IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661,p-ISSN: 2278-8727, Volume 17, Issue 1, Ver. V (Jan – Feb. 2015), PP 53-60
www.iosrjournals.org

Encryption Technique for a Trusted Cloud Computing Environment

Aishwarya Asesh
(School of Computing Science and Engineering, VIT University, India)

Abstract: *Cloud computing has reached a certain level of maturity which leads to a defined productive state. With varying amount of computing power present with everyone, it has become necessity of the hour to use cloud computing systems. It helps us to store our data within a virtual cloud structure. When we use the cloud storage mechanism, the computing power gets distributed rather than being centralised. The whole system uses the internet communication to allow linkage between client side and server side services/applications. The service providers may use the cloud platform as a web service platform or a data storage architecture. The freedom to use any device and location for cloud management is an added advantage for any user. Maintenance of such systems is also easy as installation of resources aren't required in each and every system which is using cloud services. But along with varying flexibility and multi tenancy in usage comes the question of reliability and security. As in public hosting, the client is totally unaware of the security strategies applied by the service provider, it creates a necessity for the end user to save the data from expected threats. One cannot totally rely on the quality of service (QoS) which is guaranteed by host servers. When we look at the security of data in the cloud computing, the vendor has to provide some assurance in service level agreements (SLA) to convince the customer on security factors. This paper describes a schema that ensures encryption of data using Advanced Encryption Standards. By doing so, the customer services can become quiet secured and thus can help in further enhancement of the cloud computing standards.*

Keywords: *AES, Cipher, Encryption, Rijndael, Security*

I. Introduction

Cipher is an algorithm used for encrypting and decrypting a message. It becomes difficult for a hacker if the data present in cloud is in encrypted form, as the data files or encrypted data blocks are useless for any person unless he knows the perfect method for decrypting it. Generally companies with critical data sets, encrypt the data using a proper cipher algorithm before sending it to the server. This procedure is considered to be the safest method for data security in clouds as even the service providers are not able to mess with the data they receive from client. Various cloud deployment and service models are described in the paper, so that the ideology of cloud computing can be clearly understood. Some real world issues and potential challenges to data security is then emphasized. Encryption algorithms used in earlier times like the Caesar, Vigenere, Playfair are discussed and their advantages, disadvantages are taken into account. The key point is to introduce a more secure and safe method or process which can strengthen the current system of cloud security services, so that the people using it may rely completely on them. Advanced Encryption Standards have been used for the encryption purposes in the proposed system. Rijndael being one of the safest algorithm used for encryption can result into increase in overall reliability of the cloud environment. The designed and proposed architecture can help to encrypt and decrypt the file at the user side as well as client side thus providing security to static as well as dynamic data.

II. Cloud deployment models and service models

Four different cloud deployment models are used currently along with 3 types of integration systems among them:

- **Public Clouds**

Very common type of cloud deployment model. Customers use the services offered by cloud service providers. Most of the companies are providing these services today such as sky drive, google drive and iCloud services. Customers have no idea about the infrastructure and working of the computing mechanism. Consumers can add data and retrieve the data at any moment when required. Security aspect is taken care by the service providers.

- **Private Clouds**

This type of cloud is mainly used by large companies and enterprises. People are given a completely private environment and they can use the security measures best suited to them. Disadvantages of this model is

that the cost of such deployment is too high. Mostly used in banks to provide private services to customers and employers.

- **Hybrid Clouds**

This type of deployment is basically used when we need both the private and public deployment model simultaneously. The security strategies are independent for both type of services. The cost effectiveness is less as both type of models are integrated in one system. Best example can be Amazon simple storage service.

- **Community Cloud**

More than one single infrastructure are used by this type of model. More than one organization can control single service deployment model. The control may be administered by more than a single provider also. Used when many organizations have a shared interest to use a single cloud model.

- **IaaS - Infrastructure as a service**

Cloud infrastructure services are used to maintain and monitor the cloud data, networking or network services. The requirements can be based on consumption of resources by the users.

- **PaaS - Platform as a service**

Cloud platform services are used in the development of applications and for providing cloud components in them. Framework for further development can be achieved using such services.

- **SaaS – Software as a service**

Cloud software services are used to manage third party software by the client side. Such applications can run directly using the web plugins, no downloads or such installations are required for using such services.

You can see the examples of different cloud models and their usage in Figure 1.

	SaaS	PaaS	IaaS
Private Cloud		Apprenda, Stackato	VMware, Hyper-V, OpenStack, CloudStack
Community Cloud		NYSE Capital Markets Community Platform	NYSE Capital Markets Community Platform
Pubic Cloud	Salesforce.com, QuickBooks online, Office 365	Google AppEngine, Microsoft Azure, VMware CloudFoundry.com	Amazon EC2, Rackspace
Hybrid Cloud		Custom CloudFoundry	Custom, Rackspace

Fig. 1

III. Security concerns

Due to the increasing number of cloud users per day, the amount of data stored by the service providers is also increasing rapidly. The major concern arises of data theft as all users are storing sensitive information and service providers are taking care of the security measures used to protect and safeguard the data. It becomes a major concern to know that which type of security is used by providers to protect what kind of data. Sometimes even the data usage pattern may give important information to the hackers or data thieves [1].

API changes are frequent from the service provider's end but they never inform or intimate the clients about the changes being made. Therefore the clients are not updated about the kinds of vulnerability their data may face in future. Customers are just using the data from time to time without any effective safeguard guaranteed. Along with this problem, people of different countries, use the cloud services at different rendering speeds, so many a time incomplete or corrupt data gets recorded while uploading to the server [2]. This creates a big problem whenever the customer tries to get back the data, as the server doesn't reply to the request in a proper manner.

In case of data access by some random person, the customer might never know that the data was stolen or leaked, as the service providers have the record of data access and they don't share such information with the users. The situation becomes worst when the data was being shared by multiple user [3]. If you are using a private cloud then there are less chances of such unalarmed data thefts.

Most of the public data service providers believe that encrypting the data maybe the possible way to protect the data. But there are many disadvantages of the process as encryption consumes a lot of processing power. Many a times encryption results in data getting corrupt because of lack of proper server response or service time. Even if the hacker knows the data access pattern, he may be able to decrypt the data. Cloud providers can also not be trusted to the very limit as they may sell the encrypted data to some other potential buyers. The best way to deal is by providing the data in encrypted form to the service providers so that they may not be able to mine or decrypt any useful information from those encrypted blocks of files.

Data security problems also differ by different countries and places. Such as in America the distribution of private information is considered safe and thus even big enterprises have transparency in the file and record sharing aspect whereas on the other side in Europe, people don't believe in data transparency and thus every bit of detail is kept hidden from the normal people [4].

IV. Encryption methods

Data is encrypted by the service provided when the customer/client uploads the data to the servers, using proper encryption methods they can protect the data to large extent as even if the data is stolen, hackers will not be able to extract any useful information from the encrypted data block [5]. Three of the famous encryption algorithms used are:

a) Caesar Cipher

Caesar's shift or Caesar's code is the most famous and widely used encryption algorithm used to code data into encrypted block of information. In this we replace the letter of alphabet in the original text to another letter three places down in the order, for example "ABCD" will be converted in to "DEFG". In Practical application it may be difficult to solve, but through brute force this can easily be broken as only 25 possibility of letters exist. The vulnerability increases as there is a fixed pattern of encoding, so the thief knows what to look for while decoding the whole text. The advantages are that the algorithm doesn't do runtime overhead while encrypting and decrypting. The amount of time consumed in encryption and decryption is less compared to other complex algorithms, so there are less chances of data getting corrupted while in the process. Can be best used when we need low to moderate amount of security and when the time efficiency matters more than the security aspect.

Example:
Text: ABCDEF
Encrypted text: DEFGHI

b) Vigenere Cipher

This cipher increases the level of security from Caesar cipher. Encryption is done using a sequence of different Caesar ciphers based on the letter of the keyword. The shift values are different in this type of encryption as many patterns of shifts are used to encrypt a text. A table is used to refer to the keyword description and thus it makes it easy to encrypt a text. The possibility of the code word being broken reduces. We make a table using the plaintext and keyword. Then we match the corresponding letters of both the Keyword and Plaintext to get a defined algorithm or set of rules for encryption.

Example:
Plaintext: KILLERBOY
Key: RATRATRAT
Cipher text: BIECEKSOR

Plaintext: HOWAREYOU
Key: BOY
Cipher text: ICUBFCZCS

Thus as you can see this has a very difficult chances of getting decrypted. When developed, this was considered indecipherable cipher. But if someone gets to refer the key table then it becomes easy to compute the original text message or data. So the risk factor for sensitive data storage using this method still remains unused.

c) Playfair Cipher

Playfair Cipher is also known as symmetric encryption method as generally matrices of 5x5 dimension is used in this kind of encryption. This encryption encrypts pairs of letters rather than a single unit and thus is much more complex than the Vigenere cipher. As frequency determination doesn't work for this method, so breaking or decrypting such ciphers is tougher. The probability of correct decryption increases to 600 instead of just 26 as in simple encryption methods. A matrix is formed using keyword and letters of the alphabet. Then the original text is broken in 2 letter pairs each. Then a particular encryption algorithm is used to encrypt the text.
Example:
Plaintext: THISISAMESSAGE
Key: PAPER
Cipher text: UG KN KN BJ FX OE FR

Plaintext: HOWAREYOU
Key: JOLLY
Cipher text: PC ZL SD AL VZ

The main problem while using such kinds of cipher algorithms is the time complexity involved in the process. As the time consumed is quiet high when the data to be encrypted is large.

V. Proposal

Chances of data theft increases as one tries to upload/send the data to the server or otherwise when client requests for the data from the service provider/server. Thus data security becomes essential while the data block is in transmission state. Even when the data is static, it is best suited to keep it in encrypted form, as the data leakage chances gets reduced by a large extent [6].

The following step plan is used in order to achieve this target of encryption and decryption (at a later stage) with the best ratio of security and time complexity:

a) Authorization Step:

Encryption key and other parameters need to be provisioned with dynamic security strategies. For this purpose we use Extensible Authentication Channel which acts like an extended version of normal authorization by giving secure transmission for encryption keys [7]. For general authentication of user we use a challenge handshake mechanism which prevents unwanted users from entering the data warehouse system.

b) Encryption Step:

For Encryption of data, the algorithm should be tested for high level of security. Advanced

Encryption Standard algorithm – Rijndael is chosen by the Nation Institute of Standards and Technology – United States. Being much more secured than DES and triple DES, this algorithm is a perfect balance between security standards, performance and efficiency [8]. Algorithm is based on standard symmetric key formulation. Encryption and Decryption is done using blocks of data. Iterative refinement is used in this algorithm as Rijndael uses a dynamic number of rounds based on the key or block size. The best part lies in the fact that for added security more rounds can be added at a later stage. So security standards can be increased based on the vulnerability of data [9]. The whole process is based on iterative block formation and operations are carried in different intermediate stages.

VI. Implementation

A. Authentication Channel

A secure mechanism is used to verify the authenticity of the user. When customer tries to get the data, the 'challenge' signal mechanism is used to bypass the user through the channel, thus a secured login is guaranteed.

Challenge handshake signal security is implemented using the following steps:

1. As soon as a client/customer wants some data, the mechanism creates and sends a 'challenge' signal to the client, thus verifying the data request.
2. A hash table is available with client, the client/customer replies back to the challenge signal with the correct hash value.
3. The protocol verifies/authenticates the value sent by the client ad thus accepts the request for required services or data pooling.

B. Encryption Algorithm – Advanced Encryption Standard

Rijndael supports key of variable length like 128 bits, 192 bits and 256 bits. This variation in key

selection has been recently been in use, so that the hacker may not be able to plunder over the fixed key sizes [10]. However the block and key sizes can be same, although which is not allowed under AES. Algorithm uses a dynamic combination of rounds, depending on key/block sizes, as follows:
9 rounds when key/block size is 128 bits
11 rounds when key/block size is 192 bits
13 rounds when key/block size is 256 bits

Rijndael is a substitution linear transformation block cipher, not requiring a network support. Three transformation layers used in this are linear layer, non-linear transform and key schedule mechanism. Different steps involved in Rijndael AES can be seen in the following Figure 2.

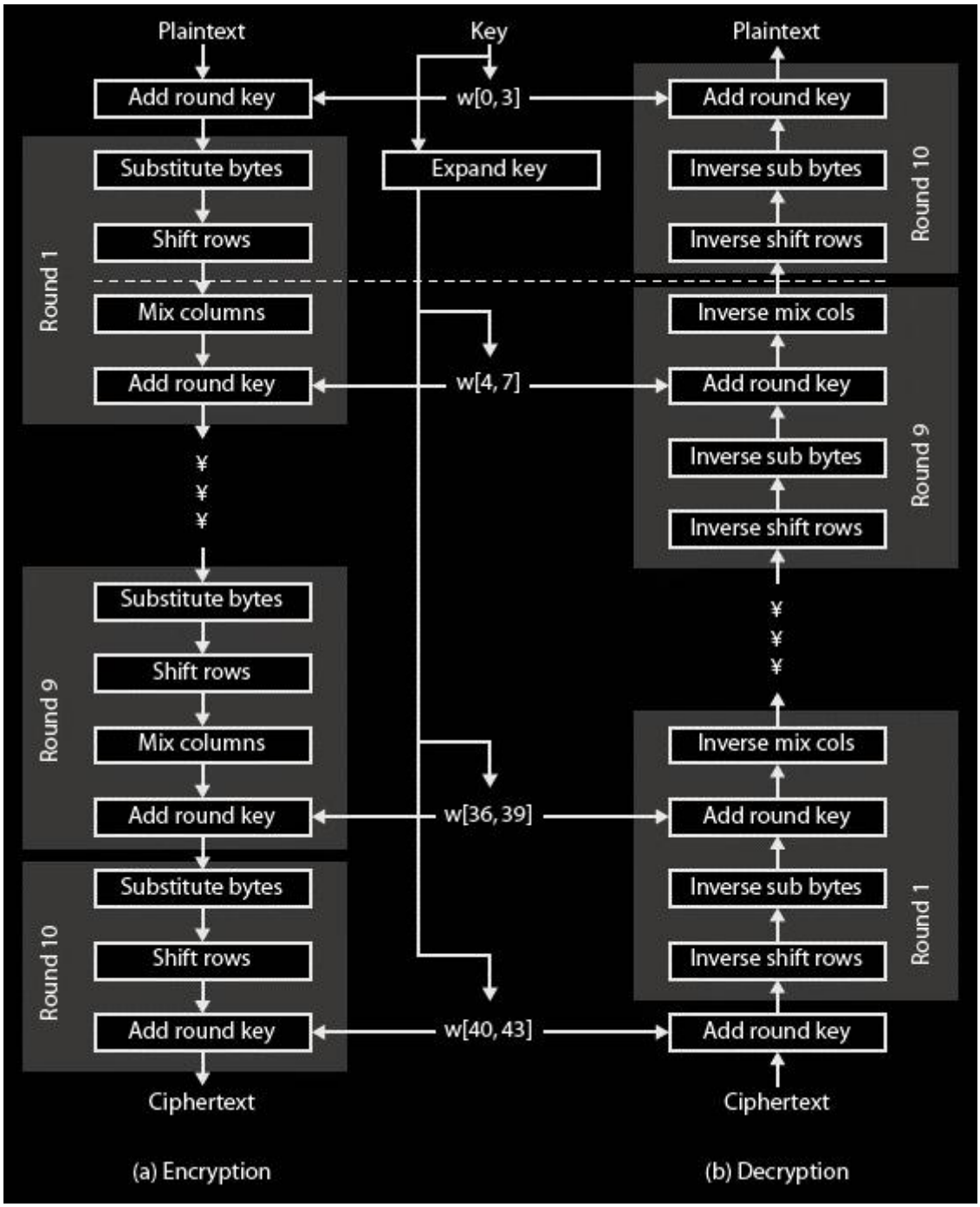

Fig. 2

Following steps are involved:

i. Byte Substitution:

Each byte is just simply substituted in this step. One Substitution Box of 16x16 bytes is used, which contains the permutation of all 256 8 bits. Each byte of state is then replaced by byte starting from row (left 4-bits) & column (right 4-bits). Example: byte {69} is replaced by byte in row 6 column 9, which has value {8A}. S-box is constructed using defined transformation of values in GF (256).The construction of S-box is done using a simple math formula of a non-linear function: 1/x.

ii. Shift Rows:

A circular byte shift in performed in each of the following

- 1st row is unchanged
- 2nd row does 1 byte circular shift to left
- 3rd row does 2 byte circular shift to left
- 4th row does 3 byte circular shift to left

Decryption of inverts is done using shifts to right. Since state is processed by columns, this part ony involves permutation bytes among the columns.

iii. Mix Columns:

Each column is processed independently. Each byte is replaced by a value dependent on all 4 bytes in the column. Effectively a matrix multiplication in GF (2^8) using prime polynomial $m(x) = x^8+x^4+x^3+x+1$ can express each column of the new state as 4 equations. Each equation is used to derive each new byte in column. Decryption requires use of inverse matrix with larger coefficients, hence is a little harder.

iv. Add Round:

XOR state with 128-bits of the round key is used. Again processed by column (though effectively a series of byte operations). Inverse is used for decryption of identical parts. Since XOR is its own inverse, with reversed keys. Add Round is therefore the iterative step.

VII. Data analytics and trade off in this mechanism

Data analysis was performed on the data being encrypted by using the mechanism. When a large amount of data is processed, there is a probability of some data getting corrupt during the compression and coding process. As you can see in Fig. 3.

X axis represents the line of data filter

Y axis represents the data timeline after processing

Z axis is the time complexity line

As we see two bands of data passing through the filter where it gets encrypted. Bits of data is passed in a continuous wave. Encryption in general is power and resource consuming. So due to this reason the data flow doesn't happen 100% accurately and thus some bits are missed during the encryption process. This creates problem at a later stage when the client requests for the data [11]. It becomes difficult for the system to render whole data files back to the user.

A proper upgrade for overcoming this problem can be by using the correct error handling system. Error correction and detection can be used in a proper manner and at a proper time in between this whole process [12]. Data backup servers should be maintained by the service providers so that the error correction can be done in the best way possible, and the discrepancies in data bits can be removed.

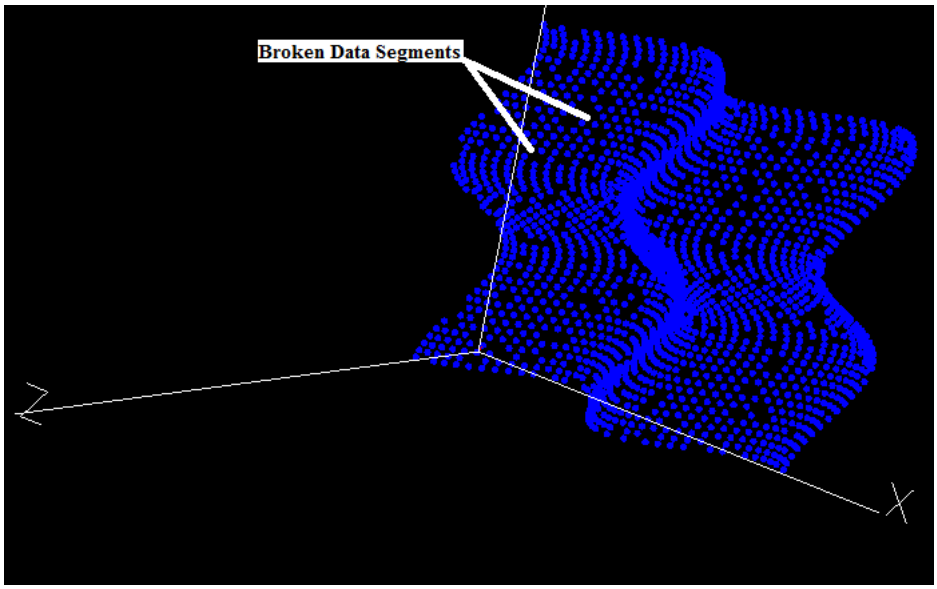

Fig. 3

VIII. Conclusion

Cloud computing is the need of the hour for today's computing requirements as the advantages of using the services are far more than the disadvantages or trade-offs we need to do. The vulnerability of data is the biggest concern for the public who store sensitive information in the cloud servers. Many of the issues have risen as a result. Shared data computation and multiple access are two major issues faced by the cloud service providers. Flexibility of choice of security measures should be given to the consumer/client so that they can benefit from the system by deciding about the perfect security strategy they need, based on the type of data stored in clouds.

If possible, consumers should upload the data in pre-encrypted form. This will eliminate the chances of service providers or other intermediate persons to hassle with the sensitive information which might be present in the uploaded data. In the SLA - Service Level Agreement the service provider should clearly mention and describe the security measures taken by them to protect the data of the user. Third party security service providers can also be used by the cloud storage provider so that they can manage separate departments. This way the risk of an immediate attack by a hacker can be reduced by a large extent. Backup data servers should be maintained so that in case of any discrepancy in data the backup may be used to get back the original details of the file. Simple encryption techniques need to be enhanced and new ways to handle cloud data should be introduced.

In the proposed mechanism server data can only be used by the client after successful signal verification by the server side data handler. Thus this takes care of any unauthorised access to cloud data. So sharing of data may be performed as required. Use of Advanced Encryption Standards justify the security reasoning of the whole encryption mechanism. Moreover Rijndael security can be further increased by adding more key rounds to the system, thus one can encrypt the data to the desired limit. The proposed mechanism can thus help the in getting a secured cloud computing environment.

Acknowledgements

This paper is dedicated to my parents and my sister, who stood by me in every frame of life. This work would not have been possible without their love and support.

References

Journal Papers:

[1]. Venkata Karthik Gullapalli and Aishwarya Asesh, Data Trawling and Security Strategies, ISSN – 2278-8727, IOSR Journal of Computer Engineering, Volume 16, Issue 6, Ver. 1, Nov - Dec 2014.
[2]. Z. Lan, V. Varadharajan and M. Hitchens, "Achieving Secure Role-Based Access Control on Encrypted Data in Cloud Storage", Information Forensics and Security, IEEE Transactions on, vol. 8, no. 12, (2013), pp. 1947-1960.
[3]. A. A. Soofi and M. I. K Fazal-e-Amin, "A Review on Data Security in Cloud Computing", International Journal of Computer Applications, vol. 94,no.5,(2014), pp.12-20.
[4]. D. E. Goldberg and J. H. Holland, Genetic Algorithms and machine learning, Guest Editorial, Machine Learning 3: 95-99, 1988 Kluwer Academic Publishers - The Netherlands.
[5]. I. Foster and C. Kesselman (eds). The Grid: Blueprint for a Future Computing Infrastructure. Morgan Kaufmann, San Francisco, USA, 1999.

[6]. K. Keahey, I. Foster, T. Freeman, and X. Zhang. Virtual workspaces: Achieving quality of service and quality of life in the Grid. Scientific Programming, 13(4):265-275, October2005.
[7]. I. Raicu, Y. Zhao, C. Dumitrescu, I. Foster, M. Wilde. "Falkon: a Fast and Light-weight tasK executiON framework", IEEE/ACM SuperComputing 2007.
[8]. Aishwarya Asesh, Tested Paradigm to Include Optimization in Machine Learning Algorithms, ISSN: 2278-0181, IJERT International Journal of Engineering Research & Technology, Vol. 4 Issue 02, February-2015.

Books:

[9]. R. Buyya, K. Bubendorfer. "Market Oriented Grid and Utility Computing", Wiley Press, New York, USA, 2008.
[10]. I. Foster, C. Kesselman, J. Nick, S. Tuecke. The Physiology of the Grid: An Open Grid Services Architecture for Distributed Systems Integration. Globus Project, 2002.

Theses:

[11]. E. Marinelli, "Hyrax: Cloud Computing on Mobile Devices using MapReduce", Master Thesis Draft, Computer Science Dept., CMU, September 2009.

Proceedings Papers:

[12]. I. Foster, C. Kesselman, L. Pearlman, S. Tuecke, and V. Welch. "The Community Authorization Service: Status and Future," In Proc. of Computing in High Energy Physics (CHEP), 2003.
[13]. M. S. Abolghasemi, M. M. Sefidab and R. E. Atani, "Using location based encryption to improve the security of data access in cloud computing", Paper presented at the Advances in Computing, Communications and Informatics (ICACCI), 2013 International Conference on. (2013, 22-25 Aug. 2013).
[14]. H. Shuai and X. Jianchuan, "Ensuring data storage security through a novel third party auditor scheme in cloud computing", Paper presented at the Cloud Computing and Intelligence Systems (CCIS), 2011 IEEE International Conference on. (2011, 15-17 Sept. 2011).
[15]. D. E. Irwin, J. S. Chase, L. E. Grit, A. R. Yumerefendi, D. Becker, and K. Yocum. Sharing networked resources with brokered leases. In Proceedings of the 2006 USENIX Annual Technical Conference (USENIX2006), Boston, USA, June2006.