Modified Division and Replication of Data in the Cloud Computing Environment for Optimal Performance and Security (MDROPS)

Vahid Ahmadi*
Assistant Professor: Department of Computer and Electornic
Taali Institute of Higher Education
Qom,Iran
v.ahmadi@taali.ac.ir

Hossein Hassanpour
MA student: Department of Computer and Electornic
Taali Institute of Higher Education
Qom,Iran
H.hasanpour@taali.ac.ir

Abdolreza Rasouli Kenari
Assistant Professor: faculty of electronic and computer science
Qom university of technology,
Qom,Iran
rasouli@qut.ac.ir

***Abstract*—The tendency of organizations to use cloud services is increasing day by day, due to the economic benefits that can be contained, the adoption of every service of information security topic is regarded as one of the most critical primitive requirements. Hence, urgent need to outsourcing data as well as the significant growth of the hacker's knowledge may be considered as a major obstacle in the path of securing the cloud services. One of the proposed techniques to secure cloud computing using network platform, is to divide and replicate data. This method suffers from problems like disclosure of information through telephone (internet) tapping, and incompleteness of encryption of location's data storage algorithm.**

Therefore, in this paper, we have tried to present a way to optimize the performance and increase the level of security, by dividing and replicating data using a methodology with the encryption of location's respondents and finally, we aim at indicating the increase of efficiency and security with implementation of the algorithm in the context of data center networks which is simulated by CLOUDSIM software, and also comparing the algorithm using the same proposed approaches.

Keywords—Cloud Computing; Data Storage; Information Security; Computer Attacks; Transmission Line

I. INTRODUCTION

As it is obvious from its name, 'cloud' means a sort of internet with a hidden abstraction base for system users. We can mention on-demand, high availability and resource integrated services as the most significant features of cloud computing which bring about increase in acceptance level of these services within various organizations. In addition to these evident advantages, we can consider low setup cost, low managerial severity level for user and high flexibility against changes as some other notable characteristics. Meanwhile, security challenges are regarded as serious obstacles in the path of accepting progress of these services.

A compound cloud-based service includes varied nodes to put processing, storage and control affairs to action. In each compound system, the highest level of security will be determined based on the lowest existing level of security among different nodes. Hence, the process of measuring security level in cloud computing is not only limited to a single node, but also all existing nodes are capable of leading their neighbors to plentiful calamities. Virtual machines can share resources among system users as the main core of cloud systems. Regarding this fact that virtual machines are temporary and we can rent them to other users, information recovery sometimes leads to information disclosure of previous users. Furthermore, virtual machines monitoring level can be intentionally or unintentionally changed (by the attacker) in a virtual environment. In this way, obstacles ahead will be eliminated to provide attacker's accessibility to other virtual machines information [2].

In such scenarios, security mechanisms should significantly increase the attacker's effort to achieve information (even in the time of a successful attack) and information disclosure should be insignificant enough to stop disclosing total information security.

All mentioned problems have roots in data storage scope and some of them have been partly removed by DROPS methodology which has been evaluated in [3]. Nevertheless, we have to note that information transmission line can be listed as a right location for disclosing secured information as the result of applying this method.

The reminder of this paper will be dedicated to the previous efforts, methods and equations. In the third section, we elaborate on algorithm, information sending and receiving scenarios, and methods to hide server's location. Eventually, the fourth section includes the results and outcomes of applying algorithm and selected similar methods.

II. THE PREVIOUS SURVEYS

In this section, we will discuss prerequisites and corresponded topics to MDROPS methodology.

A. Attack types

1) Distributed Denial of service attack

In this kind of attack, the attacker is authorized to exclude machine and network resources from user's accessibility zone. However, DOS attack can be defined and motivated differently, in general, it is striving for temporary or permanent disconnection or procrastinating the process of serving clients by a connected host to the internet.

978-1-5090-5443-5/17/$31.00 ©2017 IEEE

2) Scan port attack

Scan port attack (or port scanning attack) is one of the most popular methods applied by the attackers to identify implemented services in a target host. All connected systems to the internet or internal networks will run determined and famous services. The attacker sends data in the form of TCP and UDP packages toward system ports via special means and then, by receiving responses from these ports can identify which ports are being used (in the other words, which ports are open). In this situation, when the attacker becomes aware of open ports, he will focus on these ports and seize target system based on existing weak points [1].

3) Information sniffing attack

Information sniffing attack is one of the most common attacks in computer network platforms. In this attack, the attacker continuously examines information transmission line via special means and in this way, he will access to transmitted information to the network platform. For resolving this problem, we use encryption protocols of communicational routes. In these routes, we can identify addresses of locations within files have been transmitted and the attacker can identify information storage location based on the above address [1].

We apply data division and replication method along with submerging the location of server in our proposed plan to avoid various types of existing attacks in a cloud platform.

III. Different Methods for Securing Clouds

A. Data migration to cloud via Iris system file

The technique of data migration to cloud is done via Iris system file in [5] and then, a gateway is placed in organization based on Markle Tree method to ensure data comprehensiveness and data novelty [10]. When system undergoes a successful attack, this method poorly acts in decreasing the possibility of losing data. Furthermore, this method highly hinges on users location plan in the network in order to maintain information confidentiality.

B. Trusted third party along with an information encryption method

As it is mentioned in papers [7] and [8], we can apply trusted third party along with an information encryption method to enhance confidence level in data authentication, comprehensiveness, confidentiality and transmitting information within other sides of a trusted third party. Paper [9] demonstrates that some plans cannot be secure against data manipulation and their loss according to outcomes of virtualization and multi-tenant problems.

C. Security of encryption key

We have evaluated a secure and published location in the network platform in [10]. In this method, an encryption key is divided into n parts and then, it will be published in different sites (different locations in a cloud platform with the capability of key maintenance). In this method, the network is divided into clusters and then, the main part of each cluster which allocates keys to internal components of clusters is selected. This plan only discusses encryption key security and data is fully stored as one file in different locations via various keys. In this method, there is possibility of losing whole file when a successful attack happens.

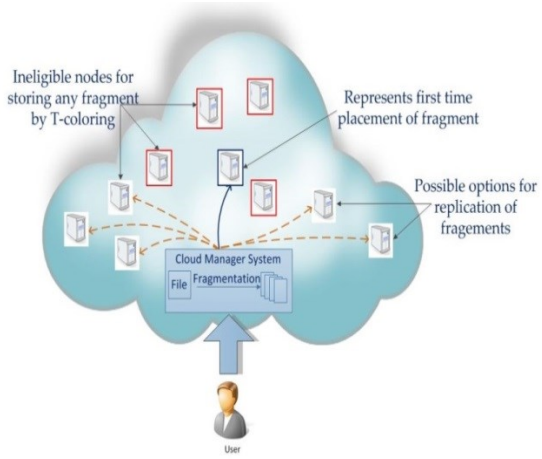

Fig. 1. DROPS Method Scenario

D. Data division and replication

In the proposed method by reference [3] information is divided into different segments and is stored in different locations regardless of encryption methods. This method discloses little information while faces a successful attack to system; because in this case, only system information which is a meaningless part of a file will be disclosed. The main problem of this method happens in the time of information transmission. In this scenario, the attacker can access to information sender addresses via information transmission line sniffing. In this way, the attacker can neutralize submergence algorithm of information location by recognizing responder node to user or transmission route of whole information and so the whole information is obtainable by a widespread attack to special nodes. Figure 1 has shown a view of DROPS method scenario.

IV. MDROPS Method

We have implemented our proposed system on an object-oriented basis. MDROPS evaluates information disclosure with the help of data division and using several nodes to store a file. In the reminder of this chapter, we comprehensively elaborate on the proposed method at first. Then, we fully examine each part and finally, we illustrate system model.

A. Definition of MDROPS method

In system infrastructure compound models, we evaluate security level based on its lowest available level among other nodes. In such infrastructures, if we apply a heterogeneous model in the process of combining systems, the security level will be enhanced; because the required effort for securing all systems will remarkably increase in this model and so, it is hardly possible to discover security errors of all systems. Storing the whole file in one point can introduce that point as a single point of failure in the network, so if we apply replication technologies for increasing system efficiency and confidence, this action may lead to more failure points in the network. Because a successful attack to each information storage can bring about information disclosure. Information security and

978-1-5090-5443-5/17/$31.00 ©2017 IEEE

disclosure are two essential factors within widespread systems and should be considered moderately so that a service level is not less than other service levels. Accordingly, in MDROPS methodology we divide a file into several meaningless segments and then expand it through network infrastructure so that each node includes one part and location of nodes is not distinguishable by the attacker, in the next step, we apply replication technology to handle parts. Placing one part in each node can increase security level. In this case, if we need information recovery in the process of transferring information, the amount of delay will increase and the attacker can also attack the whole file in the time of transferring information via sniffing communicational line. To do that, we send data in segmented form to avoid attacker's accessibility to the whole information via sniffing; because in the proposed method we don't use line encryption algorithms to increase system efficiency.

B. *Data Division*

Topics such as secrecy, confidence, multitenancy, avoiding information disclosure, destructive virtual machines and etc which are mentioned in [11, 12, 13] may decrease user's reliance on cloud computing infrastructure. We divide data to several meaningless segments in order to avoid information disclosure via information recovery by system's new tenant, restrain information disclosure by the attacker or destructive virtual machines and increase efficiency in applying system resources within MDROPS methodology.

A cloud system is made up of plentiful joint systems which are assembled heterogeneously or homogenously. We can apply discovered failure of a system to destroy other systems in the framework of homogenous systems. Accordingly, the required effort to attack subsequent nodes is less than the required effort to attack the first node. In comparison with homogenous systems, heterogeneous ones require the same level of effort for attacking the first node and subsequent nodes. Now if we store information as single files, destroying one file only requires attacking a special file [9,11].

A successful attack can intentionally or unintentionally arise from managers, plans, destructive virtual machines, and attacker's vulnerability. Accordingly, publishing customer data with the aim of processing and storing them necessitates investigation of a virtual infrastructure in order to provide secrecy and integrity before, after and during the process of running virtual machines [4].

In order to change files to parts, we consider two different methods; the first method put the obligation of determining file segmentation interval on the owner's shoulder. In this case, we argument that owner is the best candidate for producing tolerance threshold of parts; because he is aware of relevant realities as the owner of file. If the file owner doesn't determine data segmentation information, it will be possible to divide file into small parts. We try to increase user's effort in order to achieve data and decrease full information disclosure to zero in MDROPS methodology.

C. *File storage scenario in MDROPS methodology*

In MDROPS methodology, user sends file along with its division interval to cloud management system. After receiving information, the desired file will undergo the following operations:

a) Segmentation

b) Selecting the first circle of cloud nodes in which parts should be stored.

c) Selecting the second circle in order to publish parts within it

When we change files into parts, MDROPS methodology selects cloud nodes for placing parts based on the proposed criteria in order to store data. If we place all parts in nodes based on a descending order, this position will give a clue to the attacker and make him aware of other parts location. We apply T-Coloring algorithm which has been originally used to solve problem of channel determination to investigate different aspects of locating parts [14].

In order to achieve the mentioned objectives, we allocate different colors to nodes so that all nodes are uncolored and store open-color amount in their feature at first. Whenever a part is placed in a node, all its adjacent nodes allocate close-color amount to themselves. The operation of locating parts in nodes will be repeated in the same way based on T-coloring algorithm.

Algorithm (1) demonstrates part selection and replacement methodology within internal nodes. We often control part publication with the aim of increasing data availability, reliability and recovery time improvement.

D. *Sending file scenario in MDROPS methodology*

In DROPS methodology, at first the file is collected in the user responder node and then, it is transferred to user as a single file. As it is illustrated in chapter 2 the information transfer line is known as algorithm failure point. Lack of security in transferring information brings about full disclosure of information and its location address within network and this event leads algorithm to complete failure.

Now we try to store and send segmented file by applying storage submergence strategy and avoid disclosure of parts storage location via transfer platform sniffing.

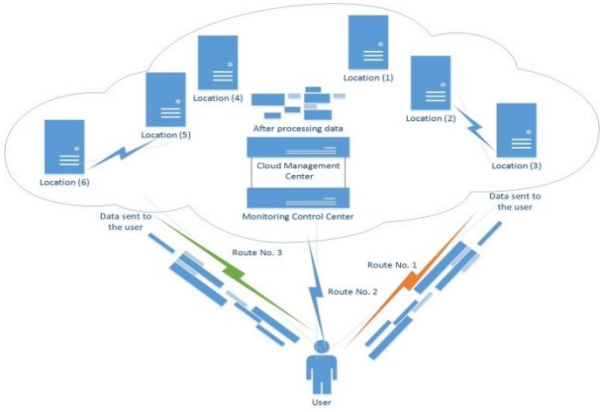

Fig. 2. Information transmission scenario

1) *submerging parts storage location*

A cloud infrastructure is made up of plentiful data centers which are attached together via high speed communicational

978-1-5090-5443-5/17/$31.00 ©2017 IEEE

links and virtualization tools are used in their internal layers in order to separate different segments.

We store parts in the existing virtualized file servers of different cloud datacenters. In this way, it is possible to access majority of various storage locations. Regarding this description and attending to the fact that there is a cache in file servers with temporary internal information to increase the velocity of responding process to user, we transfer the desired part to the existing cache in another file server via existing high-speed links before sending it in order to submerge parts storage location. Meanwhile, we use the existing interval address in the network which is not valid outside the network to avoid the attacker discovering the location of part storage server because of missing data center's internal addresses when he becomes aware of information transmission. This circumstance has been demonstrated in Figure2 in which location 5 includes file segment and it transfers information toward location 6 via transferring file segment to location6.

V. ILLUSTRATION OF RUNNING RESULTS

Let us evaluate an example with the topic of security. Imagine a cloud which is made up of M nodes along with a file which includes Z segments. Consider S as the number of successful attacks to distinguished nodes. So S should be greater than Z. the possibility of containing all file segmentations by all invaded nodes is shown as p(s, z) which is computable by formula 1 obtained by [3].

$$P(s,z) = \frac{\binom{S}{Z}\binom{m-s}{s-z}}{\binom{m}{s}}$$

Formula1. The possibility of discovering a system which includes a part by the attacker

If we consider m=30, S=10 and Z=7, the value of P(s, z) will be 0.0046. However, if we consider M=50, S=20 and Z=15, we will have P(s, z) = 0.000046. By decreasing the amount of M, the possibility of discovering location increases. Therefore, we can argument that great value of M can destroy the possibility of discovering whole information. According to this fact, in a cloud system with thousands nodes the possibility of discovering a remarkable part of information by an attacker will decrease significantly. In comparison with DROPS methodology, MDROPS methodology because of containing internal transmission stage experiences a delay in sending data toward user. This delay is negligible because of internal link's high bandwidth and also transmission among virtual machines and these transmissions are only possible by replacing virtual machines monitors and addresses.

The reminder of this chapter will be allocated to the obtained results by running MDROPS methodology in three data center architecture platforms which are called Three tire, Fat tree, and Dcell [14, 15] in CLOUDSIM stimulation plans and applying A$^{\varepsilon}$ -star, Suboptimal A-star1(sa1), suboptimal A-star2(sa2) and greedy algorithm strategies which are known as fine-grained algorithms [16] and have been used in previous strategies.

It should be noted that these strategies have been used in DROPS by the authors of [3] and have been selected in order to provide the possibility of a precise comparison.

As chart 1 demonstrates, reproduction cost (RC) includes reduction in comparison with other methods.

Chart 2 shows reproduction reduction of files segments in comparison with other selected algorithms in a data center architecture platform.

In chart 3, the effect of MDROPS method on reading and writing rate has been shown. Of course, in this chart we neglected movement rate of system internal information (in order to consider internal cache).

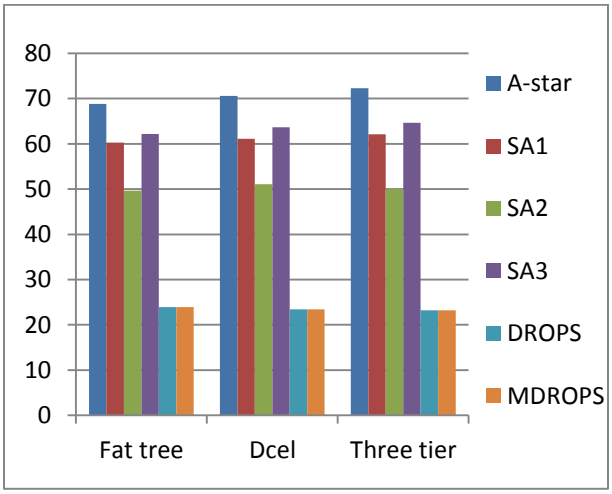

Chart. 1.Average reduction percentage of reproduction cost for increasing the number of parts

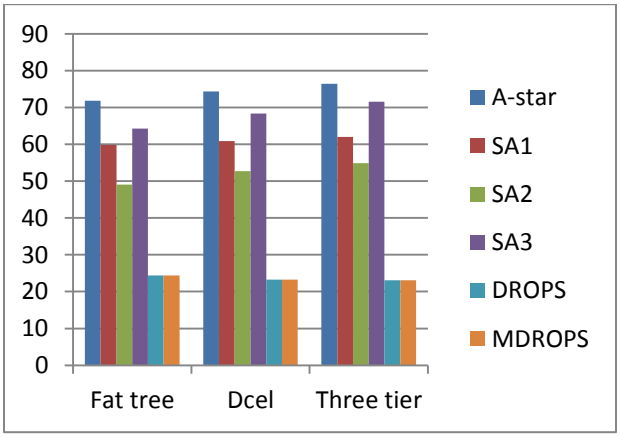

Chart. 2.Average reduction percentage of reproduction cost for increasing the number of nodes

978-1-5090-5443-5/17/$31.00 ©2017 IEEE

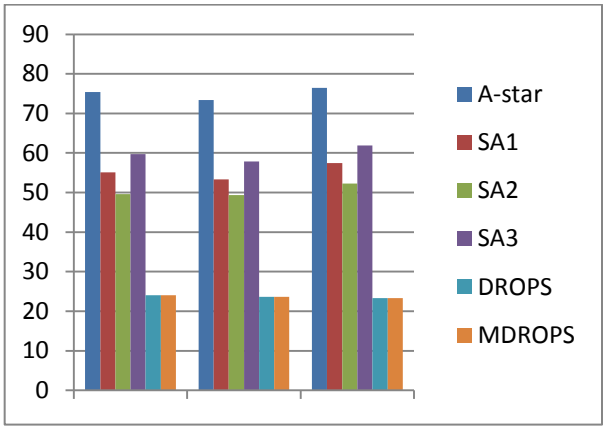

Chart. 3. Average reduction percentage of reproduction cost for increasing information reading and writing rate

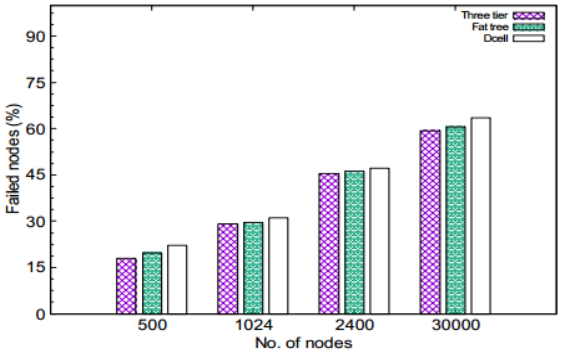

Chart. 4. Tolerance threshold of error in MDROPS

A. Investigating the effect of running algorithm within various datacenter platforms

As it has been shown in chart 4, when we increase number of file segments, the tolerance threshold of error is the same in all three data center architectures (Three tire, Fat tree and Dcell) in MDROPS and this case demonstrates algorithm capabilities while running on different architectures of cloud computing data center infrastructures.

VI. Practical Comparison of the Proposed Method with DROPS Method

In the proposed method, we tried to resolve problem of discovering information storage by the attacker - which can neutralize all submergence stages by discovering the above locations with the help of DROPS - via sniffing communicational line. In order to apply this advantage and to submerge data storage locations from attacker's informational sniffing, we have used internal server caches with different addresses and different storage locations in each segment. We have also applied changing addresses instead of transferring information to the desired segment. Applying and simulating these methods in CLOUDSIM simulator have been illustrated in charts 1, 2 and 3. In comparison with DROPS, this method registers similar rate of measured criteria in addition to resolve the above security problem.

VII. Conclusion

Cloud processing is regarded as one of the developing technologies and from this side, it faces plentiful security outcomes. It is evident that cloud computing chief service providers such as Google, Amazon and etc are still undergoing plentiful security challenges and there is no special method to overcome these challenges. In order to resolve a security challenge, we have to pay attention to service efficiency; accordingly, we try to secure storage platform and information transmission with the help of dividing data to meaningless segments, repeating them in node platform, responder submerging strategy and sending segmented information in this paper. As we mentioned above, in comparison with other existing algorithms MDROPS methodology demonstrates higher efficiency rates. It also decreases parts waiting time in queue, error rate, Energy waste rate, the occupied volume of communicational line by sending segmented information toward user.

References

[1] K.Hashizume, D.G.Rosado, E.Fernndez-Medina, and E.B.Fernandez, “An analysis of security issues for cloud computing,” Journal of Internet Services and Applications, Vol. 4, No. 1, 2013, pp. 1-13.

[2] U.A.Kashif, Z.A.Memon, A.R.Balouch, J.A.Chandi, ”Distributed Trust Protocol for IaaS Cloud Computing,” Proceedings of 2015 12th International Bhurban Conference on Applied Sciences & Technology Ubaidullah,2015,pp.275-279

[3] M.Ali, K.Bilal, S.U.Khan, B.Veeravalli, K.Li, A.Y. Zomaya, “DROPS: Division and Replication of Data in Cloud for Optimal Performance and Security,” IEEE Transactions on Cloud Computing, 2015.

[4] D.Sun, G.Chang, L.Sun, and X.Wang, “Surveying and analyzing security , privacy and trust issues in cloud computing environments”Procedia Engineering ,Vol. 15, 2011, pp . 2852-2856

[5] A.Juels, A.Opera, “New approaches to security and availability for cloud data” Communications of the ACM, Vol.56, No. 2, 2013, pp. 64-73.

[6] W.K.Hale, “Frequency assignment: Theory and applications,”Proceedings of the IEEE, Vol. 68, No. 12, 1980, pp. 1497-1514.

[7] M.Tu, P.Li, Q.Ma, I-L.Yen, and F.B.Bastani, “On the optimal placement of secure data objects over Internet,”In Proceedings of 19th IEEE International Parallel and Distributed Processing Symposium, 2005, pp. 10-14.

[8] D.Zissis, D.Lekkas,“Addressing cloud computing security issues” Future Generation Computer systems , Vol. 28, No. 3,2012, pp. 583-592.

[9] A.Mei, L.V.Mancini, S.Jajodia, “Secure dynamic fragment and replica allocation in large-scale distributed file systems,” IEEE Transactions on Parallel and Distributed Systems, Vol. 14, No. 9, 2003, pp. 885-896.

[10] Y.Tang, P.P.Lee, J.C.S.Lui, R.Perlman, “Secure overlay cloud storage with access control and assured deletion,” IEEE Transactions on Dependable and Secure Computing, Vol. 9, No. 6, Nov. 2012, pp. 903-916.

[11] A.Fox, R.Griffith, A.Joseph, R.Katz, A.Konwinski, G.Lee, D.Patterson, A.Rabkin, ,I.Stoica, "Above the clouds :A Berkeley view of cloud computing," Dept. Electrical Eng and comput Sciences, University of California,Berkeley,Rep. UCB IEECS,vol. 28,2009.

[12] T.Ristenpart, E.Tromer, H. Shacham, S.Savage, "Hey,you,get off of my cloud:exploring infonnation leak agein third–party compute clouds, " in Proceeding sof the 16th ACM conferenceon Computer and communications security,2009,pp.199-212.

978-1-5090-5443-5/17/$31.00 ©2017 IEEE

[13] K.Hashizume,D.G.Rosado,E.Fernandez Medina, E.B.Fernandez, "Ananalys is of security issues for cloud computing, " Journal of Internet Services and Applications,vol. 4,pp. 5, 2013

[14] K.Bilal, S.U.Khan, L.Zhang, H.Li, K.Hayat, S. A. Madani, N. Min-Allah, L.Wang, D.Chen, M.Iqbal, C. Z. Xu, and A.Y.Zomaya, "Quantitative comparisons of the state of the art data center architectures," Concurrency and Computation: Practice and Experience, Vol. 25, No. 12, 2013, pp. 1771-1783.

[15] K.Bilal, M. Manzano, S.U.Khan, E.Calle, K.Li, A.Zomaya, "On the characterization of the structural robustness of data center networks," IEEE Transactions on Cloud Computing, Vol. 1, No. 1, 2013, pp. 64-77.

[16] J.J.Wylie, M.Bakkaloglu, V.Pandurangan, M. W. Bigrigg,S. Oguz, K.Tew, C.Williams, G.R.Ganger, and P.K.Khosla,"Selecting the right data distribution scheme for a survivable storage system," Carnegie Mellon University, Technical Report CMU-CS-01-120, May 2001.

978-1-5090-5443-5/17/$31.00 ©2017 IEEE