Oracle RAC performance Comparison on Physical Servers and VMware VSAN Servers

Ying Wei1 2*
1Department of Computer and Science Xiamen University Tan Kah Kee College, ZhangZhou, Fujian, 363105, China
2Network Information Center Xiamen University of Technology, Xiamen, Fujian, 361024, China
*Corresponding author’s e-mail: sailwy07@xujc.com.cn

Abstract— The physical performance of the Oracle RAC hardware architecture directly affects Oracle's output performance. In this paper, we perform performance tests on three Oracle RAC hardware architectures, and the test data is simulated using real teaching management system data on campus. The simulation results show that the two new oracle RAC hardware architectures are respectively occupied by the TPM and IOPS indicators, and the overall is stronger than the traditional external storage architecture.

1. Introduction
Nowadays, the information application in the campus network is more and more common, and the application load is getting heavier and heavier, and the performance requirements of the back-end database are also increasing. In this article we want to find a suitable ORACLE RAC [1] hardware architecture to make ORACLE's performance in digital campus applications to the best, while the investment is relatively reasonable. There are currently three ORACLE RAC architectures that can be used in campus information environments in the Network Information Center at Xiamen University of Technology, China. They are External fiber storage for Oracle RAC, Internal PCIE [2] storage for Oracle RAC, and VMWare Virtual SAN [3, 4] for Oracle RAC. The Oracle RAC system is a scalable performance cluster based on general purpose or dedicated hardware. It runs on a pure private network and generally uses Linux as the operating system to improve performance and stability. It can share file read and write with ASM or VxFS [5] file system. There are usually two nodes, or you can expand to four nodes or eight nodes. Designers can scale up performance with added node machines based on actual load conditions. However, for our school size and actual situation, the two-node RAC ORACLE database is the best cost-effective configuration. In the classic configuration, two physical servers and a shared fiber storage architecture are generally used. This is the hardware architecture that our school has been using, but there are two new solutions to choose from. One is still using two the physical server, but the storage is provided by the server itself, and the data files are shared by VERITAS storage foundation software. The other is to build an ORACLE RAC environment directly on VMWare VSAN using a virtualization solution. Finally we focus on testing the performance of the all the three Oracle RAC system using the most representative teaching management information system database, “teach”, in the campus core database.
We took out a representative application in a digital campus as an experimental subject. We finally chose teaching management information system. The teaching management service platform serves various objects of teaching management in colleges and universities, including teaching resources, teaching plans, student selection, short-term exchange, experimental practice teaching, course planning, online course selection, re-examination, re-examination, examination, and grade entry. It is based on the B/S architecture of the campus network to provide a comprehensive, whole-process information-based means for the teaching management of colleges and universities, providing information management for all teachers and students. Therefore, the comprehensive comparison of the teaching management system is one of the most important systems in the digital application of the university. His ease of use and performance directly represent the results of campus information construction [6]. At the same time, the teaching management system is also the system with the highest usage rate in colleges and universities, with a large number of user actual data and large-scale user access. Using the teaching management system to test can fully reflect the performance of the system in the actual production system.

In our university (XMUT), the teaching management information system can reach nearly 2,000 visits per day. In special periods such as semester started and closed reach about 100,000 visits per day. The finally performance of the teaching management information system directly affects the user experience. There are two ways to improve the performance of the educational system. One is the concurrent performance of the web. The other is to process power of the database.

In this study, we used External fiber storage for Oracle RAC as a Benchmark hardware platform to investigate the performance of core database of teaching management information system. We also introduced two other ORACLE RAC hardware architectures for performance comparison. There are Internal PCIE storage for Oracle RAC and VMWare Virtual SAN for Oracle RAC.

2. Oracle on Physical servers
The tradition way to build oracle RAC system is to use at least two physical servers as a cluster. In our school, In order to ensure data consistency, the core of Oracle RAC is how to establish a unified data storage problem. At the same time the performance of the data storage will also ultimately determine the performance of the entire oracle database. We use two physical servers to build a two-node Oracle RAC environment. For building shared storage, we have two options, one is to use external fiber storage and the other is to use server internal storage.

![Figure 1. External fiber storage for Oracle RAC](image-url)

In the figure 1, we use the external fiber storage architecture to build the oracle RAC environment. We use two DELL R910 rack servers as two RAC nodes. Each server have two Intel® Xeon™ Processor E7-8837 with 8-core CPU 2.67GHz and 128GB memory. Also, each server have Two 900G
15K local hard drives, six 1000M Network card and two HBA card. The extra storage system we use HP EVA6100 which has 70 450G SAS hard drives. We connect the two DELL R910 rack servers and HP EVA6100 storage by two Brocade fiber switch. In the software system part, the operating system we use RHEL6.5. The database is Oracle 11g 11.2.0.4 x64 for linux [7], Stress test tool is Swingbench 2.5 [8], and The IO stress test tool is Orion 11.1 [10, 11].

![Internal PCIE storage for Oracle RAC](image)

Figure 2. Internal PCIE storage for Oracle RAC

In the figure 2, we abandoned the external fiber storage when building the system and switched to the built-in PCIE SSD internal hard drive. Our server configuration is the same as before, both are two DELL R910 rack servers. Each server have two Intel® Xeon™ Processor E7-8837 with 8-core CPU 2.67GHz and 128GB memory. Also, each server have Two 900G 15K local hard drives, six 1000M Network card and two HBA card. In addition, we added two 800GPCE SSD cards and two IB 56GHCA cards. We connect the internal storage from two DELL R910 rack servers by two Mellanox SX6012 56 GB infiniband switch. In the software system part, the operating system we use RHEL6.5. The database is Oracle 11g 11.2.0.4 x64 for Linux [7], Stress test tool is Swingbench 2.5, The IO stress test tool is Orion 11.1. [8] And Storage Foundation 6.2 is the key role of internal shared storage. By using the storage foundation, we made two PCIE SSD inside the server into the raid0 disk group. Then we can use the PCIE SSD raid0 disk group of another server to make raid1. It is provided to oracle RAC for database.

3. Oracle on vmware VIRTUAL SAN

We also build oracle RAC system on VMWare virtual SAN. VMware Virtual SAN [3, 4] is a new architecture storage system, a software-driven architecture that provides tightly integrated compute, network, and shared storage capabilities from several virtualized x86 servers. This is different from dedicated storage, and you can build a high-performance storage with just a normal server. Together with VMware VSphere, these servers can also be used to provide computing resources. Take advantage of server performance and save on investment.

A Virtual SAN for oracle RAC system was built with four DELL R740xd Rack Server. There are two Intel® Xeon™ Processor Gold 5118 with 10-core CPU 13.75M Cache 2.4GHz and 256GB 2666 DDR4 memory for one serer. Also, each server have five Samsung SAS SSD 1.6T. A raid card that supports the pass-through function, 4*10 Gigabit Ethernet cards, 4*1 Gigabit Ethernet cards. Four nodes are connected via two 10 Gigabit Ethernet switches: Big Switch 4048. One 10 Gigabit Ethernet for network service and the other is for Virtual SAN. Figure3 shows VMWare Virtual SAN for oracle RAC.

In our Virtual SAN oracle RAC system [9], The DELL R740xd server cluster will be unified for computers and storage nodes. The entire system takes full advantage of the high-speed computing power of the Intel Gold5118 CPU and the high-speed read and write capabilities of SSDs. Due to the characteristics of VSAN, all calculations and storage are provided through the network. In order to ensure that the network bandwidth is not preempted, we use four 10 Gigabit NICs in the DELL R740xd, and they are divided into two groups, and two groups in each group are bound. The external interconnect switch uses the high-performance 10 Gigabit switch from DELL BigSwitch. In addition,
we built an all-flash storage disk group in the DELL R740xd to increase the speed of disk read and write operations.

The Virtualization software are VMWare Vcenter Server and ESXi 6.0U3. The operation system is Oracle Enterprise Linux 6.8(OEL 6.8), The database is Oracle 11g 11.2.0.4 x64 for Linux[7], Stress test tool is Swingbench 2.5[8], The IO stress test tool is Orion 11.1[10, 11].

![Figure 3. VMWare Virtual SAN for oracle RAC](image)

4. Results and Discussion

4.1 Standard database establishment

In our experiment, Oracle 11gR2 was adopt for test. The core of Oracle RAC is the shared disk subsystem. All nodes in the cluster must have access to all data, redo log files, control files, and parameter files. The data disks must be globally available, allowing all nodes to access the database, and each node has it. Own redo logs and control files, but other nodes must be able to access them to recover from a system failure in that node. Running on a cluster, Oracle RAC provides the highest levels of availability, scalability, and low-cost computing power for Oracle Database. If a node in the cluster fails, Oracle will continue to run on the remaining nodes. At last, related parameters are established as showed for the three type of Oracle RAC architecture in table 1-3:

| Disk Type | The size of VMDK (GB) | Total storage (GB) | ASM/VXFS |
|-----------|-----------------------|--------------------|----------|
| Operating system and Oracle directory | 1600 x 1 | 1600 | not suitable |
| Database data disk | 500 x3 | 1500 | DATA |
| FRA | 300 x 3 | 900 | FRA |
| CRS、Voting disk | 10 x 3 | 30 | CRS |

| Test parameters | Data value | Remarks |
|-----------------|------------|---------|
| Test data volume | 300G | (Import the size of the Oracle database) |

Table 1. the description of database space

Table 2. the requirements of test configuration
User concurrent number | Up to 1200 users (up to 600 users per RAC node)
---|---
Number of simulated hosts | 6 (Independent host distribution of 600 users)

Table 3. The type and proportion of stress test tool transactions of Swingbench

| Transactions | Class Name | Short Name | Load Ratio | Active? |
|--------------|------------|------------|------------|---------|
| Customer Registration | cora_bernberg_swingscan_recommendation_11_gs | CRL | 0.1 | 0 |
| Update Customer Details | cora_bernberg_swingscan_recommendation_11_gs | UCDD | 0.0 | 0 |
| Browse Products | cora_bernberg_swingscan_recommendation_11_gs | BP | 0.5 | 0 |
| Order Products | cora_bernberg_swingscan_recommendation_11_gs | OP | 0.4 | 0 |
| Process Orders | cora_bernberg_swingscan_recommendation_11_gs | PO | 0.5 | 0 |
| Browse Orders | cora_bernberg_swingscan_recommendation_11_gs | BO | 0.5 | 0 |

4.2 TPM performance test results

We conducted TPM (Online transaction volume per minute) testing for the Oracle 11g database platform running on External fiber storage for Oracle RAC, Internal PCIE storage for Oracle RAC and VMWare Virtual SAN for Oracle RAC. We are referred to below as Scene1, Scene2, and Scene3. Our purpose is to discuss the overall performance of the database in the actual environment by testing the TPM. First, run the Swingbench [9] tool and import the corresponding test data; Second, stress testing database nodes rac1 and rac2 in different Oracle RAC architecture; third, the simulation client initiates a database request from six computers, each computer has 100 users, and observes the final results. After 10 minutes of testing, the Scene1 result of the TPM value is approximately 217626, the Scene2 is 300123, and the Scene3 is 180664. Then we change the parameter, the test simulation client initiates a database request from 12 computers, each computer has 100 users, and observes the test results. After 10 minutes of testing, the Scene1 result of the TPM value is approximately 207890, the Scene2 is 524260, and the Scene3 is 152925. Also we change the parameter again. The test simulation client initiates a database request from 18 computers, each computer has 100 users, and observes the test results. After 10 minutes of testing, the Scene1 result of the TPM value is approximately 268737, the Scene2 is 431124, and the Scene3 is 100467. The figure4 shows the Oracle database performance comparison of TPM per minute under different user numbers.

Table 4. Three hardware architecture performance comparison in TPM

| TPM | Scene1 (tpm/m) | Scene2 (tpm/m) | Scene3 (tpm/m) |
|-----|----------------|----------------|----------------|
| 600users | 217626 | 300123 | 180664 |
| 1200users | 207890 | 524260 | 152925 |
| 1800users | 268737 | 431124 | 100467 |
4.3 Query Table Database Performance Test Results

In this study, we tested the performance of query table database performance by three kinds of Oracle 11g RAC system. We run the PLSQL [12, 13] tool, query the corresponding table record query time and record the result. We mainly used three SQL statements to test the performance of the three databases.

The first SQL statement is as follows “select ‘12-16’ week nums, ‘week’||’four’ week day, ‘morning’ session,” name, jsmc classroom, zws seats, jslb class, lh nums, xqdm campus, bz remarks from jxcdxxb where jsbh from jxcdview a, jxcdxxb b where xn='2014-2015' and xq='2' and a.qsz<=16 and a.jsz>=12 and a.xqj='4' and ((a.sjd + a.skcd – 1)>=1 and b.xqdm = ‘1’) order by jsmc, zws”; And the second SQL statements is “select * from testable”; (1 million records, 16 fields). The result are in table 5:

| SQL1   | Scene1 | Scene2 | Scene3 |
|--------|--------|--------|--------|
| 10     | 2.7    | 1      |
| SQL2   | 597    | 477    | 384    |

4.4 Host system performance test results

We tested the hardware performance and storage IO performance of the database nodes (RAC nodes). The purpose is to understand the actual performance of the current host hardware through the performance test of the host, and understand the storage read and write performance of the IO test. The experimental method is to run the Orion tool and test the data according to different IO sizes. The main test items are IOPS, throughput and latency. The test results of the three scenarios are shown in Table 6.

| Orion IO test | IOPS |
|---------------|------|
| Type          | Scene1 | Scene2 | Scene3 |
| 8k block random read | 13260 | 96530 | 9553 |
| 8k block random write | 382 | 9906 | 5087 |
| 32k block random read | 5415 | 20728 | 5196 |
| 32k block random write | 286 | 8276 | 4608 |
| 64k block random read | 3175 | 15047 | 3384 |
| Block Size | Random Read | Random Write |
|------------|-------------|--------------|
| 64k        | 273         | 6319         | 2611         |
| 1M         | 360         | 2312         | 464          |
| 8k         | 120         | 995          | 201          |
| 32k        | 870         | 15235        | 6318         |
| 64k        | 671         | 12067        | 4873         |
| 1M         | 679         | 9364         | 3669         |

5. Conclusion

We compared Oracle RAC performance under the same test data in three application scenarios. The three application scenarios are External fiber storage for Oracle RAC, Internal PCIE storage for Oracle RAC, and VMWare Virtual SAN for oracle RAC. The main technical indicators we tested were the oracle performance indicator TPM and the shared disk read and write performance IOPS. From the test results, we can see that Oracle RAC is the best performer in the VSAN environment in the TPM indicator. The virtualized environment integrates computing resources and storage resources into one whole. The data exchange of ORACLE RAC is completely carried out internally, achieving the optimal query speed. In terms of disk read and write performance, Oracle RAC is the best performer in the Internal PCIE storage environment. Because the transmission speed of the PCIE interface hard disk is much higher than the transmission speed of the SAS or SATA interface. The solid state hard disk with PCIE interface is used as the shared storage of ORACLE database, which can improve the performance of ORACLE database and solve the bottleneck of storage when the oracle RAC environment has large concurrency.

ACKNOWLEDGE

Fund Project: Research on University Data Center Application Based on SDN/Open Flow (JAT170435), Science and Technology Project of Fujian Provincial Department of Education. This research was also support by CERNET Innovation Project (NGII20170620, NGII20180510).

References

[1] Hart M, Jesse S. Oracle database 10g high availability with RAC, Flashback & Data Guard[M]. McGraw-Hill, Inc., 2004.
[2] PCI Express Base Specification Revision 3.1. PCI-SIG. http://pcisig.com/specifications/pciexpress. 2014
[3] Hogan C, Epping D. Essential Virtual SAN (VSAN): Administrator's Guide to VMware Virtual SAN[M]. VMware Press, 2016.
[4] Rajendran C. Getting Started with VMware Virtual SAN[M]. Packt Publishing Ltd, 2015.
[5] https://cn.norton.com/online-threats/glossary/v/vxfs-veritas-file-system.html
[6] Kingsosf page: http://www.kingsosft.com/page/Products.html
[7] Loney K. Oracle database 11g the complete reference[M]. McGraw-Hill, Inc., 2008.
[8] Giles D. SwingBench 2.2 Reference and User Guide[J].
[9] Bose S, Mishra P, Sethuraman P, et al. Benchmarking database performance in a virtual environment[C]//Technology Conference on Performance Evaluation and Benchmarking. Springer, Berlin, Heidelberg, 2009: 167-182.
[10] Koclanes M, Reed C, Feilinger M, et al. Policy based management of storage resources: U.S. Patent Application 10/447,677[P]. 2004-12-2.
[11] Wang H S, Zhu X, Peh L S, et al. Orion: a power-performance simulator for interconnection networks[C]//Proceedings of the 35th annual ACM/IEEE international symposium on Microarchitecture. IEEE Computer Society Press, 2002: 294-305.
[12] Feuerstein S, Pribyl B. Oracle pl/sql Programming[M]. "O'Reilly Media, Inc.", 2005.
[13] Gupta S K. Oracle Advanced PL/SQL Developer Professional Guide[M]. Packt Publishing Ltd, 2012.