

Enterprise Data Backup and Validation Strategy Selection based on Cloud Computing

Fanli Meng

Jilin Vocational College of Industry and Technology

Keywords: Cloud computing; data, the backup; check.

Abstract. According to the survey, more than 75% of the database data had abuse and security risks. Data security and compliance audit for the necessity of enterprise data management enhanced increasingly. This article embarks from the status quo of information security enterprise data backup and calibration solution based on cloud computing. This scheme provided for the enterprise to obtain cost-effective data integration management method, to reduce risk, reduce cost, and improve the effect of management. Used cloud computing for the enterprise data to improve its management level enhance the competitiveness of provides the beneficial reference.

1. Introduction

The wealth of the enterprise and the most precious is data, in order to ensure the enterprise sustainable development and operation, is to protect the computer based information. Human error, the loss of the hard drive, computer virus, natural disaster and so on are likely to cause data loss, cause inestimable losses to the enterprise. For financial, securities, insurance system, due to the particularity of the industry, computer systems business data loss is a disaster, can lead to system files, the loss of customer data, business data, business will be difficult to normal. At this time, the key question is how to restore the computer system as soon as possible, make it can run normally.

2. Cloud computing

The third IT revolution caused by cloud computing, completely change people access to information, software, hardware resources. Cloud computing refers to the enterprise based on Internet will scale pool of computing resources, storage, cloud development platform to provide to the user, automation, low cost, quick to provide flexible and scalable IT services. Computing is a kind of new information technology and business services consumption and delivery model, not a single information technology innovation, and this model is the core principle of hardware and software resources and is encapsulated in the service, enterprise users can through the network on-demand access to and use, and pay according to resource usage or business scale. Cloud computing as the age of the Internet service mode, put forward by the latest data on small and medium-sized enterprises more and more use of cloud computing service. In information economy, small and medium-sized enterprise most of the employees are young people, they all grew up in the Internet age, accustomed to use the Internet, compared with the traditional way of information they are more likely to accept the cloud computing. The characteristics if cloud computing is shown in table 1.

Can be seen from the table, cloud computing has low invest, short implementation cycle, staff technical requirements is not high, strong expansibility etc., it can well meet the reality in the products and services of the information of small and medium-sized enterprises need. At the same time, cloud computing can help improve data during the process of small and medium-sized enterprise information resources efficiency and reduce the maintenance costs, and also reduces the management cost, better complete enterprise, calibration data backup and network security management.

Due to data backup occupies an important position, in general, a variety of operating systems attached backup programs have this or that defects, so if the data is reliable backup, must choose a special backup software and hardware, and formulate the corresponding backup and recovery plan.

Table 1. The characteristics of cloud computing

The way of information	Traditional information way	Cloud computing
Software and hardware investment	High cost of software, and is a one-time investment, pay for the hardware is very high and need to buy the server, network devices, storage devices, etc.	Software costs low, pay by the usage, only need to purchase a terminal PC hardware
Implementation of the cycle	Longer, at least six months	About 1 to 2 months to complete
Payment method	One-time pay	Monthly or usage billing
Maintenance convenience	The user's responsible	The responsible for Cloud service provider
The new demand	Need to secondary development, the high cost	Increase the corresponding service, convenient and simple and low cost
Expanding	Expanding low	On-demand service, extend is good
Economic management benefit	Not obvious, the traditional information meet the needs of computer operation	Obviously, cloud computing can solve the daily management of software and hardware maintenance, help to promote enterprise business innovation

3. The principle of data backup

The data backup is in order to guarantee the data consistency and integrity, eliminate worries about the system users and operators. Different applications require different solutions to adapt to, in general, a complete backup system, need to meet the following principles.

3.1 Compatible stability.

Backup products system main function is to provide a data protection method, and the stability and reliability of the product itself becomes one of the most important aspects. Firstly, the backup soft must be 100% compatible with operating system, secondly, when the accident happened, can recover data quickly and efficiently.

3.2 Support overall.

In a complex environment, the computer network may include a variety of operating platform, such as various manufacturers of UNIX, Windows, VM, and installed a variety of applications, such as ERP, database, and cluster system and so on. Choose the backup software, to support various operating systems, databases, and typical applications.

3.3 Application of automatic.

Many systems due to the nature of work, how long it took about when backup, backup has certain restrictions. From work time system load is light, suitable for backup. But this will increase the burden of system administrator, due to reasons such as mental state, will also bring potential to backup safety hidden trouble. So, backup plan should be able to provide a timing automatic backup, and use the technology of tape library to automatically change the tape. In the process of automatic backup, should have logging, and automatic alarm when abnormal situation.

3.4 Real-time performance.

Some of the key tasks is to 24 hours non-stop running, at the time of backup, there are some files may still is in a state of opening. So at the time of backup, to take measures, in real time to check the file size, to carry on the transaction tracking, to ensure that all the files in the backup system correctly.

With the continuous development of business, more and more data, update faster and faster, the rest time to back up so much content, backup during work time and affect the performance of the system. This requires that when designing a backup. Try to consider to improve the speed of data

backup, using the method of multiple concurrent operation of tape recorders. To maintain the effectiveness of business systems: Real-time backup on the performance of the business system will produce certain impact, sometimes is very large. How to take effective technical measures to avoid the backup to the server system, database system, the influence of the network system, will be very important.

3.5 The operation is simple.

Data backup is applied in different fields, carries on the data backup operators are at different levels. This requires an intuitive graphical user interface, easy operation, shorten operation personnel of learning time, reduce work pressure, operating personnel to easily set up and complete the backup job.

3.6 Disaster Recovery Considerations.

Copy a tape in the tape library, stored in a place away from the data center, in case of unpredictable disaster data center.

4. Choose a reasonable solution

At present, enterprise, government data center data backup, mostly use the third party backup software, through the key data backup software will regularly back up it to tape library equipment. Even if has a reasonable backup scheme and make the data backup, many clients still plagued by the integrity of data backup and recoverability. Traditional way backup data backup it to tape library equipment, the backup it to tape the data cannot be read at once, so we can't to check the backup data, only through the third party backup software, in the form of disaster recovery, the tape library data recovery to disk devices, to test the backup data is complete. This will undoubtedly bring a lot of trouble to the customer's maintenance and management; and the backup data is usually offline data, when you need according to the historical data to produce reports, with data mining and data analysis, historical data recovery will face many complicated problems.

Server virtualization technology created by the virtual machine is a collection of files on a physical server, And these documents collection includes some application soft wares, such as operating system, database and middle, also contains a virtual application data files on the server. Just need to the virtual machine file collection for backup respectively, can be realized the virtual machine operating system, application software, the data backup.

This way of data backup, have broken the traditional physical server backup, need to purchase a database backup module, using network data management protocol NDMP backup module requirements, and greatly simplifies the complexity of the backup. Administrators only need to virtual machine files for backup, all of the data backup can be realized.

After physical server to server virtualization environment, can not only realize the effective integration and utilization of physical server resources, but also can greatly simplify the data backup process.

4.1 The data backup strategy.

Under the environment of server virtualization, data backup strategy compared with traditional physical server; also have more choice of convenient and quick. Can simplify the backup process greatly. For the virtualization environments data backup server, can be completely separated from the third party software, only through the virtual server software included with the function, can be realized the operating system, application software, the production data backup. Can also through the integration of third-party backup software, the realization of the virtual machine to tape library off-line data backup. Server virtualization environment, can also be very good implementation, from the existing production environment of backup data validation and data statistical analysis, data disaster recovery and other supplementary functions.

(1)Virtual machine snapshot backup and advantages. Through a snapshot of the virtual server function, achieve the fast data pointer level backup. For server virtualization system, provide the database snapshot function, snapshot at some point to be able to record, within the virtual machine

operating system and application software, and the instantaneous state of the production data, and the user can continue for virtual machine snapshot, so as to realize the continuous backup of data.

If we can out of the third party software to realize data continuous fast backup and incremental backup, we only need a very short time to complete snapshot backup process and will not cause the production system load, and reduce the size of the backup data of disk space, and can save a large number of physical disk resources.

(2) Virtual machine clone backup and advantages. Virtual machine clone backup and advantages can provide Virtual server software, and provide a virtual machine clone function, which can be a virtual machine for the complete replication, and save to another virtual machine. Customers through the virtual machine clone function can realize the virtual machine backup. In the backup virtual machine can be stored in the server locally, can also be stored in a long distance through the export way, archive operating system and data offline. And these the backup cloning virtual machine, when need for data recovery, we just need to reimport to clone the virtual machine.

The software for the virtual machine can be realized by virtual machine all data (including operating systems, applications, business data) replication can out of the third party backup software. Only between different disk backup processes, several times higher than traditional tape library backup speed. Backup of the virtual machine and the data can be virtualization platform to load and read at any time, can be convenient for data validation and data recovery.

(3) With the third party backup software integration backup and advantages. In the mode of the virtual machine backup, Virtual machine can not only by its own virtual machine snapshot, cloning function of data backup, but also could be accessed by integration of third-party backup software for data backup. For physical servers, the virtual machine is only on a physical disk a collection of files, through the set for backup files, it will be able to realize the batch virtual machine backup. This way of backup, for customers is good for existing backup environment for integration, the new virtual machine backup only need new backup strategy in the backup software data backup of the virtual machine can be realized.

Does not change the existing backup environment, only need to add the data backup, backup strategy can realize virtual machine to a backup to disk, magnetic tape, such as medium, realizes the data offline archive has asked to the backup window, backup process takes up a large amount of data network or the backup data storage network bandwidth was not immediately read and validation, only through reduction data disaster recovery means, it will take a long time to recover the data.

4.2 Data recovery strategy.

After using virtual server platform for deployment, server are created for the virtual machine, and by using the data backup is a virtualization platform snapshot, cloning, and other functions. All the data backup data is within the virtual machine (contain virtual machine images, configuration and application software, the production data, etc.) no matter which part of a virtual machine data appeared abnormal or missing, we can all be a snapshot, cloning, and other functions for data fast recovery.

When the abnormal data in virtual machine, can be achieved through a snapshot restore function within the virtual machine data rollback, the restoration of all data within a virtual machine to the snapshot of arbitrary point in time, to ensure data security. Snapshots restore features, is due to record a snapshot of the data pointer position, so in the process of recovery without copy to all the data recovery, only need to restore a pointer location. If the need for data, recovery can use extreme time to complete data recovery.

All data is lost when the virtual machine appears, the all data can be recovered by virtual machine. As a result of the virtual machine clone backup way is the virtual machine to copy and save all data, so no matter how much data loss in virtual machine appears, just need to reload the cloned virtual machine to the server virtualization platform, all data can restore the virtual machine.

4.3 Data validation strategy.

For backup data checking, the traditional way of backup can't very well solve the problem. Due to the traditional way backup using third-party backup software, the form of the backup data use special backup software and encryption algorithm, this leads to the backup data cannot be directly read, And

also unable to verify the accuracy and integrity of the backup data. For server virtualization platform, with the method of virtual machine deployment and backup application, system can solve this problem very well.

Because the virtual machine in the physical disk is only one set of different files, whether using snapshots, data backup in the form of cloning which is directed at the virtual machine will back up all data, the backup of data is a complete virtual machine image file.

For data validation virtualization platform, it cannot affect under the situation of existing production system. Users only need to back up the virtual machine starts, reconfigure the IP address, Can produce an exactly the same as that of production system configuration. Through this virtual machine, the user can inspect the production data validation. In addition, the backup virtual machine can not only provide data validation function, because the virtual machine data is consistent with the data and production systems, but also to the virtual machine data for statistical analysis, generate reports and data mining, etc. Even when there was a disaster can't run production system, the virtual machine can replace the production system server, continue to provide application services, so as to realize a simple disaster, cheap strategy.

References

- [1] Liu Shuyan, Xiaoli Lu 2013, Database of disaster backup and recovery technology analysis and research ,Electronic Technology & Software Engineering. 16(2013)223-224.
- [2] Li Suolei, Oracle database backup and recovery, Information Security and Technology. 9(2011)99-102.
- [3] Tu Guoyi 2013, a virtual machine backup and restore system based on ESXi, Computer security. 11(2013)2-8.
- [4] Xie Xinyun 2014, Cloud computing environment of small and medium-sized enterprise information demand research, Chinese & Foreign Entrepreneurs. 29(2014)25.
- [5] Ying Bing 2014, Data backup and recovery technology application in the portal, Network Security Technology & Application. 2(2014)25-27.
- [6] Yin Hongbo 2011, Enterprise information security policy research, Computer and Information Technology. 3(2011) p, 28-31.