

Cloud Computing and Data Masking Techniques: A Survey

Somi Khan

Research Scholar

School of Energy & Environment Management

Rajiv Gandhi Proudyogiki Vishwavidyalaya, Bhopal, Madhya Pradesh, India

Abstract— Cloud computing is technology, which provide scalable computation capacity, low cost as well as services to enterprise on demand for expansion, cloud computing is the development of distributed computing, virtualization technologies and grid computing which describe the shape of an era. There are a lot improvement in the cloud computing, security of the data in the cloud has become the one of major phase in the cloud computing, here we study the formation of cloud architecture as well as we equate cloud computing In this paper, aim is located different model of cloud computing as well as data masking techniques.

Key words: Cloud Computing, Data Masking

I. INTRODUCTION

Cloud Computing is a technique which use the internet and central remote servers to maintain applications and data. It is a group of computers and service station linked together over the internet refers to designing manipulating and accessing the applications online also allow the consumers and businesses to use applications without installation and access their personal file from any computer with the help of internet. It also offers online data storage, infrastructure and application. It is design for providing computing services by internet on claim and pay per use access to a pool of shared assets for the network storage, services and applications data is stored and maintained in data center of cloud provider like Google, Amazon, and Microsoft etc. The resources in cloud system are transparent for the application and the user do not know the place of resource. The user can access your application from anywhere. The amount of resources provided in the cloud system for the cloud system for the client is increased when their requirements are high and decreases when their requirements are less. The cloud computing can be seen as the important change of information industry and will make more impact on the development of information technology for the society.

Cloud computing is the combination and evolution of Virtualization, Software-as-a-Service and Utility computing, (SaaS), Infrastructure-as-a-Service (IaaS) and Platform-as-a-Service (PaaS). Cloud is a metaphor to describe web as a space where computing has been pre-installed and exist as a service; data, operating systems, applications, storage and processing power exist on the web ready to be shared.

II. CLOUD COMPUTING SERVICE MODELS

A. Three types of cloud computing services models:

1) Software as a Service (SaaS)

It is the top layer provider in which customer with ready to use applications running on the infrastructure provider. SaaS can be explained as a process by which Application Service Provider (ASP) provide different software application over

the internet. SaaS applications are pattern for end users, deliver over the internet. It enables the customer to eliminate of installing and operating the application on his own computer and also get rid of the immense load of software maintenance. With SaaS a provider licenses an application to the customer as a service on demand through subscription.[2] Generally the customer is only able to modify parameters of the application that have been exposed by the provider. The customer should have knowledge protecting data against administrative access by the provider. . The customer should understand the data encryption methods which are applied to data. The customer needs to be aware of how secure data, as defined in their data classification, is to be handled in general and by configuration options[3]. Salesforce, Zoho, workday are instances of SaaS which are used for email, billing etc. SaaS applications run on a SaaS provider's servers SaaS customers have no need of hardware or software to buy, install, maintain, or update. Its access to applications is easy. User just needs an Internet connection. It delivers a single application through the browser to thousands of user using architecture. On the customer side it means no investment in servers or software licensing and on the provider side just one app to maintain, costs are low as compared to conventional hosting. Office software is the best example of businesses in SaaS. Tasks related to accounting, sales and planning can all be performed through Software as a Service. In an organization everyone who needs to access to a particular piece of software can be set up as a user, whether it is one or two people or every employee.

a) Benefits of SaaS

- SaaS helps to managed software from a central location.
- Software delivered in a 'one to many' model.
- The user can sign up and quickly start using ingenious business apps.
- There is no setup costs with SaaS, as these are available with other applications.

2) Platform as a Service(PaaS)

It is a middle layer which provide platform oriented facility. In this consumer has responsibility for application arrangement as well as offer securing access to the application itself. PaaS is especially useful for situation where multiple creator working on a development project.. Here the customer does not regulate the underlying cloud set up including network, operating systems, and storage, but it control over the organized applications and possibly configuration settings for the application-hosting environment. Google App Engine Load Storm are the instances of PaaS for executing web applications over internet.[4]

PaaS is a combination of a development platform and a solution stack, delivered as a service on demand. It provides framework on which software developers can build

new applications or extend existing ones without the cost and complexity of buying and managing the hardware and software. The customer uses a hosting environment for applications. Most cloud offerings, PaaS services are generally paid for on the basis of agreement with clients.

a) Benefits of PaaS

- Develop application and get to market faster.
- Integration with web services and databases via common standards
- Reduce complexity with middleware as a service.
- Teams in various locations can work together
- Makes development possible for non-experts

3) Infrastructure as a Service (IaaS)

IaaS can be operated by enterprise client to create cost effective and easily climbable IT solutions where the difficulties as well as cost of managing the original hardware are outsourced to the cloud provider. The user can buy the infrastructure according to the requirements instead of buying the infrastructure that might not be used for months. IaaS works on a —Pay as you go model, meant for startup or small business; one of the most problematic things to do is keep capital expenditures in control. In cloud you must have the ability to scale as if you owned your own hardware and data center that the users pay for only what they are using. Virtualization enables IaaS provides to pose almost unlimited occurrence of servers to users or make it cost-effective use of the hosting hardware

a) Benefits of IaaS

- Infrastructure scales on demand to support dynamic workloads.
- Generally include multiple users on a single piece of hardware.
- Flexible and innovative services are available on demand.
- No need to invest in your own hardware.
- Physical security of data center locations

III. DEPLOYMENT MODELS

Deployment models is define as the type of accesses towards the cloud i.e. how the cloud is situated? Cloud computing having four type of approach: - Public, Hybrid, private and community.

A. Public cloud

Public cloud which is based on standard cloud computing, services may be free or offered on a pay-per-use model.[4] The public cloud allows system and services to be easily accessible to general public. Public cloud may be less secure because it is open to everybody. Public clouds offer service, usually over an internet connection. A public cloud is lying on the internet and designed to be used by any user with an internet connection to provide a similar range of capabilities and services. Public cloud users are mostly residential clients and connect to the public through an internet service provider's network. Google, Amazon and Microsoft are examples of public cloud who offer their services to the general public.

Public cloud providers manage the infrastructure and resources required by its users. Organization can use public clouds towards their operations meaningfully more effective, for example, with the storage of non-sensitive

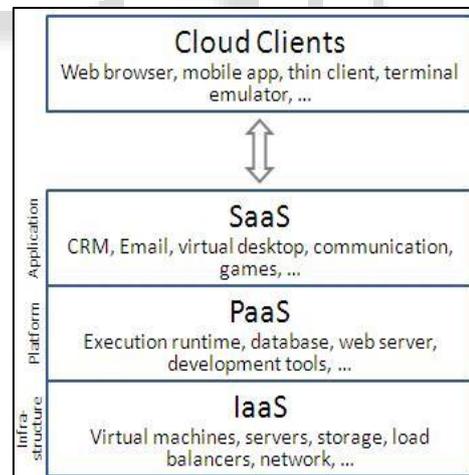
content, online document collaboration as well as webmail. Although one of the biggest problem facing public cloud computing is security, the cloud computing prototype provides opportunities for foundation in provisioning safety services that hold the outlook of improving the overall security of particular organizations. Which must need some selected public cloud computing solution should be deployed, configured as well as managed to run their security and other requirements.

The public cloud offers following benefits

- Public cloud provides ultimate scalability.
- Cloud services like IaaS, PaaS, SaaS follow the public cloud, so it is more flexible.
- It is Location independence means its services are available through internet whenever the client is required.
- Public cloud is also rate effective because it carries together resources which are shared through all subsequent cloud services.

B. Private cloud

A private cloud provides more security than public clouds. It is set up within an organization's internal enterprise data center. The scalable resources and virtual applications provided by the cloud vendor are merging together which are available for cloud users to share and use. The use of private cloud can be much more secure than that of the public cloud because of its specified internal exposure. The organization can access to operate on a specific Private cloud. Corporations are deciding that the private cloud proves less risky. The ability of Private cloud is to simulated services maximizes hardware control, ultimately dipping price and complexity. Most important resources of some organization



Architecture. Security is most important concern in cloud computing. Data masking is the process of hiding ornate its resources and its data. The major disadvantage of private cloud is its price. When comparisons are made with public cloud the rate of obtaining equipment kit, software often results in higher price to an organization in isolated cloud. Through, underneath the private cloud model, the cloud is only accessible thru a single organization provide that organization with control and privacy. A isolated cloud is named by way of Internal Cloud live thru in the company environment in addition to its access is restricted regularly to company employee and business partner.

Private cloud offers following benefits

- Flexibility and scale which meet client demands.
- Resource sharing among a large number of users.
- Payment according to use of the services.
- Use of technologies and internet protocol to access cloud resources.

C. Hybrid Cloud

A Hybrid Cloud is an integrated cloud services which use both private and public cloud to perform distinct functions within the same organization. It defined as multiple cloud systems that are joined in a system which allows programs or data to be moved effortlessly from one system to another [9]. It is a configuration of at least one private cloud and at least one public cloud. This computing model combines the security benefits of a private cloud as well as public cloud. . Hybrid Cloud provides more security to control data as well as applications which allows numerous parties to access information over the Internet.

A hybrid cloud presented in one of two ways: a merchant has a private cloud and forms a partnership with a public cloud provider, or a public cloud provider forms a partnership with a vendor that provides private cloud.[9] In hybrid cloud, an organization manages some resources in house and some out-house. Typically, the hybrid technique allows a business to take benefits of the scalability and cost- efficacy that a public cloud computing environment offers deprived of exposing data to third-party vendors.

Hybrid Cloud offers following benefits

- The hybrid cloud is responsible for security as the private cloud element of the hybrid cloud model give the security where it is needed for sensitive operations and also satisfies customer requirements for data handling and data storage where it is applicable.
- Supplies support for cloud-bursting.
- The hybrid cloud provides flexibility according to the availability of together secure resource as well as scalable cost effective public resource can provide organizations with more opportunities to explore for different operations.

IV. SYSTEM ARCHITECTURE

The architecture refers to the components and subcomponents essential for cloud computing. These components typically consist of a front end platform which consist of thin client, fat client, mobile device back end platforms which consist of servers, storage. These combined, components make up cloud computing original data with random characters or data. The purpose of data masking is to protect data that is categorized as personal identifiable data or sensitive data. In data masking data may be altered in different technologies including character stuffing, encryption, and character of word substitution. The overall practice of Data Masking at an organizational level should be tightly coupled with the Test Management, underlying Methodology and should incorporate processes for the distribution of masked test data subsets

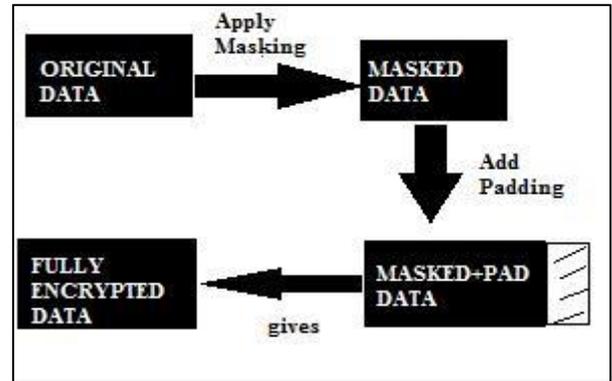
A. Working of Proposed System

This process shows how we encrypt the data so that the invader does not know what the actual data is about.

In this we use Data Masking and along with it the padding of data is applied

B. Data Security in the Cloud

1) Sender Side:

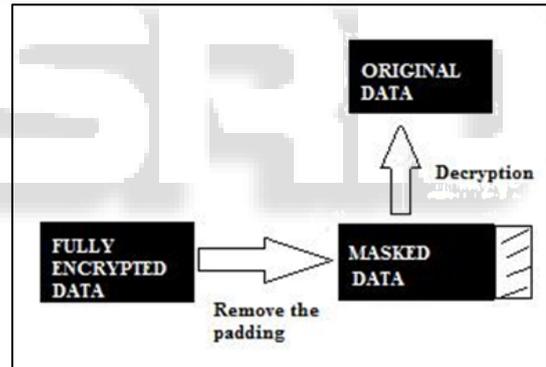


a) Explanation

In the above diagram, the encryption process takes place at the socket layer of the sender side.

The diagram shows us that the mask is applied to the data so that the original data is not being reflected to invader. After masking, the data which is masked is applied with the padding and hence the data is much more secure than it was before. It gives us the double encryption of the data and hence the data is much more securely transferred to the receiver.[6]

2) Receiver side



a) Explanation

Reverse the process of the sender happens, At the receiver side

In this as shown in the diagram, the data which is encrypted doubly is being decrypted doubly.

At first the filling of the data is removed and after the removal of the filling we obtained the masked data.

Now the masking of the data is being removed and hence the receiver obtains the original data.

b) Need of Data Masking

- When copy sensitive data outside of production environment.
- Moving the test data to cloud.
- Sending data to vendors.
- Leverage off-shore development/consultant

C. Different types of Masking

1) Static data masking:

Static data masking technique is used by most of the organization when testing in fact is the only possible masking way when using subcontract developers in a distinct location or a separate company. In these cases it is

necessary to duplicate the database. Once doing so, critical to use static data masking technique. These tools make sure that all sensitive data is masked before transfer it out of the organization. Static data masking provides a basic level of the data protection by creating testing database or an offline, using a standard ETL procedure. [7] The static data base can be updated repeatedly, for example on a daily or weekly basis. This is not a security risk, but it often has conclusion for a variety of tests and development issues.

V. DATA MASKING AND THE CLOUD

latest years, organizations developed their new applications in the cloud. The cloud solution as of now allows organization to use IaaS, PaaS, SaaS. There are various models of creating test data as well as moving it to the cloud. Data masking becomes the part of these processes in SDLC as the development environments.

VI. DATA MASKING TECHNOLOGY

A. Substitution

The Substitution is nominal method of capable to preserve the authentic aspect of the data records of data applying masking. It is completely unrelated to the actual details but technology substituting of randomly contain the contents of a column of data with information that aspect similar. For example, to sanitize surnames by substitution, a list of random last names available, Substitution data can sometimes be very hard to find in large length - but any data masking software should contain datasets of usually required items. [7] As the surnames in a user database clean through replacing the last names with surnames pinched from a largish random list. [8] list of the phone numbers must be available formerly to clean telephone one numbers. substitution technique need to apply for many of the arenas [11] telephone numbers, credit card numbers zip codes, and other card type numbers like Social Security numbers in DB structure

B. Shuffling

Shuffling is similar to replacement apart from that the substitution data is derived from the column itself. In simple terms the data is randomly shuffled with the column. Shuffling is effective for small amounts of data. Alternative consideration is the algorithm used to scuffle the data. If the shuffling method can be defined, then the data can be easily —unshuffled. [8] For example, if the shuffle algorithm simply ran down the table swapping the column data in between each group of two rows it would not take much work from an attention party to revert things to their unshuffled state. Shuffling is rarely effective when used on small amounts of data. [10] For example, only five rows in a table it perhaps will not be too difficult to figure out which of the shuffled data really belongs to which row. [10] On the other hand, if a column of numeric data is shuffled, the sum and average of the column static work out towards the same amount. It is sometimes beneficial.

C. Encryption

Encryption is most complex methods to resolution the data masking problem. The Encryption technique algorithmically mix-up the data. This usually does not leave the data looking

truthful and can sometimes make the data larger. Encode on also destroys the formatting and look and feel of the data. Encrypted data rarely looks meaningful; in fact, it usually looks like binary data. This principal to character set problems when manipulating encrypted varchar fields. Certain types of encryption levy constraints on the data format as well. [7] This means that the fields must be extended with a suitable padding character which must then be stripped off at decryption time.

VII. CONCLUSION

Mostly organization wants combination of static and dynamic database masking. In this paper we discussed about the deployment models, cloud services models, and security in cloud by using data masking techniques. Storage data on the cloud improve the way we achieve the storage of data and access the data from the cloud also mentioned about the various cryptography algorithms which help us to encrypting the data at source side as well as then transferring it to the receiver side. This paper also discover the need of data masking in present information. Data masking will allow us to achieve the following: (a) Increase protection against data theft. (b) Enforces 'need to access'. (c) Provides realistic data for testing, development and data sharing. (d) Provides a heightened sense of security to clients, employee and supplier.

ACKNOWLEDGMENT

I would like to express my gratitude to my project guide Dr. Savita Vyas, Associate Professor, School of Energy & Environment Management for guidance, suggestion, support, help and constructive criticisms throughout the period of project work. Without her able guidance it would not have been possible to complete the project in time. I would also like to thank Dr. Mukesh Pandey, HOD, School of Energy & Environment Management for School of Energy & Environment Management for providing the facilities required in completing the project work. I am grateful to other engineer staff members for their guidance throughout the project and critical comments which improved the quality of this report.

I am also thankful to my parents for their support and encouragement during my working period.

REFERENCES

- [1] Research on Cloud Data Storage Technology and Its Architecture Implementation Kun Liua, Long-jiang]Donga
- [2] An Advanced Survey on Cloud Computing and State-of-the-art Research Issues Mohiuddin Ahmed1, Abu Sina Md. Raju Chowdhury2, Mustaq Ahmed3, Md. Mahmudul Hasan Rafee4 Department of Computer Science and Information Technology, Islamic University of Technology Board Bazar, Gazipur-1704, Bangladesh
- [3] Cloud Computing Security Issues And Challenges Pranita P. Khairnar, Electronics Department, Amrutvahini College of Engineering, Sangamner Prof. V.S. Ubale Asst. Prof. Electronics Department, Amrutvahini College of Engineering, Sangamner

- [4] Various Issues & Challenges of Load Balancing Over Cloud: A Survey Palak Shrivastava¹ Sudheer Kumar Arya², Dr. Priyanka Tripathi³ Research Scholar National Institute Of Technical Teachers Training & Research Bhopal, India shrivastava, Associate Professor National Institute Of Technical Teachers Training & Research Bhopal, India.
- [5] International Journal of Advanced Research in Computer Science and Software Engineering Identity – Based Secure and Flexible Data Sharing in Cloud Based Smart Grid R. Ruhina Student, (SE), Department of Computer Science and Engineering, JNTUA College of Engineering, Anantapuramu, Andhra Pradesh, India
- [6] DATA SECURITY IN THE CLOUD :Mr. Jiten Prithiani MCA Final year Student ,V.E.S. Institute of Technology, Mumbai, India jiten.prithiani@ves.ac.in, Mrs. Dhanamma Jagli, Department of MCA, V.E.S. Institute of Technology ,Mumbai, India
- [7] Design of Data Masking Architecture and Analysis of Data Masking Techniques for Testing Ravikumar G K, Dr. B. Justus Rabi, Dr MGR University, Chennai, Tamil Nadu, INDIA
- [8] Data Masking: What You Need to Know What You Really Need To Know Before You Begin A Net 2000 Ltd. White Paper.
- [9] Swot Analysis of Mobile Cloud Computing A.M.S.Zunaitha Sulthana*, L.Clara Mary*, A.Sangeetha Department of computer science, MIET Institution, Trichy.
- [10] A Survey on Recent Trends, Process and Development in Data Masking for Testing Ravikumar G K¹ ,Manjunath T N², Ravindra S Hegadi³, Umesh I M⁴
- [11] Data Sanitization Techniques A Net 2000 Ltd. White Paper