Journal of Software Engineering and Applications, 2014, 7, 347-353
Published Online May 2014 in SciRes. http://www.scirp.org/journal/jsea
http://dx.doi.org/10.4236/jsea.2014.75031

Scientific Research

New Proposed Robust, Scalable and Secure Network Cloud Computing Storage Architecture

Fawaz S. Al-Anzi, Ayed A. Salman, Noby K. Jacob

Compute Engineering Department, Kuwait University, Kuwait City, Kuwait
Email: Fawaz.Alanzi@ku.edu.kw, Ayed.Salman@ku.edu.kw, nobyjacob2006@gmail.com

Received 20 February 2014; revised 18 March 2014; accepted 25 March 2014

Abstract

Cloud computing describes highly scalable computing resources provided as an external service via the internet. Economically, the main feature of cloud computing is that customers only use what they need, and only pay for what they actually use. Resources are available to be accessed from the cloud at any time, and from any location via the internet. There's no need to worry about how things are being maintained behind the scenes—you simply purchase the IT service you require. This new, web-based generation of computing utilizes remote servers for data storage and management. One of the challenging issues tackled in the cloud computing is the security of data stored in the service providers' site. In this paper, we propose a new architecture for secure data storage in such a way that users' data are encrypted and split into various cipher blocks and distributed among different service providers site rather than solely depend on single provider for data storage. This architecture ensures better reliability, availability, scalability and security.

Keywords

Cloud Computing, Data Storage, RAID, Security, Service Provider

1. Introduction

The paradigm shift from traditional software models to the Internet has progressively gained momentum over the last 10 years. Traditional business applications have always been very complex and costly. The amount and type of hardware and software required to run them are scary. With the arrival of cloud computing, those headaches are eliminated because we are not handling hardware and software. It is the responsibility of a proficient Service Provider. The shared infrastructure means that it works like a utility. We only pay for what we need,

How to cite this paper: Al-Anzi, F.S., *et al.* (2014) New Proposed Robust, Scalable and Secure Network Cloud Computing Storage Architecture. *Journal of Software Engineering and Applications*, **7**, 347-353.
http://dx.doi.org/10.4236/jsea.2014.75031

upgrades are automatic, and scaling up or down is easy. Businesses are running all kinds of applications in the cloud, like customer relationship management, human resources management, finance, and much more. Some of the world's largest companies moved their applications to the cloud after rigorously testing the security and reliability of the infrastructure. Most IT departments are forced to spend a significant portion of their time on frustrating implementation, maintenance, and upgrade projects that too often don't add significant value to the company's bottom line. Increasingly, IT teams are turning to cloud computing technology to minimize the time spent on lower-value activities and allow IT to focus on strategic activities with greater impact on the business. To find enough storage space to hold all the user data they have acquired is a real challenge. Some people store data in larger hard drives. Others prefer external storage devices like USB drives or external hard drives. But some are choosing to rely on a growing trend: **cloud storage**.

Cloud storage really refers to saving data to an off-site storage system maintained by a third party. Instead of storing information to your computer's hard drive or other local storage device, you save it to a remote database. The Internet provides the connection between your computer and the database. Cloud storage providers operate large data centers, and people who require their data to be hosted buy or lease storage capacity from them. They virtualize the resources according to the requirements of the customer and expose them as storage pools, which the customers can themselves use to store files or data objects. Physically, the resource may span across multiple servers and multiple locations. The safety of the files depends upon the hosting companies, and on the applications that leverage the cloud storage.

Cloud storage services may be accessed through a web service application programming interface (API) or by applications that utilize the API, such as cloud desktop storage, a cloud storage gateway or Web-based content management systems.

Cloud storage has the same characteristics as cloud computing in terms of agility, scalability, elasticity and multi-tenancy, and is available both off-premises and on-premises. The cloud storage makes data safety by divided data to small pieces and save them to different places. If data pieces in one data center or a disk crashed, the data can be resumed by left pieces. It is an important method to promote access performance and system availability.

In cloud computing environment, data are stored as public in service providers site, so data are highly insecure. Depending on a single service provider for data storage in cloud environment is not trustworthy. Cloud data storage is growing in popularity due to the benefits it provides, such as simple, anywhere access with independent geographical locations, avoidance of capital expenditure on hardware and software, the removal of the burden of in-house maintenance and management. The pioneer of Cloud Computing vendors, Amazon Simple Storage Service (S3) and Amazon Elastic Compute Cloud (EC2) are both well-known examples. It is basically the delivery of data storage as a service, from a third party provider, with access via the internet and billing calculated on capacity used in a certain period (e.g. per month). While Cloud Computing makes these advantages more appealing than ever, it also brings new and challenging security threats towards users' outsourced data. The major security challenge with clouds is that the owner of the data may not have control of where the data is placed.

Depending on a single service provider for data storage in cloud computing is insecure. In this paper, we propose a Robust, Scalable and Secure Network Storage (RSSNS) architecture which depends on multiple service providers for the secure storage of outsourced data. In order to provide better availability, reliability and security, user data are encrypted and split into various cipher blocks and distributed among available service providers. Data loss will happen due to hardware or network problem in the service provider's site. In order to recover data from any data loss due to hardware or network issues in service provider's site, we adopt a distributed parity scheme in this architecture. The second important aspect used in this architecture is that service provider site adopts RAID (Redundant Array of Inexpensive Disks) storage scheme for the better availability and reliability of data in data storage servers.

In [1], the authors discussed distributing data over multiple clouds in such a way that if an adversary is able to intrude in one network, he cannot retrieve any meaningful data; because it's complementary pieces which are stored in other network. In [2] and [3], the authors discussed the idea of RAID technology for storage in cloud computing. Cryptographic measures [4] alone cannot meet the privacy demanded by cloud computing services. It is insufficient for ensuring data privacy in cloud computing. In [5], the authors put the idea of distributing the data over multiple cloud service providers site rather than centralized distribution of data. Our approach is also similar to this approach with a change in the distribution scheme.

In [6]-[8], the authors discussed the cloud storage system structure which consists of access layer, application interface layer, basic management and physical storage layer. In [9], the authors focused on the research by the combination of private cloud and cloud storage. Wu *et al*. [10] proposed the infrastructure of cloud storage and to hide complexity of hardware and software from its users. Zhang *et al.* [11] analyses the advantage and feasibility of private cloud storage technology based on Hadoop. Zhang *et al.* [12] used Service Level Agreement (SLA) as the common standard between user and service provider to ensure data security in cloud storage system. Koletka *et al.* [13] [14] proposed architecture to securely store user data in public cloud and private cloud using encryption. Various researches of cloud storage applications are described and implemented in [15]-[17]. Liu *et al.* [18] analyse security issue in cloud storage according to cloud computing concepts and features.

In the proposed system, user data are encrypted and split into cipher blocks. The cipher blocks are distributed among available service providers site. **Figure 1** shows the proposed data storage architecture with the host machine represented as client and Service providers marked as SP1 to SPn. Not only encrypted blocks of data, but also the parity information associated with the distributed data are also stored in the service provider's data server. This parity information is not stored on single service provider server, but it is distributed among the available service providers for the efficient reconstruction of data from the available data blocks. For better availability of data, each data server in the service provider premises adopts RAID level implementation. Based on the performance comparison of various RAID levels, we suggest RAID 10 for implementation.

The RAID 10 combines the best features of striping and mirroring to yield large arrays with high performance in most uses and superior fault tolerance. RAID 10 has been dramatically increasing in popularity as hard disks become cheaper. It provides very good to excellent overall performance by combining the speed of RAID 0 with the redundancy of RAID 1 without requiring parity calculations. **Figure 2** represents the detailed architectural diagram of the proposed architecture with three service providers data.

Storage server in the service providers' location uses Raid level 10 for data storage. Let D be the original data, client wants to store in the cloud storage. The original data D is encrypted to form D' and is split into various cipher blocks A, B, C, D, E, F. and is stored in service providers sites sequentially. Suppose data block A is stored in SP1 and data block B is stored in SP2. In order to reconstruct the data blocks A and B due to any system or network failure, the parity information P associated with these data block is stored in SP3. Similarly the parity information Q associated with block C and D is stored in SP2 and R related to block E and F in SP1. Here, distributed parity scheme is used. Since we are using RAID 10, each data block and parity blocks are striped and mirrored. Data block A on SP1 is striped into two blocks as A1 A2 and mirrored copy also stored on SP1.Similarly the data blocks and parity blocks on other service providers site are also striped and mirrored.

In RAID 10 storage scheme shown in **Figure 3**, an even number of disks are required. Each disk array has a replica disk array, which is mirrored set of the former. Minimum of four disks are needed for implementing RAID 10. Since a mirrored copy of striped data is stored on dual disk, it is able to handle single disk failure. But in the case of double disk failure, we cannot recover the data in RAID 10. So in the proposed architecture, we introduce a parity scheme.

The parity scheme introduced in this architecture is explained as shown in **Figure 4**. Suppose customer data are distributed among three service providers SP1, SP2 and SP3.The parity information P related to data block A stored in SP1 and B stored in SP2 is stored in SP3. If any data loss will happen on data block A in SP1, we can reconstruct data block.

A with the help of other data block B in SP2 and parity information P in SP3. Similarly if data block B in SP2 is corrupted, we recover it with the help of data block A in SP1 and parity P on SP3. So we can effectively reconstruct the data with the help of this parity scheme, if double disk failure occurs. This scheme not only rectifies the problems related to hardware but also sorts out the data loss due to network issues in any of the service providers site. So it ensures the reliability of the proposed architecture.

2. Security Issues

Data security is one of the most critical issues related to any storage architectures. Even though cloud service providers have dominant infrastructure and security mechanisms to ensure customer's data safety and availability, several reports related to privacy of data have been outward in recent years. To ensure the security of the customer data, we distributed data among available service providers rather than storing whole data on single service provider site.

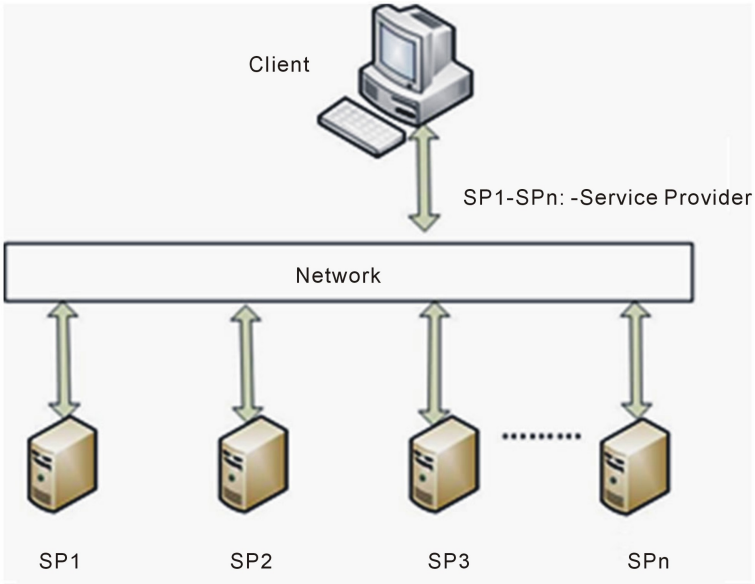

Figure 1. Proposed RSSNS architecture for cloud data storage.

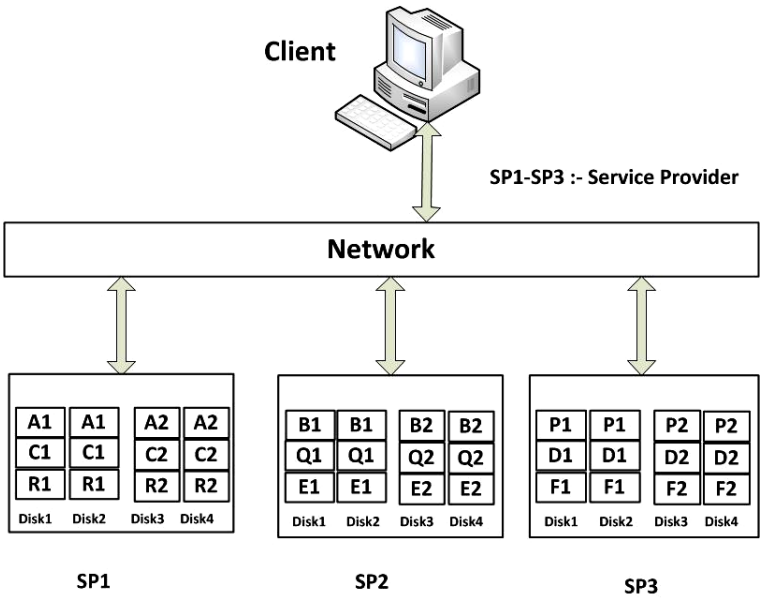

Figure 2. Proposed RSSNS architecture with three service providers.

Suppose customer data D is to be outsourced. In the centralized storage scheme, the whole data are stored on single service provider. So data are insecure in centralized storage scheme. As a security concern, in the proposed architecture, original data D is encrypted to D' and split into cipher blocks A and B as shown in **Figure 5**. Let us assume that two cloud service providers are available say SP1 and SP2. The encrypted data are distributed among service providers in such a way that cipher block A is stored on SP1 and B is stored on SP2.Proposed architecture use RAID 10 for storage. Therefore blocks A and B are again striped and mirrored (A: A1, A2 and B: B1, B2). The splitting of data blocks is done in such a way that, a single service provider cannot able to retrieve any information from the data stored in his network. The other security threat encountered is the cloud service provider might collude together to reconstruct and access the customer's stored data.

Here in this approach, the encryption and distribution is carried out in such a way that, data reconstruction is not possible, even though couple of service provider will collude each other. It guarantees the security of the

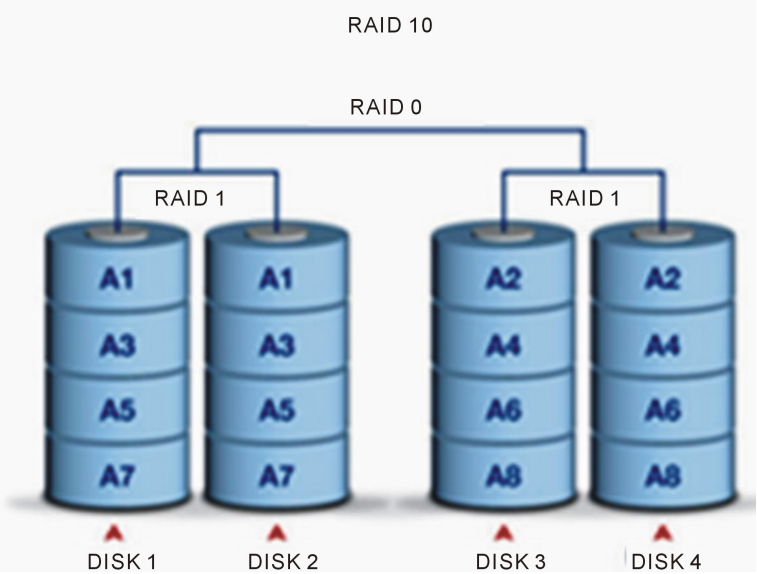

Figure 3. RAID 10 storage scheme.

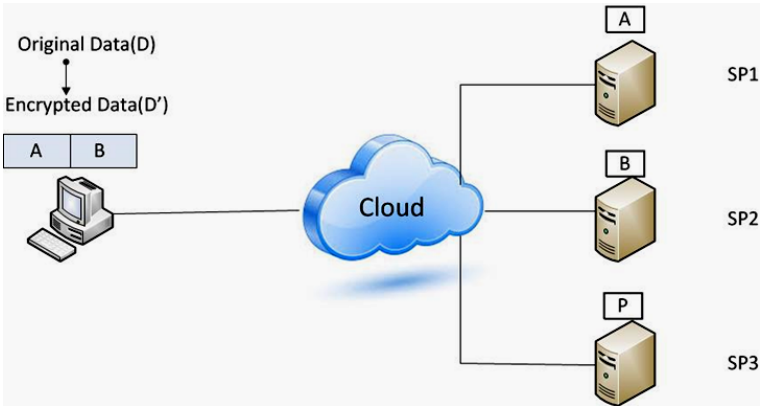

Figure 4. Parity scheme.

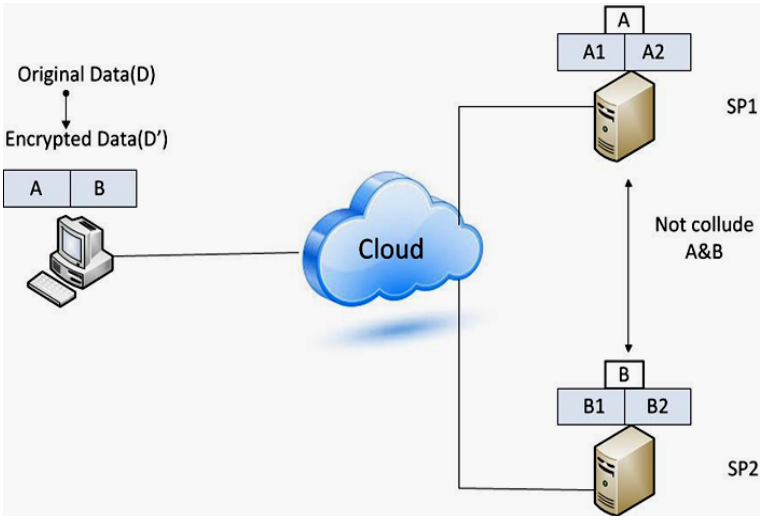

Figure 5. Security of the proposed architecture.

proposed architecture.

Data availability is achieved through redundancy involving where the data are stored and how it can be reached. Availability of the data stored on any storage device depends on how fast the data are accessed. Bandwidth of the network channel depends on availability. So high speed network cables are used for data retrieval. The RAID level implementation of the proposed architecture also offers better availability of the data.

3. Analysis of Our Proposed Scheme

In this section, we analyze security and performance properties of our proposed architecture.

Security: Ensuring the security of the data stored in the cloud storage is one of the major challenges. The Service provider might be honest, but malicious users creates security problem. This is a severe threat for critical data such as medical or financial records, as cloud service provider employees has physical access to the hosted data. To tackle the security issue we encrypt the original data and later by distributing the fragments transparently across multiple service providers. This way, none of the storage vendors is in an absolute possession of the client's data.

Availability: Management of computing resources as a service by a single Service provider implies the risk of a single point of failure. This failure depends on many factors such as hardware, software or network failure In July 2008, for instance, Amazon storage service S3 was down for 8 hours because of a single bit error. Our solution addresses this issue by storing the data on several cloud storage providers—whereby no single entire copy of the data resides in one location, and only a subset of providers needs to be available in order to reconstruct the data.

Reliability: The reliability of the proposed architecture is achieved by the parity scheme, by enabling the application to retrieve data correctly even if some of the providers corrupt or lose the entrusted data.

The proposed cloud storage architecture based on RAID technology outperforms the muti-cloud storage architecture proposed by Singh *et al.* [1] in terms of security, availability and reliability.

4. Conclusion and Future Directions

In this paper we proposed a new, web-based generation of computing utilizes remote servers for data storage and management. The model which addresses the challenging issue tackled in the cloud computing is the security of data stored in the service providers' site. The new architecture for secure data storage allows users' data to be encrypted and split into various cipher blocks and distributed among different service providers site rather than solely depend on single provider for data storage. This architecture ensures better reliability, availability, scalability and security.

Future directions of this research are to investigate the reliability of such model as well as reliability, availability, scalability, performance and robustness. Another important point to investigate is to build a business model for a fair customer charge of such storage services by the SPs.

References

[1] Singh, Y., Kandah, F. and Zhang, W. (2011) A Secured Cost-Effective Multi-Cloud Storage in Cloud Computing. *IEEE INFOCOM Workshop on Cloud Computing*, Tainan, 16-18 December 2010, 619-624.

[2] Chen, P.C., Freg, C.P., Hou, T.W. and Teng, W.G. (2010) Implementing RAID-3 on Cloud Storage for EMR System. *IEEE International Computer Symposium*, Tainan, 16-18 December 2010, 850-853.

[3] Joshi, S., Patwardhan, U. and Deshpande, P. (2010) RAID 5 for Secured Storage Virtualization. *IEEE International Conference on Data Storage and Data Engineering*, Bangalore, 9-10 February 2010, 278-282.

[4] Dijk, M. and Juels, A. (2010) On the Impossibility of Cryptography Alone for Privacy-Preserving Cloud Computing. *HotSec*'10 *Proceedings of the* 5*th USENIX Conference on Hot Topics in Security*, Article No. 1-8.

[5] Olivera, P.F., Lima, L., Barros, J. and Medard, M. (2010) Trusted Storage over Untrusted Networks. *IEEE Global Telecommunication Conference*, Miami, 6-10 Decemebr 2010, 1-5.

[6] Amazon.com (2008) Amazon Web Services (AWS). http://aws.amazon.com

[7] http://en.wikipedia.org/wiki/Cloud_computing

[8] Sun, J. and Yue, S.-S. (2011) The Application of Cloud Storage Technology in SMEs. *International Conference on E-Business and E-Government* (*ICEE* 11), Shanghai, 6-8 May 2011, 1-5.

[9] Deng, J., Hu, J., Liu, A.C.M. and Wu, J. (2010) Research and Application of Cloud Storage. 2*nd International Workshop on Intelligent Systems and Applications* (*ISA* 10), Wuhan, 22-23 May 2010, 1-5.

[10] Wu, J., Ping, L., Ge, X., Wang, Y. and Fu, J. (2012) Cloud Storage as the Infrastructure of Cloud Computing. *International Conference on Intelligent Computing and Cognitive Informatics* (*ICICCI* 10), Kuala Lumpur, 22-23 June 2010, 380-383.

[11] Zhang, D., Sun, F., Cheng, X. and Liu, C. (2011) Research on Hadoop-Based Enterprise File Cloud Storage System. 3*rd International Conference on Awareness Science and Technology* (*iCAST* 11), Dalian, 27-30 September 2011, 434-437.

[12] Zhang, X., Du, H., Chen, J., Lin, Y. and Zeng, L. (2011) Ensure Data Security in Cloud Storage. *International Conference on Network Computing and Information Security* (*NCIS* 11), Guilin, 14-15 May 2011, 284-287.

[13] Koletka, R. and Hutchison, A. (2011) An Architecture for Secure Searchable Cloud Storage. *International Conference on Information Security South Africa* (*ISSA* 11), Johannesburg, 15-17 August 2011, 1-7.

[14] Hao, L. and Han, D. (2011) The Study and Design on Secure-Cloud Storage System. *International Conference on Electrical and Control Engineering* (*ICECE* 11), Yichang, 16-18 September 2011, 5126-5129.

[15] Feel, H.T.A. and Khafagy, M.H. (2011) OCSS: Ontology Cloud Storage System. *First International Symposium on Network Cloud Computing and Applications* (*NCCA* 11), Toulouse, 21-23 November 2011, 9-13.

[16] Srinivasan, J., Wei, W., Ma, X. and Yu, T. (2011) EMFS: Email-Based Personal Cloud Storage. 6*th International Conference on Networking, Architecture and Storage* (*NAS* 11), Dalian, 28-30 July 2011, 248-257.

[17] He, Q., Li, Z. and Zhang, X. (2010) Study on Cloud Storage System Based on Distributed Storage Systems. *International Conference on Computational and Information Sciences* (*ICCIS* 11), Chengdu, 17-19 December 2010, 1332-1335.

[18] Liu, W. (2012) Research on Cloud Computing Security Problem and Strategy. 2*nd International Conference on Consumer Electronics, Communications and Networks* (*CECNet* 12), Yichang, 21-23 April 2012, 1216-1219.