Performance analysis of Internet Protocol Storage Area Network (IP SAN) and its usage in Clustered Database

**Akanksha Verma[1], Shrijee Bhattacharjee[2], Prashant Baste[3], Nandan Mirajkar[4]
Sandeep Bhujbal[5], Nilesh Deshmukh[6], Kumud Singh[7]**

**[1,6]Department of Advanced Networking and Telecommunications
IGNOU – I²IT Centre of Excellence for Advanced Education and Research
Pune, Maharashtra 411 057, India**

**[2,3,4,5]Department of Advanced Software and Computing Technologies
IGNOU – I²IT Centre of Excellence for Advanced Education and Research
Pune, Maharashtra 411 057, India**

**[7]Systems Department
IGNOU – I²IT Centre of Excellence for Advanced Education and Research
Pune, Maharashtra 411 057, India**

Abstract

In current social networking world, organizations are demanding high-speed security for client data such as images, audio video files even huge databases. Also no flexibility to connect and share remote devices and servers reduces the performance of service given by organization. To overcome this issues Storage Area Network (SAN) was introduced which is a dedicated storage network that carries I/O traffic only between servers and storage devices. The paper describes requirement of SAN in real world by implementing a SAN using Openfiler, Based on seven cases SAN's performance analysis done with help of Iometer tool and SAN's importance in clustered Database.

Keywords: *Storage Area Network (SAN), Iometer, Average I/O Response time, Database, Real Application Cluster (RAC).*

1.0 Introduction

The SAN improves the concept of data sharing. Although a typical LAN enables applications and end users to access data held in a central location, the SAN moves that data onto a much faster infrastructure. This allows multiple computers to transfer large files concurrently at rates comparable to locally attached disks over the SAN without adversely affecting the corporate LAN [17]. Openfiler is a opensource Linux operating system which can be used to design a SAN. In case of performance analysis five disks with video data are accessed by two machines and analysis is done for various cases on readings taken by Iometer tool. Oracle RAC has benefits including fault tolerance, security, load balancing, and scalability.

Unfortunately, for many shops, the price of the hardware required for a typical production RAC configuration makes this goal impossible. A small two-node cluster can cost from US$10,000 to well over US$20,000. This cost would not even include the heart of a production RAC environment, the shared storage. In most cases, this would be a Storage Area Network (SAN), which generally start at US$10,000.This paper provides a low-cost alternative to configuring an Oracle RAC 11g Release 2 system. All shared disk storage for Oracle RAC will be based on iSCSI (Internet Small Computer System Interface, an Internet Protocol (IP)-based storage networking standard for linking data storage facilities) using Openfiler running on a third node [9]. Section 2 illustrates about SAN its definition, architecture and its benefits while section 3 puts forward the results obtained for performance analysis. Section 4 describes the importance of SAN in clustered database.

2.0 Storage Area Network (SAN)

2.1 Definition

A **Network** is a collection of computers and devices which are interconnected by communication channels. These channels allow sharing of services, resources and information among it efficiently. **Storage Area Network** (SAN) is a network whose primary purpose is to transfer data between computer systems and storage elements-as defined by Storage Networking

Industry Association (SNIA) [1]. A storage area network (SAN) is a dedicated high performance network to facilitate block-level data access. It carries data between servers (hosts) and storage devices through switches [2][5].

2.2 Architecture

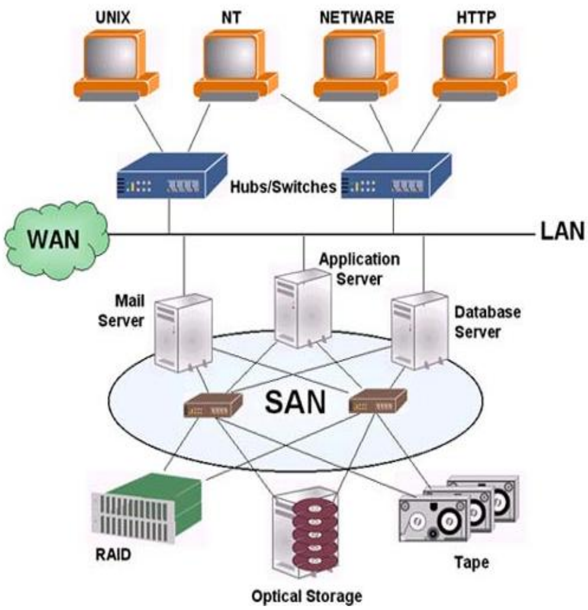


Fig 1: Architecture of Storage Area Network (SAN) [3]

The SANs are used to connect shared storage arrays and tape libraries to multiple servers, and are used by clustered servers for failover. A SAN allows direct, high-speed data transfers between servers and storage devices, potentially in any of the following three ways:
1. Server to storage: This is the traditional model of interaction with storage devices. The advantage is that the same storage device might be accessed serially or concurrently by multiple servers.
2. Server to server: A SAN might be used for high-speed, high-volume communications between servers.
3. Storage to storage: This outboard data movement capability enables data to be moved without server intervention, therefore freeing up server processor cycles for other activities like application processing [1].

The information stored in SAN can be accessed by all servers via Local Area Network (LAN) and Wide Area Network (WAN) so it becomes easy for information accessing. IP SAN uses TCP/IP as its media. The Transmission Control Protocol (TCP) and the Internet Protocol (IP) is part of the backbone of the Internet's suite of communication protocols. The advantage of IP SAN is that when it is utilized, networked storage can be available any place TCP/IP goes. Internet SCSI (iSCSI) uses the SCSI command set to communicate between the computing devices and storage, via a TCP/IP network. IP SAN uses TCP as a transport mechanism for storage over Ethernet, and iSCSI encapsulates SCSI commands into TCP packets, thus enabling the transport of I/O block data over IP networks [4].

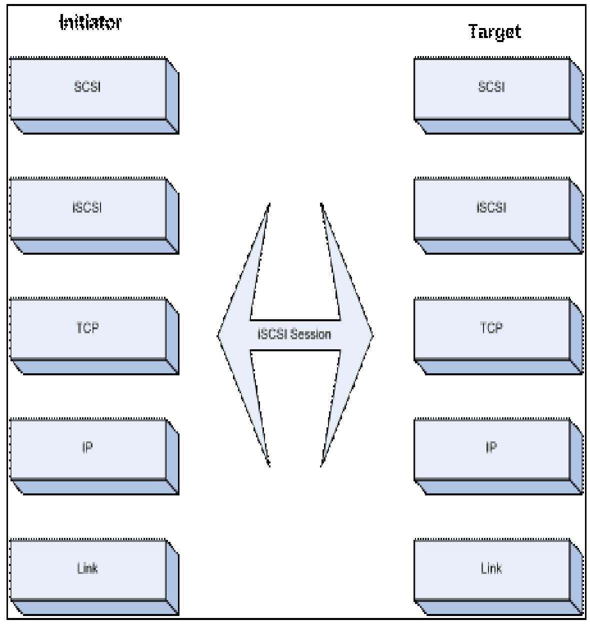


Fig 2: iSCSI and IP SAN Architecture [4]

2.3 Following are benefits of SAN:

1. Removes the distance limits of SCSI-connected disks.
2. Greater performance.
3. Increased disk utilization.
4. Higher availability to storage by use of multiple access paths.
5. Reduced data center rack/floor space.
6. New disaster recovery capabilities.
7. Online recovery.
8. Better staff utilization [18].

3.0 Performance analysis of IP SAN

Performance analysis of IP SAN can be done with the help of **Iometer**. Iometer is both a workload generator (it performs I/O operations in order to stress the system) and a measurement tool (it examines and records the performance of its I/O operations and their impact on the system). It can be configured to emulate the disk or network I/O load of any program or benchmark, or can be used to generate entirely synthetic I/O loads. It can generate and measure loads on single or multiple (networked) systems [16].

Iometer can be used for measurement and characterization of:

1. Performance of disk and network controllers.
2. Bandwidth and latency capabilities of buses.
3. Network throughput to attached drives.
4. Shared bus performance.
5. System-level hard drive performance.
6. System-level network performance.

Following are some parameters with their explanations required to study and perform analysis on data.

Total I/Os per Second: Average number of I/O operations per second, averaged over the length of the test so far.

Total MBs per Second: Average number of Megabytes read and written per second, averaged over the length of the test so far.

Average Latency:
a. Average I/O Response Time (ms): Average time between initiation and completion of an I/O operation, averaged over the length of the test so far, in milliseconds.
b. Average Read Response Time (ms): Average time between initiation and completion of a read operation.
c. Average Write Response Time (ms): Average time between initiation and completion of a write operation.
d. Average Transaction Time (ms): Average time between initiation of a request and completion of the corresponding reply. If there are no replies in the access specification, this is the same as Average I/O Response Time.

% CPU Utilization (total): Percentage of processor time spent executing threads other than the Idle thread (in other words, time spent doing useful work). Also known as % Processor Time [16].

While testing various cases some default settings were done in Iometer
1. Maximum disk size 2048000 sectors
2. Access specification : Default
3. update frequency: 10 seconds
4. Run time: 5 minutes

Sample video data was stored in 5 disks created using openfiler and with help of iSCSI initiator all disks were accessed at Machine 1 (M1) and Machine 2 (M2) following is the image for iSCSI status at target.

Fig 3: Screenshot of iSCSI status

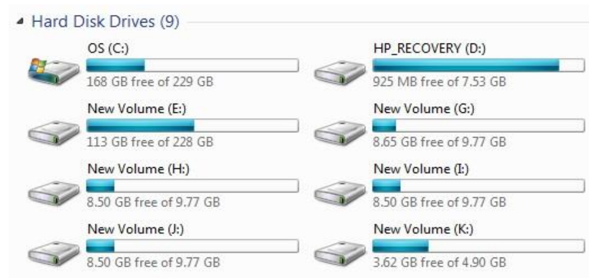


Fig 4: Screenshot of disks available on a machine on connecting to target using iSCSI initiator (G,H,I,J,K)

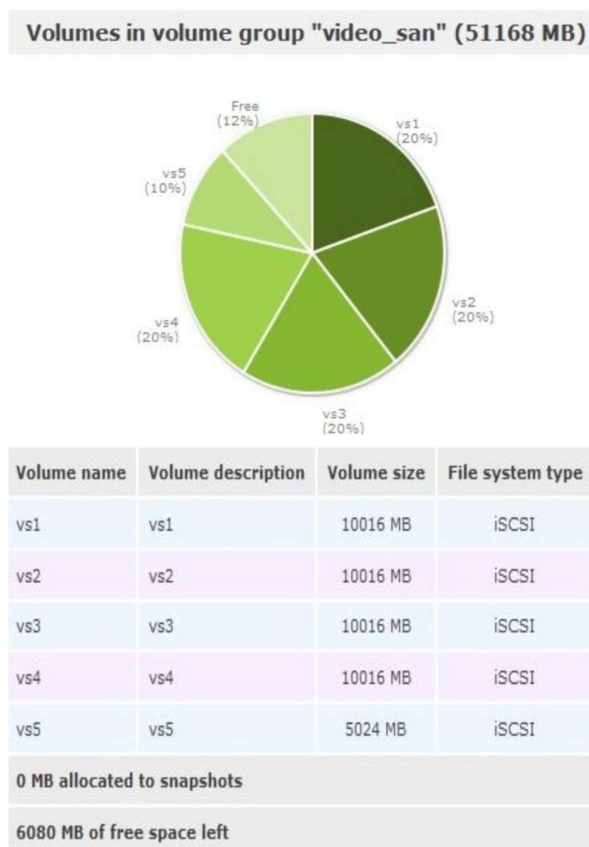


Fig 5: Screenshot of Volumes in Volume group

Readings were taken for following 7 cases

1. Default case: no video streaming
2. Multiple machines accessing one same video
3. Multiple machines accessing multiple videos (same disk)
4. Multiple machines accessing multiple videos (different disks)
5. At a time one machine accessing one video
6. At a time one machine accessing multiple videos (same disk)
7. At a time one machine accessing multiple videos (different disks)

1. For Local Area Network (LAN)

Sr .N o	Cases	Ma chi nes	Total I/Os per second	Total MBs per second	Average I/O Response Time (ms)	% CPU Utilizat ion
1	Case1	M1	160.94	0.31	31.0385	4.51%
		M2	163.30	0.32	30.6459	4.85%
2	Case2	M1	158.17	0.31	6.3193	15.77%
		M2	149.84	0.29	6.6703	12.50%
3	Case3	M1	182.81	0.36	5.4676	21.52%
		M2	191.62	0.37	5.2156	14.82%
4	Case4	M1	117.06	0.23	17.0776	20.09%
		M2	119.67	0.23	16.7087	19.72%
5	Case5	M1	431.14	0.84	2.3178	18.29%
		M2	403.18	0.79	2.4778	10.98%
6	Case6	M1	433.95	0.85	2.3024	25.01%
		M2	374.04	0.73	2.6716	23.67%
7	Case7	M1	206.84	0.40	9.6666	20.59%
		M2	196.09	0.38	10.1965	17.49%

Table 1: Readings taken when machines are connected in LAN

For LAN in Case 1 i.e the Default case: no video streaming the average I/O response time for machine 1 M1 is 31.0385 ms and for machine 2 M2 is 30.6459 which are maximum as compared to other cases. For cases 2,3,5,6 where video access is from on disk the response time is less but when video is accessed from different disks the average I/O response time shoots up as seen in cases 4 & 7. Total MBs/sec is high in cases 5 & 6 where at a time one machine is accessing one video and multiple videos from same disk respectively as compared to other cases.

Machine 1: Case 4

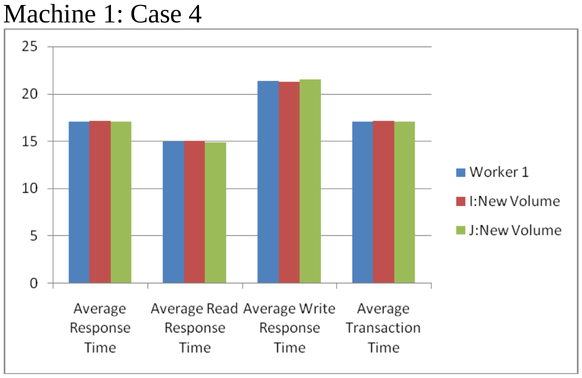


Fig 6: Bar chart for Average Latency in Case4 of Machine 1 (LAN)

Machine 2: Case 4

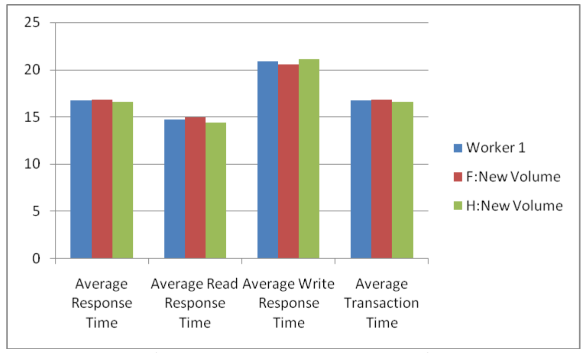


Fig 7: Bar chart for Average Latency in Case4 of Machine 2 (LAN)

2. For Wide Area Network (WAN)

Sr .N o	Cases	Ma chi nes	Total I/Os per second	Total MBs per second	Average I/O Response Time (ms)	% CPU Utilizat ion
1	Case1	M1	507.35	0.99	9.8485	2.64%
		M2	409.98	0.80	12.1943	6.27%
2	Case2	M1	203.19	0.40	4.9192	13.96%
		M2	218.66	0.43	4.5697	6.76%
3	Case3	M1	205.04	0.40	4.8748	19.76%
		M2	236.81	0.46	4.2194	42.08%
4	Case4	M1	219.77	0.43	9.0973	18.52%
		M2	257.91	0.50	7.7523	29.62%
5	Case5	M1	258.06	0.50	3.8733	16.20%
		M2	293.52	0.57	3.4046	8.00%
6	Case6	M1	232.69	0.45	4.2944	24.01%
		M2	260.18	0.51	3.8404	10.18%
7	Case7	M1	344.87	0.67	5.7970	19.34%
		M2	305.52	0.60	6.5433	10.09%

Table 2: Readings taken when machines are connected in WAN

For WAN the average I/O response time is maximum in Case 1 for both machines. And when video access from different disks is involved i.e Cases 4 & 7 the average I/O response time is high compared to other Cases. In case of WAN the Total MBs/sec is high in Case 1 as compared to other cases.

Machine 1: Case 4

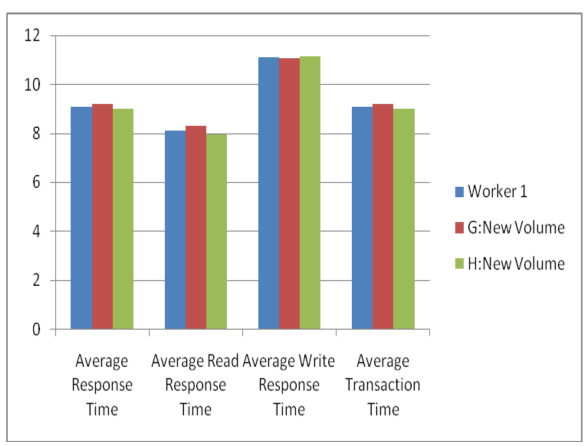


Fig 8: Bar chart for Average Latency in Case4 of Machine 1(WAN)

Machine 2: Case 4

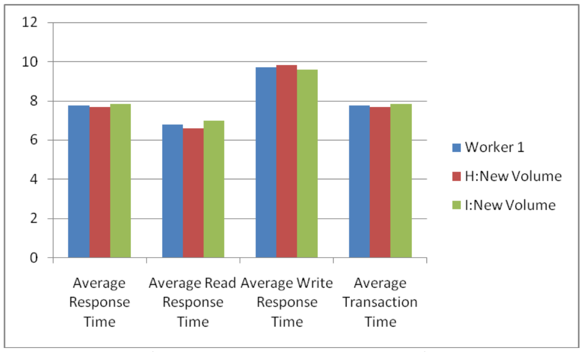


Fig 9: Bar chart for Average Latency in Case4 of Machine 2(WAN)

The Bar charts depicts average write response time in case of LAN or WAN is high as compared to average read response time or average transaction time.

4.0 Oracle Real Application Cluster (RAC)

Oracle's Real Application Clusters (RAC) supports the transparent deployment of a single database across pools of server, providing fault tolerance from hardware failures or planned outages [7]. In an Oracle RAC environment, two or more computers each with an Oracle Relational Database Management system (RDBMS) instance simultaneously access a single database which makes it possible for an application or user to connect to either computer and have access to a single coordinated set of data. The database consists of a collection of data files, control files, and redo logs located on disk. The instance comprises the collection of Oracle-related memory and operating system processes that run on a computer system [8].

4.1 Usage of SAN in Clustered environment

The Internet Small Computer System Interface (iSCSI) is an Internet Protocol (IP)-based storage networking standard for establishing and managing connections between IP-based storage devices, hosts, and clients. iSCSI is a data transport protocol defined in the SCSI-3 specifications framework and is similar to Fibre Channel in that it is responsible for carrying block-level data over a storage network. In Block-level communication data is transferred between the host and the client in chunks called blocks. Database servers depend on this type of communication [9]. Oracle Real Application Clusters is a shared everything architecture. All servers in the server pool share all storage used for an Oracle RAC database. The type of storage pool used can be network attached storage (NAS), Storage Area Network (SAN), or SCSI disks [10]. We will be using Storage Area Network (SAN) by Openfiler.

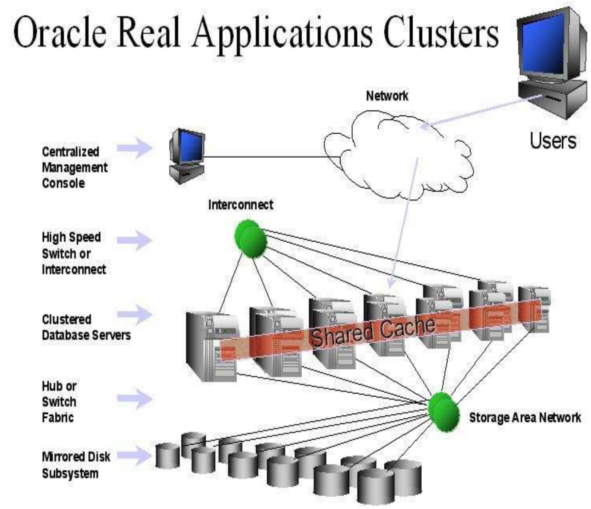


Fig 10: Architecture of Oracle RAC [10]

iSCSI Initiator: An iSCSI initiator is a client device that connects and initiates requests to some service offered by a server (in this case an iSCSI target). The iSCSI initiator software will need to exist on each of the Oracle RAC nodes (racnode1 and racnode2).
iSCSI Target: An iSCSI target is the "server" component of an iSCSI network. It is the storage device that contains the information you want and answers requests from the initiator(s). In this project Openfiler will be the iSCSI target.

4.2 Implementation of Oracle RAC

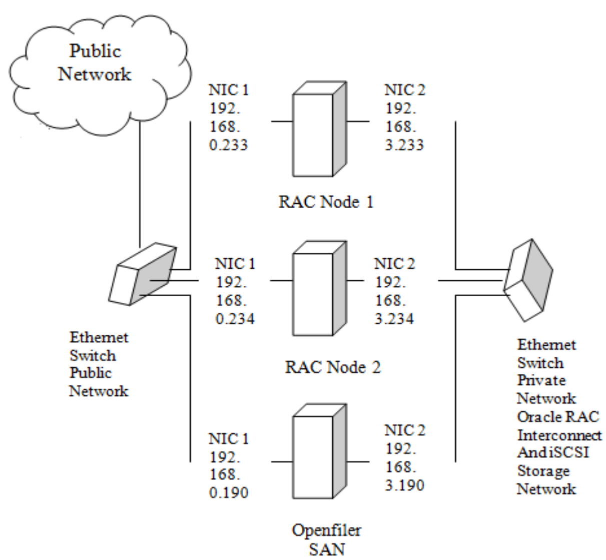


Fig 11: Architecture of Implemented Oracle RAC with IP addresses

In Oracle VM VirtualBox [12] both RAC 1 and RAC 2 machines are built having Operating system Windows server 2008 [11] and following softwares
1. Oracle Grid Infrastructure 11g Release 2
2. Oracle Database 11g Release 2 [13]

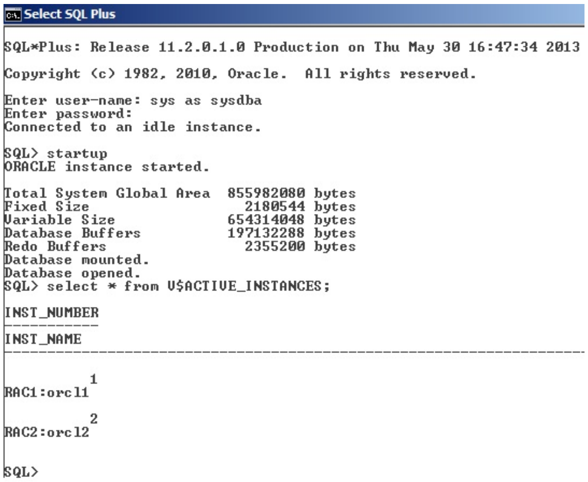


Fig 12: Screenshot of Active Instances in RAC

4.3 Time based analysis of retrieve and display data

First we will put 10 lakh records in single machine database (non-RAC environment).

create table temp (no int);

Now using PL-SQL commands to put 10 lakh records

BEGIN
FOR i in 1..1000000 LOOP
INSERT INTO temp(no) VALUES (i);
END LOOP;
COMMIT;
END;

4.3.1 Time required to put 10 lakh records in single machine database in minutes:seconds

Ta = 1:11.80

4.3.2 Time required to retrieve and display data in single machine database.

*select * from temp;*

T1 = 2:30.65
T2 = 2:30.55
T3 = 2:28.55
T4 = 2:29.55
T5 = 2:29.18

The average time **Tb = 2:29.69**

Now we will put 10 lakh records in tables temp7 from Node 1 and temp8 from Node 2 in RAC database and get retrieve and display results.

4.3.3 Time required to put 10 lakh records in RAC database from node1 and node2 not simultaneously

Nodes	Node 1	Node 2
For Table	temp7	temp8
Time	**Tc1 = 3:12.39**	**Td1 = 3:51.60**

Table 3: Time taken to put data in RAC

4.3.4 Time required to put 10 lakh records in RAC database from node1 and node2 simultaneously.

Nodes	Node 1	Node 2
For Table	temp9	temp10
Time	**Tc2 = 2:53.70**	**Td2 = 3:26.97**

Table 4: Time taken to put data in RAC simultaneously

As we can compare Ta with Tc's & Td's it proves that time required to put data in RAC environment is more than time required to put data in single machine database. From above readings it can analyzed that time difference between nodes is negligible in RAC environment hence as much as nodes increase performance remains same. In case of production environment with **Fibre channel** speed increases tremendously hence multiple nodes can insert data at same time without any data corruption.

4.3.5 Time required to retrieve and display data in RAC environment

Case 1: RAC environment in LAN (Local Area Network) and nodes accessing same data [14]

	Time	Node 1	Node 2
Not Simultaneous Access of data	T1	8:42.46	7:25.55
	T2	6:48.95	6:55.21
	T3	5:16.46	6:53.76
	T4	5:15.46	6:45.18
	T5	5:01.89	7:01.29
	Average	**Tm = 6:13.04**	**Tn = 7:00.02**
Simultaneous Access of data	T1	7:23.27	8:10.73
	T2	7:08.12	7:32.93
	T3	6:57.66	7:18.99
	T4	7:00.74	7:26.98
	T5	7:02.11	7:21.08
	Average	**To = 7:06.38**	**Tp = 7:34.14**

Table 5: Time taken to retrieve & display data from RAC (LAN + same data)

Case 2: RAC environment in LAN (Local Area Network) and nodes accessing different data [14]

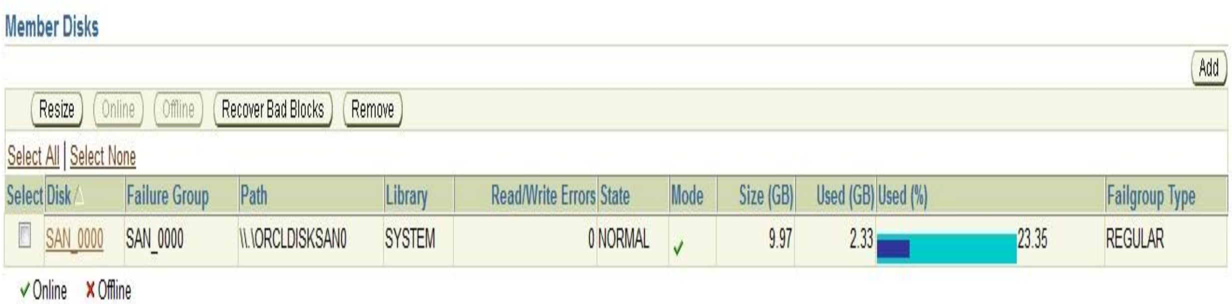
	Time	Node 1	Node 2
Not Simultaneous Access of data	T1	6:24.81	9:34.29
	T2	5:21.98	6:55.16
	T3	5:27.99	6:28.76
	T4	5:29.12	6:14.91
	T5	5:21.68	6:19.42
	Average	**Tq = 5:37.12**	**Tr = 7:06.51**
Simultaneous Access of data	T1	4:33.97	7:25.40
	T2	4:27.17	6:34.78
	T3	4:25.13	7:09.28
	T4	4:24.67	6:56.37
	T5	4:32.90	7:03.11
	Average	**Ts = 4:28.77**	**Tt = 7:01.79**

Table 6: Time taken to retrieve & display data from RAC (LAN + different data)

The above 2 cases proves even in case of retrieve and display of data, time delay is minimum, data can be accessed from many servers simultaneously without corruption. The major advantage here is that data is stored in a centralized storage and multiple nodes can perform transactions on that data through network. Hence RAC provides high availability i.e If a node in a server pool fails, the database continues to run on the remaining server in the pool [7].

5. Conclusion

The paper clearly illustrates the performance analysis of SAN considering seven cases and its importance as centralized storage access to all nodes in case of clustered database. The cost and complexity of Fibre Channel has kept SAN deployment out of reach for small and midsized businesses until the introduction of Storage over IP (SoIP) SANs based on the iSCSI protocol approved by the Internet Engineering Task Force (IETF) in 2003. In case of security Fibre Channel SANs are traditionally less secure than iSCSI. Fibre Channel SAN can see multiple LUNs on any particular disk, but iSCSI can only deal with a disk target. Consequently, iSCSI authentication is very important, and iSCSI employs advanced authentication methods to establish security, such as Challenge Handshake Authentication Protocol (CHAP). Fibre Channel does not support native encryption over the wire, but iSCSI can utilize IPSec encryption to protect data in flight [6]. In SAN if all the hosts are allowed to access all the drives the two important problems arise they are disk resource contention and data corruption. To deal with them one can isolate and protect storage devices on a SAN by using zoning and LUN (Logical Unit Number) masking, which allows to dedicate storage devices on the SAN to individual servers. Zoning: Many devices and nodes can be attached to a SAN, When data is stored in a single cloud, or storage entity, it is important to control which hosts have access to specific devices. Zoning implemented at the hardware level, isolates a single server to a group of storage devices or a single storage device, or associate a grouping of multiple servers with one or more storage devices, as required in a server cluster deployment. LUN masking: performed at the storage controller level, allows to define relationships between LUNs and individual servers. Storage controllers usually provide the means for creating LUN-level access controls that allow access to a given LUN by one or more hosts. By providing this access control at the storage controller, the controller itself enforces access policies to the devices. LUN masking provides more detailed security than zoning, because LUNs provide a means for sharing storage at the port level [15]. SAN holds its importance even in this Cloud era and will continue to provide optimum service to end users.


Member Disks

Add

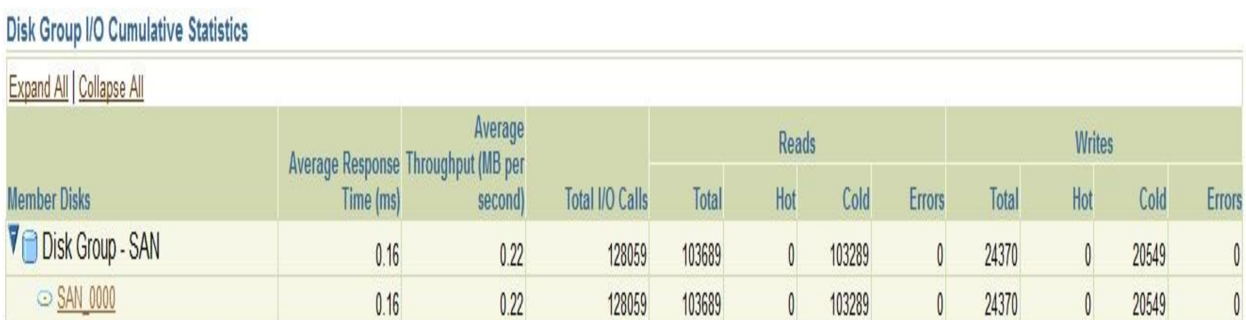
Resize Online Offline Recover Bad Blocks Remove

Select All | Select None

Select	Disk	Failure Group	Path	Library	Read/Write Errors	State	Mode	Size (GB)	Used (GB)	Used (%)	Failgroup Type
☐	SAN_0000	SAN_0000	\\.\ORCLDISKSAN0	SYSTEM	0	NORMAL	✓	9.97	2.33	23.35	REGULAR

✓ Online ✗ Offline

Fig 13: Screenshot of disk available in RAC


Disk Group I/O Cumulative Statistics

Expand All | Collapse All

Member Disks	Average Response Time (ms)	Average Throughput (MB per second)	Total I/O Calls	Reads Total	Reads Hot	Reads Cold	Reads Errors	Writes Total	Writes Hot	Writes Cold	Writes Errors
Disk Group - SAN	0.16	0.22	128059	103689	0	103289	0	24370	0	20549	0
SAN_0000	0.16	0.22	128059	103689	0	103289	0	24370	0	20549	0

Fig 14: Screenshot of disk statistics

References

[1] Introduction to Storage Area Networks and System Networking Jon Tate Pall Beck Hector Hugo Ibarra Shanmuganathan Kumaravel Libor Miklas IBM Fifth Edition (November 2012) © Copyright International Business Machines Corporation 2012. All rights reserved.

[2] http://en.wikipedia.org/wiki/Storage_area_network [last modified 6 June 2013]

[3] http://www.allsan.com/sanoverview.php3 [last modified 6 June 2013]

[4] http://www.tns.com/ip_san.asp [last modified 25 June 2013]

[5] Information Storage and Management Participant guide volume 1 of 2 student guide EMC2 July 2009

[6] http://www.cuttedge.com/files/iscsi_vs_fiberchannel_explain.pdf [last modified 24 June 2013]

[7] Oracle Real Application Cluster , Oracle Data Sheet Copyright © 2010, Oracle and/or its affiliates.All rights Reserved.

[8] http://en.wikipedia.org/wiki/Oracle_RAC [last modified 6 June 2013]

[9] http://www.oracle.com/technetwork /articles/hunter-rac11gr2-iscsi-088677.html [last modified 6 June 2013]

[10] An Oracle White Paper November 2010 Oracle Real Application Clusters (RAC) 11g Release 2 Author: Barb Lundhild, Markus Michalewicz Copyright © 2010, Oracle and/or its affiliates. All rights reserved

[11] http://www.microsoft.com/en-in/ server-cloud/windows-server/2008-r2-trial.aspx [last modified 6 June 2013]

[12] https://www.virtualbox.org/ [last modified 10 June 2013]

[13] http://www.oracle.com/technetwork/database/ enterprise-edition/downloads/index.html [last modified 10 June 2013]

[14] http://www.csgnetwork.com/timescalc.html [last modified 10 June 2013]

[15] http://technet.microsoft.com/en-us/library/ cc758640(v=ws.10).aspx [last modified 10 June 2013]

[16] http://csis.pace.edu/~lombardi/sciences /computer/systems/windows/docs/iometer.pdf [last modified 10 June 2013]

[17] http://www.eetimes.com/design/communications-design/4017917/Storage-Area-Network-Overview [last modified 23 June 2013]

[18] http://media.wiley.com/product_data/excerpt/38 /04703851/0470385138.pdf [last modified 23 June 2013]

First Author: Ms. **Akanksha Verma**, She is pursuing **M.Tech** in Advanced Information Technology with specialization in Networking and Telecommunications from IGNOU – I^{2}IT Centre of Excellence for Advanced Education and Research, Pune, India. She completed **B.E** Computer Science and Engineering (CSE) from Rajiv Gandhi Proudyogiki Vishwavidyalaya (RGVP). Her research interest includes Storage Area Network (SAN) and Networking.

Second Author: Ms. **Shrijee Bhattacharjee,** She is pursuing **M.Tech** in Advanced Information Technology with specialization in Software Technologies from IGNOU – I^{2}IT Centre of Excellence for Advanced Education and Research, Pune, India. She completed **B.Tech** Computer Science and Engineering (CSE) from North Eastern Hill University (NEHU). Her research interest includes Cloud Computing, Database Management System, Data Mining and Business Intelligence.

Third Author: Mr. **Prashant Baste,** He is pursuing **M.Tech** in Advanced Information Technology with specialization in Software Technologies from IGNOU–I^{2}IT Centre of Excellence for Advanced Education and Research, Pune, India. He completed **B.E** Computer Science and Engineering from University of Pune. His research interests include databases, data mining and cloud computing.

Fourth Author: Mr. **Nandan Nagarajappa Mirajkar,** He is pursuing **M.Tech** in Advanced Information Technology with specialization in Software Technologies from IGNOU – I^{2}IT Centre of Excellence for Advanced Education and Research, Pune, India. He is also **Teaching Assistant** in Advanced Software and Computing Technologies Department. He completed **B.E** Electronics and Telecommunications from University of Mumbai. His research interests include Cloud computing, Databases and Networking.

Fifth Author: Mr. **Sandeep Bhujbal,** He is **Sr. Research Associate** in Advanced Software and Computing Technologies Department of IGNOU – I^{2}IT Centre of Excellence for Advanced Education and Research, Pune, India. He completed **M.C.S** from University of Pune. His research interests include Operating systems, Compiler construction, Programming languages and Cloud computing.

Sixth Author: Mr. **Nilesh Deshmukh,** He is **Senior Lecturer** in Advanced Software and Computing Technologies Department of IGNOU – I^{2}IT Centre of Excellence for Advanced Education and Research, Pune, India. He completed APGP from I^{2}IT and **B.E** Computer Science from Amravati University. His research interests include Storage Area Network (SAN) and Networking.

Seventh Author: Mr. **Kumud Singh**, He is Senior Executive IT Infrastructure Services in IGNOU – I^{2}IT Centre of Excellence for Advanced Education and Research Pune, Maharashtra, India. He completed **BCA** from GIMT Delhi, India. His key interests are in Virtualization and Network security.