**Cloud Security based on Data Fragmentation and Improved Encryption for Optimal Performance**

**Anuj Kumar Gupta**

Professor & Head, CSE, Chandigarh Group of Colleges, Mohali, PB, India

**Abstract**

In adoption of cloud computing technology, the data security of the customer data is a prime research in these days. The user’s get access of the cloud resources that are hosted over the internet that can be hijacked by the attacker. In a one example, the access of the virtual machine (VM) as dedicated resource is given to the end user by the CSP and when the end user is accessing these resources through a internet connected PC, there is a possibility of security violation by the attacker and he can take full control of the data. Most of the intrusion detection and prevention system implemented over the cloud infrastructures is being rule based so called signatures and therefore they are only capable to detect the known threats. There are research challenges such as a) only using these traditional security mechanims may lead extra overhead over the cloud network b) hosting the data over the public internet may make the data directly accessible as end user is not aware about the technology. This research work addressed the security problem by a) selection of optimal data security algorithm b) data fragmentation and distribution of data blocks over the multiple cloud nodes that make difficult to guess about the actual data and its location. The major advantage of the data fragmentation in cloud network is to anonymize the originality of the data to make more harden to guess about the original data by the attacker and he will not be able to identify the real location of the data hosted over the cloud nodes. Then by applying another layer of security in a way of optimal encryption algorithm with strong key size make it worst for the attacker to find the actual piece of data. For simulation purposes, Microsoft .NET Framework using Visual Studio is being used in the implementation and its in-built cryptography library. The performance readings have been taken in an iterative manner by applying and selection of optimal algorithm for data encryption.

**Keywords:** Cloud Computing, File segmentation, Cryptography, Virtualization, Database

**1 Introduction**

Cloud Computing (CC) is an developing technology that has connection with other technology such as Grid Computing (GC) and other domains such as Distributed Computing, Cluster Computing etc [1-2]. The major aim of both GC and CC is primarily to achive full virtualization. However both the technology are synonymous to each but there are technological difference among them. The Grid Computing major objective is to achive the optimized computing capacity [3] whereas in case of Cloud Computing, aim is to provide a way to handle the need of an organizations by enabling the dynamic and scalable infrastructure to work with [4]. The Cloud Computing introduced a new paradigm of technology that help its users to develop applications and store their data somewhere on the internet and they can access their resources from anywhere and anytime by just using the internet connectivity [5]. Based on the need and requirements of the customer, Cloud Computing enable easy and modifiable services to access their work with cloud applications. The services in the form of platform, infrastructures are provided to the customer for designing applications, for storage their data and to build their own applications by using the cloud infrastructure to fulfill their route task. These resources are hosted by the CSP over the dedicated resources as a service model. A user can buy these resources by pay-basis model and merely having internet connectivity and PC [8-9].

Now the research question is by hosting and accessing these resources over the public internet what the security measurements are installed to protect the user’s data[17-18]. If improper security measures are placed at the side of CSP that may make the complete loss of the user’s data or we can say there are higher risk involved when the customer records are strored in internet instead of user’s local repository [6]. An intruder gain complete access of the user’s data transmitted over the connection between end user and remote location by just performing the man-in-middle attacks[19]. The intruder may also take the complete access of the user’s data and their sensitive information hosted on cloud infrastructure [6]. The users of cloud services are diversified that may include the expert and the user who don’t have any knowledge of technology, thereby making their data is on huge risk in such infrastructures [10-11].

From the above discussions, it is summarized that security is a prime concern with the growing use of Cloud Computing Technology and that should be address to provide the data protection of the end users [8]. In this research work, existing security mechanism placed in cloud computing is presented and a solution to protect the customer data fragmentation and its distribution by combining the optimal encryption technique is presented[15].

**2 DATA SECURITY IN CLOUD COMPUTING**

The growth and scalable model of CC comes with lots of security challenges to protect the customer’s data. Even though the technology is new and provide a platform to access the applications hosted over the internet easily, still there are challenges and issues to adopt the technology by the end user. One of such challenges that are on top priority by service provider is the security pertaining to the infrastructure and protects the user’s data hosted on such infrastructure [24]. Even though a company use to have top class security installed over it and does update their security policies from time to time, there are still the issues to be addressed to protect the customer data. In this regard, through this detailed study we propose security challenges and their solutions pertain to cloud infrastructures.

Cloud Resources

End User

Internet

Attacker

Attacker

**Figure 1: Attacker can intrude the cloud network**

When the end user access the applications and data hosted over the Cloud Network with public internet connectivity, the intruder can intrude the connection and can take the control of the user’s resources that are hosted over the internet connected infrastructures. The major security risks involved in such infrastructures are data loss and leakage [20-23].

**3**. **PROPOSED SYSTEM**

Following diagram depict the system design of developed framework during this research implementations. The registered user submit the file as data inputs and then file segmentation module is to divide the main file into multiple pieces as per similar way of Google File System (GFS) [12]. The segmented files are encrypted using optimal symmetric key encryption algorithm [13]. Here the RSA[14] encryption algorithm is used and performances are measures through changing the key size to get the best optimal results. The in-built cryptography algorithm of Microsoft .NET framework [16] is implemented and applied on the fragmented data. The main benefit of this research work is to confuse the attacker about the actual data where it is located and not able to get the actual data because it is located over the different cloud nodes and another layer of security is fragment data is encrypted.



**Figure 1: System Design Flow**

***Workflow of the implemented software code:***

Following figure 2 depict the work flow of the implementation of the software code. The developed application is user driven application in which a user can submit a file to the cloud node resources created in a software program. If the user is registered, he will get the user credentials to upload the files on the cloud nodes. The registered user is authorized to split the main files into multiple chunks and distribute them over the cloud nodes. The pieces of the main files are encrypted with encryption algorithms and distribute them over the cloud nodes. The metadata server is inbuilt in software program to maintain the indexing of the files and storage server is a database repository of SQL server to maintain the user’s logs.

Data owner

Meta data server

Storage server

**Figure 2: Workflow diagram of implemented code**

Figure 2 describe the software modules in the software program developed in Microsoft .NET framework. There are four major modules:

* User registration module
* Upload files modules
* File splitting/data fragmentation module
* Data encryption and decryption module

The user registration module is developed to give the provision to the user to get registered and get the access of the cloud resources as nodes. As mentioned above, the registered user is able to access the cloud nodes and can submit the files into cloud nodes. The submitted main is then divided into multiple data fragments to distribute over the multiple cloud nodes by file splitting program. It is a program which takes the main files and divides them into multiple pieces of same size equal to 7.

Here the numbers of file segments are fixed in a software program that can be changed by the user. Next module is data encryption module that encrypts the segmented files.

Upload files

File id

File name

Uploaded data

Download data

User account

Name

User name

Password

Address

Email id

Create account ()

Generate keys ()

Search split files

File id

File name

User data

Duplicate block added for security

Download files ()

Show duplicates ()

Upload files

File id

File name

Uploaded data to split

Generate duplicate blocks

Download data

Send to cloud ()

View file details ()

Secure

User data

User id

Encrypted data

Decrypted data

Public key

Encryption ()

Decryption ()

**Figure 3: Software modules in the code.**

**4. EXPERIMENTAL RESULTS**

Here the experimental results in the form of system time taken by applying a encryption algorithms with different key size applied on file segments is presented. After the file segmentation process, the multiple chunks of files are encrypted with optimal data encryption algorithm and distributed over the cloud nodes. As discussed above, the main advantage of fragmentation and distribution is to hide the originality of the data so that the attacker will not be able to guess about the actual data and the location of cloud node where it is located. Then by applying another layer of security in the form of data encryption algorithm with strong key size make the guess more difficult to protect the data over the cloud node. For simulation purposes, we have used the Microsoft .NET Framework using Visual Studio 2008 and its in-built cryptography library. The results measurement readings have been taken thoroughly by applying the data encryption algorithm of different key size.

**Table 1: System time taken as Performance Matrix**

|  |  |  |
| --- | --- | --- |
| **Performance Metrics** | **Attribute** | **Descriptions** |
|  | Treal | Total execution time taken by a process to execute |
| Tuser | The time where the programming was running in user space |
| Tsys | It is the time spent in CPU execution |

**Use Case 1: Data Fragmentation and distribution over the cloud after applying AES encryption with different key size.**

**Table 2 system time taken of data encryption w.r.t number of fragments**

**Result Readings:**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **System Time Taken** | No of segments=7node0000000to node0000006 | No of segments=6node000000to node000005 | No of segments=5node00000to node00004 | No of Segments=4node0000tonode0003 | No of segments=3node000tonode002 | No of segments=2node00tonode01 |
| **Treal** | 467.799 | 444.343 | 421.745 | 453.304 | 445.902 | 414.16 |
| **Tuser** | 0.032 | 0.068 | 0.064 | 0.052 | 0.032 | 0.052 |
| **Tsys** | 16.8 | 16.96 | 16.996 | 16.984 | 16.996 | 17.08 |
|  | AES-256 for 7 Segments | AES-256 for 6 Segments | AES-256 for 5 segments | AES-256 for 4 segments | AES-256 for 3 segments | AES-256 for 2 segments |
| **Treal** | 733.287 | 757.798 | 714.544 | 568.445 | 454.813 | 320.761 |
| **Tuser** | 66.6 | 60.736 | 53.496 | 43.284 | 36.688 | 28.876 |
| **Tsys** | 22.604 | 20.192 | 19.344 | 18.584 | 17.812 | 16.9 |
|  | AES-192 for 7 segments | AES-192 for 6 segments | AES-192 for 5 segments | AES-192 for 4 segments | AES-192 for 3 segments | AES-192 for 2 segments |
| **Treal** | 896.628 | 939.465 | 756.698 | 587.498 | 468.354 | 287.962 |
| **Tuser** | 64.552 | 57.496 | 49.24 | 41.972 | 34.656 | 27.584 |
| **Tsys** | 21.224 | 19.864 | 18.944 | 18.688 | 18.096 | 16.556 |
|  | AES-128 for 7 segments | AES-128 for 6 segments | AES-128 for 5 segments | AES-128 for 4 segments | AES-128 for 3 segments | AES-128 for 2 segments |
| **Treal** | 644.223 | 850.524 | 740.413 | 506.765 | 452.951 | 343.739 |
| **Tuser** | 58.56 | 50.74 | 46.104 | 40.044 | 33.656 | 26.548 |
| **Tsys** | 20.74 | 20.764 | 20.204 | 18.552 | 17.248 | 16.588 |

Table 2 describe the readings of system time taken of AES algorithm with different key cipher block corresponding to each fragments. For example Treal time taken by AES-256 algorithm in seven number of file segments is 467.799 seconds, in a similar for two file segmentation AES-256 algorithm time taken is 414.16 seconds. Then each fragmented files are encrypted with AES with different key size. This table clearly indicate that more the number of segments, more will be the system time taken, but in the following results it is clearly indicated that in AES-128 bit algorithm is performing better in overall system time taken whereas AES-192 is performing better in system user space. It can be concluded as system time taken by the data encryption algorithm also depends upon the variant of application and its CPU clock requirements.

**Result Stats 1: System time taken of AES-256 w.r.t. each fragmentation**

****

****

**Figure 4: AES-256 performance with data fragments**

**Result Stats 2: System time taken of AES-192 w.r.t. each fragmentation**

****

****

****

**Figure 5: AES-192 performance with data fragments**

**Result Stats 3: System time taken of AES-128 w.r.t. each fragmentation**

****

****

**Figure 6: AES-128 performance with data fragments**

**Use Case 2: Comparative analysis of AES with different key size w.r.t. each segmented file**



**Figure 7: Comparison of different AES with number of fragments**

**Use Case 3: Data Loss in Segmentation and replication over the cloud Nodes:**

In this use case, the percentage of data loss is depicted by applying the data fragmentation and its replication over the cloud nodes. The stats have been taken by splitting the big file into smaller pieces of equal size and then applying the data encryption over it. The data encryption algorithm is adding the data padding along with the actual size of data.

**Table 3: Effect of data fragmentation on actual data size**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **No of Segments** | **Segmented and encrypted data size (in Mbytes)** | **Combined Data Size ( In Mbytes)** | **Data Loss** | **%Data Loss** |
| 7 | 6033 | 5700 | 333 | 5.519 |
| 6 | 5266 | 5000 | 300 | 5.051 |
| 5 | 4419 | 4200 | 300 | 4.956 |
| 4 | 3428 | 3275 | 200 | 4.463 |
| 3 | 2805 | 2700 | 105 | 3.743 |
| 2 | 1998 | 1930 | 68 | 3.403 |

**Table 4: Percentage of data loss**

|  |  |
| --- | --- |
| **Segments** | **% of data loss** |
| No of Segments=7 | 5.519 |
| No of Segments=6 | 5.051 |
| No of Segments=5 | 4.956 |
| No of Segments=4 | 4.463 |
| No of Segments=3 | 3.743 |
| No of Segments=2 | 3.403 |



**Figure 8: Percentage distribution of data loss w.r.t each fragmentation**

As depicted in figure 8, more the number of fragmented data, more distribution of data loss. But the comparative data loss is smaller in more number of segmentation compare to small number of segmentation. By performing this research implementation, it is advised to set the count of data segmentation which is hard (data segmentation=7) coded in our software code to get more efficient results.

**Conclusion and future work**

The cloud computing technology has grown exponentially and very well adopted by the end users that make the technical and cultural shift from localization to globalization with inclusion of internet connected resources. A customer can access their data hosted anywhere on the global world by merely having internet connectivity. Cloud resources are maintained and controlled by the third party as Cloud Service Providers (CSP) and these pay less attention to secure the user’s data. The cloud service provider can take care of infrastructure security and protect the network those are under their control but end user who is not so technical guy has a belief that data hosted over the cloud network is full proof secured. Normally the data hosted on cloud resources are more vulnerable to attacks than the data hosted on a PC locally because data flow is always through a public internet which is always a insecure infrastructure and always prone to attacks. Moreover the volume of users of cloud computing is increasing in a exponential manner which definitely make the internet based environment under the reach of end user but it also lead to threats pertains to these resources.

In research, many authors had proposed multiple approaches to ensure the data security and protection of end users. Researchers also depicted that by only applying traditional cryptographic algorithms may not be such useful in the context of cloud computing because it increase overhead to the cloud infrastructures, moreover the security is wholly depend upon the type of encryption algorithm and key used in that. To avoid such algorithms, the researchers explored another phenomenon of data security of the user hosted on CSP. It includes the data fragmentation and its dispersal over the multiple cloud nodes.

In this research thesis, the need and requirements of protection of user’s data is presented by addressing the security problems in cloud infrastructures. Then solution is presented to solve this research problem by incorporation of novel data fragmentation techniques and distribution of data files over the cloud nodes along with advance and optimal data encryption algorithm. For the protection of end user data, we propose a schematic algorithmic novel combined approach for data security which will be involved a lightweight encryption with data fragmentation and replication techniques to ensure the data protection.

The core benefits of the proposed approach are:

* It explored the benefits and usage of data fragmentation techniques to protect the user’s data
* It incorporates the optimal advance data encryption algorithm prior to disbursal of user’s segmented files.
* Scalable implementation is performed which provide the flexibility of more pieces of files, incorporation of more optimized data encryption algorithm with different key factor.

The developed system is user-driven as user can submit the data to be hosted over the cloud. The user registration module is to give the access to the user so that a user can sign up for access the cloud nodes as resources. The registered user submit the file as data inputs and then file segmentation module is to divide the main file into multiple pieces as per similar way of Google File System (GFS). The segmented files are encrypted using optimal symmetric key encryption algorithm. Here the RSA encryption algorithm is used and performances are measures through changing the key size to get the best optimal results. The in-built cryptography algorithm of Microsoft .NET framework is implemented and applied on the fragmented data. The main benefit of this research work is to confuse the attacker about the actual data where it is located and not able to get the actual data because it is located over the different cloud nodes and another layer of security is fragment data is encrypted. The performance benchmarking is listed and presented through system time taken in the form of Treal, Tsys and Tuser after enabling the data fragmentation and data encryption processes on a system.

The developed system is tested on the limited simulated environment using Microsoft .NET framework and its in-build cryptography algorithm. The testing of AES data encryption algorithm is performed with different key size and data segmentations. The exploration and testing of more symmetric key encryption algorithm such as DES, Blowfish and more Asymmetric cryptography algorithm is left for future work.

**References**

1. Gupta, A. K., “Cloud Computing: Concepts and Challenges”, Asian Journal of Computer Science and Technology, ISSN: 2249-0701, 4(2): 27-30, July – December 2015.
2. I Foster et.al. Cloud computing and grid computing 360 degree compared. In Grid Computing Environments Workshop, 2008, IEEE, November 2008.
3. Talwar, B., Gupta, A. K., “A Survey on Cloud Computing Model Configuration Security Issues with Proposed Solutions”, International Journal of Exploration in Science and Technology, ISSN: 2456-1118, 1(2): 43-49, September 2016.
4. Ruoyu Wu, Gail-Joon Ahn, Hongxin Hu, and Mukesh Singhal. Information flow control in cloud computing. In Collaborative Computing: Networking, Applications and Worksharing (CollaborateCom), 2010 6th International Conference.
5. L.M. Vaquero, L. Rodero-Merino, J. Caceres, and M. Lindner, ACM Computer Communication review, 2008, a break in clouds
6. Shucheng Yu, Cong Wang, Kui Ren, and Wenjing Lou. Achieving secure, scalable, and fine-grained data access control in cloud computing. In IN- FOCOM, 2010 Proceedings IEEE, pages 1-9, 2010.
7. M.B. Mollah et.al. Next generation of computing through cloud computing technology. In 2012 25th IEEE CCECE, pages 1-6, May 2012.
8. K. Hwang and D. Li, "Trusted Cloud Computing with Secure Resources and Data Coloring," in IEEE Internet Computing, vol. 14, no. 5, pp. 14-22, Sept.-Oct. 2010.doi: 10.1109/MIC.2010.86
9. K. Hwang, G. Fox, and J. Dongarra, Distributed Sys­tems and Cloud Computing: Clusters, Grids/P2P, and Internet Clouds, Morgan Kaufmann, to appear, 2010.
10. K. Hwang, S. Kulkarni, and Y. Hu, “Cloud Security with Virtualized Defense and Reputation-Based Trust Management,” IEEE Int’l Conf. Dependable, Autonomic, and Secure Computing (DASC 09), IEEE CS Press, 2009.
11. J. Nick, “Journey to the Private Cloud: Security and Compliance,” tech. presentation, EMC, Tsinghua Univ., 25 May 2010.
12. <https://research.google.com/archive/gfs.html>
13. Guanrong Chen , Yaobin Mao, Charles K. Chui, A symmetric image encryption scheme based on 3D chaotic cat maps, Chaos, Solitons and Fractals 21 (2004) 749–761
14. Jawahar Thakur, Nagesh Kumar, DES, AES and Blowfish: Symmetric Key Cryptography Algorithms Simulation Based Performance Analysis, International Journal of Emerging Technology and Advanced Engineering Website: www.ijetae.com (ISSN 2250-2459, Volume 1, Issue 2, December 2011
15. Mazhar Ali, Kashif Bilal, Samee Khan, Bharadwaj Veeravalli, Keqin Li, Albert Zomaya, "DROPS: Division and Replication of Data in the Cloud for Optimal Performance and Security", IEEE Transactions on Cloud Computing, vol. , no. , pp. 0, 5555, doi:10.1109/TCC.2015.2400460
16. <https://www.microsoft.com/>
17. Keiko Hashizume, David G Rosado, Eduardo FernndezMedina, Eduardo B Fernandez,An analysis of security issues for cloud computing,Journal of Internet Services and Applications,2013.
18. Radhika Chavan, Prof. S.Y.Raut, A Review on Secure Methodology Fragmentation and Replication of data in cloud for Data Secrecy and Ideal Performance, International Journal of Advanced Research in Computer Engineering & Technology (IJARCET), Volume 4 Issue 11, November 2015
19. Ramesh Shahabadkar, S. Sai Satyanarayana Reddy, Chinthakunta Manjunath, Ugranada Channabasava, Krutika Ramesh Shahabadkar, Secure Framework of Authentication Mechanism over Cloud Environment, CSOC 2017: Software Engineering Trends and Techniques in Intelligent Systems pp 284-291
20. Harika, A.V., Haleema, P.K., Subalakshmi, R.J., Iyengar, N.Ch.S.N.: Quality based solution for adaptable and scalable access control in cloud computing. Int. J. Grid Distrib. Comput. 7(6), 137–148 (2014)
21. Sirohi, A., Shrivastava, V.: Implementing data storage in cloud computing with HMAC encryption algorithm to improve data security. Int. J. Adv. Res. Comput. Sci. Softw. Eng. 5(8), 678–684 (2015)
22. Oyeyinka, F.I., Omotosho, O.J.: A modified things role based access control model for securing utilities in cloud computing. Int. J. Innov. Res. Inf. Secur. (IJIRIS) 5(2), 21–25 (2015)
23. Talib, A.M.: Ensuring security, confidentiality and fine-grained data access control of cloud data storage implementation environment. J. Inf. Secur. 6, 118–130 (2015)
24. Supreet Kaur, Amanpreet Singh,security issues in cloud computing,  International Education and ResearchJournal, 2017.