IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661, p- ISSN: 2278-8727Volume 9, Issue 1 (Jan. - Feb. 2013), PP 18-25
www.iosrjournals.org

Secure Massive Data Storage With Consistency And Route Control On The Cloud

G.Nalinipriya[1], R.Aswin Kumar[2]
[1]*Associate Professor, Department of IT, Saveetha Engineering College, Chennai, India.*
[2] *PG Scholar, Department of IT, Saveetha Engineering College, Chennai, India.*

Abstract: *As the advent of computer has thrown light to the storage of massive data in small spaces the resources has been reduced broadly. Cloud computing in recent times has become a trend among organization to store data and retrieve it on-the-go around the globe. With this the data storage devices and data retrieval process arise a greater responsibility of security that should be provided to the client. This task provides a challenge to the service providers in terms of security and consistency. As shared systems are prone to be attacked, various counter measures have been proposed by experts to secure the stored data. In this work yet another efficient technique is introduced to the storage and retrieval process. Here the data is being encrypted by the advanced encryption algorithm which is considered to be the most competent in the present security scenario. A symmetric key is generated to the stored file. To further enhance the trust of the service provider the trust certificate of the company is being sent along with the keys to the client. The data is being encrypted by the AES-NI algorithm which is considered to be the most efficient up to date. A hash function is also generated to further enhance the security of the stored file. Also the trust of the service provider is provided to the client by sending the trust certificate of the company along with the keys to the client. All the communication process is taken place using the DH Key Exchange protocol. These techniques are performed to ensure the client that the stored data is being secured, integrated and a total control over the route in which the data is being communicated. In this research work, a new efficient technique is proposed which indicates the effectiveness, flexibility of the storage and retrieval process by a generic framework. This framework fills the gap between the security needs and challenges.*

Keywords: *Diffie Hellman, Advanced Encryption Standards – New Instructions, Cloud Service Provider, Data Encryption Standards, Message Digest, Message Authentication Code and Cloud Security Alliance group*

I. Introduction

The cloud is evolving as the modern way to slant marginal distribution models for IT competences. It is a way of allocating IT-enabled services in the form of software, infrastructure and more. This study and work observes the meaning of cloud computing and the ways it will advance. With old-fashioned desktop computing, we run facsimiles of software programs on our own computer. The documents we produce are kept on our own pc. Even though documents can be retrieved from other computers on the network, they can't be retrieved by computers exterior the network.

With cloud computing the software programs that we use aren't run from our own pc, but are rather warehoused on servers retrieved via internet. Anyone with accurate security authorizations can not only has the right to use the documents but can also control and work in partnership on those documents in real time without the need of any software and the user no longer need knowledge of, skill in or mechanism over the technology infrastructure in the cloud.

Moreover, nearly all IT resources can be distributed as a cloud service: application, power computation, storage capacity, networking, programming tools, communication facilities and collaboration tools.

The term Cloud is used as an exemplification for the internet, based on the cloud picture used to portray the internet in computer network illustrations.

Summing up, Cloud computing is comparatively new data storage and handling concept. It enables one to access, create and store files apart from accessing various applications like word processing, spread sheets etc. online from any computer with internet access, combined with the fastest processing speed regardless of the Operating System.

A. ***Cloud Computing Architecture***

When discussing about cloud computing, it's supportive to choose it into two segments: the front end and the back end. They join together through a network, usually the Internet. The front end is on the side the user, or simply the client. The back end is the cloud section of the structure.

If the cloud computing corporation has a portion of clients', there's a possible of high request for storage area. Some firms involve hundreds of digital storage maneuvers. Cloud computing systems need at least twice the

number of storage strategies it needs to keep all its clients data warehoused. That's because these strategies, like all computers seldom brake down. Cloud computing system necessities make a copy of all its clients' data and store it on other devices. The copies allow the central server to contact backup machines to recover data that or else would be inaccessible. Making copies of data as a backup is called redundancy.

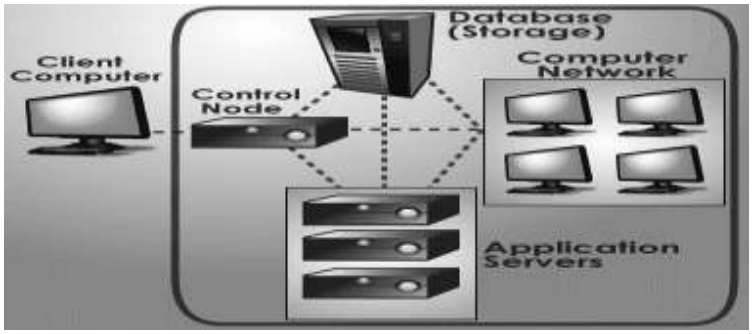

Fig. 1. A typical CC Architecture

B. Public, Private, and Hybrid Cloud

1) Public clouds:

These are managed by third parties. Works from different customers may be assorted together on the servers, storage structures, and other infrastructure in the cloud. End user does not know whose job may be executing on the same server, network, or disk as their own jobs.

2) Private clouds

This is a good option for corporations that are involved with data security and SLA. These are on-demand infrastructure operated by a single customer who maintains which application to run, and where. They possess the server, network, and disk and can choose which users are permitted to use the infrastructure.

3) Hybrid cloud

This combines the public and private cloud models. Industries own parts of the cloud and share others in a controlled way. It also offers on-demand, externally managed scale, but in addition to it the difficulty of defining how to dispense applications diagonally across these different places. While enterprises may be fascinated to these features of a hybrid cloud, these will likely be set aside for simple displaced applications that are not in need of complex databases or synchronization.

C. Applications

The applications of cloud computing are almost boundless. With the middleware, these would execute all the programs that a usual computer could do. Possibly, the whole thing from basic word processing software to tailored computer programs planned for an explicit company could work on a cloud computing system.
These applications are largely divided into categories such as

- Software as a Service (SAAS)
- Infrastructure as a Service (IAAS)
- Platform as a Service (PAAS)
- Web Services
- Managed Service Providers (*MSP*)

D. Future of Cloud Computing

One of the main motives why cloud computing is hostilely being established is the creativity or the business setting. Many industries, big and small, have come to realize the prospective of cloud computing in terms of enabling business communications without having to devote too much on supplementary organization, manpower and even time. There mere statistic that dealings in almost any form could be through online has made cloud computing a decent response to diverse business snags.

E. Concerns over CC

1) Influential Internet Link

Cloud computing will only be conceivable with a strong internet link. Cloud computing might not work in areas where link is feeble. Even though there are presentations that might work with humble dial-up connectivity, the application could easily go down particularly when there is too many data to be handled.

2) ***Interoperability***

A chief barricade to cloud computing is the interoperability of applications. Though it is conceivable to enclosure an Adobe Acrobat file into a Microsoft Word document, possessions get a little bit tackier when we discourse about web-based applications.

3) ***Security and Privacy***

The leading concern of any of initiative or commercial or end users in dealing with cloud computing is security.

Vivacious facts located in cyber space faces meaningfully greater dangers that in interior data depositories and software system. The cost of shielding that data can balance the advantages of the cloud, especially if there is a security break.

Hacks on the structure will last to be there as well. The outbreak that users experience today will also progress to adjust to different sorts of security methods. Worries can continue about loss of control over convinced sensitive data, and the lack of security for stowed kernels. Privacy is another problem. If a client can log in from any locality to access data and applications, it's likely the client's privacy could be negotiated. Cloud computing corporations will need to find methods to defend client privacy. One way is to use verification techniques such as user names and passwords. Another is to employ an authorization format -- each user can access only the data and applications applicable to his or her job.

II. Benefits of Cloud Computing

A. ***Accessible/Device and location independence***

Cloud computing authorizes users to contact systems using a web browser regardless of their place or what convenient they are using. Once as a client is related to the cloud, whatever are stored there-documents, messages, images, applications are repossessed via the Internet as infrastructure is off-site.

1) Cost effective
2) Collaborations
3) Disaster recovery
4) Powerful
5) Scalability
6) Security

B. ***Technologies that CC replaces***

- Cloud computing is substituting enormous Corporate Data Centers and redundant, costly private server infrastructure.
- Web 2.0, SaaS, Enterprise and government users are espousing cloud computing because it removes capital investment in hardware and facilities as well as decreases operations labor.
- It will replace traditional business applications like those from SAP, Microsoft, and Oracle which have been always complicated and expensive to use.

C. ***Basic Components***

Effective application of cloud computing needs proper implementation of certain components. Without any of these components, cloud computing will not be possible.

D. ***The Client/the End User***

Everything ends with the client. The hardware components, the application and the whole thing else established for cloud computing will be used in the client. Without the client, nothing will be possible.

III. Literature Survey

Yashaswi Singh, Farah Kandal and Weiyi Zhang [1] proposes about a Secured Cost Effective Multi-Cloud Storage in Cloud Computing which seeks to provide each customer with a better cloud data storage decision taking into consideration the user budget as well as providing him the best quality of service offered by available CSP.

SravanKumar.R and AshutoshSaxena [2] proposes to facilitate the client in getting a proof of integrity of the data which he wishes to store in the cloud storage server also minimizes the size of the proof of data integrity so as to reduce the network bandwidth consumption. This scheme proves advantageous to thin clients like PA's and mobile phone. The encrypting process is very much limited to only a fraction of the whole data thereby saving the computation of the client. This scheme applies only to the static storage if the data.

Anthony Bisongand and Sayed.M.Rahman [3] proposed that Deploying cloud computing in an enterprise infrastructure bring significant security concerns. Successful implementation of cloud computing in

an enterprise requires proper planning and understanding of emerging risks, threats, vulnerabilities, and possible countermeasures. We believe enterprise should analyze the company/organization security risks, threats, and available countermeasures before adopting this technology. In this paper, we have discussed security risks and concerns in cloud computing and enlightened steps that an enterprise can take to reduce security risks and protect their resources. We have also explained cloud computing strengths/benefits, weaknesses, and applicable areas in information risk management Paper is about a survey of various threat issues faced by the cloud.
Bernd Grobauer, Tobias Walloschek and Elmar Stocker[4] propose an overview of different types of vulnerabilities on the cloud. They also discuss about the state-of-art cloud offerings such as SQL Injection, command Injection and cross-site scripting. This paper deals with the overall view of the cloud and the various measures and practices that can be followed and implemented in it.

A. Existing System Description

In the existing system the authors have described about the various techniques which have been followed to ensure the privacy of the data and they have concluded that AES being the most recent and widely used is the best known algorithm. Even though AES can be used in the cloud they it has a common drawback; the time consumption. This can lead to various attacks during the encrypting process. This has to be reduced at the hour and a still more efficient algorithm is needed. Further the authors has extended their concern about more efficient algorithms are needed to have an efficient access and retrieval of the data. This needs to be done because most of the attacks on the data are only in the channel or the medium on which it is being propagated. The algorithm proposed also creates more overhead on the cloud server and results in degradation of the process.

Since the proposed mechanism are too old and DES has been broken long back; we could not rely on that as it might be too harmful. Despite of abundance of standards and products dealing with the protection of cloud computing systems, many aspects such as the third party auditing and resource allocation are still being investigated. Privacy protection issues are also a greater concern where the data are being prone to the third parties such that there is a greater chance of the sensitive data being distributed. While the other risks such as the global positioning of the data where the user is not known of the propagation of his data and is unaware of the parties who handle it. From the other perspective existing vulnerabilities in the cloud model will increase the threats from hackers. The enterprise should also verify and understand cloud security carefully analyze the security issues involved and plan for ways to resolve it before implementing the technology.

B. Proposed System Description

In the proposed system an advanced cryptographic encryption standards is being used to prevent the attack on the data. This method is necessary for the hour as the vulnerability of the data in the remote server is high. The privacy of the sensitive data that is being stored in the server is done using the most advanced algorithm; the AES-NI. Before encrypting the sensitive data a hash number along with a time stamp is being generated using MD5 such that the integrity of the stored data is confirmed. The hash number thus generated is encrypted along with the data and stored in the database. To further enhance the trust of the CSP a trust certificate of the CSP along with the hash number and the time stamp is sent to the client either through text message or to e-mail id. This hash code is retained by the client so that during the retrieval process the integrity of the data can be measured only by this. During the retrieval process the client can decrypt the data by giving the hash value that is obtained during the encrypted process. This should match the hash value stored along with the data in the database. Once it matches the data is decrypted and is shown to the client. This process is used to restrict the access of the data by others.

1) Hash Functions:

A public function that maps a message of any length into a fixed length hash value, which serves as the authenticator. We will mainly be concerned with the last class of function however it must be noted that hash functions and MACs are very similar except that a hash code doesn't require a secret key. With regard to the first class, this can be seen to provide authentication by virtue of the fact that only the sender and receiver know the key. Therefore the message could only have come from the sender. However there is also the problem that the plaintext message should be recognizable as plaintext message. The hash value is appended to the message at the source at the time when the message is assumed or known to be correct. The receiver authenticates that message by computing the hash value. Because the hash function itself is not considered to be secret, some means is required to protect the hash value.

2) AES-NI Algorithm:

AES (Advanced Encryption Standard) is an encryption standard adopted by the U.S. government starting in 2001. It is widely used across the software ecosystem to protect network traffic, personal data, and corporate IT infrastructure. AES is a symmetric block cipher that encrypts/decrypts data through several rounds.

The new 2010 Intel® Core™ processor family (code name Westmere) includes a set of new instructions, Intel® Advanced Encryption Standard (AES) New Instructions (AES-NI). The instructions were designed to implement some of the complex and performance intensive steps of the AES algorithm using hardware and thus accelerating the execution of the AES algorithms. AES-NI can be used to accelerate the performance of an implementation of AES by 3 to 10x over a completely software implementation. The AES algorithm works by encrypting a fixed block size of 128 bits of plain text in several rounds to produce the final encrypted cipher text. The number of rounds (10, 12, or 14) used depends on the key length (128b, 192b, or 256b). Each round performs a sequence of steps on the input state, which is then fed into the following round. Each round is encrypted using a subkey that is generated using a key schedule. The new AES-NI instruction set is comprised of six new instructions that perform several compute intensive parts of the AES algorithm. These instructions can execute using significantly less clock cycles than a software solution. Four of the new instructions are for accelerating the encryption/decryption of a round and two new instructions are for round key generation. The following is a description of the new instructions.

- *AESENC.* This instruction performs a single round of encryption. The instruction combines the four steps of the AES algorithm - *ShiftRows, SubBytes, MixColumns & AddRoundKe*y into a single instruction.
- *AESENCLAST.* Instruction for the last round of encryption. Combines the *ShiftRows, SubBytes, & AddRoundKey* steps into one instruction.
- *AESDEC.* Instruction for a single round of decryption. This combines the four steps of AES - *InvShiftRows, InvSubBytes, InvMixColumns, AddRoundKey* into a single instruction
- *AESDECLAST.* Performs last round of decryption. It combines *InvShiftRows, InvSubBytes, AddRoundKey* into one instruction.
- *AESKEYGENASSIST* is used for generating the round keys used for encryption.
- *AESIMC* is used for converting the encryption round keys to a form usable for decryption using the Equivalent Inverse Cipher.

3) **Benefits of using AES-NI**

The performance improvement expected with the use of AES-NI would depend on the applications and how much of the application time is spent in encryption and decryption. At the algorithm level, using AES-NI can provide significant speedup of AES. For non-parallel modes of AES operation such as CBC-encrypt AES-NI can provide a 2-3 fold gain in performance over a completely software approach. For parallelizable modes such as CBC-decrypt and CTR, AES-NI can provide a 10x improvement over software solutions. Intel continues to provide leadership in developing instruction- set extensions with recently released ISA support for Advanced Encryption Standard (AES). This paper presents the excellent performance of the AES algorithm on the Intel® Core™ i7 Processor Extreme Edition, i7-980X, using the AES New Instructions (AES-NI). Performance results for serial and parallel modes of operation are provided for all key sizes, for variable numbers of cores and threads. These results have been achieved using highly optimized implementations of the AES functions that can achieve ~1.3 cycles/byte on a single-core Intel® Core™ i7 Processor Extreme Edition, i7-980X for AES-128 in parallel modes. The paper also has a brief description of how to code to achieve these results and a reference to the complete source code. Beyond improving performance, the new instructions help address recently discovered side channel attacks on AES. AES-NI instructions perform the decryption and encryption completely in hardware without the need for software lookup tables. Therefore using AES-NI can lower the risk of side-channel attacks as well as greatly improve AES performance.

4) **Using AES-NI**

AES-NI instructions can be used in any application that uses AES for encryption. AES is very widely used in several applications such as network encryption, disk and file encryption applications. File-level and disk encryption applications use AES to protect data stored on a disk. Networking applications use encryption to protect data in flight with protocols encompassing SSL, TLS, IPsec, HTTPS, FTP, SSH, etc. There are several ways to take advantage of AES-NI in your applications, whether you are starting from scratch or optimizing existing applications. [22]

5) ***Architecture Diagram***

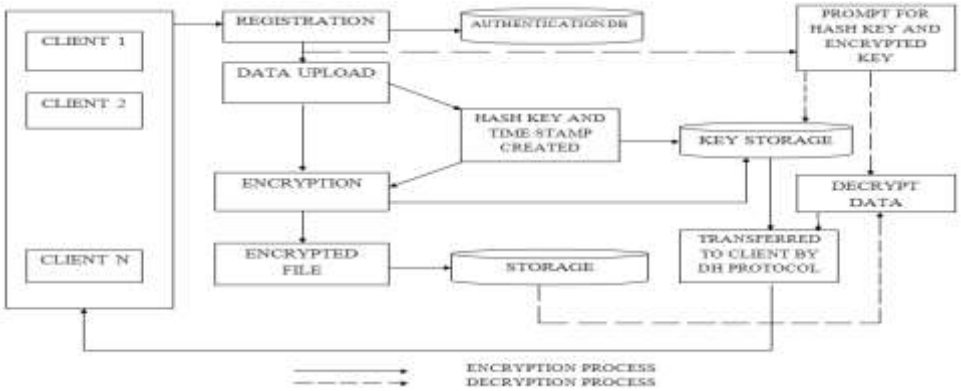

Fig. 2. Architecture diagram of the proposed massive storage data in cloud system

The clients register their credentials upon accessing the cloud to upload their data. Once their login being successful the data upload form is being displayed to upload their data. After this process a hash function is generated such that it provides a key to the user to access his AES-NI algorithm and is being stored in the data

The clients register their credentials upon accessing the cloud to upload their data. Once their login being successful the Cloud prompts for the hash value and the private key generated during the time of encryption. The client enters his key that he got which is further being checked and the data is decrypted and displayed to the client.

Here the client data is being sent to the cloud server where a hash key and a time stamp are being generated to it. This file along with the key details is being encrypted using AES-NI algorithm which is the most advanced cryptographic algorithm now. The keys that are being generated are sent through the D-H Key Exchange protocol and hence the route is being secured. After the encryption process is carried out the file is being stored in the database. When the client requests his data then the Hash keys are required to decrypt the data and once the keys matches the encryption keys are prompted and once those keys also matches the data is decrypted and retrieved.

IV. Conclusion

As there has been a rapid advancement in the storage and the retrieval of the data from the database the concentration has to be laid on the security issues. Here the data is being uploaded by the client in a way such that they can be confirmed of the place and the storage details of their file. By this way the client can be sure that his data is in a secured way and that it has the key that is equal to the key that is being generated during the upload of the data. In this phase a security protocol is being presented to secure the data files of the client in the cloud infrastructure. Here the hash key acts as a capability based model to ensure the secure access of the file only by the client. This provides further enhancement to the security concerns of the cloud. This feature also helps in providing a two-step authentication and hence the client is assured with more security for his data.

*A. **Comparision***

TABLE 1: DATA SECURITY (ENCRYPTION) IN CLOUD COMPUTING

Storage	Processing	Transmission
Symmetric encryption	Homomorphic encryption	Secret socket Layer SSL encryption
AES-DES-3DES-Blowfish-MARS…	Unpadded RSA-ElGamal…	SSL1.0-SSL3.0-SSL3.1-SSL3.2...

The above table indicates that every cloud provider encrypts the data in three types and still there is a threat in the intrusion of the stored data. Hence a more advanced encryption standard is needed to store the client's data safe.

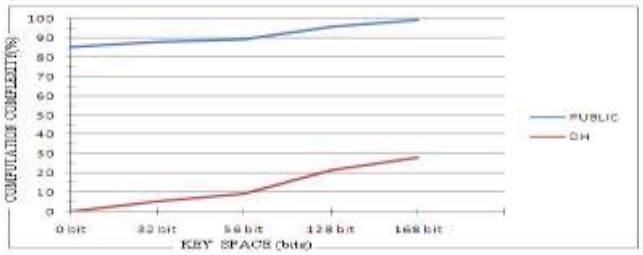

Fig. 3. Computational complexities of public keyand D-H key exchange ciphers

The above graph indicates the complexity that is being computed by the D-H key exchange protocol and other public ciphers. Usage of D-H decreases the computational time and complexity of the process

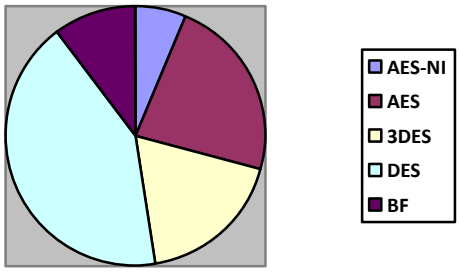

Fig. 4. Cipher Encryption performance comparison of various algorithms.

The above chart shows the performance comparison of encryption algorithms. The results shows that AES outperformed other algorithms in both the number of requests processes per second in different user loads, and in the response time in different user-load situations.

*B. **Future Work***

Since the data has been uploaded, stored and hash has been generated the future work will wholly depend upon the encryption and decryption of the data using AES-NI algorithm which is the most advanced in the cryptography field and the transfer of the keys using D-H key exchange protocol. As there are lots of stages in the encryption and decryption of the data which is being uploaded it will be practically difficult for the third party to access the data that is being stored in the cloud. It also ensures that only the person who uploads the data on the cloud will be able to access the file and also ensures that the consistency of the data is being maintained during the third pay auditing process. Since during the encryption process the latest algorithm is being used to encrypt the data that is being uploaded and the route is being controlled by the D-H protocol the system will be more secure when compared to other means of security algorithms. These methods also focus on the reduction of the complexity of the process. As these algorithms are more secure and light weight they provide a simple means in the encryption and decryption process providing lesser time for the user to secure his data.

More over real time data values are to be tested during the implementation stage. For the real time data the medical data from the hospital are planned to be taken for consideration and the security features are to be implemented in those. This will enhance the protection of the patient's records by the medical practitioners and they can be well sure that the stored patients records are maintained in a secure way.

References

[1] Yashaswi Singh, Farah Kandal and Weiyi Zhang, "A Secured Cost-Effective Multi-cloud Storage in Cloud Computing," IEEE INFOCOM 2011 Workshop on Cloud Computing.
[2] SravanKumar.R, "Data integrity proofs in cloud storage," Third International Conference on Communication Systems and Networks (COMSNETS), pp: 1 – 4, 2011.
[3] Anthony Bisong&Sayed.M.Rahman, "An Overview Of The Security Concerns In Enterprise Cloud Computing," IJNSA, Vol.3, No.1, Jan 2011.
[4] Bernd Grobauer, Tobias Walloschek and Elmar Stocker, "Understanding Cloud Computing Vulnerabilities," A white paper by SIEMENS, co-published by The IEEE Computer and Reliability Societies, March/April 2011.
[5] EmanM.Mohammed, Hatem.S.Abdelkader and Sherif El-Etriby, "Enhanced Data Security Model For Cloud Computing," INFOS2012, 14-16 May, 2012.
[6] Deyan Chen and Hong Zhao, "Data Secuity and Privacy protection Issues in Cloud Computing," IEEE International Conference of Computer Science and Electronics Engineering, 2012.
[7] National institute of Science and technology, "the NIST Definition of Cloud Computing," p.7, Retrieved July 24 2011.
[8] Mohamed Al Morsy, John Grundy and Ingo Müller, "An Analysis of The Cloud Computing Security Problem," In Proceedings of APSEC 2010 Cloud Workshop, Sydney, Australia, 30th Nov 2010.
[9] ChipurupalliSekhar, U. Nanaji, "Secure Cloud By It Auditing," International Journal Of Modern Engineering Research (IJMER) www.ijmer.com vol.1, Issue.2, pp-332-337 issn: 2249-6645.
[10] David C. Wyld, "The Cloudy Future of Government It: Cloud Computing and the Public Sector Around The World," International Journal of Web & Semantic Technology (IJWesT), Vol 1, Num 1, January 2010.
[11] Paul T. Jaeger, Jimmy Lin & Justin M. Grimes (2008): Cloud Computing and Information Policy:Computing in a Policy Cloud?," Journal of Information Technology & Politics, 5:3, 269-283.
[12] Daniel Warneke and Odej Kao, "Exploiting Dynamic Resource Allocation forEfficient Parallel Data Processing in the Cloud," IEEE Transactions On Parallel And Distributed Systems, January 2011.
[13] Andreas Berl1, ErolGelenbe, Marco di Girolamo, Giovanni Giuliani, Hermann de Meer1, Minh Quan Dang and Kostas Pentikousis, "Energy-Efficient Cloud Computing," The Computer Journal, Vol. 53 No. 7, 2010.
[14] Qian Wang, Cong Wang, KuiRen, Wenjing Lou, and Jin Li, "Enabling Public Auditability and Data Dynamics for Storage Security in Cloud Computing," Presented at The 14th EuropeanSymposium on Research in Computer Security (ESORICS'09).

[15] Ali Inan, Gabriel Ghinita, Murat Kantarcioglu, and Elisa Bertino, "A Hybrid Approach to Private Record Matching Fellow," IEEE Transactions On Dependable And Secure Computing, Vol. 9, No. 5, September/October 2012.
[16] ZhuoHao, Sheng Zhong, and Nenghai Yu, "A Privacy-Preserving Remote Data Integrity Checking Protocol with Data Dynamics and Public Verifiability," IEEE Transactions On Knowledge And Data Engineering, 2011 IEEE.
[17] VerenaKantere, Debabrata Dash, Gregory Francois, Sofia Kyriakopoulou and Anastasia Ailamaki, "Optimal service pricing for a cloud cache," IEEE Transactions on Knowledge and Data Engineering, 2011 IEEE.
[18] Website references, www.wikipedia.com, www.saleforce.com, www.ibm.com, www.sun.com
[19] Amazon EC2 and S3, online at http://aws.amazon.com/
[20] Google App Engine, Online at: http://code.google.com/appengine/
[21] S.H.Shin, K.Kobara, "Towards Secure Cloud Storage" , Demo for CloudCom 2010.
[22] http://software.intel.com/en-us/articles/intel-advanced-encryption-standard-instructions-aes-ni