Proceedings of the 2023 International Conference on Machine Learning and Automation
DOI: 10.54254/2755-2721/32/20230187

Examination of essential technologies and representative applications in RAID 6

Yuxuan Du

Sydney Smart Technology College, Northeastern University, Qinhuangdao, 066004, China

202019152@stu.neuq.edu.cn

Abstract. The evolution of the Internet of Things (IoT) has significantly intensified the interconnectivity between various entities. The robust advancement of information technology has ushered in a societal upswing while simultaneously triggering an exponential increase in data volumes. Consequently, the efficiency of access to storage systems and the reliability of data are severely challenged. Researchers are actively seeking efficient solutions to these challenges. The RAID storage system, with its commendable access performance, excellent scalability, and relative affordability, has become a preferred choice for the storage servers of numerous enterprises. This paper delves into the workings of RAID 6, erasure codes, and capacity expansion, thereby exploring the feasibility of various capacity expansion strategies. Effectively, a well-designed expansion scheme can mitigate issues related to insufficient storage capacity. Simultaneously, the configuration of the code plays a crucial role in determining the expansion time and consequently influences expansion efficiency. Overall, the information and findings presented in this study contribute to enhancing our understanding and management of storage systems in an increasingly data-intensive era.

Keyword: RAID6, erasure, expanded codes.

1. Introduction

Firstly, modern information technology is progressively assuming a central role in our lives. Its influence seeps into every corner of our existence, becoming increasingly ubiquitous. The rapid advancement of technologies such as artificial intelligence, cloud computing, big data, and 5G has driven exponential data growth, creating an increasingly noticeable bottleneck effect. The traditional computer system, initially built around the CPU and memory, is transitioning towards a memory-centric structure, leading to the evolution of storage systems into relatively independent entities [1].

Secondly, the dramatic surge in network users and the swift expansion of application domains have resulted in an unimaginable volume of data. This has imposed a significant strain on data center storage capacity. Long-term exposure to high loads makes the storage media within storage systems more prone to damage. Beyond the losses incurred from system downtime, the financial and temporal cost of data recovery is high. For many businesses that operate on real-time or near-real-time data, such blows can be severe and potentially disastrous [2]. The RAID memory offers excellent

DOI: 10.54254/2755-2721/32/20230187

accessibility and scalability and is relatively cost-effective, making it the primary choice for many corporate memory servers.

The following provides a brief overview of a common RAID 6 code, known as H-code. The H-code comprises an array of (p-1)×(p+1). H-code encoding includes two types of check chains: the reverse diagonal check chain and the horizontal check chain. The H-code uses the layout of the anti-skew check block, evenly placed along the disk array's diagonal, to enhance partial stripe write performance. The horizontal check chain of the H-code ensures optimal continuous data writing performance, and the horizontal check block possesses a special horizontal check disk, as depicted in Figure 1.

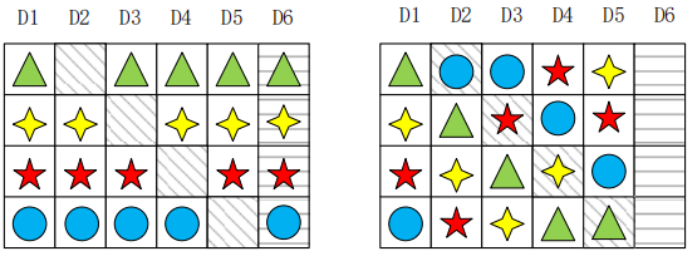

Figure 1. Horizontal check layout and anti-skew check layout of h code (photo/picture credit: original).

2. Relevant theories

2.1. RAID6 technical analysis

The full name of RAID 6 is a separate hard disk array with two independent parity data. To ensure that data is not lost when two hard disks go offline, two different verification algorithms are required [3]. In this way, when two hard disks are disconnected, the data on the disconnected hard disks can be deduced and recovered by combining the equations according to two different calibration algorithms. According to the People's Post and Tele communications (2017), the usual practice is that the first check data is generated by a traditional XOR algorithm, and the other check data is calculated by a reversible function, and the result is generated by XOR again. At present, the most common implementation is to convert the data in Galois Field, and then use the XOR operation to generate a second copy of the verification data. The heart of RAID 6 consists of two copies of the check data to ensure data safety in the event that two hard drives fail simultaneously. RAID6 also achieves better random I/O performance because it inherits the characteristics of striped and distributed parity data storage from RAID 5.

DOI: 10.54254/2755-2721/32/20230187

2.2. Erasure correction code

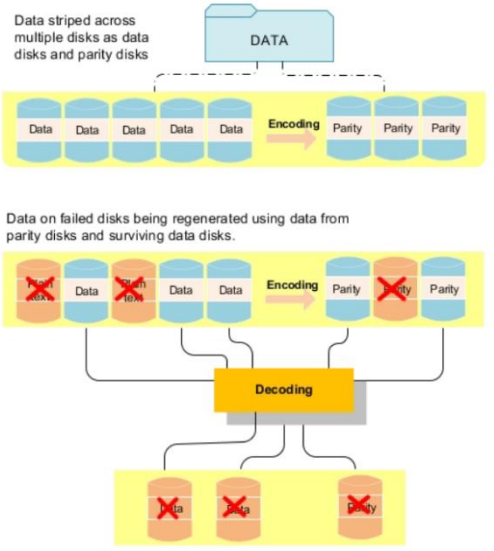

Figure 2. Erasure coding schematic diagram (Photo/Picture credit: Original).

Erasure coding is an advanced error correction technique, as illustrated in Figure 2. It is a fault-tolerant coding technology that can recalculate lost data blocks based on the remaining data blocks and parity blocks. Erasure coding has several advantages that bolster the reliability, availability, and performance of storage as compared to replication. Reliability: Objects are encoded as data and parity blocks and are distributed across multiple nodes and locations [4]. This decentralized method provides a safeguard against site and node failures. Erasure coding improves reliability over replication at a comparable storage cost. Availability: In the context of storage systems, availability refers to the ability to retrieve objects when a storage node fails or becomes inaccessible. Storage Efficiency: For instance, a 10MB object duplicated at two sites would utilize 20MB of disk space (two copies). In contrast, an object encoded between three sites using a 6+3 erasure coding scheme would only use 15MB of disk space. Despite its advantages, erasure coding has some drawbacks. It increases the number of storage areas and locations. In contrast, if only data objects are replicated, a single copy should suffice at one storage location. When erasure coding is implemented across locations distributed in different regions, retrieval delays increase. Extracting fragments of an object encoded with erasure and distributed in remote locations over a WAN connection will take longer than retrieving an object that is replicated and stored locally. Moreover, the utilization of WAN network traffic for retrieval and repair is high when using erasure coding across geographically dispersed sites, especially for objects that are frequently accessed or repaired over a WAN connection. This process also escalates the utilization of computing resources.

2.3. Expansion

To better describe the performance of the RAID expansion solution, we introduce the following basic concepts:

Data: The user's information is stored in a string, which is called data. Block: The most basic unit of storage. It can be classified into data blocks and parity blocks according to the information it carries. Data blocks carry user information, and parity blocks carry redundant information. Stripe: Divides continuous data into data blocks of the same size and stores each piece of data on different disks. Strip: The value can contain only data blocks or parity blocks, or both. Parity Chain: A computing chain consisting of a check block and the data block that generates it. According to the different coding rules of the check chain, it can be divided into: row check chain, oblique check chain and reverse oblique check chain. Encoding: According to certain calculation rules, the generation of redundant data, then

DOI: 10.54254/2755-2721/32/20230187

the process is called encoding. Decoding: The process of making use of both surviving and redundant data to recover when data is lost. Horizontal code: According to the layout rules of erasure codes, the check blocks are stored in separate check disks, then this encoding is called horizontal coding. Vertical code: According to the layout rules of erasure codes, the parity block and data block are stored in the data disk, and this coding is called vertical coding. Scaling up: Adding storage media to RAID to increase storage capacity. Scaling down: Cutting down the disks in the RAID to eliminate damaged disks and reduce energy consumption. Metadata: contains information about data, such as addresses and attributes. It is usually stored in the start location of a disk. Any operation on the data will update the metadata. Read Modify Write: The method of creating a new parity block from the original parity block and the updated data block. Reconstruction Write: An update method that generates new parity blocks from updated data blocks and old data blocks. Degraded Read: A service that continues to provide requests to users despite a disk failure is in degraded read state. Partial Stripe Write: One or more data updates that belong to the same stripe are called partial stripe write [5].

3. Research on 3 erasure code correction technology

3.1. RS coding

The lost data block can be recovered by multiplying (GT) -1 by the codeword vector.

3.2. LDPC coding

The LDPC code is a kind of packet error correcting code with a sparse checking matrix that was proposed by Robert Gallager of MIT in his doctoral thesis in 1963. Its performance is close to the Shannon limit, theoretical analysis and research is easy [6].

The LDPC code is essentially a linear block code that maps the information sequence to the sending sequence via the generation matrix G, i.e. the codeword sequence. For G, there is an exactly equivalent parity matrix H, and all codeword sequences C form the zero space H.

The low-density parity code is an improvement of the parity code.

Because the decoding of parity codes is difficult to apply, there are very suitable decoding schemes using low density. Although LDPC code is not the best code, but because of the existence of a very simple decoding scheme, it makes up for the deficiency of non-optimal code.

3.3. Array coding

Data recovery is completed by XOR operation. The mechanism of single fault tolerance, double fault tolerance and three fault tolerance is used in array coding. Data and redundant data are stored on disks. If some data is faulty, the remaining data is used to restore data. According to whether the array coding meets the optimal storage efficiency, the array coding can be divided into MDS coding and Non-MDS coding. That is, the encoding that meets the optimal storage efficiency is MDS encoding, and vice versa is Non-MDS encoding. In a RAID storage system, the disk that stores data information is called a data disk, and the disk that stores parity information is called a parity disk.

Horizontal parity codes are codes in which data is stored in one area and check information is stored in another area, and the two areas are not on the same disk. Common horizontal codes include RS code, Cauchy RS code, RDP code, Generalized RDP code, etc.

Vertical parity codes are the distribution of data and check information in different areas of the disk. Common vertical coding includes X-Code coding, H-Code coding, HDP Code coding, Short Code coding, D-Code coding, N-Code coding, H-Code coding, P-Code coding, Balance P-Code coding, and balance P-code coding. Hover Code coding, WEAVER Code coding, etc.

DOI: 10.54254/2755-2721/32/20230187

4. Research on capacity expansion technology

4.1. Traditional expansion algorithm

Capacity expansion solutions are divided into two types based on optimization policies: data migration process optimization and minimum data migration.

4.2. Data migration process optimization

This section describes a typical Round-Robin expansion solution. In this expansion solution, including migrating the old disk to the new disk and moving the old disk to the old disk. This solution has a lot of data to migrate, but once scaled out it can respond to a uniform distribution of data and respond very well to user input [7].

Optimized capacity expansion solutions based on data migration such as MDM and ALV are also available.

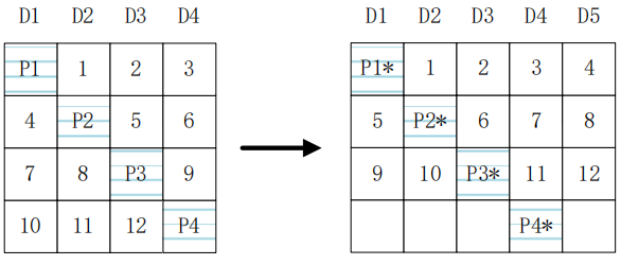

Figure 3. Schematic diagram of comparison before and after expansion (Photo/Picture credit: Original).

The Semi-RR expansion scheme significantly curtails the volume of data requiring migration. As depicted in Figure 3, within the context of the Semi-RR expansion scheme, only data blocks 4, 8, and 12 are transferred from the old disk to the new disk D5. The remaining files are left untouched, thereby minimizing data migration. When considering the least data migration cost for expansion, several methodologies stand out within different encodings, namely, PBM, RS6, Xscale, and FastScale. These techniques aim to streamline the expansion process while ensuring a minimal amount of data is transferred, resulting in significant savings in time and computational resources. The specific application of these methods is represented in Figure 4. Through such strategic management of data migration, we aim to optimize the storage expansion process and improve overall system efficiency.

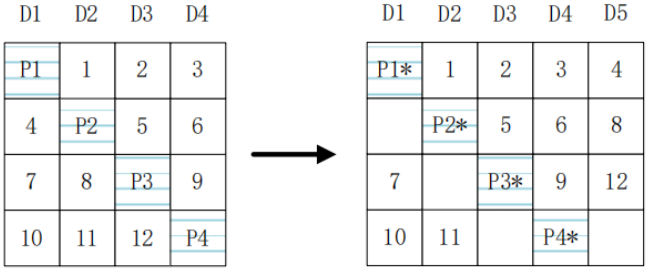

Figure 4. Schematic diagram of comparison before and after expansion (Photo/Picture credit: Original).

4.3. HS6 expansion algorithm

According to the characteristics of H-Code encoding with 2 check chains, HS6 expansion algorithm fully considers the problem of minimum data migration and minimum computing overhead, and carries out its implementation.

Example 1: As shown in Figure 5, before expansion, if p = 5, the RAID-6 storage system has 4 x 6 storage strips [8]. After expansion, the storage strips must be (p-1) x (p +1). Therefore, add disks D5 and D6 forms 6 x 8 storage strips (p = 7).

DOI: 10.54254/2755-2721/32/20230187

D0	D1	D2	D3	D4	D5
0	Q0	1	2	3	P0
4	5	Q1	6	7	P1
8	9	10	Q2	11	P2
12	13	14	15	Q3	P3
16	Q4	17	18	19	P4
20	21	Q5	22	23	P5
24	25	26	Q6	27	P6
28	29	30	31	Q7	P7
32	Q8	33	34	35	P8
36	37	Q9	38	39	P9
40	41	42	Q10	43	P10
44	45	46	47	Q11	P11

Figure 5. RAID-6 storage system with 4* 6 stripes (Photo/Picture credit: Original).

D0	D1	D2	D3	D4	D5	D6	D7
0	Q0	1	2	3			P0
4	5	Q1	6	7			P1
8	9	10	Q2	11			P2
12	13	14	15	Q3			P3
16	Q4	17	18	19			P4
20	21	Q5	22	23			P5
24	25	26	Q6	27			P6
28	29	30	31	Q7			P7
32	Q8	33	34	35			P8
36	37	Q9	38	39			P9
40	41	42	Q10	43			P10
44	45	46	47	Q11			P11

Figure 6. Add two new disks to a RAID-6 storage system (Photo/Picture credit: Original).

After a disk is added, the HS6 expansion algorithm first standardizes the storage system so that the storage strip meets the condition that (p-1) x (p +1) and p is a prime number. Figure 6. During the standardization of the storage system, the original storage strip structure is preserved to the maximum extent, and the data migration and computing costs are reduced. The standardization of HS6 expansion algorithm adopts the practice of forward tuning of the following data.

D0	D1	D2	D3	D4	D5	D6	D7
0	Q0	1	2	3			P0
4	5	Q1	6	7			P1
8	9	10	Q2	11			P2
12	13	14	15	Q3			P3
32	Q8	33	34	35			P8
36	37	Q9	38	39			P9
16	Q4	17	18	19			P4
20	21	Q5	22	23			P5
24	25	26	Q6	27			P6
28	29	30	31	Q7			P7
40	41	42	Q10	43			P10
44	45	46	47	Q11			P11

Figure 7. Standardization of RAID-6 storage systems (Photo/Picture credit: Original).

After standardization, when p = 7, the 6 x 8 storage strips in the storage system are migrated to ensure that all diagonal check blocks are on the reverse diagonal. The HS6 expansion algorithm migrates certain data from the old disk to the new disk, and the migrated data is only moved in the same check chain, which realizes the minimum calculation cost of horizontal check [9]. At the same time, after the data is migrated to the new disk, the original skew check data is retained to the greatest extent, which minimizes the calculation cost of skew check, that is, the calculation cost is minimized.

DOI: 10.54254/2755-2721/32/20230187

D0	D1	D2	D3	D4	D5	D6	D7
0	Q0	1	2	3			P0
4	5	Q1	6	7			P1
8	9	10	Q2	11			P2
12	13	14	15	Q3			P3
32		33	34	35	Q8		P8
36	37		38	39		Q9	P9
16	Q4	17	18	19			P4
20	21	Q5	22	23			P5
24	25	26	Q6	27			P6
28	29	30	31	Q7			P7
40	41	42		43	Q10		P10
44	45	46	47			Q11	P11

Figure 8. Move Q8，Q9，Q10，Q11 skew check blocks to diagonal diagonals (Photo/Picture credit: Original).

Figure 7 shows that before standardization, Q0 = 1⊕6⊕11⊕12, Q1 = 2⊕7⊕8⊕13, Q2 = 3⊕4⊕9⊕14, Q3 = 0⊕5⊕10⊕15. Two new disks are added, they are standardized into 6 x 8 storage strips. The HS6 expansion algorithm migrates certain data blocks to the new disk to maintain the integrity of the original diagonal check chain to the greatest extent. After data is migrated to the new disk, the diagonal check chain changes due to the expansion of the storage strip, as shown in Figure 8 after standardization. Q0' = 1⊕6⊕11⊕12⊕ New data ⊕36, so Q0' =Q0⊕ new data⊕36; Q1' = 2⊕7⊕8⊕13⊕32⊕37, so Q1' =Q1⊕32⊕37; The check chain controlled by the Q2', Q3' check block is all new data after standardization, and the check chain needs to be reconstructed; Q8' = 3⊕4⊕9⊕14⊕34⊕39, so Q8' =Q2⊕34⊕39; Q9' = 0⊕5⊕10⊕15⊕35⊕ New data, so Q9' =Q3⊕34⊕ new data. It can be seen that the HS6 expansion algorithm utilizes data analysis of the first 4.2.4 diagonal check chains, reduces the 5 XOR operations required for chain construction to 2 XOR operations, and reduces overhead costs. As shown in Figure 9 [10].

D0	D1	D2	D3	D4	D5	D6	D7
0	Q0	1	2			3	P0
4	5	Q1	6	7			P1
	9	10	Q2	11	8		P2
		14	15	Q3	12	13	P3
32		33	34	35	Q8		P8
36	37		38	39		Q9	P9
16	Q4	17	18			19	P4
20	21	Q5	22	23			P5
	25	26	Q6	27	24		P6
		30	31	Q7	28	29	P7
40	41	42		43	Q10		P10
44	45	46	47			Q11	P11

Figure 9. Migrate certain data to new disks for minimal computational overhead (Photo/Picture credit: Original).

5. Conclusion

This paper introduces the application of erasure codes to enhance storage reliability and proposes optimizations for H-code encoding within RAID 6 storage systems. We also seek to optimize the online capacity expansion process for RAID 6 storage systems by presenting a new HS6 capacity expansion algorithm. This unique approach ensures that during the HS6 expansion, data migration only occurs between the old and new disks, while maintaining the integrity of the parity data. Through this process, we manage to maximize the use of available storage, significantly reducing the cost of data migration and computational overheads. It's crucial to note that the primary motivation for these advancements is to improve the efficiency and reliability of RAID 6 systems, and consequently, support more robust and efficient data storage solutions in our increasingly data-intensive world. A

DOI: 10.54254/2755-2721/32/20230187

comparative analysis between this new algorithm and traditional expansion techniques clearly demonstrates the advantages of our approach. By minimizing the cost of data migration and maximizing the utilization of storage, our method can streamline the expansion process while maintaining system integrity, ultimately contributing to the overall performance and reliability of RAID 6 storage systems.

References

[1] Kamyod C. CIA analysis for lorawan communication model[C]//2021 Joint International Conference on Digital Arts, Media and Technology with ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunication Engineering. IEEE, 2021: 394-397.

[2] Yuan Z, You X, Lv X, et al. HDS: optimizing data migration and parity update to realize RAID-6 scaling for HDP[J]. Cluster Computing, 2021, 24(4): 3815-3835.

[3] Maneas S, Mahdaviani K, Emami T, et al. Operational Characteristics of {SSDs} in Enterprise Storage Systems: A {Large-Scale} Field Study[C]//20th USENIX Conference on File and Storage Technologies (FAST 22). 2022: 165-180.

[4] Kingsmore K M, Puglisi C E, Grammer A C, et al. An introduction to machine learning and analysis of its use in rheumatic diseases[J]. Nature Reviews Rheumatology, 2021, 17(12): 710-730.

[5] Cheng Y, Wei J, Tan X, et al. Research on key technologies of data-oriented intelligent campus in 5G environment[C]//2022 2nd International Conference on Consumer Electronics and Computer Engineering (ICCECE). IEEE, 2022: 203-208.

[6] Bennett H M, Stephenson W, Rose C M, et al. Single-cell proteomics enabled by next-generation sequencing or mass spectrometry[J]. Nature Methods, 2023, 20(3): 363-374.

[7] Jiang T, Zhang G, Huang Z, et al. {FusionRAID}: Achieving Consistent Low Latency for Commodity {SSD} Arrays[C]//19th USENIX Conference on File and Storage Technologies (FAST 21). 2021: 355-370.

[8] Mohsan S A H, Khan M A, Noor F, et al. Towards the unmanned aerial vehicles (UAVs): A comprehensive review[J]. Drones, 2022, 6(6): 147.

[9] Alzahrani A, Alyas T, Alissa K, et al. Hybrid approach for improving the performance of data reliability in cloud storage management[J]. Sensors, 2022, 22(16): 5966.

[10] Wu Y, Dai H N, Wang H, et al. A survey of intelligent network slicing management for industrial IoT: Integrated approaches for smart transportation, smart energy, and smart factory[J]. IEEE Communications Surveys & Tutorials, 2022, 24(2): 1175-1211.