Comparative study for Protection of Data Stored in Cloud Computing Using CSE, DACESM and S^{3}DCE

Nisha[*], Satvender

Department of Computer Science and Engineering
BRCM College of Engineering & Technology, Bahal, Bhiwani, INDIA.
[*]Corresponding Author email: rngnisha@gmail.com

Abstract

Storage on the cloud is accessible from any of assessable devices, anywhere, at any time. In terms of login, one needs to enter personal information such an email address and password to access the computer. Different cloud providers could have a variety of terms of service, privacy policies, and login methods. Due to two research difficulties, it is still unclear how to properly develop useable privacy-preserving cloud systems to manage sensitive data safely. The paper presents a comparison of Data Security for Cloud Environments with Semi-Trusted Third Parties (DACSE), Data Security for Cloud Environments with Scheduled Key Managers (DACESM), and Secure Storing and Sharing of Data in the Cloud Are All Compared in This Research (S^{3}DCE)

Keywords: *Cloud computing, DACSE, DACESM, S^{3}DCE.*

1. Introduction

Computer can be anything like desktop, laptop, smart phone or notepad. Cloud can be accessed in terms of storage anywhere, anytime from any of these devices. As for as login is concerned, one must need to provide personal credentials such as email addresses and password to get access to the computer. Off-site data storage is important because the data has been stored in provider's server and not been saved on your own computer. Uploading is transferring data from your computer to another larger computer system (The Cloud) and downloading is the receiving data from cloud to your computer. Table 1 shows the pros and cons of the off-site data storage. Table 2 differentiates the data storage benefitsof both on-premise and cloud computing.

The term "cloud" originates from the telecommunications world of the 1990s, when providers began using virtual private network (VPN) services for data communication. VPNs maintained the same bandwidth as fixed networks with considerably less cost: these networks supported dynamic routing, which allowed for a balanced utilization across the network and an increase in bandwidth efficiency, and led to the coining of the term "Tele-com Cloud" [1]. Cloud computing's premise is very similar in that it provides a virtual computing environment that's dynamically allocated to meet user needs.

Table 1 Pros and Cons of off-site data storage

Benefits	**Drawbacks**
Saves storage space	Potential security threats
Improves disaster recovery	Requires internet connection
Increases collaboration	Terms of agreement

Table 2 Difference between On-premise and Cloud computing

On-premise	**Cloud Computing**
Lack of flexibility	High flexibility
No automatic updates	Automatic software updates
Less collaboration	Teams who collaborate from widespread locations
Data cannot be accessed remotely	Data can be accessed and shared anywhere over the internet
Takes longer Implementation time	Rapid Implementation

2. Cloud Computing

Cloud Computing (CC) is associated with a new standard for the provision of computing infrastructure. This paradigm shifts the location of the infrastructure to the network that reduces the cost associated with the management of both hardware and software resources. The cloud is drawing the attention from the Information and Communication Technology (ICT) community [2]. It provides the set of services with common characteristics provided by important industry players. However, some of the existing technologies of cloud concept draws on virtualization, utility computing or distributed computing. CC has remained as a computing standard in maintaining the resources that are residing in some other location [3].

CC architecture is comprised of two parts: Front end and Back end which are connected through Internet as shown in Figure 1.

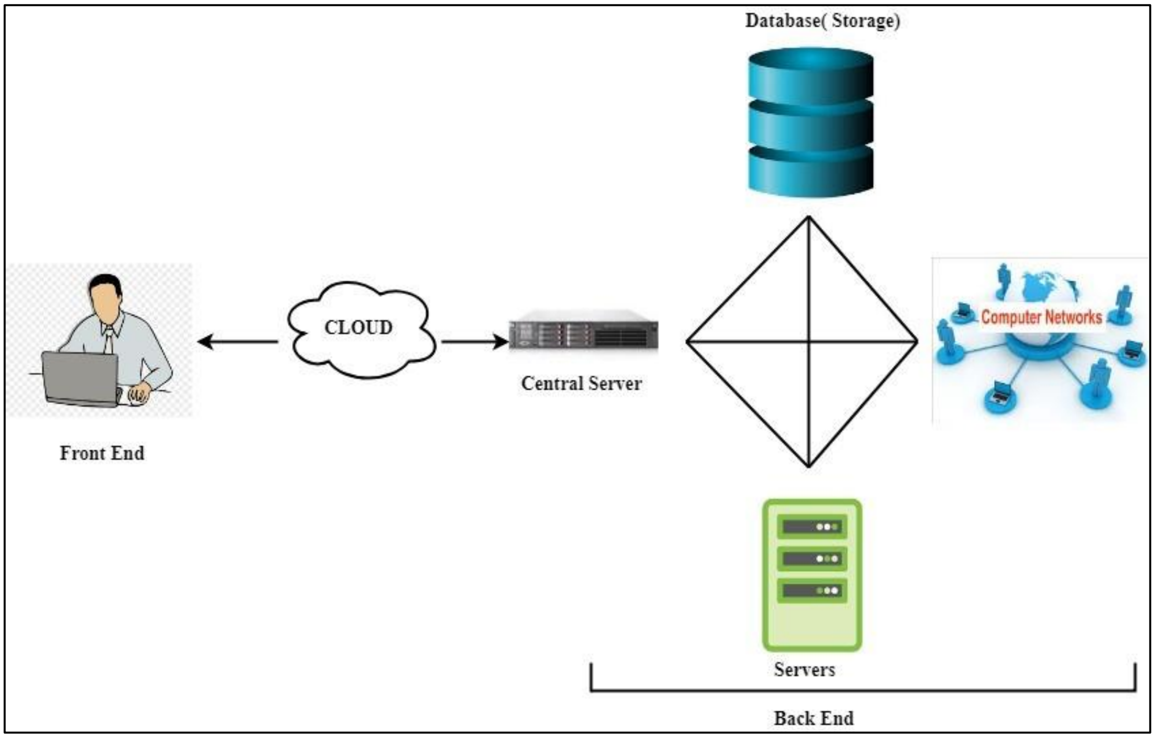

Fig. 1. Cloud Architecture

3. DSCESM: Data Security for Cloud Environment with Scheduled Key Managers

Off-site data storage is an application of cloud that relieves the customers from focusing on data storage system. So, the current organizations/industries are moving their records to cloud, instead of storing then on-premises by considering it as cost effective and of less management. However, the query arises for the security of records that are stored on off-premises i.e cloud. Providing the data security in cloud is much more complicated compared to the traditional information system [4-6].

The current industries are moving their records to cloud, instead of storing the data and records on premises, so that it is cost effective. But the query arises for the security of records stored in the off premises. There are many offers by third party cloud providers, but data leakage may occur due to some attack or virus. So we proposed the securitysystem that provides the data security to the cloud. The proposed system has 4 modules i.e., the client, the cloud, the scheduler and the Duplicate Finder as shown in Figure 2. The client is the one who wants to upload thefile to the cloud. So the client generates a secret random key and encrypts the file. Thenthe client approaches the scheduler to send the keys.

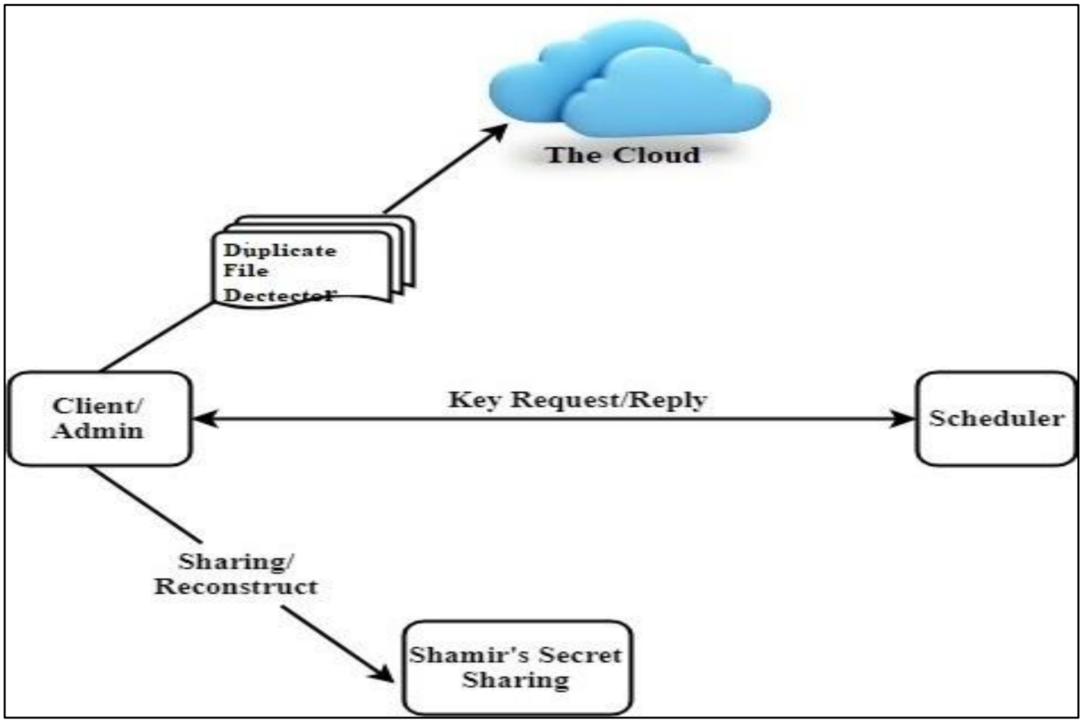

Fig. 2. DSCESM Architecture

The scheduler intern contacts the key manager and public keys are sent to the client. After receiving the number of keys from scheduler, the client divides the secret random key into number of shares by sending it tothe shamir’s secret sharing scheme [5-8] and obtain the shares as the public keys. These public keys are used to encrypt the shares and these encrypted keys are stored along withthe encrypted files. While uploading the file, the client checks whether the file is already being stored in the cloud through Duplicate File Detector. This is done by comparing the hash value of the uploading file to the hash value of the files that is already being stored. If any two file has the same value, then file will not be uploaded. If there is no similar hashvalue then the file will be uploaded.

Whenever the client wants the file, requests the cloud and gets back both the encrypted keys and file. Then the client goes to scheduler to get the private keys from the same key managers. These keys are used to decrypt the number of shares of the secret random keyand intern let it into shamir’s secret sharing scheme [6-7] to get back the secret randomkey. Later this secret random key is used to decrypt the file. By this way the client getsback the original file.

While uploading, the client checks for the file, if any duplicate file is already existing in thecloud. If the file exists, the cloud discards the file, if not the file is uploaded. Algorithm 1 describes how the duplicate file is detected while uploading. Initially, the hash value of theuploading file is calculated, and retrieve the hash values of all the files which are already in cloud. If two values are similar, then uploading file already exists so it will be not be uploaded. If no similar values, then the file will be uploaded to the cloud.

Algorithm 1 Duplicate File Detector

```
Input: File in the cloud and uploading file.
Output: Duplicate file or not.
begin
Step 1: Get the hash value of the selected file for uploading to thecloud i.e., $HV_1$.
Step 2: Get the list of the files present in the cloud.
Step 3: for each file in the list
Retrieve the hash value of file ($HV_2$)
Compare this $HV_2$ with $HV_1$if ($HV_1$ != $HV_2$)
Upload the fileelse
Not upload the fileend if
end for
end
```

4. S^3DCE: Secure Storing and Sharing of Data in Cloud Environment using User Phrase

Data centers [11] plays an important role in storing the user's data and access them through the internet. This greatly reduces the users cost on spending on the other storage devices like flash-drives, pen-drives and hard disks. The cloud service providers works on pay-as-you-use model where the providers offers the storage space so that the users can scale up/down the storage space based on their requirement. These data centers does not belong to the users, instead they are owned by the cloud service provider.

The cloud storage service providers are Google, Amazon, and Dropbox etc. The user hire some space in data center where data can be stored and accessed through internet. Since data centers are away from user's physical connection, the data can be accessed in any geographical region where internet is available. The cloud service providers make multiple copies of these data and stores in the multiple data centers, which are geographicallyapart. If one data center goes down, then the user can get their data from other data center from where the copy is stored. This is an important benefit offered by the cloud known as "*Disaster Recovery*". It is also cost benefit because the user need not to concentrate on the storage space management. So, Storage-as-a-Service is widely used among the cloud users. Even though the cloud acquires a good amount of benefits, it is still considered as third party. Therefore, it is better to protect the data before uploading it into the cloud. To protect the data, the approach used until now are the traditional encryption techniquesthat encrypts the raw data into an unreadable format at the user end and stores safely atthe cloud end. So even the attacker attacks the data center, it is impossible to read thedata unless he/she gets information used to encrypt the data.

The interactions are considered in three scenario between the user, the cloud and the owner:

1. Single Interaction.
2. One-to-One Interaction.
3. One-to-Many Interaction.

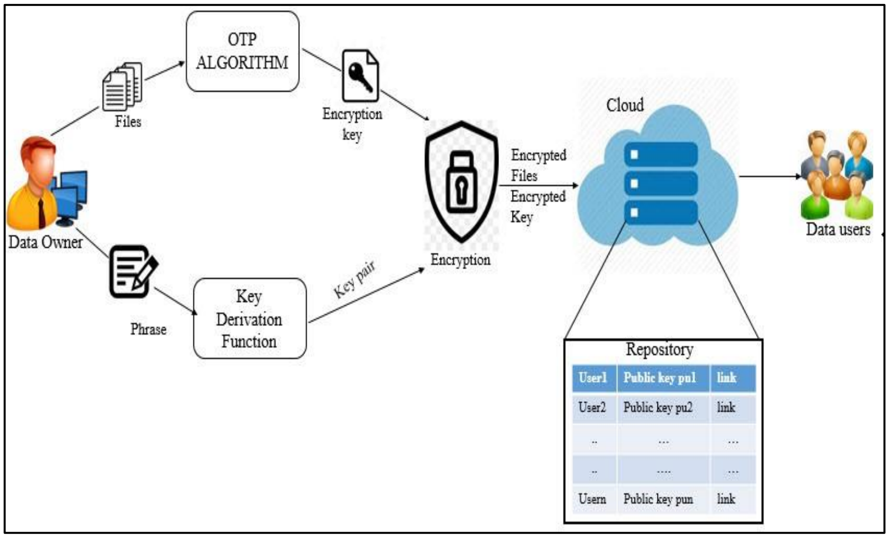

Fig. 3. S^3DCE Architecture

Algorithm 2 :One-Time Key Algorithm

```
Input: Number of digits required $K_i$.
Output: Encryption key of specified length.
begin
    Step 1:: Initialize all the characters to $W = W_0, W_1, ...., W_n$ i.e., a-z, A-Z, 0-9and Special Characters.
    Step 2:: Select a random index $i$
    Step 3:: Randomly select the number of digits required from $W$.
    Step 4:: Concatenate the digits.
    Step 5:: Shuffle the digits.
    Step 6:: The Encryption key $K_i$ with specified number of digits is generated
end
```

The performance is analysed in terms of time required to generate keys, total time to upload and download the encrypted file and encrypted secret key. The file size varies from 10kb to 1000kb that is used to analyse the performance. Figure 4 shows the time taken to upload the file. As the file size increases, the time takento upload the file also increases. By eliminating the key manager and shamir's concept, the proposed method S^{3}DCE takes much less time compared to the existing system DaSCE and DSCESM. Thus the cost incurred in storing is reduced. The uploading time reducesby 88% in S^{3}DCE compared to DaSCE and 61% in S^{3}DCE compared to DSCESM. Thetotal time taken to download the file from the cloud includes retrieving both encrypted fileand the encrypted key from the link, generating the private key, decrypting the encryptedsecret key, decrypting the file using the secret key.

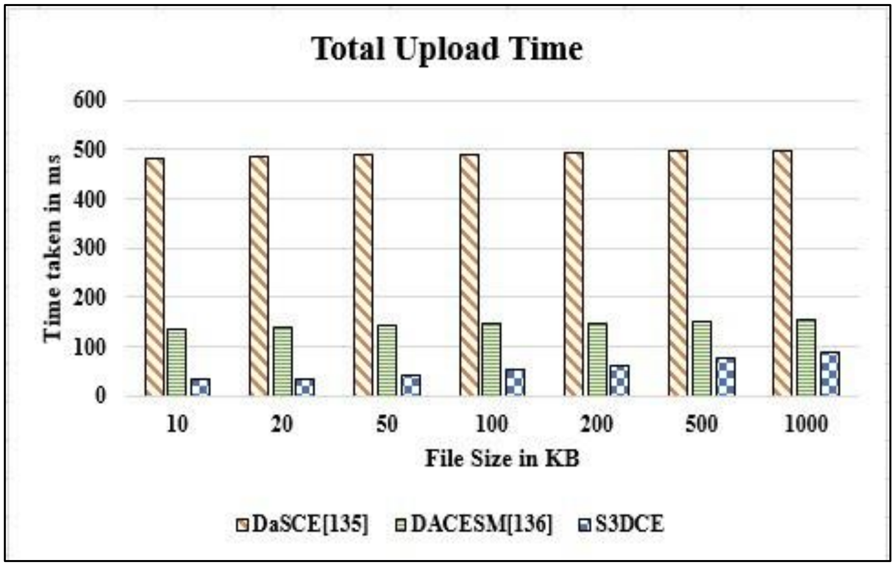

Fig. 4. Comparison of Total Upload Time of DACSE, DACESM and S^{3}DCE

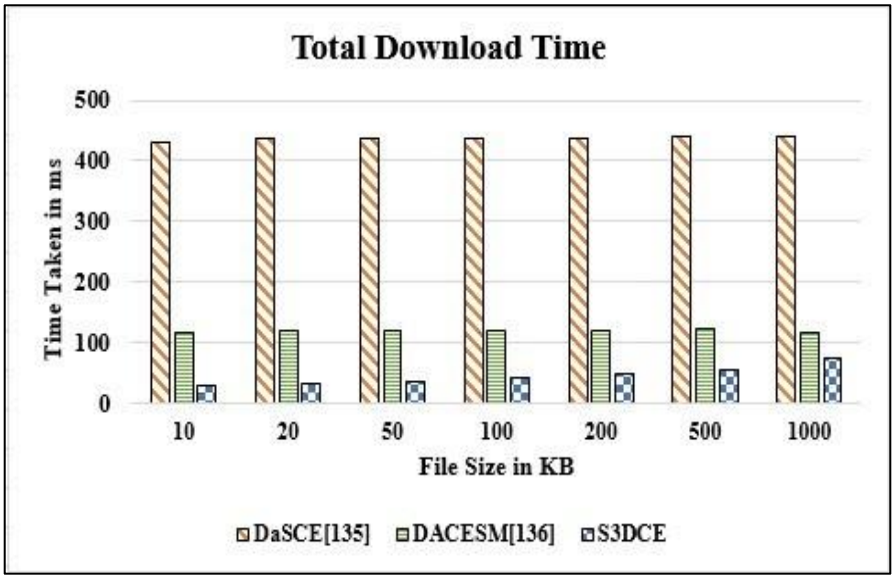

Fig. 5. Comparison of Total Download Time of DACSE, DACESM and S^{3}DCE

Figure 5 shows the time taken to download the file. As the file size increases, the time taken to download the file also increases. By eliminating the key manager and shamir's concept, the proposed method S^{3}DCE takes much less time compared to the existing sys- tem DaSCE and DSCESM. The download time reduces by 89% in S^{3}DCE compared toDaSCE and 61 % in S^{3}DCE compared to DSCESM.

5. Conclusions

A new security model "DSCESM" is proposed for cloud storage that reduces waiting time and avoids using extra space in cloud. Shamir's scheme is used in managing the keys. Scheduler is introduced to manage the task of key managers and assigns the tasks basedon workload that reduces the waiting time of the client. Deletion of files are assured basedon the policy of the client related to the file. The space is also saved by not uploading the repeated file to the cloud with the help of Duplicate File Detector.

S^{3}DCE is a security method that eliminates the concept of key manager and the shamir's concept. The OTK key generation algorithm is used to generate encryption key and it is more efficient compared to the random key generator. The ECC key pair generator provide the same key strength with smaller key size and more secure compared to RSA. This feature of ECC is very appealing with limited storage, processing power and reduced computational requirements. The performance is analysed in terms of time with respect to key generation, total upload time and total download time. It is observed that the proposed method is more efficient as compared to the existing system DACSE and DACESM.

References

[1] Loki M. Kaufman, “Data Security in the World of Cloud Computing”, IEEE Security and Privacy Magazine, vol. 7, no. 4, pp. 61-64, 2009.

[2] Luis M. Vaquero, Luis Rodero-Merino, Juan Caceres, and Maik Lindner, “A Break in the Clouds: Towards a Cloud Definition”, ACM SIGCOMM on Computer Communication Review, vol. 39, no. 1, pp. 50-55, January 2009.

[3] Jeevitha B K, Thriveni J, and Venugopal K R, “Data Storage Security and Privacy in Cloud Computing: A Comprehensive Survey”, International Journal on Computer Applications, vol. 156, no. 12, pp. 16-27, December 2016.

[4] Kire Jakimoski, “Security Techniques for Data Protection in Cloud Computing”, International Journal of Grid and Distributed Computing, vol. 9, no. 1, pp. 49-56, 2016.

[5] Mitsuru Ito, Akira Saito and Takao Nishizcki, “Secret Sharing Scheme Realizing General Access Structure”, Lecture Notes on Fundamental Electronic Science, vol. 72, issue. 9, pp. 56-64, February 2010.

[6] Mazhar Ali, Saif U. R. Malik, and Samee U. Khan, “DaSCE: Data Security for Cloud Environment with Semi-Trusted Third Party”, IEEE Transactions on Cloud Computing, vol. 5, issue. 4, pp. 642 - 655, October-December 2017.

[7] Jeevitha B K, Sindhura D, Thriveni J, and Venugopal K R, “DSCESM: Data Security for Cloud Environment with Scheduled Key Managers”, International Conference on Advances in Electronics, Electrical and Computational Intelligence, May 2019.

[8] Claudio Fiandrino, Dzmitry Kliazovich, Pascal Bouvry and Albert Y. Zomaya, “Performance Metrics for Data Center Communication Systems”, IEEE 8th International Conference on Cloud Computing, pp. 98-105, 2015.

[9] Francis Bloch, Matthew O. Jackson, and Pietro Tebaldi, “Centrality Measures in Networks”, Elsevier SSNR Publishing, June 2019.

[10] Mohammad Samadi Gharajeh, “A Dynamic Replication Mechanism in Data Grid based on a Weighted Priority-based Scheme”, i-manager’s Journal on Cloud Computing, vol. 6, no. 1, pp. 9-13, January - June 2019.

[11] Mazhar Ali, Kashif Bilal, Samee U. Khan, Bharadwaj Veeravalli, Keqin Li, and Albert Y. Zomaya, “DROPS: Division and Replication of Data in Cloud for Optimal Performance and Security”, IEEE Transaction on Cloud Computing, vol. 6, no. 2, pp. 303-315, April-June 2018.