IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 7, Issue 3, March 2018

Survey of Challenge Big Data on Cloud Computing and Meta Cloud Data Storage Architecture Efficiency Performances

R.Edwina[1], S.Kamatchi[2], M.Kavitha[3], T.Sathiyapriya[4]

II MCA, Anjalai Ammal Magalingam Engineering College, Kovilvenni, Tamilnadu[1,2,3,4]

Abstract: Implementing cloud computing empowers several paths for Web-based service contributions to meet diverse needs. However, the data security and privacy has become a significant issue that restricts many cloud applications. Many researchers have been trying to protect big data in cloud computing environment. Usually security mechanisms using encryption are neither efficient nor matched to the task of shielding big data in the Cloud. In this paper, we first discuss about challenge and potential solutions for protecting big data in cloud computing. Second, we propose Meta Cloud Data Storage Architecture efficiency performances.

Keywords: Cloud Computing, Big Data, MetaCloudDataStorage Architecture.

I.INTRODUCTION

Cloud computing can defined as five attributes such as Massive Scalability, Multi-tenancy (Shared Resources), Elasticity, Pay as You go and Self-Provisioning of resources. Cloud computing enables user to access the remote servers hosted on the internet to store and process the data. Cloud-computing provider offer their "services" according to special models, of which the three standard models per NIST are Infrastructure as a Service (IaaS), Platform as a Service (PaaS), and Software as a Service (SaaS) and different exploitation models are classified into Private, Public, and Hybrid. Due to the high accessibility of cloud to all end users, cloud computing faces more security challenges. These challenges are classified into two broad categories as security issues faced by cloud providers and security issues faced by Customers.

II.LITERATURE REVIEW

Cloud computing patterns have been universally for some time now, but the term established “widespread” one-time in October 2007 when IBM and Google revealed a corporation in that sphere. This was surveyed by IBM’s declaration of the “Blue Cloud” determination. Later then, everybody is conversation about “Cloud Computing”. Of course, there also is the unavoidable Wikipedia entry. It is thinkable that August 24, 2006 will go down as the centennial of Cloud Computing, as it was on this day that Amazon made the exam form of its Flexible Devious Cloud (EC2) public [Business Week 2006]. This offer, providing elastic IT incomes (computing capacity), marks a final landmark in lively profitable relations amongst IT operators and breadwinners. The period first developed general in 2007, to which the first admission in the English Wikipedia from arch 3, 2007 attests, which, again meaningfully, controlled a situation to usefulness calculating. Today, Cloud Totalling generates over 10.3 truckload competitions on Google. The prospect of Cloud Computing produced from humble organization services such as storing and control resources to include requests. However, this future that indications such as submission provision providing and Software as a Service would also henceforth be involved above the title of Cloud Computing.

III.BIG DATA

Big data challenge include capturing data, data storage, data analysis, search, sharing, transfer, visualization, querying, updating and information privacy. There are five scope to big data known as Volume, Variety, Velocity and the recently added Veracity and Value. Big data used in many application health care , education, natural resources, social networking and so on.

BIG DATA OPPORTUNITIES:

Recently, several US government agencies, such as the National Institutes of Health (NIH) and the National Science Foundation (NSF), ascertain that the utilities of Big Data to data-intensive decision-making have profound influences in their future developments. Consequently, they are trying to developing Big Data technologies and techniques to

facilitate their missions after US government passed a large-scale Big Data initiative. This proposal is very helpful for building new capability for exploit enlightening information and make possible decision-makers.
From the Networking Information Technology Research and Development (NITRD) program which is recently recognized by President's Council of Advisors on Science and Technology (PCAST), we know that the bridges between Big Data and knowledge hidden in it are highly crucial in all areas of national priority. This initiative will also lay the groundwork for complementary Big Data activities, such as Big Data infrastructure projects, platforms development, and techniques in settling complex, data-driven problems in sciences and engineering. Finally, they will be put into practice and benefit society. According to the report from McKinsey institute, the effective use of Big Data has the underlying benefits to transform economies, and delivering a new wave of productive growth. Taking advantages of valuable knowledge beyond Big Data will become the basic competition for today's enterprises and will create new competitors who are able to attract employees that have the critical skills on Big Data. Researchers, policy and decision makers have to recognize the potential of harnessing Big Data to uncover the next wave of growth in their fields. There are many advantages in business section that can be obtained through harnessing Big Data, including increasing operational efficiency, informing strategic direction, developing better customer service, identifying and developing new products and services, identifying new customers and markets, etc. The vertical axis denotes the percentages that how many enterprises think Big Data can help them with respect to specific purposes[1].

SECURITY CHALLENGES ASSOCIATED WITH BIG DATA IN CLOUD COMPUTING:
In order to protected big data, technique such as logging, encryption, and honey pot detection must be necessary. In many organizations, the deployment of big data security framework is very attractive and useful. Big data analytics can be used to detect and prevent the malicious intruders and advanced threats[2]. Big data security in the cloud computing is essential due to the following issues such as: 1) To defend and avert huge size of classified business, government, or regulatory data from spiteful prowler and advanced threats, 2)Lack of awareness and principles about how cloud service providers firmly maintaining the huge disk space and erase existing big data, 3) Lack of standards about auditing and reporting of big data in public cloud, 4)Users who does not even work for the organization (malicious intruders), but may have full control and visibility into history of organization data (big data)[5]. Many researchers are rising big data protection architecture and frameworks to protect huge data in cloud. For example preventing and detecting intrusion, Securing web application, Protecting confidential data in the cloud[7], Protecting GPS data from data mining based attacks and Securing the bank details in the cloud[9]. We predictable Meta Cloud Data Storage Architecture for defending Big Data in Cloud Computing Environment.

The contests of safekeeping in cloud computing surroundings can be considered into network level, user authentication level, data level, and generic issues[11].

- Network level: The experiments that can be considered under a network level deal with network rules and network security, such as distributed nodes, distributed data, Internodes communication.
- Authentication level: The experiments that can be considered under user authentication level deals with encryption/decryption techniques, authentication methods such as managerial rights for nodes, verification of applications and nodes, and logging.
- Data level: The experiments that can be characterized under data level deals with data integrity and availability such as data protection and distributed data.
- Generic types: The experiments that can be characterized under general level are traditional security tools, and use of different technologies.

IV.PROPOSED FRAMEWORK

MetaCloudDataStorage security architecture is proposed in this division to defend big data against gatecrasher with security and efficient performance. In this manner association data is stored in multiple cloud data midpoints based on the reputation and scope. Data classification is classified into three levels such as Sensitive, Critical and Normal. Each Characterized data are invented to be stored in dissimilar data center. Proposed MetaCloudDataStorage boundary can competently redirect the user request to the suitable datacenter in cloud providing by different vendors.

Although, the proposed framework will allocate all data parts in dissimilar storage service suppliers, and each supplier holds some of the data parts. In order to provide high obtain ability and robustness, the proposed framework will supply multiple copies of same data on different cloud storage providers. Though big data is divided and kept in dissimilar data center, the administrator of the whole system will keep the storage index data for each data parts. When there is a delinquent in some data parts on the cloud storage, propose framework can find another copy of the data parts according to their storage index data. Figure 1 shows how the end user will access the applications and facts in dispersed cloud.

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 7, Issue 3, March 2018

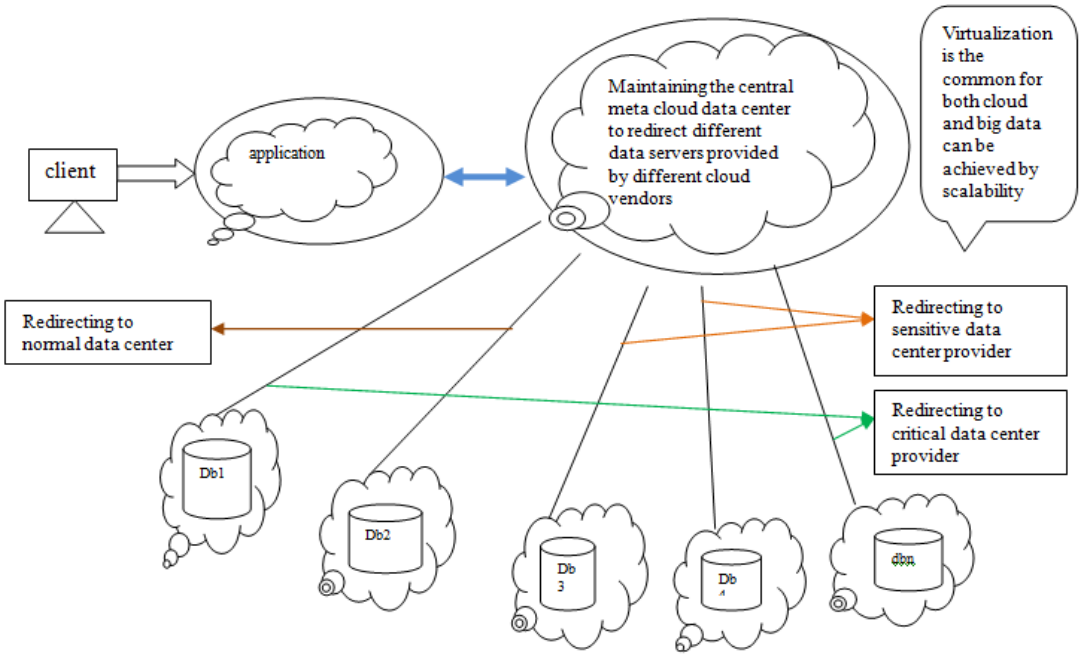

Fig.1.End User Accessing Applications and Data in Distributed Cloud

V.SECURITY FOR CLOUD DATA AND APPLICATIONS

The security architecture for the MetaCloudDataStorage in Cloud is shown in the Figure2.

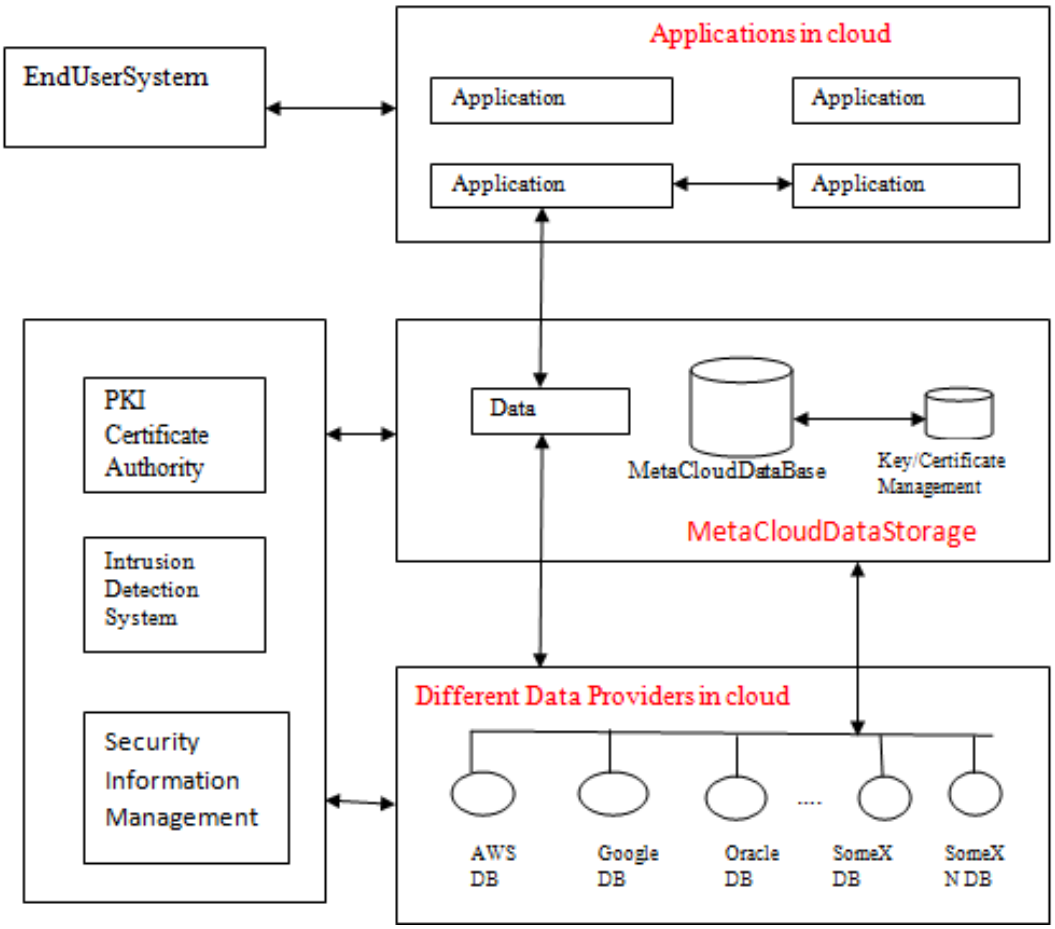

Fig.2. Security Architecture for MetaCloudDataStorage in Cloud

DOI 10.17148/IJARCCE.2018.7398

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 7, Issue 3, March 2018

VI.DATA ANALYTICS

Map Reduce Algorithm for dispensation log files in disseminated cloud data center Map Reduce is a programming model or framework that procedure tasks in similar across a vast size of schemes. It contains two functions such as Map and Reduce. Map function separations the huge size of input data into <key, value> pairs. Middle<key, value> pairs will be created bases on combining several input key value combines from the Map phase. Finally, Reduce takes the in-between key value pairs and produces the output <key, value> pairs that can be easily understood by the end user. In this proposed construction, Map Reduce framework is used to find the number of users who were logged in to the cloud data center. Proposed Map Reduce pseudo code can professionally process the huge size of log file in which it contains users who were logged in with date and the log in time length. As shown in the figure, the first process is map phase in which each date that embodies the key is consigned a value of one originally. While reduce phase, the key values are summed up to discovery out the number of users listed in.

Mapper Function
public void Map(LongWritable key, Text value, OutputCollector output, Reporter reporter)
for each key ϵ value do
Emit(term key; count 1)

Reducer Function
public void reduce(Text key, Iterator values, OutputCollector output, Reporter reporter)
sum←0
for each v ϵ value do
sum←sum + v
Emit(key, sum)

VII.CONCLUSION

In this paper we proposed MetaCloudDataStorage Structural design for defending Big Data in Cloud Computing Environment. Map Reduce framework is used to find the number of users who were charted into the cloud data center. Proposed structure shields the mapping of various data elements to each supplier using MetaCloudDataStorage interface. Though this proposed method involves high application effort, it delivers valued data for cloud computing environment that can have high impact on the next generation systems. Our future work is to encompass the proposed MetaCloudDataStorage Architecture for real time treating of streaming data.

REFERENCES

[1] C.L. Philip Chen , Chun-Yang Zhang "Data-intensive applications, challenges, techniques and technologies: A survey on Big Data" Department of Computer and Information Science, Faculty of Science and Technology, University of Macau,China
[2] Victor, N., Lopez, D., &Abawajy, J. H.. Privacy models for big data: a survey. International Journal of Big Data Intelligence, 2016 *3*(1), 61-75.
[3] Lopez, D., &Gunasekaran, M. Assessment of Vaccination Strategies Using Fuzzy Multi-criteria Decision Making. In Proceedings of the Fifth International Conference on Fuzzy and Neuro Computing (FANCCO-2015) 2015: 195-208. Springer
[4] Lopez, D., Gunasekaran, M., Murugan, B. S., Kaur, H., & Abbas, K. M. Spatial big data analytics of influenza epidemic in Vellore, India. In 2014 IEEE International Conference on Big Data (Big Data), 2014, October: 19-24. IEEE.
[5] Thilagavathi, M., Lopez, D., &Murugan, B. S. Middleware for Preserving Privacy in Big Data. Handbook of Research on Cloud Infrastructures for Big Data Analytics, IGI Global, 2014.
[6] Marchal S, Jiang X, State R, Engel T. A Big Data Architecture for Large Scale Security Monitoring. In 2014 IEEE International Congress on Big Data (BigData Congress), 2014 Jun 27: 56-63.IEEE.
[7] Hongbing C, Chunming R, Kai H, Weihong W, Yanyan L. Secure big data storage and sharing scheme for cloud tenants. Communications, China. 2015 Jun;12(6):106-15.
[8] Subashini S, Kavitha V. A metadata based storage model for Securing data in cloud environment. In CyberC 2011 Oct 10:429-434.
[9] Schmidt K, Phillips C. Programming Elastic MapReduce: Using AWS Services to Build an End-to-end Application. O'Reilly Media, Inc.; 2013 Dec 10.
[10] Siddesh GM, Hiriyannaiah S, Srinivasa KG. Driving Big Data with Hadoop Technologies. Handbook of Research on Cloud Infrastructures for Big Data Analytics. IGI Global, 2014 Mar 31:232.
[11] Pooja Bavarva1 Yash Punjabi2 Rushabh Raolji3 "Security issues associated with Big Data in cloud Computing " IJSRD || National Conference on Advances in Computer Science Engineering & Technology || May 2017 || ISSN: 2321-0613.