Data Protection as a Service (DPaaS): A Novel Approach For Data Protection in Cloud

Jeelani
Department of CSE,VTU,CIT,Gubbi

Anil Kumar .G
Department of CSE, VTU,CIT,Gubbi

Abstract – **Cloud computing has been evolved as a next generation computing and changed the way organizations see the IT infrastructure. Cloud computing is paradigm which provides pool of IT resources over internet as a service on rent. It is becoming popular day by day because it reduces operational cost by preventing upfront investment of computing resources and frees the user from maintenance. Clouds providers offer centralized data centers for storing user data. In cloud computing the primary challenge is to gain the trust of the user by ensuring them that their data is safe from unauthorized access. Data security and privacy issues has become a major obstacle to the wide adoption of cloud services. To address this problem, , we proposed the service model based on idea provided by Dawn Song, Elaine Shi, and Ian Fischer. The proposed architecture provides security to cloud data by providing access control, logging, and key management mechanisms. Encryption techniques and third party auditing service are used to provide data security and to protect data integrity.**

Index Terms-- **Cloud computing, DPaaS** ,D**ata protection as a service, Encryption, Third party auditing**

I.INTRODUCTION

Cloud computing, to put it simply, is nothing but "Internet Computing." To visualize the Internet cloud is used; hence the term "cloud computing" is used for computation done through the Internet. Cloud Computing offers its users huge database storage space and complete business application via the Internet from anywhere, any time and as long as they need, without worrying about underlying details such as maintenance or management of actual resources.

Cloud computing is computing model that make pool of IT resources like storage, networks, interfaces, servers, middleware, and applications provided over internet as a services to individual customers and business organizations on a leased basis. Cloud computing services are delivered by a third party cloud service provider who is the owner of the infrastructure.

Cloud computing model is popularly called as "pay-per-use" or "pay-for-what-you-use" model ,because it eliminates the need for an up-front investment or long term commitment, thus by reducing capital and operational costs[1]. In cloud computing data provided by the user will be stored in centralized cloud data storage at remote servers. Cloud computing is a mixture of early technologies such as Grid computing, Utility Computing and virtualization techniques. Cloud computing is often confused as the outcome off advances in grid computing but it is not true. The path for the evolution of cloud computing is paved by the Grid computing[1].

A. The key Characteristics of Cloud computing

- **Multi- Tenancy (resource sharing):** Cloud computing is based on a business model in which resources are pooled to serve multiple users ,with different virtual and physical resources are dynamically assigned and reassigned based on demand [2].
- **Rapid Elasticity:** Users can rapidly increase and decrease the resources required as needed. To the consumers the resources appear to be unlimited and be purchased in any quantity[2].
- **On –Demand service :** Resources are allocated and de-allocated dynamically on demand, without requiring human interaction with each service provider[2].
- **Pay-By-Consumption :** which means utility pricing ,that is Users pay only for the resources they use.
- **Broad Network Access:** capacities that are available over the network and can be accessed from different devices (e.g., Desktop computers, laptops, mobile phones, personal digital assistants (PDAs) and tablets) [2].

B. Cloud Computing Deployment Models

Public Clouds: In Public cloud the services are provided dynamically to the consumers by third party service provider over the internet via web applications/web services. Consumers can use it on demand with pay-as-you-use pricing model. Public clouds are less secure compare to other cloud models.

Private Clouds: Private clouds are deployed , maintained and operated within an organization's internal enterprise datacenter. private cloud can be more secure than that of the public cloud because only the organization and authorized users have access to specific Private cloud.

ISSN: 2278 – 1323

Private clouds can also managed by third parties on the premises .Due to security reasons more and more companies are choosing Private Clouds over public clouds.

Hybrid Cloud: A hybrid cloud is a combination of the of the Public cloud and private cloud. In hybrid Cloud users outsource the non critical business information and processing to public clouds and keep critical services and sensitive data in private clouds.

Community Cloud: A community cloud is used and controlled by several organizations who shared common interests, like security requirements and common mission.

C. Cloud Service Models

Software as a Service (SaaS) - Software-as-a-Service model provide the user access to the software applications over a network, typically the Internet. SaaS eliminates the up-front software licensing and infrastructure costs and also reduces the operational cost. SaaS allows to have the same software on all of your devices at once by accessing it on the cloud.

Platform as a Service (PaaS) - PaaS provide development platform for the developers to design, build ,test and manage their applications without having any knowledge about what is going on underneath the service. Paas shortens the time required for software development and delivery since software development and testing is performed on same Paas platform.

Infrastructure as a Service (IaaS): The "Infrastructure as a service" (IaaS) provides" computing infrastructure like servers, network devices, storage and backup systems on rent on pay-as-you-use basis according to hours of usage. A client or an enterprise can use these services as a platform for building their applications.

We propose a cloud computing paradigm called Data Protection as a Service (DPaaS), which provide data security and privacy, and considerably reduces the per-application development efforts required to provide data protection while still allowing rapid development and maintenance.

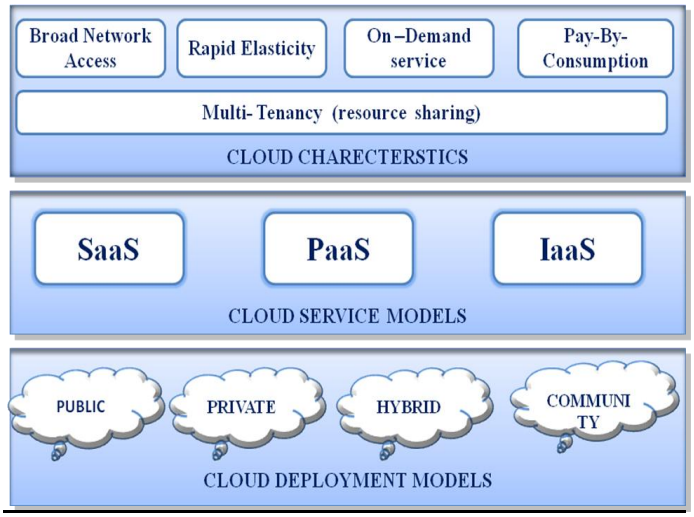

Fig 1.Cloud characteristics, service and deployment models

II.LITERATURE REVIEW

Though cloud Computing is a rapid growing technology and number of clients and cloud service providers are increasing rapidly, there is a huge concern about the protection, security and privacy of data. Although cloud computing reduces cost, provide rapid scaling, easy maintenance, and service anytime, anywhere, how to ensure and gain the confidence of the user that the cloud can handle their data securely is a key challenge . A survey by Microsoft says that most of the public and business organizations are excited about the cloud computing services, but they are slow in fully accepting it due to security issues and privacy challenges associated with it[3].

In cloud computing environment Data security and privacy is the major problem that requires the cloud service provider and Consumers to take the necessary measures to manage and secure cloud services. In cloud computing, control over personal data is shared between the user and the cloud service provider. By transferring control of data to a third party service provider, responsibility of preventing data leaks is shifted to the service provider. A good cloud service providers will provide strong encryption and data security systems.

A. The Data security issues in cloud are:

1. Data Availability : Data Availability is assuring that data will be available to the user s over the network in location independent manner through standard mechanism using various platforms(e.g. ,mobiles, desktops, laptops, PDAs , etc)[4].

2. Data Integrity : Data Integrity assurance that the messages or data is received as sent without any modification or alteration and no duplication. Integrity can be applied to single message, stream of messages or some particular field within the message. Data Integrity is compromised or violated if the received message is not exactly same as the sent message[4].

3. Data Confidentiality : Confidentiality is protection of information from unauthorized users. The eavesdroppers should not be able to view the data, not be able to guess source and destination of data[4].

B. Advantages of Cloud Storage

- A primary advantage of cloud data storage is the data can be accessed from anywhere in the world via internet connection. Offices in different locations

ISSN: 2278 – 1323

can access the same data at same time. Cloud provides continuity and effectiveness for organizations in which employees are working from multiple locations. Travelling employees or people can access data from anywhere in the world[5].

- Disaster Recovery is easier with cloud storage configuration. Organizations or companies can retrieve their data whenever needed, by assuming that their storage provider wasn't affected. If our organization or company is also affected with a disaster, copies of our data remains safe on the cloud servers[5].

- Cloud computing allows flexible storage requirements. When storage needs increases then we can have a huge amount of virtual storage available in the cloud ,which avoids adding extra hard drives. Company can scale cost depending on storage requirements and pay only for what they are using.

- As our company or organization is paying usage charges for certain amount of storage space on monthly or annual basis, it reduces the operational cost. That usages charges are much lesser than the upfront investment of hardware backup and which also need expertise in configuration and management [5].

C. Disadvantages of Cloud Storage

- With a cloud storage solution, you are completely depending on a third party cloud service provider to effectively protect your confidential data.

- Using cloud-based technologies means you need to provide your service provider with access to your private data. You do not have control over the servers. You do not knowledge about the location of those servers, and you always have chance that your cloud storage company may go out of business[5].

- Cloud has multi-tenant architecture where multiple users are hosted on the same server, a hacker might try to break into the data of other users hosted and stored on the same server

- There is a data loss threat from cloud service providers themselves. Accidental deletion of data may occur more often than many people may think and this causes damage to the reputation of consumers and the service provider[5].

- Changing of the cloud providers can be very difficult job. You need to download all data to your local storage and then upload all the data again to the new cloud provider's storage, which would require huge storage capacity, large enough to store all of the data you stored with the previous cloud storage provider [5].

The major goals of data protection service is :

- **Data Privacy:** The user's private data should be not leaked to unauthorized users.
- **Data Integrity:** The user's data must be stored faithfully and not be corrupted.
- **Access transparency**: Log of accesses to data must be transparent indicating who performed what on each access of data.
- **Ease of verification:** It should be possible to offer some level of users transparency, such that they can verify to some extent what platform or application code is running. Users may also like to verify that their privacy policies have been strictly administered by the cloud or not.
- **Provide Rich computations:** The platform must allow most of the computations on sensitive data, and can efficiently run those computations.
- **Support for development and maintenance**: Every developer faces a long list of challenges like, Errors or bugs to find and fix, frequent software upgrades, continuously changing usage patterns, and demands of users for high performance. Any reliable data protection approach must be able to handle these issues, which are usually overlooked in the literature on the topic [6].

III.EXISTING SYSTEM

Cloud Computing paradigm has emerged as a well-known technology of the IT industry that provides huge storage capacity, resource pooling and sharing, on-demand network access to computing resources on a pay-per-use basis, that can be provisioned and released rapidly with very negligible management effort or interaction of service provider.

Even though cloud computing reduces development cost, provides rapid scaling, easy maintenance, and service availability anytime, anywhere but lack of security is the major obstacle for adoption of cloud. A key challenge is how to gain users trust and build confidence that the cloud can handle user's private data securely[6].

Users want to maintain control of their data, but they also want to get the benefits from the services provided by application developers using that data. Till now the cloud providers offer little platform-level support for protecting user data ,beyond data encryption at rest.

A primary challenge in designing a platform-layer solution for many applications is ensuring that it enables rapid development and maintenance.

ISSN: 2278 – 1323

IV.PROPOSED SYSTEM ARCHITECTURE

The proposed architecture is the new paradigm for cloud Computing, named as DPaaS (Data Protection as a Service) which is meant for mass data protection in cloud. It is a collection of security primitives which imposes security and privacy on data and provide the affirmation of data security to the cloud users, even in the presence of potentially compromised or malicious applications[6].

This service model is built into the cloud computing environment , can reduce the effort required by cloud service providers to have mechanisms to protect data of the cloud users. This will increase cloud efficiency in terms of data protection.

This will obviously encourage more number of users to adopt cloud and thus cloud computing becomes even more reliable and popular technology. The architecture is developed keeping several issues in mind such as key management, access control ,IT resource sharing, rapid development, performance and maintenance.

In this approach access control polices and key management are moved into middle tier which is nothing but the computing platform that provides a common platform for all the users of cloud computing[7].

It provides cryptographic protection to data at rest and offers logging and auditing services to provide accountability.

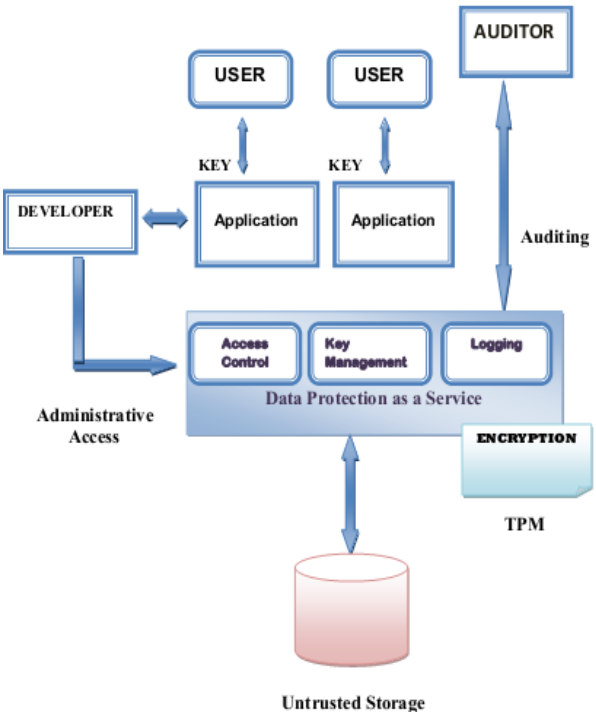

Fig 2. DPaaS Architecture

V.MODULE DESCRIPTION

1. Cloud Computing Storage Service
2. User Module
3. Third Party Auditor
4. Trusted Platform Module

1 Cloud Computing Storage Service

Cloud Computing is the provision of dynamically scalable IT resources over the internet as a services , individuals can access applications from anywhere in the world on demand**.** A cloud computing service provider offer huge amounts of storage space, and control over personal data stored in cloud is often shared between the customer and the service provider **.**

Cloud computing eliminates the need for up-front commitment or investment by users, by allowing them to request and use resources only when they are required. Today a large number of individual customers and business organizations are using the cloud services , and also the adoption rate is increasing day by day[3].

2. User Module

The user module is nothing but a user interface. It allows the user to login with his account and perform several operations such as- uploading his private data Into the cloud, checking the status of uploaded file, viewing his data or file.

User store large amount of private or sensitive data to the cloud storage and can view his data later using secure key. Admin provide secure key to user after encrypting data. Data encryption is done using Trusted Platform Module (TPM). User store his data after auditor verify data and also alter the data if required. User gets a message if auditor alters the user data.

3. Third Party Auditor

The Third party auditor is one who audits the overall performance of the System, and can track all the transactions and logins of users. In this module auditor is software that is capable of tracking user transactions. Auditor views the users data and verify data and also can change the data, after auditing the data, store data to the cloud[3]

4. Trusted Platform Module

Trusted Platform Module (TPM) acts as the crypto processor which encrypt each of the data bit before it going to the database and store cryptographic keys that protect information. It is often called as the "TPM chip" or "TPM Security Device".

ISSN: 2278 – 1323

The proposed architecture illustrated that the various technologies such as logging ,encryption , application confinement and information flow checking can be potentially combined at high level [3].

Encryption overveiw

Encryption is the process of conversion of data or messages into unreadable form in such a way that only authorized users can read it using key provided. The data or message is referred to as plaintext before encryption and ciphertext after encryption [8].

Decryption is the inverse process of encryption, after decryption ciphertext is transformed back to original data or plaintext. An encryption process uses a key for encryption, and only authorized users can decrypt the message with the key provided by the originator[8].

Encryption algorithms are used to provide security to the message or data by scrambling its contents by performing various mathematical transformations and substitutions so that it can be readable only by authorized users having encryption key to unscramble it.

There are two types of encryption:

1. Symmetric Key Encryption (Private Key Encryption)
2. Asymmetric Key Encryption (Public Key Encryption)

1.Symmetric Key Encryption (Private Key Encryption) :
In this type of encryption algorithms same key is used for both encrypting and decrypting the data or information. The key is called a secret key because it is kept as secret and shared only between the sender and receiver of the message.

The major advantage of Private Key Encryption is execution speed is high compared to asymmetric key encryption because single key is used at both the ends of the network.

2.Asymmetric Key Encryption (Public Key Encryption):
In this type of encryption algorithms different keys are used for encrypting and decrypting the data or information.

In this approach private key is kept secret and never shared between the users while public key is made available to all. If a send want to send a message to the receiver he must encrypt it using public key, at receiving side user must decrypt message using corresponding private key.

Public Key encryption is slower compared to private key encryption because of multiple key management between sender and receiver.

The private data is encrypted if it is stored into the un trusted cloud storage to ensure that the cloud service providers does not misuse or alter the data stored in the cloud.

In our proposed system we provide data security on cloud by using symmetric block encryption algorithm called Advanced Encryption Standard (AES).AES uses fixed 128 bits block length and key length can be 128,192, 256 bits. Most commonly used key length is 128 bits.

In AES algorithm data block is processed using substitutions and permutations . A number of AES transformation rounds depend on the key length. For example, if the key used is of 128 bit length then there are 10 rounds ,if it is of 192 bits then 12 rounds and for 256 bit key 14 rounds.

During each round, the following operations are performed:

SubBytes: Each byte of the block is replaced by another one, using the S-Box .

ShiftRow: Each row of state is shifted a certain amount to the left.

MixColumn: Each column vector is multiplied by a fixed matrix similar to multiplication of columns of the matrix.

AddRoundKey: Each byte of the state array is combined with a round sub key, which is a different for each round.

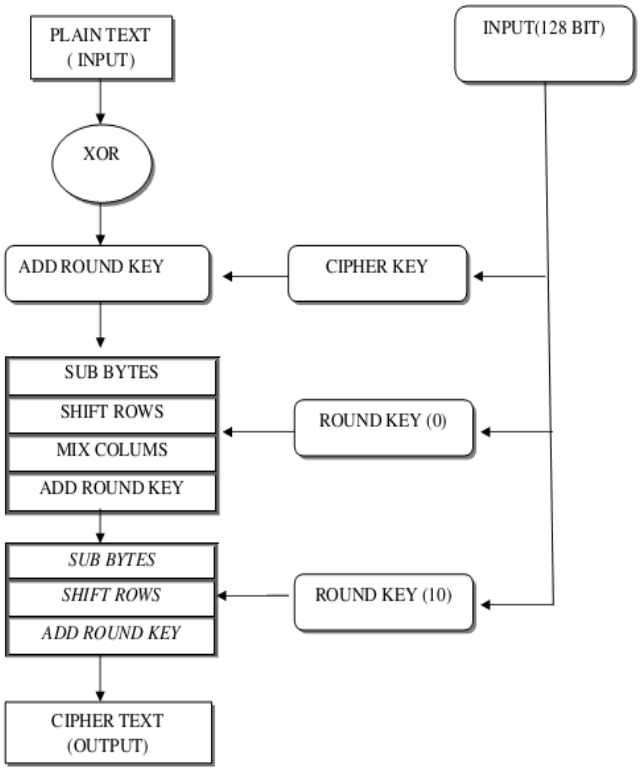

Fig 3. AES Round Structure

ISSN: 2278 – 1323

VI. CONCLUSION

Most of the organizations now a days are shifting to Cloud Computing to reduce hardware and software capital investments. However, they are also worried about the security risks of Cloud Computing, and losing direct control over their sensitive data which is the major barrier for cloud adoption.

In this paper we have discussed the cloud terminology, cloud characteristics, advantages and disadvantages of cloud storage ,security concerns, and issues regarding data security, and proposed cloud security model known as "Data protection as a service" (DPaaS). The proposed architecture provides cloud data security by providing access control, logging, and key management mechanisms. Encryption techniques are used to provide data security at rest, and third party auditing service is used to protect data integrity.

REFERENCES

[1] Cloud Computing by Dr. Kumar Saurabh, Wiley India,2011

[2] Prabhleen Singh,Ketki Arora," Security of Data in Cloud Environment Using DPaaS", *IJMER | ISSN: 2249–6645| Vol. 5 | Iss.1| Jan. 2015*

[3] K.Bhima,,S.Suresh. *",*Privacy Data Control and Data Protection as a Service in Cloud Computing*." International Journal of Advanced Research in Computer and Communication Engineering Vol. 2, Issue 9, September 2013*

[4] Neha Tirthani, Ganesan R. "Data Security in Cloud Architecture Based on Diffie Hellman and Elliptical Curve Cryptography"

[5]http://www.iwebsoftsolution.com/the-advantages-disadvantages-of-cloud-storage-solutions/

[6] Dawn Song, Elaine Shi, Ian Fischer, Umesh Shankar."Cloud Data Protection For The Masses" *Computer, vol. 45(1),Jan 2012 page(s): 39-45.*

[7] Ashok Jammi, Meena.S, D V Arjun," Design and Analysis of Data Protection as a Service (DPaaS) for Cloud Computing", *(IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 5 (1) , 2014.*

[8] http://en.wikipedia.org/wiki/Encryption.

ISSN: 2278 – 1323