A new structure of Lustre file system based on Artificial Intelligence

Hongxia Wen[1]

[1] Shengma Intelligent Technology Co., Ltd
zty_urb_iot_8@163.com

Abstract. In today's data age, the storage of big data calls for new storage technology. After more than half a century of development and continuous technological innovation, volume holographic storage technology has brought the dream of storing several terabytes of data on one optical disc one step closer to reality. This unprecedented advantage of data storage technology will lead the strong return of optical discs and bring a qualitative leap to the electronic information industry.

Keywords: Lustre, AI, Storage technology.

1 Introduction

The development of science and technology is bringing mankind from the era of information technology (it) to the era of data technology (DT). According to the statistics of American International Data Corporation (IDC), the global data volume will double every two years in the next few years and will reach 40zb (zetta bytes, 1021)) in 2020. The social demand for data storage is increasing exponentially[1]. It will be difficult to meet the traditional way of increasing the number of disks[2].

At present, the storage density of commonly used storage technologies[3], such as optical disk and hard disk technology, is limited by the size of two-dimensional planar memory storage point[4]. With the continuous development of technology, the size has gradually approached its physical limit, which is very difficult to further reduce. How to break through the shortcomings of existing storage technology and meet the needs of massive data storage in the information age, new data storage technology has become a research hotspot.

1.1 The advanced nature of the project

The development trend of storage is cloud, distributed and flash. The development of AI makes the whole computing industry face a major opportunity and challenge. Computing architecture brings many challenges, but also drives the new direction and new requirements of storage: first, intelligent storage; The second is the optimization of storage in AI application scenarios.

There are three main storage modes: semiconductor storage represented by solid-state disk, magnetic storage represented by mechanical hard disk and optical storage

EAI Urb-IoT 2021, December 20-21, Shenzhen, People's Republic of China

DOI 10.4108/eai.20-12-2021.2315029

represented by Blu ray storage. At present, the market share of solid-state disk is expanding steadily and gradually replacing the mechanical hard disk. Solid-state disk has been widely used as the underlying storage device in large-scale storage systems, and there is a trend of flash memory. Its advantage is faster reading and writing speed, but it has the disadvantages of limited erasure times and charge loss. It is not suitable for long-term storage of data, and it also faces the problem of limited storage density (550 GB / in2).

The storage technology based on magnetic storage mechanism not only has the problem of limited service life (hard disk 5 years), but also the data is stored in the two-dimensional plane of the disk surface, and the storage density (735 GB / in2) has approached the theoretical limit of its technology.

Holographic optical storage has obvious advantages in data long-term storage cost and energy consumption[5], but its disadvantage is relatively slow. Therefore, the ideal storage system is a hybrid storage system of flash memory and holographic optical storage system.

2 Technology development trend

2.1 Development of media technology

Among semiconductor memory, hard disk, magnetic tape and other storage media, optical storage has the significant advantage of low cost per unit storage capacity. The existing hard disk storage technology has gradually reached the limit of storage density; Due to the characteristics of light wave, optical storage can adopt a variety of storage technologies such as multi-layer, multi-order, multi-dimensional and nano super-resolution, which are difficult to be realized by other media, and the storage density will be significantly improved[6].

Holography is essentially a light wave recording method. The idea of using holography for data storage was first put forward by van heerdan in 1963. However, due to the limitations of photoelectric and electro-optic conversion equipment, the technology has made slow progress. In the 1990s, especially from 1995 to 2000, holographic storage ushered in a research upsurge and entered the stage of intensive laboratory research. Under the auspices of the national Storage Industry Federation of the United States, 12 units including DARPA, IBM and Stanford University jointly established a cooperative organization and implemented two holographic data storage projects. Subsequently, many volume holographic storage and application systems came out successively. Research institutes and universities all over the world have carried out research, published numerous papers and monographs. After 2000, volume holography began to move towards the stage of practical and commercial research. Large companies such as American general motors, Japanese Sony and Hitachi have carried out research on the commercialization of volume holography. Commercial companies with volume holography storage as the core technology have also appeared in Europe, America and Japan, such as inphase (now akonia holography) in the United States and optware in Japan, and have pushed out the principle prototype.

At home, Tsinghua University has been following up since the 1990s, studied a variety of principle prototypes and published a large number of high-level papers. At the same time, Beijing University of technology has also continued to carry out relevant research, made remarkable progress, studied a variety of principle prototypes, and published a monograph on volume holographic storage. In recent years, under the leadership of Tan Xiaodi, inventor of coaxial holographic storage technology, Beijing University of technology has continuously carried out the research work of holographic storage technology, and proposed a coaxial volume holographic storage system based on phase and amplitude coding

3 Key contents proposed

3.1 File system design

More and more data are stored in memory, and new data storage methods such as memory file system and memory database have emerged. The external memory represented by disk will gradually change to the medium for storing non hot data and backup data. For hybrid storage system, we should give full play to the advantages of optical storage medium, such as large capacity, long life, low cost and low energy consumption, It is also necessary to overcome the shortcomings of slow speed and low performance of optical storage system and realize the seamless connection with the existing information system.

CD is a media that can only be written once, so the file update operation can only be realized through the file system, and this feature enables the CD library to provide multi version management of files

Small file IO access performance tests the performance of metadata. On the one hand, it should support massive files, on the other hand, it should also support efficient metadata operations.

This project will design a new file system to support hybrid storage mechanism. Provide data version management.

3.2 High performance distributed file system for AI applications

For AI scenarios, distributed storage faces three major challenges: storage of massive files, access performance of small files, and directory hotspots. For massive files, the accuracy of the training model depends on the size of the data set. The larger the sample data set, the more accurate the model is. Usually, the number of files required for training tasks is in the order of hundreds of millions or billions, and the storage requirement is to be able to carry billions or even tens of billions of files.

Small files, many training models rely on pictures, audio clips and video clips. These files are basically between several KB and several MB. For some feature files, they are even only dozens to hundreds of bytes.

Read more and write less. In most scenarios, the training task only reads files and starts to calculate after reading the files. Intermediate data is rarely generated. Even if a small amount of intermediate data is generated, it will choose to write locally and

rarely choose to write back to the storage cluster. Therefore, it will read more and write less, write once and read many times.

Directory hotspots. Due to the uncontrollable data organization mode of business departments during training, the system administrator does not know how users will store data. It is likely that users will store a large number of files in the same directory, which will cause multiple computing nodes to read this batch of data at the same time during training, and the metadata node where the directory is located will become a hot spot, resulting in performance problems during training.

Based on Lustre file system, this scheme will realize the support of lustre's portals network abstraction layer for Infiniband, and add tools such as data version management and mixed media access.

Lustre file system is a distributed file system for next-generation storage devices. It comes from the coda project research work of Carnegie Mellon University. lustre uses object-based disk as the storage device of the whole file system, uses Sandia's open ponals network transmission mechanism (now renamed lnet), supports a variety of networks (Ethernet, Infiniband, etc.), and optimizes the reading and writing of large files(see Fig. 1).

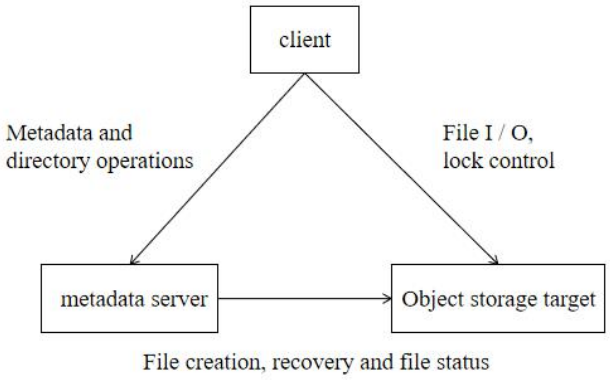

Fig. 1. Structure of Lustre file system

4 Main innovation

Data security: it has the advantages of resisting natural disasters, magnetic storms and human data deletion and modification; Independent and controllable: it can realize complete domestic intellectual property rights; Green energy saving: it can save energy consumption by more than 80% and maintenance cost by more than 50%; Long term storage: the service life of hard disk and tape is only 5-10 years, and the service life of optical storage is more than 50 years. Low procurement cost: Blu ray storage needs no procurement for at least 50 years; The procurement of air conditioning power and quantity, power supply and UPS battery will be greatly

reduced. Low operating cost: 1 / 10 of energy consumption, cooling and power consumption / reduction of equipment maintenance cost and human resource cost.

The next generation cloud storage system integrates distributed storage technology, uses standardized hardware facilities to construct storage pools, virtualizes the existing storage facility space, interconnects and breaks the data scheduling barrier; Object, block and file storage services are provided under the unified system; It has the advantages of high reliability and simple management. At the same time, the next generation cloud storage system has flexible scalability and can provide EB level storage capacity.

Holographic storage technology will be an ideal capacity and high-speed storage technology in the broadband era. Cloud storage system features: high performance, capable of providing sufficient performance, covering most of users' business needs, and meeting the needs of high concurrency or a large number of business data analysis. High availability, cloud storage systems need to meet higher requirements of high availability. High reliability of storage and data is the basic support for the continuous development of business activities. Resource dynamic expansion: the cloud storage system can support the dynamic expansion of resources and the dynamic expansion of resource pool, and can be allocated and expanded flexibly on demand. High manageability, cloud storage system can realize automation and visualization in daily deployment, management and monitoring, and improve the manageability of storage resource services. Make efficient use of existing resources, including dynamic monitoring and prediction of storage IOPs, storage throughput and the use of storage capacity, so as to facilitate managers to understand the current situation of storage and timely plan the expansion of storage in the future.

5 Conclusion

Among many new storage technologies, holographic storage technology is expected to become a strong competitor of the next generation storage technology. Holographic storage technology is completely different from today's optical storage technologies, such as DVD and Blu ray disc. DVD and Blu ray discs can only store information on a limited number of layers on the surface of the disc, while holographic storage technology can store data in a three-dimensional space including the thickness of the disc.

In a holographic disc, a hologram, or a two-dimensional pattern representing information bytes, is written into the disc in the form of volume holography, and then read by light diffraction. Since the holographic optical disc can be stored by using the thickness of the optical disc, its storage capacity is 100 times stronger than that of the prior art. Therefore, volume holographic storage will bring a new change to the optical disc.

References

1. EMMANUEL KOBINA PAYNE. Design of Distributed Energy Resources Integration for Active Demand Power Delivery as A Microgrid System[D]. JiangSu university, 2019.
2. Khalid Ibn Hassan. SPOC-based Flipped Classroom with Freshmen in Chinese EFL Context:Experiences and Percetions[D]. HuaZhong normal university, 2020.
3. Lei Wang, Ci-Hui Yang, Shan Gai, Jing Wen. Current Status and Future Prospects of Conventional Recording Technologies for Mass Storage Applications[J]. Current Nanoscience, 2014, 10(5):
4. Andrey S.Sokolov, Haider Abbas, Yawar Abbas, Changhwan Choi. Towards engineering in memristors for emerging memory and neuromorphic computing: A review[J]. Journal of Semiconductors, 2021, 42(01):37-65.
5. Ye Shen, He Shen, Dongyu Guo. Recent developments in regenerative ophthalmology[J]. Science China(Life Sciences), 2020, 63(10):1450-1490.
6. Robert Fischetti, Shenglan Xu, Oleg Makarov. High Brightness Beam for Microcrystallography at GM/CA@APS[J]. Acta Crystallogr A Found Adv,2014,70(Pt a1):