Performance Analysis of Row-Diagonal Parity (RDP) Algorithm in RAID 6 and Its Application in Cloud Computing

Zhaohuan Zhang *

College of Information Science and Technology/College of Cyber Security, Jinan University, Guangzhou, Guangdong, 510632, China

* Corresponding Author Email: Jaszhang@stu2020.jnu.edu.cn

Abstract. With the explosive growth of information, there is an increasing demand for storage technology, and cloud computing has become a major trend with the continuous development of new technologies. Therefore, this paper aims to find superior fault-tolerant coding schemes and explore their potential applications in the field of cloud computing. This paper provides a brief introduction to the background of RAID systems and fault-tolerant coding; elaborates on the principles and encoding process of the RDP algorithm in detail; deeply analyzes the application and performance of the RDP algorithm in RAID-6, and a comparison was made with RDP coding in cloud storage; and analyzes the development potential and specific application scenarios of the RDP algorithm in the field of cloud computing. Based on the above analysis and research, this paper affirms the development potential of the RDP algorithm in cloud storage based on cloud computing. It also proposes the research direction of the RDP algorithm in the field of cloud computing.

Keywords: RDP coding, RAID-6, Cloud storage.

1. Introduction

1.1. Background Information

With the development of computers, research on storage technology is crucial. The introduction of RAID (Redundant Array of Independent Disks) technology [1] can be considered a significant revolution in data storage and has become the cornerstone of virtualized storage technology. RAID was developed to combine multiple small-capacity inexpensive disks as a cost-effective alternative to larger, expensive disks, to minimize data loss in the event of disk failures. It organizes multiple disks in a specific configuration and scheme to achieve higher speed, better stability, and larger storage capacity than a single disk. With the rapid advancement of computer technology, RAID technology has permeated various fields. However, it has also brought about increasing concerns regarding data security. Although RAID technology provides a certain level of data security, catastrophic events resulting from hardware damage and human errors still lead to RAID data loss incidents. In the face of damaged RAID storage devices, RAID data recovery technology becomes the only hope for data restoration.

RAID technology utilizes different categories of techniques based on varying application requirements, known as RAID levels [2]. Each level represents a specific set of technologies. Currently, the widely recognized RAID levels in the industry are RAID-0, RAID-1, RAID-2, RAID-3, RAID-4, and RAID-5. These different levels do not indicate the superiority or inferiority of the technology. The choice of which RAID level to use depends on the user's operating environment and application needs. In this article, RAID-6 is mentioned, which is not a standard RAID level. Therefore, different vendors may have their standards, with the most famous being Intel's P+Q dual-parity RAID-6. However, the different standards do not affect the analysis in this article.

1.2. Purpose and Motivation

RDP (Row-Diagonal Parity) was first introduced to reconstruct up to two disk failures [3]. In RAID-6, the RDP algorithm is used for generating and checking parity data. It is responsible for calculating parity across data blocks in a row diagonal manner. It utilizes a combination of XOR operations and Reed-Solomon codes to create parity information. The parity data is then distributed across multiple drives in the RAID array, providing redundancy and fault tolerance. The RDP algorithm plays a crucial role in RAID-6 by allowing for the reconstruction of data in the event of disk failures. Thus, this paper researches the RDP algorithm to find a superior fault-tolerant coding scheme and explore its application in cloud computing.

1.3. Generalization

In the following second chapter, this paper will provide a brief introduction to the previous research achievements of RAID systems. In Chapter 3, the principles and encoding process of the RDP algorithm will be analyzed in detail. Chapter 4 will delve into the application and performance of the RDP algorithm in RAID-6. Chapter 5 will analyze the potential of the RDP algorithm in the field of cloud computing. In the final chapter, a summary of this paper's research will be provided along with prospects.

2. Related work

2.1. Error-Resilient Coding in RAID-6

Currently, in the application of RAID systems, a large number of coding methods have been accumulated. The earliest erasure coding techniques only allowed for the failure of a single disk. As the size of disk arrays increased, the possibility of multiple disk failures greatly increased, making the need for fault-tolerant codes capable of recovering multiple failed disks essential. Gibson et al. first proposed a one-dimensional parity grouping method that could correct any two disk failures [4]. Later, Gibson and Patterson proposed a method that used Hamming codes to correct any two disk failures [5]. However, both coding schemes required several redundant disks logarithmic to the total number of disks in the array, resulting in a large amount of redundant information and expensive costs.

During the same period, research on applying Reed-Solomon (RS) codes in storage systems had also begun, but it was not until 1997 that breakthrough progress was made. Plank [6] and Nonnenmacher [7], among others, proposed two simple software implementations of RS codes, eliminating the need for hardware implementation and directly promoting the application of RS codes in distributed storage systems. However, in terms of computational complexity, array codes are more efficient than RS codes at the same coding efficiency. Research on array codes has mostly focused on double erasure MDS array codes, such as EVENODD code [8], BCP code, RDP code [3], etc.

With the further development of storage systems and erasure coding technology, low-density parity-check (LDPC) codes have begun to be applied in disk array storage systems. For example, Liberation code [9] contains the "1"s in its generator matrix, achieving the optimal update complexity theoretically achievable for RAID-6 codes.

2.2. Comparison of RDP Algorithm and Other Coding

In this section, this paper will compare the fault-tolerant coding mentioned in the previous section with the RDP codes. First, let's consider EVENODD codes, which have the most similarity in structure with RDP codes. Both codes have the same error-correcting capability and update complexity, which means that each change in a data bit requires an average of about three updates to the parity bits. The advantage of RDP codes lies in their lower encoding and decoding complexity. In the encoding process, RDP codes first calculate the value of the first parity column and then use it as a data column to participate in the calculation of the diagonal parity columns. When the code's dimension, i.e., the number of data columns $k \in \{p-1, p-2\}$, the encoding and decoding complexity of RDP codes reaches the lower bound of $k-1$. For other values of k, the encoding and decoding complexity of RDP codes is still significantly lower than that of EVENODD codes and very close to the theoretical lower bound [10].

Furthermore, in comparing with Liberation codes, because the codewords of Liberation codes are a $p \times (p+2)$ two-dimensional array, where p is a prime number, so its update complexity is very close to 2, which is lower than the 3 of RDP codes. In terms of encoding complexity, if the dimension of the code is k, the encoding complexity of Liberation codes is$k-1+\frac{k-1}{2p}$. It can be seen that for a given k, the larger the prime number p, the closer the encoding complexity is to the theoretical lower bound of k − 1. This is exactly the opposite of RDP codes, which require choosing a prime number closest to k to minimize the encoding and decoding complexity. This characteristic of Liberation codes makes them suitable for constructing dynamically expandable RAID-6 disk arrays. Furthermore, in terms of decoding, its complexity is about 15% higher than the theoretical lower bound, which is much higher than the decoding complexity of RDP codes [11].

2.3. Implementation of RDP in Cloud Computing

In this section, I will briefly describe the application research of the RDP algorithm in the cloud computing environment. The RDP algorithm can be mainly used in cloud storage, cloud data analysis, and so on. Cloud storage usually utilizes scalable distributed storage systems deployed on a large number of inexpensive nodes to store data. As the scale of cloud storage continues to increase, it often faces the risk of node loss. To improve reliability against node loss, storage systems usually redundantly store data on multiple nodes. Erasure code is an efficient data redundancy technique that can achieve higher data reliability with the same storage overhead. Therefore, it is widely used in cloud storage systems. In the academic community, many research studies on cloud storage systems are based on erasure codes, such as private cloud storage (i.e., enterprise data centers) like Facebook's f4, IBM's Cleversafe, Hitchhiker, Giza, etc., as well as public cloud storage like Azure, Hybris, etc.

Data from a Google data center shows that disk failure has the highest probability among all causes of operational abnormalities and data loss [12]. The RDP algorithm has significant advantages in addressing disk failure issues. It has extremely low encoding and decoding complexity and low update complexity, meeting the low-cost and high-reliability requirements of cloud storage. However, RDP encoding also faces certain challenges, such as the increase in capacity without significant improvement in read performance. To solve this problem, this paper suggests to consider minimizing random data read as much as possible.

3. Basic Principles of RDP

3.1. Introduction

In this section, I will briefly introduce the basic principles and design ideas of the RDP algorithm. RDP coding is based on a simple parity coding method that uses only the XOR operation to calculate the checksum, which is concise and efficient. The array of the RDP algorithm is controlled by a prime number p, which is the number of disks-1. For easier understanding, let's take an example of an RDP array with p=5, as shown in Table 1. In a regular RDP array, there are p+1 disks. All the data blocks of the disks in a row form a stripe. In each stripe, one block stores the diagonal parity, one block stores the row parity, and the remaining p-1 blocks store the data. Each data block of the RDP algorithm belongs to a row parity set and also belongs to a diagonal parity set. In general, each stripe has a row parity set and a diagonal parity set. This algorithm can also scatter the checksum data by rotating it between disks to fully utilize the transmission bandwidth [3].

Table 1. RAID-6 applying RDP coding, where p=5 (Picture credit: Original).

	Data disk 0	Data disk 1	Data disk 2	Data disk 3	Row parity	Diagonal parity
Stripe 0	x	**x**	**x**	**x**	**x**	**x**
Stripe 1	x	**x**	**x**	**x**	**x**	**x**
Stripe 2	x	**x**	**x**	**x**	**x**	**x**
Stripe 3	x	**x**	**x**	**x**	**x**	**x**

There are two advantages to doing this. First, by calculating the checksum units in sequence during encoding, each unit only needs to perform two XOR operations, so the average encoding complexity of the array can be optimized. Secondly, when the encoding cycle is also p-1, the RPD code has one less data column compared to the EVENODD code, because the horizontal parity needs to be included in the diagonal parity calculation. Therefore, the number of units in each column of the diagonal parity is uniform, which effectively eliminates system bottlenecks by evenly distributing the update calculations to each unit.

3.2. Coding process

The detailed encoding process of the RDP algorithm is shown in Figure 1, which depicts the four stripes of the RDP array when the controlling parameter p=5. All the computational operations are XOR. The calculation formulas are as follows [13]:

$$a_{i,p-1} = \bigoplus_{j=0}^{p-2} a_{i,j} \quad \text{For } 0 \le i \le p-2 \tag{1}$$

$$a_{i,p} = \bigoplus_{j=0}^{p-1} a_{\langle i-j \rangle p,j} \quad \text{for } 0 \le i \le p-2 \tag{2}$$

The roe parity can be calculated using equation (1) and the diagonal parity using equation (2), where $\left(a_{i,j}\right)_{\substack{0 \le i \le p-2 \\ 0 \le j \le p}}$. And there is always one diagonal that only takes part in one row parity check but does not participate in any diagonal parity check. This diagonal is called the “missing diagonal”.

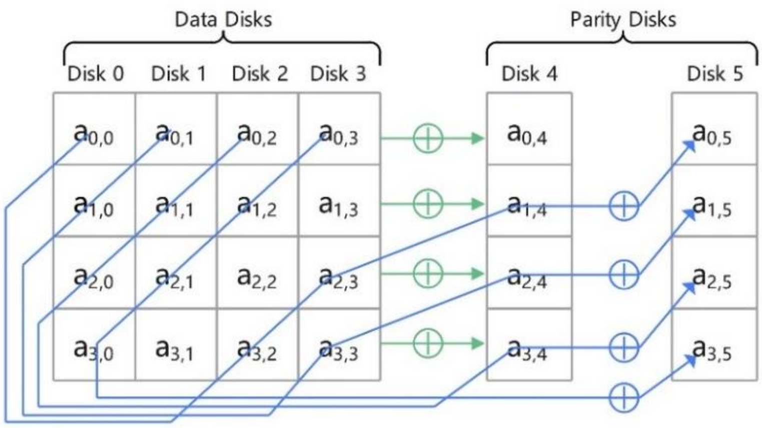

Figure 1. RAID-6 using RDP algorithm (p=5) (Picture credit: Original).

As can be seen, by using equations (1) and (2), the data of any two failed disks can be effectively recovered. Dual disk failure can be divided into two cases: ① one data disk failure + diagonal parity disk failure; ② two data disk failures or one data disk failure + row parity disk failure. The first case is relatively simple, as we only need to first calculate the missing data block using equation (1) and then calculate the diagonal parity using equation (2). In the second case, we need to alternately use the two equations based on the missing information.

4. Implementation and Performance Analysis of RDP in RADI 6

4.1. Integration of RDP coding in RAID 6

In this section, this paper discusses the concrete implementation of the RDP algorithm in RAID-6. Figure 2 shows the visualization of the disk drive to be seen on the server. In fact, it is the map of a continuous stripe set on the disk array. This set is called a slice, the slice then is distributed in several stripe groups, and each stripe group contains p-1 stripes. If the RDP algorithm is applied to a RAID-6 system, it means that each stripe group is constructed individually based on the RDP algorithm. Therefore, I want to note that all the examples discussed in Section 3 are focused on a single stripe group. In other words, implementing the RDP algorithm in RAID-6 is a process of constructing and reconstructing stripe groups using the RDP encoding concept mentioned in Section 3.

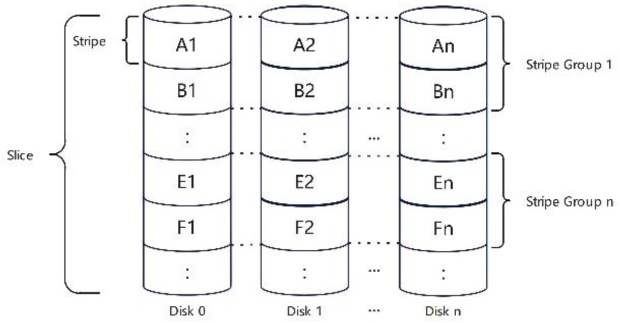

Figure 2. Structure of a disk (Picture credit: Original).

4.2. Performance analysis

In this section, I will primarily analyze the performance of the RDP algorithm in RAID-6 from two aspects: encoding complexity and update complexity. The theoretical minimum (optimal) encoding complexity for any RAID-6 code with a code length of n and MDS properties is to perform an average of $2-\frac{2}{n-2}$ XOR operations per data block during the encoding process. The minimum decoding complexity is to perform an average of (n-3) XOR operations per data block during the decoding process [14].

The biggest advantage of the RDP algorithm is that it can achieve the minimum encoding complexity of RAID-6. This can be proven by the data shown in Figure 3. Figure 3 demonstrates the encoding complexity of RDP codes for code lengths ranging from 4 to 30. All values are normalized relative to the minimum encoding complexity of RAID-6 codes with the same code length. One curve represents the case where the parameter p is always chosen as the prime number closest to n when constructing an RDP code of code length n, while the other curve represents the case where the parameter p is fixed at the prime number 31 [15].

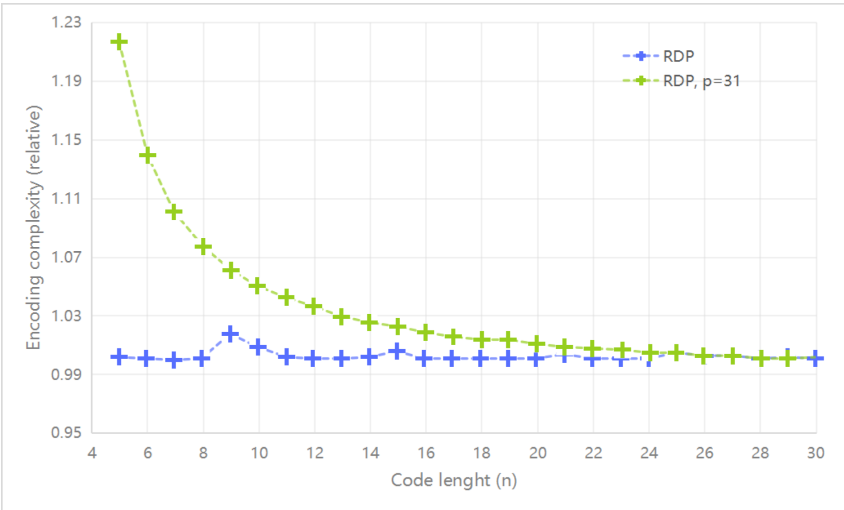

Figure 3. Encoding complexity of RDP code in RAID-6 (Picture credit: Original).

By observing the first curve, it is proven that RDP encoding achieves the minimum encoding complexity of RAID-6 at standard code lengths and is very close to the minimum value at other code lengths. Additionally, by analyzing the second curve, it can be found that the encoding complexity of the RDP algorithm increases as the code length decreases when the prime number p is fixed at 31. Therefore, to minimize the encoding complexity of the generated RAID-6 code, the parameter p needs to be the smallest possible value.

Furthermore, the minimum (optimal) update complexity of RAID-6 codes should be 2 because each data block in the RAID-6 code structure must participate in the calculation of at least two parity blocks to tolerate double disk failures [14].

The update complexity of the RDP algorithm is not optimal compared to the Liberation code, but we can minimize it by controlling the parameter p. Figure 4 shows the average update complexity of RDP codes for code lengths ranging from 4 to 30.

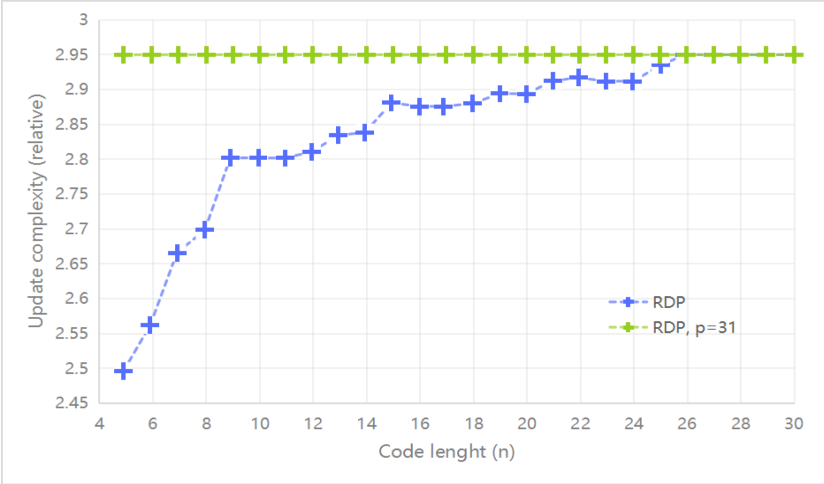

Figure 4. Update complexity of RDP code in RAID-6 (Picture credit: Original).

From the figure, it is clear that the average update complexity of RDP encoding increases with the code length and eventually approaches the value of 3. Therefore, we can reduce the update complexity by choosing the smallest possible value for the parameter p.

Through the analysis in this section, we can see that RDP codes perform excellently in RAID-6 systems.

5. Potential Applications of RDP in Cloud Computing

5.1. Background Information

Cloud computing is a new service operation model that combines traditional computer technologies such as distributed computing, network storage, virtualization, and load balancing with network technologies. In this model, various computing and storage services that traditionally reside locally are replaced by cloud computing services. Users can access various computing and storage resources provided by cloud computing through the network, eliminating the cost of purchasing, managing, and maintaining for enterprises or individual users. Figure 5 shows the fundamental model of cloud computing, where the cloud computing platform consists of centralized large-scale data centers and computing centers, forming the "cloud" side that provides various services to users. Users access the services provided by the "cloud" side through the network.

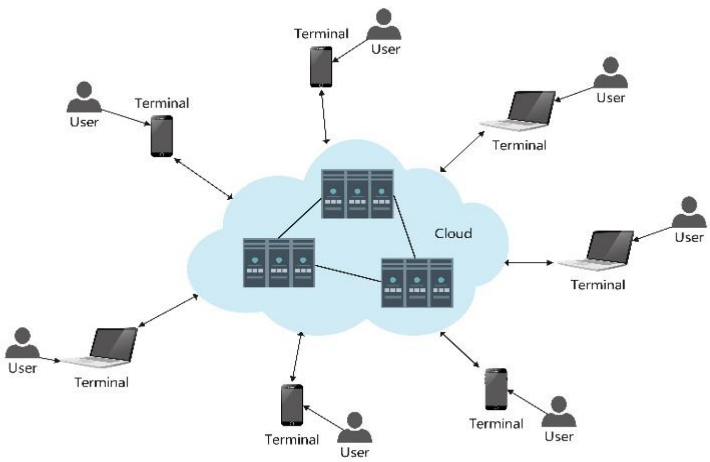

Figure 5. Cloud computing model (Picture credit: Original)

Cloud computing is a combination of multiple services, as shown in Figure 6, which illustrates its layered structure. SaaS (Software as a Service) hosts all applications. Users no longer need to worry about software installation and upgrades. For software developers, as all resources related to software are placed in the cloud, they can easily deploy and upgrade software. PaaS (Platform as a Service) involves deploying operating systems, development environments, such as programming languages and tools (e.g., Java, Python, .Net, etc.), or acquired applications on the vendor's cloud computing infrastructure. Users can develop new applications or extend existing ones on this platform. IaaS (Infrastructure as a Service) refers to providing IT infrastructure capabilities (servers, storage, computing, networks, etc.) to users over the Internet as a service.

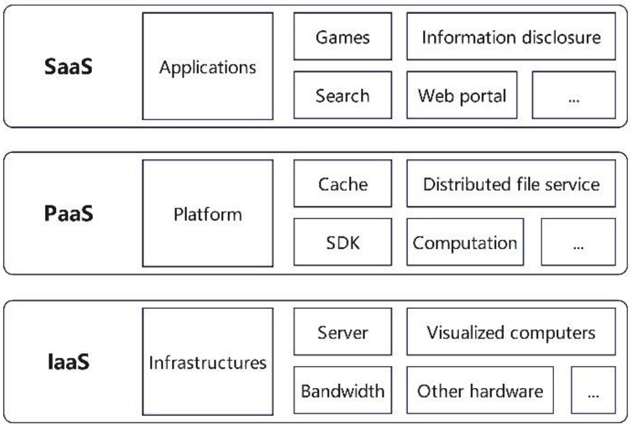

Figure 6. Layers of cloud computing (Picture credit: Original).

5.2. Adaptability Analysis of RDP

In section 2.3, this paper briefly mentioned that the main application of the RDP algorithm in cloud computing is in the field of cloud storage. As an erasure code, its main role is to ensure data integrity and reliability of the storage system. Cloud storage is a new concept generated by combining cloud computing with distributed storage technology. Compared with traditional storage technology cloud storage requires fault-tolerant techniques to be low-cost, highly reliable, and so on.

Cloud storage can be divided into storage layer, infrastructure management layer, application interface layer, and access layer. One important technology in the infrastructure management layer is fault tolerance, where the RDP algorithm can be applied. Through the analysis in section 4.2, we can understand that the RDP algorithm not only has good encoding and decoding performance but also can tolerate the failure of multiple nodes in the system.

6. Conclusion

6.1. Summary and Comparison

Over the past half century, computer technology has become fully integrated into social life, resulting in an enormous amount of information worldwide and its rapid growth. This poses demanding requirements for storage technology. At the same time, with the rapid growth of unstructured data and the booming development of new technologies such as 5G, the Internet of Things (IoT), and artificial intelligence, massive data is pouring into cloud storage, leading to an explosive development of cloud storage. Therefore, this paper provides an in-depth discussion of the application of the RDP algorithm in RAID-6 and cloud computing.

Firstly, in Chapter 3, this paper provides a detailed introduction to the basic principles of the RDP algorithm and the complete encoding process. Subsequently, in Chapter 4, the specific application methods of the RDP algorithm in RAID-6 and its performance are elaborated. Finally, in Chapter 5, this paper briefly introduces cloud computing technology, followed by the specific application scenarios and advantages of the RDP algorithm in the field of cloud computing.

6.2. Prospect and Future Research Orientation

Since its inception, the RDP algorithm has been widely used in RAID systems, and its advantages are self-evident. However, in practical applications in the field of cloud computing, most of them are still based on replica technology, and the use of erasure codes is mostly based on RS codes. Research on array codes is just beginning. Therefore, this is a new direction that has the potential to overcome existing challenges.

Compared to the multi-replica fault tolerance mechanism, the erasure code fault tolerance mechanism reduces data storage costs but causes larger network traffic and increased system data repair latency when performing data repair. Therefore, in erasure code schemes, reducing network traffic during data repair and minimizing system repair latency become challenges in data storage.

References

[1] Katz, Randy H., Garth A. Gibson, and David A. Patterson. “Disk system architectures for high-performance computing.” Proc. IEEE 77 (1989): 1842-1858.

[2] Zhongxia Wang, et al. “Demystification of RAID Data Recovery technology”, Tsinghua university press, Beijing, p.1-45, 2010.

[3] Corbett, Peter F. et al. “Row-Diagonal Parity for Double Disk Failure Correction (Awarded Best Paper!).” (2004).

[4] Gibson, Garth A., Lisa Hellerstein, Richard M. Karp, Randy H. Katz, and David A. Patterson. "Failure correction techniques for large disk arrays." ASPLOS III (1989).

[5] Gibson, Garth A. and David A. Patterson. “Designing Disk Arrays for High Data Reliability.” J. Parallel Distributed Comput. 17 (1993): 4-27.

[6] Plank, James S... “A tutorial on Reed–Solomon coding for fault-tolerance in RAID-like systems.” Software: Practice and Experience 27 (1997): n. pag.

[7] Nonnenmacher, Jörg, Ernst W. Biersack and Donald F. Towsley. “Parity-based loss recovery for reliable multicast transmission.” TNET (1997).

[8] Blaum, Mario, James Thomas Brady, Jehoshua Bruck, and Jai Menon. "EVENODD: an optimal scheme for tolerating double disk failures in RAID architectures." Proceedings of 21 International Symposium on Computer Architecture (1994): 245-254.

[9] Plank, James S... “The RAID-6 Liberation Codes.” USENIX Conference on File and Storage Technologies (2008).

[10] Chen, Yiran. "Efficiency Comparison of Row-Diagonal Parity and EVENODD Encoded Check Disk Repair Algorithms." 2022 International Symposium on Advances in Informatics, Electronics, and Education (ISAIEE) (2022): 55-58.

[11] Zhijie Huang. “Research on MDS Array Codes in Fault-Tolerant Storage Systems” [D]. Huazhong University of Science & Technology,2017.

[12] Yan Zhang. “On the Disk Failure Recovery of RDP-based Distributed Storage Systems” [D]. University of Science and Technology of China, 2015

[13] Blaum, Mario. “A Family of MDS Array Codes with Minimal Number of Encoding Operations.” 2006 IEEE International Symposium on Information Theory (2006): 2784-2788.

[14] Jin, Chao, Hong Jiang, Dan Feng and Lei Tian. "P-Code: a new RAID-6 code with optimal properties." Proceedings of the 23rd International Conference on Supercomputing (2009): n. pag.

[15] Chao Jin, et al." Code Shortening and Performance Analysis for RAID6Codes." ACTA ELECTRONICA SINICA 40.01(2012):173-178.