2015-12

Forward Error Correction for Storage Media: An Overview

Hassan, Kilavo

International Journal of Computer Science and Information Security (IJCSIS)

International Journal of Computer Science and Information Security (IJCSIS),
Vol. 13, No. 12, December 2015

Forward Error Correction for Storage Media: An Overview

Kilavo Hassan
Nelson Mandela African Institution of Science and Technology, School of Computational and Communication Science and Engineering, P.O. Box 447 Arusha, Tanzania

Kisangiri Michael
Nelson Mandela African Institution of Science and Technology, School of Computational and communication Science and Engineering, P.O. Box 447 Arusha, Tanzania

Salehe I. Mrutu
The University of Dodoma, College of Informatics and Virtual Education, P.O. Box 490 Dodoma

***Abstract*— As the adoption of Information and Communication Technology (ICT) tools in production and service rendering sectors increases, the demand for digital data storage with large storage capacity also increases. Higher storage media systems reliability and fault tolerance are among the key factors that the existing systems sometimes fail to meet and therefore, resulting into data loss. Forward error correction is one of the techniques applied to reduce the impact of data loss problem in digital data storage. This paper presents a survey conducted in different digital data storage companies in Dar es Salam, Tanzania. Data were collected and analyzed using Statistical Package for Social Sciences (SPSS). Secondary data were captured from user and manufacturer technical reports. It was revealed that data loss is still a predominant challenge in the digital data storage industry. Therefore, the study proposes the new storage media FEC model using locked convolutional encoder with the enhanced NTC-Viterbi decoder.**

***Index Terms*—Storage Media, FEC, NTC, Viterbi, RS**

I. Introduction

A storage media is any device on which data or information can be electronically stored, kept, and retrieved when needed [1, 2]. Storage media are devices that store user information (data) and application. The media can be categorized as optical data storage, magnetic hard disk drives, magnetic tape drives and flash disk drives[3]. The prevention of the record, information, and data stored electronically and the ability to retrieve them later or in future, require more than the safe keeping of the storage media [4]. There is little published scholarly work regarding the failure pattern of storage media and the key factors that affect their lifetime. Most of the information reported comes from storage manufacturers themselves. Since there is a huge amount of information stored and transferred among many storage media, data loss due to disk failure is a major issue that affects the reliability of the system. Reliability, fault tolerance and performance of storage media are the biggest concerns [3]. The demand for storage media increases everyday and it is estimated that over 90% of all the information and data produced in the world is being stored on magnetic media, and most of them are stored on hard disk drives [5]. On the other hand, data storage industry faces technological challenges due to the increase in demand and consistency. Currently, digital data have become one of the most important parts nowadays and create the increasing demand of data storage systems[6]. The world has entered the information era and any area of life is immersed in information.

The demand for digital data storage with large storage capacity, higher reliability and fault tolerance, easier accessibility, better scalability and cheap management poses a remarkable challenge on the storage industry[7]. Essentially, the demand for data storage becomes more and more growing and data storage system needs to have high density, short access time, and fast input and output transfer rate[6]. International Data Corporation (IDC) shows that data are expanding at approximately 50% to 80% per year and in other places the growth rate is closer to 100% per year. Some of the catalysts for this growth include databases, Enterprise Resource Planning (ERP), Supply Chain Management, E-procurement, Content Management, Data Mining, Customer Relationship Management (CRM), Electronic Document Management (EDM), emails, social media and multimedia[8, 9]. Seemingly, the demand for bigger, and faster memories has led to significant improvement in the conventional memory technologies like hard disk drives and optical disks. However, there is a strong proof that these two dimensional storage technologies are approaching their fundamental limits[2]. Almost all vital data are now stored on external disk storage subsystems. An average usable capability is approximately 2.18 peta bytes which is up to 12.8% year after year. Factors such as growth in storage requirement, larger capacity disks and subsystems, and affordable racing have led to larger storage configuration [3]. Microsoft has introduced the functionality of storage pools and storage spaces to Microsoft Windows server 2012 and Microsoft Windows 8. This shows that we still need to have an improvement in the storage media[10, 11]. Increasingly, it is not viable to talk about a storage media these days without talking about solid state or flash storage [12, 13]. The use of portable storage media has increased, thus, there is a need for improving the security mechanism for such devices. A secure solution that provides more security and user convenience is needed to avoid direct data loss and compromising the security of that data[14].

More or less, one and all who have suffered data loss know the importance of reliable data storage. The error on a storage device might be random or burst. There are some common causes of error in the storage devices. These error sources could be a scratch on the disc, the error from read/write failure, dust, or controller [7, 15]. It was observed that some of the challenges include managing storage growth, lack of skilled cloud technology professionals, lack of knowledge about forward error correction in storage media, designing and managing backup recovery and archive solution, and finally conversing the management to adopt the cloud. Figure 1 shows the distribution of hard disk problems at the Internet archive. In Figure 1, storage media problems such as disk failure, disk error and other disk subsystem problems consist of 60% of all the recorded hardware problems[16]. The common problem is the disk failure which covers 42% of all problems.

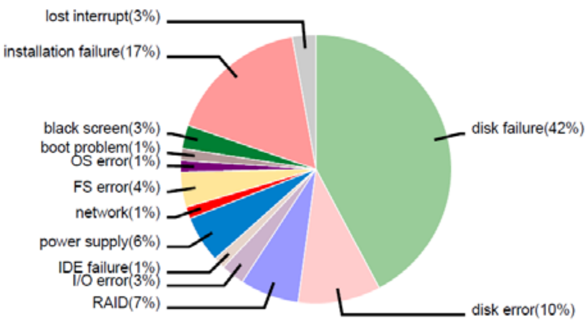

Figure 1: Hardware Failures Distribution at Internet Archive: Source : [16].

The recent research conducted in some companies in Dar es Salaam shows common error facing the storage media. In Figure 2, the recent study findings reveal that the most common error is the disk failure.

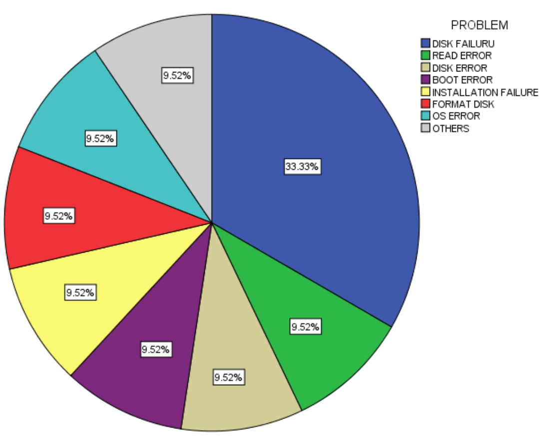

Figure 2: Common Storage Media Error for Different Companies

Reliability and performance of storage media are the biggest concerns. This implies that there is a need for the research to improve storage technologies in the existing systems.

II. STATE OF THE ARTS

In Spite of the problems discussed in the previous section, digital data storage industry applies defferent mechanisms such as backup and recovery systems to avoid data loss during failures. This section discusses efforts undertaken .

A: Backup System

Backup is the process of copying files or databases in order to preserve data in case of equipment failure or any other calamity. Backup or archiving data are later used to recover systems to their original state after failure. Backup has two main goals. The primary goal is to recover data after a loss by either deletion or corruption, whether intentionally or accidentally. The secondary goal is to restore the system to its earliest condition. There are three common types of backup systems which are full backup, incremental backup, and differential backup. In full backup, all the files are backed up every time you run a backup system. Incremental backup brings back the files that have changed since the last backup was done. Finally, differential backup provides backup of the files that have changed since a full backup was performed. However, different studies indicate that the pace of data growth is very high per year and even higher for the companies having large data and intensive application or distributed data centres. This means all data need to be backed up and duplicated. This fact makes the system more expensive and time consuming with this solution[17]. The crucial data should always be backed up to protect them from loss or corruption. Saving only one backup file sometimes might not be enough to preserve the information. In same vein, in order to safeguard crucial data, one needs to keep at least three copies of any such data, that is, one primary file and two backups. The two backup files are supposed to be kept in different media storage types to protect different types of hazards and making sure one file is kept offsite[18-20]. The importance of sensitive data to be backed up cannot be avoided because data storage media reliability can be low. With respect to this scenario, new techniques are introduced from time to time to secure vital data. For example, Oracle introduced sun ZFS appliance for Oracle Exadata Database which can restore up to 7TB/hour[21, 22]. Disaster planning would not be necessary if we were sure that nothing will ever happen in our environment such as hardware problem, viruses, and user errors. However, if a disaster is anticipated to occur, making backup and recovery systems is important to any vital data stored into the system. Essentially, the goal here is to create a way that crucial data will always be restored within acceptable time frame, within the budget, and without unnecessary shock on normal day to day activities[23]. However, one has to remember that it is very expensive because to back up data one needs extra devices and time to make backup, especially when there is large data.

B:Recovery System

Data recovery is a process of restoring or rescuing unintentionally deleted, corrupted, or inaccessible data to the storage media. Normally, this action is performed when we have physical damage or logical damage of the storage devices that prevent the file system of the devices to be mounted by the host system[24, 25]. Data recovery was introduced since 1974 when the magnetic data storage devices were commercially introduced into the market. And this was

https://sites.google.com/site/ijcsis/
ISSN 1947-5500

after the users discovered the vulnerability of their data stored on those devices and quickly to overcome this problem data recovery was introduced to protect the data stored on those devices[8]. Any event that causes an interruption or destruction in the daily operation or processing for a period of time and affects that operation is called a disaster. This disaster should be addressed to recover the operation intended and this can be done by the disaster recovery which must ensure the continuation of all processes in the event when a disaster occurs [26]. In the world where we live today, tera bytes or peta bytes of data are not enough for storing large chunks of databases and therefore data recovery systems become more challenging especially in a distributed system[27]. The importance of a recovery system cannot be ignored and that is why even in the application such as Oracle you will find that the application for recovery system is included; for example, Oracle Recovery Manager (RMAN) utility from Oracle that was designed for online backup and recovery of Oracle database files[28]. Due to huge increase in the electronic data, large volume of storage devices is required to store those data. Currently, consumers prefer to store their data in cloud computing. However, if the cloud gets corrupted or damaged, then the consumers lose their important data. This led to the introduction of some mechanisms to back up data so that they could be restored at any time when the cloud fails. There is a technique like the plain data backup though it has many security and reliability problems. To overcome plain data backup and recovery problems, we can use a system like Redundant Array Independent Disk (RAID) [29]. Again we can recover data from Window search in which we obtain the record from search database either via carving or via extensible storage Engine API which provides a potential source of evidence about the files that we cannot access and the reason why it is not accessible[30]. The cost for keeping data safe could be taken into consideration because it has some financial implication. In this case, a company or an organization must decide on the best practices that will support the solution in such a way that it does not compromise the financial aspect[8, 9]. Data recovery is one of the fields whereby once you make a mistake it might lead you to irrecoverable data. So it is very important that you perform operations which you are familiar with and do not do any action which you are not familiar with. The importance of data recovery cannot be ignored as it is very significant for our data and that is why even in the Window system, the introduction of data recovery is always there. Remarkably, it is very expensive to have these systems, even in the case of softwares, some are very expensive. As long as people still lose their important data, something has to be worked on regardless the presence of these other technologies to recover the data when something happens to our systems.

C: FEC Associated

Forward Error Correction (FEC) is a digital signal processing technique used to improve data reliability[31] . FEC makes the use of error correction codes to detect and correct errors automatically at the receiver. Error correction coding is the way whereby errors which may be introduced into digital data can be corrected based upon receiving data.. Error detection coding is the means whereby errors can be detected based upon receiving the data. Collectively, error correction and error detection coding are error control coding[32, 33]. Forward Error Correcting (FEC) codes grant algorithms for encoding and decoding data bits, and help to accomplish data rates nearer to Shannon's Limit[34]. Similarly, error coding is a method of providing reliable digital data transmission and storage signal to noise ratio[35]. Error coding is used for fault tolerant computing in computer memory, magnetic and optical data storage media, satellite and deep space communications, network communications, cellular telephone networks, and almost any other forms of digital data communication[36, 37]. The design of efficient error correcting codes needs a complete understanding of the error mechanisms and error distinctiveness [38]. There are several types of codes. The first major classifications are linear and non linear. A linear code is an error correction code for which any linear combination of codes is also a codeword. Linear codes are partitioned into block codes and convolutional codes. Linear codes allow for more efficient encoding and decoding algorithm than other codes [39].

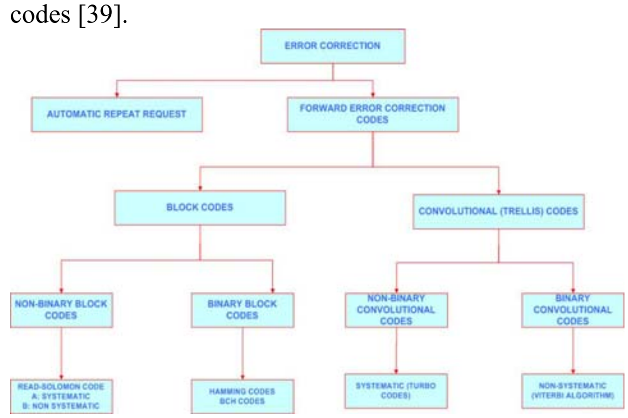

Figure 3: Error Correction Block

The linear codes are encoded using the method of linear algebra and polynomial arithmetics. If C is a linear code that a vector space over the field F has dimension k, then we say that C is an (n,k) linear code over F, or an (n,k) code, in short. Since linear codes allow for more and efficient encoding and decoding algorithm, then we will focus on linear codes and these are block codes and convolutional codes. Increasingly, a convolutional code operates on streams of data bits continuously, inserting redundant bits used to detect and correct errors. Convolutional codes are extensively used for real time error correction[40]. On the other hand, Block codes data are encoded in discrete blocks but not continuously. Convolutional codes are processed on a bit by bit basis. They are particularly suitable for the implementation in hardware, and the Viterbi decoder is one of the best algorithms in convolutional codes that allow optimal decoding. Additionally, Block codes are processed on a block by

https://sites.google.com/site/ijcsis/
ISSN 1947-5500

block basis and the best algorithm is the Reed Solomon codes [41]. There are different algorithms which are used for error correction in the storage media. Currently, most of the error correction in the storage media is done by block codes and this is because block codes have the power of correcting burst errors. The convolutional codes are more powerful than the block codes besides there more computational complexity than block codes. Mrutu and his colleagues showed that Viterbi algorithm decoder has less computational complexity than other convolutional decoders[42]. Furthermore, the introduction of non transmittable code words technique to assist Viterbi decoders has enabled it to overcome bust errors [42]. This fact motivated the researcher to find out the effectiveness of the technique in storage media.

Reed Solomon.

Reed Solomon (RS) is an error correcting code that addresses multiple error correction especially burst errors in the storage media (hard disk drives, CD, and DVD), wireless and mobile communication units, satellite links, and digital communication. Reed Solomon is among the best block coding techniques of which the data stream to be transmitted is broken down into blocks and redundant data is then added to each block. The size of the block and the amount of check data added to each block is either specified for a particular application or can be user defined for closed system[43, 44]. RS codes are non binary cyclic codes with symbols made up of m bit sequences, where m is any positive integer having values greater than 2. Reed Solomon are block codes which are represented as RS (n,k) where n is the size of the code word length and k is the number of data symbols. The size of the code word length n is given by 2t + k where 2t is the purity symbol.

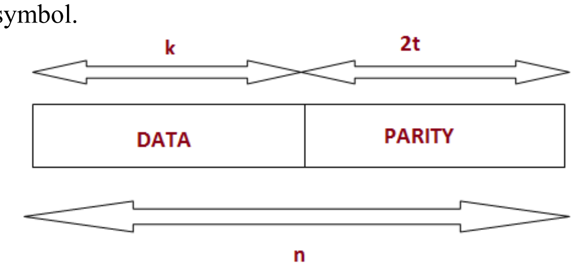

Figure 4: Structure of Reed Solomon
K=data symbol
2t=parity symbol
N=symbol for RS code

Reed Solomon (RS) code was founded by Irving S. Reed and Gustave Solomon who by then were staff members at MIT Lincoin Laboratory. RS t was subsequently presented to the world on paper in 1960[45]. The RS code construction is based on the Galois Field (GF (2^W)) operation for W is positive integer. To encode Reed Solomon, we split the polynomial that representing the message by an irreducible generator polynomial, then the reminder is the RS code which we attach to the original message. The first commercial application for RS appeared in 1982 in compact disk (CD). This was the first storage media device that used these error correction codes mechanisms [46]. RS illustrates a systematic way of building codes that may well detect and correct multiple random errors for different applications and they have strong burst and erasure error correction capabilities [47]. Also RS codes are able to correct multiple errors up to **t** errors, and they can be extended to correct errors up to **2t** errors, just making sure that up to **t** errors the positions are known [48]. RS codes make them especially suited to the applications where burst error occurs and when there is a single bit error. This is because it does not matter to the code how many bits in a symbol are incorrect, if multiple bits in a symbol are corrupted, it only counts as a single error. This means for single bit errors, RS code is usually a poor choice. Seemingly, RS operates in two sides, in the encoder and in the decoder. The code operates on 8 bits and the number of symbols in the encoder block is $n=2^m-1$. The generator polynomial generates redundant symbols and they are appended to the message symbol. The decoder side is the one which identifies the location and magnitude of the error. The same generator polynomial is used to do the identification and the correction is applied to the received code.

Fundamentally, RS is suitable for much application because it has very high coding rate and low complexity. Besides being suitable for storage media, we still face problems in our storage media, and this alerts the researchers that there is a need to work more in improving the error correction mechanisms in the storage media. The Reed Solomon coding currently remains the optimal code for the smallest storage systems[49]. Generally, Reed Solomon is considered not very scalable[7]. Altogether, Reed-Solomon coding employs the same methodology. There are **n** data words and m element column vector. The product is an n data words + m element column vector representing the coding words[45, 50].

Convolutional Codes

Convolutional code is one of the error correcting codes which generate parity symbols through the sliding application of Boolean polynomial function G(z) to the data stream. The sliding application is the representation of the convolution and its nature facilitates trellis decoding. More often Convolutional codes are termed as continuous. However, they have arbitrary block length, rather than being continuous, since most of the encoding is performed on blocks of data. Convolutional codes were introduced in 1955 by Peter Elias[51-53]. The main challenge in the convolutional codes is to find out the method for constructing codes of a given rate and minimizing their complexity[54]. Convolutional codes belong to FEC which makes use of the algorithm to automatically detect and correct error. They are one of the powerful error correcting codes that have a lower code rate (k/n). The common modulation scheme which is used is BPSK and QPSK, and the redundant bits are used to determine the error. The convolutional codes have two parts; one is the convolutional encoder and the second is the convolutional decoder. In the encoder, it uses encoder parameters which are n, k and m codes. Parameter ‘k’ is input, ‘n’ is output, and ‘m’ is memory. Usually k<n and ‘m’ must be

large to achieve low error probability. The value of these encoder parameters can range from 1 to 8 for k and n, and from 2 to 10 for m[53, 55]. The decoder part is very important because the performance of the convolutional code is determined by the decoding algorithm and distance properties of the code, and that is why it is very important to determine the best decoding algorithm. For example, for small value of k, the Viterbi algorithm is commonly used. Viterbi decoder is one the most widely used decoding algorithm[53]

III: Methodology

It is through a combination of literature review, technical reports, and data obtained from selected employees in data storage sections through interviews, in some companies operating in Tanzania that we could assess the awareness of data storage techniques and error correction mechanisms. The primary data for this study were collected from digital data storage companies in Dar es Salaam, Tanzania. Secondary data were also collected from published papers and industrial technical reports. In this study, seven companies were surveyed to identify the kind of storage media they use, the reliability and failure rate, the main course of error or failure of storage devices, the possible measures to be taken when an error or failure occurs, the possible solution being suggested to avoid data loss when an error or failure occurs, and the future plan to avoid error in the storage media. The collected data were analyzed using Statistical Package for the Social Sciences (SPSS).

IV. RESULT AND THE PROPOSED MODEL

Analysis of reliability and failure behavior of the storage media from different companies from Dar es Salaam was conducted. All possible measures taken when a disk with error fails and the suggestion for the improvement for solving the problem were collected and analysed. When conducting a survey, the discussion with IT professionals working in the data storage departments was conducted. It was observed that most of the existing data storage media in the existing systems use Read Solomon [43, 44, 56]. However, the survey shows that data loss is still a problem in the industry. Therefore, there is a need of improving the algorithm to minimize the residual error during data reading process from digital storage media. It was also perceived that most users of digital storage media are not aware of FEC technologies in the media. In brief, the following are the observations from the literature review and data collected from different companies in Dar es Salaam.

- There is little scholarly published work on the error and failure pattern of the storage media and the key factors that affect their lifetime. Most of the information comes from storage manufacturers
- In spite of having RS codes for Forward Error Correction in storage media, the industry is still facing problems of data and mostly disk error and disk failure.
- Most of the error correction in the storage media is done by using RS codes which work best on Optical disk drive.
- RS is relevant in correcting burst errors and not appropriate in correcting single error because the errors are corrected in blocks so it is the wastage of resources.
- Convolutional codes are more powerful than Block codes; however, they are not preferred as a solution due to their computational complexity which causes results of data processing to delay.
- The introduction of Locked convolutional codes with Non Transmittable Codewords (NTC) at the decoding side can solve the problem of computational complexity in Viterbi decoder.
- Majority of the people are not aware and are not interested in improving the Forward Error Correction codes for storage media; rather they are interested in improving the backup systems and data recovery software.
- The hard disk drive is one of the storage media, which is mostly used to store data.

The following results in figures 5, 6, 7 and 8 are obtained from the data collected in seven companies in Dar es Salaam.

Figure 5 displays the reliability of storage media in some companies in Tanzania. The reliability ranges from daily to weekly basis in most of them. This demonstrates that the reliability is very low and that is why the backup is done on daily and weekly basis and very few of them do it yearly.

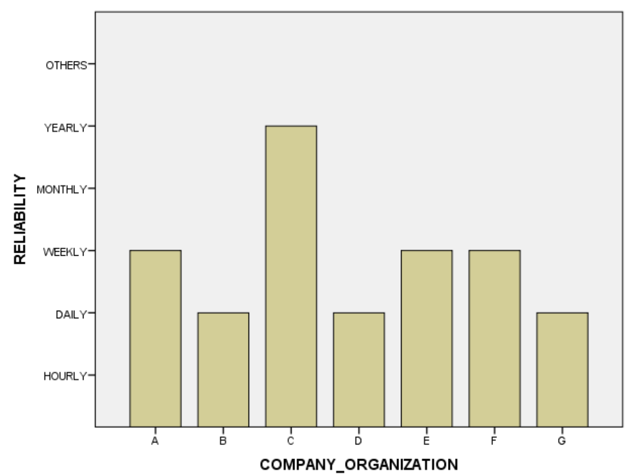

Figure 5: Reliability of storage media for some company in Tanzania

Figure 6 shows the failure rate of storage media in the surveyed companies. The failure rate is indicated for five years in the majority of the companies. This response to the findings indicated the reliability. If the reliability is low, then the failure rate is high. 85% revealed that the failure rate is not more than five years and only 14% is showed more than five years.

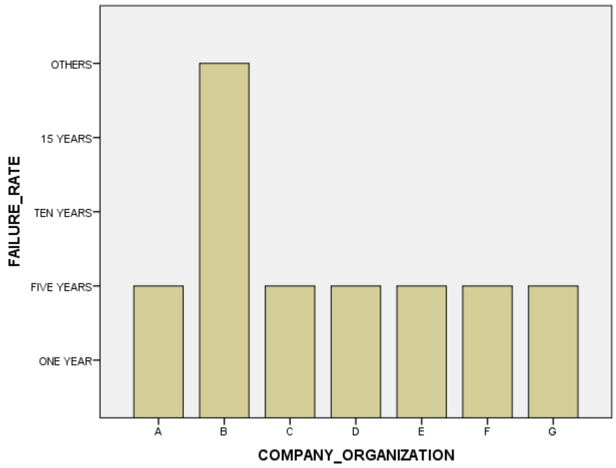

Figure 6: Failure Rate of Storage Media for Some Company in Tanzania

Figure 7 illustrates that most IT professionals in the industry think that the best way to improve data storage system is through backup system and few through data recovery and restore system. It was revealed that 71.43% prefer to improve backup system, 14.29% prefer to improve on data recovery software and improve other techniques respectively. But 0% prefers the improvement of Forward Error Correction. The findings imply that most people are not aware of the importance of improving these algorithms. However, improving these algorithms can improve the reliability of the storage media. In this case, if deploying a backup system is very expensive then having a reliable storage system can help a lot, especially for those who do not store large volumes of data. Similarly, the individual case will also be improved as we all know that this issue can go down to individual level.

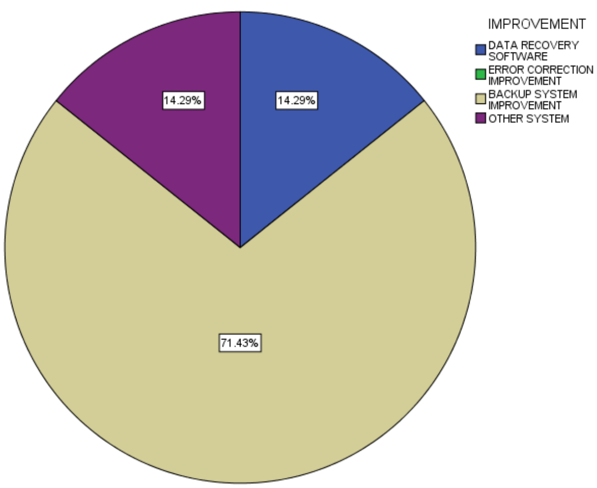

Figure 7: How to improve data storage loss

Figure 8 demonstrates that in most cases when there is data loss, very often the data are successfully recovered, and this is through backup system and data recovery software. But the analysis revealed that not always all the data are recovered. It is only 14.29% who are sure that the data will always be recovered and 85.71% remarked that most of the time they can recover their data but not always. 0% never lost everything when they tried to recover the lost data, but the question is for those who do not have a backup system; thus it goes to an individual level.

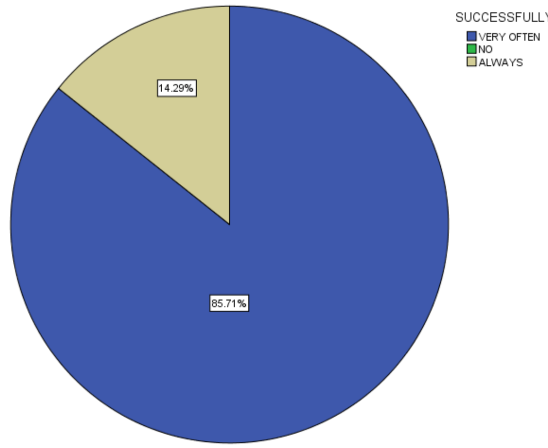

Figure 8: Data recovery after failure

Proposed Model

This study proposes improvement in the Forward Error Correction (FEC) in the storage media using Locked Convolutional Encoder and the enhanced Viterbi Algorithm with Non transmitted code words. In writing, the source data will be modulated, then encoded through locked convolutional encoder and at last written to the storage media. During reading the data will be decoded through Non transmitted code words (NTC)-Viterbi decoder then demodulated and finally the data can be retrived. It is over two decades since Iterative Detection Read Channel (IDRC) technology was adapted in hard drive design[57]. This approach is delivers enhancement to signal and noise ratio that are unique and improves the reliability, resiliency, and overall storage capacity of the storage drives. The proposed model adopts the mentioned IDRC architecture as shown in Figure 8. This architecture has three levels of coding where modulation and locked convolutional codes are concatenated in the media writing process. A reverse concatenation is done on the reading process. A study on the design and evaluation of the proposed architecture is our next task. Error correction in magnetic field started in 1960s, with the application of Fire codes. Later on Reed Solomon (RS) took the major role in this area. RS is currently the principal error correcting codes being used in optical disc memories. RS stretched out when Compact Disc (CD) was introduced and the CD ROM were introduced at the beginning of 1980s [58]. However, data loss problem due to storage media failure is still a challenge in the industry, leaving a door open for researchers to find a more relevant solution

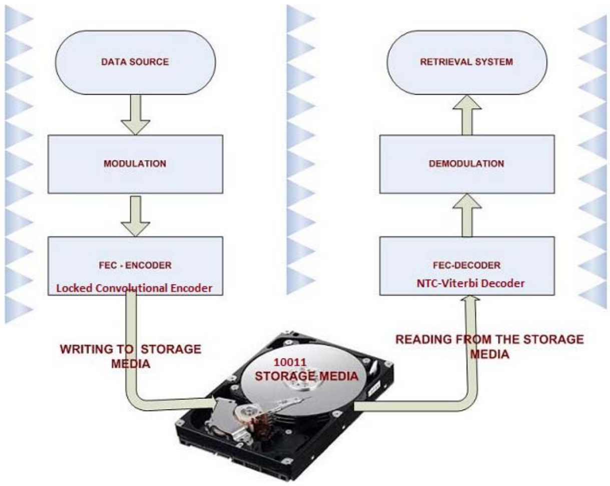

Figure 9: Proposed model

IV: CONLUSION

This paper reviewed the challenges facing the industry on storage media. It has been revealed that data reliability can be increased by adopting or introducing powerful Forward Error Correction Code (FEC) in digital storage media. It has been witnessed that storage industries are confronted with disasters of varying degrees; yet, people in the industries are not interested in improving the storage devices reliability and fault tolerance, but are interested in improving the backup and recovery systems which have cost implication. Observation on introducing powerful FEC suggested the use of new technique which uses locked convolutional encoder and NTC-Viterbi decoder. The technique shows that if improved and adapted, may give better results compared to the current techniques. Therefore we recommend a further study on improving the new technique so that we can improve the current situation facing the storage industries.

REFERENCES

1. Martin, M.P. and R. Pileggi, *A quantitative analysis of Propolis: a promising new storage media following avulsion.* Dental traumatology, 2004. **20**(2): p. 85-89.
2. Goradia, I., J. Doshi, and K. Deulkar, *A Review Paper on 3D Optical Data Storage.* 2014.
3. Report, E., *MANAGING STORAGE: TRENDS, CHALLENGES, AND OPTIONS (2013-2014).* 2014.
4. Van Bogart, J.W., *Magnetic Tape Storage and Handling: A Guide for Libraries and Archives*. 1995: ERIC.
5. Pinheiro, E., W.-D. Weber, and L.A. Barroso. *Failure Trends in a Large Disk Drive Population*. in *FAST*. 2007.
6. Nguyen, D.-C.L., Jaejin, *Performance Improvement Using Iterative Two-Dimensional Soft Output Viterbi Algorithm Associated with Noise Filter for Holographic Data Storage Systems.* 2014. **39**(3): p. 121-126.
7. Puttarak, N., *Coding for storage: disk arrays, flash memory, and distributed storage networks.* 2011.
8. Mayer, P., *Data Recovery: Choosing the Right Technologies.* Datalink White Paper, 2003.
9. Rao, Y.N., *DATA RECOVERY: CHOOSING THE RIGHT TECHNOLOGIES.* International Journal of Emerging Technology in Computer Science & Electronics (IJETCSE), 2014. **11**(4): p. 27-29.
10. Liu, J., et al. *PLC-cache: Endurable SSD cache for deduplication-based primary storage*. in *Mass Storage Systems and Technologies (MSST), 2014 30th Symposium on*. 2014. IEEE.
11. Yoshiyama, M. and K. Toriyama, *Fujitsu's Approach to Server System Development.* FUJITSU Sci. Tech. J, 2011. **47**(2): p. 123-129.
12. Peters, M., *EMC's Comprehensive Flash Strategy.* 2014.
13. Yaakobi, E., et al., *Error correction coding for flash memories.* Flash Memory Summit, 2009. **160**.
14. Agwankar, N., et al., *Security For Portable Data Storage media*. 2014.
15. Tsai, T., N. Theera-Ampornpunt, and S. Bagchi. *A study of soft error consequences in hard disk drives*. in *Dependable Systems and Networks (DSN), 2012 42nd Annual IEEE/IFIP International Conference on*. 2012. IEEE.
16. Xin, Q., *Understanding and coping with failures in large-scale storage systems*. 2007, Technical Report UCSC-SSRC-07-06.
17. Xu, Q., et al., *YuruBackup: a space-efficient and highly scalable incremental backup system in the cloud.* International Journal of Parallel Programming, 2013. **43**(3): p. 316-338.
18. Crighton, I.P., *Data backup system*. 2001, Google Patents.
19. Gold, S., et al., *Data backup and recovery systems.* 2004, Google Patents.
20. Jewell, T.R., *Data backup system and method*. 2008, Google Patents.
21. Clarke, J., *Backup and Recovery*, in *Oracle Exadata Recipes*. 2013, Springer. p. 189-238.
22. Bach, M., et al., *Recovering Exadata*, in *Expert Oracle Exadata*. 2015, Springer. p. 303-339.
23. Chow, W.S. and W. On Ha, *Determinants of the critical success factor of disaster recovery planning for information systems.* Information Management & Computer Security, 2009. **17**(3): p. 248-275.
24. McCormack, G.D. and R.F. Talaga Jr, *Multiple channel adaptive data recovery system*. 2002, Google Patents.
25. Gabber, E., et al., *Redundant data storage and data recovery system*. 2004, Google Patents.
26. Martin, B.C., *Disaster Recovery Plan Strategies and Processes.* SANS Reading, 2002.
27. V. S. Karwande, D.S.S.L., Prof. R. A. Auti, *The Data Recovery File System for Hadoop Cluster -Review Paper*. (IJCSIT) International Journal of Computer

Science and Information Technologies, 2015. **6**(1): p. 365-367.

28. corporation, E., *MCD Data Domain Boost for Oracle Recovery Manager (RMAN)*. 2015: EMC Corporation.
29. Sharma, K. and K.R. Singh, *Online data back-up and disaster recovery techniques in cloud computing: A review.* International Journal of Engineering and Innovative Technology (IJEIT), 2012. **2**(5): p. 249-254.
30. Chivers, H. and C. Hargreaves, *Forensic data recovery from the Windows Search Database.* digital investigation, 2011. **7**(3): p. 114-126.
31. Puri, R., et al., *Forward error correction (FEC) codes based multiple description coding for Internet video streaming and multicast.* Signal Processing: Image Communication, 2001. **16**(8): p. 745-762.
32. Jadhao, M.V.G. and P.D. Gawande, *Performance Analysis of Linear Block Code, Convolution code and Concatenated code to Study Their Comparative Effectiveness.* 2012.
33. Toghuj, W. and G.I. Alkhatib, *Improved Algorithm for Error Correction.* International Journal of Information Technology and Web Engineering (IJITWE), 2011. **6**(1): p. 1-12.
34. Mitchell, G., *Investigation of Hamming, Reed-Solomon, and Turbo Forward Error Correcting Codes*. 2009, DTIC Document.
35. Jadhao, V.G. and P.D. Gawande, *Development Of Error Control Coding Technique In Digital Communication System.* International Journal of Managment, IT and Engineering, 2012. **2**(4): p. 304-316.
36. Jadhao, M.V.G.G., Prafulla D, *Performance Analysis of Linear Block Code, Convolution code and Concatenated code to Study Their Comparative Effectiveness.* 2012.
37. Moon Todd, K., *Error correction coding: mathematical methods and algorithms. 2005 by John Wiley & Sons*. 2005, ISBN 0-471-64800-0.
38. Yaakobi, E., et al. *Characterization and error-correcting codes for TLC flash memories*. in *Computing, Networking and Communications (ICNC), 2012 International Conference on*. 2012. IEEE.
39. Ryan, W. and S. Lin, *Channel codes: classical and modern*. 2009: Cambridge University Press.
40. Kumar, S. and R. Gupta, *Performance Comparison of Different Forward Error Correction Coding Techniques for Wireless Communication Systems.* 2011.
41. Morelos-Zaragoza, R.H., *The art of error correcting coding*. 2006: John Wiley & Sons.
42. Mrutu, S.I., A. Sam, and N.H. Mvungi, *Assessment of Non Transmittable Codewords Enhancement to Viterbi Algorithm Decoding.* International Journal of Communication Science and Information Security (IJCSIS), 2014. **12**: p. 9.
43. Singh, S. and S. Sujana. *ASIC Implementation Of Reed Solomon Codec For Burst Error Detection And Correction*. in *International Journal of Engineering Research and Technology*. 2013. ESRSA Publications.
44. Shrivastava, P. and U.P. Singh, *Error Detection and Correction Using Reed Solomon Codes.* International Journal of Advanced Research in Computer Science and Software Engineering, 2013. **3**.
45. Moon, Y.S.J. and N. Kaplan, *Introduction to Reed-Solomon Codes*. 2011, August.
46. van de Pavert, L., *REED-SOLOMON ENCODING AND DECODING.* 2011.
47. Revathy, M. and R. Saravanan, *HDL IMPLEMENTATION OF ALGEBRAIC SOFT DECISION ALGORITHM FOR RS CODES.* 2013.
48. Varsamou, M., I. Zacharias, and T. Antonakopoulos. *Correlated noise estimation and error correction in parallel storage channels*. in *Signal Processing and Information Technology (ISSPIT), 2013 IEEE International Symposium on*. 2013. IEEE.
49. Plank, J.S. and L. Xu. *Optimizing Cauchy Reed-Solomon codes for fault-tolerant network storage applications*. in *Network Computing and Applications, 2006. NCA 2006. Fifth IEEE International Symposium on*. 2006. IEEE.
50. Plank, J.S., *Optimizing Cauchy Reed-Solomon codes for fault-tolerant storage applications.* University of Tennessee, Tech. Rep. CS-05-569, 2005.
51. WIKIPEDIA. *Convolutional code*. 2015; Available from: https://en.wikipedia.org/wiki/Convolutional_code.
52. Mrutu, S.I., A. Sam, and N.H. Mvungi, *Forward Error Correction Convolutional Codes for RTAs' Networks: An Overview*. International Journal of Computer Network and Information Security (IJCNIS), 2014. **6**(7): p. 19.
53. NEHA, G.M.R., *CONVOLUTION ERROR CONTROL CODING -A REVIEW*, in *IRF International Conference*. 2015: Pune, India.
54. García Planas, M.I., E.M. Souidi, and L.E. Um, *Decoding algorithm for convolutional codes under linear systems point of view*. 2014.
55. Grassl, M. and M. Rötteler. *Quantum block and convolutional codes from self-orthogonal product codes*. in *Information Theory, 2005. ISIT 2005. Proceedings. International Symposium on*. 2005. IEEE.
56. Raj, V., *Performance study on the suitability of Reed Solomon codes in communication systems.* CT International Journal of Information & Communication Technology, 2013. **1**(1).
57. Oenning, R.G.a.T., *Iterative Detection Read Channel Technology in Hard Disk Drives*. 2008: 2012 HGST, a Western Digital company.

58. Costello Jr, D.J., et al., *Applications of error-control coding.* Information Theory, IEEE Transactions on, 1998. **44**(6): p. 2531-2560.

Author Profile

Kilavo Hassan: Currently he is the Ph.D. student at Nelson Mandela African Institution of Science and Technology in the School of Computational and Communication Science and Engineering. He received his B.Sc. In Computer Science at the University of Dar Es Salaam, Tanzania and His M.Sc. in Information and Communication Science Engineering majoring Telecommunication Engineering at the Nelson Mandela African Institution of Science and Technology Arusha Tanzania. His research interest includes Forward Error Correction, Embedded System and Automation System.

Kisangiri Michael, received his Ph.D. degree in Telecommunication Engineering (2008) and Master of Science in Telecommunication Engineering (2002), both from Wrocław University of Technology - Poland. Since then he has been working as academician with several universities. Currently he works with Nelson Mandela African Institution of Science and Technology located in Arusha, Tanzania. His research interests include evolutionary computation in Telecommunication Networks, Antenna design and Triangular mesh modeling.

Salehe I. Mrutu received his B.Sc. and M.Sc. degrees in Computer Science from The International University of Africa in 2003 and the University of Gezira in 2006 respectively. He received his Ph.D.in Information and Communication Science and Engineering in 2014 at Nelson Mandela African Institution of Science and Technology (NM-AIST) in Arusha, at the school of Computational and Communication Sciences and Engineering. He is currently serving as lecturer at the University of Dodoma (UDOM) under the school of Informatics since 2007. His research interests include forward error correction codes, algorithms analysis and designing, quality-of-service provisioning and resource management for multimedia communications networks.

https://sites.google.com/site/ijcsis/
ISSN 1947-5500