p-ISSN: 2301-5373
e-ISSN: 2654-5101

Multistage Approach for Securing Data of Returned Disk On Data Center Environment

Lompo Ramos Emakarim[a1]
[a]Fakultas Teknik, Universitas Cokroaminoto Makassar
Jalan Perintis Kemerdekaan No.7, Indonesia
[1]lompo_ramos@ucm-si.ac.id (Corresponding author)

Abstract

Once disk was either proactively or correctively replaced in a data center, it still contained data. With warranty or contractual agreement, the disk will be sent back to disk vendor outside of data center. Data with sensitive and confidential information will be risky leaved on disk sent outside data center without any effort to anticipate. By using sophisticated and latest technique, this information can be extracted for misused or criminal activity. We propose RAID striping, disk encryption, and dedicated shredding machine as multistage approach to secure the information. RAID striping will augment data while encryption will scramble information. Dedicated shredding machine will erase data by implementing simple machine and universal adaptor. By applying these combined multistage approaches, stakeholder will be no worried about the replaced disk sent outside their data center. By applying these three methods, accessible replaced disk's data will be meaningless, unreadable then erased. Hence leaving disk with useless bits. For customer's view on inaccessible disk, RAID striping and encryption will be more enough without erasing due to limited access on it. This approach shows three of four simulation which data on disk sent outside data center is secured.

Keywords*—Secured Data, Replaced Disk, Data Center, RAID, Shredding, Encryption*

1. Introduction

Data center is a place with facilities where data of any institutions are maintained in computers and related equipment [1]. It retains the data on disks. Disks depend on the vendor and are managed accordingly so when a disk failed, the data on it is then relocated to another normal spared disk or reconstructed to a new disk [5]. For a company having service contract or warranty [3], once sparing data is completed the failed disk is replaced then returned to the vendor then to manufacturer to be analyzed and considered to be recycled. The returned disk may still contain data which then put the data in risk to be obtained by someone outside the data center.

According to [2] disk is defined to be failed when the disk is deemed to be replaced by data center technician because the disk cannot operate properly. A disk can also be replaced by using predictive analysis such ASLDP even before the disk totally failed [18]. It means the data on the disk is still recoverable by any forensic tool or low-level access [7][8][9]. Even if the data on file system is already deleted before replaced it is still recoverable since the file deletions only unlink the data [7]. Unlinking data only deletes the metadata on the disk but the data itself is still there.

In this paper we describe three stages approach with one stage modified to secure data on the disk with minimum possibility any data can be recovered.

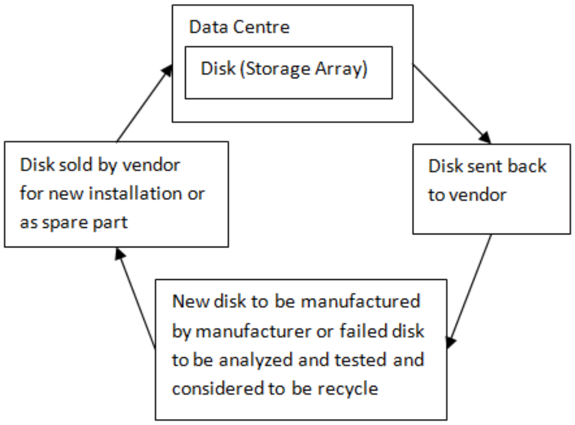
Data Centre
Disk (Storage Array)
Disk sent back to vendor
New disk to be manufactured by manufacturer or failed disk to be analyzed and tested and considered to be recycle
Disk sold by vendor for new installation or as spare part

Figure 1. Disk Cycle Based on Service Contract or Warranty

2. Methods

Multistage approach that we propose is combining RAID striping data, disk encryption, and dedicated shredding schema. We then simulate a word in this approach whether the word can be recovered after the three stages. RAID Striping data is simulated by separating the word into two pieces. Disk Encryption is simulated by using DES Algorithm with cipher block chaining mode. Dedicated shredding schema is simulated by replacing all the bits with bits representing word A repeatedly.

2.1. RAID Striping Data

Basically, RAID technology is used to make \the storage devices reliable and giving I/O performance. Either by using RAID 0, RAID 1, RAID10, RAID5, or RAID 6, users have options of level of availability and level of performance. What needs to be considered from its security concern in our discussions is the parity concept in the RAID technology. From [4][6] we can see striping fragments and spreads the data on several disks. On user perspective, it can be seen a comprehensive data located on a LUN, but the LUN itself is imported from fragmented logical device which is united from separated locations on several disks. It is illustrated on figure2.

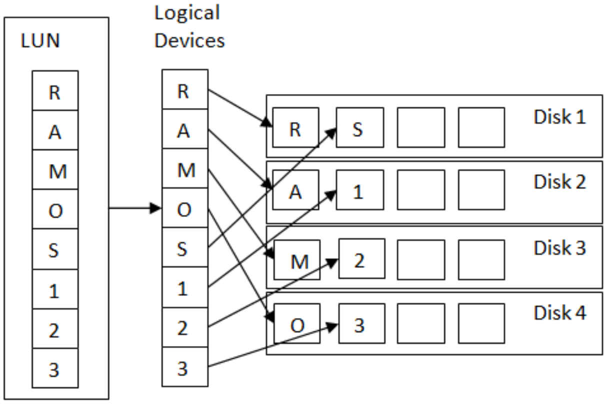

Figure 2. Data Striping Illustration

Based on figure 2, we can see the comprehensive-meaningful-usable data from user perspective is RAMOS123. However, this RAMOS123 is augmented to small parts and spread across Disk 1 to Disk 4. Each disk contains data which should be meaningless and useless from user perspective and not representing the main information. Hence once one of 4 disks get failed then replaced and sent back outside data center, the stakeholder no need to worry much about its confidential data on the failed disk. This shows us a RAID technology not only gives reliability and performance but also increases security of confidential data.

One thing that need to be aware is if the comprehensive data size is smaller than stripe size on the disk. For example, if a file's size is only 20KB and the stripe size is 32KB then the file is enough to be located on one disk and no need to be spread. It puts the file on risk to be recovered from outside the data center. Both username and password are examples file with small size. Figure 3 draws this condition. On a condition where disk 2 failed then it will contain full representative information of file 4. If file 4 is a confidential or privacy or critical data, then returning disk 2 outside data center is risky.

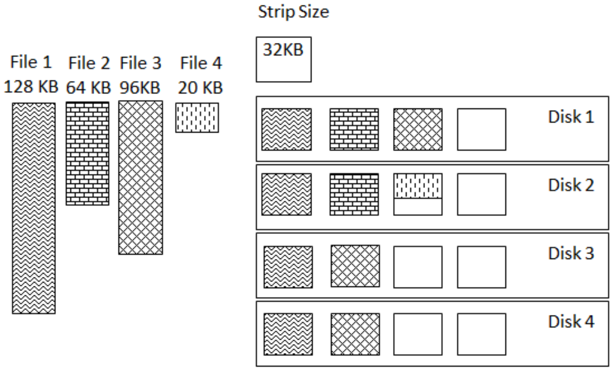

Figure 3. File Fits Strip Size

2.2. Full Disk Encryption

Full disk encryption is encrypting the whole hard drive. Full disk encryption can be performed through software or hardware [13]. For hardware-based encryption, many storage array providers will equip this encryption feature on their product and can be used by customer by purchasing license of it [10][11][12]. For this hardware based, they will keep the keys files on their service processor and shared or local memory [10][12]. Generally, by encrypting disk, the original information couldn't be taken by any person outside data center if they didn't know the algorithm or decryption key. Key for encrypting disk can be different for each disk and for safety it can be destroyed once the disk is replaced and sent back outside data center [12] as illustrated on figure 4.

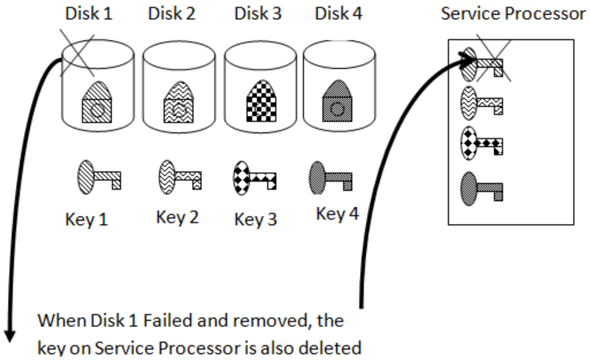
Disk 1
Disk 2
Disk 3
Disk 4
Service Processor
Key 1
Key 2
Key 3
Key 4
When Disk 1 Failed and removed, the key on Service Processor is also deleted

Figure 4. Encryption schema and disk removal procedure

2.3. Dedicated Shredding Schema

Shredding is a process of disk sanitation by erasing data using wiping algorithm [14]. Simple concept of sanitation is by resetting all bits on disk to 0 ,1, or combination of both as illustrated on figure 5. Such algorithm is NIST, Guttman, NISPOM,and DoD 5220.22-M [14]. On [16] *purgefs* is used which based on FiST language and can be used on any OS to overwrite data. Storage array vendor provides this feature as well by selecting any format of wiping on volume selected [15]. Off course, this can be done if the disk condition is still accessible. Not all disks were inaccessible when they were replaced. Some disks were indicated failed when exceeding S.M.A.R.T parameters and producing alerts [18]. Thus, the disk is proactively replaced although it is still accessible. By using this shredding, we can erase data contained on it and as a last approach before the disk is sent outside data center.

One issue to consider is that customer itself must convince that the disk is fully erased. Because of it, we propose that stakeholder of a data center can provide a small and simple machine that can shred the replaced disk. They then can choose any algorithm published or using their own to convince that the disk is safe to be sent outside. A low-end server or even a cheap workstation can be chosen. The important thing is an adaptor that connects the machine and the disk, so that accessible disk can be read. The universal adaptor needs to be discussed on separated research which can interface any port and protocol of disk. By using this dedicated machine concept, we will get benefit that we will not utilize storage array resource and of course no need to purchase disk shredding license and customer can see their self of the content of the disk

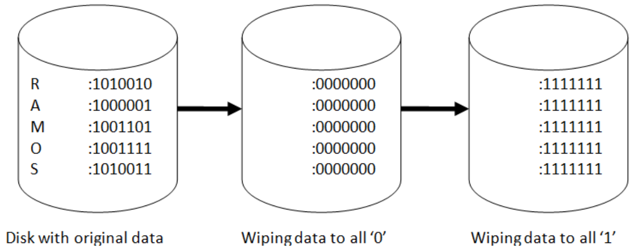

Figure 5. Shredding schema

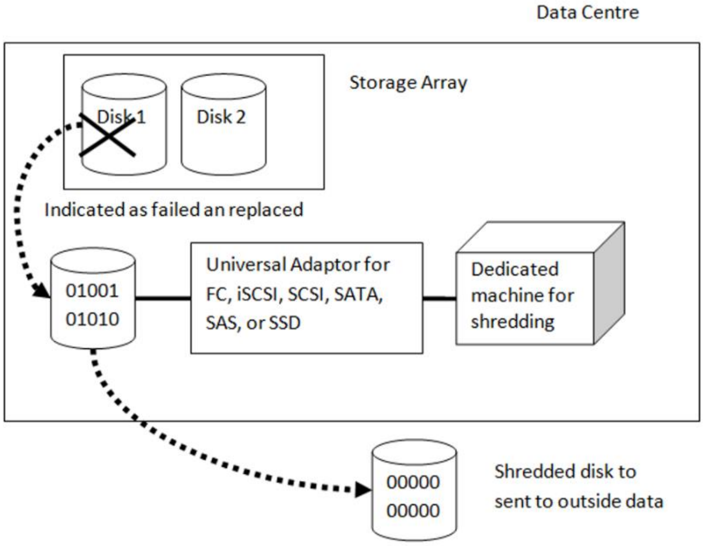
Data Centre
Storage Array
Disk 1
Disk 2
Indicated as failed an replaced
01001
01010
Universal Adaptor for FC, iSCSI, SCSI, SATA, SAS, or SSD
Dedicated machine for shredding
00000
00000
Shredded disk to sent to outside data

Figure 6. Dedicated Machines for Shredding

3. Results and Discussion

We simulate a password “Ramos812@” when saved on disk with four scenarios.

Table 1. Sequence of the first scenario

Sequences	Words	Bits
Original password	Ramos812@	01010010 01100001 01101101 01101111 01110011 00111000 00110001 00110010 01000000
Striped on 2nd disk	mos812@	01101101 01101111 01110011 00111000 00110001 00110010 01000000
Encrypt with DES	khAORtk/mY2FpCah7aIwKg==	01001110 01110010 01000011 01011000 01010000 01100110 00111001 01001101 01110011 01101000 01001101 01100001 01010111 01001101 01101111 01100111 01101001 01000111 01110101 00110010 01110110 01110111 00111101 00111101
Shredding with word A	AAAAAAAAAAAAAAAAAAAAAAAA	01000001 01000001 01000001 01000001 01000001 01000001 01000001 01000001 01000001 01000001 01000001 01000001 01000001 01000001 01000001 01000001 01000001 01000001 01000001 01000001 01000001 01000001 01000001 01000001
Decrypt with DES	ԁ�G;ԁ�G;	00011010 00011001 00000110 101000000LLL00 1111111111111101 01000111 00111011 00011010 00011001 00000110 10100000000 111111111111101 01000111 00111011

Table 2. Sequence of the second scenario

Sequences	Words	Bits
Original password	Ramos812@	01010010 01100001 01101101 01101111 01110011 00111000 00110001 00110010 01000000
Striped on 2nd disk	mos812@	01101101 01101111 01110011 00111000 00110001 00110010 01000000
Encrypt with DES	khAORtk/mY2FpCah7aIwKg==	01001110 01110010 01000011 01011000 01010000 01100110 00111001 01001101 01110011 01101000 01001101 01100001 01010111 01001101 01101111 01100111 01101001 01000111 01110101 00110010 01110110 01110111 00111101 00111101
Decrypt with DES	mos812@	01101101 01101111 01110011 00111000 00110001 00110010 01000000

The first scenario is when password Ramos812@ is striped into 2 disks. First disk holds word “Ra” and second disk holds word “mos812@”. The second disk is then indicated will be failed after some times and will be proactively replaced with a new disk. As the encryption is implemented the word “mos812@” will be encrypted as listed on table 1. Before the disk sent outside data center, dedicated shredding wipes the bits as the disk is still accessible. When someone is trying to decrypt data on the disk outside data center he will get error even though he has encryption key.

The Second scenario is when the second disk is failed and inaccessible. Then shredding cannot be implemented as listed on table 2. When someone is able to make the disk accessible and try to recover the data then decrypt the disk he will only get word “mos812@” as a wrong password.

Third scenario is when the password is not striped due to the small size of the words. One strip is enough to hold the whole password. The disk is still accessible before proactively replaced. Hence, Shredding is able to be implemented. The decryption then will get error as first scenario.

Table 3. Sequence of the third scenario

Sequences	Words	Bits
Original password	Ramos812@	01010010 01100001 01101101 01101111 01110011 00111000 00110001 00110010 01000000
Striped on one disk	Ramos812@	01101101 01101111 01110011 00111000 00110001 00110010 01000000
Encrypt with DES	XAw8uXIZJUJHa0dPwT/MBEh5Fui1eZ49	01011000 01000001 01110111 00111000 01110101 01011000 01001001 01011010 01001010 01010101 01001010 01001000 01100001 00110000 01100100 01010000 01110111 01010100 00101111 01001101 01000010 01000101 01101000 00110101 01000110 01110101 01101001 00110001 01100101 01011010 00110100 00111001
Shredding with word A	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	01000001 01000001 01000001 01000001 01000001 01000001 01000001 01000001 01000001 01000001 01000001 01000001 01000001 01000001 01000001 01000001 01000001 01000001 01000001 01000001 01000001 01000001 01000001 01000001 01000001 01000001 01000001 01000001 01000001 01000001 01000001
Decrypt with DES	d�G;d�G;	00011010 00011001 00000110 101000000LLL00 1111111111111101 01000111 00111011 00011010 00011001 00000110 10100000000 111111111111101 01000111 00111011

The fourth scenario is when the third scenario is not running when the disk is inaccessible before replaced and shredded. Only by having the encryption key and the ability to make the failed disk accessible, someone can gain the password.

Table 4. Sequence of the fourth scenario

Sequences	Words	Bits
Original password	Ramos812@	01010010 01100001 01101101 01101111 01110011 00111000 00110001 00110010 01000000
Striped on one disk	Ramos812@	01101101 01101111 01110011 00111000 00110001 00110010 01000000
Encrypt with DES	XAw8uXIZJUJHa0dPwT/MBEh5Fui1eZ49	01011000 01000001 01110111 00111000 01110101 01011000 01001001 01011010 01001010 01010101 01001010 01001000 01100001 00110000 01100100 01010000 01110111 01010100 00101111 01001101 01000010 01000101 01101000 00110101 01000110 01110101 01101001 00110001 01100101 01011010 00110100 00111001
Decrypt with DES	Ramos812@	01101101 01101111 01110011 00111000 00110001 00110010 01000000

From four scenarios above, the password is secured on three scenarios. One remaining scenario left the password vulnerable decrypted only if four conditions are fulfilled together i.e., first the password located in one disk only, second the encryption key is gained, third the disk is inaccessible before leaving data center, and fourth someone can make the disk accessible outside data center.
Summary of the approach is shown is table 5

Table 5 Three approaches combination

<table>
<tr><th>RAID striping</th><th>Encryption disk</th><th>Dedicated shredding</th></tr>
<tr><td>not secure small size data</td><td>data not secure if the key is stolen</td><td>only for customer accessible disk</td></tr>
<tr><td colspan="2">approach before the disk become inaccessible</td><td>cover RAID and Encryption, not utilize storage array resource and license</td></tr>
</table>

4. Conclusion

This approach shows three of four scenarios which data on disk sent outside data center is secured. Before disk is inaccessible, dedicated shredding will be the last powerful approach to erase the data. However, after disk is inaccessible, data can be previously secured by using RAID and encryption disk. With this multistage approach disk, disk will be securely sent outside data center. A data sensitive related institution such as Banking, Government, or Mass People Database will secure their confidential data contained on a replaced disk.

This multistage approach is simulated only. RAID is already implemented in storage array by storage vendors from performance and availability perspective. Encryption is as well from security perscpective. Dedicated shredding machine is separated from vendor of storage array. This machine is an opportunity to be implemented.

REFERENCES

[1] Pawlish, Michael et al. The Greening of Data Centers with Cloud Technology - In International Journal of Cloud Applications and Computing. 5. 1-23. 10.4018/IJCAC.2015100101. 2015
[2] Sidi Lu, Bing Luo et al. Making Disk Failure Predictions SMARTer!. Proceedings of the 18th USENIX Conference on File and Storage Technologies. 2020
[3] HPE Storage Global Limited Warranty and Technical Support. Part Number: P01457-402. Hewlett Packard Enterprise. 2020
[4] HPE MSA 1060/2060/2062 Storage Management Guide. Part Number: R0Q73-90009. Hewlett Packard Enterprise. Edition September 2020
[5] Ramkumar, M P et al. Recovery of Disk Failure in RAID-5 Using Disk Replacement Algorithm. International Conference on Innovations in Engineering and Technology. Volume: 3. 2014
[6] Rahman, P A and Shavier, G D'K Novikova Freyre. Reliability model of disk arrays RAID-5 with data striping. IOP Conference Series Materials Science and Engineering 327(2):022087. 2018
[7] Reardon, Joel et all. SoK: Secure Data Deletion. IEEE Symposium on Security and Privacy. 2013
[8] Reardon, Joel. Robust Key Management for Secure Data Deletion. Springer International Publishing. 10.1007/978-3-319-28778-2_11. 2016
[9] OliveiraJr, Edson et al. Towards a conceptual model for promoting digital forensics experiments. Forensic Science International: Digital Investigation. 35. 301014. 10.1016/j.fsidi.2020.301014. 2020
[10] HPE XP7 Encryption User Guide. Hewlett Packard Enterprise. 2019
[11] Hitachi Virtual Storage Platform 5000 Series SVOS RF 9.8.6 System Administrator Guide. Hitachi. 2023
[12] Dell PowerMax Family Security Configuration Guide: PowerMaxOS 10. DellEMC. 2023
[13] Hasan, Shiza et al. Full Disk Encryption: A Comparison on Data Management Attributes. In Proceedings of the 2nd International Conference on Information System and Data Mining (ICISDM '18). Association for Computing Machinery, New York, NY, USA, 39–43. https://doi.org/10.1145/3206098.3206118. 2018
[14] Nasaruddin, Siti Hajar et al. Securing Data with Hard Disk Shredding. International Journal of Computer Science and Electronics Engineering (IJCSEE) Volume 1, Issue 3, 2013.
[15] HPE XP7 Volume Shredder for Open and Mainframe Systems User Guide. Hewlett Packard Enterprise. 2018
[16] Joukov, Nikolai and Zadok, Erez. Adding Secure Deletion to Your Favorite File System. IEEE 3th Security in Storage Workshop, 2005.
[17] Gaber, Shiri et al. HDD failures from compound SMART attributes, Proceedings of the 10th ACM International Systems and Storage Conference. DOI: 10.1145/3078468.3081875. 2017
[18] Zhou, Yang et al. “A Disk Failure Prediction Method based on Active Semi-supervised Learning”, ACM Transactions on Storage Vol. 18 N0. 4, 2022