets, is becoming increasingly important for competition. It's affecting all sectors industry and academia but also the public sector while the economy is exploring Big Data as a new gold mine, politicians are fighting over the problem of data capitalism .Whereas science tackles questions of cross-disciplinary benefits, as well as on the challenges and the likely consequences for technology, innovation and society. The main key characteristics that define big data are volume, velocity and variety as described in figure.



IV. BIG DATA IN THE CLOUD

In order to overcome the volume issue, big data requires technologies that store vast amount of data in a scalable fashion and provide distributed approaches to querying or finding that data. The cloud computing model is a perfect match for big data since it provides unlimited resources on-demand. With cloud, no longer need to

purchase hardware or high cost software. On the other hand generally data intensive or data driven applications generate and process massive data sets usually stored in the cloud. These applications have large data processing requirements and are engineered with custom algorithms to run on scalable infrastructure. The characteristics of the cloud computing it's provided big data acquisition, and software data processing strategies. Gartner has estimated that 50% of data will be stored on the cloud by 2016 (Schouten, 2012). The availability of cloud based solutions has dramatically lowered the cost of storage, amplified by the use of commodity hardware even on a "pay as-you-go" basis that is directed to effectively and timely processing large data sets. The big data could be delivered "as -a -service". Google Big Query is an example of applying big data in a cloud based environment. The convergence of two key technological areas is having far reaching implications.

A. ISSUES AND CHALLENGES

The issues in Big Data are some of the conceptual points that should be understood by the organization to implement the technology effectively. Big data Issues are need not be demented with problems but they are important to know and crucial to handle. Issues related to the Characteristics Data Volume As data volume increases, the value of different data records will decrease in proportion to age, type, richness, and quantity among other factors. The social networking sites existing are themselves producing data in order of terabytes every day and this amount of data is definitely difficult to be handled using the existing traditional systems.

The challenges in Big Data are usually the real implementation hurdles which require immediate attention. Any implementation without handling these challenges may lead to the failure of the technology implementation and some unpleasant results. Privacy and Security It is the most important challenges with big data which is sensitive and includes conceptual and technical as well as legal significance. The personal information (e.g. in database of a merchant or social networking website) of a person when combined with external large data sets, leads to the inference of new facts about that person and it's possible that these kinds of facts about the person are secretive and the person might not want the data owner to know or any person to know about them.

V. BIG DATA AND THE CLOUD

Independently, both big data analytics and cloud computing are invaluable tools. With big data paralytics, companies are able to gain insight from unstructured and semi structured data - information that would otherwise remain inaccessible .A company has discovered ways to improve internal operations, audience outreach and a host of other critical processes.

The cloud's advantages are massive. With the cloud, a firm can upgrade its data storage and data warehousing capabilities in a scalable, cost-efficient manner. Computing possibilities are significantly improved with the use of cloud services.

Consequently, the news source noted that cloud-based big data services can have tremendous benefits for companies of all kinds. "Cloud-based solutions purchased from big data providers allow companies to access processing power and computer applications installed on remote networks," Kane asserted. "End-users can then access cloud-based applications via web browsers or mobile apps and, since the data is stored remotely, businesses can manage and monetize their data resources more effectively."

VI. IMPACT OF CLOUD IN INDIA

The Cloud has become the way that people live now. It has become the part and way of life. If an example of Face book is taken about 500 million people today keep their photo albums in the Cloud. Over 87,000 companies in the world use the Cloud. In India Cloud computing grow at 36 percent this year.

Some of them are tiny little start-ups; others are some of the largest organizations in the world. Dell, for instance has deployed our collaboration tool across 100,000 employees in recent days. Accordingly, it could be said that Cloud computing means, the technology being used at the other end is invisible and irrelevant as far as the customer's concerned.

The Cloud is not about technology, it is the abstraction of technology for delivering pure services. The Cloud has finally led to the civilization of services because, in the

case of Cloud computing, what you pay for is what you get.

VII. ACKNOWLEDGMENT

The authors would like to thank the editor, mysterious reviewers for their valuable suggestions that appreciably improved the quality of this paper, especially to Project Guide Dept. of MCA for their constant encouragement and providing us the overwhelming support and guidance to write this paper. Finally also thankful our IMCOST Educational Trustees, Dear Principal Dr. J.N. Shah , Teaching Faculties, and Non- Teaching faculties of Department of MCA IMCOST Institute , Thane, India.

VIII. CONCLUSION

Cloud environment is widely used in industry and research aspects. Today's increase the impact of big data lots of issues is affecting the Indian industry .Attacker is easily get any data from the cloud therefore security is an important aspect for organizations running on these cloud environments. Using proposed approaches, cloud environments can be secured for complex business operations.

REFERENCES

[1].http://www.theatlantic.com/sponsored/ibm-cloud-rescue/archive/2012/09/changing-the-world-big-data-and-the-cloud/262065/

[2].http://www.newindianexpress.com/cities/chennai/Cloud-computing-in-India-to-grow-at-36-per-cent-this-year/2013/09/25/article1801561.ece

[3]. "Introduction to Big Data: Infrastructure and Networking", White Paper Juniper Networks 2012.

[4]. Hao, Chen, and Ying Qiao. "Research of Cloud Computing based on the Hadoop platform." Chengdu, China: 2011.

[5]. http://airccse.org/journal/nsa/6314nsa04.

[6]. http://irdindia.in/journal_ijacte/pdf/vol3_iss5/5.

[7].http://hpc-asia.com/4-ways-big-data-will-impact-india-2015/

[8].http://www.slideshare.net/HarshMishra3/harsh-big-data-seminar-report

Author: Pratima Vishwakarma is currently studying in Mumbai University at IMCOST thane since 2012, currently pursuing MCA with excellent academic. She is having interest in studying new technology.

