

## The Research of Digital Library Mass Information Storage System Architecture

Lihua Geng<sup>1, a</sup>

<sup>1</sup> Changchun University of Science and Technology, Changchun, China

<sup>a</sup>glhchangchun@163.com

### Abstract

The university digital library has changed the existence and the service mode of traditional library, and combined with the campus network platform, realized retrieval and sharing of library literatures for users networked and simplification, and realized the library office automation. In this paper, combining with the mainstream mass information storage technology that includes DAS (Direct Attached Storage), NAS (Network Attached Storage), SAN (Storage Area Network) and cloud storage, to design the university digital library network information storage system architecture that data resources are hierarchical distributed storage. This architecture meets the digital library storage system's requirements of high capacity, easy expansion and data resources safety.

*Key Words: university digital library; DAS; NAS; SAN; cloud storage; hierarchical distributed storage*

Digital library is a management system for broadband multimedia networks and massive information. It is an extension of traditional library functions. Its storage object and technology has been far beyond the traditional library. Digital library data resources are not necessarily stored centrally in a specific library, which can be distributed in different fields. Therefore, the accurate definition of the digital library is: digital library is a very large scale, distributed storage, and can the huge amounts of digital information resource for library information retrieval. Digital library implements the digital storage of information resources, information resources sharing, information transmission network, and the user as the center of the service mode . In view of the positive significance of the digital library, create your own digital library is desperately needed .

## 1 The function of the digital library

Data of the digital library mainly include two categories, the business data and electronic data. Data of the digital library mainly include two categories, the business data and electronic data. Among them, the business data is the user's borrowing record book, bibliographic information and billing, etc., the electronic data is electronic books, electronic journals, such as audio and video digital resources. According to different management and operation data object, the digital library functions mainly include resource management, library management and customer service, as shown in figure 1.

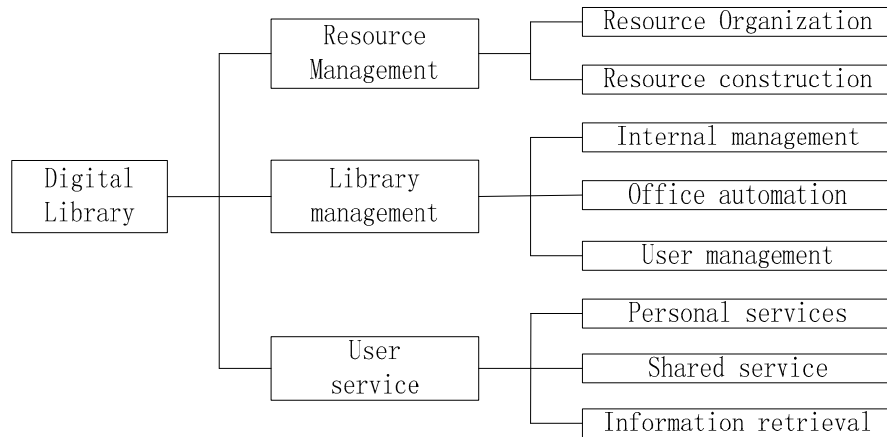


Figure 1 digital library system function structure

Among them, the resource management subsystem of resource organization and resource construction is conducted in the digital library storage system. Therefore, the data resources storage is the core of digital library management system .

## 2 Commonly used massive information storage technology

Now commonly used massive information storage technology, including DAS, NAS, SAN and cloud storage<sup>[1]</sup>, each technology has its unique way of implementation and the advantages and disadvantages.

DAS (Direct Attached Storage) is a Storage device through the SCSI

interface or fibre channel directly connect to the server. Its simple structure, low cost, but not in the true sense of sharing network load, increase the difficulty is big, high maintenance costs.

NAS (Network Attached Storage) is a special kind of private Network Storage devices, embedded system software, can provide a cross-platform file sharing of large capacity Storage devices. NAS implementation of separate storage device and the server, the effective use of bandwidth, network storage mode is to improve the overall performance of the network. NAS improve the access efficiency, reduce the storage cost, reduce the load of server I/O, and have a strong scalability. The disadvantage is the network bandwidth consumed much.

SAN (Storage Area Network) through a separate Storage Area Network connects Storage devices and hang on the LAN server group is connected, the data through the SAN network in the high speed transmission between servers and storage devices, the server can access any storage device on the SAN. SAN system has a good network performance, at the same time, the SAN system storage capability is strong, easy to expand capacity, any server can be connected to the storage array, improve the data sharing. However, SAN price is quite expensive.

Cloud storage is a new concept derived from the concept of cloud computing<sup>[2]</sup>. Cloud storage is a service, not a specific device, is made up of many storage devices and servers. Compared to other storage technologies, cloud storage of digital resources has a higher security, reliability, and virtual<sup>[3]</sup>.

Four types of massive information storage technology above can be used in the digital library system, but should be based on the actual situation of their own library, Such as digital resource capacity<sup>[4]</sup>, network architecture, the amount of funds and future planning. Combined with the advantages and disadvantages of various storage technology, create a digital library reasonable efficient storage system architecture

### **3 The storage system of Digital Library**

Library's network storage system using a network attached storage device, the establishment of data centric, provides the user with massive and shared storage services, to realize separation system of computing resources and storage resources<sup>[5]</sup>. Due to the type of digital library information stored there, a large amount of data storage, and high complexity. Therefore, the realization of efficient storage, efficient management, quick access to digital information is facing the challenge of Digital Library. Digital library is a requirement of storage system to store the lowest cost to meet the performance efficiency, capacity, easy extensibility and data security.

Existing form of storage to a certain extent, can satisfy the concentration of data storage and sharing. But also each have advantages and disadvantages. Alone with any kind of way is unable to fully solve the problems faced by digital library storage system. So use a variety of storage, constructing the optimal digital library storage system.

#### **3.1 Digital library storage system architecture**

Performance of the digital library storage system displays in: the storage system security and reliability, in order to ensure safe and reliable operation of the books management system management system 24 hours a day; storage system scalability<sup>[6]</sup>, the amount of data that is in the growing of digital resources, and the size are unpredictable, so storage system should support the expansion of the hardware and software upgrades; universal storage system, the storage system must be applied to a variety of system platform; store the maintainability of the system, the storage system can not only manage the massive data resources, but also to the storage device management simplification, data resource management simple operation; efficient system, guarantee the speed of access to data network system of library users<sup>[7]</sup>. In order to meet the system performance requirements and design based on the framework of digital library storage system of mass data storage as shown in figure 2.

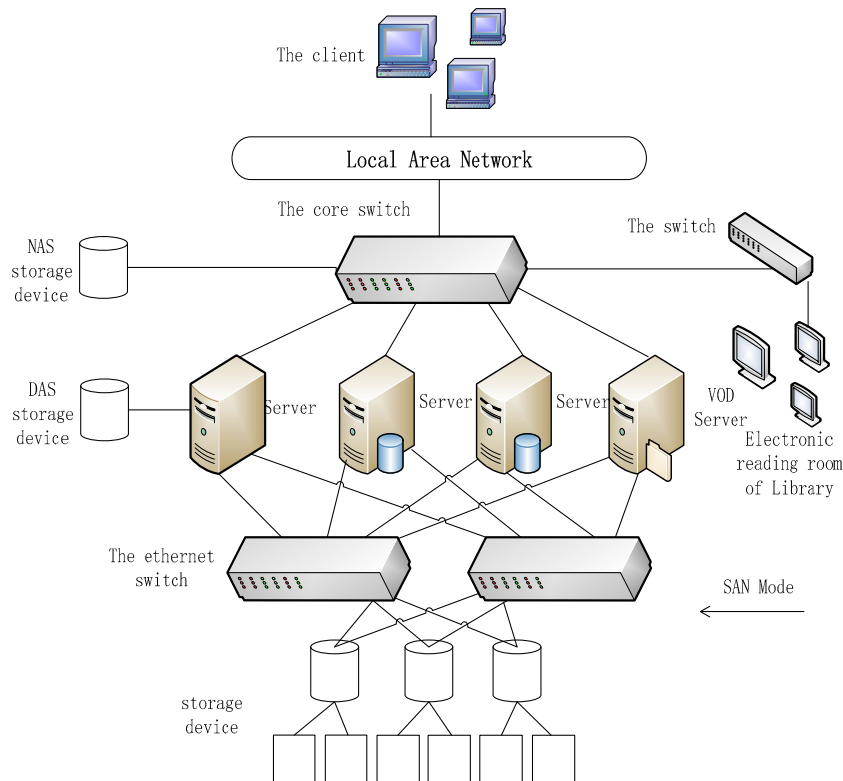


Figure 2 digital library storage system architecture

Library automation integrated system ILAS data using the way of DAS stored in a storage device. Because of ILAS is the core part of the digital library, use the DAS model to improve the reliability of system operation, and use the function of machine hot standby guarantee the reliability of the system<sup>[8]</sup>. NAS mode in the whole storage system instead of a file server, use the NAS the advantages of low cost, flexible light traffic data storage<sup>[9]</sup>. In the storage system is used in the IP\_SAN network storage technology, realizes the data block level data storage.

Digital library storage system, the choice of reasonable storage solutions and appropriate storage devices, this is just the first step to build a storage system, but also perfect storage management technology application to perfect unifies in together, both would then use the cloud storage. Digital library storage system

based on cloud storage architecture is shown in figure 3.

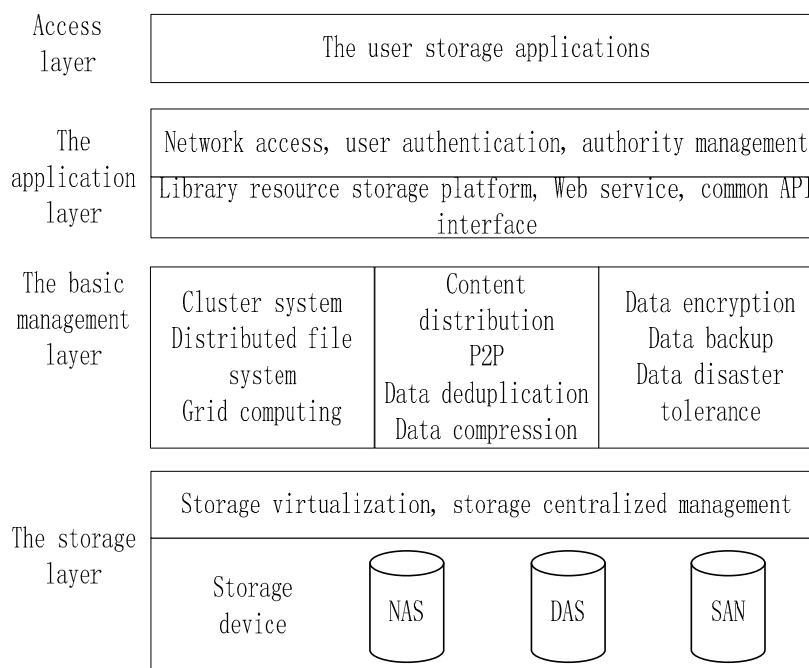


Figure 3 digital library storage system based on cloud storage architecture

The whole storage system is the bottom of the storage layer, provide basic network environment for the whole storage system, the physical storage resources and logical storage resources, etc. The storage device is NAS, DAS and SAN equipment through the network connection. The top is the access layer, any access to the digital library users, can use mobile phones, PC connected to the Internet terminal equipment through the digital library storage platform, meet the demand of users, such as information retrieval and storage.

### 3.2 The storage system architecture

Discovered by figure 2 storage system architecture, in the storage of digital library network topology with two Ethernet switches build SAN storage network, realizes the two switches are using a backup. NAS storage device connected directly to the core switch, and SAN storage network connectivity, the business

data on the NAS storage devices can be in SAN network backup storage devices, improved the security of the NAS data resources.

Hierarchical storage in a storage system in two aspects: The first on the one hand, the digital library business data and access to the very low frequency data resources stored in the NAS storage equipment, access to the high frequency data resources stored in SAN storage devices. On the other hand, SAN network storage system using hierarchical storage technology, the access of high frequency data resources stored in the first level of disk array, visit frequency is low or stale data resources stored in the second storage, the hierarchical storage to improve the user's access speed. SAN storage in the network backup important data resources stored in secondary storage devices, improve the reliability of key data..

This storage system, because of the SAN storage network "data core" in the network center, distributed memory implementation of data resources. This way of storage devices can also be used for the rest of the network data storage, avoid network storage devices of dispersion and repeated investment. So, distributed storage can improve the utilization rate of storage space, save the storage system funds, more convenient for data management and maintenance.

Because of cloud storage is not specific to a device, is a set composed of many servers and storage devices, its core is to realize the transformation of the storage device to the storage service through the application software. So, the problem of data resource storage cloud storage characteristics determine the storage architecture flexible to solve the digital library faces. In short, the performance index of the digital library network topology structure classification of cloud storage distributed storage based on satisfaction, applicable to the construction of storage system.

#### **4 Conclusion**

This paper from the concept of digital library, the digital library function and the now massive information storage technology mainstream start, combined with the actual demand of digital library system, describes the architecture of Digital Library Based on DAS, NAS and SAN storage model of massive

information storage system. In this paper, design of storage system based on cloud storage architectures of hierarchical distributed storage, data resources for digital library storage capacity and capacity expansion easily size according to your request, embodies the high scalability of the storage system. DAS, NAS, SAN and cloud storage technology integration in digital library network storage architecture, realize complement each other, build hierarchical distributed storage system, to build cost-effective data storage network resources. With the mature of cloud storage, the digital library based on cloud storage architectures will enter a new stage of development.

## References

- [1]Fahai Yao.The mainstream technology of data storage and management of Digital Library [J]. Modern library and information technology,2004,(115):16-18.
- [2]Hongli Zhang.Research on framework model of Distributed Digital Library Based on cloud computing platform[J]. Information Science ,2013,31(3):40-44.
- [3]Aiqin Li,etc. Study on the resources of cloud storage model of Digital Library [J]. Modern Information.2012,32(2):48-50.
- [4]Ra´ul Alonso-Calvo,Jose Crespo,Miguel Garc´ia-Remesal,Alberto Anguita,Victor Maojo .On distributing load in cloud computing: A real application for very-large image datasets. Procedia Computer Science[J], 2010:2669–2677.
- [5]Dongying Jiang,etc. Discussion of the data storage security library [J]. Library Work in Colleges and Universities,2011,(3):41-43.
- [6]Lijuan Hou. Discussion on the selection of network storage system in Digital Library [J]. Sci-Tech Information Development & Economy,2010, 20(15):41-43.