ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 1, January 2017

A Survey on Securing Access and Security Issues of Maintenance for Cloud Database

Swati Vithal Khidse[1], Dr. Santosh S. Lomte[2]

Assistant Professor, CSMSS Chh. Shahu College of Engineering Aurangabad, Maharashtra, India[1]

Principal, Vilasrao Deshmukh Foundation, School of Engg. & Tech., New MIDC Airport Road Latur, Maharashtra, India.[2]

ABSTRACT: Cloud computing emerges as one of the vigorous topic in information technology and widely adopted due to easy accessibility and availability of data from remote locations. In cloud computing environment clients access the data through virtual resources in cloud. Because of exponential growth of data in cloud there are various aspects that need to be addressed by cloud service provider like how to satisfy the large number of request that are made to the cloud database, how to secure the access of the customers while accessing the cloud database. Database as a service has become boon in cloud computing, it reduces the overhead of client for installing and maintaining database on their machine, it is also provided on pay per usage. This paper provides survey of security challenges, security issues of maintenance faced by the database service provider and customer, and the solutions they implemented.

KEYWORDS: Cloud computing, Confidentiality, Encryption, Database as a service, Secure access, Security.

I. INTRODUCTION

Cloud computing is emerging technology. National Institute of Standards and Technology (NIST) defines cloud computing as follows: “Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to the shared pool of configurable computing resources (e.g. networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimum management efforts or interaction with the cloud service provider.”[1]. The Cloud computing model composed of the five characteristics: On-demand self-service, broad network access, resource pooling, rapid elasticity, measured service [1]. Using cloud computing businesses, IT industry and various clients are storing and accessing their important data through internet without worrying about how the data is stored and retrieved, hardware management in the cloud. The data can be accessed from anywhere via the internet in the cloud. This cloud data storage architecture is given in Fig.1. The service and maintenance for the data is provided by the cloud database service provider following the five characteristics of the cloud computing model.

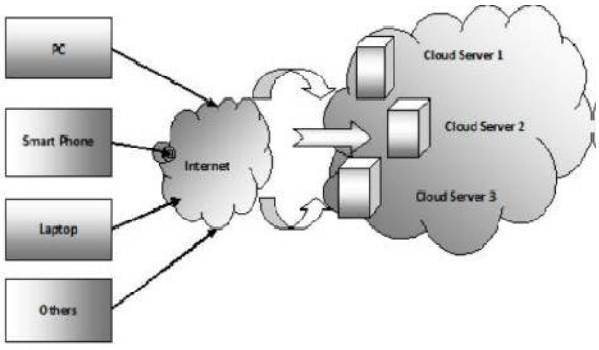

Fig.1. Cloud Data Storage Architecture

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 1, January 2017

II. RELATED WORK

The exponential growth of the Internet has resulted in an explosion of data sources, creating storage and data-usability problems. Furthermore, an increase in the variety of data types has created challenges for storing and manipulating unstructured data.

These issues have lead corporations, firms and open source communities to build new tools, called as NoSQL systems or "key-value-store" systems, which aim to offer, on an enormous scale, on demand services and simplified application development and deployment. A NoSQL database is useful for applications that deal with very large semi structured and unstructured data. NoSQL databases can be categorized into two different groups, based on elasticity level. First group having truly elastic databases (such as MongoDB) new nodes can be added to the cluster. Second groups containing rigidly defined BigTable-based NoSQL databases (such as Cassandra and HBase), when new nodes are added to the cluster significant downtime is observed.

A survey of 15 popular NoSQL databases [2] has conducted by authors, providing overview of the system, storage platform, license type, data-handling techniques, billing practices, programming language used for writing source code of NoSQLs database.

1. *HBase* –It is open source, distributed, versioned, column-oriented database.
2. *Cassandra*-Apache Cassandra database offers good scalability, fault-tolerance and high availability without compromising with the performance.
3. *MongoDB* –It is open source, schema free, full index support, data availability, document-oriented storage and scalable NoSQL database system. It also provides a complex query language and implementation of MapReduce.
4. *Pnuts* –It is a massive-scale hosted, centrally managed database system. It provides data management as a service.
5. *BigTable* -Google's BigTable maps any two random string values (i.e. row key and column key) and a time stamp (creating 3D mapping) into a corresponding arbitrary byte array. It is characterized as a light, distributed, multidimensional sorted map.

The success of the cloud database can be strictly related to terms of service availability, security, scalability, and data confidentiality. Multi-User relational Encrypted Data Base (MuteDB) [3], is an architecture for the cloud database services that guarantees data confidentiality by executing SQL encryption algorithms and data isolation by access control enforcement using encryption and key derivation techniques. It combines data encryption, key management, authentication and authorization solutions, and addresses the threat issues including risk of information leakage for cloud database services. MuteDB architecture's performance and scalability is comparable to unencrypted cloud services.

With the rapid growth of cloud computing in business sector and IT industries, data has escaped from IT department's control, moving into the wider reaches of cloud-based services, mobile devices, and social networking sites. As a result it has increased the need for reliable cloud-database services. Cassandra, HBase, and MongoDB are the most widely used cloud databases and they represent most of the NoSQL world.

Every cloud service has issues and majority of them are concerned with the security. Database as a service (DaaS) is also having several major issues and concerns [4], such as data security, trust, expectations, regulations, and performance issues. In this paper author has discussed proposed solutions for these issues, including risk management, better contractual agreements between the service provider and customer which makes it a key aspect of the service, database encryption, and authenticity techniques. The Service-Level Agreement (SLA) is the legal document between the service provider and customer. It describes how to manage risks, understand the assurance level available, and discover and deal with the insecurities

Data security is of prime concern when the data is migrated to cloud DBMS. Database encryption [5], is an approach in which the columns carrying the sensitive data are encrypted before they are stored in the cloud. It has been proposed as a mechanism for addressing data security concerns.

It is categorized into three broad approaches:

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 1, January 2017

1. *Homomorphic Encryption*-This is specially designed encryption scheme that directly allows computation over encrypted data.
2. *Client-Server based Approaches*- Homomorphic encryption is performed at the server and client, which is assumed to be trusted.
3. *Trusted Hardware based Approaches*- In this approach a secure and tamper-proof hardware is used at the server for securely decrypting the data and performing computations.

The data that is migrating to the cloud database is often more complex. So, firstly it should be preprocessed before performing any operation i.e. to clean the data. For this efficient algorithms should be in place for detecting errors. Errors in the data can be detected as a violation of constraints (data quality rules), such as functional dependencies, denial constraints, and conditional functional dependencies [6]. Authors provide incremental algorithms, detecting the violations of CFDs for vertically partitioned data and horizontally partitioned data, and shown that the algorithms are optimal.

In a cloud computing environment any number of clients can access the data from different companies, IT industry and etc. at a same time. In order to fulfill their requests, resources are shared among different clients. Intelligently managing and allocating the resources from cloud among various clients is very important for system providers in cost-effective manner satisfying the client service level agreements (SLAs). In this paper, authors [7] addressed how to intelligently manage the resources in a shared cloud database system and present a cost aware resource management system SmartSLA. Two main components of SmartSLA are: System modeling module and the resource allocation decision module. The system modeling module addresses the answer to the question "How to judge system performance accurately?" It uses machine learning techniques to learn a model (data driven approach). The model captures the relationships between the systems resources and database performance.

The resource allocation decision module, adjusts the resource allocations dynamically for achieving optimum profits from the learned model. The performance evaluation of SmartSLA is performed under TPC-W benchmark with workload from real-life systems. The results shows that SmartSLA can compute successfully predictive models under various hardware resource allocations (e.g. CPU and memory), and database specific resources (e.g. number of replicas) in the database systems. SmartSLA also provides intelligent service differentiation depending on factors like variable workloads; cost of resources, SLA levels, and delivers improved profit margins.

In cloud database users cannot control and audit their sensitive data by themselves. As a result there is a need for security and privacy for data management and query processing in the cloud. It is critical for better and broader uses of the cloud. In [8] authors mentions two important aspects of data security and privacy i.e. data confidentiality, and query access privacy for sensitive data in the cloud.

1. *Security and Privacy Threats*-
   When data oriented services are deployed in the cloud environment, the cloud provider or any unauthorized parties can monitor and control the data and activities. So there is a need for ensuring data confidentiality and access privacy in the cloud.
2. *Data Confidentiality*-
   Data confidentiality can be ensure by Encryption and querying encrypted data and trusted computing (encrypted data is stored in the cloud but decrypt and process the plaintext data in a secure trusted container in the cloud).
3. *Access Privacy*-
   When both the data and the queries are encrypted and when data is queried in cloud database, queries may reveal the partial information about data. Hence there is a need for ensuring access privacy. It is achieved with Private Information Retrieval, Oblivious RAM, covered search and index shuffling for protecting accesses to encrypted index and hybrid approaches.

In cloud computing, database server is owned by a third party. As a result of query to database, data are generated by multiple sources. The correctness of the data can be ensured from authorized transactions. So, there is a need to ensure trustworthiness of data from cloud database. In [9] authors had addressed the problem of: Ensuring trustworthiness of data from untrusted server with transactional updates that directly run on the database. The solution is implemented as a prototype system built on top of Oracle with no database internal modifications. It detects any failures on the server for

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 1, January 2017

faithfully hosting a transactional database with multiple independent users, provides indemnity for honest server and assured provenance.

The basic functionality of the Automatic Test Equipment (ATE) community is to support a multi-tiered maintenance. It is a three tiered system, which consist of: organizational maintenance (O-level), intermediate maintenance (I-level), depot maintenance (D-level) organizations. The primary goal of the ATE is to:

1. Detect and isolate fault, quickly and accurately
2. Software tools for analyzing historical data, and
3. Gather, manage, and distribute the accurate and reliable maintenance information for the failed Unit Under Test (UUT).

In [10] authors provide improvement of the ATE testing and maintenance capabilities by using the cloud computing model for building globally linked ATE maintenance system.

III. THE CLOUD PARADIGM

Once the large amount of data is stored on the cloud by the customers, now the next question is how to use it and what are the different operations that can be performed in the cloud. In cloud computing four deployment models are used:

A. *Private cloud*- In private cloud model, infrastructure is provided for exclusive use by a single organization. It is use for protecting highly sensitive data.
B. *Public cloud*- The cloud infrastructure is available for use by everyone.
C. *Community cloud*- Cloud infrastructure is exclusively used by a community of users having same requirements and concerns.
D. *Hybrid cloud*- When two or more distinct cloud infrastructures are used together a hybrid cloud is formed.

The cloud service provider offers the following resource types, i.e. types of service provider models which create a classification of cloud types: infrastructure, platform, or software.

A. *Infrastructure as a service* (IaaS) - In this environment customers are allowed to have their own virtual machines (server) ,raw (block) and file-based storage, firewalls, load balancers in the cloud. Customer can build, load, and run their virtual machines. It is ensured by the provider that machine will have necessary computing power.
B. *Platform as a service* (PaaS) - Cloud providers, provides a computing platform (application environment), it include operating system, programming language execution environment, database. The provider ensures to run and maintain the programs in a Web environment. This paradigm is also called as Web 2.0.
C. *Software as a service* (SaaS) - In this customer is provided with access to application software and databases. Cloud providers run the applications (infrastructure, platform), and provides Internet access to it.
Database as a service (DaaS) is a prime example of SaaS. Service provider picks the database management software installs it, runs, and manages it. These cloud computing service models are represented in the Fig.2.

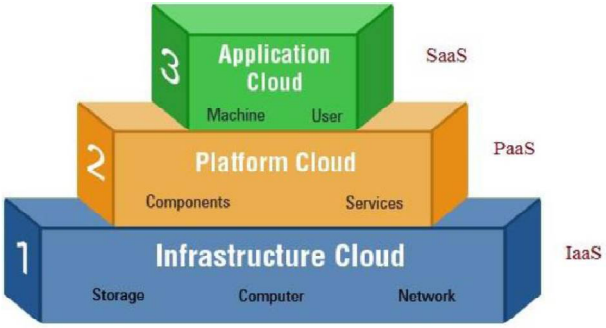

Fig.2. Cloud Computing Service Models

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 1, January 2017

IV. CONCLUSION

With the exponential growth of cloud services by business organizations, IT industry massive amount of data is migrating to databases into cloud environment. This provides several benefits to the clients like affordability, flexibility and scalability, increased efficiency through mobility. It also provides services like Infrastructure, Platform, and Software to the users and also database as a service for the cloud. But on the other hand it brings a number of security challenges for their clients like data security, data loss, data integrity, availability, access control, privileged user access, auditing and monitoring issues. For handling these various security challenges, different solutions are provided like encryption, key management, authorization, authentication (Integrity and Completeness), querying encrypted data, disaster recovery, segregation of customer data, auditing of the service provider. This research paper provides outlines of what cloud computing is, various cloud models and overview of solutions for secure access and security issues of maintenance in cloud database.

REFERENCES

1. Peter Mell ,Timothy Grance, "The NIST Definition of Cloud Computing", National Institute of Standards and Technology U.S. Department of Commerce, September 2011.
2. Ganesh Chandra Deka, "A Survey of Cloud Database Systems", IEEE Computer Society, pp.50-57, 2014.
3. Luca Ferretti, Fabio Pierazzi, Michele Colajanni and Mirco Marchetti, "Scalable architecture for multi-user encrypted SQL operations on cloud database services", IEEE Transactions on cloud computing, pp.1-14, 2013.
4. Joel Weis and Jim Alves-Foss, "Securing Database as a Service", IEEE, COPUBLISHED BY THE IEEE COMPUTER AND RELIABILITY SOCIETIES, pp. 49-55, 2011.
5. Arvind Arasu, Ken Eguro, Raghav Kaushik, Ravi Ramamurthy, "Querying Encrypted Data", IEEE, ICDE Conference, pp. 1262-1263, 2013.
6. Wenfei Fan, Jianzhong Li, Nan Tang, and Wenyuan Yu, "Incremental Detection of Inconsistencies in Distributed Data", IEEE Transactions on Knowledge and Data Engineering, VOL. 26, NO. 6, pp.1367-1383, JUNE 2014.
7. Pengcheng Xiong, Yun Chi, Shenghuo Zhu, Hyun Jin Moon, Calton Pu, Hakan Hacıg¨um¨us, "Intelligent Management of Virtualized Resources for Database Systems in Cloud Environment", IEEE, ICDE Conference, pp. 87-98, 2011.
8. Divyakant Agrawal, Amr El Abbadi, Shiyuan Wang, "Secure and Privacy-Preserving Database Services in the Cloud", IEEE, ICDE Conference, pp. 1268-1271, 2013.
9. Rohit Jain, Sunil Prabhakar, "Trustworthy Data from Untrusted Databases", IEEE, ICDE Conference, pp.529-540, 2013.
10. Dale D. Reitze, "Using Cloud Computing to Enhance Automatic Test Equipment Testing and Maintenance Capabilities", IEEE 2013.