



AENSI Journals

Australian Journal of Basic and Applied Sciences

ISSN:1991-8178

Journal home page: www.ajbasweb.com



An Amplified Self-Destructing Data System derived from Active Storage for Grid or Cloud Services

¹Victor Jose, M. and ²V. Seenivasagam

¹Department of CSE, Noorul Islam Centre for Higher Education, Kumaracoil, Tamil Nadu, India

²Department of CSE, National Engineering College, Kovilpatti, Tamil Nadu, India

ARTICLE INFO

Article history:

Received 10 October 2014

Received in revised form

22 November 2014

Accepted 28 November 2014

Available online 1 December 2014

Keywords:

Grid privacy, Active storage framework, Self-destructing data system, Distributed load pairing technique, Load balancing.

ABSTRACT

This paper describes a new self-destructing data system with an efficient load balancing mechanism so that it meets the challenges of privacy management in user data and also provides distributed load managing schemes in a distributed grid environment. To increase the user's data privacy an active storage framework of self-destructing data system is used. Without user intervention, all the data and their copies become destructed or unreadable after a user-specified time and the decryption key is also destructed. The load balancing mechanism called distributed load pairing technique has been used in the system to distribute the files as uniformly as possible among the distributed nodes such that no node manages an excessive number of files and to reduce network traffic by rebalancing the loads of nodes as much as possible. The uploading and downloading algorithms of distributed load pairing is explained and self-destructing data system uploading procedure has also been explained.

© 2014 AENSI Publisher All rights reserved.

To Cite This Article: Victor Jose, M. and V. Seenivasagam., An Amplified Self-Destructing Data System derived from Active Storage for Grid or Cloud Services. *Aust. J. Basic & Appl. Sci.*, 8(18): 122-127, 2014

INTRODUCTION

Grid/ Cloud computing is the delivery of computing services over the heterogeneous networks. The Grid services allow individuals and businesses to use resources that are managed by service providers. The grid computing model allows an access to information and computer resources from anywhere a network connection is available. Grid computing provides a shared pool of resources, including data storage space, networks, computer processing power, and specialized corporate and user applications. In general, a grid can provide a consistent way to balance the loads on a wider federation of resources (Yu and Buyya, 2005) such as CPU, storage, and other types of resources on a grid. In a grid environment, secure resource sharing is a challenging problem among the various online and offline attacks. Since security is a much more important factor in planning and maintaining a grid than in conventional distributed computing. Furthermore, it is important to understand the issues involved in data privacy and security in the grid. Now, many of the organizations such as government sector, private sectors and scientists for research work have been concentrating on the grid computing environments as shown in Fig. 1. Data privacy concerns are present, wherever personal data or sensitive data are stored in the digital form. The Control disclosure should be the reason for privacy issue. Sharing personal information is the challenge issue in data privacy. In grid, heterogeneous resources with different privacy policy are interconnected and data is shared. The Privacy protection technology in IT systems is communication and enforcement. Web Service Privacy may specify how privacy policy information can be embedded in the SOAP envelope. The Shamir's algorithm (Shamir, 1979) is used to distribute equally divided key (Lingfang et al., 2013) to users in the object storage system. A proof-of-concept (Lingfang et al., 2013) prototype is also implemented.

MATERIALS AND METHODS

A Self-Destructing Data System Based on Active Storage Framework (SeDas) for protecting user data's privacy is proposed by (Lingfang et al., 2013). Without user intervention, all data and their copies become destructed after a user-specified time. In addition, the decryption key is also destructed. A novel approach is the leveraging of active storage framework based on T10 OSD (Object-based Storage Device) standard. The feasibility of approach by presenting SeDas, a proof-of-concept prototype (Lingfang et al., 2013) based on

Corresponding Author: Victor Jose, M., Department of CSE, Noorul Islam Centre for Higher Education, Kumaracoil, Tamil Nadu, India.
E-mail: mvictorjose@yahoo.com, Tel: +91 9487112584, Fax: +91 4651 257266

object-based storage techniques. SeDas clearly inadequate in a large-scale, failure-prone environment because the cloud service providers is put under considerable workload that is linearly scaled with the system size, and may thus become the performance bottleneck and the single point of failure occurs. The searchable symmetric encryption for secure storage and retrieval of sensitive information from cloud is proposed by (Rajeshkumar and Rubasoundar, 2014) to eliminate the problem of searching on encrypted data. In this two-round searchable encryption scheme with multi-keyword top-k retrieval scheme is used to eliminate the Server-side leakage of data privacy. In this encryption scheme the idea of vector space model and homomorphic encryption are used. The vector space model helps to provide sufficient search accuracy, and the homomorphic encryption enables the users to involve in the ranking while the majority of computing work is done on the server side. The problem of ensuring the integrity of data storage in cloud computing is studied by (Qian et al., 2011) and introduced a third party auditor to overcome the involvement of the client auditing, whether the data is stored properly or not, and verify the integrity of the dynamic data stored. In this protocol the existing proof of storage model has improved by manipulating the classic Merkle Hash Tree construction for block tag authentication and efficient multiple auditing tasks can be performed using bilinear aggregate signature technique. A review on the feasible security qualities by making use of multiple distinct clouds is done by (Jens-Matthias et al., 2013). Various separate architectures are introduced and discussed based on their security and privacy capabilities and prospects. From the examinations performed based on that one approach is using multi cloud approach with sound data encryption for both technical and regulatory requirements. The second approach is using homomorphic encryption with secure multiparty computation protocols for both technical security and regulatory compliance. A protocol to fulfill the goals of correctness, privacy, robust cheating resistance and high efficiency for outsourcing of matrix inversion computation to a malicious cloud is designed by (Xinyu et al., 2013). The idea to protect the privacy is converting the original matrix into encrypted matrix which is sent to the cloud, and then transforming the result into the original matrix. The challenges for parallel data processing in clouds is discussed by (Daniel and Odej, 2011) and proposed modified Nephele framework with dynamic resource provisioning of Infrastructure-as-a-Service (IaaS). The modified framework, perform extended evaluations of map reduce-inspired processing jobs on an IaaS cloud system. The intrusion detection system architecture for identifying inside and outside intrusions is proposed by (Hulisi, 2013). This architecture combines configuration and waiting time decisions and illustrate the solution procedure for waiting time and configuration decision under an optimal policy. It is suggested that configuration decision is more important than waiting time decision to decrease the cost of operating intrusion detection systems. A model that extends the incorporated idea of risk assessment process and trust of the system is proposed by (Nathalie and James, 2013). In this, if user's trust falls below a certain threshold, that privileges are identified and removed. This threshold is calculated based on the inference of unauthorized information. A mechanism for investigation intrusion detection is suggested by (Karen et al., 2012). This mechanism separates abnormal behavior from normal behavior using incorporated idea of hidden Markov model with k-means. A distributed Host based Intrusion Detection System (HIDS) for associate management to optimum cost is discussed by (Carol et al., 2012). Each HIDS evaluates false positive and false negative rate of its neighboring HIDS opinions using Bayesian learning, and aggregates these opinions using a Bayesian decision model. A detection method that identifies malicious analysts who crack decisions through intelligence reports are suggested by (Eugene et al., 2012). This method is based on each analyst's working style and degree of consistency of their activities. Based on the idea, normal hypothesize is generated. In this, inconsistency is caused by malicious action and it is identified easily. A community anomaly detection framework to identify insider threats using access logs are suggested by (You-Chen et al., 2012). It consists of two components such as relational pattern extraction for deriving community structures, and anomaly prediction, for determining when users deviated from communities. And it is extended into Meta community anomaly detection system for semantics of subjects to achieve improved performance.

Proposed work:

A new self-destructing data system is introduced with an efficient load balancing mechanism called Distributed Load Pairing Technique (DLPT), so that it meets the challenges of privacy management in user data and also provides distributed load managing schemes in a distributed environment on active storage framework. For the distribution of equally divided key to clients, Shamir's algorithm is used in object storage system. A proof-of-concept prototype is also implemented. It supports safety erasing files and random encryption keys stored. The technique also aims to reduce network traffic (or movement cost) caused by rebalancing the loads of nodes as much as possible.

The DLPT is used to maintain the load balancing techniques at the time of distribution of files to different cloud/ grid service providers. Whenever a key is assigned to the client, it is associated with a set of SeDas parameters including the client's secret key survival time parameter. Each client is associated with separate secrete keys; the administrator of the grid/cloud manages the file manager. It is also responsible for storing the

active objects into the server. After authenticated by the administrator, the file manger will create keys based on objects for the encryption and to the distribution of secret keys to the client.

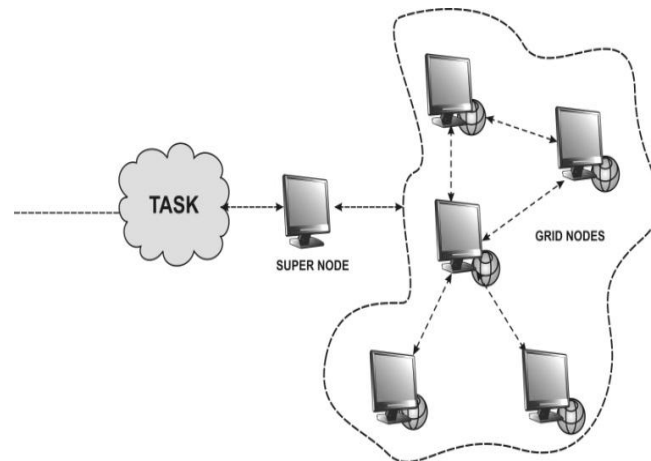


Fig. 1: Grid computing Environment.

Block Diagram:

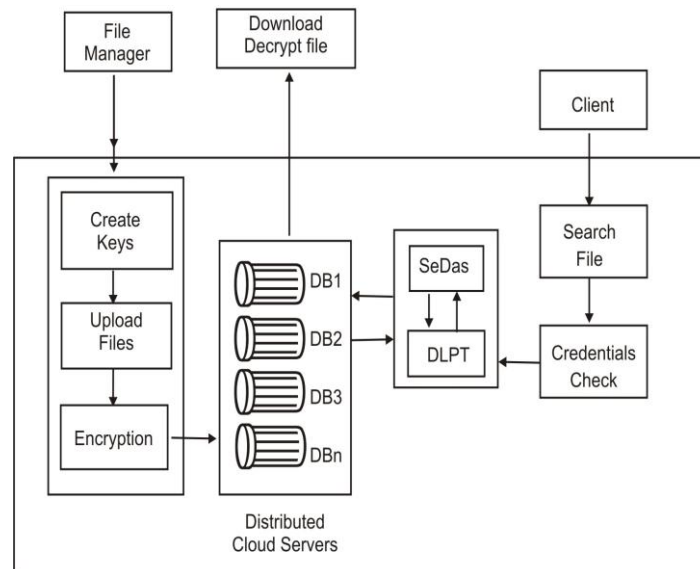


Fig. 2: Block Diagram of Proposed Work.

The file manager, then encrypt and upload the file to different active objects and stored in the storage server. The encrypted files are distributed to different Service Providers based on the load of each Service Provider. The DLPT is performed here for the purpose of distributing the load evenly to all Service Providers. The Service Provider with fewer loads is given the first priority and higher load is given the least priority. The client is self registered with the system and it must be authenticated by the file manager. The file manager authenticates the client and assigns a secret key to the client for the decryption of files. The authenticated client logs in to the system and creates SeDas parameters associated with their active storage object. The registered parameters will destruct the accessed file details and also decrypt the decryption keys after the survival time specified by the clients. When a file request is generated from the client the file manger views the request and processes the views request to the client. The clients then download the file using their secret key and decrypt them.

Algorithm for DLPT uploads operation

BEGIN

// count the no. of nodes as n

for i from 1 to n

//read the load of each node

Total node = k

```

//Repeat this for all k nodes
load size1 = s
load size2 = j
// pairing
If (s < j)
No of file = p
Upload file p to s.
else
Upload file p to j.
end If
//end loop
END
Algorithm for DLPT downloading operation
BEGIN
//count the no. of request as r
If (r = client need)
// categorize and evaluate
node size = n
load = 1
// based on properties
Ongoing respond = q
If (q = 0)
Ideal, hand over request,
else if (q = 1) start download
Wait
Stop responding
end if
end else if
END
SeDas uploading procedure
PROCEDURE Upload file (data, key, ttl)
data: data read from this file to be uploaded
key: data read from the key uploaded
ttl: time-to-live of the key
//encrypt the input data with key
BEGIN
Buffer=ENCRYPT (data, key);
Connect to a data storage server
If failed, then return fail
Create a file in the data storage server and writes buffer into it;
// k is the count of data servers in the SeDas system
Shared keys [1...k] =Shamir secret sharing split (n, k, key);
for i from 1 to k then
Connect to Ds[i];
If successful, then create-object (shared keys[i], ttl);
else
for j from 1 to i then
Delete key shares created before this one;
end for
return fail;
end if; end for
return successful;
END

```

RESULTS AND DISCUSSION

Grid/cloud services are becoming more and more important for people's life. People are more or less requested to submit or post some personal private information to the grid/cloud by the Internet. When people do this, they subjectively hope service providers will provide security policy to protect their data from leaking. So, other people will not invade their privacy. As people rely more and more on the internet and grid/cloud technology, security of their privacy takes more and more risks. When the data is being processed, transformed

and stored by the current computer system or the network must cache, copy or archive it. These copies are essential for systems and network. However, the people have no knowledge about these copies and cannot control them, so these copies may get their privacy leaked. On the other hand, their privacy also can be leaked through Service Providers negligence, hackers' intrusion or some legal actions. These problems present formidable challenges to protect people's privacy. Performance Evaluation

The result is compared with the existing native system and proposed new seDas system with the following operations.

- Upload operations.
- Download operation.

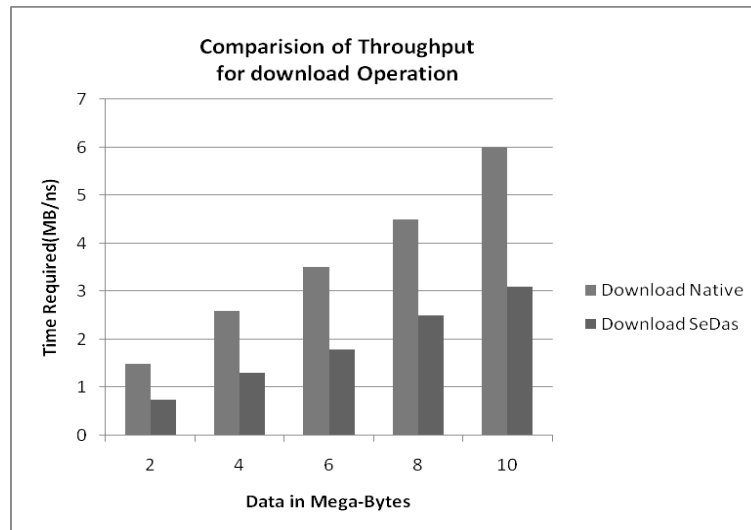


Fig. 3: Throughput for upload operations.

Fig 3 shows the throughput results for the proposed new scheme with the native scheme for uploading the data to the grid/cloud. The sample data of size in 2, 4, 6, 8 and 10 MB are used to test the case. The graph is plotted for throughput for data in MB versus through speed in nanoseconds. The raw data is shown in Table I.

Table I: Raw data values for upload operation.

File Size	2	4	6	8	10
Upload Native	2	3	4	5.2	5.9
Upload newSeDas	1	1.5	2	2.9	3.4

Fig 4 shows the throughput results for the proposed new scheme with the native scheme for downloading the data from the grid/cloud. The sample data of size in 2,4,6,8 and 10 MB are used to test the case. The graph is plotted for throughput for data in MB versus through speed in nanoseconds.

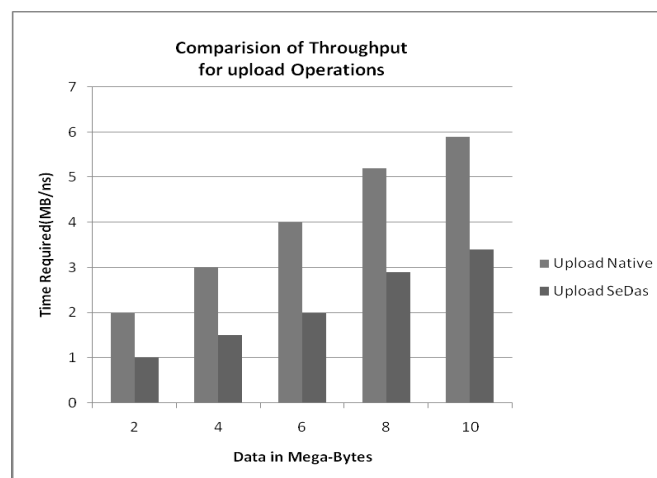


Fig. 4: Throughput for download operations.

The raw data is shown in Table II. From fig 3, the new SeDas reduces the throughput over the Native system by an average of 50% for the uploading. From fig 4, the new SeDas reduces the throughput over the Native system by an average of 50% for the downloading.

Table II: Raw data values for download operation.

File Size	2	4	6	8	10
Download Native	1.5	2.6	3.5	4.5	6
Download newSeDas	0.75	1.3	1.8	2.5	3.1

Conclusion:

This algorithm is a reversible or re-generative of the sin-NOT. The Data privacy has become increasingly important in the grid/cloud environment. This paper has introduced a new approach for protecting data privacy from attackers who retroactively obtain, through legal or other means, a user's stored data and private decryption keys. A novel aspect of this approach is the leveraging of the essential properties of active storage framework based on T10OSD standard. The feasibility of the approach is presenting new SeDas, a proof-of-concept prototype based on object-based storage techniques. The new SeDas causes sensitive information, such as account numbers, passwords and notes to irreversibly self-destruct, without any action on the user's part. The measurement and experimental security analysis shows that this scheme is the best and optimum approach. The current SeDas system will help the researchers with further valuable experience for grid or cloud services.

REFERENCES

- Carol J. Fung, Jie Zhang and Raouf Boutaba, 2012. Effective Acquaintance Management Based on Bayesian Learning for Distributed Intrusion Detection Networks. *IEEE Transactions on network and service management*, 9(3): 320-332.
- Daniel Warneke and Odej Kao, 2011. Exploiting Dynamic Resource Allocation for Efficient Parallel Data Processing in the Cloud. *IEEE Transactions on parallel and distributed systems*, 22(6): 985-987.
- Eugene Santos, Hien Nguyen, Fei Yu, Keum Joo Kim, Deqing Li, John T. Wilkinson, Adam Olson, Jacob Russell and Brittany Clark, 2012. Intelligence Analyses and the Insider Threat. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, 42(2): 331-347.
- Hulisi Ougt, 2013. The Configuration and Detection Strategies for Information Security Systems, *Computer and Mathematics with applications*, 65: 1234-1253.
- Jens-Matthias Bohli, Nils Gruschka, Meiko Jensen, Luigi Lo Iacono and Ninja Marnau, 2013. Security and Privacy Enhancing Multicloud Architectures. *IEEE Transactions on dependable and secure computing*, 10(4): 212-224.
- Karen A. Garcia, Raul Monroy, Luis A. Trejo, Carlos Mex-Perera and Eduardo Aguirre, 2012. Analyzing log Files for Postmortem Intrusion Detection. *IEEE Transactions on systems, man, and cybernetics-part c: applications and reviews*, 42(6): 1690-1704.
- Lingfang Zeng, Shibin Chen, Qingsong Wei and Dan Feng, 2013. A Self-Destructing Data System Based on Active Storage Framework. *IEEE Transactions on Magnetics*, 49(6): 2548-2554.
- Nathalie Baracaldo and James Joshi, 2013. An Adaptive Risk Management and Access Control Framework to Mitigate Insider Threats. *Computers and security*, 39: 237-254.
- Qian Wang, Cong Wang, Kui Ren and Wenjing Lou, 2011. Enabling Public Auditability and Data Dynamics for Storage Security in Cloud Computing. *IEEE Transactions on parallel and distributed systems*, 22(5): 847-859.
- Rajeshkumar, C. and K. Rubasoundar, 2014. Retrieval of encrypted cloud data using multi-keyword. *International Journal of Innovative Research in Computer and Communication Engineering*, 2(1): 2138-2144.
- Shamir, A., 1979. How to share a secret. *Communications of the ACM*, 22(11): 612-613.
- Xinyu Lei, Xiaofeng Liao, Tingwen Huang, Huaqing Li and Chunqiang Hu, 2013. Outsourcing Large Matrix Inversion Computation to a Public Cloud. *IEEE Transactions on cloud computing*, 1(1): 78-87.
- You Chen, Steve Nyemba and Bradley Malin, 2012. Detecting Anomalous Insiders in Collaborative Information Systems. *IEEE Transactions on dependable and secure computing*, 9(3): 332-334.
- Yu, J. and R. Buyya, 2005. A Taxonomy of Workflow Management Systems for Grid Computing. *Journal of Grid Computing*, 3(3-4): 171-200.