International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2013): 6.14 | Impact Factor (2014): 5.611

Survey: Using Multiple Storage Clouds for Data Security on Cloud

Sandeep Nehe[1], M. B. Vaidya[2]

[1]Pune University, AVCOE Sangamner

[2]Professor, Pune University, AVCOE Sangamner

Abstract: ***Currently there are several cloud systems which provide storage as a service which are PaaS systems. They provide access to their services via web services. Can we create a unification framework that integrates different IaaS cloud layers in order to collaborate storage space of two or more storage clouds? We have one question always in mind i.e. How can we provide security to the files being uploaded on cloud storage? People are always worried about the security for the data being uploaded to storage cloud. One way of doing this is encrypting the files being uploaded. This ensured that the files cannot be opened by any other person. But if once decryption key is known then the data can be easily. In this paper we proposes a middleware system that authenticates users and uses storage cloud systems such as DropBox, Box.net and OneDrive as backend storage. We provide a layer of security above these storage systems by distributing the storage to each cloud. We also encrypt each segment being distributed. Our middleware split a file being uploaded into segments, encrypt them and then upload each segment to each backend cloud storage. The system maintains all these segmentation, encryption and distribution information and user information to the systems database. When user want to download a file, system again recreate the original uploaded file and allow downloading it. We use web services or web APIs provided storage clouds for implementing this system.***

Keywords: Data Security, Data slicing, Cloud computing, Service Computing.

1. Introduction

Today, commodity cloud providers especially from the Infrastructure as-a-Service (IaaS) cloud expose their service APIs which allows enterprises to create workflow applications. This paper introduces an application that is a middleware-layer that handles the authentication process on behalf of the user in real time. The middleware consume data from/to IaaS Storage clouds such as Dropbox, Box.net and OneDrive. Further, the middle-ware employs Personal Login to identify the end-user. The deployment of the middleware enforces additional data protection. We further provide one layer of security by splitting the file into segments, encrypting each segment and uploading them to cloud storage.

Cloud computing which is the era that has witnessed the virtualization of internet based services such as applications, storage, databases, and networks, and so on has long been categorized into three major taxonomies. There are: Platform as a Service (PaaS) allowing customers to utilize development environments from entity providers (e.g., Google App Engine and Microsoft Azure), Infrastructure as a Service (IaaS) allowing data consumers and providers to utilize virtualized servers, storage, and network for computing space (e.g., Amazon Web Services, Dropbox, GDrive), and Software as a Service (SaaS) allowing consumers and providers to access cloud hosted software and applications (e.g. social media). The benefits of cloud computing are enormous ranging from enterprise business agility and system scalability to data maintainability. However, the services cost charged by the cloud services providers is deemed economical which gives enterprise stakeholders the breathing space to cut down on internal IT infrastructural budget. But as a challenge, stakeholders have raised questions regarding the safety, privacy, and potential security risks of adopting cloud computing. This question is further fueled for public cloud usage. In cases where enterprises have huge economic muscles, they have deployed private cloud architectures thereby enforcing internal security and data protection. Also, some enterprises proposed hybrid-cloud architecture techniques whereby parts of the enterprise data and business infrastructure is handled by internal clouds and the other parts managed on public clouds. The public cloud services providers are also doing their very best to enforce data and information security on their clusters [2]. In this paper, we focus on the peculiar way to existing consumer cloud providers from the IaaS layer enforce security in their environments: Dropbox, Box.net, OneDrive etc. These IaaS layers require the data consumer (requester) to provide credentials such as access key, secret access key, harsh signature, and sometimes session id in order to access the stored digital assets which are mostly files and documents.

In this paper we presents a middleware-oriented framework that integrates different IaaS Storage Clouds. Based on the registered users access token, the middleware does the authentication with the IaaS cloud frameworks. The middleware relies on a service level manager which decides how to split a file being uploaded, its encryption and decryption followed by merger for download. The middleware supports the integrated authentication of all cloud platforms. The evaluation of the framework shows high improvement in the security as we are distributing the storage and also encrypting the data.

Licensed Under Creative Commons Attribution CC BY

2. Literature Survey

2.1 Cloud Storage

In Cloud storage digital data is stored in logical pools, the physical storage contains multiple servers or locations, and the physical environment is generally owned and managed by a hosting company. These cloud storage providers keep the data accessible and available, and the physical environment protected and running. Storage capacity is given on lease to people and organizations to store different type of data. Cloud storage services can be accessed via web services (API) or by applications that utilize the API or Web-based content management systems. Amazon Web Services introduced their cloud storage service AWS S3 is a one of the first cloud storage supplier. Other cloud storage services are popular services like Google Drive, Google Cloud Platform, Smugmug, Dropbox, Box and OneDrive.

Cloud storage is based on highly virtualized infrastructure. It is a multitenant system with multiple storage devices inside. A storage cluster is composed of storage rich nodes constructed from commodity hardware and connected by commodity interconnect. As common for cloud infrastructures, the storage cloud is built from low cost components, ensuring reliability in the software, and building advanced functionality on top of this foundation.

The Cloud storage can be public, private or hybrid. Private cloud storage services provide a dedicated environment behind an organization‟s firewall. Hybrid cloud storage includes components from at least one private cloud and one public cloud infrastructure. [2]

2.2 Web Services

Almost all of today‟s network based applications such as email, notification services, online file and document sharing, social media, and so on are modeled as web services with standards such as SOAP, WSDL, UDDI, XSD, and REST to ensure data availability and access at real time . Simple Object Access Protocol (SOAP) is explained as a Web specification that provides interoperability for heterogeneous Web services. This is achievable because SOAP messages can be exchanged over multiple communication protocols using HTTP and other transport protocols. Also, the use of HTTP for transporting SOAP messages has reduced the challenge for building services that can run on the Internet. Though the standard protocol used by SOAP is HTTP, it can use other protocols as well [1].

However there have been some challenges associated with SOAP in modern architectural designs of web services. While some perceived it as a complex way of de-sign, others saw that the simple way of turning legacy applications into web services can be misused. Further, SOAP uses the POST method for all operations and this leads to caching challenges especially when the protocol becomes too heavy in a wireless network. As a way to improve on bandwidth utilization and the use of clear semantics for system operations, some studies have proposed the REST design.

REpresentational State Transfer (REST) is another alternative to SOAP. REST is an architectural principle that uses the Web as a platform for distributed computing. The REST design has certain architectural principles. Everything is a resource: All the identifiable entities must be considered as a resource and should be assigned an ID.

Identification of resources through URI: The entities should be given a URI to aid interactions within the system and allow resources to be searched for globally.

Uniform interface: Resources can be manipulated through representations using HTTP verbs such as:
GET to fetch a resource,
HEAD to read the metadata of the resource,
POST to push or create a new resource,
PUT to change (update) the state of a resource and
DELETE removes the specified resource while
OPTIONS check the functionality of a web server.
There are other methods as well such as TRACE, CONNECT, and PATCH.

Stateless interactions: While resources are stateful, their interactions should be kept stateless. At the end of every transaction, resources should have information about themselves but not how the last interaction was done. This is the reason REST architectural design achieves high scalability.

Hypermedia as the engine of application state: In order to navigate between resources, URIs such as hypertext can be used in resource representation. [1]

2.3 Oauth

Open Authorization (OAuth) is an open protocol to al-low secure authorization from third-party applications in a simple and standardized way. The OAuth protocol provides an authorization layer for HTTP-based service APIs, typically on top of a secure transport layer, such as HTTP-over-TLS (i.e., HTTPS). OAuth defines three main roles in the above scenario:

- The User (U) is the entity which generates some sort of information;
- The Service Provider (SP) hosts the information generated by the users and makes it available through APIs;
- The Service Consumer (SC), also referred to as “client application,” accesses the information stored by the SP for its own utilization.

In order to comply with the security and privacy requirements must issue an explicit agreement that some client application can access information on user‟s behalf. For this client is granted a access token which containing user‟s and service consumer‟s identities, which must be exhibited in every request as an authorization proof. The OAuth 2.0 protocol is evolved from original OAuth protocol and aims at improving the client development simplicity by defining scenarios for authorizing web, mobile, and desktop applications. At SP,s side implementing OAuth is a com complicated and time consuming task which involves registration of both client applications and users and also

Licensed Under Creative Commons Attribution CC BY

authentication services. [3].

3. Proposed System Architecture

There are some open questions that need to be solved such as:

1) How do we access data from storage clouds? ,
2) How can we provide security to the files being up-loaded on cloud storage? [6]
3) How can we distribute our storage [5] to different cloud storage nodes in order to give security to the cloud data?

To answer above questions we present a middleware framework. Figure 1 shows the middleware framework for our proposed solution.

The framework is a three tier architecture with multiple sub layers. The architecture consist of User, IAAS layer and Middleware system. The mode of communication between these devices is HTTP.

A. User

The user can access the system from the browser from any computer. The system is hosted on the webserver and the users are expected to put URL in the browser. The user can login to the system from any device which supports browser and internet connection.

B. Data Slicing and Merging

The User when uploads the File, the middleware split the file into four segments. The data slicing algorithm is a simple algorithm which divides the total file size by four. The each file is created by file I/O system calls. New three files are created after this process. The merging of different files is created by opening one file in append mode and rest of the data from each other file is copied to it.

C. Encryption and decryption

We can use AES symmetric cryptography or any good encryption algorithm to encode and decode each segment file created by above step.

D. IAAS Layer oriented cloud services which specific to this are Dropbox, Box.net and OneDrive facilities. The facilities are employed primarily as repositories where documents in different formats are stored. The motivation for using four facilities is necessitated by our use case: Most of people are subscribed only for the 2GB free space. However, the idea is to extend storage by allowing saving document to the user on DropBox .When DropBox account is created, the customer is given an Access Key Id and a Secret Access Key. When a customer wants to access a resource on the Dropbox, she/he has to add to the HTTP request headers the access key id and secret key id. If the authentication test is passed, the DropBox facility will allow the request and issue a response back to the customer otherwise, the request is denied. It is interesting to note that, the other facilities also follow the same authentication workflow as the DropBox.

E. Middleware

The middleware is the hub that links all the components of the framework. The middleware has full control over the management of the security issues.
The middleware has three interfaces which are exposed to the following components:

1. Data Slicing/Merging, Data
2. Encryption/Decryption,
3. Upload/Download from Storage clouds.

To be able to access data on any of the IaaS facilities, the requester has to first establish communication with the middleware through the interface which is basically a publicly available link. When a request comes through this interface, the middleware redirects the user to an authentication page where the user can choose authentication mechanism. Our middleware system also proposes to split the file into segments .Then each segment is encrypted and then uploaded to each cloud storage at the backend. This method has following benefits:

1) The data cannot be recovered by single cloud storage.
2) Only our system can recover the data, because data is encrypted which provides one more security layer.

When user wants to upload the file, the file is first pass through our middleware. The middleware applies slicing algorithm which splits the file in to four segments. Then these Segments are encrypted by encryption algorithm. These segments are then further uploaded to each of the cloud storage by accessing their APIs. The user is asked to authorize by their account in the cloud storage, to upload the file and the user‟s access token are saved into the user database. Then when the user want to download the file the middleware fetches the database and find the segments on the cloud storage and download the segment using API call. The downloaded segments are decrypted and then merged by the merging algorithm to recreate the original file and return it to the user.

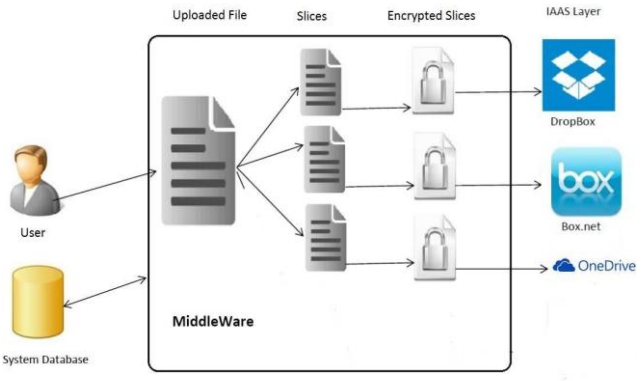

Figure 1: Architecture

4. Conclusion

Thus we can give more security over and above the encryption for the data on the cloud storage. We can use multiple public cloud storage providers for distributing the

Licensed Under Creative Commons Attribution CC BY

different segments of files on different cloud storages. Then it becomes difficult to recover files using any one of these cloud storages.

References

[1] Richard K. Lomotey and Ralph Deters, "Reliable Consumption of Web Services in a Mobile-Cloud Ecosystem Using REST" 2013 IEEE seventh International Symposium on Service Oriented System Engineering. Caves, Multinational Enterprise and Economic Analysis, Cambridge University Press, Cambridge, 1982. (book style)
[2] Richard K. Lomotey and Ralph Deters, "SaaS Authentication Middleware for Mobile Consumers of IaaS Cloud" 2013 IEEE ninth world congress on Services.H.H. Crokell, "Specialization and International Competitiveness," in Managing the Multinational Subsidiary, H. Etemad and L. S, Sulude (eds.), Croom Helm, London, 1986. (book chapter style)
[3] Simone Cirani, Marco Picone, Pietro Gonizzi, Luca Veltri, and Gianluigi Ferrari, Senior Member, IEEE,"IoT-OAS: An OAuth-Based Authorization Service Architecture for Secure Services in IoT Scenarios" IEEE SENSORS JOURNAL, VOL. 15, NO. 2, FEB-RUARY 2015
[4] Bharath Balasubramanian1, Tian Lan2, and Mung Chiang1, "SAP Similarity Aware Partitioning for Efficient Cloud Storage: IEEE Conference on Computer
[5] Zhangjie Fu, Member, IEEE, Xinyue Cao, Jin Wang, Xingming Sun ,"Secure storage of Data in Cloud Computing" 2014 inter-national Conference on Intelligent Information Hiding and multimedia signal Processing.

Author Profile

Sandeep Nehe received the B.E. degrees in Computer Engineering from Pune University and he is doing his post-graduation from AVCOE, Pune University.

Prof. M.B. Vaidya has completed his BE and Post Graduation in Computer Engineering . He is currently working as professor at AVCOE sangamner , Pune University.

Licensed Under Creative Commons Attribution CC BY