A new generation cloud storage system integrates distributed storage technology

Yuyang Feng[1]

[1] Shenzhen Tuoyuan Energy Technology Co., Ltd.， Shenzhen 518172, China
zty_urb_iot_10@163.com

Abstract. The next generation cloud storage system starts from the practical needs of the industry, faces the private cloud scenario of the data center, realizes large-scale and high-capacity storage resource pool, integrates and replaces existing storage facilities, and supports all kinds of OLTP or OLAP business applications. In order to provide effective support for various decision support systems and R & D test systems; Break through the performance bottleneck of random access to massive data; Solve the problems of data security and smooth storage expansion.

Keywords: First Keyword, Second Keyword, Third Keyword.

1 Introduction

1.1 Research and development abroad

Holography is essentially a light wave recording method[1]. 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[2]. 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[3]. 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.

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

DOI 10.4108/eai.20-12-2021.2315031

The European Computer Manufacturers Association (ECMA), a well-known standardization organization, recently announced that it would establish the 44th Technical Committee (tc44) to formulate the standards of his (holographic information storage). These companies include: CMC, CMC magnetics Corporation; Fuji Photo Film Co., Ltd. Palus Technology Co., Ltd. (PU). In 2006, optware released 30g holographic memory card and driver. Strategic Media Technology. Optoware's in-line holographic technology.

Inphase's holographic storage drives and optical discs were put into OEM production at the beginning of this year. Now they are finally put on the market in batches, and have won many large enterprises and government customers, including Locke Martin, Turner Broadcasting under Time Warner, the U.S. Geological Survey, Deutsche Bank, the European Space Agency, Siemens Medical Volkswagen and others are also interested in adopting this new technology.

In 2018, GE Global R & D center announced that it has successfully developed a new type of micro holographic storage material optical disc[4]. This holographic disc with ordinary CD size can reach 500GB, which is equivalent to the capacity of 20 Blu ray discs, 100 DVD discs and the whole hard disk of a high-capacity desktop computer.

In 2011, the City University of Hong Kong and Shanghai Institute of Optics and mechanics jointly proposed a phase only common optical path storage scheme, which effectively realizes the homogenization of the spectral plane through special coding and the use of high pass filter. In 2013, Japanese researchers proposed the "self referential" method, which can realize holographic recording and reading without using reference light. In 2016, the Japanese research team proposed to use digital holography to realize phase storage and multi gray scale amplitude storage. The storage structure of this method is also similar to the coaxial structure. The difference is that phase SLM is adopted[5]. The phase distribution uploaded by SLM needs to be calculated by digital holography technology according to the data to be stored.

In terms of polarization holographic storage, in 2013, the research team of Taiwan Jiaotong University reported the research on polarization holographic storage using 2 mm thick PQ / PMMA material. It was found that using circularly polarized light storage, the material will show better modulation and sensitivity. In 2016, researchers from Utsunomiya University in Japan reported the recording method of coaxial polarization holography, and realized the recording of polarization holography on the basis of coaxial structure for the first time.

Although the use of phase and polarization can increase the manipulation dimension of volume holographic storage and bring some unique characteristics, the use of phase and polarization can increase the storage density and solve the storage density bottleneck, which needs to be further studied.

1.2 Domestic research, development and industrialization

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.

In recent years, domestic enterprises rise and gradually break the monopoly of foreign enterprises. The core technology of holographic optical storage management system has laid a solid foundation for future industrial development. In order to make China a world leader in new optical storage technology, greater efforts will be invested in capacity, system software and application software for more in-depth research and development.

A number of technical indicators of domestic massive holographic disc inventory products with independent intellectual property rights are in a leading position; The optical storage technology and application management software developed in this project have been industrialized. Including: cold data storage application management software (rzdz No. 1267301); Blu ray storage file system software platform software (rzdz No. 1636033); Big data multiple storage classification - Software (ruozhdz No. 1539623) and other intellectual property rights.

1.3 Market demand

At present, the mainstream long-term storage methods of information rely on magnetic recording technology. The reliable storage time of hard disk is only 5 years, and its normal operation consumes a lot of energy. As data backup and archive storage, the service life of tape is only 10 years, and it should be transferred every 2 years or so to prevent data loss caused by tape mildew and adhesion. In this case, a last resort is to transfer the information to a new medium every few years. This transfer method has a huge workload, is prone to operational errors, and has huge cost and energy consumption. Under the trend of increasing data volume, it puts more and more pressure on data management.

Automatic tape library is required for long-term storage of large amount of data in the data center, and automatic optical disk library is also required for storage of large amount of information on optical disk. However, compared with the capacity of hard disk and tape (hundreds of GB to 1TB), the capacity of early optical storage devices is too small. Moreover, small devices only support hundreds of optical discs, which is difficult to compete with the huge capacity of tape libraries.

However, with the emergence of new optical storage technology and the sharp increase of capacity, this situation is changing. The storage capacity and volume density of optical storage have reached the same level as that of tape library, and the ultra long-term storage capacity is far unmatched by tape. As an information base for ultra long-term storage, the main indicators are capacity, volume density and reliable storage time. The volume density of optical storage devices reaches the same level as that of tape library, and the storage time is much better than that of tape library. Optical storage devices will soon replace tape library and become the mainstream of ultra long-term storage of huge amounts of data[6].

China's data centers are developing rapidly, with a total of more than 400000, and the annual power consumption exceeds 1.5% of the power consumption of the whole society. The pue of most data centers is still generally greater than 2.2, which is far behind the international advanced level, and has great energy-saving potential. At the same time, the data center produces a lot of greenhouse gas emissions and consumes a lot of water resources. Its equipment is abandoned, resulting in great pollution, which brings great challenges to resources and environment.

The Ministry of industry and information technology issued the pilot work plan of national green data center and the guiding opinions on the construction layout of data center, and put forward the layout guidance and requirements that should be observed in the construction of data center. Create 100 pilot green data centers around key areas, improve the energy efficiency of pilot data centers by more than 8% on average, formulate 4 national standards related to green data centers, promote advanced and applicable technologies of green data centers, and formulate guidelines for the construction of green data centers. The project products meet this demand.

2 Overall scheme architecture of cloud storage system

The core of the next generation cloud storage system is to uniformly manage storage resources, face the cloud platform and provide diversified data services. The next generation cloud storage system decouples the application from the underlying storage and does not depend on the binding of traditional devices and application manufacturers. In the future, in the process of comprehensive transformation of data center and overall cloud, realize the linkage of storage, computing and network resources, and comply with the transfer of data value chain to service.(see Fig. 1)

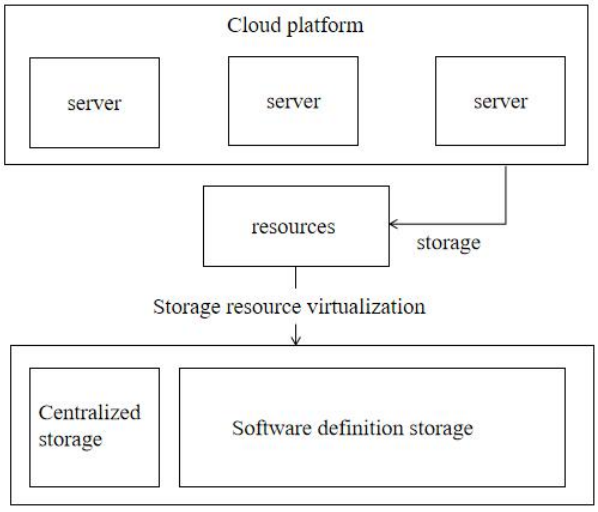

Fig. 1. Architecture diagram of new generation cloud storage system

The next generation cloud storage system is mainly composed of software defined storage system based on distributed architecture and lightweight heterogeneous storage unified management components.

The software defined storage based on distributed architecture runs on the standard x86 server, uses virtualization technology to virtualize the storage resources in the cluster into a storage pool, and provides block device, file and object storage services upward. At the same time, software definition storage has high performance and can easily meet the requirements of various high load management, including business critical applications and core business systems; The application of multi copy and strong consistency technology provides high availability; Strong horizontal expansion capability provides great flexibility and convenience for management and maintenance brought by business expansion.

The lightweight heterogeneous storage unified management component realizes the unified automatic management of distributed storage and centralized storage. The distributed software defines storage and realizes the monitoring, operation and maintenance of the storage system by opening the control interface of the storage system for the storage unified management component. Through the open interface, the heterogeneous storage unified management component can realize the resource division and service arrangement of the distributed storage system, and divide the centralized storage devices into virtual volumes based on different QoS policies to serve the cloud platform, so as to realize the linkage with computing and network.

3 Overall scheme architecture of cloud storage system

Software defined storage technology based on distributed architecture provides a variety of storage services including objects, blocks, and files. It has the advantages of high reliability, simple management, flexible scalability, and can provide PB to EB level storage capacity.

The software defined storage technology based on distributed architecture virtualizes the hard disks of all servers into several resource pools, provides functions such as creation / deletion and snapshot of virtual volumes, and provides storage services in the form of northbound virtual volumes.

The software defined storage system is divided into five layers: hardware device layer, engine driver layer, feature function layer, service interface layer and monitoring management layer.(see Fig. 2)

The software defined storage system based on distributed architecture provides different levels of IOPs, bandwidth and other service performance through standard x86 server and different disk media, including traditional mechanical disk HDD, sata-ssd and pcie-ssd. At the same time, the wide application of 10GE network card also makes the system faster in the process of transmission and reconstruction.

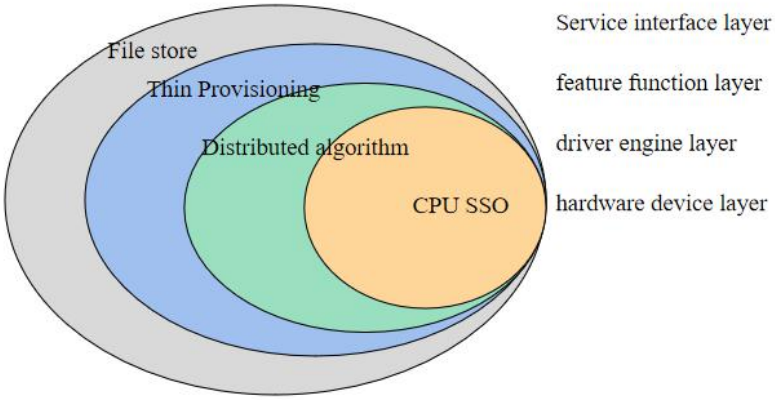

Fig. 1. Software defined storage system hierarchy diagram

4 Conclusion

The software defined storage system based on distributed architecture uses distributed algorithms (such as crush, DHT, etc.) to randomly distribute the data among all disks in the cluster, avoiding the generation of data storage hotspots. The storage of data provides high availability through multiple replicas, each replica is scattered on different servers, and can follow strong consistency according to business requirements.

The software defined storage system based on the distributed architecture can realize thin configuration, that is, it supports dividing the size of storage volumes in advance, but automatically increases according to the size of data writing when adding and distributing, so as to save available storage space. At the volume level, real-time QoS can be realized and the restriction attributes attached to the volume can be adjusted. At the same time, for the needs of business, the system also supports online capacity expansion and reduction to ensure that other volumes can obtain enough space.

References

1. Rao Fu, Liangui Deng, Zhiqiang Guan, Sheng Chang. Zero-order-free meta-holograms in a broadband visible range[J].Photonics Research,2020,8(05):723-728.
2. Will Wai-kit Ma, Robert Andersson, Karl-Oskar Streith. Examining user acceptance of computer technology: an empirical study of student teachers[J]. Journal of Computer Assisted Learning,2005,21(6):
3. Suresh K. Aggarwal. Introduction of the Indian Society for Mass Spectrometry[J]. Mass Spectrometry Letters, 2012, 3(1):
4. EMMANUEL KOBINA PAYNE. Design of Distributed Energy Resources Integration for Active Demand Power Delivery as A Microgrid System[D]. JiangSu University, 2019.

5. Dan Dan, Baoli Yao, Ming Lei. Structured illumination microscopy for super-resolution and optical sectioning[J]. Chinese Science Bulletin, 2014, 59(12):1291-1307.
6. Karim Khan, Ayesha Khan Tareen, Muhammad Aslam. Challenges, and Prospects in Two-Dimensional Photo-Catalyst Materials and Environmental Remediation[J]. Nano-Micro Letters, 2020, 12(12):33-109.