ISSUES AND CHALLENGES IN CLOUD COMPUTING AND BIG DATA

[1]*MR.PARAG D PATEL*

[1,] Department of Computer Engineering, C. U. Shah Technical Institute of Diploma Studies, C. U. Shah University, Surendranagar, India.

patel_paragkumar@gtu.edu.in

ABSTRACT:

This paper introduces a detailed analysis of between big data and cloud computing security issues and challenges focusing on the cloud computing types and the service delivery types. However, big data entails a huge commitment of hardware and processing resources, making adoption costs of big data technology prohibitive to small and medium sized businesses. It advantages includes scalability, resilience, flexibility, efficiency and outsourcing non-core activities. The use of big data could provide sufficient benefit to a small to medium sized company to the extent that the business would be willing to commit resources to implement big data technology in-house. However, the level of benefit is difficult to determine without some experience.

KEY WORDS: Cloud Computing, Scalability, Infrastructure, IT.

1. INTRODUCTION

The concept of cloud computing is founded in early 1960s when John McCarthy an American computer scientist predicted the eventual convergence of computing infrastructure, which would allow a great degree of versatility in the distribution of IT resources in order to meet fluctuating and unpredictable business demands. It is an extremely successful paradigm of service oriented computing which has revolutionized the way computing infrastructure is abstracted and used. The popularity and versatility of the internet enabled companies like Amazon to modernized their data centers and adopt the cloud architecture which allowed them to outsource computing power to external customers. This service was later launched in 2006, after which numerous companies took the initiative to create cloud-based infrastructure. Cloud Computing is a term used to describe a new class of network based computing that takes place over the Internet or a model that relies on a large, centralized data center to store and process a great wealth of information [1].It can be defined as a collection of integrated and networked hardware, software and Internet infrastructure called a platform i.e. using the Internet for communication and transporting hardware, software and networking services to clients. This platform hides the complexity and details of the underlying infrastructure from users and applications by providing very simple graphical interface or API (Applications Programming Interface) and also provides on-demand services that are always on, anywhere, anytime and anyplace. Cloud computing is a way to increase the capacity or add capabilities dynamically without investing in new infrastructure, training new personnel, or licensing new software. But as more and more information are placed in the cloud, concerns begin to grow about the security of the cloud environment. Security issues in cloud computing has played a major role in slowing down its acceptance.

2. CLOUD COMPUTING SERVICE DELIVERY MODELS

Cloud deployment solutions provide services that businesses would otherwise not be able to afford under the traditional hardware and software acquisition method. Cloud computing revolutionizes the way information is handled, the typical deployment models for cloud computing includes: infrastructure as a service (IaaS), platform as a service (PaaS), software as a service (SaaS) and hardware as a service (HaaS).

i. Infrastructure as a Service (IaaS):-

In the infrastructure as a Service (IaaS) model, consumers are given full freedom to manage their data on the server. Here the service provider is only responsible for raw storage, computing power, networks, firewalls, and load balancers and this is often manifested as a virtual machine. A client business pays on a per-use basis whenever the equipment is used to support computing operations such as: storage, hardware, servers, and networking equipment [2]. Infrastructure as a service is a cloud computing model that has received most attention from the market, with an expectation of 25% of enterprises planning to adopt a service provider for IaaS [3].
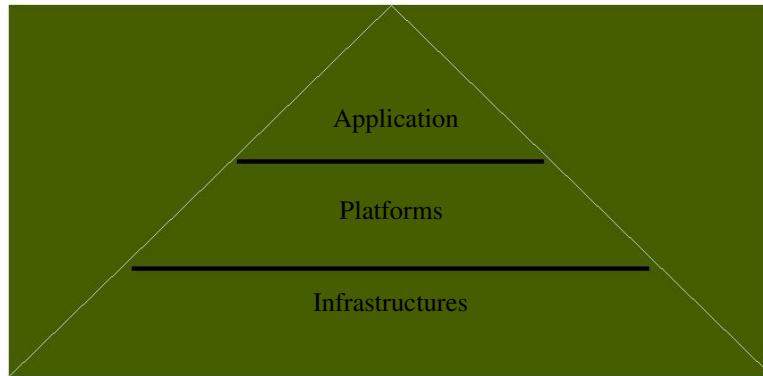
Fig.1. Cloud Computing Service delivery models

ii.   Platform as a service (PaaS):-

Platform as a Service is a level above Infrastructure as a service (IaaS). In the PaaS model, consumers are provided with an operating system, programming language execution environment, database, and web server. They are not concern with the cost and management in the hardware and software layers. PaaS is the use of cloud computing to provide platforms for the development and use of custom applications [4]. The PaaS solutions include application design and development tools, application testing, versioning, integration, deployment and hosting, state management, and other related development tools [5]. Businesses attain cost savings using PaaS through standardization and high utilization of the cloud-based platform across a number of applications [6]. As related to big data, PaaS provides companies a platform for developing and using custom applications needed to analyze large quantity of unstructured data at a low cost and low risk in a secure environment. Therefore maintaining the integrity of applications and enforcing accurate authentication checks during the transfer of data across the entire networking channels is fundamental.

iii.   Software as a Service:-

Software as a service (SaaS) is the level above Platform as a service (PaaS). In this model, consumers are given access only to the application software, which can be run remotely from the data centers of the cloud service provider. The provider is responsible for the maintenance and support of the infrastructure and operating platforms i.e. it provides businesses with applications that are stored and run on virtual servers in the cloud [8], Since cloud service providers specialize in one area, they can provide reliable service at a fraction of the cost. The business is not charged for hardware, only for the bandwidth for the time and number of users necessary. The main advantage of SaaS is that this solution allows businesses to shift the risks associated with software acquisition while moving IT from being reactive to proactive [9]. Benefits of using SaaS include: easier software administration, automatic updates and patch management, software compatibility across the business, easier collaboration, and global accessibility [10] Software as a Service provides companies analyzing big data proven software solutions for data analysis. The difference between SaaS and PaaS in this case is that SaaS does not provide a customized solution whereas PaaS will allow the company to develop a solution tailored to the company's needs (could we change this to: whereas PaaS allows clients to develop a customized solutions based on their specific needs?) . SaaS is becoming an increasingly prevalent delivery model as underlying technologies that support web services and service-oriented architecture (SOA) mature and new developmental approaches become popular. SaaS is also often associated with a pay-as-you-go subscription licensing model. Meanwhile, broadband service has become increasingly available to support user access from more areas around the world. SaaS is most often implemented to provide business software functionality to enterprise customers at a low cost while allowing those customers to obtain the same benefits of commercially licensed, internally operated software without the associated enabled multi-tenancy (specifically designed to support many concurrent users. at once). SaaS applications are accessed using web browsers over the Internet therefore web browser security is vitally important. Information security officers will need to consider various methods of securing SaaS applications. Web Services (WS) security, Extendable Markup Language (XML) encryption, Secure Socket Layer (SSL) and available options which are used in enforcing data protection transmitted over the Internet.

iv.   Hardware as a service (HaaS):-

In addition cloud computing allows computing to be treated as a commodity. In the past a company in need of computing power and storage is required to purchase its own processors and servers and maintain them. Idle capacity and economic waste is recorded whenever those IT resources are not in use. However, with the advent of cloud computing, a company can purchase its exact computing needs. If maximum computing power is only

needed occasionally the company does not have to pay for those IT resources to just stay idle. The outsourcing inherent in cloud computing reduces the cost of maintaining IT and that reduction in cost is the bottom line benefit. That benefit is not limited to companies needing business applications. Cloud computing has a variety of potential applications in other fields that can potentially be explored.

3. BIG DATA AND THE CLOUD

Big Data is a data analysis methodology enabled by recent advances in technologies and architecture which support high-velocity data capture, storage, and analysis. However, big data entails a huge commitment of hardware and processing resources, making adoption costs of big data technology prohibitive to small and medium sized businesses. Cloud computing offers the promise of big data implementation to small and medium sized businesses. The term big data is derived from the fact that the datasets involved are so large that typical database systems are not able to store and analyze the datasets the characteristics of big data present data storage and data analysis challenges to businesses. Analyzing big data is done using a programming paradigm called Map Reduce. The Map Reduce paradigm requires that huge amounts of data be analyzed. The mapping is done concurrently by each separate NAS device; the mapping requires parallel processing. The parallel processing needs of Map Reduce are costly, and require the configuration noted previously for storage. The processing needs can be met by cloud-service providers. Data storage using cloud computing is a viable option for small to medium sized businesses considering the use of Big Data analytic techniques. Cloud computing is on-demand network access to computing resources which are often provided by an outside entity and require little management effort by the business

4. TYPES OF CLOUDES

The Cloud Computing model has three types of clouds model which are – the public cloud, the private cloud, and the hybrid cloud.
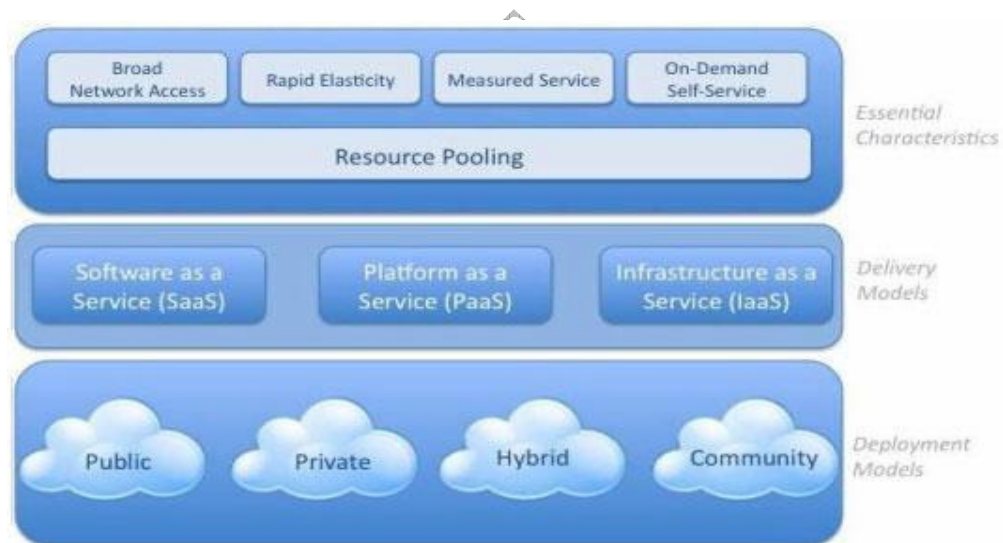

Fig. 2. NIST Visual Model of Cloud Computing Definition

i. Public cloud:-
A public cloud is the pay- as-you-go services available to the general public. In this configuration, a business does not own the core technology resources and services but outsource these to service providers. Public cloud is also considered to be an external cloud. Public cloud describes cloud computing in the traditional mainstream sense, whereby resources are dynamically provisioned on a fine-grained, self-service basis over the Internet, via web applications/web services, from an off-site third-party provider who shares resources and bills on a fine-grained utility computing basis. Public clouds are less secure than the other cloud models because it places an additional burden of ensuring all applications and data accessed on the public cloud are not subjected to malicious attacks.

ii. Private cloud:-
A private cloud is internal data center of a business that is not available to the general public but uses cloud structure. In this configuration, resources and services are owned by the business, with the services accessible within the business through the intranet and since the technology is owned and operated by the business, this

type of cloud is more expensive than a public cloud A private cloud is an internal cloud residing inside the company's firewall and managed by the company.

iii. Hybrid cloud:-

Hybrid cloud is a combination of both public and private cloud, when a company uses a hybrid cloud; it uses a public cloud for some tasks and a private cloud for other tasks. In this model, a company uses the public cloud to expedite extra tasks that cannot be easily run in the company's data center or on its private cloud [1]. The private cloud portion of the hybrid cloud is accessed by company employees, both in the company and on the go, and is maintained by the internal technology group. The private cloud part of the hybrid cloud is also accessed by the company employees but is maintained by external service providers. Every Portion of hybrid cloud are connect with each other.

iv. Community cloud :-

Community cloud is a private cloud that is shared by several customers with similar security concerns and the same data and applications sensitivity.

5. CLOUD COMPUTING CHALLENGES

Cloud computing is related with numerous challenges and the major challenges that prevent Cloud Computing from being adopted are as follows:

i. Security: -

Security issue plays the most important role in hindering Cloud computing acceptance. Security issues such as data loss, phishing, botnet poses serious threats to organization's data and software. For example, hackers can use Cloud to organize botnet as Cloud often provides more reliable infrastructure services at a relatively cheaper price for them to start an attack.

ii. Costing Model: -

Cloud consumers must consider the tradeoffs amongst computation, communication, and integration. While migrating to the Cloud can significantly reduce the infrastructure cost, it does raise the cost of data communication, i.e. the cost of transferring an organization's data to and from the public and community Cloud and the cost per unit of computing resource used is likely to be higher. This problem is particularly prominent if the consumer uses the hybrid cloud deployment model where the organization's data is distributed amongst a number of public/private (in-house IT infrastructure)/community clouds.

iii. Charging Model:-

The elastic resource pool has made the cost analysis a lot more complicated than regular data centers, which often calculates their cost based on consumptions of static computing. Moreover, an instantiated virtual machine has become the unit of cost analysis rather than the underlying physical server.

iv. Service Level Agreement (SLA):-

Since cloud consumers do not have control over the underlying computing resources, they do need to ensure the quality, availability, reliability, and performance of these resources when consumers have migrated their core business functions onto their entrusted cloud. It is therefore vital for consumers to obtain guarantees from providers on service delivery. Typically, these are provided through Service Level Agreements (SLAs) negotiated between the providers and consumers. The very first issue is the definition of SLA specifications in such a way that has an appropriate level of granularity, namely the tradeoffs between expressiveness and complicatedness, so that they can cover most of the consumer expectations and is relatively simple to be weighted, verified, evaluated, and enforced by the resource allocation mechanism on the cloud.

6. BENEFITS OF CLOUD COMPUTING

There are lots of benefits in using Cloud computing to render or access computing resources. Presently a lot of people use

Cloud computing without even knowing what it means. For example, Gmail, Yahoo mail, YouTube, and Skype users…are all in the Cloud. Increasingly companies and organizations are becoming aware of the huge benefits that Cloud computing provides. Some of these benefits include:

i. Flexibility and storage: -

With Cloud computing Files are stored in the "Cloud". This allows for development in the organization because workers no longer have to worry about the storage of documents. Also, workers can access office files from wherever and whenever. Workers can also work together virtually even when they are not at the same place at

the same time. Various documents can be viewed simultaneously provided Internet connection is available.

ii.   Time saving:-

Alongside easy collaboration, Cloud computing also aids the easy access to information. Easy access in this context could be seen in how fast it is to access Yahoo mail,Gmail, mailboxes in general. It is fast and easy in contrast to the time it would take to download and install software.

iii.   Reduced Cost:-

Cloud computing puts a stop to the illegal reproduction and distribution of software. Some software on the Cloud is free. For example, most SaaS solutions have a pay-as-you-go pricing model instead of a large up-front investment. Such pricing models allow end users to pay only for what they use thus freeing up resources such as time and money for other more important (core) business activities. Cloud computing is therefore cheaper and less labor intensive for companies. There is no need to buy and install expensive software. There is no need to acquire, track and manage software licenses.

7.   CONCLUSION

Despite the benefits enumerated, it is surprising that not many companies and organizations are rushing to leverage the advantages of Cloud computing, especially in developing countries because the benefits of cloud computing are tempered by two major concerns – security and loss of control. Although Big data and Cloud computing is a new phenomenon which is set to revolutionize caution must be exercised in the way we use the Internet. Although it will be extremely difficult to provide a complete solution to securing the Cloud in relative terms thereby completely erasing the anxiety and disquiet which the idea of outsourcing key and crucial business operations, and technology may create. Security measures such as data encryption is advised and should be employed by end-users to reduce the fear and risk of data invasion. If the fear of Cloud computing are addressed, Cloud computing will revolutionize the world of information technology, even in developing countries Cloud computing promises real benefits to companies seeking a competitive edge in today's economy.

REFERENCES

[1]   Armbrust, M., Fox, A., Griffith, R., Joseph, A. D., Katz, R., Konwinski, A., Lee, G…Zaharia, M. (2010, April). A view of cloud computing. Communications of the ACM, 53(4), 50-58. DOI: 10.1145/1721654.1721672.

[2]   Rouse, M. (2010b, August). Infrastructure as a Service. Retrieved from http://searchcloudcomputing.techtarget.com/definition /Infrastructure-as-a -Service-IaaS

[3]   Cisco. (2009). Infrastructure as a Service: Accelerating time to profitable new revenue streams. Retrieved from http://www.cisco.com/en/US/solutions/collateral/ns341/ns991/ns995/IaaS_BDM_WP.pdf

[4]   Salesforce.com. (2012). The end of software: Building and running applications in the cloud. Retrieved from http://www.salesforce.com/paas/

[5]   Géczy, P., Izumi, N., &Hasida, K. (2012). Cloudsourcing: Managing cloud adoption. Global Journal of Business Research, 6(2), 57-70.

[6]   Oracle. (2012). Oracle platform as a service. Retrieved from http://www.oracle.com/ us/technologies/cloud/oracle-platform-as-a-service-408171.html

[7]   Jackson, K. L. (2012). Platform-as-a-service: The game changer. Retrieved from http://www.forbes.com/sites/kevinjackson/2012/01/25/platform-as-a-service-the-gamechanger/

[8]   Cole, B. (2012). Looking at business size, budget when choosing between SaaS and hosted ERP. E-guide: Evaluating SaaS vs. on premise for ERP systems. Retrieved from http://docs.media.bitpipe.com/io_10x/io_104515/item_548729/SAP_sManERP_IO%23104515_EGuide_061212.pdf

[9]   Carraro, G., & Chong, F. (2006, October). Software as a service: An enterprise perspective. Retrieved from http://msdn.microsoft.com/en-us/library/aa905332.aspx #enterprisetw.topic3

[10]   Rouse, M. (2010a, August). Software as a service. Retrieved from .techtarget.com/definition/Software-as-a-Service.