ISSN (Online) : 2319 - 8753
ISSN (Print) : 2347 - 6710

International Journal of Innovative Research in Science, Engineering and Technology

Volume 3, Special Issue 3, March 2014

2014 International Conference on Innovations in Engineering and Technology (ICIET'14)
On 21st & 22nd March Organized by

K.L.N. College of Engineering, Madurai, Tamil Nadu, India

Assessment on Storage Security Progenies and Prospective Solutions of Cloud Computing

Balasubramanian.C [#1], Prasanthi.T [#2]

[#1] Department of CSE, P.S.R. Rengasamy College of Engineering for Women, Sivakasi, Tamilnadu, India

[#2] Department of CSE, P.S.R. Rengasamy College of Engineering for Women, Sivakasi, Tamilnadu, India

ABSTRACT— Cloud computing is an emerging technology of computing environment in the upcoming themes. The cloud provides an outstanding amenity to the business people, individuals, organizations, etc., with flexible infrastructure and storage as a service. Cloud computing is a type of computing that relies on sharing resources among the cloud users from a shared pool of computing resources (network, services, applications, server, databases, etc.). This computing move the application software's, databases and user files to the large data centers, where the management of these data may not be fully trustworthy. Cloud is the fullest innate with numerous security challenges, which cannot be well understood by the cloud users. Allowing users to store a large amount of data, including the sensitive information in the cloud stimulates the highly skilled attackers, thus creating a need for security to be considered as one of the top issues that exists within cloud computing. This paper depicts the survey on security issues raises in cloud computing and how the cryptographic functions are involved to preserve the integrity, confidentiality and availability of the user data.

KEYWORDS— cloud computing, Data Confidentiality, Data Integrity, Data Security

I. INTRODUCTION

Cloud computing is a model for enabling everywhere, well-located, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, applications, and services) that can be rapidly provisioned and released with minimal management effort. Cloud is used to describe a new class of network based computing that takes place over the Internet. Cloud computing environment provides conventional services to be met are infrastructure and service providers. The Cloud platforms are managed and charter out the resources to users on-demand, by the infrastructure providers. The resources from infrastructure providers are sent as loan to the end users. The giant companies like Google, Microsoft, and Amazon has attracted towards cloud computing and considered as a great influence in today's Information Technology industry. Business owners are paying more attention to cloud computing concepts because of several features.

The features are as follows:

- Lower initial investment
- Scalability
- Deploy faster
- Location independent
- Device independent
- Reliability
- Security

Although cloud computing has shown considerable amount of opportunities to the IT industry, but still there is a number of challenges to be addressed carefully. The main aim is to provide a better understanding of cloud computing and the ongoing research in this tremendously thriving arena of computer science.

II. CLOUD COMPUTING OVERVIEW

A. *Cloud Computing Definitions*

Cloud computing is a technology where the users dig up resource (Application software's, databases, platform etc.,) as a service from the Cloud by paying per-service basis, over the internet.

Buyya's definition states that "A cloud is a type of parallel and distributed system consisting of a collection of interconnected and virtualized computers that are dynamically provisioned and presented as one or more unified computing resources based on service-level agreements established through negotiation between the service provider and consumers". The Cloud model is composed with the following essential characteristics, three service models and four deployment models.

B. *Essential Characteristics of Cloud*

Cloud computing exhibits five essential characteristics defined by NIST (National Institute of Standards and Technology).

1. *on-demand Self-Service*. A consumer can acquire unilateral computing resources by pay per-service basis.
2. *Broad Network Access.* The resources are available over the network and accessed through standard mechanisms that endorse the usage by various thin or thick client platforms.
3. *Resource Pooling*. The service provider's pooled the computing resources to serve diverse users, with different the physical and virtual resources dynamically. The resources should be assigned and reassigned according to the consumers demand.
4. *Selection of Provider*. A good service provider is the key to good service. So, it is crucial to select the right service provider. One must make sure that the provider is consistent, well-reputed for their customer service and should have a verified path record in IT- related ventures.
4. *Rapid Elasticity*. Capabilities can be rapidly and elastically provisioned to the cloud users dynamically and automatically.
5. *Measured Service*. Cloud systems automatically control and optimize resource use by leveraging a metering capability at some level of abstraction appropriate to the type of service.

C. *Cloud Deployment Models*

Private cloud is a cloud infrastructure operated solely for a single organization, whether managed internally or by a third-party and hosted internally or externally. Undertaking a private cloud project requires a significant level and degree of engagement to virtualized business environment, and requires the organization to reevaluate decisions about existing resources.

Public cloud is a cloud where the services are rendered over a network that is open for public use. The users can obtain services from this cloud from the owned organizations.

Hybrid cloud is a combination of one or more clouds. Such composition expands deployment options for cloud services, allowing IT organizations to use public cloud computing resources to meet temporary needs.

Community cloud involves sharing of computing infrastructure in between organizations of the same community. For example all Government organizations within the state of California may share computing infrastructure on the cloud to manage data related to citizens residing in California.

D. *Offered Cloud Services*

Cloud provides the various service models to the users such as,

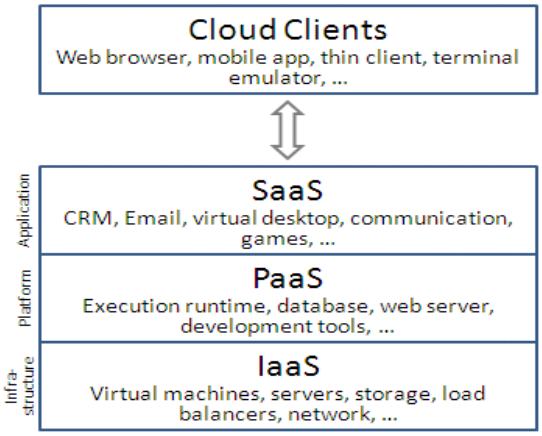

Fig1 Cloud Services

Infrastructure as a Service (IaaS) is providing general on-demand computing resources such as virtualized servers or various forms of storage (block, key/value, database, etc.) as metered resources. This can often be seen as a direct evolution of shared hosting with added on-demand scaling via resource virtualization and use-based billing.

Platform as a Service (PaaS) is providing an existent managed higher-level software infrastructure for building particular classes of applications and services. The platform includes the use of underlying computing resources, typically billed similar to IaaS products,

although the infrastructure is abstracted away below the platform.

Software as a Service (SaaS) is providing specific, already-created applications as fully or partially remote services. Sometimes it is in the form of web-based applications and other times it consists of standard non-remote applications with Internet-based storage or other network interactions.

III. Study Of Security Issues And The Prospective Solutions

A. Issues in Cloud Storage Security

Security in the cloud is one of the major areas of research. The survey shows that, the researchers are focused on efficient algorithms and the traditional encryption techniques to enhance the data storage security in the cloud.

Recent days many of the Organizations or individuals stored their data in cloud in order to acquire an instant access to the stored data over the internet. But with the stored data, there is a loss of integrity and confidentiality over the network. For Examples., the Red Hat servers in cloud were corrupted in 2008, still they remain highly confident that their systems and processes prevents the intrusion from compromising RHN and accordingly made customers to believe that their systems updated using Red Hat Network are not at risk.

Qian Wang et al., proposed a novel model for dependable and secure data storage with dynamic integrity assurance [7]. Based on the proposed principle of secret sharing and erasure coding, a hybrid share generation and distribution scheme is proposed to achieve a reliable and fault tolerant initial data storage by providing redundancy for original data components. To further dynamically ensure the integrity of the distributed data shares, we then propose an efficient data integrity verification scheme exploiting the techniques of algebraic signature and spot-checking. The proposed scheme enables individual sensors to verify in one protocol execution the correctness of all the pertaining data shares simultaneously in the absence of the original data.

Ateniese et al., stated the model for Provable Data Possession (PDP) to ensure the possession of a file at untrusted storages [3]. The public key based homomorphic tags are utilized for auditing the user's data file. However, the precomputation of the tags imposes heavy computation overhead that can be pricey for an entire file. In their subsequent work in 2008, PDP scheme used symmetric key based cryptography. This method shows a lower-overhead than their previous proposed scheme and also allows for block updates, deletions and appends to the stored file. This scheme focuses only on the single server scenario and does not provide the assurance of data availability against server failures and thus left both the distributed scenario and data error recovery issues uncharted.

Juels and Kaliski present proofs of retrievability (PORs), focusing on static archival storage of large files [6]. Their scheme's effectiveness rests largely on preprocessing steps the client conducts before sending a file F, to the server: "sentinel" blocks are randomly inserted to detect corruption, F is encrypted to hide these sentinels, and error-correcting codes are used to recover from corruption. As expected, the error-correcting codes improve the error-resiliency of their system. Unfortunately, these operations prevent any efficient extension to support updates, beyond simply replacing F with a new file F′. Furthermore, the number of queries a client can perform is limited, and fixed a priori. Let N be an RSA modulus. The verifier stores k = F mod φ (N) for file F (suitably represented as an integer).To challenge the prover to demonstrate retrievability of F, the verifier transmits a random element H ∈ φ_N. The prover returns s =$H^F \ mod \ N$, and the verifier checks that H^k mod N = s. This protocol has the drawback of requiring the prover to exp onentiate over the entire file F .

Cong wang et al., projected the ideas for generating the proofs of data storage in the cloud by homomorphic identication protocols [7].It allows the client to verify whether the cloud server loyally stores the copy of a file. In prior, the proof of storage was developed from the homomorphic linear authenticator (HLA) but later it was generated from the signatures and message authentication codes (MAC).

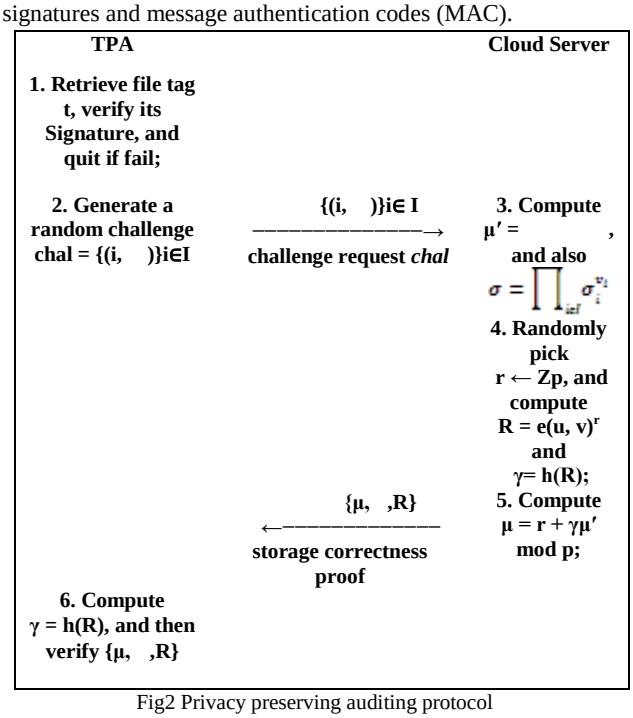

Fig2 Privacy preserving auditing protocol

This scheme commonly translates any HLA to verifiable proof of storage where the complexity of communication and the client's state are autonomous of the file size.

Sadiaet al., implements an extensive authentication protocols with RSA to enhance the security of cloud computing [8]. This scheme is to ensure the storage correctness of the users file previously stored in the cloud. It proposes the two sections, one of which analyzing the attackers involved in accessing the client data and the other uses Challenge-Handshake Authentication Protocol (CHAP) for authentication , RSA for encryption of users data. In this research, the asymmetric key encryption (RSA) encryption algorithm is used for cryptography. In RSA, anyone can encrypt messages using the public key, but only the holder of the paired private key can decrypt. Security depends on the secrecy of that private key and two parts of the key pair are mathematically linked.

Lanxiang and Gongde, have discussed the RemoteData Checking (RDC) scheme, that incorporates the mechanisms to mitigate arbitrary amounts of data corruption in cloud server [23]. In particular, protection against small corruptions (i.e., bytes or even bits) ensures that attacks modify a few bits do not destroy an encrypted file or invalidate authentication information. The initial solution to the integrity problem is RSA based hash functions to hash the entire file at every challenge. This technique solves the initial problem of storing data to the external server that verify the sever continually and faithfully stores entire file. But the problem with this technique is providing exponential over the entire file F and accessing the entire file's blocks. This is clearly prohibitive for the server whenever the file is large. Early RDC schemes have focused on static data, whereas later schemes such as DPDP support the full range of dynamic operations on the outsourced data, including insertions, modifications, and deletions.

Wayne stated the benefits of cloud computing and the basic security issues associated with the cloud services [9]. The security issues faced by end users or clients are mandatory. The researchers and professionals generate the strong security policies to make sure that the data is safe and prohibited from an unauthorized access, in both corporate data centers and in the cloud servers. The key security issues recognized and addressed in this paper are end user trust, Insider Access, Visibility, Risk Management, Client-Side Protection, Server-Side Protection, Access Control and Identity management. The main work is to identify and discuss the cloud security issues and in this they didn't proposed any tool or framework to address these issues.

B. Prospective Solutions

To boost the security in cloud server data by using the cryptographic techniques:

1. Cryptographic Encryptions

Encrypting data is one of the solutions to secure cloud data , but it limits the efficiency of the cloud. This is because encrypted documents must first be decrypted before they can be searched or manipulated. Furthermore, cloud data must be encrypted before storing. Performing encryption and decryption on large data sets can be prohibitively expensive and time consuming. The data loss in server will violate the data integrity and affects the storage correctness in cloud.

DES: (Data Encryption Standard), was the first encryption standard to be recommended by NIST (National Institute of Standards and Technology). Since many attack that exploit the weaknesses of DES, which made it an insecure block cipher.

3DES (Triple DES), the enhancement of DES is, the 3DES. In this standard, the encryption method is similar to the original DES but the encryption level is increased by 3 times and slower than other block cipher methods.

AES: (Advanced Encryption Standard), is the new encryption standard recommended by NIST to replace DES. Rijndael (pronounced Rain Doll) algorithm was selected in 1997 after a competition to select the best encryption standard. Brute force attack is the only efficient attack known against AES, since the attacker tries all the possible combinations of characters or values to unlock the encryption.

BLOWFISH: It is one of the most common public domain encryption algorithms provided by Bruce Schneier. Blowfish is a variable key length of 64 bit block cipher. This algorithm can be optimized in both hardware applications and software applications. Although it suffers from weak key problems, no attack is known to be successful against this standard.

TABLE1
Comparison of Algorithms Based Upon the Four Characteristics

Characteristics / Algorithms	Block Size	Key Size	Speed	Security
AES	128	128,192 and 256	High	More secure
DES	64	56	Low	Less secure when speed increases
3DES	128	112 or 168	Low	Less secure
Blowfish	64	32 to 448	Faster except when changing keys	More secure

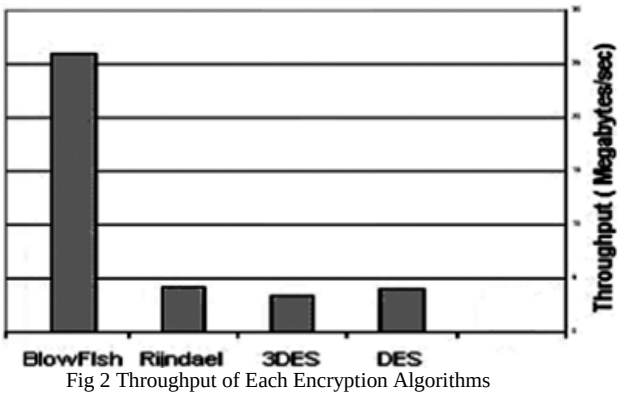

Fig 2 Throughput of Each Encryption Algorithms

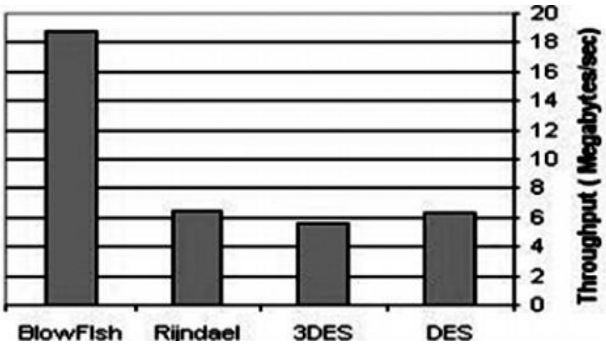

Fig 3 Throughput of Each Decryption Algorithm

2. Homomorphic Metadata Generation

The homomorphic metadata are generated based on the message authentication codes, where these codes are constructed based on the hash functions. In hashing algorithm, mathematical function is applied to the data, an alphanumerical value is produced. There are many different hash functions, SHA-1, 256, 512, MD5, CRC etc., that each returns a unique value. A set of data will produce the same hash value in most cases, as long as the same mathematical function is used and nothing in the data has been changed. The value that is produced can be thought of as a "digital fingerprint".

TABLE 2
Comparative Study Of Hashing Algorithms

Parameters / Algorithms	*DIGEST SIZE*	*ROUNDS*	*MB/SEC*	*CYCLE/ BYTES*	*ATTACK*
MD5	128	64	255	6.8	COLLISION / PREIMAGE
SHA-1	160	80	153	11.4	COLLISION
SHA-256	256	64	111	15.8	NONE
SHA-512	512	80	99	17.7	NONE
TIGER	192	24	214	8.1	COLLISION / PREIMAGE

It is possible for two different sets of data to come up with an identical hash value. When two sets of data produce the same hash value, it is called a "Collision". For a collision resistant hash function H, it is computationally infeasible to find out any two data's D and D' such that H (D) = H (D') where D = D'. Then the authentication tag φ is generated based on hashing algorithms.

The file is divided into blocks

$$D \longrightarrow \sum B_1 + B_2 \ldots + B_n \quad \text{------- (1)}$$

Authentication tag generated for each block

$$B_1, B_2 \ldots B_n \longrightarrow \varphi_1, \varphi_2 \ldots \varphi_n \quad \text{------- (2)}$$

Tags and blocks are moved to cloud

$B_1, B_2 \ldots B_n$

$+ \quad \Longrightarrow$ CSP

$\varphi_1, \varphi_2 \ldots \varphi_n$

SHA1, for example, a collision rate of 2^69, (5.9 x 10^20).Hash algorithms have different Collision rates. Using multiple hashing algorithms also reduces the chance of having a collision.

3. Auditing

In cryptography, a Trusted Third Party (TTP) is an entity which facilitates secure interactions between two parties who both trust this third party. The scope of a TTP within an Information System is to provide end-to-end security services, which are scalable, based on standards and useful across different domains, geographical areas and specialization sectors. In cloud environment includes the Trusted Third Party Auditor (TTPA) for verifying the data storage correctness of cloud server in a timely manner. As described by Castell, ''A Trusted Third Party is an impartial organization delivering business confidence, through commercial and technical security features, to an electronic transaction. It supplies technically and legally reliable means of carrying out, facilitating, producing independent evidence about and/or arbitrating on an electronic transaction. Its services are provided and underwritten by technical, legal, financial and/or structural means''. Mainly auditing scheme involves the following algorithms:

KeyGen(δ) $\longrightarrow$ (pr_x, sk_a, sk_h). It takes input as secret parameters (δ) of the user or data. It randomly chooses the secret key and the secret hash key $sk_a, sk_h \in z_n$.

AuthGen(D, sk_a, sk_h) $\longrightarrow$ *A*. The authentication tag is generated based on the data D, the secret key sk_a and the secret hash key sk_h. It selects r random values $y_1, y_2, y_3 \ldots, y_r \in z_n$ and also computes $v_i = f_1^{x_j} \in M_1$ for j ∈ [1,t]. The authtag is computed as,

$$a_i = (\text{h}(sk_h, Q_i). \prod_{j=1}^{t} u_j^{d_{ij}})\, sk_a$$

Where Q_i =ID‖ i (the "‖" denotes the concatenation operation), in which ID is the identifier of the data.

Test (D,A)$\longrightarrow \varphi$. The test consists of authentication proof AP. The authproof is generated as

$$\text{AP} = \prod_{j \in z} a_j^{u_j}.$$

Result (R) $\longrightarrow$ *{"success", "failure"}:* The verifier checks the validity of the response or result(R). If it is valid, then output will be a "success" one, otherwise the function outputs be a "failure".

III. CONCLUSIONS

Cloud Computing is an extension of distributed computing systems. For example, "VPN tunneling can be used for secure communication; existing encryption methods can be used to ensure protection of data on the cloud; and existing user-centric authentication methods, such as Open ID, can be used to authenticate with cloud services". There are number of inbuilt features in cloud computing, such as "resource pooling, multitenancy, rapid elasticity, broad network access, and on-demand self-service", where the existing security techniques are not adequate to deal with cloud security risks. This paper, presents a survey on cloud computing and the cryptographic techniques are involved to address the cloud computing security issues. The security techniques are used by the clients and end users on data based on PKI and hashing functions, in order to achieve the integrity, confidentiality, privacy for the stored data etc.

REFERENCES

1. P. Mell and T. Grance, "The NIST definition of cloud computing," National Institute of Standards and Technology, Tech. Rep., 2009.
2. M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. H. Katz, A. Konwinski, G. Lee, D. A. Patterson, A. Rabkin, I. Stoica, and M. Zaharia, "A view of cloud computing," Commun. ACM, vol. 53, no. 4, pp. 50–58, 2010.
3. Ateniese G, Burns R, Curtmola R, Herring J, Kissner L, Peterson Z, and Song D, "Provable data possession at untrusted stores," in Proc. of CCS'07. New York, NY, USA: ACM, 2007, pp. 598–609.
4. Ateniese G, Pietro R.D, Mancini L.V, and Tsudik G, "Scalable and efficient provable data possession," in Proc. of SecureComm'08. New York, NY, USA: ACM, 2008, pp.1-10.
5. Bowers K.D, Juels A, and Oprea A, "Proofs of retrievability: Theory and implementation," Cryptology ePrint Archive, Report 2008/175, 2008.
6. Juels A and Kaliski B.S, Jr., "Pors: proofs of retrievability for large files," in Proc. of CCS'07. New York, NY, USA: ACM, 2007, pp. 584–597.
7. Cong Wang, Sherman S.-M. Chow, Qian Wang, Kui Ren, and Wenjing Lou, "Privacy-Preserving Public Auditing for Secure Cloud Storage," IEEE Transactions On Cloud Computing, Year 2013.
8. Sadia Marium, Qamar Nazir, Aftab Ahmed, Saira Ahthasham ,Mirza Aamir Mehmood "Implementation of Eap with RSA for Enhancing The Security of Cloud Computing," International Journal of Basic and Applied Sciences, 2012, pp. 177-183.
9. Wayne A. Jansen, "Cloud Hooks: Security and Privacy Issues in Cloud Computing," 44th Hawaii International Conference on System Sciences 2011.
10. C. Erway, C. Papamanthou, and R. Tamassia, "Dynamic provable data possession," in ACM Conference on Computer and Communications Security , E. Al-Shaer, S. Jha, and A. D.Keromytis, Eds. ACM, 2009, pp. 213–222.
11. Q. Wang, C. Wang, K. Ren, W. Lou, and J. Li, "Enabling public auditability and data dynamics for storage security in cloud computing," IEEE Trans. Parallel Distrib. Syst., vol. 22, no. 5, pp. 847–859, 2011.
12. Wang, Q. Wang, K. Ren, and W. Lou, "Privacy-preserving public auditing for data storage security in cloud computing," in INFOCOM. IEEE, 2010, pp. 525–533.
13. J. Walker, M. Kounavis, S. Gueron and G. Graunke "Recent Contribution to Cryptographic Hash Functions," Intel Technology Journal, vol-13, issue-2, 2009, pp- 80-95.
14. S.M. Bellovin, E.K. Rescorla," Deploying a New Hash Function," presented at first NIST Workshop", 2005. Available at http://www.csrc.nist.gov/pki/HashWorkshop/2005/Oct31_Presentations/Bellovin.new-hash.pdf.
15. J. Li, M. N. Krohn, D. Mazieres, and D. Shasha, "Secure untrusted data repository (sundr)," in Proceedings of the 6th conference on Symposium on Operating Systems Design & implementation, Berke-ley, CA, USA, 2004, pp. 121–136.
16. Y. Zhu, H. Wang, Z. Hu, G.-J. Ahn, H. Hu, and S. S. Yau, "Dynamic audit services for integrity verification of outsourced storages in clouds," in SAC , W. C. Chu, W. E. Wong, M. J. Palakal, and C.-C. Hung, Eds. ACM, 2011, pp. 1550–1557.
17. K. Zeng, "Publicly verifiable remote data integrity," in ICICS , ser.Lecture Notes in Computer Science, L. Chen, M. D. Ryan, and G. Wang, Eds., vol. 5308. Springer, 2008, pp. 419–434.
18. G. Ateniese, S. Kamara, and J. Katz, "Proofs of storage from homomorphic identification protocols," in ASIACRYPT, ser. Lecture Notes in Computer Science, M. Matsui, Ed., vol. 5912. Springer, 2009, pp. 319–333.
19. Yamamoto, S. Oda, and K. Aoki, "Fast integrity for large data," in Proceedings of the ECRYPT workshop on Software Performance Enhancement for Encryption and Decryption. Amsterdam, the Netherlands: ECRYPT, June 2007, pp. 21–32.
20. M. A. Shah, M. Baker, J. C. Mogul, and R. Swaminathan, "Auditing to keep online storage services honest," in HotOS, G. C. Hunt, Ed.USENIX Association, 2007.
21. C. Wang, K. Ren, W. Lou, and J. Li,"Toward publicly auditable secure cloud data storage services," IEEE Network , vol. 24, no. 4, pp. 19–24, 2010.
22. Cloud Security Alliance. Security guidance for critical areas of focus in cloud computing. http://www.cloudsecurityalliance.org/,April 2009.
23. Lanxiang Chen, Gongde Guo, "An Efficient Remote Data Possession Checking in Cloud Storage," JDCTA: International

Journal of Digital Content Technology and its Applications, Vol. 5, 2011, pp. 43-50.
24. Top 7 threats to cloud computing DOI = www.net-security.org/secworld.php?id=8943
25. Qiu Xiu-feng, Liu Jian-Wei, Zhao Peng-Chuan. "Secure Cloud Computing Architecture on Mobile Internet", IEEE 2011.