

Crucial Data Backup and Recovery Strategy for Campus Network

Liu Wei

Network Center
Shenyang Jianzhu University
Shenyang China
e-mail: liuwei@sjzu.edu.cn

Xu Jingke

Graduate School
Shenyang Jianzhu University
Shenyang China
e-mail: njnke@sjzu.edu.cn

Yang Yi

Building Science and Technology Company
Shenyang Jianzhu University
Shenyang China
cs_5066@sjzu.edu.cn

Liu Xidong

Assets and Equipment Department
Shenyang Jianzhu University
Shenyang China
xd1@sjzu.edu.cn

Abstract—With the rapid development of the campus network date and high-availability, security requirements, the data backup and disaster recovery technology has been widely used, and this technology has become a key aspect of campus IT infrastructure. This paper analyzed the date backup technology, and proposed a plan of campus net establishment data disaster backup system and it is proved well in practical application.

Keywords- campus network; database; snapshot; backup and recovery

I. INTRODUCTION

The rapid development of information technology has greatly promoted the progress of human society. Online information exchange, electronic commerce, office automation, automatic control technology, information technology has greatly reduced the people's labor intensity, easy for people to work, study and life [1]. However, subsequent problem such as data backup, disaster recovery solution has been paid more and more attention.

Disaster recovery data backup is the foundation where aims to prevent the operating system error or system failure resulting in data loss, and all or part of data collection from the application host's hard disk or array that is copied to other storage media in the process [2-5]. Traditional data backup is mainly used in internal or external tape drive for cold backup. However, this approach can prevent operational errors such as human error, and its recovery time is very long. As technology continues to evolve, the massive increase in data, many companies began to use network backup. Usually through a professional network backup data storage management, the software with the appropriate hardware and storage devices can be achieved [6]. The important data inside a computer, files, or record, whether for business users or for individual users is critical. Once it is lost, it will cause incalculable damage, ranging from the hard work hard to pay the accumulated east, and would seriously affect the

normal operation, to scientific research, production and caused a huge loss. In order to protect the production, sale and development of normal operation, advanced business users should take effective measures to backup data. The rest of the paper is organized as follows: Backup method is discussed in section 2. The data backup strategy is described in section 3, and section 4 presents the key technology disaster recovery. This section is the application of data backup and data disaster recovery. Finally, section 5 concludes the paper.

II. BACKUP METHOD

The common methods include three kinds of data backup: LAN backup, LAN Free backup and SAN Server-Free backup [7,8]. LAN-based backup method requires installation on each host. The native tape backup system use the LAN backup strategy when the data is not large, and centralized backup can be used. A central backup server will be installed on a LAN, then the application server and the workstation is configured for the backup server to the client. Central backup server running on the client receives the backup agent's request, and the data is passed to it via LAN management, which is connected to a local tape drive resources. This approach provides a centralized, easy way to manage backup programs, and it improves the efficiency through shared network drive resources.

When we need a large amount of data backed up, the backup time window is tight, and the network prone to clogging. In a SAN environment, storage network can be LAN-Free backup, and the backup server connected to a SAN tape drive in the LAN-Free backup client software trigger. The read data to be backed up will get to a shared tape drive through the SAN backup. This independent network not only allows LAN traffic to be transferred, and its CPU resources require operating the LAN mode. Because fiber Channel connection does not require the server's TCP/IP stack, and some level of error checking

can depend on the fiber channel internal hardware. LAN free backup on the need will take to back up host CPU resources, if the backup process can be completed within the SAN, and a large number of data streams without flow through the server, we can greatly reduce the backup on production systems.

III. DATA BACKUP STRATEGY

Data backup strategies determine the need to back up, recovery. From a backup strategy, now it is the backup which can be divided into three types: full backup, incremental backup, differential backup. There are three data backup methods totally. Full backup is a copy of a given computer or all the files on the file system, regardless of whether it was changed. Daily full backup of the system: the advantage of this backup strategy is that when a data loss disaster, we just use a tape (that is, the day before the disaster backup tapes). Then we can recover the lost data [9]. However, its drawback lies in that: As the daily full backup of the entire system, resulting in substantial duplication of backup data. The duplication of data takes up a lot of tape space, which means increased user costs. Second, the amount of data to back up, so the backup time will be longer. Full backup will have a lot of data movement, customers can choose daily full backups to tape media often directly connected to each computer on. Incremental backups work only after the increase of the last backup, changing the part of the data. It can be divided into multi-level incremental backup and each incremental backup that are derived from changes in the last part. Conduct a full backup on Sunday, and on the next period, six days only need new or modified data backup. The advantage of this strategy is to save the backup tape space, reducing the backup time. But its drawback is that when a disaster occurs, data recovery is too much trouble. In addition, this backup reliability is very poor. In this backup mode, the tape is like the relationship between the same chains, a ring set a ring, and any of a tape will cause a problem with the whole chain out of line. Differential backups work only in the last full backup after changing some of the data. Administrator first need full backup on Sunday once the system starts. Then in the next few days, the administrator will get different data (new or modified) to be backed on tape. Differential backup strategy is used to avoid the above shortcomings of the two strategies, but it also has all of their advantages. First, it does not need to do every day full backup of the system, so the backup time is short, and save tape space; secondly, its disaster recovery is also very convenient. System administrators need two tapes that the tape on Sunday and the day before the disaster the tape, we can restore the system. Generally we use the combination of three strategies: full backup, full backup plus incremental backups and full backup plus differential backup.

IV. KEY TECHNOLOGY OF DISASTER RECOVERIES

Data disaster recovery refers to a remote data system, which is a critical application data locally available copy. In the local data and the application of a disaster, the system ensures at least in the offsite preservation of a

critical business data available. The data can be completed with local real-time copy of production data, and local data can also be more than slightly behind, but it must be available. The main technique is the use of data backup and data replication technology. In the establishment of disaster recovery system will involve multiple technologies, the current key technologies include: remote mirroring, snapshots, virtual storage, IP-based SAN interconnection technology.

Remote mirroring is the main data center and disaster recovery center that is used for data backup [10]. Mirrored in two or more disks or disk subsystems will generate image data with a view of the information stored procedures One system is called the primary mirror, and the other is called from the mirror system. By the master from a mirrored storage system location it can be divided into local mirroring and remote mirroring. Remote mirror is called remote replication, and disaster recovery is the core technology. It maintains data synchronization and remote disaster recovery basis. Remote mirror image of the host by the need will get request of a remote mirror site to confirm the information that is divided into synchronous remote mirroring and asynchronous remote mirroring. Synchronous remote mirroring is the remote mirroring software, by fully synchronized local data to be copied to the remote position; every local I/O transactions are required to wait for the completion of remote copy confirmation and side release.

Synchronous mirroring means a remote copy of the local machine that can always copy the contents of the required match. When the primary site fails, the user is applied to switch to a backup alternative site, and the remote copy can be mirrored to ensure that business continues to perform without loss of data. But it spreads there from a long delay caused by the shortcomings, only relatively short distance applications. Asynchronous remote mirroring ensures the updated view of the remote storage to local storage system, to complete the basic I/O operations. The local storage system gets the request image host I/O operation to complete confirmation. Remote synchronous data replication is the way of background, which makes the performance impact on the local system is very small, with long distance transmission and a small network bandwidth requirements. However, many remote storage sub-systems get used to write the slave that has not been confirmed. Some factors when data transmission fails, the potential data consistency problems. Snapshot technology is a software to backup data quickly and scans the disk subsystem to create a snapshot of the data to back up the logical unit number LAN and snapshot cache [11]. In the fast scan, the backup process will block while the data to be modified quickly copied to the snapshot cache. Snapshot LUN is a set of pointer to snapshot cache and disk subsystems in the same block (in the backup process). In normal operations, while the use of the original data snapshot LAN to achieve a complete backup. It allows users not to be affected in case of normal business, real-time data extraction of current online business. The backup window” close to zero, and the system can greatly increase business continuity, to

achieve a real operating system providing a guarantee. Early primary data center and backup data centers, data backup, mainly based on SAN for remote replication, through the Fibre Channel. Internet technology is used to connect two SAN, remote mirroring. When a disaster occurs, the backup data center alters to the main data center to ensure the system T for continuity. This remote disaster recovery mode, there are some short-comings, such as: achieving the high cost of equipment, poor interoperability across short geographic distance. These factors hindered its further promotion and application.

Virtual storage system provides a dynamic change in the size of the logical functions. It can increase or decrease the system number of physical disks in the logical volume to change the cluster size. This feature provides users with a volume capacity of the immediate requirements of dynamic change. In addition, the storage volume can easily change the volume, move and replace them. After installing the system, only the minimum allocation for each logical volume capacity leaves the remaining disk space. With the development of business, it can be expanded as needed with the remaining space logical volume.

V. APPLICATIONS OF DATA DISASTER RECOVERY

The key problem is how to build enterprise information from reality and build a business to meet the automation of digital systems development needs of disaster. We need to continue research and exploration. Enterprise information management is based on the current practical application of network and data security needs. In this paper, we establish disaster recovery system which is IP-based SAN for remote disaster recovery system, using the disaster recovery industry has reached a mature technology with high cost. The system consists of the primary data center and disaster recovery centers in the protection of data integrity, timeliness and service continuity, basically to meet a variety of business-critical data and information database disaster recovery applications.

There are many disaster recovery technologies on the market proven to choose from, and these disaster recovery technologies are the most important technical differences between the launch platform for data replication and accepted platform. In the implementation of disaster recovery systems engineering, we need to develop detailed disaster recovery plan. Through the development of disaster recovery plans, according to their business model to determine the disaster recovery system design requirements, the system analysis will determine the disaster recovery system based on design parameters. According to the business systems area network environment, we select the appropriate disaster recovery technology. Domestic common disaster recovery solutions are based on the same copy of data between storage subsystems way, generally installed in the storage subsystem using the data replication software to keep business systems, stored data logical volume data storage and backup systems logical consistency. Storage

subsystem disaster recovery technology option is used for replication or asynchronous replication.

There is a major portal set in order to increase customer reach and expand customer base. It paid to individuals and corporate customers to provide efficient mail services. The e-mail system in the process of building, crucial data storage devices made using a simple, easy management, expansion of flexible, easy to be maintained, supporting data sharing, disaster recovery and other supporting data requirements. The disaster recovery system theory is shown in Fig.1.

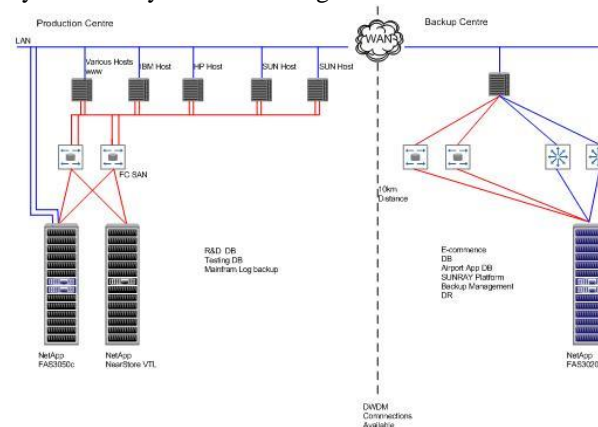


Figure 1. Basic theory of disaster recovery principle

Among them, the crucial data are stored in center fabric switch, disk array, and the backup disk library composed of SAN network storage systems. Primary data center connects to disaster recovery system configuration host according to business needs may include databases, e-mail, WEB, OA servers and other application servers, software components by installing ROSE HA reliability multi-machine environment. Enterprise data is stored in the main data center storage disk array. Meanwhile, they are in the same offsite data center configuration structure of the storage array and two or more backup servers. Through a dedicated disaster recovery software can automate the primary data center and disaster recovery center data, real-time full backup. Two data centers through the use of optical devices composed of single or multimode fiber optic lines are connected to two data centers to provide high-bandwidth, with highly reliable network connectivity support. The system topology for the future of the enterprise to provide high storage scalability, and support for heterogeneous operating system and disk system that can protect existing investments and later storage solutions combine to create from a scalable, easy management, flexibility to adapt to a variety of storage needs of network storage environments.

Ever bright datacenter is the largest commercial server, to focus on its core business model to be used. In order to ensure the continuity of critical business to improve disaster recovery capabilities, every network center disaster recovery program are is necessary. The disaster recovery system using the following topology, as is shown in Fig.2. Database disaster recovery technology uses asynchronous transfer mode, and business center to support multiple backup center database replication

requirements, or multiple business centers to copy a backup center requirement. Database disaster recovery technology also includes the type of storage subsystem, and platform-independent business system server. The database version has a certain relation-ship, with database disaster recovery solutions, and has good use of flexibility. Database disaster recovery technology is only used as a disaster recovery solution for database applications. If we need additional unstructured data, disaster recovery, we also need other disaster recovery techniques as a supplement.

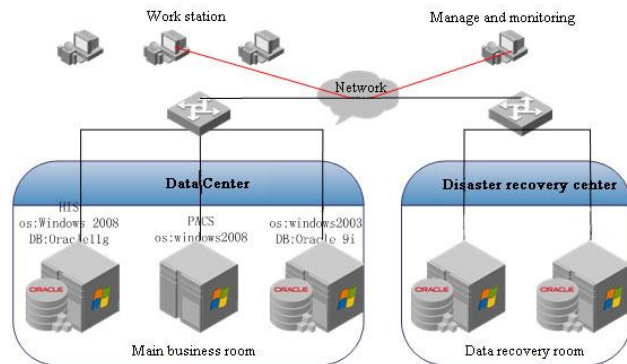


Figure 2. A simple disaster recovery topology graph of campus

In the campus network in local area network, the recovery performance test, service switching, remote replication system for disaster recovery system. Test environment as shown in table 1. From the diagram we can see, the disaster recovery system are more than the original system data storage time, and with the increase of the amount of stored data, temporal growth rate remained stable. Thus, it has perfect performance on the impact the original system of the disaster tolerance system backup.

TABLE I. TEST RESULTS

Environment	CPU /Ghz	RAM	Ethernet	OS
Application server	1.70	128M	RTL8139	Windows server 2003
Main disaster recovery center	2.40	512M	RTL8139	Windows server 2003
Disaster recover control center	1.80	2048M	RTL8139	Windows server 2008
Back application servers	1.70	1024M	RTL8139	Linux centos 5.5

VI. CONCLUSION

With the continuous development of computer information systems and business information to continuously improve the level of the user's core business is increasingly dependent on reliable operation of information systems. The information systems in business-critical data have become a user of the most important asset, with more many companies establishing enterprise-based institutions, the system disaster recovery issues into account. Disaster recovery is a systems engineering, scientific and rational manner with only

disaster recovery system design and deployment, to the corporate network safe operation of the system play a fundamental role in safeguarding. Therefore, how to address the characteristics of enterprises, from the disaster recovery system is based on theory, to meet the requirements of sound corporate network security, and disaster recovery system, with a complete disaster recovery solution or a good disaster recovery system. In practice, constant assessment repeated tests at any time can adjust and improve the company's future development and the application has far-reaching practical significance.

ACKNOWLEDGMENT

This work was supported by Youth Foundation of Shenyang Jianzhu University (NO. 2015031)

REFERENCES

- [1] Yu Jing, Mai Shaohui and Li Lidong, "Data backup scheme in power system EMS," *Automation of Electric Power Systems*, vol.33, Sep.2009, pp. 101-104.
- [2] Javaraiah Vijaykumar, "Backup for cloud and disaster recovery for consumers and SMBs," *Proc. International Symposium on Advanced Networks and Telecommunication Systems*, IEEE Press, Nov. 2011, pp. 15-21, doi:10.1109/ANTS.2011.6163671.
- [3] Bang Jewan, Lee Changhoon and Lee, "Damaged backup data recovery method for Windows mobile," *Journal of Supercomputing*, vol. 66, Aug. 2013, pp. 875-887, doi: 10.1007/s11227-013-0919-6.
- [4] Bang Jewan, Lee Changhoon and Lee Kyungho, "Damaged backup data recovery method for Windows mobile," *Journal of Supercomputing*, vol. 66, Apr. 2013, pp. 875-887, doi: 10.1007/s11227-013-0919-6.
- [5] Li Zhonghua and Zhang Wei, The model construction of incident recovery autonomous management in e-commerce data disaster backup, "Journal of Information and Computational Science," vol. 6, Jun. 2009, pp. 1479-1486.
- [6] Yang YiXian, Yao WenBin and Chen, Zhao, "Review of disaster backup and recovery technology of information system," *Journal of Beijing University of Posts and Telecommunications*, vol.33, Apr. 2010, pp. 1-6.
- [7] Hu Qiaolin, Sun Yipin, Su Jinshu, BAR-BGP: Achieving high reliability interdomain routing through backup as-address advertisement and recovery forwarding, *Computer Research and Development*, vol.48, Dec. 2011, pp. 2242-2252.
- [8] Shuang Kai, and Xie Jing, "Disaster recovery backup system for P2P based VoIP application," *Journal of Computational Information Systems*, vol. 9, Oct.2013, pp. 8099-8109, doi: 10.12733/jcis6851.
- [9] Yu Young Jin, Shin DongIn and Yeom Heon Young, "Backup Metadata As Data: DPC-tolerance to commodity file system," *Journal of Information Science and Engineering*, vol. 27, Jul. 2011 pp. 1193-1218.
- [10] Kermarrec Anne-Mari, Le Merrer Erwan, "Performance evaluation of a peer-to-peer backup system using buffering at the edge," *Computer Communications*, vol. 52, Oct. 2015, pp. 71-81, doi:10.1016/j.comcom.2014.06.002.
- [11] Zhang Dongyan, Zhang Tao and Du Pingxin, "A real-time data backup model and methods based on peer-to-peer network," *International Journal of Information and Communication Technology*, vol.7, Feb.2015, pp. 202-217.