International Journal of Control and Automation
Vol.8, No.11 (2015), pp.325-336
http://dx.doi.org/10.14257/ijca.2015.8.11.31

Research and Implementation of CDP Disaster Recovery System based on Oracle

Yongjian Ren[ab], Jiaolong Ye[ab], Jilin Zhang[abc], Li Zhou[ab], Jue Wang[cd] and Lei Zhang

[a]*School of Computer Science and Technology, Hangzhou Dianzi University, Hangzhou, 310018, China*
[b]*Key Laboratory of Complex Systems Modeling and Simulation , Ministry of Education, China* [c]*College of Electrical Engineering, Zhejiang University, Hangzhou, 310058, China*[d]*Supercomputing Center of Computer Network Information Center, Chinese Academy of Sciences, Beijing, China* [e]*Computer Science Department, Beijing University of Civil Engineering and Architecture Beijing, China*
jilin.zhang@hdu.edu.cn

Abstract

The OraCDP disaster recovery system uses CDP disaster recovery technology based on the combination of block level, combining with the communication coupling between the underlying I/O and Oracle database. The experimental results show that the OraCDP system can not only maintain the normal operation of the original system without any negative influence, but also guarantee the recovery of data and service at specific time quickly, fully and accurately.

Keywords: *Oracle, CDP, Disaster recovery system*

1. Introduction

1.1. Introduction of Disaster Recovery Technology

The disaster recovery system can be said to be the highest level of data storage and backup, consult to the common principle of disaster backup grading of international disaster backup industry, the disaster backup system can be divided into the following four levels from low to high, according to the different amount of data, the degree of difference between the different data and production data, as well as the complete extent of the disaster recovery environment.

The zeroth level, no backup center. In fact, this level has no disaster recovery capability, it only backup local data locally, rather than sending it to the remote;

The first level, local backup to tape, remote backup. Key data backed up locally and then sent remotely. System and data will be recovered according to the predetermined data recovery program soon after the disaster. This scheme has low cost, and easy configure, however, problems of storage medium management rises as the amount of data increases, what is more, a large amount of data may be difficult to recover in time when the disaster happened. So key data should be restored prior to the restoration of noncritical data;

The second level, hot backup the site backup, establishing a hot point in remote backup and then backup the data through network, which means the main site data backup to the backup site through the network in synchronous or asynchronous mode, usually, the backup site backup data only and it does not undertake the business. The backup site

ISSN: 2005-4297 IJCA

replaces the primary site of business so as to maintain the continuity of business operation when the disaster comes;

The third level, activities backup center. Usually, two data centers will be established in the long distance but they have mutual backup data while they are in working condition. As a result, when the disaster happens in a data center, another data center can replace its task.

As the product of the development of data storage and backup to a certain extent, disaster recovery system will ensure the security of user data and provide uninterrupted application service without affecting the operation of original system. Local cluster system, hot backup mechanism as well as multi-level network failover mechanism should be established when high availability disaster backup system is to be constructed.

That is to say, the application should be ensured to run without interruption no matter which node has any fault occurred when one or more applications are running across multiple server nodes; Once a fault occurs, the system will switch the application and system to other service node as soon as possible and keep the normal operation with the help of the fault diagnosis and the switching strategy of powerful formulation mechanism.

The enterprise needs to construct the specific disaster backup system according to their area, business type and scale of data and other differences so as to avoid unnecessary cost. Consider the following process in the formulation of disaster backup system solutions:

(1) No affecting the normal operation of the original business system;
(2) Maintaining synchronization of the data state;
(3) The disaster tolerance system maintenance work must be as convenient as possible;
(4) The system recovery time should be as short as possible;
(5) The business subsystems can be switched and cut back;
(6) Various combinations of technologies to choose.

1.2. Measure Index of Disaster Recovery Technology

Two technical indicators data protection method in the theory of measure in disaster recovery are RPO (Recovery Point Objective) and RTO (Recovery Time Objective) [4]. RPO data recovery point objectives, mainly refers to the loss that can be tolerated business system data. RTO is the recovery time objective, mainly refers to the longest service-stopping time can tolerate, which means the shortest time required to restore the service functions from the disaster short time.

RPO aims at data loss while the RTO aims at the loss of service and the two do not have inevitable relevance. Determination of RTO or RPO should be made after the risk analysis and business impact analysis according to the corresponding businesses. For the same kind of business of different enterprises, RTO and RPO requirements may be different. However, RTO has high request of cost requirements, and unfortunately, the return seems not proportional to its pay. It is impossible for the enterprise to put unrestricted funds to a disaster recovery system. For financial institutions like bank securities are very close to deal with every online transaction accident and even every penny counts, every data needs to be restored and in this matter, RPO is clearly more appropriate. However, the purpose of disaster recovery is to ensure business continuity, using either RTO or RPO, to move closer towards the core.

According to the different RPO, common block level data protection technology today can be divided into regular backup (Periodical Backup), (Snapshot) a snapshot and continuous data protection (Continuous Data Protection, CDP) three class. Continuous data protection (Continuous Data Protection, CDP) technology can recover data to any moment in history whenever the data is lost or damaged through the real-time track and record all data updating-records without affecting the normal data service. The realization mechanism of traditional continuous data protection technology of the block level is to save all the update of each data block with a single write requests for the unit according to the time sequence.

When the data recovery is needed, what we have to do is only to replace all the time stamps in the recovery time point before the data block can be. Compared with the previous two kinds of technology, continuous data protection technology in the data block level can recover from the failure to a single write request for any moment in history of the unit, having the minimum amount of data loss.

1.3. Current Research Status of CDP Technology

Definition of CDP by continuous data protection group (CDP SIG) of Storage Networking Industry Association (SNIA) is: "continuous data protection is a set of methods, which can capture and monitor data changes, and separately stored out of the production data to ensure that the data can be recovered to any point in time. Continuous data protection system can realized based on the block, file or application and provide fine enough granularity to restore object, achieving a virtually infinite number of recovery time"[8].

At present, researches on CDP technology at home and abroad scholars have made some achievements. TH-CDP [1] study on the concrete realization of the block level CDP system; SnapCDP[4] is the realization mechanism for embedding the CDP module into a mechanism implemented in the LVM; ST-CDP [3] is a block level CDP, it works to ensure a fast and reliable data recovery by periodically inserting Snapshot parity check record of the changed data blocks; CB-CDP[5] it is a distributed CDP system scalable file level.

Continuous exploration of disaster tolerance technology by the major scientific research institutions and large enterprise IT has promoted the development of CDP technology. The representative of the commercial products includes Symantec VERITAS Global Cluster Manager, VERIATS Cluster Server, IBM HAGEO, XRC, HP MC/Service Guard and EMC SRDF[9]. These representative products and solutions have bright and the respective advantages by satisfying the different demands of different industries and companies ,and almost all of them are based on iSCSI technology, FC technology and remote replication technology, although very powerful, function relatively complete, Special optical fiber links needed to deploy such a complete set of disaster recovery system may rises the fees of the whole system, the price is very high, Therefore, the implementation of the remote disaster tolerant system needs to have very large inputs, and the distance between the data production center and remote disaster recovery center is limited, for far distance leads to a sharp increase in cost while too short distance cannot achieve the purpose of remote disaster recovery, the general will be 10 kilometers as the deployment of the system standard. In addition, the disaster recovery products have been integrated in each company's hardware, so there are some defects in the flexibility, generality of software and hardware aspects. While the software solution has some limitations, such as Symantec SF/VVR replication mechanism, although it is based on the IO level, but the use of log volume for the relevant records have great limitations.

According to relevant experts study, there is still no mature CDP technology which can ensure the real-time data recovery and fast recovery service. Version history of consistency and integrity of data search accurate will take some time, not to mention the need to restore the process time overhead. Therefore, using of backup system of CDP technology in the construction of high availability to improve the reliability and security of enterprise data information storage has the certain practical significance.

1.4. The Structure of this Paper

The structure of this paper is as follows: the second part mainly introduces the overall architecture of OraCDP disaster recovery system, the deployments of the various functional modules, the working process and the theoretical analysis of the system; the third part introduces a test in view of system performance and data integrity and analyzes

the experimental results; the fourth part makes a summary of the this paper. All printed material, including text, illustrations, and charts, must be kept within the parameters of the 8 15/16-inch (53.75 picas) column length and 5 15/16-inch (36 picas) column width. Please do not write or print outside of the column parameters. Margins are 1 5/16 of an inch on the sides (8 picas), 7/8 of an inch on the top (5.5 picas), and 1 3/16 of an inch on the bottom (7 picas).

2. Design and Implementation of CDP Disaster Recovery System

2.1. Architecture and Function Module of Disaster Recovery System OraCDP

The basic principle of CDP is to copy each writing data and to add additional time marker, in addition, it requests the copy must be stored independently. According to the data replication mechanisms, the formations of several different architectures are host (Host-Based), network (Network-Based), storage terminal (Storage-Based). However, the OraCDP is designed to achieve the purpose of disaster recovery and reduce the limitations brought by the hardware by software solutions, so we choose the host architecture, namely in the host installation agents to be responsible for monitoring disk and copy transaction data work. Agent will capture every data written to the disk and restore a copy together with a time stamp into the buffer, and then transfer the data to the disaster recovery center through the SAN network.

The OraCDP system consists of 4 modules, including the CDP agent module, configuration management module, storage virtualization module, the implementation of CDP module. These modules were deployed in the host, control center, disaster recovery center, specific deploy as shown in table 1 according to different function points.

Specific OraCDP system architecture is shown in figure 1. Disaster Recovery Center for maps logical volume to the client host end to form a cache volume as a mirrored volume; client host raw volume data to a mirrored volume Synchronously and the control center configures the storage pool size and CDP implementation cycle and other information; disaster recovery center will restore system corresponding to the normal CDP point to the original customers, bare or virtual host machine quickly.

The CDP agent module id deployed in the host, mainly responsible for monitoring the disk I/O requests and marking timestamp logo in these requests, sending these I/O request to the original volume and saving the CDP volume, synchronous raw disk volumes data to a mirrored volume, and performing cache refresh work according to CDP cycle configuration.

Configuration management module is deployed in the control center, mainly responsible for storage pool configuration of the disaster recovery center, unified management of the client host, configuration of CDP implementation cycle, the specific choice of what kind of data transmission links to perform backup restoration work and so on.

Storage virtualization module is deployed in disaster recovery center and mainly responsible for transferring the physical storage equipment disaster recovery center into the virtual storage pool, grouping into a block of logical volume according to the customer the host end disk volume size so as to maximum the unified management and allocation and take greater use of storage space.

The implementation of CDP module is deployed in the disaster recovery center, mainly based on the CDP implementation cycle, combined with the timestamp acquired and wrote to CDP save volume in a snapshot technology ROW write redirection (Redirect-On-Write) mode.

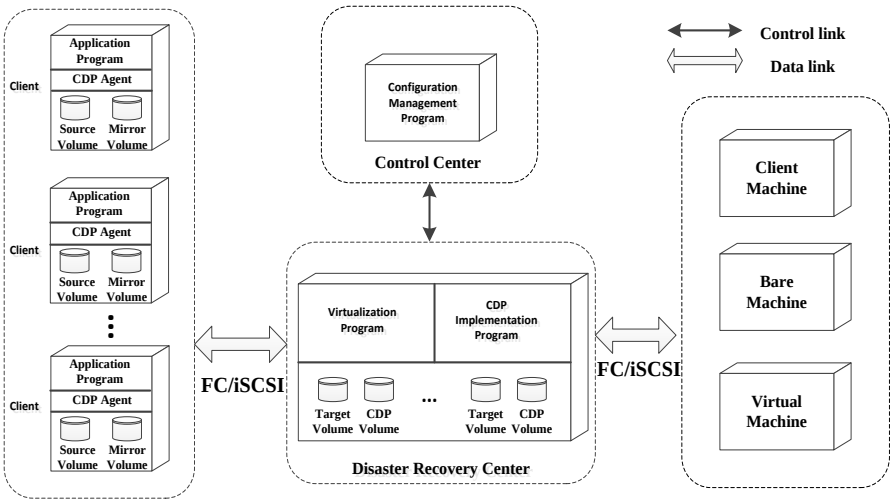

Figure 1. The Architecture of OraCDP System

Table 1. Function Module and the Deployment Location

Function module	Deployment location
CDP agent module	Host Side
configuration management module	Control Center
storage virtualization module	Disaster Recovery Center
implementation of CDP module	Disaster Recovery Center

2.1.1. The One-Way Mirror Synchronization Strategy: Synchronous mirroring technology is one of the important parts of CDP technology. It requires to backup system disk of the environment data to another disk, and make real-time backup, and never affect the original volume performance in the general case. This technology performs two main operations: the establishment of mirror image relationship, coping volume data to a mirrored volume synchronously.

As a consequence, we create a mirror synchronization strategy to achieve the above results which is shown in Figure 2.

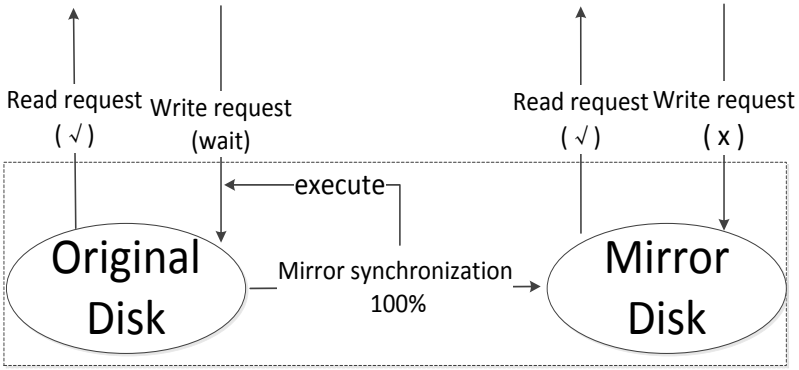

Figure 2. Schematic Diagram of Reading and Writing Processes

(1) Establishment of mirror image relationship. Once the relationship is established, protection of mirrored volume should be added, not supporting the client host write request and the mirrored volume opens only read permission.

(2) Synchronous raw volume data to a mirrored volume. Both sides need to sync data in the establishment mirror image relationship, but to ensure data consistency not to affect the performance and upper layer application of the mirror disk.

2.1.2. The Proxy Process Based On Oracle: Copy protection data is a core part of CDP disaster recovery technology, however, the integrity and consistency of database data is the priority among priorities. The traditional approach is to use the database backup and recovery mechanism itself and make disaster recovery operation alone, which obviously cannot meet the current needs of the majority of enterprises. The best results of enterprises expecting is to restore the system and application of database data, reduce or even avoid the loss soon after a disaster. Therefore, we choose the disk data backup database combining with the way of agency to meet the needs of enterprises with less cost.

CDP Agent module, which responsible for refreshing the application cache and sending the CDP SCSI command to a storage server through the SCSI subsystem, communicate and percept the Oracle database from the Oracle application, the storage server create the CDP snapshot point after receiving a CDP command. The working principle of the module is shown below:

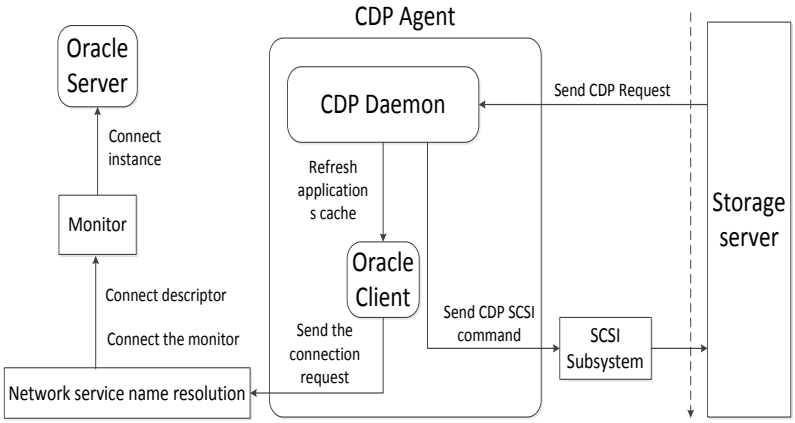

Figure 3. CDP Agent Principle

The whole process is described as follows:

(1) the CDP Agent service starts and reads the configuration information;

(2) the storage server sends a CDP request which contains the volume ID information needed to create the CDP snapshot point;

(3) the CDP Agent daemon receives a CDP request and finds the configuration information of the storage volume. If the configuration information of the storage volume is in the configuration file, the CDP daemon will call the Oracle Client API link Oracle Server according to relevant configuration parameters (network database service name, login name, password) and refresh the database cache after a successful connection, however, if the connection fails, then, the error information is written to the log file;

(4) Sending CDP SCSI command to the storage server through disk SCSI Pass interface after refreshing the database cache, and the storage server receives a CDP command to create the snapshot point after CDP.

In theory, this solution largely addressed the needs of business users can guarantee the integrity of the database data without affecting the normal operation after the recovery system.

2.2. The System Flow

The OraCDP system process mainly comprises a storage pool capacity allocation, allocation of mirror image relationship, relationship of initialization, data protection, data recovery. Flow chart of specific is shown in Figure 4.

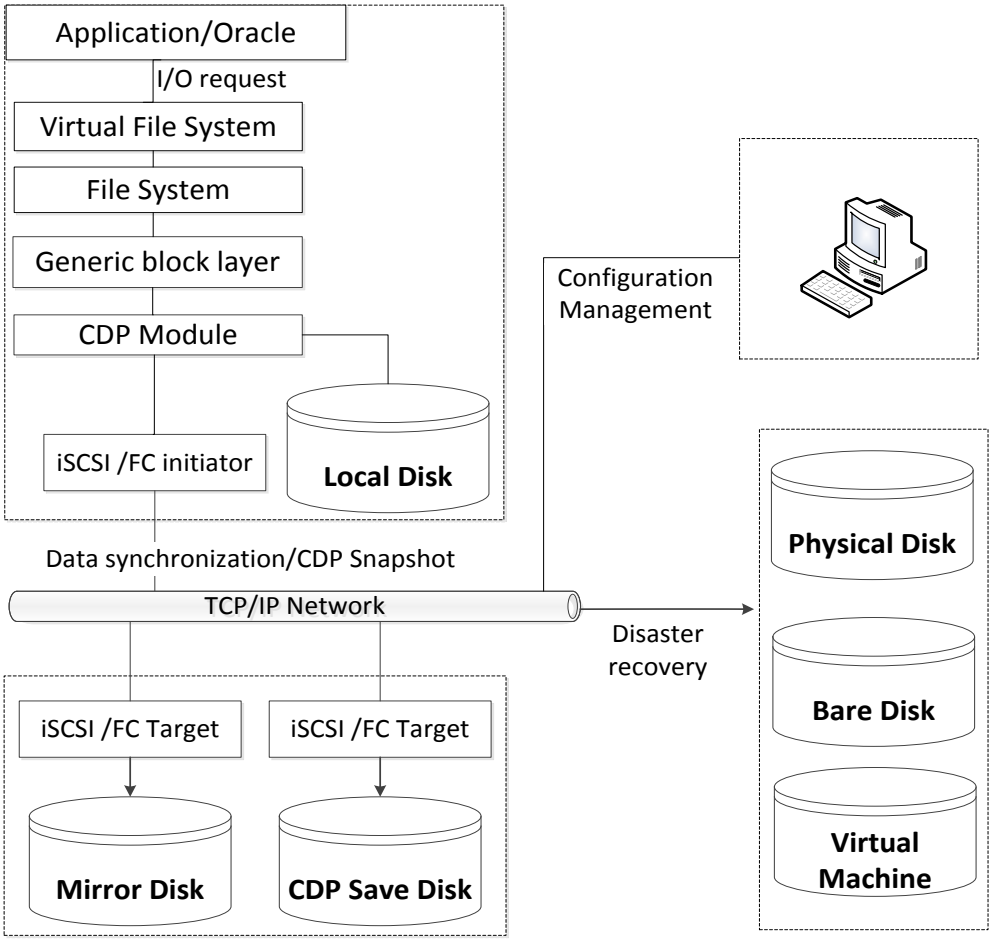

Figure 4. Flow Chart of OraCDP System

(1) The storage pool capacity configuration

The system manages the physical storage devices unitedly by the storage pool configuration of the control center and the virtual pool technology of the disaster recovery center, virtualizing it into logical storage pool, and according to different client host disk configuration, the storage pool is cut into a piece the size specified in the logical volume, cutting the storage pool into the specified-sized pieces in the logical volume.

Through the establishment of, technology and unified management of physical storage equipment disaster recovery center through the virtual pool, its virtual into logical storage pool according to different client host disk configuration.

(2) A mirroring relationship configuration.

Control center comes up with the need for protected disk, mapping the backup capacity to the client host by the disaster recovery center in the form of virtual volumes, forming a mirrored volume. And then cut off a piece of CDP from the storage pool to save the volume to save the disk CDP snapshot point. Finally, record the original volume with a mirrored volume mirroring relationship by the CDP module of the client host to ensure the mirrored volume is not written.

(3) The relationship between the initialization.

The control center initializes the relationship of the specified disk, CDP module of the client host synchronize the original volume data to the mirror volume. In the meantime, it makes synchronous operation prior to considering the original volume write I/O request.

(4) Data protection.

After completing the mirror synchronization, OraCDP system makes CDP snapshot according to the CDP cycle specified by policy control center. Combined with the time stamp information, the system will update every piece of data stored in CDP volume in the form of snapshot, completing data protection on the disk.

(5) Data recovery

Disaster recovery center can roll back the data or system to state of any point during the specified time slot, and quickly return to the original disc, disc or virtual machine by the data blocks stored on the CDP with a time stamp information.

2.3. Theoretical Analysis

The OraCDP system is realized mainly by combining with communication coupling of Oracle database application mainly by the underlying I/O. In theory, the Oracle database can be accurately recovered to the disaster happened before through the data block with the time stamp information stored on disk in CDP without affecting t the normal operation of the original system.

3. Analysis of System Testing and Experiment

In order to test the effect of OraCDP system to the original system on read and write performance and the data integrity verification after disaster recovery, this article will carried on the test and analysis separately.

3.1. Experimental Environment

The OraCDP system deploys the experiment environment and simulates the real environment disaster. The experimental environment includes three client hosts: a disaster recovery center and a control center, among which the client and disaster recovery center are deployed on the disk array, while the control center is deployed in individual PC. The physical node hardware and software configuration are shown in table 2.

Table 2. The Physical Node Configuration Table

Role	CPU	Memory	Disk Size	Operating System	CDP Installed?
Control Center	Intel Core 2.98GHz CPU*2	2G	500G	Windows 7 Sp1	No
Client 1	Intel Core 2.98GHz CPU*2	2G	100G	Redhat Enterprise 6.0	No

Client 2	Intel Core 2.98GHz CPU*2	2G	100G	Redhat Enterprise 6.0	Yes
Client 3	Intel Core 2.98GHz CPU*2	2G	100G	Redhat Enterprise 6.0	Yes
Disaster Recovery Center	Intel Core 2.98GHz CPU*4	4G	1T	Redhat Enterprise 6.0	No

3.2. Effect of the performance by OraCDP

In order to test the effect of the read and write performance to the original system by the OraCDP system, this paper chooses the client 1 and client 2 as a comparison reference. The testing tool, iometer, is installed in the host and disk of client 1 client 2 hosts in the same size volume, test those disks read and write performance of system with or without the CDP. The results are shown in figure 5.

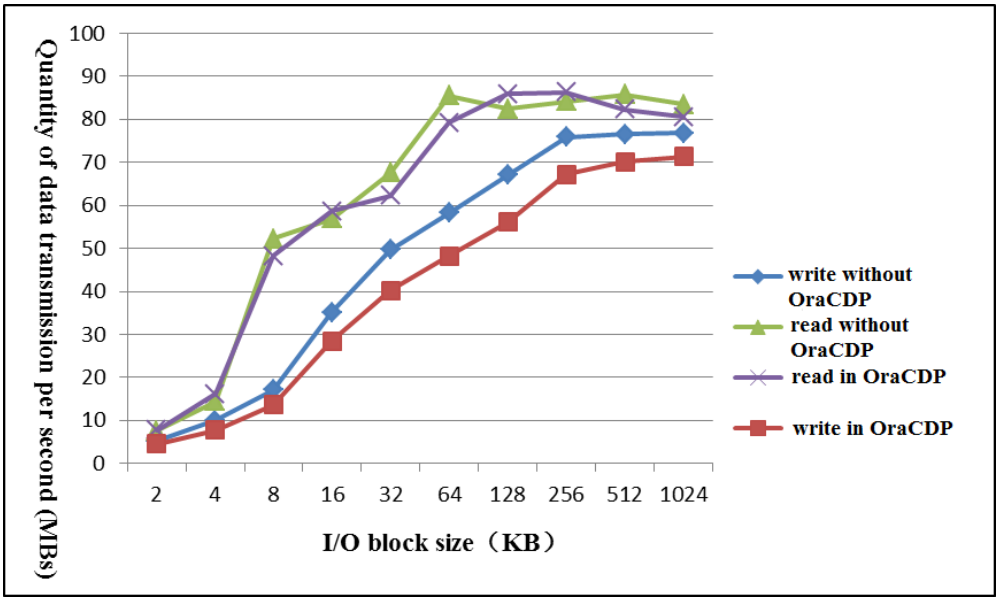

Figure 5. Comparison Chart Performance of Reading and Writing

The Figure 5 shows, read performance is better than write performance of the original volume of the client host. And there is almost no difference whether the OraCDP is installed. The original volume write performance has little effect, but the proportion is not high. Therefore, to ensure the normal operation of the original can roll system.

3.3. Data Integrity Testing

Verifying the feasibility of Oracle oriented agent by testing recovery of the OraCDP system after disaster. Insert a table in the data table space of client 2 and client 3 with OraCDP system and the Oracle database and then implement the CDP snapshot and recover the original volumes with mirror image relationship to two physical machines with the same configuration, after this, restart it. Test whether the two physical machines

can start Oracle databases and verify whether the data tablespace table is complete and correct. Results are shown in table 3.

Table 3. Oracle Database Data Validation

Client	Normal startup before Recovery	Normal startup after Recovery	Data tablespace data integrity after Recovery
client 2	Yes	Yes	100%
client 3	Yes	Yes	100%

We can see from table 3, two client hosts can startup normally after the recovery of the OraCDP system, data tablespace data integrity also can reached 100% after recovery, verifying the feasibility for Oracle agent.

3.4. Result Analysis

In the test results above, OraCDP system can guarantee the normal operation of the system, although leaving a small effect to the original write performance, this overhead is within what the users can bear; The normal operating Oracle database, together with the corresponding complete data table space after recovery verified the feasibility of database for the Oracle agent. Thus, it can be seen, the design and implementation of CDP disaster recovery system for Oracle database introduced in this paper is feasible.

Acknowledgements

This paper is supported by Natural Science Fund of China under grant No.61202094 Zhejiang Provincial Natural Science Foundation under grant no. LY13F02004, No.LY16F020018, China Postdoctoral Science Foundation Funded Project under grant No. 2013M541780. Key Projects in the National Science & Technology Pillar Program (No. 2014BAK14B00).

References

[1] Y. Sheng and D. Wang, "TH-CDP: An Efficient Block Level Continuous Data Protection System[C]", IEEE International Conference on Networking, Architecture, and Storage, 2009.
[2] Q. Yang, W. Xiao and J. Ren, "TRAP-Array: A Disk ArrayArchitecture Providing Timely Recovery to Any Point-in-Time[C]", In Proceedings of ISCA'06: The 33rd Annual International Symposium on Computer Architecture, pp. 289-301, 2006.
[3] J. Yang and Q. Cao, "ST-CDP: Snapshots in TRAP for Continuous Data Protection[J]", IEEE Transactions on Computers,2011.
[4] F. Wang, "SnapCDP: Design for Asynchronous and Real-Time Remote Replication System[J]", Journal of Computer Research and Development. 46(Suppl.): 114-121, 2009.
[5] C. Wang, A. Chai, Z. Wang and D. Ju, "CB-CDP: A Cloud Based Continuous Data Protection System[C]", Consumer Electronics, Communications and Networks (CECNet), 2013 3rd International Conference on, pp. 188-191, 2013.
[6] J. Liu, T. Yang, Z. Li and K. Zhou, "TSPSCDP: A Time-Stamp Continuous Data Protection Approach Based on Pipeline Strategy[C]", 2008 Japan-China Joint Workshop on Frontier of Computer Science and Technology, 2008.
[7] SNIACDPSIG. http://www.snia.org/forums/dmf/programs/data_protect_init/cdp/, 2008.
[8] X. Li, C. Xie and Q. Yang, "Optimal Implementation of Continuous Data Protection (CDP) in Linux Kernel", Proc. the Int'l Conf. Networking, Architecture, and Storage (NAS '08),pp. 28-35, 2008.
[9] Y. Zhu, "Data backup and disaster recovery [M]", Beijing：China Machine Press, 2004：8-10.
[10] W. Xiao, J. Ren and Q. K. Yang, "A Case for Continuous DataProtection at Block Level in Disk Array Storages", IEEE Transactions onParallel and Distributed Systems, vol. 99, 2008.
[11] Tntel, "IoMeter: Performance Analysis Tool", Iometer, http://www.iometer.org/, 2011.

[12] D. Yibing, C. Shanguang, H. Wei, G. Feng and L. Chuanyi, "A novel block-level continuous data protection system[C]", International Conference on New Trends in Information Science and Service Science (NISS 2010),2010.
[13] M. Pokharel, S. Lee, J. S. Park, "Disaster Recovery for System Architecture using Cloud Computing[C]", International Symposium on Applications and the Internet. 2010.
[14] M. Lu, D. Simha and T. Chiueh, "Scalable Index Update for Block-Level Continuous Data Protection[C]", International Conference on Networking, Architecture, and Storage (NAS 2011). 2011.
[15] X. Zhang, K. Liang and X. Zhang, "Research on the Recovery Strategy of IncrementalData-Based Continuous Data Protection[C]", International Conference on Computer Science and Electronics Engineering, 2012.
[16] B. I. Ismail, M.N.M. Mydin and M.F. Khalid, "Architecture of Scalable Backup Service For Private Cloud[C]", IEEE Conference on Open Systems (ICOS). 2013.

Authors

Yongjian Ren, he received the PhD degree in Computing Application Technology from Florida Atlantic University, in 1998. He is currently a professor in software engineering in Hangzhou Dianzi University, China. His research interests include Cloud Computing and Mass Storage.

Jiaolong Ye, he is now M.S. in School of Computer Science and Technology in Hangzhou Dianzi University, China. His research interests include High Performance Computing and Cloud Computing.

Jilin Zhang, he received the PhD degree in Computer Application Technology from University of Science Technology Beijing, Beijing, China, in 2009. He is a PostDoc at Zhejiang University. He serves as an assistant professor of software engineering in Hangzhou Dianzi University, China. His research interests include High Performance Computing and Cloud Computing.

Li Zhou, she serves as an assistant professor of software engineering in Hangzhou Dianzi University, China. His research interests include High Performance Computing and Cloud Storage.

Jue Wang, he is currently working as associate research fellow in the supercomputing center of Chinese Academy of Science. The motivation behind his work is to improve soft systems by increasing the productivity of programmers and by increasing software performance on modern architectures including many cores clusters and GPU.

Lei Zhang, she received the Ph. D. degree in School of Information and Communication Engineering from Beijing University of Posts and Telecommunications, in 2009.She is currently lecturer in the Computer Science Department, Beijing University of Civil Engineering and Architecture. Her areas of interest are in wireless sensor networks, data security, routing and multicasting protocols, and data mining. She is also a Member of China Computer Federation (CCF).